Toward Distributed/Decentralized DC Optimal Power Flow Implementation in Future Electric Power Systems

Amin Kargarian, *Member, IEEE*, Javad Mohammadi, *Student Member, IEEE*, Junyao Guo, *Student Member, IEEE*, Sambuddha Chakrabarti, *Student Member, IEEE*, Masoud Barati, *Member, IEEE*, Gabriela Hug, *Senior Member, IEEE*, Soummya Kar, *Member, IEEE*, and Ross Baldick, *Fellow, IEEE*

***Abstract*—This paper reviews distributed/decentralized algorithms to solve the optimal power flow (OPF) problem in electric power systems. Six decomposition coordination algorithms are studied, including analytical target cascading, optimality condition decomposition, alternating direction method of multipliers, auxiliary problem principle, consensus+innovations, and proximal message passing. The basic concept, the general formulation, the application for dc-OPF, and the solution methodology for each algorithm are presented. We apply these six decomposition coordination algorithms on a test system, and discuss their key features and simulation results.**

***Index Terms*—Distributed/decentralized optimization, optimal power flow, power systems operation, resource coordination.**

I. Introduction

In smart electric power grids, the system's physical structure is becoming more distributed. In such energy infrastructures, different entities take the control responsibility of different parts of the system. Although these entities might be independent, any decision made by a control entity affects decisions made by other entities as they are physically interconnected via transmission lines. This means that the entities need to collaborate with each other to achieve effective and reliable operation of the entire grid. However, the existing centralized control algorithms and energy management functions may no longer be appropriate for controlling and operating such a distributed energy system.

Optimal power flow (OPF) is an essential energy management function in electric power systems. The OPF seeks the generation settings which yield the least generation cost to supply a given power demand while taking into consideration system and equipment constraints [1]. A vast amount of research has been done on OPF solution methods in conventional centralized power systems with one control entity [2]–[10]. In order to modify and adjust the OPF (and generation scheduling) solution methods for the power grids including multiple control entities, various distributed/decentralized OPF (and economic dispatch) approaches have recently been presented in [11]–[35]. Generally, in most papers that study the OPF problem in smart grids, an OPF is formulated for each control entity with respect to shared information between neighboring entities. Then, iterative distributed or decentralized algorithms are developed to find the reliable and optimal operation of the entire grid. Note that we categorize approaches presented in the literature into two classes: distributed and decentralized solution algorithms. Although distributed and decentralized approaches have similarities, there are key differences between these two types of solution algorithms.

A. Distributed and Decentralized Approaches

In distributed approaches, as shown in Fig. 1a, there is a central coordinator that coordinates independent control entities. Each control entity only communicates with the coordinator, and there are no direct communication links between the entities. Thus, one can say that a distributed approach has (at least) two management levels, where level one includes the control entities and level two consists of the coordinator. We also can say that in the distributed OPF solution algorithms,

Manuscript received March 9, 2016; revised June 9, 2016 and August 22, 2016; accepted September 23, 2016. Date of publication October 3, 2016; date of current version June 19, 2018. The work of A. Kargarian was supported by the Louisiana Board or Regents under Grant LEQSF(2016-19)-RD-A-10. The work of J. Mohammadi, J. Guo, and S. Kar were supported by the National Science Foundation under Grant ECCS-1408222. The work of S. Chakrabarti and R. Baldick were supported by the National Science Foundation under Grant ECCS-1406894. Paper no. TSG-00306-2016. *(Corresponding author: Amin Kargarian.)*

A. Kargarian is with the Division of Electrical and Computer Engineering, Louisiana State University, Baton Rouge, LA 70803 USA (e-mail: kargarian@lsu.edu).

J. Mohammadi, J. Guo, and S. Kar are with the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: jmohamma@andrew.cmu.edu; junyaog@andrew.cmu.edu; soummyak@andrew.cmu.edu).

S. Chakrabarti and R. Baldick are with the Department of Electrical and Computer Engineering, University of Texas at Austin, Austin, TX 78712 USA (e-mail: sambuddha.chakrabarti@gmail.com; baldick@ece.utexas.edu).

M. Barati is with the Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77004 USA (e-mail: mbarati@uh.edu).

G. Hug is with the Information Technology and Electrical Engineering Department, Swiss Federal Institute of Technology, Zurich, Switzerland (e-mail: ghug@ethz.ch).

Digital Object Identifier 10.1109/TSG.2016.2614904

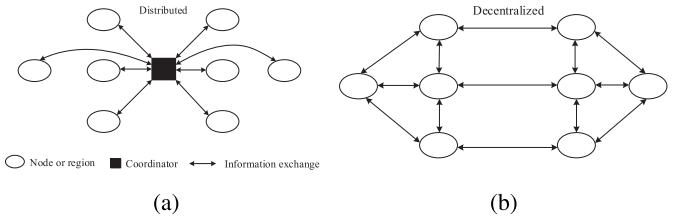

Fig. 1. Distributed and decentralized approaches.

the communication network differs from the physical electric power topology.

Unlike distributed approaches, decentralized methods do not have a central coordinator, and the control entities directly communicate and exchange information with their neighbors, as depicted in Fig. 1b. One can say that in the decentralized solution algorithms, all the control entities are autonomous and in the same management level. It is worth to mention that the communication network and the physical electric power topology do not have to be the same in the decentralized algorithms even though this is often the case.

B. Objectives of This Paper

In this paper, we overview distributed/decentralized approaches for optimal power flow solutions in smart electric power systems including multiple control entities where each entity is responsible to control its physical region. A physical region might include multiple buses connected by transmission lines or a single bus. In order to coordinate the OPF solutions of the entities, we study six decomposition coordination algorithms: analytical target cascading (ATC), alternating direction method of multipliers (ADMM), proximal message passing (PMP), auxiliary problem principle (APP), optimality condition decomposition (OCD), and consensus+innovations (C+I). We explain the key features of each algorithm and its application and solution method for the DC-OPF problem. Note, while ATC, OCD, ADMM and APP have been previously applied to AC OPF [18], [26], [27], [30], [36], C+I and PMP have only been used to solve DC-OPF problems. Therefore, for consistency and comparison of the algorithms, we present the application to the DC-OPF problem. The algorithms are applied to the IEEE 48-bus test system and simulation results are discussed.

Note that there are two existing papers which compare several algorithms on distributed/decentralized optimization in power systems [11], [37]. However, [11] is already fairly dated and [37] covers only two decomposition algorithms [37]. Hence, the main contribution of this paper is to provide an updated comparison of six algorithms which have been developed/studied in more recent years.

C. Brief Review of Other Methods

In addition to the six methods discussed in this paper, there are additional approaches which do not fall into any of these six categories [15]–[17], [38]–[44]. A few examples for such methods are provided here. In [15], a distributed implementation of the DC-OPF is proposed, where a problem is decomposed based on decoupling its first-order optimality conditions similar to OCD and C+I. However, to check the convergence after each iteration, a master coordinator receives local convergence flags from the subproblems and then returns a total convergence flag to every subproblem which is not the case in OCD and C+I. A heterogeneous decomposition algorithm is proposed in [38] and [39] for coordinated transmission and distribution economic dispatch. Locational marginal prices of the boundary buses are exchanged between the transmission and distribution systems. The heterogeneous decomposition algorithm falls into the optimality condition decomposition family. In addition, similar to ATC, the heterogeneous decomposition algorithm is applicable to multilevel hierarchical systems with heterogeneous entities. Reference [16] introduces a distributed dynamic stochastic OPF control algorithm based on the adaptive critic designs. This algorithm is a two-level control scheme, where lower-level controllers control their own region and a top-level controller adjusts the power flow in tie-lines connecting the adjacent regions. Since one top-level controller is needed, one can say that this method is similar to the classical ADMM and ATC which are distributed algorithms with a central coordinator. In [17], an approach is developed to implement a distributed OPF, which takes into account continuous and discrete control variables in power systems. The ordinal optimization strategy is applied to select a good enough discrete control variables solution, and a supervisor (similar to ATC and the classical ADMM that need a coordinator), called a root subsystem, is needed to monitor the convergence of the algorithm. Reference [45] considers a disaggregated bundling method for market clearing application. In this work, it has been shown that using the bundle method for the dual update can converge faster than the schemes of the cutting plane method and plain subgradient ascent. The PMP method is also based on the same concept of dual decomposition combined with augmented Lagrangian and proximal functions. Thus, it will be interesting to observe the effect of combining PMP with the method presented in [45].

D. Notes on the Paper Organization

The general versions of ATC and ADMM, which are distributed algorithms (i.e., with a coordinator), are discussed in this paper. Also, the general versions of PMP, APP, OCD and C+I, which are decentralized algorithms, are discussed. Out of these algorithms, PMP is a version of ADMM. While ATC, ADMM, PMP, and ATC use augmented Lagrangian, OCD and C+I solve the KKT conditions in a decentralized fashion. Thus, the organization of the paper is as follows: A general DC-OPF problem formulation is given in Section II. The analytical target cascading, the alternating direction method of multipliers, the proximal message passing, the auxiliary problem principle, the optimality condition decomposition, and the consensus+innovations are respectively illustrated in Sections III–VIII. The simulation studies are provided in Section IX. A summary and comparison of the six algorithms are provided in Section X. Section XI provides possible future directions of research and development.

E. Notation

We use bold italic symbols to denote vectors of variables, and italic symbols for scalar variables. Superscript † denotes a matrix or vector transpose. Superscript k is the iteration index and x^{*k} means that the value of variable x in iteration k is given. $\|\cdot\|^2$ is the square of the ℓ_2 norm, e.g., $\|x\|^2$ is the sum of the squares of all components of x. $\bar{x}$ and $\underline{x}$ indicate upper and lower bounds of x. Note that, in this paper, lowercase of $\boldsymbol{x}$ denotes variables and uppercase X refers to the reactance of lines.

II. Optimal Power Flow Formulation

A. DC-OPF Formulation

The objective of the optimal power flow is to determine the generation dispatch that yields the lowest cost to supply the load while all system and equipment constraints are satisfied. In this paper, we use the DC approximation of the OPF that leads to the following problem formulation:

$$\min_{(\boldsymbol{p},\boldsymbol{\theta})} \sum_{u\in\Omega_G} f(p_u) \tag{1}$$

s.t.

$$h(\boldsymbol{x}):\begin{cases} p_i - d_i = \sum_{j\in\Omega_i} \dfrac{\theta_i-\theta_j}{X_{ij}} \quad \forall i\in\{1,\ldots,N_b\} & (2)\\ \theta_{ref}=0 & (3)\end{cases}$$

$$g(\boldsymbol{x}):\begin{cases} \underline{P}_u \le p_u \le \overline{P}_u \quad \forall u\in\Omega_G & (4)\\ \underline{PL}_{ij} \le \dfrac{\theta_i-\theta_j}{X_{ij}} \le \overline{PL}_{ij} \quad \forall ij\in\Omega_L & (5)\end{cases}$$

$$\boldsymbol{x}=\{\boldsymbol{p},\boldsymbol{\theta}\}$$

where $\boldsymbol{x}$ includes generating units' outputs, $\boldsymbol{p}$, and the voltage angles of buses, $\boldsymbol{\theta}$. Functions $h(\boldsymbol{x})$ and $g(\boldsymbol{x})$ are sets of equality and inequality constraints, respectively. $f(p_u)$ is the quadratic cost function for generating unit u, d_i denotes the load at bus i, Ω_G denotes the set of all the generating units, Ω_i denotes the set of all the buses connected to bus i, and Ω_L denotes the set of all branches. Nodal power flow balances are enforced by (2). Equality (3) sets the voltage angle of the reference bus to zero. Inequalities (4) and (5) ensure that power output of generating unit u and power flow in transmission line ij are always within their upper and lower bounds.

B. OPF Formulation for a Multi-Region Power System

In a multi-region power system where the regions are interconnected through tie-lines, OPF formulations of the regions are linked together. There are two common ways to reformulate the multi-region OPF for individual regions, one uses coupling variables while the other uses coupling constraints. In the former formulation, different regions may contain the same variables but the constraint sets are separated, while in the latter formulation, the variable sets are separated but there are constraints that contain variables from two or more regions. Based on the two formulations, different decomposition methods can be applied to solve the problem. In the following, we present the subproblem formulation associated with coupling variables and coupling constraints, respectively [46].

In terms of coupling variables, in a general form, one can formulate the OPF problem of region n as follows:

$$\min_{(\boldsymbol{x}_n,\boldsymbol{\Phi})} \sum_{u\in\Omega_{Gn}} f(p_u) \tag{6}$$

s.t.

$$h_n(\boldsymbol{x}_n,\boldsymbol{\Phi})=0 \tag{7}$$

$$g_n(\boldsymbol{x}_n,\boldsymbol{\Phi})\le 0 \tag{8}$$

$$\boldsymbol{x}_n=\{\boldsymbol{p}_n,\boldsymbol{\theta}_n\}$$

where $f(p_u)$ is the cost function of a generating unit u located in region n. Ω_{Gn} is the set of all generating units of region n. $\boldsymbol{x_n}$ denotes the local variables of region n, which includes generator power outputs $\boldsymbol{p}_n$ and voltage angles $\boldsymbol{\theta}_n$. $\boldsymbol{\Phi}_n$ denotes the coupling variable between region n and its neighboring regions. As the coupling variable $\boldsymbol{\Phi}$ appears in the constraints of more than one region and thus affects the feasible space of the solution to the OPF problem in multiple regions, the OPF problems of those connected regions cannot be solved independently. A distributed/decentralized algorithm is required to coordinate the value of $\boldsymbol{\Phi}$ determined by different regions. In the considered DC-OPF problem, $\boldsymbol{\Phi}$ corresponds to the voltage angles of the cross-bordering buses between two regions. This will be explained further in Section III.

In terms of coupling constraints, the centralized optimization problem can be reformulated with respect to sets of variables associated with each region as follows:

$$\min_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n,\ldots,\boldsymbol{x}_N} \sum_{u\in\Omega_G} f(p_u) \tag{9}$$

s.t.

$$h_n^c(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n,\ldots,\boldsymbol{x}_N)=0 \tag{10}$$

$$g_n^c(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n,\ldots,\boldsymbol{x}_N)\le 0 \tag{11}$$

$$h_n^s(\boldsymbol{x}_n)=0 \tag{12}$$

$$g_n^s(\boldsymbol{x}_n)\le 0 \tag{13}$$

$$\boldsymbol{x}_n=\{\boldsymbol{p}_n,\boldsymbol{\theta}_n\}$$

where N is the total number of regions (or subproblems), n is the index of the region, and $\boldsymbol{x}_n$ includes the variables of region n. Constraints (10) and (11) are the coupling constraints in region n, and (12) and (13) are the non-coupling constraint included in region n. In the considered DC-OPF problem, coupling constraints include the power balance equations at the buses placed at the boundaries of the subsystems (i.e., (2) where buses i and j belong to different regions) and the limits on tie-line flows (i.e., (5) where buses i and j belong to different regions), while other constraints are considered as non-coupling constraints.

Note that in general different decomposition methods vary in the way they handle the coupling variables or the coupling constraints, and the specific coordination carried out by different methods will be introduced in the following sections. Also, note that depending on the considered problem, it is possible to transform a problem with coupling constraints into a problem with coupling variables and vice versa.

