Diagonal Quadratic Approximation for Decentralized Collaborative TSO+DSO Optimal Power Flow

Ali Mohammadi, *Student Member, IEEE,* Mahdi Mehrtash, *Student Member, IEEE,* and Amin Kargarian, *Member, IEEE*

***Abstract*- Collaborative operation of electricity transmission and distribution systems improves the economy and reliability of the entire power system. However, this is a challenging problem given that transmission system operators (TSOs) and distribution system operators (DSOs) are autonomous entities that are unwilling to reveal their commercially sensitive information. This paper presents a decentralized decision-making algorithm for collaborative TSO+DSO optimal power flow (OPF) implementation. The proposed algorithm is based on analytical target cascading (ATC) for multilevel hierarchical optimization in complex engineering systems. A local OPF is formulated for each TSO/DSO taking into consideration interactions between the transmission and distribution systems while respecting autonomy and information privacy of TSO and DSOs. The local OPF of TSO is solved in the upper-level of hierarchy, and the local OPFs of DSOs are handled in the lower-level. A diagonal quadratic approximation (DQA) and a truncated diagonal quadratic approximation (TDQA) are presented to develop iterative coordination strategies in which all local OPFs are solved in a parallel manner with no need for a central coordinator. This parallel implementation significantly enhances computations efficiency of the algorithm. The proposed collaborative TSO+DSO OPF is evaluated using a 6-bus and the IEEE 118-bus test systems, and promising results are obtained.**

***Index Terms*- Collaborative transmission and distribution operation, analytical target cascading, diagonal quadratic approximation, decentralized optimization, parallel algorithm.**

NOMENCLATURE

A. Indices, Sets, and Parameters

a, b	Index for border buses in TSO side.
a', b'	Index for border buses in DSOs side.
i, j	Index for subproblem j in level i.
k	Outer loop iteration index.
l	Inner loop iteration index.
f^*	Operating cost determined by the centralize algorithm.
f^{ATC}	Operating cost determined by the decentralized algorithm.
f_{ij}	Operating cost function of subproblem j located in level i.
g_{ij}	Set of inequality constraints of subproblem j located in level i.
h_{ij}	Set of equality constraints of subproblem j located in level i.
N_d	Number of DSO in level two.
P_l^*	Power mismatch in tie-line l determined by the centralize algorithm.
P_l^{ATC}	Power mismatch in tie-line l determined by the decentralized algorithm.
X	Set of variables.
$\pi(.)$	Penalty function.
P_{mis}	Relative power mismatch in a tie-line.
rel	Relative distance of the operating cost.

B. Variables

r_{ij}	Response of subproblem$j$ in level i.
t_{ij}	Targets of subproblem$j$ in level i.
v_a	Voltage magnitude at bus a.
δ_a	Voltage angle at bus a.
$\tilde{v}\angle\tilde{\delta}$	Voltage phasor corresponding to response variables.
$v\angle\delta$	Voltage phasor corresponding to target variables.
α, β	EPF penalty multipliers.
$\delta^{k,l}$	Voltage angle in outer loop k and in inner loop l.
Γ	Step size.
λ, w	AL_BCD's penalty multipliers.
$\lambda_{ij,\delta}^T, w_{ij,\delta}$	DQA's penalty multipliers related to the voltage angle of subproblem$j$ in level i.
$\lambda_{2j,v}^T, w_{ij,v}$	DQA's penalty multipliers related to the voltage magnitude of subproblem j in level i.
τ	Tuning parameter.

I. INTRODUCTION

EMERGING active distribution grids (ADGs), which include distributed energy resources, is reshaping power systems paradigm. Unlike passive distribution grids, an ADG is capable of locally supplying power for end-users with its distributed generators (DGs). Incorporation of ADGs into power system management potentially enhances the overall system performance in terms of economic and security [1, 2]. This has motivated the recent interests in collaborative management of transmission and distribution grids [3-5].

The transmission system is operated by a transmission system operator (TSO), and the distribution system is controlled by a distribution system operator (DSO). Since the transmission and distribution grids are parts of an interconnected system, any decisions made by TSO (DSOs) affects the DSOs' (TSO's) operation and decisions. On the other hand, TSO and DSOs are autonomous control entities with their own rules, policies, and objectives. While one entity aims at minimizing its own costs, the objective of another entity might be reliability maximization with respect to its local operational constraints. Furthermore, TSO and DSOs might compete with each other to achieve their objectives.

This work was supported by the Louisiana Board of Regents under grant LEQSF(2016-19)-RD-A-10.

The authors are with the Department of Electrical and Computer Engineering, Louisiana State University, Baton Rouge, LA 70803 USA (e-mail: amoha39@lsu.edu, mmehrt3@lsu.edu, kargarian@lsu.edu).

Thus, although TSO and DSOs are parts of an interconnected system, they are unwilling to share their commercially sensitive data with each other. This is, preserving the privacy of TSO and DSOs' is critical. Hence, a central scheduling framework, in which TSO and DSOs need to share all their information with a central control authority, may not be appropriate for the entire power system operation [6]. Even if TSO and DSOs share their information with a central control authority and allow this entity to perform the decision-making, solving the resulting integrated large-scale optimization problem is challenging. In addition, failures and cyber-attacks could have a devastating impact on the functionality of a centralized control approach.

