A Survey of Distributed Optimization Algorithms for Optimal Power Flow Problems

Daniel K. Molzahn, *Member, IEEE*

***Abstract*—Historically, centrally computed algorithms have been the primary means of power system optimization. Independent System Operators dispatch large generators in a transmission system by centrally solving an optimal power flow (OPF) problem using full knowledge of the system parameters. With increasing penetrations of distributed energy resources requiring optimization of distribution networks, distributed optimization algorithms have been the subject of significant research interest. This paper surveys the literature of distributed optimization algorithms with applications to OPF problems.**

***Index Terms*—Distributed optimization, electric power systems**

I. Introduction

CENTRALIZED computation has traditionally been the primary way that optimization algorithms have been applied to electric power systems. Most notably, independent system operators (ISOs) seek a minimum cost generation dispatch for large-scale transmission systems by solving an optimal power flow (OPF) problem. (See [1]–[8] for related literature reviews.) The OPF formulation includes network parameters, such as line impedances, system topology, and flow limits; generator parameters, such as cost functions and output limits; and load parameters, such as an estimate of the expected load demands. The OPF problem minimizes the total system cost subject to engineering limits related to these parameters and the physical constraints dictated by the power flow equations. The ISO collects all the necessary parameters and performs a central computation to solve the OPF problem.

With increasing penetrations of distributed energy resources (e.g., rooftop PV generation, battery energy storage, plug-in vehicles with vehicle-to-grid capabilities, demand response resources, etc.), the centralized optimization paradigm most prevalent in current power systems will potentially be augmented with distributed optimization algorithms. Rather than collecting all problem parameters and performing a central calculation, distributed optimization algorithms are computed by many agents that obtain certain problem parameters via communication with a limited set of neighbors. Depending on the specifics of the distributed optimization algorithm and the application of interest, these agents may represent individual buses or large portions of a power system.

Distributed optimization has several potential advantages over centralized approaches. The computing agents only have to share limited amounts of information with a subset of the other agents, which can provide potential advantages with respect to cybersecurity and data privacy as well as reduce the expense of the necessary communication infrastructure. Distributed optimization also has advantages in robustness with respect to failure of individual agents. Further, with the ability to perform computations in parallel, distributed optimization algorithms have the potential to be computationally superior to centralized algorithms, both in terms of solution speed and the maximum problem size that can be addressed.

This paper surveys the literature of distributed optimization techniques applied to optimal power flow problems. Many distributed optimization techniques have been developed concurrently with new representations of the physical models describing power flow physics (i.e., the relationship between the complex voltage phasors and the power injections). The characteristics of a power flow model can have a large impact on the theoretical and practical aspects of an optimization formulation. Accordingly, this survey is segmented into sections based on the power flow model considered by each distributed optimization algorithm.

Note there have also been significant efforts to develop distributed optimization algorithms that neglect the power flow model entirely to solve the so-called "economic dispatch" problem. See, for instance, [9]–[12]. Other approaches incorporate economic and/or network congestion information into distributed approaches for solving frequency control problems [13]–[15]. See also [16], [17] for reviews of distributed control techniques. Further related work develops distributed solution approaches for other power system optimization problems, such as state estimation [10], [18], [19] and unit commitment [20]. With a focus on the OPF problem, this paper does not further emphasize these related topics.

This paper is organized as follows. Section II overviews background material, including the the power flow equations (along with various relaxations and approximations), the OPF problem, and common distributed optimization techniques. Sections III, IV and V review applications of distributed optimization techniques that use linear, non-linear convex, and non-convex power flow models, respectively. Section VI concludes the paper.

II. Overview of Background Material

This section overviews the power flow equations, presents the OPF problem formulation, and summarizes several distributed optimization techniques.

A. Power Flow Representations

This section summarizes the power flow equations and some relaxations and approximations which are relevant to existing

Argonne National Laboratory, Energy Systems Div.: dmolzahn@anl.gov

distributed optimization techniques. See [21] for a detailed survey of various power flow representations.

Consider an n-bus electric power system, where $\mathcal{N} = \{1, \ldots, n\}$ denotes the set of buses. Let $\mathcal{L}$ denote the set of lines. The network admittance matrix containing the electrical parameters and topology information is denoted $\mathbf{Y} = \mathbf{G} + \mathbf{j}\mathbf{B}$, where $\mathbf{j} = \sqrt{-1}$. Define $\overline{(\cdot)}$ as the complex conjugate.

For notational brevity and to match the development of many of the distributed optimization approaches that are reviewed in this paper, the power flow equations given here use a balanced single-phase-equivalent network representation. An unbalanced three-phase representation is more appropriate for some applications, such as models of distribution networks. Many of the algorithms surveyed in this paper could be extended to an unbalanced three-phase power flow model.

Each bus has an associated voltage phasor as well as active and reactive power injections. The voltage phasors are denoted $V \in \mathbb{C}^n$, with polar coordinate representation $|V|\, e^{\mathbf{j}\theta} = |V| \angle\theta$, where $|V| > 0 \in \mathbb{R}^n$ and $\theta \in (-180^\circ, 180^\circ]^n$. Each bus $i \in \mathcal{N}$ has active and reactive power injections $P_i + \mathbf{j}Q_i$, $P, Q \in \mathbb{R}^n$.

The power flow equations are

$$P_i + \mathbf{j}Q_i = V_i \sum_{k=1}^{n} \overline{\mathbf{Y}}_{ik} \overline{V}_k. \tag{1a}$$

Squared voltage magnitudes are

$$|V_i|^2 = V_i \overline{V}_i. \tag{1b}$$

Splitting real and imaginary parts of (1) and using polar voltage coordinates yields

$$P_i = |V_i| \sum_{k=1}^{n} |V_k| \left(\mathbf{G}_{ik} \cos(\theta_i - \theta_k) + \mathbf{B}_{ik} \sin(\theta_i - \theta_k)\right) \tag{2a}$$

$$Q_i = |V_i| \sum_{k=1}^{n} |V_k| \left(\mathbf{G}_{ik} \sin(\theta_i - \theta_k) - \mathbf{B}_{ik} \cos(\theta_i - \theta_k)\right) \tag{2b}$$

As an alternative to (2), balanced radial distribution networks can be represented using the *DistFlow* model [22]. Define the active and reactive power flows on line $(i,k) \in \mathcal{L}$ as P_{ik} and Q_{ik}, respectively, and denote the squared magnitude of the current flow from bus i to bus k as ℓ_{ik}. The DistFlow model is

$$P_{ik} + P_{ki} = r_{ik}\ell_{ik} \qquad \forall\,(i,k) \in \mathcal{L} \tag{3a}$$

$$Q_{ik} + Q_{ki} = x_{ik}\ell_{ik} \qquad \forall\,(i,k) \in \mathcal{L} \tag{3b}$$

$$|V_i|^2 = |V_k|^2 - 2\left(r_{ik}P_{ik} + x_{ik}Q_{ik}\right) + \left(r_{ik}^2 + x_{ik}^2\right)\ell_{ik} \qquad \forall\,(i,k) \in \mathcal{L} \tag{3c}$$

$$\ell_{ik}\,|V_i| = P_{ik}^2 + Q_{ik}^2 \qquad \forall\,(i,k) \in \mathcal{L} \tag{3d}$$

where $r_{ik} + \mathbf{j}x_{ik}$ is the series impedance of the line connecting buses i and k.[1] The power injections at each bus are given by the net flows out of the bus.

[1]Note that the DistFlow model can be extended to more general line models that include shunt admittances, non-zero phase shifts, and off-nominal voltage ratios.

The DistFlow model (3) fully represents the power flows for a balanced radial network. However, (3) is a relaxation for mesh network topologies due to the lack of a constraint ensuring consistency in the voltage angles.

Use of either power flow model (2) or (3) results in non-convex optimization problems that can be difficult to directly handle in distributed optimization algorithms. Therefore, many algorithms have focused on linear approximations and convex relaxations of the power flow equations.

The most commonly used linear approximation is the DC power flow model, which is based on several assumptions:

a. Reactive power flows can be neglected.
b. The lines are lossless (i.e., $\mathbf{G} = 0$) and shunt elements can be neglected.
c. The voltage magnitudes at all buses are approximately equal, so $|V_i| \approx 1$ at all buses $i \in \mathcal{N}$.
d. Angle differences between connected buses are small such that $\sin(\theta_i - \theta_k) \approx \theta_i - \theta_k$, $\forall\,(i,k) \in \mathcal{L}$.

