Component-Based Dual Decomposition and ADMM in the OPF Problem

Sleiman Mhanna
School of Electrical and Information Engineering
The University of Sydney
Sydney, Australia
Email: sleiman.mhanna@sydney.edu.au

Gregor Verbič
School of Electrical and Information Engineering
The University of Sydney
Sydney, Australia
Email: gregor.verbic@sydney.edu.au

Archie C. Chapman
School of Electrical and Information Engineering
The University of Sydney
Sydney, Australia
Email: archie.chapman@sydney.edu.au

***Abstract*—This paper proposes a component-based dual decomposition of the nonconvex AC optimal power flow (OPF) problem, where the modified dual function is solved using the alternating direction method of multipliers (ADMM). This paper is the first to conduct extensive numerical analysis resulting in the identification and tabulation of the algorithmic parameter settings that are crucial for the convergence of ADMM to near-optimal (possibly globally optimal) solutions on a vast array of test instances, despite the nonconvexity of the OPF problem. Moreover, this work provides a deeper insight into the geometry of the modified Lagrange dual function of the OPF problem and highlights the conditions that make this function differentiable.**

Notation

A. Input data and operators

$\mathcal{B}$	Set of buses in the power network.
$\mathcal{B}_i$	Set of buses connected to bus i.
b_i^{sh}	Shunt susceptance (p.u.) at bus i.
g_i^{sh}	Shunt conductance (p.u.) at bus i.
b_{ij}^{ch}	Charging susceptance (p.u.) in the π-model of line ij.
$c0_i^g$	Constant coefficient (\$) term of generator g's cost function.
$c1_i^g$	Coefficient ($\$/\text{MW}$) of the linear term of generator g's cost function.
$c2_i^g$	Coefficient ($\$/\text{MW}^2$) of the quadratic term of generator g's cost function.
$\mathcal{G}$	Set of all generators (g, i) in the power network such that g is the generator and i is the bus connected to it.
$\mathcal{G}_i$	Set of all generators connected to bus i.
i	Imaginary unit.
$\mathcal{L}$	Set of all transmission lines ij where i is the "from" bus.
$\mathcal{L}_t$	Set of all transmission lines ij where i is the "to" bus.
$p_i^{\text{d}}/q_i^{\text{d}}$	Active/reactive power demand (MW/MVAr) at bus i.
$\overline{s}_{ij}$	Apparent power rating (MVA) of line ij.
$\underline{\theta}_{ij}^{\Delta}$	Lower limit of the difference of voltage angles of buses i and j.
$\overline{\theta}_{ij}^{\Delta}$	Upper limit of the difference of voltage angles of buses i and j.
θ_i^{shift}	Phase shift (Radians) of phase shifting transformer connected between buses i and j ($\theta_i^{\text{shift}} = 0$ for a transmission line).
τ_{ij}	Tap ratio magnitude of phase shifting transformer connected between buses i and j ($\tau_{ij} = 1$ for a transmission line).
T_{ij}	Complex tap ratio of a phase shifting transformer ($T_{ij} = \tau_{ij}\mathrm{e}^{\mathrm{i}\theta_i^{\text{shift}}}$).
Y_{ij}	Series admittance (p.u.) in the π-model of line ij.
$\Im\{\bullet\}$	Imaginary value operator.
$\Re\{\bullet\}$	Real value operator.
$\underline{\bullet}/\overline{\bullet}$	Minimum/maximum magnitude operator.
$\lvert\bullet\rvert$	Magnitude operator/Cardinality of a set.
$\bullet^*$	Conjugate operator.
k	Iteration number.
ρ	ADMM penalty parameter.

B. Decision variables

p_i^g/q_i^g	Active/reactive power (MW/MVAr) generation of generator g at bus i.
$p_{i_{(i)}}^g$	Duplicate of p_i^g at bus i.
$q_{i_{(i)}}^g$	Duplicate of q_i^g at bus i.
p_{ij}/q_{ij}	Active/reactive power (MW/MVAr) flow along transmission line ij.
$p_{ij_{(i)}}$	Duplicate of p_{ij} at bus i.
$q_{ij_{(i)}}$	Duplicate of q_{ij} at bus i.
V_i	Complex phasor voltage (p.u.) at bus i ($V_i = \lvert V_i\rvert \angle\theta_i = v_i\angle\theta_i$).
$v_{i_{(ij)}}$	Duplicate of v_i at line ij such that $j \in \mathcal{B}_i$.
$\theta_{i_{(ij)}}$	Duplicate of θ_i at line ij such that $j \in \mathcal{B}_i$.
$\boldsymbol{\lambda}$	Vector of Lagrange multipliers.

C. Acronyms

AC	Alternating current.
ADMM	Alternating direction method of multipliers.
GNLP	Global nonlinear programming.
KKT	*Karush-Kuhn-Tucker* conditions.
NLP	Nonlinear programming.
OCD	Optimality conditions decomposition.

OPF Optimal power flow.

I. Introduction

The alternating current (AC) power flow equations, which model the steady-state physics of power flows, are the linchpins of a broad spectrum of optimization problems in electrical power systems. Unfortunately, these nonlinear equations are the main sources of nonconvexity, which makes these problems notorious for being extremely challenging to solve using global nonlinear programming (GNLP) solvers. Therefore, the research community has focused on improving interior-point-based nonlinear optimization methods (IPM) to compute feasible solutions efficiently [1], [2]. Although these methods only (theoretically) guarantee local optimality, they have been shown, thanks to tight convex relaxations [3]–[7], to reach near-optimal (possibly globally optimal) solutions on all the known test cases in the literature. This paper capitalizes on this to numerically show that the proposed distributed method solves the modified dual problem of the nonconvex AC OPF problem to near optimality, possibly to global optimality, on all the considered test cases.

A. State-of-the-art

There is a plethora of existing works on distributed OPF. These can be broadly classified into three categories, dual decomposition methods [8]–[17], optimality conditions decomposition (OCD) methods [18]–[22] and sparse SDP decomposition methods [23], [24]. The dual decomposition techniques underlying the dual-decomposition-based distributed OPF methods in the literature can in turn be classified into two categories: region-based decompositions [8]–[13], [16], [25],[1] and component-based decompositions [14], [15], [17]. The focus of this study revolves around the latter decomposition techniques because they preserve privacy with respect to *all* components (generators, transformers, loads, buses, lines etc.) and are flexible enough to incorporate discrete decision variables to suit a wide variety of optimization applications in power system operations such as optimal transmission switching, capacitor placement, transmission and distribution network expansion planning, optimal feeder reconfiguration, power system restoration, and vulnerability analysis, to name a few. On the downside, dual-decomposition-based AC OPF methods have no theoretical guarantee of convergence because the (primal) OPF problem is nonconvex. Nonetheless, this paper numerically shows that under the right conditions, the proposed distributed method can converge to near-optimal (possibly globally optimal) solutions. Unlike [14]–[16], [23], [24], [24], which solve a convexified version of the OPF problem, this paper tackles the nonlinear nonconvex AC OPF directly. Convex relaxations are appealing because they are computationally conducive but their main disadvantage is that they do not always yield feasible solutions. Furthermore, in contrast to [17], the work in this paper conducts extensive numerical analysis and specifies the algorithmic parameter settings that are crucial for the convergence of the proposed component-based dual decomposition method on a vast array of test instances. On the other hand, OCD methods [18]–[22] rely on matrix factorization [26] to parallelize the computation of the *Karush-Kuhn-Tucker* (KKT) conditions. However, as of yet, these methods are not amenable to decompositions in terms of components.

