Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 53-2 (2020) 13287–13292

A distributed approach to the Optimal Power Flow problem for unbalanced and mesh networks ⋆

Giulio Ferro * **Michela Robba** * **David D'Achiardi** **
Rabab Haider ** **Anuradha M. Annaswamy** **

* *Department of Informatics, Bioengineering, Robotics and Systems Engineering, University of Genoa, Genoa, Italy (email: giulio.ferro@edu.unige.it)*

** *Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA, 02141 USA (email: aanna@mit.edu).*

Abstract: In the present paper we introduce a new distributed optimization approach to the solution of general Optimal Power Flow problem for unbalanced and mesh networks. A new convex formulation, based on McCormick Envelopes, is proposed with a decomposition profile and a distributed approach based on proximal coordination. The resulting algorithm is shown to converge with a rate of $o(1/\tau)$, where τ is the number of iterations. The approach is validated on a modified IEEE 13 bus network, with added distributed energy resources including distributed generation and demand response.

Keywords: Smart grids, modeling of power systems, optimal operation and control of power systems, optimization and control of large-scale network systems, decentralized and distributed control.

1. INTRODUCTION AND STATE OF THE ART

Over the last decade, the power grid has undergone transformation with the increased penetration of small-scale distributed energy resources (DERs), primarily in the distribution network. These include renewable energy resources (RERs), distributed generation (DG), demand response (DR), and storage devices. The large number of DERs pose a challenge to grid operators, in which the state of art is to employ centralized decision making. With such a large number of DERs, these centralized decision making tools, which typically include optimization problems for resource dispatch, become intractable. To efficiently integrate these DERs, it is necessary to develop models and methods for optimizing power injections into the grid which are able to deal with large scale systems. As the DERs are situated throughout the distribution network, detailed and accurate models of the electric networks are needed, while still ensuring tractability within an optimization framework. A detailed representation of the power grid is provided by load flow equations (Kersting, 2006).

Optimal Power Flow (OPF) denotes the method utilized for optimizing power injections or flows within the network, subject to constraints that correspond to the power physics of the grid (Dommel and Tinney, 1968). Generally, the OPF is challenging to solve since it is a non-convex nonlinear problem. To address this, convex relaxations are often employed to render the problem easier to solve. Of these, the most popular power system models are the bus injection model (BIM) based on semidefinite programming (SDP), and the branch flow model (BF) based on second order cone programming (SOCP) (Farivar et al., 2011; Lavaei and Low, 2012). It should be noted that the BF model guarantees optimality for a limited set of networks, namely networks with radial topology and balanced structure, with limits on how active the network can be (Farivar and Low, 2014; Christakou et al., 2017). This is highly problematic for distribution systems where lines are unbalanced and consist of many single-phase loads.

To address these challenges and limitations, we first propose a new convex formulation of the OPF problem which can be used for any general distribution grid (including meshed and unbalanced topologies). The formulation is based on the current injection (CI) method, and uses McCormick Envelopes (MCE) (Mccormick, 1976) to convexify bilinear constraints. We then reformulate the OPF problem posed using the CI model, in the form of a distributed algorithm based on the Proximal Atomic Coordination method (PAC) Romvary (2018) so as to make it computationally tractable.

In the current literature there are many works that apply distributed optimization to power distribution networks. In (Li et al., 2012) DR is studied in a radial distribution network, by formulating it as an OPF minimizing energy costs and power line losses, subject to the power flow constraints, with the BF model, and operating constraints. The results therein are derived based on assumptions that include a balanced structure, radial topology and network's passivity (i.e. inverse power flow). In contrast, our proposed approach removes all of these assumptions.

⋆ This work was supported in part by the Department of Energy under Award Number DE-IA0000025 for UI-ASSIST Project.

10.1016/j.ifacol.2020.12.159

In (Dall'Anese et al., 2013) and (Robbins and Dominguez-Garcia, 2016) the authors model an unbalanced distribution network modifying the BF model and solving the proposed OPF through the alternating direction method of multipliers (ADMM) algorithm. In these papers, it is assumed that the reciprocal unbalance of the phases are small, which implies that the relative difference between the angles of each phase are ignored.

The main contribution of the proposed paper can be summarized as below:

- A new convex relaxation of the OPF problem based on MCE, that can model distribution grids that may be unbalanced, or have a meshed topology.
- The development of a distributed algorithm, based on the PAC method, with clear articulation of conditions for proof of convergence.

The organization of this paper is as follows. Section 2 introduces the mathematical notation, Section 3 describes the proposed relaxed OPF problem. In Section 4 we state the distributed algorithm and prove its convergence properties. In Section 5 we show results of the proposed CI model and distributed algorithm for the voltage regulation problem on a modified IEEE 13 bus network, with DERs. Some concluding remarks are provided in Section 6.

2. MATHEMATICAL NOTATION

We denote scalars as lowercase characters (μ or y), vectors as lowercase boldfaced characters (e.g., $\boldsymbol{\mu}$ or $\mathbf{y}$), and matrices as uppercase English characters (e.g., $\mathbf{G}$). Functions are represented as italicized lowercase characters (e.g., f). For matrices $\mathbf{G} \in \mathbb{R}^{n\times n}$, we let $\lambda_{\mathbf{min}}(\mathbf{G})$, $\hat{\lambda}_{\mathbf{min}}(\mathbf{G})$ and $\lambda_{\mathbf{max}}(\mathbf{G})$ represents it smallest, smallest non-zero and largest eigenvalue, respectively. Sets are then represented as uppercase specialized characters (e.g., $\mathcal{S}$ or L). In addition, we denote n-dimensional vectors with all zero entries by $\mathbf{0_n}$ and we utilize the overbar $\overline{\mathbf{x}}$ and underbar $\underline{\mathbf{x}}$ notation to denote upper and lower limits for a variable $\mathbf{x}$, phasors are denoted as x.

We define $[Y] \triangleq \{1,\dots,Y\}$ for any $Y \in \mathbb{N}$. For any collection of $[Y]$-subsets $\mathsf{W} = \{W_\mathbf{j}\}_{\mathbf{j}\in[K]}$ s.t. $W_\mathbf{j} \subseteq [Y]$, we use the following cardinality convention:

- $N_\mathbf{j}^\mathsf{W} \triangleq |W_\mathbf{j}|$ is the size of the $\mathbf{j}$th subset;
- $N^\mathsf{W} \triangleq \sum_{\mathbf{j}\in[K]} N_\mathbf{j}^\mathsf{W}$ is the sum of all such subset sizes.