Applying these assumptions to (2) yields the DC power flow model:

$$\sum_{(i,k)\in\mathcal{L}} \mathbf{B}_{ik}\left(\theta_i - \theta_k\right) = P_i \qquad \forall i \in \mathcal{N} \tag{4}$$

Distribution networks typically violate these assumptions, which motivates the development of alternate linearizations. One approach that is relevant to distributed optimization techniques performs a linearization around the "no-load" voltage profile under the assumptions of negligible shunt impedances and near-nominal voltage magnitudes. The voltage magnitudes can then be approximated as functions of the active and reactive power injections:

$$|V| \approx 1 + \mathbf{R}P + \mathbf{X}Q \qquad \forall i \in \mathcal{N} \tag{5}$$

where $\mathbf{Y}^{-1} = \mathbf{R} + \mathbf{j}\mathbf{X}$. See [23], [24] for further details.

Alternatively, another linear approximation can be formulated by neglecting the losses in the DistFlow model (3) to obtain the so-called *Linearized DistFlow* model:

$$P_{ik} = P_k + \sum_{m:(k,m)\in\mathcal{L}} P_{km} \qquad \forall\,(i,k) \in \mathcal{L} \tag{6a}$$

$$Q_{ik} = Q_k + \sum_{m:(k,m)\in\mathcal{L}} Q_{km} \qquad \forall\,(i,k) \in \mathcal{L} \tag{6b}$$

$$|V_k|^2 = |V_i|^2 - 2\left(R_{ik}P_{ik} + X_{ik}Q_{ik}\right) \qquad \forall\,(i,k) \in \mathcal{L} \tag{6c}$$

The linearizations (4), (6) and (5) approximate the power flow equations. Alternative approaches form *convex relaxations* of the power flow equations. Convex relaxations enclose the non-convex feasible spaces associated with the power flow equations in a larger space. Convex relaxations bound the optimal objective value for the original non-convex problem and provide sufficient conditions for certifying problem infeasibility. Certain convex relaxations also yield the globally optimal decision variables for some optimization problems.

This section next presents two convex relaxations of the power flow equations: a semidefinite programming (SDP) relaxation [25] and a second-order cone programming (SOCP) relaxation [26].

The SDP relaxation is derived by formulating a rank-one matrix $\mathbf{W} = VV^H \in \mathbb{C}^{n\times n}$, where $(\cdot)^H$ is the complex conjugate transpose operator. The power flow equations (1) are linear in the entries of $\mathbf{W}$. Let $e_i \in \mathbb{R}^n$ denote the i^{th} standard basis vector. Define the matrices

$$\mathbf{H}_i = \frac{\mathbf{Y}^H e_i e_i^\intercal + e_i e_i^\intercal \mathbf{Y}}{2} \tag{7a}$$

$$\tilde{\mathbf{H}}_i = \frac{\mathbf{Y}^H e_i e_i^\intercal - e_i e_i^\intercal \mathbf{Y}}{2\mathbf{j}} \tag{7b}$$

where $(\cdot)^\intercal$ is the transpose operator. A semidefinite relaxation of (1) is formed by relaxing the non-convex rank constraint to a positive semidefinite matrix constraint:

$$P_i + \mathbf{j}Q_i = \mathrm{tr}\left(\mathbf{H}_i\mathbf{W}\right) + \mathbf{j}\mathrm{tr}\left(\tilde{\mathbf{H}}_i\mathbf{W}\right) \tag{8a}$$

$$|V_i|^2 = \mathrm{tr}\left(e_i e_i^\intercal \mathbf{W}\right) \tag{8b}$$

$$\mathbf{W} \succeq 0 \tag{8c}$$

where $\mathrm{tr}(\cdot)$ is the matrix trace operator and $\succeq$ indicates positive semidefiniteness. If a solution to an associated optimization problem has $\mathrm{rank}(\mathbf{W}) = 1$, the SDP relaxation is *exact* and yields globally optimal decision variables. Specifically, let η denote a unit-length eigenvector of $\mathbf{W}$, rotated such that $\angle\eta_1 = 0$, with associated non-zero eigenvalue λ. The globally optimal voltage phasors are then $V^* = \sqrt{\lambda}\,\eta$. If $\mathrm{rank}(\mathbf{W}) > 1$, the SDP relaxation does not directly provide globally optimal decision variables, but does yield a bound on the optimal objective value of the non-convex problem.

An SOCP relaxation can be formulated in terms of the DistFlow model (3) variables [26]. The DistFlow model is non-convex due to the equality constraint (3d). To construct a convex SOCP relaxation of (3), replace the equality constraint (3d) with an inequality:

$$\ell_{ik}\,|V_i| \geq P_{ik}^2 + Q_{ik}^2 \qquad \forall\,(i,k) \in \mathcal{L} \tag{9}$$

The SOCP relaxation considered in this paper is (3a)–(3c) and (9). For single-phase models of radial networks that satisfy certain technical conditions, the SOCP relaxation is guaranteed to be exact [27].[2] Note that the technical conditions which guarantee exactness of the SOCP relaxation are non-trivial. For systems with Π-circuit line models that have non-negligible shunt capacitances or realistic device models, the SOCP relaxation may not be exact and thus fail to provide a physically meaningful solution [28].

B. Optimal Power Flow Formulation

The OPF problem optimizes system performance according to a specified objective function. Typical objective functions C are based on generation cost (i.e., $C = \sum_{i\in\mathcal{G}} c_{2,i}P_{Gi}^2 + c_{1,i}P_{Gi} + c_{0,i}$, where $c_{2,i} \geq 0$, $c_{1,i}$, and $c_{0,i}$ are scalar coefficients associated with the generator at bus i, P_{Gi} is the generation at bus i, and $\mathcal{G}$ is the set of generator buses), losses (i.e., $C = \sum_{i\in\mathcal{N}} P_{Gi}$), proximity to a desired voltage profile (i.e., $C = \sum_{i\in\mathcal{N}}\left(|V_i|^2 - |V_i^\bullet|^2\right)^2$, where $|V_i^\bullet|$ denotes a desired voltage profile), or some combination of these.

[2]The SDP relaxation (8) is at least as tight as the SOCP relaxation and therefore inherits this exactness guarantee.

Engineering constraints in the OPF problem limit the power injections and voltage magnitudes, and the power flow equations must be satisfied. Flow limits (typically based on apparent power, active power, or current magnitude) are also generally enforced. The specific line flow formulation depends on the power flow model and type of flow. Denote $f_{ik}(V_i, V_k)$ as the appropriate flow function for line $(i,k) \in \mathcal{L}$, with the specific function descriptions excluded for brevity.

The OPF problem considered in this paper is

$$\min \quad C \quad \text{subject to} \tag{10a}$$

$$P_i^{min} \leq P_i \leq P_i^{max} \qquad \forall i \in \mathcal{N} \tag{10b}$$

$$Q_i^{min} \leq Q_i \leq Q_i^{max} \qquad \forall i \in \mathcal{N} \tag{10c}$$

$$\left(V_i^{min}\right)^2 \leq |V_i|^2 \leq \left(V_i^{max}\right)^2 \qquad \forall i \in \mathcal{N} \tag{10d}$$

$$f_{ik}(V_i, V_k) \leq I_{i,k}^{max} \qquad \forall\,(i,k) \in \mathcal{L} \tag{10e}$$

$$\text{A power flow model} \tag{10f}$$

where "max" and "min" denote specified upper and lower limits on the corresponding quantities and the power flow model (10f) may be

- a non-convex formulation (1), (2), or (3);
- the DC power flow formulation (4), in which case the reactive power and voltage magnitude constraints (10c) and (10d) are ignored;
- the linear power flow representation (5) from [23], [24];
- the linearized DistFlow model (6);
- the SDP relaxation (8);
- the SOCP relaxation (3a)–(3c) and (9).

C. Summary of Distributed Optimization Techniques

This section next summarizes several distributed optimization techniques. Adopting from the exposition in [29], the first set of distributed optimization techniques are based on augmented Lagrangian decomposition. These include Dual Decomposition, the Alternating Direction Method of Multipliers with Proximal Message Passing, Analytical Target Cascading, and the Auxiliary Problem Principle. The second set of techniques are based on decentralized solution of the Karush-Kuhn-Tucker (KKT) necessary conditions for local optimality [30]. These include Optimality Condition Decomposition, Consensus+Innovation, and Gradient Dynamics.