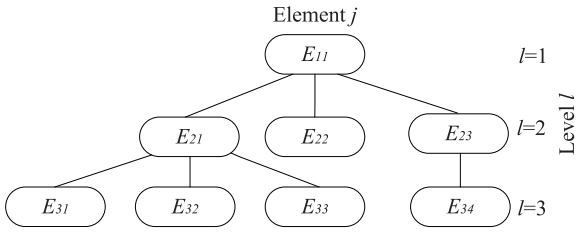

Fig. 2. An example of a hierarchical problem structure.

III. Analytical Target Cascading

A. Overview

Analytical target casting (ATC) is a model-based method for multilevel hierarchical optimization problems [47]–[49]. In the ATC method, the entire system (optimization problem) is divided into a set of subsystems (subproblems or elements) that are hierarchically connected as shown in Fig. 2. It means that the ATC structure includes several levels, each consisting of at least one subsystem. Subsystems in higher levels (also called parents) are hierarchically connected to subsystems in lower levels (also called children). Note that in the hierarchical ATC structure, there are no coupling variables between subsystems in the same level. However, parents and children are linked through coupling variables. The coupling variables are named target variables from parents' perspective, and response variables (which are indeed duplications of coupling variables) from children's perspective. Parents set the values of target variables and propagate them down to their children. The responses determined by the children define how close they are to the targets [50].

A set of consistency constraints is introduced to model the impact of the target and response variables in the parents' and children's optimization problems. The consistency constraints are then relaxed by including them as penalty functions in the objective function of each subsystem, i.e., parents and children. The ATC method provides certain flexibilities on the coordination of subproblems and the selection of penalty functions [51], [52]. An iterative solution procedure is implemented to obtain the optimal operating point of the entire system.

For a convex optimization problem, i.e., subproblems with convex local objective functions and constraints, the ATC algorithm has proven to converge to the optimal point [53]. For a nonconvex problem, selecting a quadratic penalty function to relax the consistency constraints in the subproblems, where quadratic terms act as local convexifiers, mitigates nonconvexity [50]. Note that to achieve convergence in the ATC algorithm, subproblems that share variables with each other need to be solved in a sequential manner, and they cannot be solved in parallel [50].

The ATC is applied in [26] to solve a decentralized OPF problem in active distribution grids encompassing microgrids. Distribution companies and microgrids are considered independent systems with their local OPF problems. An iterative solution algorithm is then presented to find the optimal and reliable operating point of the entire grid with respect to the information privacy of the independent systems.

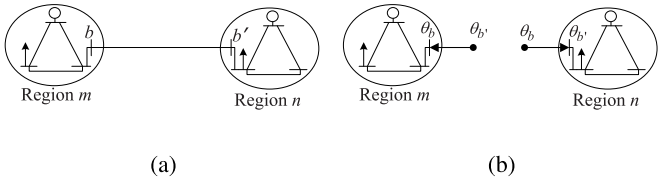

Fig. 3. (a) Two regions connected through tie-line bb', and (b) shared variables between the regions.

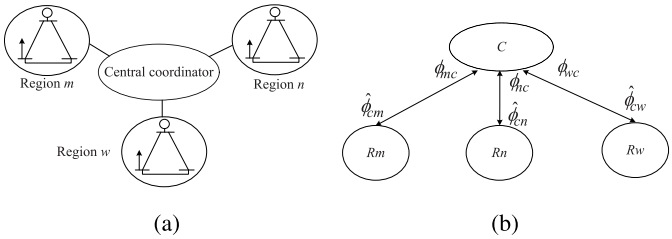

Fig. 4. (a) Virtually decoupling regions by a central coordinator, and (b) a two-level hierarchical ATC structure.

In [49] and [52]–[54], the ATC is used to find the optimal generation dispatch for day-ahead scheduling. The solution algorithm developed in [54] is a decentralized approach where the transmission and distribution operators communicate directly. In [51], the day-ahead scheduling is decomposed into several subproblems, and a central coordinator is introduced to parallelized the solution procedure of the subproblems. Introduction of the coordinator makes the problem an ATC-based two-level optimization problem which can be categorized as a distributed approach. In [57], ATC is used to develop a decentralized optimal power flow in three-phase unbalanced power distribution network. The objective is to minimize power loss by scheduling reactive power of rooftop photovoltaics generators.

B. ATC-Based Distributed OPF Algorithm

In this section, we describe how ATC is applied to the OPF problem. Fig. 3 shows regions n and m that are linked together through a tie-line connecting buses b and b'. As the power flow on this tie-line influences the operating conditions in both regions, voltage angles of buses b and b', i.e., θ_b and $\theta_{b'}$ (indicated by $\Phi_{bb'}$), are shared variables of regions n and m.

We define the system structure such that the corresponding subproblems can be solved in parallel. This is achieved by duplicating the voltage angles at both ends of the tie line, i.e., θ_b and $\theta_{b'}$ which are now coupling variables indicated by $\hat{\Phi}_{bb'}$. Additionally, a central entity is introduced to coordinate the values for the coupling variables (see [51] for more details). Fig. 4(a) depicts a multi-region power system with a central coordinator. This structure is indeed an ATC structure with two levels where the central coordinator (parent) forms the upper level, and the regions (children) form the lower level. The ATC-based structure and information exchange between the coordinator and regions are illustrated in Fig. 4(b). Note that as there is no direct links between any two children, their optimization problems can be solved in parallel.

A set of consistency constraints, i.e., $\hat{\boldsymbol{\Phi}}_{bb'} - \boldsymbol{\Phi}_{bb',n} = 0$, is formulated for region n in Fig. 4(b). These consistency constraints indicate that the target variables $\hat{\boldsymbol{\Phi}}_{bb'}$ of the parent (coordinator) and the corresponding response variables of child n ($\boldsymbol{\Phi}_{bb',n}$) need to have the same value (Note that we have a set of shared variables for each tie-line connecting two neighboring areas. We have previously indicated this by subscript bb' for $\boldsymbol{\Phi}$, however, we now merge all of these coupling variables into the vector $\boldsymbol{\Phi}$.). The consistency constraints are then relaxed using Augmented Lagrangian Relaxation, i.e., a second order penalty function, and the following local optimization problem is formulated for region n at iteration $k+1$:

$$\min_{\left(\boldsymbol{x}_n^{k+1}, \boldsymbol{\Phi}_n^{k+1}\right)} \sum_{u \in \Omega_{Gn}} f\left(p_u^{k+1}\right) + \left(\boldsymbol{\lambda}^{k\dagger}\left(\hat{\boldsymbol{\Phi}}_c^{*k} - \boldsymbol{\Phi}_n^{k+1}\right) + \left\|\boldsymbol{\beta}^k \circ \left(\hat{\boldsymbol{\Phi}}_c^{*k} - \boldsymbol{\Phi}_n^{k+1}\right)\right\|^2\right) \quad (14)$$

s.t.

$$\boldsymbol{h}_n\left(\boldsymbol{x}_n^{k+1}, \boldsymbol{\Phi}_n^{k+1}\right) = 0$$

$$\boldsymbol{g}_n\left(\boldsymbol{x}_n^{k+1}, \boldsymbol{\Phi}_n^{k+1}\right) \le 0$$

$$\boldsymbol{x}_n^{k+1} = \left\{\boldsymbol{p}_n^{k+1}, \boldsymbol{\theta}_n^{k+1}\right\}, \boldsymbol{\Phi}_n^{k+1} = \begin{bmatrix}\boldsymbol{\theta}_b^{k+1} \\ \boldsymbol{\theta}_{b'}^{k+1}\end{bmatrix}, \hat{\boldsymbol{\Phi}}_c^{*k} = \begin{bmatrix}\hat{\boldsymbol{\theta}}_b^{*k} \\ \hat{\boldsymbol{\theta}}_{b'}^{*k}\end{bmatrix}$$

where the symbol $\circ$ denotes the Hadamard product, and superscript $\dagger$ denotes transpose. $\boldsymbol{x}_n$ is the set of local variables of region n, and $\boldsymbol{\Phi}_n$ is the set of shared variables between region n and other regions. The term $\hat{\boldsymbol{\Phi}}_n^*$ indicates variables for which the values are determined by the central coordinator, and they are modeled as fixed values in the optimization problem of each region n (i.e., child n). All regions solve their own local subproblem (14), to determine the values of the coupling variables.

The central coordinator, as the parent, receives the values for the shared variables from its children. It then solves an optimization problem given by

$$\min_{\hat{\boldsymbol{\Phi}}_c^{k+1}} \sum_n \left(\boldsymbol{\lambda}^{k\dagger}\left(\hat{\boldsymbol{\Phi}}_c^{k+1} - \boldsymbol{\Phi}_n^{*k+1}\right) + \left\|\boldsymbol{\beta}^k \circ \left(\hat{\boldsymbol{\Phi}}_c^{k+1} - \boldsymbol{\Phi}_n^{*k+1}\right)\right\|^2\right) \quad (15)$$

which minimizes the difference between its own values for the coupling variables and the values provided by the children. Hence, the objective function of the central coordinator is indeed a set of penalty functions where its variables are the auxiliary variables $\hat{\boldsymbol{\Phi}}$, and $\boldsymbol{\Phi}^*$ are the values predetermined by the children. Assume that tie-line bb' connects areas n and m. The values of $\hat{\theta}_b^{k+1}$ (and $\hat{\theta}_{b'}^{k+1}$) determined by the central coordinator should be consistent for both areas n and m. In this regard, we can use one variable in (15) for bus b (and one variable for bus b'), and then exchange this variable with both areas n and m. This guarantees the feasibility of the results as the voltage angle of bus b must be the same in subproblems n and m, and indeed indicates that the central coordinator functions as a node with no power generation or consumption.

The iterative algorithm to solve the ATC-based distributed OPF where the subproblems are solved in parallel is presented in Algorithm 1. Each region n formulates its own local DC-OPF problem as in (14). The objective is to minimize the summation of region ns generation costs and the penalty functions subject to the local equality (i.e., (2) and (3)) and inequality (i.e., (4) and (5)) constraints of that region. In the first step of the solution procedure, all the voltage angles of the border buses (i.e., shared variables sent by the coordinator to the regions, $\hat{\boldsymbol{\Phi}}$) and the penalty multipliers $\boldsymbol{\lambda}$ and $\boldsymbol{\beta}$ corresponding to the penalty functions are initialized. Then, each region solves its local DC-OPF problem. Note this solution step is not sequential and all the regions can solve their local optimization problems in parallel. When all local OPF problems are solved, each region n sends the updated values of the shared variables with its neighbors (note that only a few shared variables may appear in the optimization problem of each region) to the central coordinator. Upon receiving all the shared variables from all the regions, the central coordinator solves (15) to coordinate the shared variables. Then, the coordinator checks whether convergence is achieved (i.e., difference between the shared variables is less than a threshold). If so, the coordinator propagates the updated values of the shared variables toward the regions and the solution procedure stops; otherwise, the penalty multipliers $\boldsymbol{\lambda}$ and $\boldsymbol{\beta}$ are updated according to Steps 7 and 8 of Algorithm 1 and the local OPF problems are solved again using the updated values of the shared variables and penalty multipliers. This procedure iteratively continues until achieving convergence (see [54] for more details). Note that the tuning parameter α in Algorithm 1 needs to be no smaller than one in order to achieve convergence.

Algorithm 1 Solution Algorithm of ATC-OPF

1: **Initialization** Initialize $\hat{\boldsymbol{\Phi}}^*$, $\boldsymbol{\lambda}$, $\boldsymbol{\beta}$, and set $k = 0$
2: **while** $|\hat{\boldsymbol{\Phi}}^{*k} - \boldsymbol{\Phi}^{*k}| > \epsilon$, $k = k + 1$ **do**
3: Solve (14) for all regions in parallel and determine the values of $\boldsymbol{x}_n$ and $\boldsymbol{\Phi}^{*k}$
4: Send $\boldsymbol{\Phi}^{*k}$ up to the central coordinator
5: Determine $\hat{\boldsymbol{\Phi}}^{*k}$ by (15)
6: Propagate $\hat{\boldsymbol{\Phi}}^{*k}$ down to the regions
7: Update $\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + 2(\boldsymbol{\beta}^k)^2(\hat{\boldsymbol{\Phi}}^{*k} - \boldsymbol{\Phi}^{*k})$
8: Update $\boldsymbol{\beta}^{k+1} = \boldsymbol{\alpha} \cdot \boldsymbol{\beta}^k$
9: **end while**

IV. Alternating Direction Method of Multipliers

A. Overview

The alternating direction method of multipliers (ADMM) is a well-known method for the solution of convex optimization problems in a distributed manner, in particular large-scale problems such as OPF and unit commitment (UC) in power grids [58]. ADMM is based on a decomposition-coordination process, in which a large-scale problem is divided into smaller subproblems and the solution procedures of the small subproblems are coordinated such that the overall process achieves the optimal solution of the original problem [27], [58]–[63].

In a nutshell, ADMM basically minimizes the augmented Lagrangian using a single Gauss-Seidel pass instead of a usual joint optimization [58]. In order to implement the ADMM method, we first derive the augmented Lagrangian function of the original optimization problem. Then, we decompose the obtained augmented function, and we minimize the corresponding augmented Lagrangian functions over sequential Gauss-Seidel iterations. At each iteration, at first, each subproblem (decomposed augmented function) is only minimized with respect to a subset of primal variables over which it is decomposed. Then, the dual variables are updated using the updated primal values. The iterative process continues until the convergence criteria is satisfied. Note that here we implement the most popular version of the ADMM [58] which distributes the computation between several areas and requires coordination between all areas. However, some efforts attempt to relax the need of a central coordinator and make ADMM a decentralized algorithm [27], [64]–[66]. In this paper, we have focused on the most common implementation of ADMM [58].

ADMM is applied to several energy management functions in power systems, e.g., optimal generation scheduling [27], [63], [67]–[73], demand response [62], and state estimation [60]. References [27] and [70] present decentralized OPF algorithms using ADMM, where subproblems directly communicate with each other with no need of a central coordinator. In [73], a decentralized OPF is proposed based on ADMM combined with sequential convex approximations. The centralized OPF is decomposed into subproblems corresponding to each bus, where solutions of the subproblems are coordinated with a light communication protocol among the buses.

B. ADMM-Based Distributed OPF Algorithm

We again refer to Fig. 3 which illustrates two areas that are connected through a tie line (connecting buses b and b'). The ADMM method introduces copies of coupling variables for each area to decouple the coupling constraints, while enforcing agreement between the copies of the coupling variables using consistency constraints. The stack of the coupling variables is again denoted by $\boldsymbol{\Phi}$ (shared vector). The components of this coupling vector are updated in a distributed fashion.

In the DC-OPF problem, the nodal power balance equation is a coupling constraint since it involves voltage phase angles of physically connected neighboring buses. In order to decouple the power balance equation over tie-lines, each area is holding a copy of voltage angles associated with neighboring buses that are located in neighboring areas. Introducing copies of shared variables enables an autonomous solution approach for each subproblem. Therefore, we can assign each subproblem to an area. However, a coordinating entity has to ensure that all local copies are reaching consensus over the course of iterations. For bus b located in area n, we introduce B_{bn} as the set of buses of neighboring areas that are physically connected to bus b. The set B_{bn} also includes bus b itself. The vector of voltage angle copies associated with buses in B_{bn} is denoted by $\boldsymbol{\theta}_{bn} \in \boldsymbol{R}^{|B_{bn}|}$. Also, we consider $\boldsymbol{\Phi}$ as network variables (consisting of voltage angles of the border buses). Note that all areas have access to the copies of network variables. The consistency constraints are given by

$$\boldsymbol{\theta}_{bn} = \boldsymbol{M}_{bn}\boldsymbol{\Phi} \tag{16}$$

where $\boldsymbol{M}_{bn} \in \boldsymbol{R}^{|B_{bn}|\times B}$ is called sectionalization matrix. An element of $\boldsymbol{M}_{bn}$ is equal to 1 in the case that bus b is in region (subsystem) n, and zero otherwise.