In addition, we denote the elements of each subset $W_\mathbf{j} = \left\{w_\mathbf{k}^{[\mathbf{j}]}\right\}_{\mathbf{k}\in N_\mathbf{j}^\mathsf{W}}$, where the subscript $\mathbf{k}$ indicates that $w_\mathbf{k}^{[\mathbf{j}]}$ is the $\mathbf{k}$th element of the subset $W_\mathbf{j}$ and the superscript $[\mathbf{j}]$ signifies that $w_\mathbf{k}^{[\mathbf{j}]}$ belongs to the $\mathbf{j}$th subset.

We then succinctly represent vectors and matrices using this notation:

- $\mathbf{v} = \{v_\mathbf{r}\}_{\mathbf{r}\in[M]}$ represents the vector with elements $v_\mathbf{r}$;
- $\mathbf{G} = \{g_{\mathbf{r},\mathbf{c}}\}_{\mathbf{r}\in[M],\mathbf{c}\in[Y]} \in \mathbb{R}^{M\times Y}$ represents the matrix with elements $g_{\mathbf{r},\mathbf{c}}$;
- $\mathbf{v}_{R_\mathbf{i}} \triangleq \{v_\mathbf{r}\}_{\mathbf{r}\in R_\mathbf{i}}$ represents a submatrix of $\mathbf{v}$ with entries from rows $R_\mathbf{i} \subseteq [M]$;
- $\mathbf{G}_{R_\mathbf{i},C_\mathbf{j}} \triangleq \{g_{\mathbf{r},\mathbf{c}}\}_{\mathbf{r}\in R_\mathbf{i},\mathbf{c}\in C_\mathbf{j}}$ represents the submatrix of $\mathbf{G}$ with entries from rows $R_\mathbf{i} \subseteq [M]$ and columns $C_\mathbf{j} \subseteq [Y]$.

In addition, if we want to represent a submatrix or subvector with all rows or all columns of $\mathbf{G}$, we use the subscript '$-$'. E.g., $\mathbf{G}_{R_\mathbf{i},-} \triangleq \mathbf{G}_{R_\mathbf{i},[Y]}$ represents the submatrix of $\mathbf{G}$ with entries at specific rows $R_\mathbf{i}$ and all columns.

3. THE RELAXED OPF PROBLEM

The distribution network is modelled as an undirected graph $G(N,E)$, with N set of nodes and E set of branches. For a general 3-phase network, each variable is a vector with 3 components, denoting the phases $k \in K$ with $K = a,b,c$). The impedance matrix $\mathbf{Z}_{r,s} \in \mathbb{R}^{3\times3}$ for a line between nodes r and s, represents the magnetic coupling between phases k and k', such that $z_{k,k'} = R_{k,k'} + jX_{k,k'} \forall k,k' \in K$. The OPF problem aims at minimizing a performance index subject to constraints that describe the physics of the power system. The performance index can be, for example, to minimize line losses, cost for power production, or deviation of voltage from nominal setpoints.

We model the current injections per phase at each node in the network. A general formulation of a 3-phase CI-OPF is given by:

$$\min_{(\mathbf{i},\mathbf{v},\mathbf{s})} f(\mathbf{i},\mathbf{v},\mathbf{s}) \tag{1}$$

$$\mathbf{v} = \mathbf{Z}\mathbf{i} \tag{2}$$

$$\mathbf{s} = \mathbf{v}\mathbf{i}^H \tag{3}$$

$$\underline{\mathbf{v}} \le \mathbf{v} \le \overline{\mathbf{v}} \tag{4}$$

$$\underline{\mathbf{s}} \le \mathbf{s} \le \overline{\mathbf{s}} \tag{5}$$

where $\mathbf{i}$, $\mathbf{v}$ and $\mathbf{s}$ are vectors ($\in \mathbb{C}$) for nodal current injection, voltages, and apparent power injection respectively, $\mathbf{Z}$ is the network's impedance matrix, and H denotes the Hermitian matrix. In this CI-OPF formulation, equation (1) describes general objective function typical of power systems minimizing a function of current voltages or power, (2) is the definition of the nodal analysis method for an AC system, (3) is the definition of apparent power, and (4)-(5) are upper and lower bounds for voltage and apparent power injected at each bus respectively.

As can be seen from (1)-(5), the source of non-convexity is the bilinear term $\mathbf{v}\mathbf{i}^H$. To this end we use MCE (Mccormick, 1976) as convex relaxation to our bilinear problem, which denotes the convex hull of a bilinear form $w = xy$. We denote this as $\text{MCE}(w,\underline{x},\overline{x},\underline{y},\overline{y}) = \{w = xy : x \in [\underline{x},\overline{x}], y \in [\underline{y},\overline{y}]\}$, and formally define it as:

$$M(w,\underline{x},\overline{x},\underline{y},\overline{y}) = \begin{cases} w \ge \underline{x}y + x\underline{y} - \underline{x}\underline{y} \\ w \ge \overline{x}y + x\overline{y} - \overline{x}\overline{y} \\ w \le \underline{x}y + x\overline{y} - \overline{x}\underline{y} \\ w \le \overline{x}y + x\underline{y} - \underline{x}\overline{y} \end{cases} \tag{6}$$

Using the relaxation described above, we can rewrite the convex optimization problem as:

$$\min_{(\mathbf{i},\mathbf{v},\mathbf{s})} f(\mathbf{i},\mathbf{v},\mathbf{s}) \tag{7}$$

$$\mathbf{v} = \mathbf{Z}\mathbf{i} \tag{8}$$

$$\mathbf{s} \geq \underline{\mathbf{v}}\mathbf{i}^H + \mathbf{v}\underline{\mathbf{i}}^H - \underline{\mathbf{v}}\underline{\mathbf{i}}^H \tag{9}$$

$$\mathbf{s} \geq \overline{\mathbf{v}}\mathbf{i}^H + \mathbf{v}\bar{\mathbf{i}}^H - \overline{\mathbf{v}}\bar{\mathbf{i}}^H \tag{10}$$

$$\mathbf{s} \leq \underline{\mathbf{v}}\mathbf{i}^H + \mathbf{v}\bar{\mathbf{i}}^H - \overline{\mathbf{v}}\underline{\mathbf{i}}^H \tag{11}$$

$$\mathbf{s} \leq \overline{\mathbf{v}}\mathbf{i}^H + \mathbf{v}\underline{\mathbf{i}}^H - \underline{\mathbf{v}}\bar{\mathbf{i}}^H \tag{12}$$

$$\underline{\mathbf{v}} \leq \mathbf{v} \leq \overline{\mathbf{v}} \tag{13}$$

$$\underline{\mathbf{i}} \leq \mathbf{i} \leq \bar{\mathbf{i}} \tag{14}$$

$$\underline{\mathbf{s}} \leq \mathbf{s} \leq \bar{\mathbf{s}} \tag{15}$$

where equation (14) describes suitable current limits introduced to use the MCE.