Given its widespread use, the Alternating Direction Method of Multipliers is given a more detailed overview, while other decomposition techniques have a more summary treatment. These techniques have a broad conceptual similarity in that each considers distributed agents that pass information among one another and perform local computations to solve the overall problem. However, the details of the mathematical structure (which information is shared, how the algorithms ensure consistency between different subproblems, the specific computations performed by each agent, etc.) lead to differences in practical performance and theoretical properties. See, e.g., [29], [31] or the references below for more detailed discussions of these techniques. See also [29], [32], [33] for numerical comparisons between different distributed optimization techniques in the context of power system optimization problems, including empirical analyses of convergence rates.

1) Dual Decomposition: The Lagrangian functions for some optimization problems have a separable structure that can be exploited using dual decomposition techniques [31], [34], [35]. Consider an optimization problem of the form

$$\min_{x} \sum_{i=1}^{N} f_i\left(x_i\right) \text{ subject to} \tag{11a}$$

$$\sum_{i=1}^{N} \mathbf{A}_i x_i = b \tag{11b}$$

where, for $i = 1, \ldots, N$, $f_i\left(\cdot\right)$ is a specified function; $x_i \in \mathbb{R}^{n_i}$ is the length n_i vector of decision variables associated with the function f_i; $\mathbf{A}_i \in \mathbb{R}^{m \times n_i}$ is a specified matrix; and $b \in \mathbb{R}^m$ is a specified vector. The Lagrangian for (11) is

$$L\left(x, y\right) = \sum_{i=1}^{N} \left(f_i\left(x_i\right) + y^T \mathbf{A}_i x_i - (1/N) y^T b\right) \tag{12}$$

where $y \in \mathbb{R}^m$ is the vector of dual variables and $\left(\,\cdot\,\right)^T$ is the transpose. A decomposable mathematical structure in this form can often be constructed by duplicating variables shared by multiple functions f_i along with additional equality constraints that ensure consistency among the duplicated variables.

Dual decomposition methods use an iterative method called "dual ascent":

$$x_i^{k+1} = \operatorname*{argmin}_{x_i} L_i\left(x_i, y^k\right) \tag{13a}$$

$$y^{k+1} = y^k + \alpha^k \left(\sum_{i=1}^{N} \left(\mathbf{A}_i x_i^{k+1}\right) - b\right) \tag{13b}$$

where k is the iteration counter and $\alpha^k > 0$ is the specified step size at iteration k. Observe that each update of (13a) can be performed independently, which enables a distributed implementation of this step. (The dual variable update step (13b) requires a central coordinator.) Note that the convergence of dual decomposition techniques is generally not guaranteed, even for convex problems, and depends on the the step size α^k and problem characteristics.

2) Alternating Direction Method of Multipliers: Many distributed optimization approaches are based on the Alternating Direction Method of Multipliers (ADMM) algorithm or related variants. Similar to dual decomposition, ADMM has minimization and dual variable update steps. However, ADMM uses an *augmented* Lagrangian function. This section provides an overview of ADMM. See [35] for a detailed tutorial.

ADMM is applicable to optimization problems of the form

$$\min_{x,z} \ f\left(x\right) + g\left(z\right) \quad \text{subject to} \tag{14a}$$

$$\mathbf{A}x + \mathbf{B}z = c \tag{14b}$$

where x and z are decision variables, $\mathbf{A}$ and $\mathbf{B}$ are specified matrices, c is a specified vector, and $f\left(x\right)$ and $g\left(z\right)$ are specified functions. The ADMM algorithm is based on the augmented Lagrangian for (14):

$$L_\rho = f\left(x\right) + g\left(z\right) + y^T\left(\mathbf{A}x + \mathbf{B}z - c\right) + \frac{\rho}{2}\left|\left|\mathbf{A}x + \mathbf{B}z - c\right|\right|_2^2 \tag{15}$$

where $\rho > 0$ is a specified penalty parameter and $||\cdot||_2$ is the two-norm. Observe that (15) is the Lagrangian of (14) augmented with a weighted squared norm of the constraint residual. The ADMM algorithm iteratively minimizes the augmented Lagrangian by performing the following updates:

$$x^{k+1} = \operatorname*{argmin}_{x} \ L_\rho\left(x, z^k, y^k\right) \tag{16a}$$

$$z^{k+1} = \operatorname*{argmin}_{z} \ L_\rho\left(x^{k+1}, z, y^k\right) \tag{16b}$$

$$y^{k+1} = y^k + \rho\left(\mathbf{A}x^{k+1} + \mathbf{B}z^{k+1} - c\right) \tag{16c}$$

where superscripts indicate the iteration index and y is a dual variable. Since the x and z updates in (16a) and (16b) are independent, they can be performed in a distributed fashion.

If the functions $f\left(x\right)$ and $g\left(z\right)$ are convex, the ADMM algorithm (16) is guaranteed to converge to the solution of (14). Typically, the iterations converge quickly to moderate accuracy but can be slow to converge to high accuracy. The convergence rate depends on the choice of ρ, and different strategies have been proposed for adaptatively choosing this parameter [35]. The ADMM algorithm can be applied to non-convex problems, but there is no guarantee of convergence.

The flexibility afforded by the choice of the functions $f\left(x\right)$ and $g\left(z\right)$ allows for consideration of optimization problems with non-linear constraints. Consider, for instance, the optimization problem $\min_x f\left(x\right)$ s.t. $g_i\left(x\right) \geq 0$, $i = 1, \ldots, m$. ADMM can be applied to this problem using the reformulation $\min_{x,z} f\left(x\right) + h\left(z\right)$ s.t. $x = z$, where $h\left(z\right)$ is the indicator function $h\left(z\right) = \begin{cases} 0 & g_i \geq 0,\ i = 1, \ldots, m \\ \infty & \text{otherwise} \end{cases}$. The variations among the ADMM algorithms considered in this survey are often related to different choices for the decomposition between $f\left(x\right)$ and $g\left(x\right)$.

As described above, ADMM algorithms require a central coordinator to manage the dual variable update step (16c). However, a modification known as Proximal Message Passing (PMP) facilitates a fully decentralized algorithm. At each iteration of the proximal message passing algorithm, each agent evaluates a "prox" function:

$$\text{prox}_{f,\rho}\left(v\right) = \operatorname*{argmin}_{x}\left(f\left(x\right) + (\rho/2)\left|\left|x - v\right|\right|_2^2\right). \tag{17}$$

The vector x contains the decision variables (which themselves are chosen based on the power flow model) and the dual variables.[3] The vector v contains the average values of the variables in x for all neighboring nodes. The function $f\left(x\right)$ is the local objective for a specific agent with respect to the decision variables in x. The scalar ρ is a tuning parameter. Thus, the prox function optimizes the agent's local objective $f\left(x\right)$ while minimizing the weighted mismatch to the primal and dual variables from the agent's neighbors. The agents pass the results of the prox algorithm (i.e., their local copy of the variable x) to their neighbors such that each agent can compute the average value v to execute the next iteration. The algorithm convergences when the agents agree on common values for x.

[3] In other words, the vector x in (17) for each agent contains local copies of both the primal variables (x or z in (14), depending on the agent) and the dual variables y in the notation of (16).

The Proximal Message Passing algorithm is a special case of ADMM and thus inherits the convergence guarantees for convex problems. See [36] for further details, including a derivation of (17) from (14).

3) Analytical Target Cascading: Analytical Target Cascading (ATC) is a hierarchial, iterative approach for distributed solution of an optimization problem. The optimization problem is split into subproblems which are related by a tree structure. Parent and children subproblems in this tree share optimization variables, with the coupling modeled using penalty functions that are modified at each iteration. If all subproblems are convex, the algorithm is guaranteed to converge to the solution. Note that ATC algorithms require a central coordinator to manage the distributed computations. See, e.g., [37], [38] for further details.

4) Auxiliary Problem Principle: Similar to the previous techniques, the Auxiliary Problem Principle (APP) technique decomposes an optimization problem into subproblems with shared variables [39]. The subproblems each correspond to a region of the system with shared variables at the tie-lines connecting to neighboring regions. An augmented Lagrangian approach is again used to ensure consistency between the subproblems for neighboring regions. The key difference for APP techniques is that the cross-terms in the two-norm expression employed in the augmented Lagrangian (15) are linearized rather than modeled directly as in ADMM and ATC techniques. This decouples the subproblems such that no central coordinator is required for APP techniques. Convergence is guaranteed if all subproblems are convex.