[1]Note that the OPF problem in [12] and [25] is decomposed in terms of buses, which can be thought of as the maximum number of regions in a power network.

B. Contributions of this work

Against this background, this paper is the first to conduct extensive numerical analysis on the application of ADMM to solve the augmented Lagrange dual function of the AC OPF problem. The paper also numerically shows that the proposed algorithm converges to the same solutions obtained from the centralized IPM. In more detail, this paper advances the state of the art in the following ways:

- Extensive numerical simulations on 72 test cases from MATPOWER [2], PEGASE [27] and NESTA v6 [28] instances show that the proposed algorithm converges to the same solutions obtained from the centralized IPMs.
- The algorithmic parameter settings that are crucial for convergence are identified and tabulated.
- A deeper insight into the geometry of the modified Lagrange dual function of the OPF problem shows that this function can be nonsmooth for small values of the ADMM penalty parameter.

C. Notation

All vectors are column vectors unless otherwise specified, and $\mathbf{1}$ is an all-ones vector of length depending on the context. The inner product of two vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^n$ is delineated by $\langle \boldsymbol{x}, \boldsymbol{y} \rangle := \boldsymbol{x}^T \boldsymbol{y}$, where $\boldsymbol{x}^T$ is the transpose of $\boldsymbol{x}$. The Euclidean norm of a vector $\boldsymbol{x} \in \mathbf{R}^n$ is denoted by $\|\boldsymbol{x}\| := \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$ and the nonnegative orthant in $\mathbf{R}^n$ is denoted by $\mathbf{R}^n_+$.

II. The OPF problem

In a power network, the OPF problem consists of finding the least-cost dispatch of power from generators to satisfy the load at all buses in a way that is governed by physical laws, such as Ohm's Law and Kirchhoff's Law, and other technical restrictions, such as transmission line thermal limit constraints. Knowing that $\Re\left\{V_i V_j^*\right\} := v_i v_j \cos\left(\theta_i - \theta_j\right)$ and $\Im\left\{V_i V_j^*\right\} := v_i v_j \sin\left(\theta_i - \theta_j\right)$, the OPF problem in *polar form* can be written as

$$\underset{\substack{p_i^g, q_i^g, v_i, \theta_i, \\ p_{ij}, q_{ij}, p_{ji}, q_{ji}}}{\text{minimize}} \sum_{(g,i)\in\mathcal{G}} f_i^g\left(p_i^g\right) \tag{1a}$$

subject to

$$\underline{p}_i^g \le p_i^g \le \overline{p}_i^g, \qquad (g,i)\in\mathcal{G} \tag{1b}$$

$$\underline{q}_i^g \le q_i^g \le \overline{q}_i^g, \qquad (g,i)\in\mathcal{G} \tag{1c}$$

$$\underline{v}_i \le v_i \le \overline{v}_i, \qquad i\in\mathcal{B} \tag{1d}$$

$$\underline{\theta}_{ij}^{\Delta} \le \theta_i - \theta_j \le \overline{\theta}_{ij}^{\Delta}, \qquad (i,j)\in\mathcal{L} \tag{1e}$$

$$\sum_{(g,i)\in\mathcal{G}} p_i^g - p_i^{\mathrm{d}} = \sum_{j\in\mathcal{B}_i} p_{ij} + g_i^{\mathrm{sh}} v_i^2, \qquad i\in\mathcal{B} \tag{1f}$$

$$\sum_{(g,i)\in\mathcal{G}} q_i^g - q_i^{\mathrm{d}} = \sum_{j\in\mathcal{B}_i} q_{ij} - b_i^{\mathrm{sh}} v_i^2, \qquad i\in\mathcal{B} \tag{1g}$$

$$p_{ij} = g_{ij}^c v_i^2 - g_{ij} v_i v_j \cos(\theta_i - \theta_j) + b_{ij} v_i v_j \sin(\theta_i - \theta_j), \qquad (i,j)\in\mathcal{L} \tag{1h}$$

$$q_{ij} = b_{ij}^c v_i^2 - b_{ij} v_i v_j \cos(\theta_i - \theta_j) - g_{ij} v_i v_j \sin(\theta_i - \theta_j), \qquad (i,j)\in\mathcal{L} \tag{1i}$$

$$p_{ji} = g_{ji}^c v_j^2 - g_{ji} v_j v_i \cos(\theta_j - \theta_i) + b_{ji} v_j v_i \sin(\theta_j - \theta_i), \qquad (i,j)\in\mathcal{L} \tag{1j}$$

$$q_{ji} = b_{ji}^c v_j^2 - b_{ji} v_j v_i \cos(\theta_j - \theta_i) - g_{ji} v_j v_i \sin(\theta_j - \theta_i), \qquad (i,j)\in\mathcal{L} \tag{1k}$$

$$\sqrt{p_{ij}^2 + q_{ij}^2} \le \overline{s}_{ij}, \qquad (i,j)\in\mathcal{L}\cup\mathcal{L}_t \tag{1l}$$

where, $g_{ij}^c := \Re\left\{\frac{Y_{ij}^* - \mathrm{i}\frac{b_{ij}^{\mathrm{ch}}}{2}}{|T_{ij}|^2}\right\}$, $b_{ij}^c := \Im\left\{\frac{Y_{ij}^* - \mathrm{i}\frac{b_{ij}^{\mathrm{ch}}}{2}}{|T_{ij}|^2}\right\}$, $g_{ij} := \Re\left\{\frac{Y_{ij}^*}{T_{ij}}\right\}$, $b_{ij} := \Im\left\{\frac{Y_{ij}^*}{T_{ij}}\right\}$, $g_{ji}^c := \Re\left\{Y_{ji}^* - \mathrm{i}\frac{b_{ji}^{\mathrm{ch}}}{2}\right\}$, $b_{ji}^c := \Im\left\{Y_{ji}^* - \mathrm{i}\frac{b_{ji}^{\mathrm{ch}}}{2}\right\}$, $g_{ji} := \Re\left\{\frac{Y_{ji}^*}{T_{ji}^*}\right\}$ and $b_{ji} := \Im\left\{\frac{Y_{ji}^*}{T_{ji}^*}\right\}$, and $f_i^g(p_i^g) := c2_i^g(p_i^g)^2 + c1_i^g(p_i^g) + c0_i^g$. The OPF in (1) is a nonconvex nonlinear optimization problem that is proven to be NP-hard [29]. The nonconvexities stem from equality constraints (1f)–(1k), which include nonconvex voltage bilinear terms multiplied by nonconvex sine and cosine functions of the angles, and a quadratic function of the voltage, which is also nonconvex in this equality constraint setting as it describes the boundary of the set $\{v^2 | v\in[\underline{v},\overline{v}]\}$.[2]