The ADMM approach is applied to the modified OPF problem that is the original OPF problem with additional consistency constraints (16). The ADMM basically minimizes the decomposed augmented Lagrangian functions associated with each area in an iterative manner. The augmented Lagrangian of the modified OPF problem which includes the consistency constraint (16) is given by

$$L_\rho(\boldsymbol{p},\boldsymbol{\theta},\boldsymbol{\Phi},\boldsymbol{\lambda}) = \sum_n \sum_{u\in\Omega_{Gn}} f(p_u) + \sum_n \sum_{b\in B_{bn}} \left(\boldsymbol{\lambda}_{bn}^{\dagger}(\boldsymbol{\theta}_{bn} - \boldsymbol{M}_{bn}\boldsymbol{\Phi}) + \frac{\rho}{2}\|\boldsymbol{\theta}_{bn} - \boldsymbol{M}_{bn}\boldsymbol{\Phi}\|_2^2\right) \tag{17}$$

where $\boldsymbol{\lambda}_n$ is the dual variable associated with (16) and ρ is called the penalty parameter. In the first term of (17), the first summation indicates the number of areas and the second summation is on the nodes in each area. We let $\boldsymbol{\lambda}_{bn}^k$ and $\boldsymbol{\Phi}^k$ denote the value of the dual and global variables at iteration k. In particular, $\boldsymbol{\theta}_{bn}^{k+1}$ is obtained as part of the solution of the following constrained problem solved in each area n:

$$\min_{\left(\boldsymbol{p}_n^{k+1},\boldsymbol{\theta}_n^{k+1}\right)} \sum_{u\in\Omega_{Gn}} f\left(p_u^{k+1}\right) + \sum_{b\in B_{bn}} \left(\boldsymbol{\lambda}_{bn}^{k\dagger}\left(\boldsymbol{\theta}_{bn}^{k+1} - \boldsymbol{M}_{bn}\boldsymbol{\Phi}^{*k}\right) + \frac{\rho}{2}\left\|\boldsymbol{\theta}_{bn}^{k+1} - \boldsymbol{M}_{bn}\boldsymbol{\Phi}^{*k}\right\|_2^2\right) \tag{18}$$

where the constraint set includes local power flow constraints and limits on the generation and line flows within area n. Given the dual variables $\boldsymbol{\lambda}_{bn}^k$ and $\boldsymbol{\theta}_{bn}^{k+1}$, the network variable $\boldsymbol{\Phi}$ update in iteration $k+1$ is obtained from the solution of the following unconstrained problem:

$$\min_{\boldsymbol{\Phi}^{k+1}} \sum_n \sum_{b\in B_{bn}} \boldsymbol{\lambda}_{bn}^{k\dagger}\left(\boldsymbol{\theta}_{bn}^{*k+1} - \boldsymbol{M}_{bn}\boldsymbol{\Phi}^{k+1}\right) + \frac{\rho}{2}\sum_n \sum_{b\in B_{bn}} \left\|\boldsymbol{\theta}_{bn}^{*k+1} - \boldsymbol{M}_{bn}\boldsymbol{\Phi}^{k+1}\right\|_2^2 \tag{19}$$

Moreover, we can find an analytic solution to (19) and $\boldsymbol{\Phi}^{k+1}$ can therefore be expressed as,

$$\boldsymbol{\Phi}^{k+1} = \left(\sum_n \sum_{b\in B_{bn}} \boldsymbol{M}_{bn}^T\boldsymbol{M}_{bn}\right)^{-1} \times \sum_n \sum_{b\in B_{bn}} \boldsymbol{M}_{bn}^T\left(\boldsymbol{\theta}_{bn}^{*k+1} + \frac{1}{\rho}\boldsymbol{\lambda}_{bn^k}\right) \tag{20}$$

The iterative steps of the ADMM are given in the pseudo code in Algorithm 2. Each region n formulates its local DC-OPF problem as (18). The objective is to minimize the

Algorithm 2 Solution Algorithm of ADMM-OPF

1: **Initialization** Initialize dual variable $\boldsymbol{\lambda}_b^{n(0)}$ and network phase angle vector $\boldsymbol{\Phi}^{(0)}$, and set $k = 0$
2: **while** Not converged, $k = k + 1$ **do**
3: Solve (18) for all regions in parallel and determine the values of $\boldsymbol{p}^k$ and $\boldsymbol{\theta}^k$
4: Solve (19) (or use (20)) to update $\boldsymbol{\Phi}^k$
5: $\boldsymbol{\lambda}$ update: $\boldsymbol{\lambda}_{bn}^k = \boldsymbol{\lambda}_{bn}^{k-1} + \rho\left(\boldsymbol{\theta}_{bn}^k - \boldsymbol{M}_{bn}\boldsymbol{\Phi}^k\right)$, for $b \in \boldsymbol{B}_{bn}$
6: **end while**

summation of region ns generation costs and the augmented Lagrangian functions corresponding to the shared variables modeling (a penalty term is needed for each share variable). The constraints of the optimization problem of region n are the local equality (i.e., (2) and (3)) and inequality (i.e., (4) and (5)) constraints. In the first step of the solution procedure, all the voltage angles of the border buses (i.e., network phase angle vector $\boldsymbol{\Phi}$) and the penalty multipliers $\boldsymbol{\lambda}$ corresponding to the augmented Lagrangian functions are initialized. Then, the regions solve their local DC-OPF problem in parallel. When all local OPF problems are solved, each region n sends the updated values of the shared variables with its neighbors to the central coordinator. Upon receiving all the shared variables from all the regions, the central coordinator solves (19) (or (20)) to coordinate the shared variables. Then, the convergence criteria are evaluated to check whether the difference between each pair of the shared variables (note that a shared variable is computed by two regions, hence, we have a pair for each shared variable) less than a predetermined threshold. If so, the coordinator propagates the updated values of the shared variables toward the regions and the solution procedure stops; otherwise, the penalty multipliers $\boldsymbol{\lambda}$ are updated according to Step 5 of Algorithm 2 and the local OPF problems are solved again using the updated values of the shared variables and penalty multipliers. This procedure iteratively continues until achieving convergence (we refer to [58] for more details).

V. Proximal Message Passing

A. Overview

The proximal message passing algorithm is a particular application of ADMM where minimizing at each Gauss-Seidel pass amounts to evaluating the prox functions with respect to the particular decision vector partition (please refer to Section IV for more detail on ADMM) [59]. The classic ADMM is not really a parallel algorithm, because the multipliers updates are performed in a sequential manner. However, the ADMM based proximal message passing (PMP) presented in this section is completely decentralized and highly scalable with respect to the size of the network. The PMP algorithm has been applied to the OPF and security constrained OPF (SCOPF) problems in [59] and [61], respectively. In this approach, the power grid is first split into a bipartite graph consisting of devices, nets, and terminals connecting the former two entities. Estimated values for power injections and voltage angles for devices and nodes at a particular iteration, called messages, are then exchanged between neighboring nodes and devices. The iterations proceed in a manner such that differences between messages are gradually minimized while operation costs of all devices over a given time horizon are optimized.

In the proximal message passing algorithm, prox functions are defined as follows:

$$\mathbf{prox}_{f,\rho}(v) = \arg\min_x \Big(f(x) + (\rho/2)\|x - v\|_2^2\Big). \tag{21}$$

In the above equation, the right hand side is the prox function of $f(x)$ (which represents the device cost function; fuel cost if it's a generator and cost to violate the constraints if it's a load or a transmission line) with respect to x. ρ is the tuning parameter for the augmented Lagrangian. The term v consists of the previous values of x, which in our application is power/voltage angle, node average or mismatch of the corresponding variables and the average node price (or dual variables) for the corresponding variables. The messages are the updates of the primal and dual variables from each Gauss-Seidel pass that are passed to the next stage or next iteration. Evaluating the prox function for any device implies to come up with a set of decision variable updates for the device such that they are closest to the previous iterate as well as closest to satisfying the node power mismatch and angle mismatch constraints, while also minimizing the cost function of the device. Node power mismatch constraint encodes the fact that the sum of all the incoming and outgoing powers on all the devices connected to any particular node will be zero. If not, there is a mismatch, which needs to be mitigated over successive iterations. Similarly, angle mismatch constraint entails that the voltage phase angles of all the devices connected to any particular node be all equal to each other.

B. PMP-Based Decentralized OPF Algorithm

As discussed in the previous section, in the OPF, at each iteration, we are evaluating the prox-functions. We actually split the objective function across the different devices (generators, transmission lines, and loads). The decision variables are the power injections and voltage angles at each terminal of those devices (one for generators and loads, two for transmission lines). Although the Gauss-Seidel passes are sequential, elimination of some intermediate variables (refer to [61] and [59] for the details), makes the algorithm parallel over all devices, followed by a sequential update of dual variables for power balance and voltage angle consistency at the nodes (all nodes in parallel). The $f(x)$ function for the generators are their actual cost functions whereas for loads and transmission lines, it's just the indicator function corresponding to satisfying Kirchhoff's laws. We state below (without proof) just the real power counterparts of the update equations:

1) *Proximal plan updates:*

$$p_j^{k+1} = \mathbf{prox}_{f_j,\rho}\Big(p_j^k - \bar{p}_j^k - u_j^k\Big) \tag{22}$$

1) *Scaled price updates:*

$$u_{N_i}^{k+1} := u_{N_i}^k + \bar{p}_{N_i}^{k+1} \tag{23}$$

where p, $\bar{p}$, u, k, ρ, j, and N_i are respectively the real power, average of real power, real power nodal price, iteration count, tuning parameter, device subscript, and i^{th} node subscript. The details of the different equations used for the calculations at each step of the iteration are as follows:

$$\left(p_{g_q t_v}^{(k+1)}, \theta_{g_q t_v}^{(k+1)}\right) = \mathbf{prox}_{f(\mathbf{p}),\rho}\left(p_{g_q t_v}^{(k)} - \overline{p}_{g_q t_v}^{(k)} - u_{g_q t_v}^{(k)}, \overline{v}_{g_q t_v}^{(k-1)} + \overline{\theta}_{g_q t_v}^{(k)} - v_{g_q t_v}^{(k)}\right) \quad \forall g_q \in \Omega_G \tag{24}$$

$$\left(p_{T_l t_v}^{(k+1)}, \theta_{T_l t_v}^{(k+1)}, p_{T_l t_{v'}}^{(k+1)}, \theta_{T_l t_{v'}}^{(k+1)}\right) = \mathbf{prox}_{F+X,\rho}\left(p_{T_l t_v}^{(k)} - \overline{p}_{T_l t_v}^{(k)} - u_{T_l t_v}^{(k)}, \overline{v}_{T_l t_v}^{(k-1)} + \overline{\theta}_{T_l t_v}^{(k)} - v_{T_l t_v}^{(k)}\right) \quad \forall T_l \in \Omega_L \tag{25}$$

$$\left(p_{D_d t_v}^{(k+1)}, \theta_{D_d t_v}^{(k+1)}\right) = \mathbf{prox}_{-p_{D_d},\rho}\left(\overline{v}_{D_d t_v}^{(k-1)} + \overline{\theta}_{D_d t_v}^{(k)} - v_{D_d t_v}^{(k)}\right), \forall D_d \in \Omega_D \tag{26}$$

$$u_{N_i t_v}^{(k+1)} = u_{N_i t_v}^{(k)} + \overline{p}_{N_i t_v}^{(k+1)}, \quad \forall N_i \in \Omega_B \tag{27}$$

$$v_{N_i t_v}^{(k+1)} = \tilde{v}_{N_i t_v}^{(k)} + \tilde{\theta}_{N_i t_v}^{(k+1)}, \quad \forall N_i \in \Omega_B \tag{28}$$

where Ω_G, Ω_L, Ω_D, and Ω_B are the sets of generators, transmission lines, and loads, buses/nets/nodes, and θ denotes the bus voltage phase angles. v represent the Lagrange multipliers corresponding to the voltage phase angle consistency constraints at the nodes, which states that the voltage phase angle at each terminal connected to a particular node is equal to the average voltage phase angle of that particular node. Indices q, d, and l respectively indicate generator, load, and transmission line. θ is the bus voltage angle. Function f is the cost function of the generators, and function F represents indicator functions corresponding to satisfying the transmission line limits as well as the relation between power flow on lines, and voltage angles at the ends. t_v represents the v^{th} terminal. For instance, $\tilde{v}_{N_i t_v}^{k}$ indicates the Lagrange multiplier corresponding to the power balance constraint at node N_i associated with the terminal, t_v at iteration k. According to the above set of equations, the devices perform their updates in parallel according to equations (24)-(26) and then the nets or nodes do their updates in parallel according to the equations (27) and (28).

C. Stopping Criterion

We can define primal and dual residuals for the prox-project message passing algorithm as

$$r^k = \bar{p}^k \tag{29}$$

$$s^k = \rho\left(\left(p^k - \bar{p}^k\right) - \left(p^{k-1} - \bar{p}^{k-1}\right)\right) \tag{30}$$

Here p^k is interpreted as a power plan, i.e., the power iterate at the k^{th} iteration and $\bar{p}^k$ is the average power of a node in the previous iteration. There exists exactly the same counterparts for the bus voltage angle also (which, for the sake of brevity, is not shown here). So, the primal residual effectively indicates what is the node power and angle mismatch and the dual residual indicates how much that mismatch varies from the previous iteration to the present. In case of convergence, we would expect both the metrics to gradually approach zero. A simple terminating criterion for prox-project message passing is when

$$\left\|r^k\right\|_2 \leq \epsilon^{\text{pri}}, \qquad \left\|s^k\right\|_2 \leq \epsilon^{\text{dual}} \tag{31}$$

where ϵ^{pri} and ϵ^{dual} are respectively primal and dual tolerances.

Algorithm 3 Solution Algorithm of PMP-OPF

1: **Initialize** ρ, u, v, p, voltage angles, and set $k = 0$
2: **while** $\|r^k\| \geq \epsilon^{pri} \& \|s^k\| \geq \epsilon^{dual}$, $k = k + 1$ **do**
3: **Broadcast** iterates to devices (Generators, Transmission lines, Loads)
4: Solve generators' optimization problems for updated power and angle
5: Solve lines' optimization problems for updated power and angle
6: Solve loads' optimization problems for angle
7: Gather iterates at nodes
8: Calculate average power, average voltage angle, u, and v for each node, and compute r and s and go back to step 2
9: **end while**

The iterative algorithm to solve the PMP-based decentralized OPF is presented in Algorithm 3. In PMP, the OPF is decomposed at the device level, where each device (generator, transmission line, load, etc.) solves its local DC- OPF problem as evaluating prox functions shown in equations (24)-(26). The objective is to minimize the summation of device's objective functions. This is a summation of actual generation costs and augmented Lagrangian penalty terms for not satisfying the constraints corresponding to the Kirchoff's laws for power balance at the buses and the equality of bus voltage phase angles for the different devices connected to a particular bus. The constraints of the optimization problem for each device are the power flow limit for transmission lines, power generation limits for generators, constancy of the power demand for loads, and the voltage phase angle-power flow relationship for all the devices. In the first step of the solution procedure, all the power and voltage angle variables as well as the penalty multipliers u and v corresponding to the constraints are initialized. Then, the devices evaluate their local prox functions in parallel. When all local problems are solved, each device sends the updated values of the power and angle variables along with its neighbors to the respective node(s) to which it is connected, through a process called gather. Upon receiving all the variable updates from all the devices, the particular node concerned solves (27) and (28) to evaluate the quality of the device's decision variable updates. Then, the convergence criteria (29)-(30) are evaluated to check whether a consensus has been reached regarding the values of the decision variables. If so, the solution procedure stops; otherwise, the nodes or buses first sends out the updated values of u and v to the devices that are connected to them through a process called broadcast and then we go back to Step 2 and repeat the entire process until convergence. We refer to [59] and [61] for more details.