5) Optimality Condition Decomposition: Rather than duplicating shared variables as in the previous techniques, the Optimality Condition Decomposition (OCD) technique assigns each primal and dual variable to a specific subproblem [40]. Each agent considers a subproblem under the condition that only its assigned variables are allowed to change (i.e., all variables that are assigned to other subproblems are fixed to their previous values). The couplings for the variables assigned to other subproblems are modeled using linear penalties that are added to the objective. The coefficients for these linear penalties are defined by the Lagrange multipliers resulting from other subproblems. At each iteration, each agent applies one step of a Newton-Raphson method to the KKT conditions for its subproblem and then shares the resulting primal and dual values with its neighboring agents. Thus, the OCD technique is effectively an approach for distributed solution of the KKT conditions for an optimization problem. Note that ODC techniques do not require a central coordinator. A sufficient condition for convergence holds when the coupling between subproblems is relatively weak (i.e., there is a small number of sparsely connected subproblems) [40]. A modified OCD algorithm using "correction terms" improves the convergence rate at the cost of a some additional communication between agents [41], [42].

6) Consensus+Innovation: The Consensus+Innovation (C+I) technique [10], [43] is similar to the OCD technique in that both perform a distributed solution of the KKT conditions. However, rather than assigning each variable to a certain subproblem as in the OCD technique, the C+I technique uses an iterative algorithm that allows all variables in a subproblem to vary. A limit point of the iterative algorithm satisfies the KKT conditions. For convex problems, any limit point of this iterative algorithm is therefore an optimal solution [43]. Since each step of the iterative algorithm can be performed using only local and neighboring information, computations in the C+I technique can be performed in a distributed fashion without the need for a central coordinator. Unlike OCD techniques, the C+I technique is applicable at any level of partitioning: an individual agent could potentially represent a single bus or a large region of the network. Various modifications of C+I speed convergence via additional communication links [44] and facilitate consideration of communication delays [45].

7) Gradient Dynamics: First proposed in [46] with more recent treatments in [47], [48], the Gradient Dynamics (GD) technique embeds the KKT conditions in a dynamical system. The equilibria of the dynamical system correspond to KKT points for the original OPF problem. Assuming the satisfaction of certain technical conditions, the approach in [49] and [50] constructs a formulation which ensures that only the optima of the OPF problem are locally stable, with other KKT points being unstable. Thus, the OPF problem can be solved by integrating the dynamical system. This technique inherits the decomposibility associated with the network structure: when integrating the dynamical system, each bus can serve as a computing agent that only communicates with its neighbors.

III. Distributed Algorithms for Linear Approximations of the OPF Problem

This section reviews distributed algorithms developed for optimization problems that employ the DC power flow model (4) for transmission systems and two power flow linearizations applicable to distribution systems, (5) and (6).

A. Distributed Optimization with a DC Power Flow Model

Following the exposition in [29], this section summarizes distributed optimization approaches for DC OPF problems categorized by the associated solution technique discussed in Section II-C. See [29] for an extensive review with detailed mathematical descriptions for many relevant algorithms and formulations.

1) Applications of Dual Decomposition to DC OPF Problems: Early work [51] in distributed approaches for solving DC OPF problems employs a dual decomposition technique that adds fictitious buses at the interconnections between independently coordinated areas. Note that the approach in [51] augments the DC power flow model (4) with an approximation of the line losses. Other work that applies dual decomposition techniques includes [52], which incorporates discrete decision variables. The approach in [52] uses so-called "ordinal optimization" techniques that aim to achieve "good enough" choices for the discrete variables while using a dual decomposition for the continuous variables. Recent publications [53], [54] study the integration of demand response resources, including privacy considerations and multiple time periods. Other recent work [55] applies the dual decomposition approach to the DC OPF problem (with a quadratic line loss approximation) in an electricity market context.

2) Applications of ADMM to DC OPF Problems: ADMM techniques have recently been applied to a variety of power system optimization problems. Reference [56] presents a mathematical treatment of ADMM in the context of DC OPF problems, including the consideration of asynchronous updates (i.e., only some of subproblems are updated at each iteration of the ADMM algorithm). The proximal message passing variant of ADMM (see Section II-C2) is applied to DC OPF problems in [57] and security-constrained DC OPF problems in [58]. Proximal message passing eliminates the need for a central coordinator to perform the dual update step, thus enabling a fully decentralized implementation. Note that [57] considers a multi-period formulation with many possible device types (HVDC lines, storage devices, controllable loads, etc.).

3) Applications of ATC to DC Unit Commitment Problems: Studies of ATC techniques with DC power flow models have been conducted in the context of security-constrained unit commitment problems [20], [59]. The approach in [20] has one central coordinator with multiple lower-level workers, each associated with a region of the transmission network. The approach in [59] models a transmission system with multiple connected distribution systems, with the decomposition occurring at the boundary substations. Note that a DC power flow model is used for the distribution systems, which is generally not appropriate. However, the general decomposition approach could conceivably be applied using a more realistic power flow model for the distribution systems.

4) Applications of APP to DC Unit Commitment Problems: An APP technique is applied in the context of the unit commitment problem in [60] using a two-level generalized Bender's decomposition approach. The top level determines a generator schedule by solving a conventional unit commitment problem. Multi-period DC OPF subproblems, each decomposed regionally using the APP technique, provide cuts for the master problem. This improves computational tractability and protects private utility data. Reference [61] also uses the APP technique to solve a two-stage stochastic unit commitment problem which considers wind uncertainty with geographically distributed reserves.

5) Applications of OCD to DC OPF Problems: DC OPF problems were among the first applications of OCD techniques. Rather than adding fictitious border buses, [62] uses OCD to decompose the DC OPF problem at the tie lines to neighboring regions. Reference [62] also demonstrates the capabilities of OCD techniques using a 583-bus model of the Balkan system. Demonstration on a network of computers is presented in [63], which includes some modifications that require a central coordinator to check for convergence.

A so-called Heterogeneous Decomposition (HGD) algorithm related to OCD techniques is used in [64] to jointly model transmission and distribution systems, decomposed at the boundary substations. The transmission system sends LMPs at the boundary buses to the distribution system, while the distribution system passes power consumption back to the transmission system. The approach in [64] uses a DC power flow model for both transmission and distribution systems, with the consideration of possible modifications to account for voltage constraints.

Improvements in the convergence speed of OCD techniques can be achieved by computing linear sensitivities for the dual variables passed to each subproblem [65]. A similar approach is applied in [66], which extends [64] by computing the sensitivities of LMPs to load injections at the boundaries between transmission and distribution systems.

6) Applications of C+I to DC OPF Problems: The C+I decomposition technique has solely been applied to DC OPF problems [10], [43]. At each iteration, the buses send their phase angle, power generation, and dual variables for the power balance and line flow constraints to their neighbors. Each bus then uses these shared variables to analytically compute an update for the next iteration. The C+I technique is guaranteed to converge to the DC OPF solution. Improvements made to the C+I technique include faster convergence rates via communicating with buses beyond immediate neighbors [44], the consideration of asynchronous updates [45], and incorporation of security constraints [67].

B. Distributed Optimization with Linearized Power Flow Models for Distribution Networks

While generally well suited for transmission systems, the DC power flow model is typically inappropriate for distribution systems. The other power flow linearizations discussed in Section II-A (i.e., (5) and (6)) are better models for distribution networks. This section next surveys the literature of distributed optimization algorithms that use these power flow models.

1) Applications of Dual Decomposition to Linear Power Flow Models: Reference [68] uses the linearized DistFlow model in concert with dual decomposition. Specifically, [68] considers a decentralized two-level stochastic optimization problem, with the first level representing the decisions for a microgrid and the second level representing the decisions for the distribution network operator. The microgrids are coupled by penalty functions that are iteratively determined by the distribution network operator.