III. Component-based dual decomposition

The OPF problem in its native form in (1) is not separable in terms of components. Moreover, relaxing the coupling constraints in (1f) and (1g) is not enough to bestow a component-based separability. Towards this aim, the following variables are duplicated

$$p_i^g = p_{i_{(i)}}^g, \qquad (g,i)\in\mathcal{G}, \tag{2}$$

$$q_i^g = q_{i_{(i)}}^g, \qquad (g,i)\in\mathcal{G}, \tag{3}$$

$$p_{ij} = p_{ij_{(i)}}, \qquad (i,j)\in\mathcal{L}\cup\mathcal{L}_t, \tag{4}$$

$$q_{ij} = q_{ij_{(i)}}, \qquad (i,j)\in\mathcal{L}\cup\mathcal{L}_t, \tag{5}$$

$$v_{i_{(ij)}} = v_i, \qquad (i,j)\in\mathcal{L}\cup\mathcal{L}_t, \tag{6}$$

$$\theta_{i_{(ij)}} = \theta_i, \qquad (i,j)\in\mathcal{L}\cup\mathcal{L}_t, \tag{7}$$

and the OPF problem now becomes

$$\underset{\boldsymbol{x},\boldsymbol{z}}{\text{minimize}} \sum_{(g,i)\in\mathcal{G}} f_i^g(p_i^g) \tag{8a}$$

$$\text{subject to (1b), (1c), (1l), (2)–(7)} \tag{8b}$$

$$\underline{v}_i \le v_{i_{(ij)}} \le \overline{v}_i, \qquad (i,j)\in\mathcal{L}\cup\mathcal{L}_t \tag{8c}$$

$$\underline{\theta}_{ij}^{\Delta} \le \theta_{i_{(ij)}} - \theta_{j_{(ji)}} \le \overline{\theta}_{ij}^{\Delta}, \qquad (i,j)\in\mathcal{L} \tag{8d}$$

$$\sum_{(g,i)\in\mathcal{G}} p_{i_{(i)}}^g - p_i^{\mathrm{d}} = \sum_{j\in\mathcal{B}_i} p_{ij_{(i)}} + g_i^{\mathrm{sh}} v_i^2, \qquad i\in\mathcal{B} \tag{8e}$$

$$\sum_{(g,i)\in\mathcal{G}} q_{i_{(i)}}^g - q_i^{\mathrm{d}} = \sum_{j\in\mathcal{B}_i} q_{ij_{(i)}} - b_i^{\mathrm{sh}} v_i^2, \qquad i\in\mathcal{B} \tag{8f}$$

$$p_{ij} = g_{ij}^c v_{i_{(ij)}}^2 - g_{ij} v_{i_{(ij)}} v_{j_{(ji)}} \cos\left(\theta_{i_{(ij)}} - \theta_{j_{(ji)}}\right) + b_{ij} v_{i_{(ij)}} v_{j_{(ji)}} \sin\left(\theta_{i_{(ij)}} - \theta_{j_{(ji)}}\right), \qquad (i,j)\in\mathcal{L} \tag{8g}$$

$$q_{ij} = b_{ij}^c v_{i_{(ij)}}^2 - b_{ij} v_{i_{(ij)}} v_{j_{(ji)}} \cos\left(\theta_{i_{(ij)}} - \theta_{j_{(ji)}}\right) - g_{ij} v_{i_{(ij)}} v_{j_{(ji)}} \sin\left(\theta_{i_{(ij)}} - \theta_{j_{(ji)}}\right), \qquad (i,j)\in\mathcal{L} \tag{8h}$$

$$p_{ji} = g_{ji}^c v_{j_{(ji)}}^2 - g_{ji} v_{j_{(ji)}} v_{i_{(ij)}} \cos\left(\theta_{j_{(ji)}} - \theta_{i_{(ij)}}\right) + b_{ji} v_{j_{(ji)}} v_{i_{(ij)}} \sin\left(\theta_{j_{(ji)}} - \theta_{i_{(ij)}}\right), \qquad (i,j)\in\mathcal{L} \tag{8i}$$

$$q_{ji} = b_{ji}^c v_{j_{(ji)}}^2 - b_{ji} v_{j_{(ji)}} v_{i_{(ij)}} \cos\left(\theta_{j_{(ji)}} - \theta_{i_{(ij)}}\right) - g_{ji} v_{j_{(ji)}} v_{i_{(ij)}} \sin\left(\theta_{j_{(ji)}} - \theta_{i_{(ij)}}\right), \qquad (i,j)\in\mathcal{L} \tag{8j}$$

where

$$\boldsymbol{x} := \left[\left(p_i^g, q_i^g\right)_{(g,i)\in\mathcal{G}}, \left(p_{ij}, q_{ij}, v_{i_{(ij)}}, \theta_{i_{(ij)}}\right)_{(i,j)\in\mathcal{L}\cup\mathcal{L}_t}\right],$$

and

$$\boldsymbol{z} := \left[\left(p_{i_{(i)}}^g, q_{i_{(i)}}^g\right)_{(g,i)\in\mathcal{G}}, \left(p_{ij_{(i)}}, q_{ij_{(i)}}, v_i, \theta_i\right)_{(i,j)\in\mathcal{L}\cup\mathcal{L}_t}\right].$$

Now, by relaxing the consensus constraints (2)–(7), the (partial) Lagrangian function of problem (8) is given by

$$L(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\lambda}) := \sum_{(g,i)\in\mathcal{G}} \left(f_i^g(p_i^g)\right) + \langle\boldsymbol{\lambda}, \boldsymbol{x}-\boldsymbol{z}\rangle, \tag{9}$$

where

$$\boldsymbol{\lambda} := \left[\left(\lambda_{p,i}^g, \lambda_{q,i}^g\right)_{(g,i)\in\mathcal{G}}, \left(\lambda_{p_{ij}}, \lambda_{q_{ij}}, \lambda_{v_{ij}}, \lambda_{\theta_{ij}}\right)_{(i,j)\in\mathcal{L}\cup\mathcal{L}_t}\right],$$

is the vector of Lagrange multipliers associated with consensus constraints (2)–(7). Accordingly, the Lagrange dual function is

$$D(\boldsymbol{\lambda}) := \underset{\boldsymbol{x},\boldsymbol{z}}{\text{minimize}}\ L(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\lambda}), \quad \text{subject to (1b), (1c), (1l), (8c)–(8j).} \tag{10}$$