We note that for brevity, the formulation of the optimization problem has been condensed; the complete formulation includes the separation of the complex variables, $\mathbf{i}, \mathbf{v}, \mathbf{s}$ into real and imaginary components, where $\mathbf{s} = \mathbf{p} + j\mathbf{q}$, and $\mathbf{p}$ and $\mathbf{q}$ are the real and reactive powers respectively.

4. PROXIMAL ATOMIC COORDINATION ALGORITHM

In this section, we summarize the results of Romvary (2018) for the sake of completeness and comprehension.

We start by considering the following *global standard optimization (GSO):*

$$\min_{\mathbf{y}\in\mathbb{R}^Y} \left\{ f(\mathbf{y}) \triangleq \sum_{\mathbf{k}\in[B]} f_{\mathbf{k}}(\mathbf{y}) \right\} \tag{16}$$

$$\mathbf{G}\mathbf{y} = \mathbf{0}_{\mathbf{M}} \tag{17}$$

where:

- $\mathbf{y} \in \mathbb{R}^Y$ represents the Y optimization variables;
- $f_{\mathbf{k}} : \mathbb{R}^Y \to \mathbb{R}$ represents an objective function summand;
- $\sum_{\mathbf{k}\in[B]} f_{\mathbf{k}}(\mathbf{y})$ represents the total objective function;
- $\mathbf{G} = \left[\mathbf{g}_\mathbf{1}^T; \cdots; \mathbf{g}_\mathbf{M}^T\right] \in \mathbb{R}^{M\times Y}$ is a matrix s.t. $\mathbf{null}(\mathbf{G})$ represents the feasibility region of the GSO and each row $\mathbf{g}_\mathbf{i} \in \mathbb{R}^Y$ represents the feasibility constraints [1].

We can decompose the GSO into K different coupled sub-optimization problems by using the following decomposition profile $\mathcal{D} = (\mathsf{L}, \mathsf{C}, \mathsf{S}, \mathsf{O}, \mathsf{T})$:

- $\mathsf{L} = \{L_\mathbf{j}\}_{\mathbf{j}\in[K]}$ represents the partition of $[Y]$ with $L_\mathbf{j} \subseteq [N]$ being the components of $\mathbf{y}$ that each $\mathbf{j}$th atom "owns";
- $\mathsf{C} = \{C_\mathbf{j}\}_{\mathbf{j}\in[K]}$ represents the partition of $[M]$ with $C_\mathbf{j} \subseteq [M]$ representing the rows of $\mathbf{G}$ that each $\mathbf{j}$th atom "owns";
- $\mathsf{S} = \{S_\mathbf{j}\}_{\mathbf{j}\in[K]}$ represents the partition of $[B]$ with $S_\mathbf{j} \subseteq [B]$ representing the objective summands of f that each $\mathbf{j}$th atom "owns";
- $\mathsf{O} = \{O_\mathbf{j}\}_{\mathbf{j}\in[K]}$ has each $O_\mathbf{j} \subseteq [Y]$ representing the "copies" of variables of $\mathbf{y}$ that each $\mathbf{j}$th atom needs to satisfy the *scope* of both $\mathbf{G}_{C_\mathbf{j},-}$ and $f_{S_\mathbf{j}}(\mathbf{y}) \triangleq \sum_{\mathbf{k}\in S_\mathbf{j}} f_\mathbf{k}(\mathbf{y})$; [2]
- $\mathsf{T} = \{T_\mathbf{j}\}_{\mathbf{j}\in[K]}$ represents the atomic partitioning of the GSO where each $T_\mathbf{j} = L_\mathbf{j} \cup O_\mathbf{j} \subseteq [Y]$.

Using $\mathcal{D}$, we obtain the following *atomized standard optimization (ASO)*:

$$\min_{\mathbf{a}_\mathbf{j}\in\mathbb{R}^{N_\mathbf{j}^\mathsf{T}}} \left\{ \sum_{\mathbf{j}\in[K]} \tilde{f}_\mathbf{j}(\mathbf{a}_\mathbf{j}) \right\}$$
$$\text{subj. to: } \begin{cases} \tilde{\mathbf{G}}_\mathbf{j}\mathbf{a}_\mathbf{j} = \mathbf{0}_{\mathbf{N}_\mathbf{j}^\mathsf{C}}, & \text{for all } \mathbf{j} \in [K] \\ \mathbf{B}_{\mathbf{j},-}\mathbf{a} = \mathbf{0}_{\mathbf{N}_\mathbf{j}^\mathsf{O}}, & \text{for all } \mathbf{j} \in [K] \end{cases}, \tag{18}$$

with each atom's variables (both owned and copied) being represented by $\mathbf{a}_\mathbf{j} \in \mathbb{R}^{N_\mathbf{j}^\mathsf{T}}$ and:

- $\tilde{f}_\mathbf{j}(\mathbf{a}_\mathbf{j}) = \sum_{\mathbf{k}\in\left[N_\mathbf{j}^\mathsf{S}\right]} f_\mathbf{k}\left(\mathbf{\Pi}_{\mathbf{T}_\mathbf{j}}\mathbf{y}\right)$ for all $\mathbf{k} \in \left[N_\mathbf{j}^\mathsf{S}\right]$ and $\mathbf{y} \in \mathbb{R}^Y$, where $\mathbf{\Pi}_{\mathbf{T}_\mathbf{j}} \in \mathbf{R}^{Y\times Y}$ is given by:
$$\mathbf{\Pi}_{\mathbf{T}_\mathbf{j}} \triangleq \operatorname{diag}\left\{\delta_{1\in T_\mathbf{j}}, \ldots, \delta_{Y\in T_\mathbf{j}}\right\},$$
with:
$$\delta_{\mathbf{n}\in T_\mathbf{j}} \triangleq \begin{cases} 1, \text{ if } \mathbf{n} \in T_\mathbf{j} \\ 0, \text{ otherwise} \end{cases}.$$
- $\tilde{\mathbf{G}}_\mathbf{j} \triangleq \mathbf{G}_{C_\mathbf{j},T_\mathbf{j}} \in \mathbb{R}^{N_\mathbf{j}^\mathsf{C}\times N_\mathbf{j}^\mathsf{T}}$ represents the submatrix of $\mathbf{G}_{C_\mathbf{j},-}$ obtained by removing all zero columns;
- $\mathbf{B} \in \mathbb{R}^{N^\mathsf{O}\times N^\mathsf{T}}$ represents the adjacency matrix of the directed graph with nodes signifying the atomic variables with edges $\left(y_\mathbf{k}^{[\mathbf{i}]}, y_\mathbf{k}^{[\mathbf{j}]}\right)$, where $y_\mathbf{k}^{[\mathbf{i}]}$ with $\mathbf{k} \in L_\mathbf{i}$ is defined to be the variable of $\mathbf{y}$ that is "owned" by the $\mathbf{j}$th atom and $y_\mathbf{k}^{[\mathbf{j}]}$ with $\mathbf{k} \in O_\mathbf{j}$ defined to be the variable of $\mathbf{y}$ that is "copied" by the $\mathbf{j}$th atom;
- $\mathbf{B}_{\mathbf{j},-} \triangleq \mathbf{B}_{\tilde{O}_\mathbf{j},-} \in \mathbb{R}^{N_\mathbf{j}^\mathsf{O}\times N^\mathsf{T}}$ represents the relevant incoming edges of the directed graph with adjacency matrix $\mathbf{B}$ of the variables "copied" by the $\mathbf{j}$th atom, with $\tilde{O}_\mathbf{j} = [|O_\mathbf{j}|] + \sum_{\mathbf{k}\in[\mathbf{j}-1]} |O_\mathbf{k}|$;
- $\mathbf{B}_{-,\mathbf{j}} \triangleq \mathbf{B}_{-,\tilde{T}_\mathbf{j}} \in \mathbb{R}^{N^\mathsf{O}\times N_\mathbf{j}^\mathsf{T}}$ represents the relevant out-going edges of the directed graph with adjacency matrix $\mathbf{B}$ of the variables "owned" by the $\mathbf{j}$th atom, with $\tilde{T}_\mathbf{j} = [|T_\mathbf{j}|] + \sum_{\mathbf{k}\in[\mathbf{j}-1]} |T_\mathbf{k}|$.

We notice how the ASO of (18) is an augmented version of the GSO of (16), sharing the same optimal solutions. Specifically, we relate the optimal ASO $\mathbf{a}^*$ to optimal GSO $\mathbf{y}^*$ via $\mathbf{y}^* = \mathbf{\Pi}^\mathsf{L}\mathbf{a}^*$, where each $\mathbf{\Pi}^\mathsf{L} : \mathbb{R}^Y \to \mathbb{R}^{N^\mathsf{T}}$ represents the projection from $\mathbf{a}$-space into $\mathbf{y}$-space that consists of all owned variables in each $\mathbf{a}_{\mathbf{j}^*}$ being placed correctly in the resultant $\mathbf{y}^{*\prime} = \mathbf{\Pi}^\mathsf{L}\mathbf{a}^*$.

As we will see in the next section, we prefer to use the ASO formulation to solve the underlying problem because it allows us to compute the primal update in a distributed fashion, i.e., by using the ASO we can iterate over all $\mathbf{a}_\mathbf{j}$ concurrently. The cost of this parallelization, however, is the addition of another constraint, termed *coordination*, that needs to be satisfied for every atom $\mathbf{j} \in [K]$:

$$\mathbf{B}_{\mathbf{j},-}\mathbf{a} = \mathbf{0}_{\mathbf{N}_\mathbf{j}^\mathsf{O}}.$$

[1] It should be pointed out that a general formulation of an optimization problem also includes inequality constraints $\mathbf{h}\mathbf{y} \leq \mathbf{0}$ that represent the domain of optimization variables. The discussions here can be easily extended to this case as well by suitably restricting the variables' feasible set.

[2] By *scope*, we mean the dimensions of $\mathbf{y}$ over which either $\mathbf{G}_{C_\mathbf{j},-}\mathbf{y}$ or $f_{S_\mathbf{j}}(\mathbf{y})$ vary their value. For example, the scope of $\mathbf{G}_{C_\mathbf{j},-}$ is precisely the set of its columns which have at least one nonzero value.

This coordination constraint can be interpreted as requiring all atomic copied variables in a given $\mathbf{j}$th atom (given by $O_{\mathbf{j}}$) to equal the value of their corresponding owned variable in another non $\mathbf{j}$-atom.

4.1 Algorithm Specifications

We begin by forming the atomic Lagrangian function for (18):

$$\begin{aligned}\mathcal{L}(\mathbf{a},\boldsymbol{\mu},\boldsymbol{\nu}) &= \sum_{\mathbf{j}\in[K]}\left[\hat{f}_{\mathbf{j}}(\mathbf{a}_{\mathbf{j}})+\boldsymbol{\mu}_{\mathbf{j}}^T\tilde{\mathbf{G}}_{\mathbf{j}}\mathbf{a}_{\mathbf{j}}+\boldsymbol{\nu}_{\mathbf{j}}^T\mathbf{B}_{\mathbf{j},\_}\mathbf{a}\right]\\ &= \sum_{\mathbf{j}\in[K]}\left[\hat{f}_{\mathbf{j}}(\mathbf{a}_{\mathbf{j}})+\boldsymbol{\mu}_{\mathbf{j}}^T\tilde{\mathbf{G}}_{\mathbf{j}}\mathbf{a}_{\mathbf{j}}+\boldsymbol{\nu}^T\mathbf{B}_{\_,\mathbf{j}}\mathbf{a}_{\mathbf{j}}\right]\\ &\triangleq \sum_{\mathbf{j}\in[K]}\mathcal{L}_{\mathbf{j}}(\mathbf{a}_{\mathbf{j}},\boldsymbol{\mu}_{\mathbf{j}},\boldsymbol{\nu}).\end{aligned}\tag{19}$$

We can then apply the prox-linear approach of ((Chen and Teboulle, 1994)) to (19) and obtain the *proximal atomic coordination (PAC)* algorithm:

$$\mathbf{a}_{\mathbf{j}}[\tau+1]=\underset{\mathbf{a}_{\mathbf{j}}\in\mathbb{R}^{N_{\mathbf{j}}^{\intercal}}}{\mathbf{argmin}}\left\{\begin{array}{l}\mathcal{L}_{\mathbf{j}}(\mathbf{a}_{\mathbf{j}},\bar{\boldsymbol{\mu}}_{\mathbf{j}}[\tau],\bar{\boldsymbol{\nu}}[\tau])\\ +\frac{1}{2\rho}\|\mathbf{a}_{\mathbf{j}}-\mathbf{a}_{\mathbf{j}}[\tau]\|_2^2\end{array}\right\},\tag{20}$$

$$\boldsymbol{\mu}_{\mathbf{j}}[\tau+1]=\boldsymbol{\mu}_{\mathbf{j}}[\tau]+\rho\gamma\tilde{\mathbf{G}}_{\mathbf{j}}\mathbf{a}_{\mathbf{j}}[\tau+1],\tag{21}$$