A variety of online algorithms employ the linearized DistFlow model to control distribution systems using gradient projection methods. Requiring limited or no external communications, these algorithms can be viewed as feedback controllers that use real-time voltage measurements to update controllable device setpoints. Reference [69] proposes a gradient projection algorithm that solely uses local voltage measurements to update the reactive power outputs of inverters, with numerical results showing a favorable comparison to alternative droop-controller strategies. Reference [70] uses the power system itself as a real-time "power flow solver" to provide continuous feedback for the proposed controllers. Reference [71] uses similar ideas to derive a model-free "Extremum Seeking" algorithm for optimizing reactive power supply on distribution systems.

Other online algorithms developed for distribution system optimization are proposed in [72]–[74]. References [72] and [73] solve distributed optimization problems to determine reactive power injections using an alternative linearization of the power flow equations. Agents are assumed to have access to local measurements and can communicate with neighbors. Reference [74] uses the power flow linearization (5) from [23], [24] to design feedback controllers that update voltages to continuously seek the time-varying OPF solution.

2) Applications of ADMM to Linear Power Flow Models: The linearized DistFlow model (6) is used as the basis of the work in [75]–[77]. The approach in [75] minimizes power losses in a distribution system subject to limits on voltage magnitudes and inverter reactive power capabilities. ADMM is found to outperform a dual decomposition method for this problem. The approach in [76] uses a stochastic programming approach to consider uncertainty in distribution systems using ADMM, with problems decomposed over each bus and each scenario in the stochastic program. Likewise [77] also considers uncertainty, with this reference using an online regret minimization approach.

As an alternative to the linearized DistFlow equations, the approach in [78] uses the power flow linearization (5) from [23], [24] to optimize distribution systems with large penetrations of solar PV. The ability to regulate voltage magnitudes is validated using test cases with realistic solar generation data.

The power flow model employed in [79] is based on a linearization of the DistFlow model (3) about a specified operating point. The approach in [79] provides an optimal reactive power dispatch for voltage regulation in unbalanced radial distribution systems.

IV. Distributed Algorithms for Non-Linear Convex Approximations of the OPF Problem

Convex relaxations based on semidefinite programming [25] and second-order cone programming [26] have shown promise for a variety of power system optimization problems. Related research efforts have developed distributed approaches for solving these relaxations. This section reviews these efforts.

A. Distributed Optimization with the SDP Relaxation

As formulated in Section II-A, the positive semidefinite constraint (8c) in the SDP relaxation couples the variables associated with all buses. There exists an equivalent, sparsity-exploiting reformulation of this constraint that results in a mathematical structure that more closely represents the network topology [80], [81]. Specifically, the positive semidefinite constraint on the $n \times n$ matrix $\mathbf{W}$ in (8c) can be decomposed into positive semidefinite constraints on certain submatrices of $\mathbf{W}$. (The submatrices are determined by the maximal cliques of a chordal extension of the network graph. See [81] for further details.) This helps facilitate the application of various decomposition techniques. This section reviews applications of dual decomposition and ADMM techniques used in the context of SDP relaxations of the power flow equations.

1) Applications of Dual Decomposition to the SDP Relaxation: Reference [82] proposes two decompositions for the SDP relaxation derived from the primal and dual problem formulations. Computing agents solve SDP subproblems corresponding to small regions of the network and share primal or dual variables with the other connected subregions. The updates can be performed asynchronously. Reference [83] applies related techniques to the voltage regulation problem for distribution systems.

2) Applications of ADMM to the SDP Relaxation: ADMM techniques are applied to solve OPF problems for three-phase unbalanced models of radial distribution networks in [84], which shows improved convergence relative to dual decomposition approaches. A similar ADMM approach is applied in [85] to optimize distribution systems with large quantities of solar PV generation. Reference [86] applies ADMM to OPF problems for balanced mesh network models suitable for transmission systems. The heart of the approach in [86] consists of eigenvalue computations that can be performed in parallel. Reference [87] proposes an ADMM algorithm for unbalanced three-phase models of distribution systems. In the key step for the algorithm in [87], each of the agents' problems reduces to evaluating either a closed form expression or the eigendecomposition of a 6×6 matrix.

B. Distributed Optimization with the SOCP Relaxation

Reference [88] applies an ADMM technique to the SOCP relaxation (i.e., (3a)–(3c) and (9)) that creates subproblems associated with each bus. An analytical solution for each subproblem yields favorable computational characteristics for the approach in [88]. Related work [89] considers methods for tuning the ADMM parameter ρ in (16), which can have a large impact on the convergence rate.

V. Distributed Algorithms for the Non-Convex OPF Problem

Other than the C+I technique, all other decomposition techniques described in Section II-C have been applied to non-convex formulations of the OPF problem. Note that the theoretical guarantees associated with convex formulations (i.e., the linear approximations reviewed in Section III and the relaxations reviewed in Section IV) are generally not available for non-convex formulations. However, the papers reviewed below demonstrate that various distributed optimization techniques are capable of solving certain practical non-convex OPF problems.

1) Applications of Dual Decomposition to the Non-Convex OPF Problem: Early work [90] applies a dual decomposition method that dualizes the coupling constraints associated with the tie lines between regions. Each subproblem is a non-convex OPF problem with a penalization term in the objective associated with the coupling constraints. The approach in [90] uses an interior point algorithm in combination with cutting plane methods to solve these subproblems. In more recent work, [91] proposes a dual decomposition based algorithm for balanced radial networks using an augmented Lagrangian approach. The algorithm in [91] can be implemented asynchronously and has associated theory claiming a convergence guarantee.

2) Applications of ADMM to the Non-Convex OPF Problem: Reference [92] applies ADMM to a decoupled power flow model, which independently considers the active power/voltage angle and reactive power/voltage magnitude couplings. The algorithm in [92] decomposes the active and reactive power flows between regions.

Recent ADMM-based research efforts [41], [93]–[95] model the fully coupled AC power flow equations in terms of the voltage phasors. Reference [93] decomposes coupling constraints on the rectangular voltage components (i.e., e_i and f_i where the voltage phasor $V_i = e_i + \mathbf{j}f_i$). The dual variable updates (16c) can be computed locally by each agent in this approach. Reference [94] also describes an ADMM approach that regionally decomposes subproblems based on shared rectangular voltage coordinates. Subsequent work [95] describes a decomposition using auxiliary variables that represent the sums and differences of voltage phasors between the terminals of lines that are shared by multiple regions. The sums and differences of the voltage phasors more closely represent the expressions found in the power flow equations, which results in improved convergence characteristics. Under the assumption that the solver applied to each subproblem is reliable in finding a local solution, an approach for updating the penalty parameter (ρ in (16)) gurantees convergence of the ADMM algorithm. In order to apply ADMM techniques to large problems, [41] proposes a spectral partitioning technique for determining the regional decomposition. In combination with the coupling approach proposed in [95] and a strategy for updating the penalty parameter, the spectral partitioning technique results in tractability for large problems (e.g., the 2383-bus Polish system in MATPOWER [96]) [41].

The algorithm in [97] uses a power flow formulation that includes variables for both current and voltage phasors. Each iteration of the algorithm in [97] solves a quadratic program derived via applying linearization techniques, with the overall algorithm yielding a solution that satisfies the non-linear power flow equations.

3) Applications of ATC to the Non-Convex OPF Problem: A two-level ATC algorithm is applied in [98] to coordinate the operation of a distribution grid that contains micrgrids. The voltage magnitudes and angles at the boundaries of the distribution system and microgrid subproblems are coupled using an exponential penalty formulation [99].

4) Applications of APP to the Non-Convex OPF Problem: Early work in distributed optimization techniques for OPF problems includes the APP-based approach in [100]. The OPF problem is decomposed regionally using "dummy generators" whose active and reactive power outputs and voltage phasors model the neighboring regions. Subsequent work [101] demonstrates the capabilities of this decomposition using a 2587-line model of ERCOT. Case studies with multiple regions are presented in [102], which also provides guidance regarding the choice of penalty parameters in the APP formulation.

5) Applications of OCD to the Non-Convex OPF Problem: OCD techniques were first proposed in the context of the non-convex OPF problem [40], with a more detailed description and analysis of the convergence characteristics presened in [103]. Several advances are presented in [104], including parameter tuning and better consideration of the reference angle. The approach in [105] considers the coordinated operation of FACTS devices using an overlapping regional decomposition. In order to speed convergence rates, [42] proposes the use of "correction terms" that require some additional sharing of information between buses which are not directly connected in the power system network. Reference [106] describes a partitioning method based on a spectral analysis that results in computational improvements for the OCD approach.