The dual can now be decomposed in terms of components as follows

$$D(\boldsymbol{\lambda}) := \sum_{(g,i)\in\mathcal{G}} D_i^g(\boldsymbol{\lambda}_i^g) + \sum_{i\in\mathcal{B}} D_i^{\mathrm{b}}\left((\boldsymbol{\lambda}_i^g)_{g\in\mathcal{G}_i}, (\boldsymbol{\lambda}_{ij})_{j\in\mathcal{B}_i}\right) + \sum_{(i,j)\in\mathcal{L}} D_{ij}^{\mathrm{l}}(\boldsymbol{\lambda}_{ij}, \boldsymbol{\lambda}_{ji}), \tag{11}$$

where $\boldsymbol{\lambda}_i^g := \left[\lambda_{p,i}^g, \lambda_{q,i}^g\right]$, $\boldsymbol{\lambda}_{ij} := \left[\lambda_{p_{ij}}, \lambda_{q_{ij}}, \lambda_{v_{ij}}, \lambda_{\theta_{ij}}\right]$ and $\boldsymbol{\lambda}_{ji} := \left[\lambda_{p_{ji}}, \lambda_{q_{ji}}, \lambda_{v_{ji}}, \lambda_{\theta_{ji}}\right]$. Finally, the Lagrange dual problem is given by

$$\max_{\boldsymbol{\lambda}} \quad D(\boldsymbol{\lambda}). \tag{12}$$

The main reasons for solving the Lagrange dual function (10) instead of the primal (8) is that, first, the former is the pointwise infimum of a family of affine functions in $\boldsymbol{\lambda}$ and is therefore concave, even though the primal problem (8) is nonconvex. Subsequently, first-order methods from convex optimization can be applied to solve the dual. Second, the dual is separable in terms of components and can therefore be solved in a distributed fashion, thus preserving privacy. However, in this case, since the objective functions in (11) are neither finite nor strictly convex (and also the Lagrange function in (9) is unbounded below in $\boldsymbol{x}$ and $\boldsymbol{z}$), the dual function in (9) is unbounded.

[2] The method in this paper was also applied to the OPF in *rectangular form* but the results are not documented here because they were not significantly different than the polar form ones.

IV. MODIFIED DUAL FUNCTION AND ADMM

To make the Lagrangian function finite and strictly convex, it is modified as follows

$$L_\rho\left(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\lambda}\right) := L\left(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\lambda}\right) + \frac{\rho}{2}\left\|\boldsymbol{x}-\boldsymbol{z}\right\|^2, \tag{13}$$

which is also known as the augmented Lagrange function, and the augmented Lagrange dual function would be

$$D_\rho\left(\boldsymbol{\lambda}\right) := \underset{\boldsymbol{x},\boldsymbol{z}}{\text{minimize}}\ L_\rho\left(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\lambda}\right) \tag{14a}$$

$$\text{subject to (1b), (1c), (1l), (8c)–(8j).} \tag{14b}$$

To preserve the separability of the problem, ADMM is used to approximate (14). More specifically, generators now solve

$$D^g_{\rho,i}\left(\boldsymbol{\lambda}_i^{g,k}\right) = \underset{\boldsymbol{x}_i^g}{\text{minimize}} \sum_{g\in\mathcal{G}_i}\Bigg(f_i^g\left(p_i^g\right) + \left\langle \boldsymbol{\lambda}_i^{g,k}, \boldsymbol{x}_i^g\right\rangle + \frac{\rho_{pq}}{2}\left(\left(p_i^g - p_{i_{(i)}}^{g,k}\right)^2 + \left(q_i^g - q_{i_{(i)}}^{g,k}\right)^2\right)\Bigg) \tag{15a}$$

$$\text{subject to (1b), (1c),} \tag{15b}$$

where $\boldsymbol{x}_i^g := \left[p_i^g, q_i^g\right]$, and transmission lines solve

$$D^{\mathrm{l}}_{\rho,ij}\left(\boldsymbol{\lambda}_{ij}^k, \boldsymbol{\lambda}_{ji}^k\right) = \underset{\boldsymbol{x}_{ij}^{\mathrm{l}}}{\text{minimize}}\ \Bigg\{\left\langle\left[\boldsymbol{\lambda}_{ij}^k, \boldsymbol{\lambda}_{ji}^k\right], \boldsymbol{x}_{ij}^{\mathrm{l}}\right\rangle + \sum_{(l,m)\in\{(i,j)\cup(j,i)\}}\Bigg(\frac{\rho_{v\theta}}{2}\left(\left(v_l^k - v_{l_{(lm)}}\right)^2 + \left(\theta_l^k - \theta_{l_{(lm)}}\right)^2\right) + \frac{\rho_{pq}}{2}\left(\left(p_{lm} - p_{lm_{(l)}}^k\right)^2 + \left(q_{lm} - q_{lm_{(l)}}^k\right)^2\right)\Bigg)\Bigg\} \tag{16a}$$

$$\text{subject to}\quad \underline{v}_i \le v_{i_{(ij)}} \le \overline{v}_i,\ \underline{v}_j \le v_{j_{(ji)}} \le \overline{v}_j, \tag{16b}$$

$$\underline{\theta}_{ij}^{\Delta} \le \theta_{i_{(ij)}} - \theta_{j_{(ji)}} \le \overline{\theta}_{ij}^{\Delta}, \tag{16c}$$

$$p_{ij} = g_{ij}^c v_{i_{(ij)}}^2 - g_{ij} v_{i_{(ij)}} v_{j_{(ji)}} \cos\left(\theta_{i_{(ij)}} - \theta_{j_{(ji)}}\right) + b_{ij} v_{i_{(ij)}} v_{j_{(ji)}} \sin\left(\theta_{i_{(ij)}} - \theta_{j_{(ji)}}\right), \tag{16d}$$

$$q_{ij} = b_{ij}^c v_{i_{(ij)}}^2 - b_{ij} v_{i_{(ij)}} v_{j_{(ji)}} \cos\left(\theta_{i_{(ij)}} - \theta_{j_{(ji)}}\right) - g_{ij} v_{i_{(ij)}} v_{j_{(ji)}} \sin\left(\theta_{i_{(ij)}} - \theta_{j_{(ji)}}\right), \tag{16e}$$

$$p_{ji} = g_{ji}^c v_{j_{(ji)}}^2 - g_{ji} v_{j_{(ji)}} v_{i_{(ij)}} \cos\left(\theta_{j_{(ji)}} - \theta_{i_{(ij)}}\right) + b_{ji} v_{j_{(ji)}} v_{i_{(ij)}} \sin\left(\theta_{j_{(ji)}} - \theta_{i_{(ij)}}\right), \tag{16f}$$

$$q_{ji} = b_{ji}^c v_{j_{(ji)}}^2 - b_{ji} v_{j_{(ji)}} v_{i_{(ij)}} \cos\left(\theta_{j_{(ji)}} - \theta_{i_{(ij)}}\right) - g_{ji} v_{j_{(ji)}} v_{i_{(ij)}} \sin\left(\theta_{j_{(ji)}} - \theta_{i_{(ij)}}\right), \tag{16g}$$

$$\sqrt{p_{ij}^2 + q_{ij}^2} \le \overline{s}_{ij},\ \sqrt{p_{ji}^2 + q_{ji}^2} \le \overline{s}_{ij}, \tag{16h}$$