VI. Auxiliary Problem Principle

A. Overview

Auxiliary Problem Principle was developed in [74] to find the solution of an optimization problem by solving a sequence of auxiliary problems. Since then, APP has been widely used in the literature to solve optimization problems in a decentralized and iterative manner. For example, in [11], [12], [36], [75], and [76], APP is applied to enable a decentralized OPF implementation in large-scale power systems. First, the power grid is divided into several geographical regions. Then, dummy buses are added in all tie-lines that connect the region to its neighbors. A set of variables for voltage magnitude and angle, active and reactive power flows at these dummy buses is assigned to each region. A local OPF is formulated for each region, and an iterative solution algorithm is introduced to find the optimal operating point of the entire grid. The corresponding variables at the dummy buses need to be equal once the algorithm converged. Reference [77] presents a decentralized generation scheduling approach for energy market clearing in multi-regional power systems based on APP. At first, a conventional unit commitment problem is solved to determine the on/off states of generating units. Then, the APP algorithm is applied for the decentralized OPF implementation among multiple geographical regions.

Assume that we decompose an optimization problem with coupling variables into subproblems. The coupling variables are again duplicated and then one of the duplicated copies is assigned to each of the subproblems that are coupled by these coupling variables. As each of the two corresponding duplicated variables must have the same value, a consistency constraint is imposed on the subproblems forcing the difference between the coupling variable and its corresponding duplicated variable to be zero. We now again relax the consistency constraints by an Augmented Lagrangian approach. Then, according to the auxiliary problem principle, we linearize the cross-terms in the Augmented Lagrangian [74]. To coordinate the coupling variables between the subproblems and determine the optimal operating point of the original optimization problem, an iterative solution procedure is carried out where the subproblems exchange the values for the duplicated variables with each other at the end of every iteration. In each iteration of the APP approach, a subproblem needs to know the values of the coupling variables determined by other subproblems in the previous iteration. Therefore, the APP can be implemented in a parallel manner.

Convergence properties of the APP are discussed in [74]. If the objective function of each subproblem is convex and differentiable, the APP algorithm converges to the solution of the original optimization problem regardless of the starting point. Note that the APP can also be used for practical nonconvex optimizations. In this case, if the APP algorithm converges, necessary conditions of overall optimality are satisfied [74].

B. APP-Based Decentralized OPF Algorithm

For the application of APP to the considered DC-OPF problem, consider again Fig. 3 where regions m and n are linked together through a tie-line connecting buses b and b'. Angles θ_b and $\theta_{b'}$ are the coupling variables between the regions. We duplicate these variables to produce two sets of coupling variables. One set is assigned to region n and the other set to region m. These sets are denoted by $\boldsymbol{\Phi}_{bb',m}$ and $\boldsymbol{\Phi}_{bb',n}$. Note that we have a set of shared variables for each tie-line connecting two neighboring areas. In order to simplify notation, we neglect the subscript bb' in the following DC-OPF problem formulation corresponding to region m at iteration $k+1$:

$$\min_{\left(\boldsymbol{x}_m^{k+1},\boldsymbol{\Phi}_m^{k+1}\right)} \sum_{u\in\Omega_{Gm}} f\left(p_u^{k+1}\right) + \left(\frac{\boldsymbol{\rho}}{2}\left\|\left(\boldsymbol{\Phi}_m^{k+1}-\boldsymbol{\Phi}_m^{*k}\right)\right\|^2 + \boldsymbol{\gamma}\boldsymbol{\Phi}_m^{(k+1)\dagger}\left(\boldsymbol{\Phi}_m^{*k}-\boldsymbol{\Phi}_n^{*k}\right)+\boldsymbol{\lambda}^{k\dagger}\boldsymbol{\Phi}_m^{(k+1)}\right) \tag{32}$$

s.t.

$$\boldsymbol{h}_m\left(\boldsymbol{x}_m^{k+1},\boldsymbol{\Phi}_m^{k+1}\right)=0$$

$$\boldsymbol{g}_m\left(\boldsymbol{x}_m^{k+1},\boldsymbol{\Phi}_m^{k+1}\right)\le 0$$

$$\boldsymbol{x}_m^{k+1}=\left\{\boldsymbol{p}_m^{k+1},\boldsymbol{\theta}_m^{k+1}\right\},\ \boldsymbol{\Phi}_m^{k+1}=\begin{bmatrix}\boldsymbol{\theta}_{b,m}^{k+1}\\ \boldsymbol{\theta}_{b',m}^{k+1}\end{bmatrix},\ \boldsymbol{\Phi}_n^{*k}=\begin{bmatrix}\boldsymbol{\theta}_{b,n}^{*k}\\ \boldsymbol{\theta}_{b',n}^{*k}\end{bmatrix}$$

where $\boldsymbol{x}_m$ is the set of local variables, namely power outputs of generators and voltage angles of internal buses, of region m. $\boldsymbol{\lambda}^k$ is the vector of Lagrange multipliers at iteration k and $\boldsymbol{\rho}$ and $\boldsymbol{\gamma}$ are suitable positive constants. $\boldsymbol{\Phi}_n^{*k}$ and $\boldsymbol{\Phi}_m^{*k}$ indicate the values of variables $\boldsymbol{\Phi}_n$ and $\boldsymbol{\Phi}_m$ that are respectively predetermined in iteration k by region n and m. The objective function for region n can be similarly formulated as:

$$\min_{\left(\boldsymbol{x}_n^{k+1},\boldsymbol{\Phi}_n^{k+1}\right)} \sum_{u\in\Omega_{Gn}} f\left(p_u^{k+1}\right) + \left(\frac{\boldsymbol{\rho}}{2}\left\|\left(\boldsymbol{\Phi}_n^{k+1}-\boldsymbol{\Phi}_n^{*k}\right)\right\|^2 - \boldsymbol{\gamma}\boldsymbol{\Phi}_n^{(k+1)\dagger}\left(\boldsymbol{\Phi}_n^{*k}-\boldsymbol{\Phi}_n^{*k}\right)-\boldsymbol{\lambda}^{k\dagger}\boldsymbol{\Phi}_n^{(k+1)}\right) \tag{33}$$

The penalty multiplier $\boldsymbol{\lambda}$ needs to be updated in each iteration according to

$$\boldsymbol{\lambda}^{k+1}=\boldsymbol{\lambda}^k+\alpha\left(\boldsymbol{\Phi}_m^{*k+1}-\boldsymbol{\Phi}_n^{*k+1}\right) \tag{34}$$

where α is a suitable positive value. Note that the value of the Lagrange multiplier $\boldsymbol{\lambda}$ in each iteration corresponds to the cost to maintain the consistency constraint. The above formulation can be generalized for a power system including multiple regions.

In Algorithm 4, the pseudo code for the implementation to solve the APP-based decentralized OPF is given. Each region formulates its own local DC-OPF problem. The objective is to minimize the summation of the regions generation costs and the penalty functions subject to the local equality (i.e., (2) and (3)) and inequality (i.e., (4) and (5)) constraints of that region. In the first step of the solution procedure, all the voltage angles of the border buses (i.e., shared variables among the neighboring regions, $\boldsymbol{\Phi}^*$) as well as the set of penalty

Algorithm 4 Solution Algorithm of APP-OPF

1: **Initialization** Initialize $\boldsymbol{\Phi}_m^*$, $\boldsymbol{\Phi}_n^*$, $\boldsymbol{\lambda}$, α, $\boldsymbol{\rho}$, $\boldsymbol{\gamma}$, and set $k = 0$
2: **while** $\boldsymbol{\Phi}_m^{*k} - \boldsymbol{\Phi}_n^{*k} > \epsilon$, $k = k + 1$ **do**
3: Solve local OPF of all regions in parallel, and determine the optimal values of $\boldsymbol{x}_n$ and $\boldsymbol{\Phi}^{*k}$
4: Exchange $\boldsymbol{\Phi}_m^{*k}$ and $\boldsymbol{\Phi}_n^{*k}$ between the regions
5: Update $\boldsymbol{\lambda}^k = \boldsymbol{\lambda}^{k-1} + \alpha(\boldsymbol{\Phi}_m^{*k} - \boldsymbol{\Phi}_n^{*k})$
6: **end while**

multipliers $\boldsymbol{\lambda}$ and positive constants α, $\boldsymbol{\gamma}$, and $\boldsymbol{\rho}$ corresponding to the penalty functions are initialized. Then, each region solves its local DC-OPF problem. Note this solution step is not sequential and all the regions can solve their local optimization problems in a parallel way. When all local OPF problems are solved, each region sends the updated values of the shared variables to its immediate neighboring regions. If the differences among each shared variable computed by a region and its corresponding shared variable computed by another region are less than a predetermined threshold, the convergence is achieved. Otherwise, the set of penalty multipliers $\boldsymbol{\lambda}$ are updated according to Step 5 of Algorithm 4 and the local OPF problems are solved again using the updated values of the shared variables and penalty multipliers. This procedure iteratively continues until achieving convergence. Reference [36] provides more details on the solution procedure and discusses some sufficient conditions for this iterative algorithm to converge to the optimal solution.

VII. Optimality Condition Decomposition

A. Overview

Optimality condition decomposition (OCD) (see [13], [18], [78], [79]) is an extension of the Lagrangian relaxation method which has shown improved convergence properties over Lagrangian Relaxation and Augmented Lagrangian methods. References [13] and [78] describe the OCD method applied to solve a multi-area optimal power flow problem in a decentralized manner. The proposed method is an iterative solution procedure where in each iteration the OPF problem of each region is formulated and solved independently. Variables appearing in the subproblem of a region but within the responsibility of another area are fixed to the values obtained in the previous iteration. In [18], the OCD method is applied and extended to overlapping areas in power systems for the coordination of FACTS devices. The OCD approach is applied to find a decentralized solution for the multi-area interconnected DC-OPF in [14], where the OPF problem is decoupled at tie-lines connecting neighboring areas.

OCD is suitable for solving optimization problems with coupling constraints. In OCD, the first-order optimality constraints associated with the overall problem are decomposed and solved by subproblems. At each iteration, the variables are updated by solving for the optimality conditions using a Newton-Raphson step, and the updated variables associated with the boundary buses are exchanged among the subproblems.

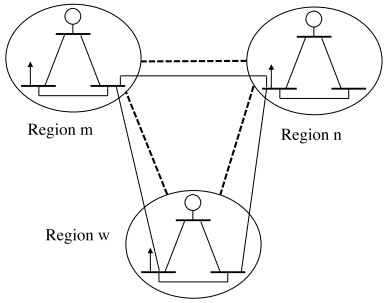

Fig. 5. Decentralized structure for implementation of OCD.

The subsystems in OCD only interact with their directly connected neighboring subsystems; i.e., no centralized coordinator is needed. An illustrative structure for implementing OCD is given in Fig. 5 where the solid lines denote the tie-lines connecting subsystems and the dotted lines denote the communication links among subsystems. However, it should be mentioned that OCD, same as actually any other decentralized method, can also be implemented in the current centralized structure of the power system as shown in Fig. 4(a) where the centralized coordinator/controller can act as a relay node to deliver the information from one subsystem to the other.

B. OCD-Based Decentralized OPF Algorithm

1) Basic OCD Method: In this section, we provide the detailed algorithm of OCD and explain how it can be applied to solving the OPF problem. To implement OCD, the first step is to assign the coupling constraints to a specific subproblem and formulate the nth subproblem as follows:

$$\min_{\boldsymbol{x}_n} \sum_{u\in\Omega_{Gn}} f(p_u) + \sum_{m=1,m\neq n}^{N} \boldsymbol{\lambda}_m^{*\dagger} c_m\left(\boldsymbol{x}_1^*, \ldots, \boldsymbol{x}_n, \ldots, \boldsymbol{x}_N^*\right) \tag{35}$$

s.t.

$$h_n^c\left(\boldsymbol{x}_1^*, \ldots, \boldsymbol{x}_n, \ldots, \boldsymbol{x}_N^*\right) = 0 \tag{36}$$

$$g_n^c\left(\boldsymbol{x}_1^*, \ldots, \boldsymbol{x}_n, \ldots, \boldsymbol{x}_N^*\right) \leq 0 \tag{37}$$

$$h_n^s(\boldsymbol{x}_n) = 0 \tag{38}$$

$$g_n^s(\boldsymbol{x}_n) \leq 0 \tag{39}$$

$$\boldsymbol{x}_n \in \left\{\boldsymbol{p}_n, \boldsymbol{\theta}_n\right\}$$

where N is the total number of subproblems or areas, n is the index of the subproblem, and x_n includes the variables assigned to subproblem n. Constraints (36) and (37) are the coupling constraints assigned to subproblem n, and (38) and (39) are the non-coupling constraints included in subproblem n. All the coupling constraints assigned to other areas ($c_m = \{h_m^c(.), g_m^c(.)\}, m \neq n$) are included in the objective function of area n and $\boldsymbol{\lambda}_m, m \neq n$ denotes the Lagrange multipliers of those coupling constraints. The inequality constraints can be handled with an Interior Point approach. The variables with an asteroid indicate that their values are fixed to their last updated values.

Subproblem (35) to (39) can be solved by first deriving its associated KKT conditions and then applying the Newton-Raphson method. In particular, subproblem n solves

the following system of equations at each iteration to find the update Δz_n:

$$\Delta z_n = -\boldsymbol{H}_{nn}^{-1} \cdot \boldsymbol{r}_n \tag{40}$$

$z_n = \{\boldsymbol{x}_n, \boldsymbol{\lambda}_n\}$ denotes all the primal and dual variables to be updated by subproblem n and, $\boldsymbol{r}_n$ denotes the KKT conditions to be fulfilled in subproblem n, and $\boldsymbol{H}_{nn}$ denotes the Jacobian of the KKT conditions in region n. It is worth mentioning that only a small amount of variables need to be exchanged which are the variables associated with the buses at the boundaries of subsystems including the voltage angles, and the Lagrangian Multipliers of the power balance constraints at the border buses. A sufficient condition for OCD to converge is given in [13] as follows:

$$c = \rho\left(\boldsymbol{I} - \widetilde{\boldsymbol{H}}^{-1}\boldsymbol{H}\right) < 1 \tag{41}$$

where $\boldsymbol{I}$ is the identity matrix and $\rho(\boldsymbol{M})$ denotes the spectral radius of matrix $\boldsymbol{M}$. $\boldsymbol{H}$ and $\widetilde{\boldsymbol{H}}$ have the following form:

$$\boldsymbol{H} = \begin{pmatrix} H_{11} & \dots & H_{1N} \\ \vdots & \ddots & \vdots \\ H_{N1} & \dots & H_{NN} \end{pmatrix} \tag{42}$$

$$\widetilde{\boldsymbol{H}} = \begin{pmatrix} H_{11} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & H_{NN} \end{pmatrix} \tag{43}$$

$$\boldsymbol{H}_{mn} = \nabla^2 L_{z_m, z_n}; m, n = 1, \dots, N. \tag{44}$$

Notice that z_n is used to indicate both the primal variables $\boldsymbol{x}_n$ and Lagrange multipliers $\boldsymbol{\lambda}_n$ in subproblem n. $\boldsymbol{H}$ is the Hessian matrix of the Lagrangian function L of the entire optimization problem evaluated at the optimal point where the variables are arranged according to the subproblems they are assigned to. $\widetilde{\boldsymbol{H}}$ is acquired by setting the off-diagonal blocks in $\boldsymbol{H}$ to zeros which removes the coupling between subproblems. Convergence criterion (41) holds with most geographical partitions of the power systems where the coupling among areas is relatively weak [13] and [80]. Also, a smaller c indicates less computational coupling between subproblems and generally results in fewer iterations until convergence is reached.