6) Applications of GD to the Non-Convex OPF Problem: The Gradient Dynamics approach is applied to solve the non-convex OPF problem in [49]. Neighboring buses communicate in order to integrate a dynamical system whose equilibria correspond to KKT points of the OPF problem. Theoretical analyses of the proposed approach and comparison to convex relaxation techniques are presented in [50], [107], [108].

VI. Conclusions

After summarizing various power flow models and techniques for distributed optimization, this paper has surveyed the literature regarding distributed optimization algorithms for the optimal power flow problem. Algorithms based on Dual Decomposition, the Alternating Direction Method of Multipliers, Analytical Target Cascading, the Auxiliary Problem Principle, Optimality Condition Decomposition, Consensus+Innovation, and Gradient Dynamics have shown promise in solving a variety of OPF problems.

References

[1] Z. Qiu, G. Deconinck, and R. Belmans, "A Literature Survey of Optimal Power Flow Problems in the Electricity Market Context," in *Power Systems Conference and Exposition, 2009. PSCE '09. IEEE/PES*, March 2009, pp. 1–6.

[2] J. Momoh, R. Adapa, and M. El-Hawary, "A Review of Selected Optimal Power Flow Literature to 1993. I. Nonlinear and Quadratic Programming Approaches," *IEEE Trans. Power Syst.*, vol. 14, no. 1, pp. 96–104, Feb. 1999.

[3] J. Momoh, M. El-Hawary, and R. Adapa, "A Review of Selected Optimal Power Flow Literature to 1993. II. Newton, Linear Programming and Interior Point Methods," *IEEE Trans. Power Syst.*, vol. 14, no. 1, pp. 105–111, Feb. 1999.

[4] S. Frank, I. Steponavice, and S. Rebennack, "Optimal Power Flow: A Bibliographic Survey, Parts I and II," *Energy Syst.*, vol. 3, no. 3, pp. 221–289, 2012.

[5] A. Castillo and R. O'Neill, "Survey of Approaches to Solving the ACOPF (OPF Paper 4)," US Federal Energy Regulatory Commission, Tech. Rep., Mar. 2013.

[6] B. Stott and O. Alsaç, "Optimal Power Flow–Basic Requirements for Real-Life Problems and their Solutions," in *SEPOPE XII Symposium, Rio de Janeiro, Brazil*, 2012.

[7] P. Panciatici, M. Campi, S. Garatti, S. Low, D. Molzahn, A. Sun, and L. Wehenkel, "Advanced Optimization Methods for Power Systems," *18th Power Syst. Comput. Conf. (PSCC)*, 18-22 Aug. 2014.

[8] F. Capitanescu, J. M. Ramos, P. Panciatici, D. Kirschen, A. M. Marcolini, L. Platbrood, and L. Wehenkel, "State-of-the-Art, Challenges, and Future Trends in Security Constrained Optimal Power Flow," *Electric Power Systems Research*, vol. 81, no. 8, pp. 1731 – 1741, 2011.

[9] S. Yang, S. Tan, and J. Xu, "Consensus Based Approach for Economic Dispatch Problem in a Smart Grid," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4416–4426, Nov. 2013.

[10] S. Kar, G. Hug, J. Mohammadi, and J. Moura, "Distributed State Estimation and Energy Management in Smart Grids: A Consensus + Innovations Approach," *IEEE J. Sel. Topics Signal Proc.*, vol. 8, no. 6, pp. 1022–1038, Dec. 2014.

[11] G. Binetti, A. Davoudi, F. Lewis, D. Naso, and B. Turchiano, "Distributed Consensus-Based Economic Dispatch With Transmission Losses," *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1711–1720, July 2014.

[12] G. Binetti, A. Davoudi, D. Naso, B. Turchiano, and F. Lewis, "A Distributed Auction-Based Algorithm for the Nonconvex Economic Dispatch Problem," *IEEE Trans. Industrial Informatics*, vol. 10, no. 2, pp. 1124–1132, May 2014.

[13] R. Mudumbai, S. Dasgupta, and B. Cho, "Distributed Control for Optimal Economic Dispatch of a Network of Heterogeneous Power Generators," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 1750–1760, Nov. 2012.

[14] C. Zhao, E. Mallada, S. Low, and J. Bialek, "A Unified Framework for Frequency Control and Congestion Management," in *19th Power Syst. Comput. Conf. (PSCC)*, June 2016.

[15] F. Dörfler and S. Grammatico, "Amidst Centralized and Distributed Frequency Control in Power Systems," in *American Control Conf. (ACC)*, July 2016, pp. 5909–5914.

[16] M. Yazdanian and A. Mehrizi-Sani, "Distributed Control Techniques in Microgrids," *IEEE Trans. Smart Grid*, vol. 5, no. 6, pp. 2901–2909, Nov. 2014.

[17] H. Han, X. Hou, J. Yang, J. Wu, M. Su, and J. Guerrero, "Review of Power Sharing Control Strategies for Islanding Operation of AC Microgrids," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 200–215, Jan. 2016.

[18] V. Kekatos and G. Giannakis, "Distributed Robust Power System State Estimation," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1617–1626, May 2013.

[19] H. Zhu and G. Giannakis, "Power System Nonlinear State Estimation Using Distributed Semidefinite Programming," *IEEE J. Sel. Topics Signal Proc.*, vol. 8, no. 6, pp. 1039–1050, Dec 2014.

[20] A. Kargarian, Y. Fu, and Z. Li, "Distributed Security-Constrained Unit Commitment for Large-Scale Power Systems," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1925–1936, July 2015.

[21] D. Molzahn, "A Survey of Relaxations and Approximations of the Power Flow Equations," *In preparation*, 2016.

[22] M. Baran and F. Wu, "Optimal Capacitor Placement on Radial Distribution Systems," *IEEE Trans. Power Del.*, vol. 4, no. 1, pp. 725–734, Jan. 1989.

[23] S. Bolognani and S. Zampieri, "On the Existence and Linear Approximation of the Power Flow Solution in Power Distribution Networks," *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 163–172, Jan. 2016.

[24] S. Dhople, S. Guggilam, and Y. Chen, "Linear Approximations to AC Power Flow in Rectangular Coordinates," in *53rd Annu. Allerton Conf. Commun., Control, and Comput.*, Sept. 2015, pp. 211–217.

[25] J. Lavaei and S. Low, "Zero Duality Gap in Optimal Power Flow Problem," *IEEE Trans. Power Syst.*, vol. 27, no. 1, pp. 92–107, Feb. 2012.

[26] S. Low, "Convex Relaxation of Optimal Power Flow–Parts I: Formulations and Equivalence," *IEEE Trans. Control Network Syst.*, vol. 1, no. 1, pp. 15–27, Mar. 2014.

[27] ——, "Convex Relaxation of Optimal Power Flow–Part II: Exactness," *IEEE Trans. Control Network Syst.*, vol. 1, no. 2, pp. 177–189, Jun. 2014.

[28] K. Christakou, D.-C. Tomozei, J.-Y. Le Boudec, and M. Paolone, "AC OPF in Radial Distribution Networks – Part I: On the Limits of the Branch Flow Convexification and the Alternating Direction Method of Multipliers," *arXiv:1503.06809*, July 2016.

[29] A. Kargarian, J. Mohammadi, J. Guo, S. Chakrabarti, M. Barati, G. Hug, S. Kar, and R. Baldick, "Toward Distributed/Decentralized DC Optimal Power Flow Implementation in Future Electric Power Systems," to appear in *IEEE Trans. Smart Grid*, 2016.

[30] J. Nocedal and S. Wright, *Numerical Optimization*, 2nd ed. Springer, 2000.

[31] E. Castillo, R. Mínguez, A. Conejo, and R. Garcia-Bertrand, *Decomposition Techniques in Mathematical Programming*. Springer Berlin, 2006.

[32] B. Kim and R. Baldick, "A Comparison of Distributed Optimal Power Flow Algorithms," *IEEE Trans. Power Syst.*, vol. 15, no. 2, pp. 599–604, May 2000.

[33] M. Arnold, S. Knopfli, and G. Andersson, "Improvement of OPF Decomposition Methods Applied to Multi-Area Power Systems," in *IEEE Lausanne PowerTech*, July 2007, pp. 1308–1313.

[34] H. Everett III, "Generalized Lagrange Multiplier Method for Solving Problems of Optimum Allocation of Resources," *Operations Research*, vol. 11, no. 3, pp. 399–417, 1963.