where $\boldsymbol{x}_{ij}^{\mathrm{l}} := \left[p_{ij}, q_{ij}, v_{i_{(ij)}}, \theta_{i_{(ij)}}, p_{ji}, q_{ji}, v_{j_{(ji)}}, \theta_{j_{(ji)}}\right]$. On the other hand, buses solve

$$D^{\mathrm{b}}_{\rho,i}\left(\left(\boldsymbol{\lambda}_i^{g,k}\right)_{g\in\mathcal{G}_i}, \left(\boldsymbol{\lambda}_{ij}^k\right)_{j\in\mathcal{B}_i}\right) = \underset{\boldsymbol{z}_i^{\mathrm{b}}}{\text{minimize}} \Bigg\{\sum_{g\in\mathcal{G}_i}\Bigg(-\left\langle\boldsymbol{\lambda}_i^{g,k}, \left[p_{i_{(i)}}^g, q_{i_{(i)}}^g\right]\right\rangle + \frac{\rho_{pq}}{2}\left(\left(p_i^{g,k+1} - p_{i_{(i)}}^g\right)^2 + \left(q_i^{g,k+1} - q_{i_{(i)}}^g\right)^2\right)\Bigg) + \sum_{j\in\mathcal{B}_i}\Bigg(-\left\langle\boldsymbol{\lambda}_{ij}^k, \left[p_{ij_{(i)}}, q_{ij_{(i)}}, v_i, \theta_i\right]\right\rangle + \frac{\rho_{pq}}{2}\left(\left(p_{ij}^{k+1} - p_{ij_{(i)}}\right)^2 + \left(q_{ij}^{k+1} - q_{ij_{(i)}}\right)^2\right) + \frac{\rho_{v\theta}}{2}\left(\left(v_i - v_{i_{(ij)}}^{k+1}\right)^2 + \left(\theta_i - \theta_{i_{(ij)}}^{k+1}\right)^2\right)\Bigg)\Bigg\}, \tag{17a}$$

$$\text{subject to (8e), (8f).} \tag{17b}$$

where

$$\boldsymbol{z}_i^{\mathrm{b}} := \left[\left(p_{i_{(i)}}^g, q_{i_{(i)}}^g\right)_{(g,i)\in\mathcal{G}}, v_i, \theta_i, \left(p_{ij_{(i)}}, q_{ij_{(i)}}\right)_{j\in\mathcal{B}_i}\right].$$

The effect of adding the ADMM penalty term (with $\rho > 0$) is twofold. First, it makes the local cost functions finite and strictly convex and therefore the modified dual function bounded. Second, it makes the modified dual function smooth for large values of ρ. For small values of ρ, the concave modified dual function $D_\rho\left(\boldsymbol{\lambda}\right)$ is typically nondifferentiable. Indeed, using *Danskin*'s theorem [30]–[32], the subdifferentials of $D_\rho\left(\boldsymbol{\lambda}\right)$ are $\partial D_\rho\left(\boldsymbol{\lambda}\right) := \left\{\boldsymbol{x}-\boldsymbol{z} : D_\rho\left(\boldsymbol{\lambda}\right), \boldsymbol{x}\in\mathcal{X}, \boldsymbol{z}\in\mathcal{Z}\right\}$, where $\mathcal{X}$ is the feasible set defined by constraints (1b), (1c), (1l), (8c), (8d) and (8g) to (8j) and $\mathcal{Z}$ is the feasible set defined by constraints (8e), (8f). More specifically, as the (nonconvex) transmission line subproblems in (16) can have multiple (globally) optimal solutions for a given vector $\boldsymbol{\lambda}$, the subdifferentials $\partial D_\rho\left(\boldsymbol{\lambda}\right)$ may be not be unique and the modified dual function $D_\rho\left(\boldsymbol{\lambda}\right)$ can be nonsmooth. In more detail,

$$\boldsymbol{g}_\rho := \left[\boldsymbol{x}-\boldsymbol{z}\right] \in \partial D_\rho\left(\boldsymbol{\lambda}\right),$$

which is a subgradient of $D_\rho\left(\boldsymbol{\lambda}\right)$, may not be unique when ρ is small. On the other hand, for large values of ρ, $\boldsymbol{g}_\rho$ is unique and is therefore a gradient of $D_\rho\left(\boldsymbol{\lambda}\right)$, *i.e.* $\boldsymbol{g}_\rho = \nabla D_\rho\left(\boldsymbol{\lambda}\right)$. Finally, at each iteration k, each bus $i\in\mathcal{B}$ updates its (local) Lagrange multipliers using the subgradient projection method as follows,

$$\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \boldsymbol{\rho}\boldsymbol{g}_\rho^k, \tag{18}$$

where $\boldsymbol{\rho} := \left[\left(\rho_{pq}\right)_{2\times|\mathcal{G}|}, \left(\rho_{pq}, \rho_{pq}, \rho_{v\theta}, \rho_{v\theta}\right)_{|\mathcal{L}\cup\mathcal{L}_t|}\right]$.

The component-based ADMM algorithm is described in Algorithm 1.

Algorithm 1: Distributed algorithm

Initialization: $\boldsymbol{\lambda}^1 = \mathbf{0}$, $\rho >> 0$, $\epsilon \le 10^{-4}$ and, for all $i\in\mathcal{B}$, $\boldsymbol{z}_i^{\mathrm{b},1} = \left[\frac{\underline{p}_i^g + \overline{p}_i^g}{2}, \frac{\underline{q}_i^g + \overline{q}_i^g}{2}, 0, 0, 1, 0\right]$.
while $\left\|\boldsymbol{g}_\rho^k\right\| \ge \epsilon$ **do**
  Generators and lines solve (15) and (16) respectively in parallel, and send $\boldsymbol{x}_i^{g,k+1}$ and $\boldsymbol{x}_{ij}^{\mathrm{l},k+1}$ to adjacent buses.
  Buses solve (17) in parallel and update their (local) Lagrange multipliers as in (18).
  Buses send $\boldsymbol{z}_i^{\mathrm{b},k+1}$, $\boldsymbol{\lambda}_i^g$, $\boldsymbol{\lambda}_{ij}$ and $\boldsymbol{\lambda}_{ji}$ to adjacent generators and lines .
  $k \leftarrow k+1$.
end while

Definition 1: Let $P_{\mathrm{IPM}}^{\dagger}$ be a feasible primal solution computed centrally by an IPM solver and let $D_\rho\left(\boldsymbol{\lambda}^{\dagger}\right)$ be a solution of the approximate modified dual function computed in a distributed fashion by Algorithm 1, initialized with the same algorithmic starting point used to find $P_{\mathrm{IPM}}^{\dagger}$. Then the gap between the feasible primal solution $P^{\dagger}_{\text{IPM}}$ and its associated approximate modified dual function optimal value $D_{\rho}\left(\boldsymbol{\lambda}^{\dagger}\right)$ is given by

$$\text{AMDgap} := \left(\frac{P^{\dagger}_{\text{IPM}} - D_{\rho}\left(\boldsymbol{\lambda}^{\dagger}\right)}{P^{\dagger}_{\text{IPM}}}\right) \times 100.$$

Note that in Definition 1, if $P^{\dagger}_{\text{IPM}}$ is globally optimal, then $D_{\text{AMD}}(\boldsymbol{\lambda}^{\dagger})$ is an accurate approximation of the modified dual function.