$$\bar{\boldsymbol{\mu}}_{\mathbf{j}}[\tau+1]=\boldsymbol{\mu}_{\mathbf{j}}[\tau+1]+\rho\gamma\tilde{\mathbf{G}}_{\mathbf{j}}\mathbf{a}_{\mathbf{j}}[\tau+1],\tag{22}$$

$$\text{Communicate } \{\mathbf{a}_{\mathbf{j}}\}_{\mathbf{j}\in[K]} \text{ within network},\tag{23}$$

$$\boldsymbol{\nu}_{\mathbf{j}}[\tau+1]=\boldsymbol{\nu}_{\mathbf{j}}[\tau]+\rho\gamma_{\mathbf{j}}\mathbf{B}_{\mathbf{j},\_}\mathbf{a}[\tau+1],\tag{24}$$

$$\bar{\boldsymbol{\nu}}_{\mathbf{j}}[\tau+1]=\boldsymbol{\nu}_{\mathbf{j}}[\tau+1]+\rho\hat{\gamma}_{\mathbf{j}}[\tau+1]\mathbf{B}_{\mathbf{j},\_}\mathbf{a}[\tau+1],\tag{25}$$

$$\text{Communicate } \{\bar{\boldsymbol{\nu}}_{\mathbf{j}}\}_{\mathbf{j}\in[K]} \text{ within network},\tag{26}$$

where we utilized the previously mentioned prox-linear method to ensure parallel computation of each primal step.

4.2 Structural Assumptions

We next make the following assumptions on the structure of the GSO and ASO formulations:

Assumption 4.1. Each $\hat{f}_{\mathbf{k}}\in\left\{\tilde{f}_{\mathbf{j}}\right\}_{\mathbf{j}\in[K]}$ is a closed, convex and proper function with $\text{dom}\left(\hat{f}_{\mathbf{k}}\right)=\mathbb{R}^Y$.

Assumption 4.2. There exists a non-trivial optimal GSO solution $\mathbf{y}^*\in\mathbb{R}^N$. The optimal ASO solution $\mathbf{a}^*\in\mathbb{R}^{N^{\intercal}}$ is related to $\mathbf{y}^*\in\mathbb{R}^N$ via: $\mathbf{y}^*=\boldsymbol{\Pi}^{\mathbf{L}}\mathbf{a}^*$.

Assumption 4.3. Let the PAC parameters satisfy:

$$1>\rho^2\gamma\lambda_{\mathbf{max}}\left(\tilde{\mathbf{G}}^T\tilde{\mathbf{G}}+\mathbf{B}^T\mathbf{B}\right)$$

4.3 Algorithm's Convergence

Defining the following quantities:

$$\begin{aligned}\tilde{\mathbf{G}}&=\mathbf{diag}\left\{\tilde{\mathbf{G}}_{\mathbf{1}},\ldots,\tilde{\mathbf{G}}_{\mathbf{K}}\right\}\\ \mathbf{V_1}&=\tilde{\mathbf{G}}^T\tilde{\mathbf{G}}+\mathbf{B}^T\mathbf{B},\\ \tilde{\mathbf{V}}_{\mathbf{1}}(\gamma)&=\gamma\mathbf{V_1}\\ \tilde{\mathbf{V}}_{\mathbf{2}}(\rho,\gamma)&=\frac{1}{\rho^2}\mathbf{I}_{\mathbf{N}^{\intercal}}-\tilde{\mathbf{V}}_{\mathbf{1}}(\gamma)\end{aligned}$$

Using the above GSO and ASO optimization models, we now state the main result of the paper:

Theorem 4.4. Let the PAC parameters satisfy $\rho>0$ and $\gamma>0$. Further, let:

$$\mathbf{a}[\tau]=[\mathbf{a_1}[\tau];\cdots;\mathbf{a_K}[\tau]],$$

represent the PAC trajectory of (20)-(26) under zero-initialization. Then if Assumptions 4.1, 4.2 and 4.3 holds, then there exists an optimal ASO solution $\mathbf{a}^*$ s.t.:

$$\lim_{\tau\to\infty}\{\mathbf{a}[\tau]\}=\mathbf{a}^*,$$

with convergence rate satisfying, for all $\tau\in\mathbb{N}$:

$$\left\{\|\mathbf{a}[\tau+1]-\mathbf{a}[\tau]\|_{\mathbf{V_2}(\rho,\gamma)}\right\}_{\tau\in\mathbb{N}}=o\left(\frac{1}{\tau}\right).$$

Proof We shall proceed in a similar fashion to the proofs of Lemmas 2.1 and 2.2 in ((Chen and Teboulle, 1994)) . From dual update equations of (21) and (24), we have:

$$\boldsymbol{\mu}[\tau]=\rho\gamma\tilde{\mathbf{G}}_{\mathbf{j}}\left[\sum_{s=0}^{\tau}\mathbf{a}[s]\right],\tag{27}$$

$$\boldsymbol{\nu}[\tau]=\rho\gamma\mathbf{B}\left[\sum_{s=0}^{\tau}\mathbf{a}[s]\right].\tag{28}$$

Using both (27) , (28) in conjunction with the necessary condition for optimality of (20), we have:

$$\begin{aligned}\mathbf{0}_{\mathbf{N}^{\intercal}}=\frac{1}{\rho}\mathbf{c}_{\hat{f}}^{[\tau+1]}&+\rho\hat{\mathbf{V}}_1(\gamma)\sum_{s=0}^{\tau}\mathbf{a}[s]\\ &+\tilde{\mathbf{V}}_{\mathbf{2}}(\rho,\gamma)(\mathbf{a}[\tau+1]-\mathbf{a}[\tau])\end{aligned}\tag{29}$$

where $\mathbf{c}_{\hat{f}}^{[\tau+1]}\in\partial\hat{f}(\mathbf{a}[\tau+1])$. Assumption 4.3 gives us:

$$\begin{aligned}1>\rho^2\lambda_{\mathbf{max}}\left(\tilde{\mathbf{V}}_{\mathbf{1}}(\gamma)\right)&\Rightarrow\lambda_{\mathbf{min}}\left(\frac{1}{\rho^2}\mathbf{I}_{\mathbf{N}^{\intercal}}-\tilde{\mathbf{V}}_{\mathbf{1}}(\gamma)\right)>0\\ &\Rightarrow\lambda_{\mathbf{min}}\left(\tilde{\mathbf{V}}_{\mathbf{2}}(\rho,\gamma)\right)>0,\end{aligned}$$

from which we get that $\tilde{\mathbf{V}}_{\mathbf{2}}(\rho,\gamma)$ is p.d.. Since $\tilde{\mathbf{V}}_{\mathbf{1}}(\gamma)$ is p.s.d when $\gamma_{\mathbf{j}}\succ 0$ for each $\mathbf{j}\in[K]$, we can then define its symmetric "square root" $\mathbf{R}\in\mathbb{R}^{N^{\intercal}\times N^{\intercal}}$ according to:

$$\mathbf{RR}=\tilde{\mathbf{V}}_{\mathbf{1}}(\gamma)=\tilde{\mathbf{G}}^T\tilde{\mathbf{G}}+\mathbf{B}^T\mathbf{B}\,^3.$$