[35] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," *Found. Trends Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[36] M. Kraning, E. Chu, J. Lavaei, and S. Boyd, "Dynamic Network Energy Management via Proximal Message Passing," *Found. Trends Optimiz.*, vol. 1, no. 2, pp. 70–122, 2013.

[37] S. Tosserams, L. Etman, P. Papalambros, and J. Rooda, "An Augmented Lagrangian Relaxation for Analytical Target Cascading using the Alternating Directions Method of Multipliers," *Struct. Multidisc. Optim.*, vol. 31, no. 3, pp. 176–189, 2006.

[38] S. DorMohammadi and M. Rais-Rohani, "Comparison of Alternative Strategies for Multilevel Optimization of Hierarchial Systems," *Appl. Math.*, vol. 3, no. 10A, pp. 1448–1462, 2012.

[39] G. Cohen, "Auxiliary Problem Principle and Decomposition of Optimization Problems," *J. Optimiz. Theory Appl.*, vol. 32, no. 3, pp. 277–305, Nov. 1980.

[40] A. Conejo, F. Nogales, and F. Prieto, "A Decomposition Procedure based on Approximate Newton Directions," *Math. Program.*, vol. 92, pp. 495–515, 2002.

[41] J. Guo, G. Hug, and O. Tonguz, "Intelligent Partitioning in Distributed Optimization of Electric Power Systems," to appear in *IEEE Trans. Smart Grid*, 2016.

[42] K. Baker, J. Guo, G. Hug, and X. Li, "Distributed MPC for Efficient Coordination of Storage and Renewable Energy Sources Across Control Areas," *IEEE Trans. Smart Grid*, no. 2, pp. 992–1001, Mar. 2016.

[43] J. Mohammadi, G. Hug, and S. Kar, "Distributed Approach for DC Optimal Power Flow Calculations," *arXiv:1410.4236*, Oct. 2014.

[44] ——, "Role of Communication on the Converence Rate of Fully Distribued DC Optimal Power Flow," *IEEE Int. Conf. Smart Grid Comm.*, pp. 43–48, 2014.

[45] ——, "Asynchronous Distributed Approach for DC Optimal Power Flow," *IEEE Eindhoven PowerTech*, June 2015.

[46] K. Arrow, L. Hurwicz, and H. Uzawa, *Studies in Linear and Non-Linear Programming*. Stanford University Press, 1958.

[47] D. Feijer and F. Paganini, "Stability of Primal-Dual Gradient Dynamics and Applications to Network Optimization," *Automatica*, vol. 46, no. 12, pp. 1974–1981, Dec. 2010.

[48] J. Wang and N. Elia, "A Control Perspective for Centralized and Distributed Convex Optimization," in *IEEE 50th Annu. Conf. Decis. Control (CDC)*, Dec. 2011, pp. 3800–3805.

[49] X. Ma and N. Elia, "A Distributed Continuous-Time Gradient Dynamics Approach for the Active Power Loss Minimizations," in *51st Ann. Allerton Conf. Commun., Control, Comput. (Allerton)*, Oct. 2013, pp. 100–106.

[50] ——, "Convergence Analysis for the Primal-Dual Gradient Dynamics Associated with Optimal Power flow Problems," in *European Control Conf. (ECC)*, July 2015, pp. 1261–1266.

[51] A. Conejo and J. Aguado, "Multi-Area Coordinated Decentralized DC Optimal Power Flow," *IEEE Trans. Power Syst.*, vol. 13, no. 4, pp. 1272–1278, Nov. 1998.

[52] C. Lin and S. Lin, "Distributed Optimal Power Flow With Discrete Control Variables of Large Distributed Power Systems," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1383–1392, Aug. 2008.

[53] N. Gatsis and G. Giannakis, "Decomposition Algorithms for Market Clearing With Large-Scale Demand Response," *IEEE Trans. Smart Grid*, vol. 4, no. 4, pp. 1976–1987, Dec. 2013.

[54] J. Joo and M. Ili'c, "Multi-Layered Optimization Of Demand Resources Using Lagrange Dual Decomposition," *IEEE Trans. Smart Grid*, vol. 4, no. 4, pp. 2081–2088, Dec. 2013.

[55] E. Miehling and D. Teneketzis, "A Decentralized Mechanism for Computing Competitive Equilibria in Deregulated Electricity Markets," in *American Control Conf. (ACC)*, July 2016, pp. 4107–4113.

[56] A. Abboud, R. Couillet, M. Debbah, and H. Siguerdidjane, "Asynchronous Alternating Direction Method of Multipliers Applied to the Direct-Current Optimal Power Flow Problem," in *IEEE Int. Conf. Acoustics, Speech Signal Proc. (ICASSP)*, May 2014, pp. 7764–7768.

[57] M. Kraning, E. Chu, J. Lavaei, and S. Boyd, "Dynamic Network Energy Management via Proximal Message Passing," *Found. Trends Optimiz.*, vol. 1, no. 2, pp. 73–126, 2014.

[58] S. Chakrabarti, M. Kraning, E. Chu, R. Baldick, and S. Boyd, "Security Constrained Optimal Power Flow via Proximal Message Passing," in *Clemson Univ. Power Syst. Conf. (PSC)*, 2014.

[59] A. Kargarian and Y. Fu, "System of Systems Based Security-Constrained Unit Commitment Incorporating Active Distribution Grids," *IEEE Trans. Power Syst.*, vol. 29, no. 5, pp. 2489–2498, Sept. 2014.

[60] K. Chung, B. Kim, and D. Hur, "Distributed Implementation of Generation Scheduling Algorithm on Interconnected Power Systems," *Energy Conversion and Management*, vol. 52, no. 12, pp. 3457 – 3464, 2011.

[61] A. Ahmadi-Khatir, A. Conejo, and R. Cherkaoui, "Multi-Area Unit Scheduling and Reserve Allocation Under Wind Power Uncertainty," *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1701–1710, July 2014.

[62] A. Bakirtzis and P. Biskas, “A Decentralized Solution to the DC-OPF of Interconnected Power Systems,” *IEEE Trans. Power Syst.*, vol. 18, no. 3, pp. 1007–1013, Aug. 2003.

[63] P. Biskas, A. Bakirtzis, N. Macheras, and N. Pasialis, “A Decentralized Implementation of DC Optimal Power Flow on a Network of Computers,” *IEEE Trans. Power Syst.*, vol. 20, no. 1, pp. 25–33, Feb. 2005.

[64] Z. Li, Q. Guo, H. Sun, and J. Wang, “Coordinated Economic Dispatch of Coupled Transmission and Distribution Systems Using Heterogeneous Decomposition,” *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4817–4830, Nov. 2016.

[65] X. Lai, L. Xie, Q. Xia, H. Zhong, and C. Kang, “Decentralized Multi-Area Economic Dispatch via Dynamic Multiplier-Based Lagrangian Relaxation,” *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 3225–3233, Nov. 2015.

[66] Z. Li, Q. Guo, H. Sun, and J. Wang, “A New LMP-Sensitivity-Based Heterogeneous Decomposition for Transmission and Distribution Coordinated Economic Dispatch,” to appear in *IEEE Trans. Smart Grid*, 2016.

[67] J. Mohammadi, G. Hug, and S. Kar, “Agent-Based Distributed Security Constrained Optimal Power Flow,” to appear in *IEEE Trans. Smart Grid*, 2016.

[68] Z. Wang, B. Chen, J. Wang, and J. Kim, “Decentralized Energy Management System for Networked Microgrids in Grid-Connected and Islanded Modes,” *IEEE Trans. Smart Grid*, vol. 7, no. 2, pp. 1097–1105, Mar. 2016.

[69] H. Zhu and H. Liu, “Fast Local Voltage Control Under Limited Reactive Power: Optimality and Stability Analysis,” *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3794–3803, Sept. 2016.

[70] L. Gan and S. Low, “An Online Gradient Algorithm for Optimal Power Flow on Radial Networks,” *IEEE J. Sel. Areas Commun.*, vol. 34, no. 3, pp. 625–638, Mar. 2016.

[71] D. Arnold, M. Negrete-Pincetic, M. Sankur, D. Auslander, and D. Callaway, “Model-Free Optimal Control of VAR Resources in Distribution Systems: An Extremum Seeking Approach,” *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3583–3593, Sept. 2016.