Definition 2: Let $P^{\dagger}_{\text{ADMM}} = f\left(p_i^{g,\dagger}\right)$ be a feasible primal solution computed in a distributed fashion by Algorithm 1, initialized with the same algorithmic starting point used to find $P^{\dagger}_{\text{IPM}}$. Then the gap between the feasible primal solution $P^{\dagger}_{\text{IPM}}$ and $P^{\dagger}_{\text{ADMM}}$ is given by

$$\text{ROgap} := \left(\frac{P^{\dagger}_{\text{IPM}} - P^{\dagger}_{\text{ADMM}}}{P^{\dagger}_{\text{IPM}}}\right) \times 100.$$

V. Numerical evaluation

Algorithm 1 is implemented in MATLAB and the interfacing between AMPL and MATLAB is made possible by AMPL's application programming interface. The simulations are all carried out on a computing platform with 10 Intel Xeon E5-2687W v3 CPUs at 3.10GHz, 64-bit operating system, and 128GB RAM. In all simulations, AMPL [33] is used as a frontend modeling language for the optimization problems along with KNITRO 10.2 [34] as a backend solver for the nonconvex original OPF problem in (1) and the nonconvex line and bus subproblems in (16) and (17) respectively. Generators have closed form solutions as in [15].

The centralized IPM solutions, shown under $P^{\dagger}_{\text{IPM}}$ in Tables II and III, are initialized with $\boldsymbol{x}_i^{g,0} = \left[0.5\left(\underline{p}_i^g + \overline{p}_i^g\right), 0.5\left(\underline{q}_i^g + \overline{q}_i^g\right)\right]$ for all $(g,i) \in \mathcal{G}$, $\boldsymbol{x}_{ij}^{\text{l},0} = [0,0,1,0,0,0,1,0]$ for all $(i,j) \in \mathcal{L}$, and $\boldsymbol{z}_i^{\text{b},0} = \left[0.5\left(\underline{p}_i^g + \overline{p}_i^g\right), 0.5\left(\underline{q}_i^g + \overline{q}_i^g\right), 0, 0, 1, 0\right]$ for all $i \in \mathcal{B}$. This initialization is the same one used as a starting point for the IPM solver at each iteration k in Algorithm 1. The parameter settings of Algorithm 1 are summarized in Table I and the results are shown in Tables II and III for MATPOWER [2], [27] and NESTA v6 [28] instances respectively. The NESTA test cases are designed specifically to incorporate key network parameters such as line thermal limits and small angle differences, which are critical in optimization applications.

Tables II and III show that for $\epsilon = 10^{-4}$, after careful individualized tuning of parameters, Algorithm 1 converges to feasible solutions with negligible AMDgap and ROgap on all the 72 test cases.[3]

There are three key contributors behind the convergence of Algorithm 1 on all the 72 cases. First, parameters ρ_{pq} and $\rho_{v\theta}$ are set to high values, typically in the ranges $[100, 20000]$ and $[100, 500000]$, respectively. Second, most test cases require setting $\rho_{v\theta}$ to at least 3 orders of magnitude larger than ρ_{pq}.

[3]Note that the stopping criterion in Algorithm 1 requires a central authority to compute the norm of the subgradient; nonetheless, if a central authority is unavailable, the stopping criterion can be defined as in [12] or [14].

TABLE I
Summarized parameter settings of Algorithm 1.

Setting	ρ_{pq}	$\rho_{v\theta}$	ϵ
A	300	300000	1.E-04
B	400	40000	1.E-04
C	400	400	1.E-04
D	400	4000	1.E-04
E	100	10000	1.E-04
F	1000	1000	1.E-04
G	100	100	1.E-04
H	100	100000	1.E-04
I	500	500000	1.E-04
J	400	400000	1.E-04
K	400	800	1.E-04
L	10000	100000	1.E-04
M	1000	10000	1.E-04
N	20000	200000	1.E-06
O	1000	100000	1.E-04
P	2000	20000	1.E-04
Q	10000	100000	1.E-06

TABLE II
Convergence of Algorithm 1 on MATPOWER instances.

	Objective ($)			Gap (%)			
Case	$P^{\dagger}_{\text{IPM}}$	$P^{\dagger}_{\text{ADMM}}$	$D_{\rho}\left(\boldsymbol{\lambda}^{\dagger}\right)$	AMD	RO	I	set
5	17551.89	17551.66	17551.89	-1.59E-05	1.30E-03	1229	A
6ww	3143.97	3143.93	3143.98	-8.79E-05	1.47E-03	972	A
9	5296.69	5296.90	5296.71	-4.39E-04	-4.02E-03	454	B
14	8081.52	8081.43	8081.53	-1.80E-05	1.14E-03	627	C
24	63352.20	63352.20	63352.20	-1.25E-06	8.52E-06	1290	C
30	576.89	576.93	576.92	-4.37E-03	-7.05E-03	1838	B
30	8906.14	8906.44	8906.14	-5.92E-07	-3.37E-03	894	C
39	41864.18	41864.17	41864.19	-3.03E-05	1.96E-05	3478	D
57	41737.79	41734.71	41737.79	-4.29E-07	7.37E-03	856	D
89	5819.81	5820.04	5819.99	-3.08E-03	-3.95E-03	8929	E
118	129660.69	129660.97	129660.71	-1.55E-05	-2.09E-04	635	D
300	719725.10	719727.91	719725.10	-5.75E-07	-3.90E-04	6202	D

Third, this disproportion in setting ρ_{pq} and $\rho_{v\theta}$ is reflected exactly in setting the values of the step size in the multiplier update (18). More specifically, the step sizes for updating the voltage and angle multipliers are also 3 orders of magnitude larger than the step sizes for the active and reactive power multipliers. To see the significance of this, all the test instances with this specific parameter tuning would otherwise either diverge or require more than 10^6 iterations to converge. Some notoriously difficult cases are MATPOWER's case 5, NESTA's cases 30_fsr_API and 189_API for which very few other parameter settings, besides the corresponding ones in Table II, seem to make Algorithm 1 converge. Furthermore, even after exhaustive parameter tuning, the convergence on some test instances (like NESTA's 300 bus systems) is substantially slower than others. Nonetheless, this still suggests that Algorithm 1 converges even on these *difficult* test instances.

It is evident from Tables II and III that there is no principled way of setting the parameters of Algorithm 1. However,

TABLE III
CONVERGENCE OF ALGORITHM 1 ON NESTA V5 INSTANCES.