Connecting $\mathbf{R}$ to $\mathbf{a}^*$ and the ASO-feasibility of (18), we have:

$$\mathbf{a}\text{ is ASO-feasible iff }\mathbf{Ra}=\mathbf{0}_{\mathbf{N}^{\intercal}}.\tag{30}$$

Sufficiency of (30) follows directly while necessity of (30) follows from:

$$\begin{aligned}&\mathbf{Ra}=\mathbf{0}_{\mathbf{N}^{\intercal}}\Rightarrow\mathbf{RRa}=\mathbf{0}_{\mathbf{N}^{\intercal}}\Rightarrow\tilde{\mathbf{G}}^T\tilde{\mathbf{G}}\mathbf{a}=-\mathbf{B}^T\mathbf{Ba}\\ &\Rightarrow\left\|\tilde{\mathbf{G}}\mathbf{a}\right\|_2^2=-\|\mathbf{Ba}\|_2^2\Rightarrow\tilde{\mathbf{G}}\mathbf{a}=\mathbf{0}_{\mathbf{N}^{\mathrm{c}}}\text{ and }\mathbf{Ba}=\mathbf{0}_{\mathbf{N}^{\mathrm{o}}}\\ &\Rightarrow\mathbf{a}\text{ is ASO-feasible.}\end{aligned}$$

Using (30), we see that the ASO of (18) is equivalent to the following *auxiliary standard optimization (XSO)* problem:

$$\min_{\rho\mathbf{Ra}=\mathbf{0}_{\tilde{N}}}\left\{\hat{f}(\mathbf{a})\right\},\tag{31}$$

whose necessary condition for optimality satisfies:

$$\mathbf{Rr}^*+\frac{1}{\rho}\mathbf{c}_{\hat{f}}^*=\mathbf{0}_{\mathbf{N}^{\intercal}},\tag{32}$$

[3] We note that such a *root matrix* exists for $\tilde{\mathbf{V}}_{\mathbf{1}}(\gamma)$ since the latter is square p.s.d. matrix. Specifically, if $\tilde{\mathbf{V}}_{\mathbf{1}}(\gamma)=\mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^T$ is a suitable eigenvalue decomposition, then $\mathbf{R}=\left[\mathbf{U}\boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{U}^T\right]$((Makhdoumi and Ozdaglar, 2017))

where $\mathbf{r}^*$ is the optimal dual variable, $\mathbf{a}^*$ is ASO-optimal, and $\mathbf{c}^*_{\hat{f}} \in \partial \hat{f}(\mathbf{a}^*)$. Since (31) is equivalent to (18), we have that the dual summands of the associated Lagrangians for both are equal for all $\tau > 0$ and, using (27) and (28), we have:

$$(\mathbf{r}[\tau])^T (\rho \mathbf{R}) \mathbf{a}[\tau] = (\boldsymbol{\mu}[\tau])^T \tilde{\mathbf{G}} \mathbf{a}[\tau] + (\boldsymbol{\nu}[\tau])^T \mathbf{B} \mathbf{a}[\tau] =$$
$$\left(\sum_{s=0}^{\tau} \mathbf{a}[s]\right)^T \left(\rho \tilde{\mathbf{V}}_{\mathbf{1}}(\gamma)\right) \mathbf{a}[\tau] = \left(\sum_{s=0}^{\tau} \mathbf{a}[s]\right)^T (\rho \mathbf{R}\mathbf{R}) \mathbf{a}[\tau]$$
$$= \left(\mathbf{R}\left[\sum_{s=0}^{\tau} \mathbf{a}[s]\right]\right)^T (\rho \mathbf{R}) \mathbf{a}[\tau].$$

from which we conclude that $\mathbf{r}[\tau] \triangleq \mathbf{R}\left(\sum_{s=0}^{\tau} \mathbf{a}[s]\right)$. Using (32) and noting that:

$$(\mathbf{a}^* - \mathbf{a}[\tau+1])^T \mathbf{c}^*_{\hat{f}} = (\mathbf{a}^* - \mathbf{a}[\tau+1])^T [-(\rho \mathbf{R}) \mathbf{r}^*]$$
$$= \rho (\mathbf{r}^*)^T \mathbf{R} \mathbf{a}[\tau+1] = \rho (\mathbf{r}^*)^T (\mathbf{r}[\tau+1] - \mathbf{r}[\tau]),$$

we can then exploit the monotonicity property of the subdifferential operator $\partial \hat{f}$ and use (29) to obtain:

$$0 \leq (\mathbf{a}[\tau+1] - \mathbf{a}^*)^T \left(\frac{1}{\rho}\mathbf{c}^{[\tau+1]}_{\hat{f}} - \frac{1}{\rho}\mathbf{c}^*_{\hat{f}}\right)$$
$$+ (\mathbf{r}^*)^T (\mathbf{r}[\tau+1] - \mathbf{r}[\tau])$$
$$\leq (\mathbf{a}[\tau+1] - \mathbf{a}[\tau])^T \tilde{\mathbf{V}}_2(\rho, \boldsymbol{\Gamma}) (\mathbf{a}^* - \mathbf{a}[\tau+1])$$
$$+ (\mathbf{r}[\tau+1] - \mathbf{r}[\tau])^T \mathbf{I}_{\mathbf{N}^\intercal} (\mathbf{r}^* - \mathbf{r}[\tau+1]). \quad (33)$$

It can be shown that (33) is equivalent to:

$$0 \leq \|\mathbf{q}[\tau] - \mathbf{q}^*\|^2_{\mathbf{W}} - \|\mathbf{q}[\tau+1] - \mathbf{q}^*\|^2_{\mathbf{W}}$$
$$- \|\mathbf{q}[\tau] - \mathbf{q}[\tau+1]\|^2_{\mathbf{W}}, \quad (34)$$

where:

$$\mathbf{W} = \begin{bmatrix} \tilde{\mathbf{V}}_{\mathbf{2}}(\rho, \gamma) & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{\mathbf{N}^\intercal} \end{bmatrix}$$
$$\mathbf{q}[\tau] = [\mathbf{a}[\tau]; \mathbf{r}[\tau]]$$
$$\mathbf{q}^* = [\mathbf{a}^*; \mathbf{r}^*].$$