2) OCD With Correction Terms: In the procedure of OCD, the variable update ignores the coupling between subproblems, hence results in some error in the searching direction of the Newton Raphson iteration which may lead to a large number of iterations until convergence. To compensate for such error, an extended OCD, namely, OCD-C is proposed and applied to various case studies [80]–[82] where an additional correction term will be calculated and exchanged among subproblems. In OCD-C, the new updates can be calculated by

$$\Delta z_n = \boldsymbol{H}_{nn}^{-1} \cdot \left(-\boldsymbol{r}_n + \hat{\boldsymbol{r}}_n\right) \tag{45}$$

Here, $\hat{\boldsymbol{r}}_n$ is the correction term and it can be calculated by

$$\hat{\boldsymbol{r}}_n = \sum_{m=1, m\neq n}^{N} \boldsymbol{H}_{nm}\boldsymbol{H}_{mm}^{-1} \cdot \boldsymbol{r}_m \tag{46}$$

where each term in the summation can be calculated by one subsystem m and sent to subsystem n. The full derivation of the correction term and the convergence criterion of OCD can be found in [82]. Due to the sparsity of $\boldsymbol{H}_{mn}$ and $\boldsymbol{r}_m$, the correction term only contains few non-zero elements which results in little additional information to be exchanged. However, by adding the correction term, OCD-C in many cases converges in significantly fewer iterations than OCD as can also be seen in the following case study.

Algorithm 5 Solution Algorithm of OCD and OCD-C

1: **Initialization** Initialize all primal and dual variables, i.e., $z = \{\boldsymbol{x}, \boldsymbol{\lambda}\}$, for all regions
2: **while** $|\boldsymbol{r}| \geq \epsilon$ **do**
3: Solve (40) (or (45) for OCD-C) for all regions in parallel and determine Δz
4: Update $z \leftarrow z + \Delta z$
5: **if** OCD-C is implemented **then**
6: Calculate $\hat{\boldsymbol{r}}$ for all regions in parallel according to (46)
7: **end if**
8: Exchange among regions the updated z
9: **if** OCD-C is implemented **then**
10: Exchange among regions the correction term $\hat{\boldsymbol{r}}$
11: **end if**
12: **end while**

The pseudo code for OCD and OCD-C is given in Algorithm 5. At first, all primal and dual variables need to be initialized. Then, at each iteration, each region n formulates its subproblem (35)-(39) and carries out one Newton-Raphson step to update its variables z_n. For OCD-CïijŇan additional step is to calculate the correction terms $\hat{\boldsymbol{r}}$ for neighboring regions using (46). Note that in both OCD and OCD-C, local computations of the regions can be performed in parallel, i.e., while a region is solving its problem, others do not need to stay idle. After these local computations, the updated values of the boundary variables z (and the correction terms $\hat{\boldsymbol{r}}$ if OCD-C is used) are exchanged among neighboring regions in order to formulate the subproblems for the next iteration. This procedure continues until convergence is achieved, i.e., $|\boldsymbol{r}_n| \leq \epsilon \ \forall n$ (see [13] for more details).

VIII. Consensus+Innovations

A. Overview

The idea of exploiting neighborhood consensus potential to solve the optimal power management problems has recently received tremendous attention [83], [84]. Early examples include [33], [85]–[87], where authors proposed decentralized solutions for the economic dispatch problem (neglecting line constraints) by enforcing an agreement of the marginal cost through an iterative process using a consensus-based algorithm. However, since a line congestion would render the marginal costs unequal, a solely consensus type algorithm would not be able to handle the OPF problem in which we need to consider the line constraints. To handle this issue and find a decentralized solution for the optimal power flow problem, [88] and [89] suggested an innovation based approach. The core of this method is based on iterative

updates which solve the first order optimality conditions associated with the OPF problem. Hence, it reduces the original optimization problem into solving a coupled system of equations. Note, this procedure relays on the fact that optimality conditions merely involves local information. To be more specific, in this method, the optimality conditions are integrated as *innovation* updates in the variable updates. In the following, we will refer to this method as the C+I method as it reduces to a consensus update in case of an uncongested system. Note, the first order optimality conditions merely involve local information, hence, the variable updates in C+I for the considered problem could be performed in a fully decentralized fashion. In this approach, each bus just exchanges information with its physical neighbors, i.e., other buses that are physically connected through transmission lines. It should be mentioned that C+I is inherently different from decomposition theory methods [11]. The C+I solves directly the first order optimality conditions of OPF, while the decomposition theory approaches generally decompose the optimization problem, leaving each entity with an optimization problem to solve.

To improve the convergence speed of the algorithm presented in [89], the reference [90] proposed adding additional communication links between non-physically connected buses. These additional communication links speed up the information spread over the network, hence, lead to an improvement of the C+I's convergence rate. Moreover, [91] presented a more realistic version of the decentralized algorithm presented in [89], i.e., C+I, since it handles the communication delay challenge by clustering the buses into areas. The proposed method in [91] allows intra and inter-area communications. The intra-area information exchange indicates the communication between the buses in the same area which happens after each variable update iteration, while inter-area communications occurs between neighboring areas only every few iterations.

B. Consensus+Innovation Based Decentralized OPF Algorithm

The first order optimality conditions of the optimization problem for DC-OPF are given as:

$$2a_i p_i + b_i - \lambda_i + \mu_i^+ - \mu_i^- = 0 \tag{47}$$

$$\lambda_i \cdot \sum_{j\in\Omega_i} \frac{1}{X_{ij}} - \sum_{j\in\Omega_i} \lambda_j \frac{1}{X_{ij}} + \sum_{j\in\Omega_i} \big(\mu_{ij} - \mu_{ji}\big)\frac{1}{X_{ij}} = 0 \tag{48}$$

$$\theta_{ref} = 0 \tag{49}$$

$$p_i - d_i = \sum_{j\in\Omega_i} \frac{\theta_i - \theta_j}{X_{ij}} \tag{50}$$

$$\underline{P}_i \le p_i \le \overline{P}_i \tag{51}$$

$$\underline{PL}_{ij} \le \frac{\theta_i - \theta_j}{X_{ij}} \le \overline{PL}_{ij} \tag{52}$$

for all $i \in \{1, \ldots, N_B\}$ and $ij \in \Omega_L$. Parameters a_i and b_i are the quadratic and linear cost for generator i. If the bus is not connected to any generator, its associated p_i is zero. λ_i is the Lagrange multiplier associated with the power flow balance at bus i, i.e., (2), μ_i^+ and μ_i^- are the multipliers associated with the generation limits of generator i, i.e., (4), and μ_{ij} and μ_{ji} are the multipliers associated with the thermal line limits on line ij, i.e., (5). Additionally, the optimality conditions also include the complementary slackness conditions for the inequality constraints and the positivity constraints on the μ's. Since, DC-OPF is a convex problem with strong duality, the set of variables and Lagrangian multipliers that fulfills the optimality conditions is the optimal solution to the problem. Therefore, solving the discussed system of linear optimality conditions yields the optimal solution to the DC-OPF problem. The aforementioned optimality conditions form the basis for the C+I decentralized approach applied to DC-OPF [89].

The C+I needs each bus to update its estimate of generation settings and all Lagrangian multipliers associated with that bus (see [89] for more details). The Lagrange multipliers λ_i update is given as:

$$\lambda_i^{k+1} = \lambda_i^k - \beta \cdot \left(\lambda_i^k \sum_{j\in\Omega_i} \frac{1}{X_{ij}} - \sum_{j\in\Omega_i} \lambda_j^k \frac{1}{X_{ij}} + \sum_{j\in\Omega_i} \Big(\mu_{ij}^k - \mu_{ji}^k\Big)\frac{1}{X_{ij}} \right) - \alpha \cdot \left(\sum_{n\in\Omega_{G_i}} p_n^k - d_n - \sum_{j\in\Omega_i} \frac{\theta_i^k - \theta_j^k}{X_{ij}} \right), \tag{53}$$

where $\alpha, \beta > 0$ are constant tuning parameters, and Ω_{G_i} denotes the set of all generating units located on bus i. In (53), the first term, which corresponds to (48), maintains the coupling between the Lagrange multipliers and in case of uncongested lines (all $\mu_{ij} = 0$) reduces to a consensus term for the λ_i, while the second term adds the power balance equation (50) as an *innovation* term.

Given the value of λ_i, the update for generators $p_{G_n}, n \in \Omega_{G_i}$ can be carried out:

$$p_{G_n}^{k+1} = \left[\frac{\lambda_i^{k+1} - b_n}{2a_n} \right]_{\underline{P}_{G_n}}^{\overline{P}_{G_n}}, \tag{54}$$

where $[\cdot]_{\underline{P}_{G_n}}^{\overline{P}_{G_n}}$ is the projection operator which projects the value of p_{G_n} onto the feasible space, i.e., $[\underline{P}_{G_n}, \overline{P}_{G_n}]$. The usage of $[\cdot]_{\underline{P}_{G_n}}^{\overline{P}_{G_n}}$ in (54) is equivalent to the presence of multipliers μ_n^+ and μ_n^- in (47).

The bus angles are updated according to

$$\theta_i^{k+1} = \theta_i^k - \gamma \left(- \sum_{n\in\Omega_{G_i}} p_{G_n}^k + d_i + \sum_{j\in\Omega_i} \frac{\theta_i^k - \theta_j^k}{X_{ij}} \right), \tag{55}$$

with $\gamma > 0$ being a tuning parameter. Note, the power balance equation (50) is exploited to update θ_i in (55). The Lagrange multipliers μ_{ij} corresponding to the line limit of the line that connects bus i to j is updated according to

$$\mu_{ij}^{k+1} = \left[\mu_{ij}^k - \delta \cdot \left(\overline{P}_{ij} - \frac{\theta_i^k - \theta_j^k}{X_{ij}} \right) \right]^+ \tag{56}$$

Algorithm 6 Solution Algorithm of C+I OPF

Initialization Initialize variables $\boldsymbol{\lambda}$, $\boldsymbol{\theta}$, $\boldsymbol{\mu}$, and $\boldsymbol{p}_G$
while convergence criteria is not satisfied **do**
  for i=1:number of agents **do**
    Update $\boldsymbol{\lambda}$ using (53)
    Update $\boldsymbol{p}_G$ using (54)
    Update $\boldsymbol{\theta}$ using (55)
    Update $\boldsymbol{\mu}$ using (56)
    Communicate $\boldsymbol{\lambda}$ and $\boldsymbol{\theta}$ to neighboring agents
  end for
end while

Here $\delta > 0$ is a tuning parameter, and $\lceil\cdot\rceil^+$ represents the projection operator that guarantees the positivity of the μ_{ij}^{k+1}. The inequality (52) is used as an *innovation* term to update μ_{ij} in (56). Reference [89] additionally provides an analytical results showing that any fixed point of the C+I algorithm, (53)–(56), fulfills all of the optimality conditions (47)-(52) of the OPF problem. Given that DC-OPF is a convex problem with strong duality, any solution of the algorithm (limit point of the C+I update procedure) constitutes an optimal solution for the DC-OPF problem (see [89, Th. 2]).

The pseudo code for the C+I approach is given in Algorithm 6. To be more specific, each region (or agent) solves the first order optimality conditions of the underlying DC-OPF that is associated with that region. Given that we are using cold start in this paper, in the first step, each region initializes its local variables, i.e., variables associated with its own region, as well as the shared variables between that region and its neighboring regions equal to zero. Then, each region (or agent) updates its associate variables λ, p_G, θ, μ using distributed updates (53)-(56). In the meantime, neighboring regions exchange the updated values of the shared variables after every few iterations. This iterative scheme continues until certain convergence criteria are met, e.g., mismatches of power balance equations are less than a pre-specified threshold. It should be mentioned that the obtained distributed solution satisfies the first order optimality conditions of the optimization problem. We refer to [89] and [92] for more details on problem formulation and solution procedure.

IX. Simulation Studies

A. Test System and Convergence Measurements

The test system is the IEEE 48-bus test system [93]. The system consists of two regions which are connected via three tie-lines. We indicate the elements of the first region by prefix 1, and prefix 2 for the second region. For example, G107 means generating unit 7 of region one, and bus 207 means bus 7 of region 2. The three tie-lines connect buses 107, 113, and 123 respectively to 203, 215, and 217.

We evaluate the performance of the algorithms with three indices (which are widely used in [59] and [88]): 1) Euclidean norm of mismatch between the power flow in tie-line bb' determined by algorithm A, i.e., $P_{bb'}^A$, and the optimal value $P_{bb'}^{opt}$ obtained by solving the centralized OPF, as given by

$$P_{mis} = \left\| \frac{P_{bb'}^{opt} - P_{bb'}^A}{P_{bb'}^{opt}} \right\| \tag{57}$$

2) the relative distance of the total cost function f^A determined by algorithm A from the optimal value f^{opt} determined by solving the centralized problem as defined by

$$rel = \frac{\left| f^{opt} - f^A \right|}{f^{opt}} \tag{58}$$

and 3) the residual of the total load balance as given by

$$res = \sum_{i \in \Omega_B} \sqrt{\mathcal{RB}_i^2} \tag{59}$$

where $\mathcal{RB}$ is the residual of the power balance at bus i and Ω_B is the set of all buses.

In order to allow for a fair comparison, we use the same initial values for the primal variables in the simulations for all of the six algorithms. The starting points of the primal variables are set to zero (i.e., a cold start). Note that one can play with the starting values to obtain different convergence performances. All computations are carried out using Matlab on a 3.1GHz personal computer with 32GB of RAM. The computation times heavily depend on how the implementation code is written.[1] In addition, the algorithms' performances highly depend on the choice of the initial values of the variables and starting point. One can play with the initial values to obtain different results. Hence, we do not compare total computation times (i.e., the total computation time is the total time that an algorithm takes to converge to the optimal result over the course of iterations) but focus on the effort that it takes to carry out one iteration (for the case of parallel computation) and the solution quality associated with each algorithm.