	Objective ($)			Gap (%)			
Case	$P^{\dagger}_{\text{IPM}}$	$P^{\dagger}_{\text{ADMM}}$	$D_{\rho}\left(\boldsymbol{\lambda}^{\dagger}\right)$	AMD	RO	I	set
Normal Operating Conditions							
3	5812.64	5811.90	5812.64	-1.18E-06	1.28E-02	901	C
4	156.43	156.54	156.43	-9.84E-04	-6.93E-02	823	F
5	17551.89	17551.66	17551.89	-2.03E-05	1.31E-03	1302	A
6_c	23.21	23.21	23.21	-1.12E-03	-4.60E-04	620	G
6_ww	3143.97	3143.76	3143.97	3.50E-06	6.72E-03	1870	H
9	5296.69	5296.90	5296.71	-4.39E-04	-4.02E-03	454	B
14	244.05	244.06	244.12	-2.58E-02	-4.10E-03	8195	H
24	63352.20	63352.20	63352.20	-1.22E-06	7.89E-06	1290	C
29	29895.49	29895.64	29895.93	-1.46E-03	-4.89E-04	7394	H
30_as	803.13	803.10	803.13	-7.54E-04	3.79E-03	6982	H
30_fsr	575.77	575.80	575.80	-6.16E-03	-5.62E-03	3599	H
30	204.97	204.98	204.97	-2.38E-03	-4.96E-03	18061	H
39	96505.52	96506.35	96505.52	-5.09E-07	-8.61E-04	709	H
57	1143.27	1143.31	1143.27	-1.76E-04	-3.50E-03	27085	H
73	189764.08	189758.34	189764.08	-1.58E-06	3.03E-03	4264	H
89	5819.81	5820.01	5819.99	-3.09E-03	-3.44E-03	8932	E
118	3718.64	3718.60	3718.71	-1.91E-03	8.49E-04	3547	E
162	4230.23	4230.38	4230.23	-8.34E-06	-3.57E-03	23978	H
189	849.29	849.26	849.29	1.07E-05	3.74E-03	2444	G
300	16891.28	16885.83	16893.58	-1.36E-02	3.23E-02	289611	I
Congested Operating Conditions (API)							
3	367.74	367.61	367.74	8.53E-04	3.42E-02	2475	D
4	767.27	767.25	767.27	6.20E-04	3.77E-03	133	D
5	2998.54	2998.74	2998.61	-2.45E-03	-6.83E-03	3282	D
6_c	814.40	814.29	814.40	1.05E-03	1.35E-02	299	D
6_ww	273.76	273.73	273.77	-1.93E-03	1.12E-02	1070	J
9	656.60	656.63	656.60	-2.45E-04	-5.08E-03	3923	D
14	325.56	325.77	325.56	-2.40E-04	-6.45E-02	4166	K
24	6421.37	6422.37	6422.87	-2.33E-02	-1.56E-02	1308	B
29	295782.68	295778.85	295784.89	-7.47E-04	1.29E-03	10799	J
30_as	571.13	570.30	571.11	3.80E-03	1.45E-01	22147	D
30_fsr	372.14	372.04	372.20	-1.59E-02	2.52E-02	95048	Q
30	415.53	415.84	415.51	3.78E-03	-7.65E-02	9222	D
39	7466.25	7466.48	7466.20	7.18E-04	-3.02E-03	20692	D
57	1430.65	1430.78	1430.66	-1.08E-04	-8.83E-03	2607	K
73	20123.98	20119.88	20124.75	-3.83E-03	2.03E-02	34083	J
89	4288.02	4286.91	4289.44	-3.31E-02	2.61E-02	31924	J
118	10325.27	10327.60	10326.17	-8.71E-03	-2.25E-02	10247	J
162	6111.68	6111.94	6111.69	-2.30E-04	-4.21E-03	9034	M
189	1982.82	1983.42	1983.66	-4.20E-02	-3.03E-02	162007	N
300	22866.01	22864.20	22866.05	-1.63E-04	7.94E-03	83484	O
Small Angle Difference Conditions (SAD)							
3	5992.72	5991.61	5992.72	-3.06E-08	1.85E-02	2573	H
4	324.02	324.01	324.02	-7.64E-05	2.71E-03	711	H
5	26423.32	26422.31	26423.33	-4.55E-05	3.83E-03	2472	H
6_c	24.43	24.43	24.44	-3.13E-02	-1.44E-02	6704	H
6_ww	3149.51	3149.11	3149.51	2.19E-05	1.25E-02	6779	H
9	5590.09	5589.08	5590.09	4.93E-06	1.82E-02	6430	H
14	244.15	244.23	244.27	-5.05E-02	-3.48E-02	16453	H
24	79804.96	79805.84	79804.99	-3.51E-05	-1.10E-03	106967	Q
29	46933.26	46915.71	46933.88	-1.31E-03	3.74E-02	20144	H
30_as	914.44	913.83	914.44	1.61E-04	6.72E-02	46319	H
30_fsr	577.73	577.46	577.73	-2.87E-04	4.63E-02	4876	H
30	205.11	205.15	205.12	-1.49E-03	-1.93E-02	20242	H
39	97219.04	97218.00	97219.04	-2.46E-06	1.07E-03	5249	H
57	1143.88	1144.07	1143.88	3.54E-04	-1.60E-02	39382	H
73	235241.70	235249.73	235241.74	-1.98E-05	-3.42E-03	5799	L
89	5827.01	5827.05	5827.02	-2.19E-04	-6.09E-04	33386	H
118	4324.17	4323.01	4324.19	-5.20E-04	2.68E-02	22618	H
162	4369.19	4369.73	4369.78	-1.35E-02	-1.23E-02	18401	L
189	914.61	914.61	914.61	-2.27E-04	-1.26E-04	10850	M
300	16910.23	16904.58	16910.29	-3.91E-04	3.34E-02	52670	P

extensive simulations show that they can be clustered in a summarizing table of plausible parameter settings. Some parameter settings, like H for example, seem to work on the most number of test cases. This stands in contrast to settings Q and P which are tailored specifically to their respective test cases in Table III.

VI. CONCLUSION

The method is numerically demonstrated to converge to feasible near-optimal (if not optimal) solutions to the nonconvex AC OPF problem, corroborated by tight convex relaxations, on all the 72 considered test cases. Despite the absence of a principled way to set up the parameters of the algorithm, this work demonstrates that the ADMM penalty parameters should be set to at least 100, in order to ensure differentiability of the modified dual function. This not only affects the speed of convergence, but can mean the difference between convergence and divergence. More interestingly, in most cases, convergence can only be witnessed if the penalty parameters for the voltages and angles are set to at least one order of magnitude higher than those of the active and reactive power.

REFERENCES

[1] R. Jabr, A. Coonick, and B. Cory, "A primal-dual interior point method for optimal power flow dispatching," *Power Systems, IEEE Transactions on*, vol. 17, no. 3, pp. 654–662, Aug 2002.

[2] R. Zimmerman, C. Murillo-Sánchez, and R. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *Power Systems, IEEE Transactions on*, vol. 26, no. 1, pp. 12–19, Feb 2011.