[72] S. Bolognani and S. Zampieri, “A Distributed Control Strategy for Reactive Power Compensation in Smart Microgrids,” *IEEE Trans. Automatic Control*, vol. 58, no. 11, pp. 2818–2833, Nov. 2013.

[73] S. Bolognani, R. Carli, G. Cavraro, and S. Zampieri, “A Distributed Control Strategy for Optimal Reactive Power Flow with Power and Voltage Constraints,” in *IEEE Int. Conf, Smart Grid Commun. (SmartGridComm)*, Oct. 2013, pp. 115–120.

[74] E. D. Anese and A. Simonetto, “Optimal Power Flow Pursuit,” to appear in *IEEE Trans. Smart Grid*, 2016.

[75] P. Šulc, S. Backhaus, and M. Chertkov, “Optimal Distributed Control of Reactive Power Via the Alternating Direction Method of Multipliers,” *IEEE Trans. Energy Conv.*, vol. 29, no. 4, pp. 968–977, Dec. 2014.

[76] M. Bazrafshan and N. Gatsis, “Decentralized Stochastic Programming for Real and Reactive Power Management in Distribution Systems,” in *IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Nov. 2014, pp. 218–223.

[77] W. Ma, J. Wang, V. Gupta, and C. Chen, “Distributed Energy Management for Networked Microgrids Using Online Alternating Direction Method of Multipliers with Regret,” to appear in *IEEE Trans. Smart Grid*, 2016.

[78] S. Guggilam, E. DallAnese, Y. Chen, S. Dhople, and G. Giannakis, “Scalable Optimization Methods for Distribution Networks With High PV Integration,” *IEEE Trans. Smart Grid*, vol. 7, no. 4, pp. 2061–2070, July 2016.

[79] B. Robbins, H. Zhu, and A. Domnguez-Garca, “Optimal Tap Setting of Voltage Regulation Transformers in Unbalanced Distribution Systems,” *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 256–267, Jan. 2016.

[80] R. Jabr, “Exploiting Sparsity in SDP Relaxations of the OPF Problem,” *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 1138–1139, May 2012.

[81] L. Vandenberghe and M. Andersen, “Chordal Graphs and Semidefinite Optimization,” *Found. Trends Optimiz.*, vol. 1, no. 4, pp. 241–433, 2015.

[82] A. Lam, B. Zhang, and D. Tse, “Distributed Algorithms for Optimal Power Flow Problem,” in *51st Annu. Conf. Decis. Control (CDC)*, Dec. 2012.

[83] B. Zhang, A. Lam, A. Domnguez-Garca, and D. Tse, “An Optimal and Distributed Method for Voltage Regulation in Power Distribution Systems,” *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1714–1726, July 2015.

[84] E. Dall’Anese, H. Zhu, and G. B. Giannakis, “Distributed Optimal Power Flow for Smart Microgrids,” *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1464–1475, Sept. 2013.

[85] E. DallAnese, S. Dhople, B. Johnson, and G. Giannakis, “Decentralized Optimal Dispatch of Photovoltaic Inverters in Residential Distribution Systems,” *IEEE Trans. Energy Conv.*, vol. 29, no. 4, pp. 957–967, Dec. 2014.

[86] R. Madani, S. Sojoudi, and J. Lavaei, “Convex Relaxation for Optimal Power Flow Problem: Mesh Networks,” *IEEE Trans. Power Syst.*, vol. 30, no. 1, pp. 199–211, Jan. 2015.

[87] Q. Peng and S. Low, “Distributed Algorithm for Optimal Power Flow on Unbalanced Multiphase Distribution Networks,” *arXiv:1512.06482*, Dec. 2015.

[88] Q. Peng and S. H. Low, “Distributed Optimal Power Flow Algorithm for Radial Networks, I: Balanced Single Phase Case,” to appear in *IEEE Trans. Smart Grid*, pp. 1–11, 2016.

[89] W. Zheng, W. Wu, B. Zhang, H. Sun, and Y. Liu, “A Fully Distributed Reactive Power Optimization and Control Method for Active Distribution Networks,” *IEEE Trans. Smart Grid*, vol. 7, no. 2, pp. 1021–1033, Mar. 2016.

[90] J. Aguado and V. Quintana, “Inter-Utilities Power-Exchange Coordination: A Market-Oriented Approach,” *IEEE Trans. Power Syst.*, vol. 16, no. 3, pp. 513–519, Aug. 2001.

[91] K. Christakou, D.-C. Tomozei, J.-Y. Le Boudec, and M. Paolone, “AC OPF in Radial Distribution Networks – Part II: An Augmented Lagrangian-based OPF Algorithm, Distributable via Primal Decomposition,” *arXiv:1503.06809*, July 2016.

[92] X. Wang and Y. Song, “Apply Lagrangian Relaxation to Multi-Zone Congestion Management,” in *IEEE Power Engineering Society Winter Meeting*, vol. 2, 2001, pp. 399–404.

[93] A. Sun, D. Phan, and S. Ghosh, “Fully Decentralized AC Optimal Power Flow Algorithms,” in *IEEE Power Energy Society General Meeting*, July 2013.

[94] T. Erseghe, “Distributed Optimal Power Flow Using ADMM,” *IEEE Trans. Power Syst.*, vol. 29, no. 5, pp. 2370–2380, Sept. 2014.

[95] ——, “A Distributed Approach to the OPF Problem,” *EURASIP J. Advances Signal Proc.*, vol. 45, 2015.

[96] R. Zimmerman, C. Murillo-Sánchez, and R. Thomas, “MATPOWER: Steady-State Operations, Planning, and Analysis Tools for Power Systems Research and Education,” *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, 2011.

[97] S. Magnússon, P. Weeraddana, and C. Fischione, “A Distributed Approach for the Optimal Power-Flow Problem Based on ADMM and Sequential Convex Approximations,” *IEEE Trans. Control Netw. Syst.*, vol. 2, no. 3, pp. 238–253, Sept. 2015.

[98] A. Marvasti, Y. Fu, S. DorMohammadi, and M. Rais-Rohani, “Optimal Operation of Active Distribution Grids: A System of Systems Framework,” *IEEE Trans. Smart Grid*, vol. 5, no. 3, pp. 1228–1237, May 2014.

[99] S. DorMohammadi and M. Rais-Rohani, “Exponential Penalty Function Formulation for Multilevel Optimization using the Analytical Target Cascading Framework,” *Struct. Multidisc Optim.*, vol. 47, no. 11, pp. 559–612, 2013.

[100] B. Kim and R. Baldick, “Coarse-Grained Distributed Optimal Power Flow,” *IEEE Trans. Power Syst.*, vol. 12, no. 2, pp. 932–939, 1997.

[101] R. Baldick, B. Kim, C. Chase, and Y. Luo, “A Fast Distributed Implementation of Optimal Power Flow,” *IEEE Trans. Power Syst.*, vol. 14, no. 3, pp. 858–864, 1999.

[102] D. Hur, J. Park, and B. Kim, “Evaluation of Convergence Rate in the Auxiliary Problem Principle for Distributed Optimal Power Flow,” *IEE Proc.–Generation, Transmission and Distribution*, vol. 149, no. 5, pp. 525–532, Sept. 2002.

[103] F. Nogales, F. Prieto, and A. 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.

[104] P. Biskas and A. Bakirtzis, “Decentralised OPF of Large Multiarea Power Systems,” *IEE Proc.–Generation, Transmission and Distribution*, vol. 153, no. 1, pp. 99–105, Jan. 2006.

[105] G. Hug-Glanzmann and G. Andersson, “Decentralized Optimal Power Flow Control for Overlapping Areas in Power Systems,” *IEEE Trans. Power Syst.*, vol. 24, no. 1, pp. 327–336, Feb. 2009.

[106] J. Guo, G. Hug, and O. Tonguz, “Enabling Distributed Optimization in Large-Scale Power Systems,” *arXiv:1605.09785*, May 2016.

[107] X. Ma and N. Elia, “Existence of Strong Lagrange Duals to Certain Optimal Power Flows,” in *22nd Mediterranean Conf. Control Automation (MED)*, June 2014, pp. 640–645.

[108] ——, “A Sufficient Saddle Point Characterization for the Lagrangian Associated with General OPF Problems,” in *53rd IEEE Conf. Decis. Control*, Dec. 2014, pp. 1119–1124.