B. Simulation Results

1) ATC: We analyze the performance of the ATC-based distributed OPF on the IEEE 48-bus test system. Three tie-lines interconnect areas one and two. Therefore, there are six shared variables, namely the voltage angles of the buses at the end of the tie-lines, between regions one and two. The central coordinator solves an optimization problem including six variables. The initial values for $\hat{\boldsymbol{\theta}}_b$ and $\hat{\boldsymbol{\theta}}_{b'}$ ($b \in 107, 113, 123$, $b' \in 203, 215, 217$,) are set to zero, $\boldsymbol{\lambda} = \boldsymbol{\beta} = 1$, $\boldsymbol{\alpha} = 1$, and the convergence thresholds ϵ is 0.0001. The ATC-based distributed OPF converges after 96 iterations. We use the quadratic programming (QP) solver of Matlab, and the average simulation time of each iteration is 0.08 seconds. The angle values of the border buses obtained in region 1, region 2, and the central coordinator differ by less than 0.0001 in iteration 96. Figure 6 shows the values of the Euclidean norm of the power flow mismatch, the relative distance, and the residual of the load balance equations over the course of iterations. Note that the values of the three convergence measures are zero at the optimal point.

[1] Note that the algorithms have been implemented by a variety of researchers all of which are authors of this paper.

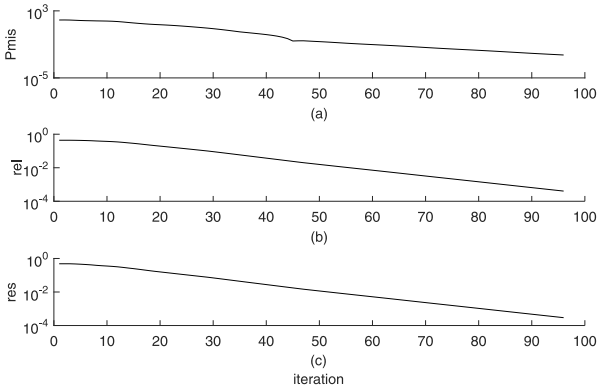

Fig. 6. The convergence property of ATC.

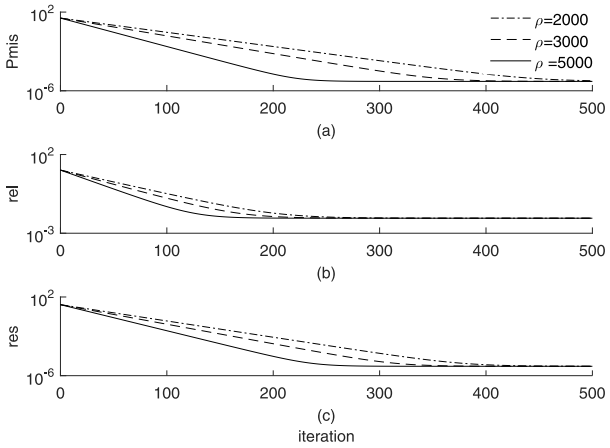

Fig. 7. The convergence property of ADMM.

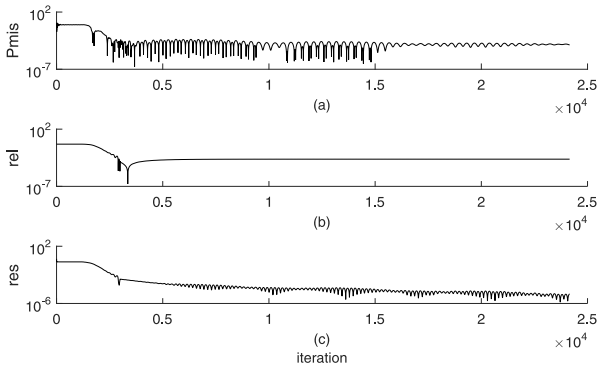

Fig. 8. The convergence property of PMP.

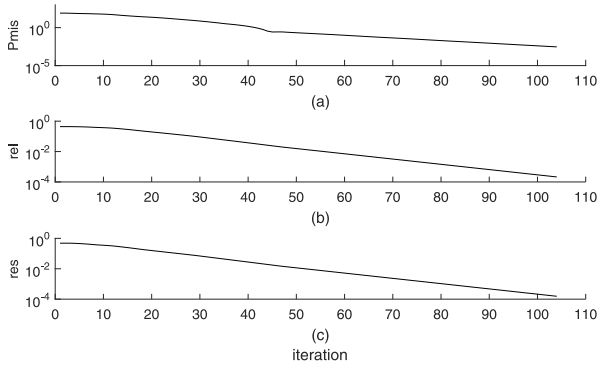

Fig. 9. The convergence property of APP.

2) ADMM: The convergence behavior of the ADMM-based distributed OPF depends on the value of penalty parameter ρ. The simulation results are carried out using three different values of ρ, 2000, 3000, and 5000. When $\rho = 2000$, the algorithm convergence after about 500 iterations. Increasing parameter ρ to 3000 reduces number of iterations to about 400. The ADMM algorithm converges after 250 iterations if ρ is set to 5000. We utilize MATLAB's QP solver, and the average simulation time of each iteration is around 0.08 seconds. The results show that lager ρ values lead to faster convergence. However, selecting a large value for parameter ρ potentially reduces the accuracy of the results determined by the ADMM algorithm. The three convergence measures (57)-(59) are shown in Fig. 7. These three indices become closer to zero as more iterations are carried out.

3) PMP: Similar to ADMM, the convergence in the PMP algorithm is highly dependent on the value of the parameter ρ. A high ρ value ensures faster convergence in terms of primal residual, but it slows down the convergence in terms of dual residual, whereas a lower ρ value does the opposite. The prox term in the objective function ensures that the iterates vary smoothly from the previous values.

We modified the parameter ρ using two different techniques. For the first 3000 iterations we change it such that ρ times the primal residual equals the dual residual, while for the rest of the iterations, we fix the value of ρ at the latest value. The tolerances have been taken as $\epsilon^{\text{pri}} = 0.006$ and $\epsilon^{\text{dual}} = 0.006$. The number of iterations to converge is 24157. The average computation time of each iteration is around 0.02 seconds. We have plotted the metrics as described by equations (57)-(59), i.e., convergence measures, in Fig. 8. The three indices become closer to zero over the course of iterations.

4) APP: The initial values $\boldsymbol{\theta}_{b,m}$, $\boldsymbol{\theta}_{b',m}$, $\boldsymbol{\theta}_{b,n}$, and $\boldsymbol{\theta}_{b',n}$ are set to zero. Parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ are set to one, and $\boldsymbol{\rho}$ is set to 2, and the convergence thresholds ϵ is 0.0001. The APP-based decentralized OPF converges after 104 iterations. We utilize MATLAB's QP solver, and the average computation time per iteration is around 0.05 seconds. The difference between the optimal value of the angle of border buses determined by the two regions is less than the pre-specified convergence threshold after 104 iterations. Figure 9 illustrates the convergence measures P_{mis}, *rel*, and *res*. These values get closer to zero, i.e., optimal point, as more iterations are performed.

5) OCD: The convergence criterion is that the maximum mismatch in all KKT conditions of all subsystems should be smaller than 0.0001. Both OCD and OCD-C converge to the exact centralized solution for which the objective function value is \$113,214; i.e., there is no gap between the centralized and decentralized solutions. Figure 10 shows the evolution of the convergence measures with both methods. It can be seen that OCD-C converges faster with almost no oscillation and can achieve a 10^{-5} accuracy in around 34 iterations, which significantly improves the convergence rate of OCD in this case. The average simulation time of OCD per iteration is slightly less than 0.01 seconds, and it is slightly more than

TABLE I
SUMMARY OF THE RESULTS

Algorithm	ATC	ADMM ($\rho = 2000$, $\rho = 3000$, $\rho = 5000$)	PMP	APP	OCD (OCD-C)	C+I
Iteration	96	(500, 400, 250)	24157	104	76 (34)	250
Average computation time per iteration	0.08	0.08	0.02	0.05	<0.01 (>0.01)	0.01

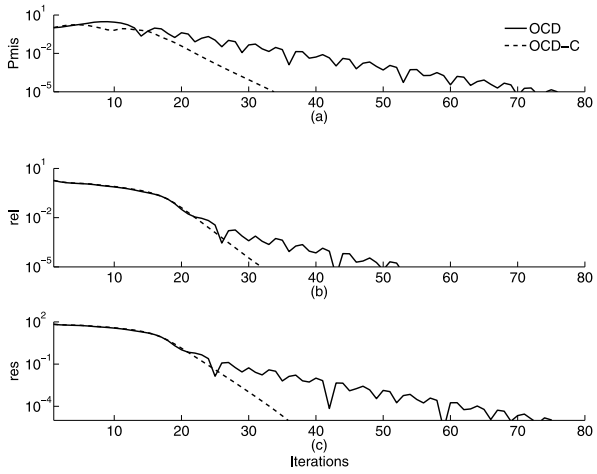

Fig. 10. The convergence property of OCD.

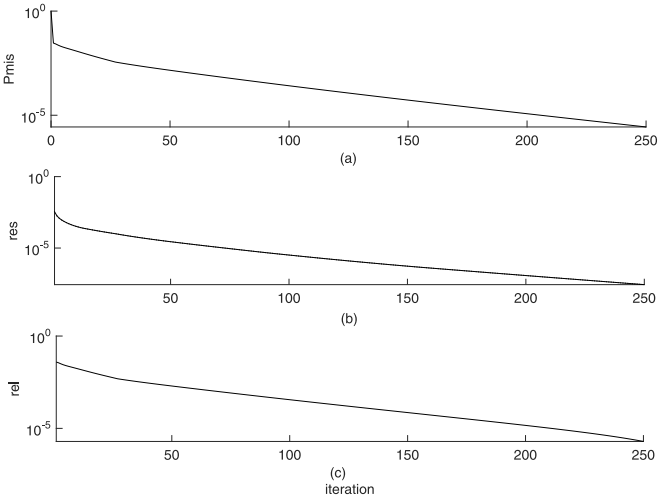

Fig. 11. The convergence property of Consensus+Innovations based method.

0.01 for OCD-C. Note that OCD-C has a correction step and its calculation time increases the average simulation time of each iteration. However, it reduces the total number of iterations and hence the total simulation times over the course of iterations.

6) C+I: We again use the 48-bus test system to evaluate the performance of the C+I approach. We use a cold start for the simulations, i.e., all the variables except the Lagrange multipliers λ_i are set to zero at the start. The initial value of λ is the electricity price in the previous time interval (i.e., $t = 0$) and it is assumed to be 40 \$/MWh. The tuning parameters are adjusted based on the guidelines presented in [89]. We use *res*, i.e., mismatch of power balance equations, as the stoping criterion. Figure 11 depicts convergence measurements P_{mis}, *rel*, and *res* over the course of 250 outer iterations, i.e., iterations between the areas. The average computation time per iteration is around 0.01 seconds. From the simulation results, it is clear that the proposed decentralized algorithm gets closer to the optimal solution as more iterations are carried out (it can also be mathematically derived, see [89] for more details). Note that the value of the residual of the load balance equations and consequently *res* are equal to zero at the optimal point. In addition, P_{mis}, i.e., mismatch between the total generation and load, becomes almost zero after 250 iterations.

Table I gives a summary over the number of iterations of each algorithm. Note that as the algorithms are implemented on different platforms and computing machines, we do not compare the simulation times.

X. SUMMARY AND DISCUSSION

Given the large size of the electric power grid, the responsibility for the control of this system has traditionally been shared by many entities and distributed/decentralized approaches can provide the means to optimally coordinate among these entities. With the increased penetration of distributed resources as well as demand response, it can be expected that the importance of such approaches further increases as they can provide the means to handle the large number of controllable elements in the system and optimally operate and manage the entire system.

In this paper, we distinguished between the terms distributed and decentralized approaches. While we used the term distributed for algorithms which need a centralized coordinator, we employed the term decentralized for approaches that do not need any form of central coordinator. We reviewed distributed/decentralized algorithms for the optimal power flow solution in electric power systems. This paper mainly focused on six popular decomposition coordination algorithms: analytical target cascading (ATC), alternating direction method of multipliers (ADMM), proximal message passing (PMP), auxiliary problem principle (APP) optimality condition decomposition (OCD), and consensus + innovations (C+I). For each algorithm, we explained the basic concept, the general formulation, the application to DC-OPF, and the solution procedure. We applied these six decomposition coordination algorithms on the IEEE-48 test system.

The specific features and characteristics of each of the six algorithms are explained in Sections III–VIII. While the general concepts of ATC, ADMM, PMP, and APP are similar and based on augmented Lagrangian relaxation, OCD and C+I solve the KKT conditions in a decentralized fashion. Figure 12 gives a diagram demonstrating how different decomposition structures are obtained. In addition, Table II summarizes characteristics and features of the six algorithms.

TABLE II
SUMMARY OF THE SIX ALGORITHMS FOR THE DC-OPF APPLICATION

Algorithm	Category	Foundation	Central coordinator	Shared information	Computational effort per iteration	Amount of data exchange per iteration	Stopping criteria
ATC	Distributed	Augmented Lagrangian relaxation	yes	voltage angles of border buses	high	low	mismatch between angles of border buses
ADMM	Distributed	Augmented Lagrangian relaxation	yes	voltage angles of border buses	high	low	mismatch between angles of border buses
PMP	Decentralized	Augmented Lagrangian relaxation	no	average node power mismatch and voltage angles, Lagrange Multipliers of node power mismatch and voltage angle inconsistency	low	high	Primal and dual residuals
APP	Decentralized	Augmented Lagrangian relaxation	no	voltage angles of border buses	high	low	mismatch between angles of border buses
OCD	Decentralized	KKT conditions	no	voltage angles and Lagrangian Multipliers of power balance constraints at border buses	high	medium	mismatch of KKT conditions
C+I	Decentralized	KKT conditions	no	voltage angles and Lagrangian Multipliers of power balance constraints at border buses	low	high	mismatch of power balance

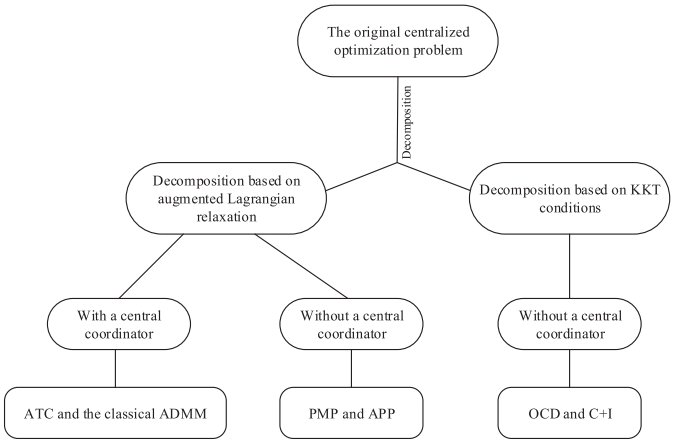

Fig. 12. A diagram demonstrating how different decomposition structures are obtained.

The implementation and performance of ATC and the classic ADMM are similar. They are distributed algorithms which need a central coordinator. We can refer to the OPF problem corresponding to each geographical area as the slave subproblems, and the problem solved by the central coordinator as the master problem. A control entity solves its own slave subproblem (note that the subproblems are solved in parallel) and then sends some specific information, obtained from the optimization, to the master problem. Receiving the required information from all slave subproblems, the master problem coordinates the control entities. In each iteration of ATC and ADMM, only voltage angles of border buses need to be shared between the master coordinator and the slave subproblems. Thus, the amount of data exchange per iteration is low and the communication structure is simple as each subproblem communicates only with the coordinator. Increasing the number of subproblems and shared variables increases the computational efforts of the ATC and the classical ADMM algorithms. Thus, ATC and the classical ADMM are well-suited for area-based OPF rather than nodal OPF. Although the computational effort to solve each subproblem (OPF problem of each area) per iteration is high (depending on size of the subproblem), ATC and ADMM algorithms need a low amount of data exchange between the subproblems per iteration.