[3] B. Kocuk, S. S. Dey, and X. A. Sun, "Strong SOCP relaxations for the optimal power flow problem," *Operations Research*, 2016.

[4] C. Coffrin, H. Hijazi, and P. Van Hentenryck, "Strengthening convex relaxations with bound tightening for power network optimization," in *Principles and Practice of Constraint Programming*, ser. Lecture Notes in Computer Science, G. Pesant, Ed. Springer International Publishing, 2015, vol. 9255, pp. 39–57.

[5] H. Hijazi, C. Coffrin, and P. V. Hentenryck, "Polynomial SDP cuts for optimal power flow," in *2016 Power Systems Computation Conference (PSCC)*, June 2016, pp. 1–7.

[6] C. Coffrin, H. Hijazi, and P. V. Hentenryck, "Strengthening the SDP relaxation of AC power flows with convex envelopes, bound tightening, and valid inequalities," *IEEE Transactions on Power Systems*, vol. PP, no. 99, pp. 1–1, 2016.

[7] B. Kocuk, S. S. Dey, and X. A. Sun, "Matrix minor reformulation and SOCP-based spatial branch-and-cut method for the AC optimal power flow problem," *arXiv preprint arXiv:1703.03050*, 2017.

[8] B. H. Kim and R. Baldick, "Coarse-grained distributed optimal power flow," *IEEE Transactions on Power Systems*, vol. 12, no. 2, pp. 932–939, May 1997.

[9] R. Baldick, B. H. Kim, C. Chase, and Y. Luo, "A fast distributed implementation of optimal power flow," *IEEE Transactions on Power Systems*, vol. 14, no. 3, pp. 858–864, Aug 1999.

[10] B. H. Kim and R. Baldick, "A comparison of distributed optimal power flow algorithms," *IEEE Transactions on Power Systems*, vol. 15, no. 2, pp. 599–604, May 2000.

[11] P. N. Biskas and A. G. Bakirtzis, "Decentralised OPF of large multi-area power systems," *IEE Proceedings - Generation, Transmission and Distribution*, vol. 153, no. 1, pp. 99–105, Jan 2006.

[12] S. Magnússon, P. C. Weeraddana, and C. Fischione, "A distributed approach for the optimal power flow problem based on ADMM and sequential convex approximations," *IEEE Transactions on Control of Network Systems*, vol. 2, no. 3, pp. 238–253, Sept 2015.

[13] T. Erseghe, "Distributed optimal power flow using ADMM," *IEEE Transactions on Power Systems*, vol. 29, no. 5, pp. 2370–2380, Sept 2014.

[14] M. Kraning, E. Chu, J. Lavaei, and S. Boyd, "Dynamic network energy management via proximal message passing," *Found. Trends Optim.*, vol. 1, no. 2, pp. 73–126, Jan. 2014.

[15] Q. Peng and S. H. Low, "Distributed optimal power flow algorithm for radial networks, i: Balanced single phase case," *IEEE Transactions on Smart Grid*, pp. 1–11, to be published.

[16] E. Dall'Anese, H. Zhu, and G. B. Giannakis, "Distributed optimal power flow for smart microgrids," *IEEE Transactions on Smart Grid*, vol. 4, no. 3, pp. 1464–1475, Sept 2013.

[17] P. Scott and S. Thiébaux, "Distributed multi-period optimal power flow for demand response in microgrids," in *Proceedings of the 2015 ACM Sixth International Conference on Future Energy Systems*, ser. e-Energy '15. New York, NY, USA: ACM, 2015, pp. 17–26.

[18] F. J. Nogales, F. J. Prieto, and A. J. Conejo, "A decomposition methodology applied to the multi-area optimal power flow problem," *Annals of Operations Research*, vol. 120, no. 1, pp. 99–116, 2003.

[19] A. G. Bakirtzis and P. N. Biskas, "A decentralized solution to the DC-OPF of interconnected power systems," *IEEE Transactions on Power Systems*, vol. 18, no. 3, pp. 1007–1013, Aug 2003.

[20] G. Hug-Glanzmann and G. Andersson, "Decentralized optimal power flow control for overlapping areas in power systems," *IEEE Transactions on Power Systems*, vol. 24, no. 1, pp. 327–336, Feb 2009.

[21] J. Guo, G. Hug, and O. K. Tonguz, "Intelligent partitioning in distributed optimization of electric power systems," *IEEE Transactions on Smart Grid*, vol. 7, no. 3, pp. 1249–1258, May 2016.

[22] A. Minot, Y. M. Lu, and N. Li, "A parallel primal-dual interior-point method for dc optimal power flow," in *2016 Power Systems Computation Conference (PSCC)*, June 2016, pp. 1–7.

[23] A. Y. S. Lam, B. Zhang, and D. N. Tse, "Distributed algorithms for optimal power flow problem," in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, Dec 2012, pp. 430–437.

[24] R. Madani, A. Kalbat, and J. Lavaei, "Distributed Computation for Sparse Semidefinite Programming with Applications to Power Optimization Problems." [Online]. Available: https://pdfs.semanticscholar.org/4ffd/961255807e453f2c60e1f2676e25ad305f8e.pdf

[25] A. X. Sun, D. T. Phan, and S. Ghosh, "Fully decentralized AC optimal power flow algorithms," in *2013 IEEE Power Energy Society General Meeting*, July 2013, pp. 1–5.

[26] S.-L. Lin and J. E. V. Ness, "Parallel solution of sparse algebraic equations," in *Conference Proceedings Power Industry Computer Application Conference*, May 1993, pp. 380–386.

[27] C. Josz, S. Fliscounakis, J. Maeght, and P. Panciatici, "AC power flow data in MATPOWER and QCQP format: iTesla, RTE snapshots, and PEGASE," *arXiv preprint arXiv:1603.01533*, 2016.

[28] C. Coffrin, D. Gordon, and P. Scott, "NESTA, the NICTA energy system test case archive," *CoRR*, vol. abs/1411.0359, 2014.

[29] J. Lavaei and S. Low, "Zero duality gap in optimal power flow problem," *Power Systems, IEEE Transactions on*, vol. 27, no. 1, pp. 92–107, Feb 2012.

[30] J. M. Danskin, *The Theory of Max-Min and Its Applications to Weapons Allocation Problems*. New York: Springer-Verlag, 1967.

[31] P. Bernhard and A. Rapaport, "On a theorem of Danskin with an application to a theorem of Von Neumann-Sion," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 24, no. 8, pp. 1163 – 1181, 1995.

[32] D. P. Bertsekas, *Nonlinear programming*. Athena Scientific, 1999.

[33] R. Fourer, D. M. Gay, and B. Kernighan, *Algorithms and Model Formulations in Mathematical Programming*, S. W. Wallace, Ed. New York, NY, USA: Springer-Verlag New York, Inc., 1989.

[34] R. H. Byrd, J. Nocedal, and R. A. Waltz, "Knitro: An integrated package for nonlinear optimization," in *Large Scale Nonlinear Optimization, 3559, 2006*. Springer Verlag, 2006, pp. 35–59.