From (34), we then have, for any $\mathbf{q}[\tau+1] \neq \mathbf{q}[\tau]$:

$$\|\mathbf{q}[\tau] - \mathbf{q}^*\|^2_{\mathbf{W}} \geq$$
$$\|\mathbf{q}[\tau+1] - \mathbf{q}^*\|^2_{\mathbf{W}} + \|\mathbf{q}[\tau] - \mathbf{q}[\tau+1]\|^2_{\mathbf{W}}$$
$$> \|\mathbf{q}[\tau+1] - \mathbf{q}^*\|^2_{\mathbf{W}}. \quad (35)$$

where we used that $\mathbf{W}$ is p.d.[4] From (35), we have that $\left\{\|\mathbf{q}[\tau] - \mathbf{q}^*\|^2_{\mathbf{W}}\right\}_{\tau>0}$ is monotonically decreasing. Since this sequence has a lower bound of 0, we can thus conclude that it converges to a fixed value:

$$\|\mathbf{q}[\tau] - \mathbf{q}^*\|^2_{\mathbf{W}} \underset{\tau\to\infty}{\to} \zeta \in \mathbf{R}_+. \quad (36)$$

From (36) we have that $\{\mathbf{q}[\tau]\}_{\tau\geq 0}$ deterministically converges to a vector $\overline{\mathbf{q}}$ which is a distance ζ away from $\mathbf{q}^*$. Due to the nature of PAC, we can then conclude that $\overline{\mathbf{q}}$ is a fixed point of PAC and also satisfies the KKT Conditions of the XSO (31). By Assumption 4.1, $\hat{f}$ is convex and $\mathrm{dom}\left(\hat{f}\right) = \mathbb{R}^N$. Utilizing this along with Assumption 4.2, we can then conclude that Slater's Condition is satisfied. Hence, $\overline{\mathbf{q}}$ is an optimal point of the XSO (31). Since our choice of choice $\mathbf{q}^*$ was initially arbitrary in the sense that it only had to be an (not necessarily the) optimal solution of the XSO, we can then simply set $\mathbf{q}^* = \overline{\mathbf{q}}$ and therefore conclude that the PAC trajectory obtains asymptotic convergence to an optimal XSO solution:

$$\lim_{\tau\to\infty} \{\mathbf{q}[\tau]\} = \mathbf{q}^* \Rightarrow \lim_{\tau\to\infty} \{\mathbf{a}[\tau]\} = \mathbf{a}^*.$$

As for the rate, we can follow the approach of Lemma 2.2 of ((Deng et al., 2017)) and use the monotonicity of $\partial \hat{f}$ and the first-order condition of optimality to get, if $\Delta\mathbf{a}[\tau+1] \triangleq \mathbf{a}[\tau+1] - \mathbf{a}[\tau]$:

$$0 \leq (\Delta\mathbf{a}[\tau+1])^T \left(\frac{1}{\rho}\mathbf{c}^{[\tau+1]}_{\hat{f}} - \frac{1}{\rho}\mathbf{c}^{[\tau]}_{\hat{f}}\right)$$
$$\leq -\frac{1}{2}\|\Delta\mathbf{q}[\tau+1]\|^2_{\mathbf{W}} + \frac{1}{2}\|\Delta\mathbf{q}[\tau]\|^2_{\mathbf{W}}, \quad (37)$$

where we used the identity for $\mathbf{Z} \in \mathbb{R}^{n\times n}$ and $\mathbf{z_1}, \mathbf{z_2} \in \mathbb{R}^n$:

$$2(\mathbf{z_1})^T \mathbf{Z}(\mathbf{z_2}) \leq \|\mathbf{z_1}\|^2_{\mathbf{Z}} + \|\mathbf{z_2}\|^2_{\mathbf{Z}}.$$

From (37) we can thus conclude that $\|\Delta\mathbf{q}[\tau]\|^2_{\mathbf{W}}$ monotonically non-increasing in τ, i.e.:

$$\|\Delta\mathbf{q}[\tau+1]\|^2_{\mathbf{W}} \leq \|\Delta\mathbf{q}[\tau]\|^2_{\mathbf{W}}.$$

Since $\lim_{\tau\to\infty} \{\mathbf{q}[\tau]\} = \mathbf{q}^*$, we can conclude using (34) that $\sum_{\tau=1}^{\infty} \|\Delta\mathbf{q}[\tau]\|^2_{\mathbf{W}} < 0$. Hence, we use Lemma 1.1 of ((Deng et al., 2017)) to determine that PAC converges with $\|\Delta\mathbf{q}[\tau+1]\|^2_{\mathbf{W}} = o\left(\frac{1}{\tau}\right)$. ■

5. CASE STUDY APPLICATION

The proposed relaxed CI-OPF formulation is solved with the proposed PAC algorithm, using a modified IEEE 13 bus network (see Figure 1a). To test the application of the proposed methods to modern distribution grids, DERs are introduced into the network. Demand response is introduced at nodes 3, 6, and 11, for reduction in real power load of up to 20%, 10%, and 30% respectively. A dispatchable DG is introduced at node 9, with $\overline{P} = 100\text{kW}, \underline{P} = 0\text{kW}$ and $\overline{Q} = 50\text{kVAr}, \underline{Q} = -50\text{kVAr}$. The capacitor banks in the original IEEE network are modelled as reactive power generators.

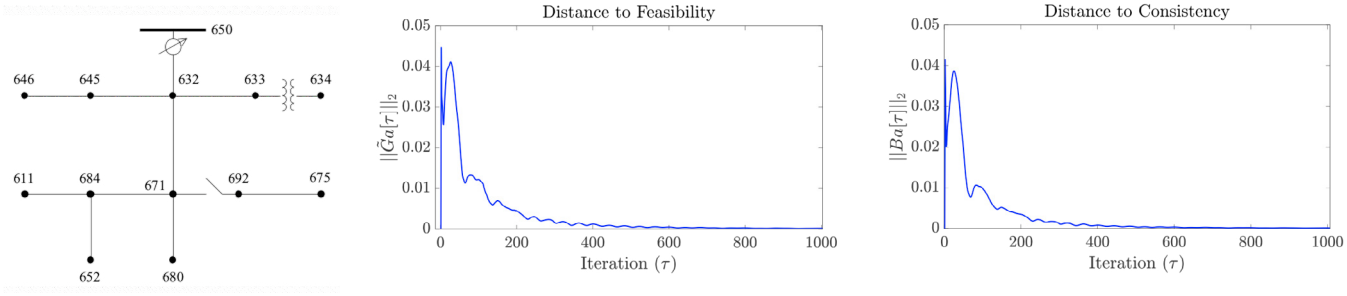
We consider the problem of voltage regulation for this active distribution network. The objective function is as follows:

$$f(\mathbf{i}, \mathbf{v}, \mathbf{s}) = \sum_{i\in N}\sum_{k\in K}[(v^R_{ik} - \hat{v}^R_{ik})^2 + (v^I_{ik} - \hat{v}^I_{ik})^2] \quad (38)$$

where v^R_{ik} and v^I_{ik} are the real and imaginary part of the voltage phasor v_{ik}, and $\hat{v}^R_{ik}$ and $\hat{v}^I_{ik}$ are the desired setpoints treated as the reference value. In our case study we take $\hat{v}^R_{ik} = 1$ and $\hat{v}^I_{ik} = 0$, $\forall i \in N \setminus i = 1, k \in K$. We note that node 1 which is the point of common coupling (PCC) to the transmission grid is treated as a slack node, with $v^R_1 = 1$ and $v^I_1 = 0$, $\forall k \in K$.

We then use the PAC algorithm to solve the OPF problem of (7)-(15), with the above objective function. All simulations were performed using a 2.3 GHz Intel Core i7 with MATLAB and the YALMIP interface (Löfberg, 2004). The chosen PAC parameters are $\gamma = 1.730$ and $\rho = 0.03035$. In Figure 1b-1c, we evaluate the algorithm's performances as measured by two different metrics. These are the distance to global feasibility: $\left\|\tilde{\mathbf{G}}\mathbf{a}[\tau]\right\|_2$; and the distance to coordination: $\|\mathbf{B}\mathbf{a}[\tau]\|_2$. These metrics describe how feasible the

[4] This follows since $\tilde{\mathbf{V}}_{\mathbf{2}}(\rho, \gamma)$ is p.d..



(a) IEEE13 network topology (b) Convergence of feasibility metric (c) Convergence of consistency metric

Fig. 1. Results of PAC employed on CI-OPF model of IEEE 13 bus network

solution is (whether power flow equations are satisfied), and how far the variable copies are from their true values (which were introduced to distribute the problem). Results show the algorithm converges to the optimal solution in less than 800 iterations. The maximum and average time required for each primal update for PAC is 0.0096s and 0.0023s, respectively.

6. CONCLUSIONS AND FUTURE WORK

In this paper we define a new model and method for solving the OPF problem for distribution grids with meshed topology and unbalanced structure. First, we propose a new MCE-based convex relaxation for the OPF problem using current injections. Second, we state a distributed optimization algorithm and its related decomposition profile to solve the OPF in a fully distributed manner, with provable convergence rate $o(1/\tau)$ (where τ is the number of iterations) for convex problems. The developed approach was tested on a modified IEEE 13 bus network, with added DERs and lines to create a meshed topology. We show convergence of the PAC algorithm to the optimal solution in under 800 iterations. Future work will include the comparison of the proposed model with alternate power systems models. We will also focus on the development of new methodologies to improve the MCE relaxation to obtain a higher quality lower bound of the OPF solution, and further tuning of the PAC parameters for faster convergence result. Research will also revisit the PAC algorithm to reduce number of iterations required for convergence, focusing on the primal and dual updates.

REFERENCES

Chen, G. and Teboulle, M. (1994). A proximal-based decomposition method for convex minimization problems. *Mathematical Programming*, 64(1), 81–101. doi:10.1007/BF01582566. URL `https://doi.org/10.1007/BF01582566`.

Christakou, K., Tomozei, D.C., Boudec, J.Y.L., and Paolone, M. (2017). Ac opf in radial distribution networks part 1: On the limits of the branch flow convexification and the alternating direction method of multipliers. *Electric Power Systems Research*, 143, 438–450. doi:10.1016/j.epsr.2016.07.030.

Dall'Anese, E., Zhu, H., and Giannakis, G.B. (2013). Distributed optimal power flow for smart microgrids. *IEEE Transactions on Smart Grid*, 4(3), 1464–1475. doi:10.1109/tsg.2013.2248175.

Deng, W., Lai, M.J., Peng, Z., and Yin, W. (2017). Parallel multi-block admm with o(1 / k) convergence. *Journal of Scientific Computing*, 71(2), 712–736. doi:10.1007/s10915-016-0318-2. URL `https://doi.org/10.1007/s10915-016-0318-2`.

Dommel, H.W. and Tinney, W.F. (1968). Optimal power flow solutions. *IEEE Transactions on Power Apparatus and Systems*, PAS-87(10), 1866–1876. doi:10.1109/TPAS.1968.292150.

Farivar, M., Clarke, C.R., Low, S.H., and Chandy, K.M. (2011). Inverter var control for distribution systems with renewables. *2011 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. doi:10.1109/smartgridcomm.2011.6102366.

Farivar, M. and Low, S. (2014). Branch flow model: Relaxations and convexification. *2014 IEEE PES TD Conference and Exposition*. doi:10.1109/tdc.2014.6863260.

Kersting, W.H. (2006). *Distribution system modeling and analysis*. CRC press.

Lavaei, J. and Low, S.H. (2012). Zero duality gap in optimal power flow problem. *IEEE Transactions on Power Systems*, 27(1), 92–107. doi:10.1109/tpwrs.2011.2160974.

Li, N., Chen, L., and Low, S.H. (2012). Demand response in radial distribution networks: Distributed algorithm. In *2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, 1549–1553. IEEE.

Löfberg, J. (2004). Yalmip : A toolbox for modeling and optimization in matlab. In *In Proceedings of the CACSD Conference*. Taipei, Taiwan.

Makhdoumi, A. and Ozdaglar, A. (2017). Convergence rate of distributed admm over networks. *IEEE Transactions on Automatic Control*, 62(10), 5082–5095. doi:10.1109/TAC.2017.2677879.

Mccormick, G.P. (1976). Computability of global solutions to factorable nonconvex programs: Part i — convex underestimating problems. *Mathematical Programming*, 10(1), 147–175. doi:10.1007/bf01580665.

Robbins, B.A. and Dominguez-Garcia, A.D. (2016). Optimal reactive power dispatch for voltage regulation in unbalanced distribution systems. *IEEE Transactions on Power Systems*, 31(4), 2903–2913. doi:10.1109/tpwrs.2015.2451519.

Romvary, J. (2018). *A proximal atomic coordination algorithm for distributed optimization in distribution grids*. Ph.D. thesis, Massachusetts Institute of Technology.