One can say that the classic ADMM has only two management levels, upper-level coordination and lower-lever subproblems. In contrast, although the ATC-based OPF presented in this paper has two levels similar to the classical ADMM, the general form of ATC is not limited to only two levels and can have multiple levels as shown in Fig. 2. In general, ATC is a model-based hierarchical optimization method for multilevel system design. Each level of ATC may include several subproblems that can be solved in parallel. Thus, ATC is well-suited for hierarchical design optimization problems, and it can be used for multilevel management of power systems with multiple voltage or management levels. However, because of the hierarchical decision-making in ATC, implementation of ATC is potentially a bit more complex than implementation of ADMM. In addition, ADMM uses a second-order penalty term to model inconsistency between the coupling variables, however, an advantage of ATC is that it is not limited to a second-order penalty function for relaxation. In ATC, a user has the option to use other forms of penalty functions (e.g., linear, quadratic, and exponential functions) to relax consistency constraints [52], [94]. This significantly enhances the flexibility of ATC. For instance, an exponential penalty function provides the user with the option to use second-order solution methods (or even higher order methods) to solve the optimization problem as the derivative of an exponential function is again an exponential function. Consider the second-order penalty function as used in this paper for ATC and ADMM (see (14) and (18)). When in ADMM the Lagrange multiplier associated with the linear term of the penalty function is updated in each iteration, the tuning parameter of the quadratic term is constant. However, in ATC, both Lagrange multipliers associated with the linear and quadratic terms need

to be updated iteratively with respect to the difference between the shared variables. Despite the similarities and differences between ATC and the classic ADMM, the existence of a central coordinator can potentially increase the vulnerability of these two approaches, for example, against a cyber-attack or data manipulation in the coordination level (i.e., central coordinator).

The foundation of APP relies upon the augmented Lagrangian relaxation and linearization of the augmented Lagrangian according to the auxiliary problem principle. The implementation and convergence performance of APP is similar to ATC and ADMM. The shared variables in these three algorithms are the voltages at the border buses. The computation effort and the amount of data exchanged between the subproblems in each iteration of APP is similar to ATC and ADMM. However, unlike ATC and the classic ADMM, in APP, the cross-terms of the augmented Lagrangian are linearized and the solution of the OPF problem is found by solving a sequence of auxiliary problems. In APP, all subproblems can be solved in parallel which reduces the overall computation time. In APP, similar to ADMM and unlike ATC, the tuning parameter of the quadratic term of the penalty function is constant and the Lagrange multipliers associated with the linear terms need to be updated in each iteration. Furthermore, ATC and ADMM, at iteration k, use values computed by the coordinator at iteration k. However, in APP, each subproblem in iteration k uses the shared variables calculated by other subproblems in iteration $k-1$. This further means that the APP does not require a coordinator to solve the OPF problem in a distributed fashion. Thus, APP is a decentralized algorithm that has only one management level, which consists of OPF subproblems of the regions. This potentially reduces the sensitivity of APP to a cyber-attack or data manipulation compared with ATC and the classical ADMM. However, as peer-to-peer communication between the subproblems is needed, the communication structure is not as simple as in the case for ATC and ADMM (in which a two-level communication structure is needed). Furthermore, an increasing number of shared variables increases the computational effort of APP. Thus, APP is more efficient to solve an area-based OPF problem rather than a nodal OPF problem (similar to ATC and the classical ADMM). Although in APP the computational effort of an OPF subproblem (OPF problem of each area) per iteration is high (depending on size of the subproblem), it needs a low amount of data exchange between the subproblems per iteration.

PMP is a particular application of ADMM where minimizing at each Gauss-Seidel pass amounts to evaluating the prox functions with respect to the particular decision vector partition. Although the classic ADMM is not really a parallel algorithm, because the multiplier updates are performed in a sequential manner, PMP is a fully decentralized algorithm. PMP is able to distribute computation efforts across each device when the decision vector is fully decomposable across the devices, which is ideally the case for the DC OPF or the linearized version of the OPF. In this paper, PMP has been applied to solve a nodal OPF. The merits of PMP are that it is fully decentralized and highly fine-grained (breaks the problem down to the individual device level). It is particularly suited for large-scale multi-time step dispatch problems, and in the case that highly distributed or parallel computing resources are available. The demerit is that, as it solves a nodal OPF, it takes a large number of iterations to converge. However, the solution process of the nodal subproblems is much faster than solving the area-based subproblems (because of the low computational effort in the nodal subproblems compared with the high computational effort in the area-based subproblems). Also, the convergence time can potentially be reduced by using custom solvers, like CVXGEN in conjunction with a comparatively lower level language as C++, as we have done here. Another advantage of PMP, which is also the case for ADMM, is that its convergence is highly dependent on the value of the tuning parameter ρ. A high tuning parameter ensures faster convergence in terms of primal residual, but it slows down the convergence in terms of dual residual, whereas a low tuning parameter does the opposite.

The advantage of OCD is that it can be used to decompose general optimization problems, which do not necessarily need to be convex, and its convergence holds as long as condition (41) is fulfilled. In addition, at each iteration, only one Newton-Raphson step is carried out instead of solving the entire subproblem, which leads to light computation cost. However, condition (41) is usually hard to check before solving the optimization problem. While empirical results show that condition (41) generally holds in test systems with mild partitions, it may not hold in large-scale power systems which are more complex and meshed, especially for solving non-convex problems. In the situation where condition (41) does not hold, additional methods such as GMRES or preconditioning need to be deployed, which involves more computation and centralized coordination.

The C+I algorithm needs a limited information exchange between the subproblems in each iteration. Note that the C+I needs all subproblems to reach consensus on a common variables, hence, it takes more iterations compared with Lagrangian-based methods. However, each iteration's solution process is faster than Lagrangian-based methods as C+I method technically solves a set of linear equations.

The OCD and C+I algorithms both directly solve the KKT conditions in a distributed fashion and the computation effort at each iteration is fairly low. However, these two methods handle the KKT conditions differently where OCD conducts one Newton-Raphson step to compute the updates of variables instead of finding the exact solutions to the KKT conditions, while C+I consecutively updates primal and dual variables using explicit updates. Another main difference between C+I and OCD is in the level of distributedness that they are offering. The C+I can be implemented at the nodal level as well as the regional level. This method is flexible in terms of selecting size and number of regions. However, the OCD is only suitable to solve an optimization problem at the regional level due to the violation of its convergence criterion when the number of subproblems is large. Furthermore, OCD can be generalized to solve the AC OPF problem which is non-convex, while the applicability of C+I to non-convex problems needs further studies.

Finally, comparing with the introduced Lagrangian methods, i.e., ADMM, ATC, APP, and PMP, the C+I directly solves the KKT conditions, and does not solve optimization problems during the course of iterations. Also, compared with the primal-dual update procedure where each variable update solely uses its associated gradients terms, the C+I update rules are based on utilizing a linear combination of gradient terms. Note that using a linear combination of the gradient terms could potentially improve the convergence speed. Furthermore, the C+I solution approach does not require updating all Lagrangian multipliers as it uses projection operators to ensure feasibility with respect to thresholding constraint, e.g., upper and lower limits for generators.

XI. Developments and Future Directions

In general, the distributed/decentralized algorithms might be used to speed up the solution process of large-scale optimization problems which are hard to solve in a centralized fashion, and to manage systems with multiple independent control entities which are not willing to share their information with other parties. Future research directions on the distributed/decentralized algorithms can be categorized into two main directions: accurate modeling and implementation of power system problems by distributed/decentralized algorithms, and improvement of convergence properties of existing or new algorithms.

The distributed/decentralized OPF algorithms presented in this paper are based on the basic DC power flow equations. However, the AC power flow equations more accurately model the physical properties of power systems. Many of the distributed/decentralized algorithms have been or can be extended to the nonlinear AC optimal power flow problem. OPF is a single period (e.g., one hour) scheduling problem. However, a day-ahead unit commitment, which includes integer variables, is usually solved in electricity markets to schedule the system. Thus, the distributed/decentralized algorithms can be applied to solve the unit commitment problem to determine on/off status and power dispatch of generating units. Furthermore, in order to ensure system security and reliability, possible contingencies and disturbances, e.g., line and generation outages need to be considered in system operation. Thus, security-constrained optimal power flow (SCOPF) and security-constrained unit commitment (SCUC) problems need to be studied. Both SCOPF and SCUC problems are computationally expensive and hence the distributed/decentralized algorithms can potentially reduce the SCOPF and SCUCs computation efforts.

Another possible direction of research is to improve the convergence rate of the presented algorithms. This might be achieved through enhancement of the algorithms, i.e., the terms in the penalty functions and the procedure of updating parameters and multipliers over iterations. The convergence might also be improved by adding extra communication links between non-neighboring areas. Furthermore, most of the distributed/decentralized algorithms are sensitive to the choice of initial values of the shared variables and multipliers, and it is highly desirable to reduce this sensitivity. Another open issue is the system partitioning. It is worth studying how one can identify the best partition of the system into subregions such that the performance of distributed methods can be optimized in terms of minimum number of iterations, shortest convergence time, and minimum communication costs. In terms of the implementation of distributed/decentralized methods, there are practical issues to be considered. One important issue to be studied is the inclusion of communications in the evaluation of the distributed methods. The communication infrastructures for implementing distributed methods need to be carefully designed, and the impact of communication delays and failures on the performance of distributed methods need to be investigated.

Acknowledgment

The authors would like to thank Dr. Mahdi Kefayati and Patryk Radyjowski (The University of Texas at Austin), Dr. Matt Kraning and Dr. Eric Chu, (both Ph.D. graduate from Stanford University), and Prof. Stephen Boyd (Stanford University) for help with software and mathematical modeling related issues.

References

[1] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation, and Control*. New York, NY, USA: Wiley, 2012.

[2] H. W. Dommel and W. F. Tinney, "Optimal power flow solutions," *IEEE Trans. Power App. Syst.*, vol. PAS-87, no. 10, pp. 1866–1876, Oct. 1968.

[3] D. I. Sun, B. Ashley, B. Brewer, A. Hughes, and W. F. Tinney, "Optimal power flow by Newton approach," *IEEE Trans. Power App. Syst.*, vol. PAS-103, no. 10, pp. 2864–2880, Oct. 1984.

[4] D. Gan, R. J. Thomas, and R. D. Zimmerman, "Stability-constrained optimal power flow," *IEEE Trans. Power Syst.*, vol. 15, no. 2, pp. 535–540, May 2000.

[5] M. Huneault and F. D. Galiana, "A survey of the optimal power flow literature," *IEEE Trans. Power Syst.*, vol. 6, no. 2, pp. 762–770, May 1991.

[6] J. A. Momoh, R. Adapa, and M. E. El-Hawary, "A review of selected optimal power flow literature to 1993. Part I. Nonlinear and quadratic programming approaches," *IEEE Trans. Power Syst.*, vol. 14, no. 1, pp. 96–104, Feb. 1999.

[7] J. A. Momoh, R. Adapa, and M. E. El-Hawary, "A review of selected optimal power flow literature to 1993. Part II: Newton, linear programming and interior point methods," *IEEE Trans. Power Syst.*, vol. 14, no. 1, pp. 105–1114, Feb. 1999.

[8] H. Wang, C. E. Murillo-Sanchez, R. D. Zimmerman, and R. J. Thomas, "On computational issues of market-based optimal power flow," *IEEE Trans. Power Syst.*, vol. 22, no. 3, pp. 1185–1193, Aug. 2007.

[9] K. S. Pandya and S. K. Joshi, "A survey of optimal power flow methods," *J. Theor. Appl. Inf. Technol.*, vol. 4, no. 5, pp. 450–458, 2008.

[10] J. Lavaei and S. H. Low, "Zero duality gap in optimal power flow problem," *IEEE Trans. Power Syst.*, vol. 27, no. 1, pp. 92–107, Feb. 2012.

[11] B. H. 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.

[12] D. Hur, J.-K. Park, and B. H. Kim, "Evaluation of convergence rate in the auxiliary problem principle for distributed optimal power flow," *IEE Proc. Gener. Transm. Distrib.*, vol. 149, no. 5, pp. 525–532, Sep. 2002.

[13] A. J. Conejo, F. J. Nogales, and F. J. Prieto, "A decomposition procedure based on approximate Newton directions," *Math. Program.*, vol. 93, no. 3, pp. 495–515, 2002.

[14] A. G. Bakirtzis and P. N. 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.

[15] P. N. Biskas, A. G. Bakirtzis, N. I. Macheras, and N. K. 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.

[16] J. Liang, D. D. Molina, G. K. Venayagamoorthy, and R. G. Harley, "Two-level dynamic stochastic optimal power flow control for power systems with intermittent renewable generation," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2670–2678, Aug. 2013.

[17] C.-H. Lin and S.-Y. 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.

[18] 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.

[19] A. Y. S. Lam, B. Zhang, and D. N. Tse, "Distributed algorithms for optimal power flow problem," in *Proc. 51st IEEE Conf. Decis. Control*, Maui, HI, USA, Dec. 2012, pp. 430–437.

[20] M. D. Ilic, L. Xie, and J.-Y. Joo, "Efficient coordination of wind power and price-responsive demand—Part I: Theoretical foundations," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 1875–1884, Nov. 2011.

[21] M. D. Ilic, L. Xie, and J.-Y. Joo, "Efficient coordination of wind power and price-responsive demand—Part II: Case studies," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 1885–1893, Nov. 2011.

[22] Z. Li, W. Wu, B. Zhang, H. Sun, and Q. Guo, "Dynamic economic dispatch using Lagrangian relaxation with multiplier updates based on a quasi-Newton method," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4516–4527, Nov. 2013.

[23] S.-S. Lin, S.-C. Horng, and C.-H. Lin, "Distributed quadratic programming problems of power systems with continuous and discrete variables," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 472–481, Feb. 2013.

[24] N. Li, "Distributed optimization in power networks and general multi-agent systems," Ph.D. dissertation, Eng. Appl. Sci., California Inst. Technol., Pasadena, CA, USA, 2013.

[25] D. Zhu and G. Hug, "Decomposed stochastic model predictive control for optimal dispatch of storage and generation," *IEEE Trans. Smart Grid*, vol. 5, no. 4, pp. 2044–2053, Jul. 2014.

[26] A. K. 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.

[27] T. Erseghe, "Distributed optimal power flow using ADMM," *IEEE Trans. Power Syst.*, vol. 29, no. 5, pp. 2370–2380, Sep. 2014.

[28] Y. Xu, W. Zhang, G. Hug, S. Kar, and Z. Li, "Cooperative control of distributed energy storage systems in a microgrid," *IEEE Trans. Smart Grid*, vol. 6, no. 1, pp. 238–248, Jan. 2015.

[29] B. Zhang, A. Y. S. Lam, A. D. Domínguez-García, 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, Jul. 2015.

[30] B. A. Robbins and A. D. Domínguez-García, "Optimal reactive power dispatch for voltage regulation in unbalanced distribution systems," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2903–2913, Jul. 2016.

[31] M. H. Amini, B. Nabi, and M.-R. Haghifam, "Load management using multi-agent systems in smart distribution network," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Vancouver, BC, Canada, Jul. 2013, pp. 1–5.

[32] M. H. Amini, R. Jaddivada, S. Mishra, and O. Karabasoglu, "Distributed security constrained economic dispatch," in *Proc. IEEE Innov. Smart Grid Technol. Asia (ISGT ASIA)*, Bangkok, Thailand, Nov. 2015, pp. 1–6.

[33] G. Binetti, A. Davoudi, D. Naso, B. Turchiano, and F. L. Lewis, "Distributed consensus-based economic dispatch with transmission losses," *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1711–1720, Jul. 2014.

[34] G. Binetti, A. Davoudi, D. Naso, B. Turchiano, and F. L. Lewis, "A distributed auction-based algorithm for the nonconvex economic dispatch problem," *IEEE Trans. Ind. Informat.*, vol. 10, no. 2, pp. 1124–1132, May 2014.

[35] H. K. Nguyen, H. Mohsenian-Rad, A. Khodaei, and Z. Han, "Decentralized reactive power compensation using nash bargaining solution," *IEEE Trans. Smart Grid*, to be published.

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

[37] D. Phan and J. Kalagnanam, "Some efficient optimization methods for solving the security-constrained optimal power flow problem," *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 863–872, Mar. 2014.

[38] 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.*, to be published.

[39] Z. Li, Q. Guo, H. Sun, and J. Wang, "A new LMP-sensitivity-based heterogeneous decomposition for transmission and distribution coordinated economic dispatch," *IEEE Trans. Smart Grid*, to be published.

[40] F. Dörfler, J. W. Simpson-Porco, and F. Bullo, "Breaking the hierarchy: Distributed control and economic optimality in microgrids," *IEEE Trans. Control Netw. Syst.*, vol. 3, no. 3, pp. 241–253, Sep. 2016.

[41] X. Wu, F. Dörfler, and M. R. Jovanović, "Input-output analysis and decentralized optimal control of inter-area oscillations in power systems," *IEEE Trans. Power Syst.*, vol. 31, no. 3, pp. 2434–2444, May 2016.

[42] A. Cherukuri and J. Cortés, "Distributed generator coordination for initialization and anytime optimization in economic dispatch," *IEEE Trans. Control Netw. Syst.*, vol. 2, no. 3, pp. 226–237, Sep. 2015.

[43] D. Richert and J. Cortés, "Robust distributed linear programming," *IEEE Trans. Autom. Control*, vol. 60, no. 10, pp. 2567–2582, Oct. 2015.

[44] H. Zhu and G. B. Giannakis, "Power system nonlinear state estimation using distributed semidefinite programming," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 6, pp. 1039–1050, Dec. 2014.

[45] Y. Zhang, N. Gatsis, and G. B. Giannakis, "Disaggregated bundle methods for distributed market clearing in power networks," in *Proc. IEEE Glob. Conf. Signal Inf. Process. (GlobalSIP)*, Austin, TX, USA, 2013, pp. 835–838.

[46] A. J. Conejo, E. Castillo, R. Minguez, and R. García-Bertrand, *Decomposition Techniques in Mathematical Programming*. Heidelberg, Germany: Springer, 2006.

[47] N. Michelena, H. M. Kim, and P. Y. Papalambros, "A system partitioning and optimization approach to target cascading," in *Proc. 12th Int. Conf. Eng. Design*, Munich, Germany, Aug. 1999, pp. 1109–1112.

[48] H. M. Kim, N. F. Michelena, P. Y. Papalambros, and T. Jiang, "Target cascading in optimal system design," *J. Mech. Design*, vol. 125, no. 3, pp. 474–480, 2003.

[49] S. Tosserams, L. F. P. Etman, and J. E. Rooda, "An augmented Lagrangian decomposition method for quasi-separable problems in MDO," *Struct. Multidiscip. Optim.*, vol. 34, no. 3, pp. 211–227, 2007.

[50] S. Tosserams, L. F. P. Etman, P. Y. Papalambros, and J. E. Rooda, "An augmented Lagrangian relaxation for analytical target cascading using the alternating direction method of multipliers," *Struct. Multidiscip. Optim.*, vol. 31, no. 3, pp. 176–189, 2006.

[51] 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, Jul. 2015.

[52] S. DorMohammadi and M. Rais-Rohani, "Exponential penalty function formulation for multilevel optimization using the analytical target cascading framework," *Struct. Multidiscip. Optim.*, vol. 47, no. 4, pp. 599–612, 2013.

[53] N. Michelena, H. Park, and P. Y. Papalambros, "Convergence properties of analytical target cascading," *AIAA J.*, vol. 41, no. 5, pp. 897–905, May 2003.

[54] 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, Sep. 2014.

[55] A. Kargarian, Y. Fu, and H. Wu, "Chance-constrained system of systems based operation of power systems," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3404–3413, Sep. 2016.

[56] A. Kargarian, Y. Fu, P. Liu, and C. Wang, "A system of systems engineering approach for unit commitment in multi-area power markets," in *Proc. IEEE PES Gen. Meeting*, National Harbor, MD, USA, Jul. 2014, pp. 1–5.

[57] A. R. Malekpour, A. Pahwa, A. Malekpour, and B. Natarajan, "Hierarchical architecture for integration of rooftop pv in smart distribution systems," *IEEE Trans. Smart Grid*, to be published.

[58] 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 Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2010.

[59] 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. 70–122, 2013.

[60] V. Kekatos and G. B. Giannakis, "Distributed robust power system state estimation," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1617–1626, May 2013.

[61] S. Chakrabarti, M. Kraning, E. Chu, R. Baldick, and S. Boyd, "Security constrained optimal power flow via proximal message passing," in *Proc. IEEE PES Clemson Power Syst. Conf.*, Clemson, SC, USA, 2014, pp. 1–8.

[62] Z. Tan, P. Yang, and A. Nehorai, "An optimal and distributed demand response strategy with electric vehicles in the smart grid," *IEEE Trans. Smart Grid*, vol. 5, no. 2, pp. 861–869, Mar. 2014.
[63] P. Šulc, S. Backhaus, and M. Chertkov, "Optimal distributed control of reactive power via the alternating direction method of multipliers," *IEEE Trans. Energy Convers.*, vol. 29, no. 4, pp. 968–977, Dec. 2014.
[64] J. F. C. Mota, J. M. F. Xavier, P. M. Q. Aguiar, and M. Puschel, "Distributed basis pursuit," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1942–1956, Apr. 2012.
[65] W. Shi, Q. Ling, K. Yuan, G. Wu, and W. Yin, "On the linear convergence of the ADMM in decentralized consensus optimization," *IEEE Trans. Signal Process.*, vol. 62, no. 7, pp. 1750–1761, Apr. 2014.
[66] I. D. Schizas, A. Ribeiro, and G. B. Giannakis, "Consensus in ad hoc WSNs with noisy links—Part I: Distributed estimation of deterministic signals," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 350–364, Jan. 2008.
[67] M. J. Feizollahi, M. Costley, S. Ahmed, and S. Grijalva, "Large-scale decentralized unit commitment," *Int. J. Elect. Power Energy Syst.*, vol. 73, pp. 97–106, Dec. 2015.
[68] Y. Zhang and G. B. Giannakis, "Efficient decentralized economic dispatch for microgrids with wind power integration," in *Proc. 6th Annu. IEEE Green Technol. Conf. (GreenTech)*, Corpus Christi, TX, USA, Apr. 2014, pp. 7–12.
[69] Q. Peng and S. H. Low, "Distributed algorithm for optimal power flow on a radial network," in *Proc. IEEE 53rd Annu. Conf. Decis. Control (CDC)*, Los Angeles, CA, USA, Dec. 2014, pp. 167–172.
[70] A. X. Sun, D. T. Phan, and S. Ghosh, "Fully decentralized AC optimal power flow algorithms," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PES)*, Vancouver, BC, Canada, Jul. 2013, pp. 1–5.
[71] 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, Sep. 2013.
[72] 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 *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Florence, Italy, May 2014, pp. 7764–7768.
[73] 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 Trans. Control Netw. Syst.*, vol. 2, no. 3, pp. 238–253, Sep. 2015.
[74] G. Cohen, "Auxiliary problem principle and decomposition of optimization problems," *J. Optim. Theory Appl.*, vol. 32, no. 3, pp. 277–305, Nov. 1980.
[75] R. Baldick, B. H. 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, Aug. 1999.
[76] A. Ahmadi-Khatir, A. J. 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, Jul. 2014.
[77] K. H. Chung, B. H. Kim, and D. Hur, "Distributed implementation of generation scheduling algorithm on interconnected power systems," *Energy Convers. Manag.*, vol. 52, no. 12, pp. 3457–3464, 2011.
[78] F. J. Nogales, F. J. Prieto, and A. J. Conejo, "A decomposition methodology applied to the multi-area optimal power flow problem," *Ann. Oper. Res.*, vol. 120, no. 1, pp. 99–116, 2003.
[79] 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.
[80] J. Guo, G. Hug, and O. Tonguz, "Impact of partitioning on the performance of decomposition methods for AC optimal power flow," in *Proc. IEEE Innov. Smart Grid Technol. Conf. (ISGT)*, Washington, DC, USA, 2015, pp. 1–5.
[81] 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*, vol. 7, no. 2, pp. 992–1001, Mar. 2016.
[82] J. Guo, G. Hug, and O. K. Tonguz, "Intelligent partitioning in distributed optimization of electric power systems," *IEEE Trans. Smart Grid*, vol. 7, no. 3, pp. 1249–1258, May 2016.
[83] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007.
[84] A. G. Dimakis, S. Kar, J. M. F. Moura, M. G. Rabbat, and A. Scaglione, "Gossip algorithms for distributed signal processing," *Proc. IEEE*, vol. 98, no. 11, pp. 1847–1864, Nov. 2010.
[85] Z. Zhang and M.-Y. Chow, "Convergence analysis of the incremental cost consensus algorithm under different communication network topologies in a smart grid," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 1761–1768, Nov. 2012.
[86] S. Kar and G. Hug, "Distributed robust economic dispatch in power systems: A consensus + innovations approach," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, San Diego, CA, USA, Jul. 2012, pp. 1–8.
[87] S. Yang, S. Tan, and J.-X. 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.
[88] S. Kar, G. Hug, J. Mohammadi, and J. M. F. Moura, "Distributed state estimation and energy management in smart grids: A consensus+innovations approach," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 6, pp. 1022–1038, Dec. 2014.
[89] J. Mohammadi, S. Kar, and G. Hug, "Distributed approach for dc optimal power flow calculations," *arXiv:1410.4236*.
[90] J. Mohammadi, G. Hug, and S. Kar, "Role of communication on the convergence rate of fully distributed DC optimal power flow," in *Proc. IEEE Int. Conf. Smart Grid Commun.*, Venice, Italy, 2014, pp. 43–48.
[91] J. Mohammadi, G. Hug, and S. Kar, "Asynchronous distributed approach for DC optimal power flow," in *Proc. IEEE Eindhoven PowerTech*, Eindhoven, The Netherlands, 2015, pp. 1–6.
[92] J. Mohammadi, G. Hug, and S. Kar, "Agent-based distributed security constrained optimal power flow," *IEEE Trans. Smart Grid*, 2016, to be published.
[93] C. Grigg *et al.*, "The IEEE reliability test system-1996. A report prepared by the reliability test system task force of the application of probability methods subcommittee," *IEEE Trans. Power Syst.*, vol. 14, no. 3, pp. 1010–1020, Aug. 1999.
[94] S. DorMohammadi and M. Rais-Rohani, "Comparison of alternative strategies for multilevel optimization of hierarchical systems," *Appl. Math.*, vol. 3, no. 10A, pp. 1448–1462, 2012.

Amin Kargarian (S'10–M'14) received the B.S. degree in electrical and computer engineering from the University of Isfahan, Iran, in 2007, the M.S. degree in electrical and computer engineering from Shiraz University, Iran, in 2010, and the Ph.D. degree in electrical and computer engineering from Mississippi State University, USA, in 2014. He joined the ECE Department, Carnegie Mellon University as a Post-Doctoral Research Associate in 2014. He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, Louisiana State University. His research interests include power system optimization and economics, and renewable energy integration.

Javad Mohammadi (S'10) received the B.S. degree from Shahid Bahonar University of Kerman, Iran, in 2008, and the M.S. degree from the University of Tehran, Iran, in 2011, both in electrical engineering. He is currently pursuing the Ph.D. degree in electrical engineering with Carnegie Mellon University, Pittsburgh, PA, USA. His research interests include power system optimization and control.

Junyao Guo received the B.S. degree in electronic engineering from Tsinghua University in 2013. She is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, Carnegie Mellon University, under the supervision of Prof. G. Hug and Prof. O. Tonguz. Her current research focuses on the distributed optimization and communications in the smart grid.

Sambuddha Chakrabarti (S'14) received the B.Tech. degree in electrical engineering from the National Institute of Technology (NIT), Allahabad, India. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Texas at Austin, under the supervision of Prof. R. Baldick. His present research focuses on application of distributed optimization to optimal power flow problem.

Masoud Barati received the Ph.D. degree in electrical engineering from the Illinois Institute of Technology, Chicago, in 2013. He is currently an Instructional Assistant Professor with the Department of Electrical and Computer Engineering, University of Houston. His research interests include microgrid operation and planning, microeconomics, mathematical modeling, and multiple infrastructure assessment.

Gabriela Hug (S'05–M'08–SM'14) was born in Baden, Switzerland. She received the M.Sc. degree in electrical engineering and the Ph.D. degree from the Swiss Federal Institute of Technology (ETH), Zürich, Switzerland, in 2004 and 2008, respectively. She worked in the Special Studies Group of Hydro One in Toronto, Canada, and from 2009 to 2015, she was an Assistant Professor with Carnegie Mellon University, Pittsburgh, USA. She is currently an Associate Professor with the Power Systems Laboratory, ETH Zurich. Her research is dedicated to control and optimization of electric power systems.

Soummya Kar received the B.Tech. degree in electronics and electrical communication engineering from the Indian Institute of Technology, Kharagpur, India, in 2005, and the Ph.D. degree in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, USA, in 2010. From 2010 to 2011, he was with the Electrical Engineering Department, Princeton University, Princeton, NJ, USA, as a Post-Doctoral Research Associate. He is currently an Associate Professor of Electrical and Computer Engineering with Carnegie Mellon University. His research interests include decision-making in large-scale networked systems, stochastic systems, multiagent systems and data science, with applications to cyber-physical systems and smart energy systems. He has published extensively in the above areas with over 140 articles in journals and conference proceedings and holds multiple patents. He was a recipient of the 2016 O. Hugo Schuck Best Paper Award from the American Automatic Control Council, the 2016 Dean's Early Career Fellowship from CIT, Carnegie Mellon, and the 2011 A.G. Milnes Award for Best Ph.D. Thesis in Electrical and Computer Engineering, Carnegie Mellon University.

Ross Baldick (F'07) received the B.Sc. degree in mathematics and physics and the B.E. degree in electrical engineering from the University of Sydney, Sydney, Australia, and the M.S. and Ph.D. degrees in electrical engineering and computer sciences from the University of California, Berkeley, CA, USA, in 1988 and 1990, respectively. From 1991 to 1992, he was a Post-Doctoral Fellow with the Lawrence Berkeley Laboratory, Berkeley. In 1992 and 1993, he was an Assistant Professor with Worcester Polytechnic Institute, Worcester, MA, USA. He is currently a Professor with the Department of Electrical and Computer Engineering, University of Texas, Austin.