Transmission and distribution systems collaboration have seen increased interest recently [7, 8]. Active distribution system can provide various services for the transmission system. These two systems can cooperate to achieve a better grid performance in terms of, for instance, voltage security, operation, planning, etc. our main focus in this paper is one short-term operation. We aim at finding a distributed solution for optimal power flow (OPF) problem for a power system that potentially includes multiple DSOs and a TSO. Here, we briefly review several papers in the field of distributed power system management and TSO+DSO operation.

Iterative approached have been presented in [9-19] to solve OPF in a distributed fashion. Alternating direction method of multipliers (ADMM) [20], heterogeneous decomposition algorithm [21], auxiliary problem principle (APP) [22, 23], optimality condition decomposition [14, 24], consensus+innovations technique [6, 25], proximal message passing [26], and analytical target cascading (ATC) [27-29] are among the most popular approaches to solve OPF in a distributed manner. Most of the existing papers focus on multi-area power transmission systems or microgrid clusters. However, few papers exist that focus on collaborative TSO and DSOs optimal power flow. An iterative master-slave algorithm is presented in [30-32] to manage power transmission and distribution systems in a collaborative manner. In [30], a heterogeneous decomposition is presented to solve a collaborative ACOPF for transmission and distribution systems. This decomposition approach, which works similar to optimality condition decomposition, solves first-order KKT conditions in a decentralized manner. This is a sequential solution procedure in which DSO is solved first, and then TSO is solved. In [27-29], we presented a decentralized algorithm for collaborative day-ahead scheduling of TSO and DSOs. The coordination strategy is based on ATC, which is for multilevel distributed optimization of hierarchical complex engineering systems. References [27, 28] apply augmented Lagrangian block coordinate descent (AL-BCD) while [29] utilizes an exponential penalty function (EPF) formulation. Although AL-BCD and EPF formulations effectively coordinate TSO and DSOs, their main drawback is their sequential solution procedure. In other words, at each iteration, TSO (DSOs) needs the updated values of the shared variables received from DSOs (TSO) at the same iteration. This degrades the computational efficiency of the decentralized algorithm, as the computation time is a summation of the subproblems' solution time. The main goal of this paper is to address the drawback of collaborative ATC-based TSO+DSO operation by enabling a *parallel execution* of subproblems' optimization.

In this paper, we present a decentralized collaborative two-level TSO+DSO optimal power flow solution. The proposed algorithm is based on analytical target cascading (ATC) method and allows a *fully parallel* implementation TSO+DSO OPF. A local OPF problem is formulated for TSO and each DSO which accounts for interactions between the transmission and distribution systems. A limited amount of information is exchanged among TSO and DSOs which is in line with respecting the information privacy of the autonomous control entities. While the transmission OPF problem is formulated and solved in the upper-level of hierarchy, the distribution OPF is handled in the lower-level. Two coordination strategies, namely diagonal quadratic approximation (DQA) and truncated diagonal quadratic approximation (TDQA), are presented to coordinate the local OPF problems in a parallel manner. While DQA needs two loops, one inner loop, and one outer loop, TDQA follows an iterative procedure with one loop. A 22-bus test system and the IEEE 118-bus transmission system are used for simulation studies.

The contributions of the paper are summarized as follows:

- The power system is modeled as a system of systems (SoS) in which TSO and DSOs are autonomous entities with their local policies and rules. A collaborate two-level OPF is presented with respect to a) interdependencies of transmission and distribution systems and b) the information privacy of TSO and DSO.
- Interdependencies between TSO and DSOs are modeled by a set of hard constraints. Quadratic penalty terms are utilized to relax the hard constraints in the local objective of each entity. A technique is presented to make non-separable quadratic terms of augmented Lagrangian penalty functions separable.
- A fully parallel solution algorithm is presented which has two loops: an inner loop to enhance the accuracy of the solution and outer loop to force the algorithm to converge. At each iteration of the proposed parallel procedure, TSO (DSOs) needs the updated values of the shared variables received from DSOs (TSO) obtained at the previous iteration. Hence, compared with the sequential algorithm, the computation time of each iteration decreases. This can significantly improve the convergence speed as the number of levels increases.

The main differences between this paper and [28] are as follows

- We dealt with a unit commitment problem in [28]; however, in this paper, we deal with the OPF problem.
- In [28], the concept of shift factor is used to formulate DC power flow for transmission and distribution systems; however, in this paper, AC-OPF is formulated using voltage magnitudes and angles.
- In [28], pseudo generations and loads are used to model energy exchange between TSO and DSOs; however, in this paper, voltage magnitudes and angles of border buses are modeled as shared variables. This shared variable modeling paradigm enables a user to handle cases with a loop between TSO and DSO without degrading information privacy. However, if a loop exists between TSO and DSO, a coordinator is needed that gathers information from the whole network to calculate the shift factor values. This degrades the information privacy.
- The solution algorithm presented in [28] is a sequential procedure in which while TSO (DSOs) is solving its subproblem, DSOs (TSO) should stay idle. However, the solution procedure presented in this paper is a fully parallel

approach that allows parallel (and simultaneous) solution of all OPF subproblems. This parallel approach reduces the computational time of each iteration.

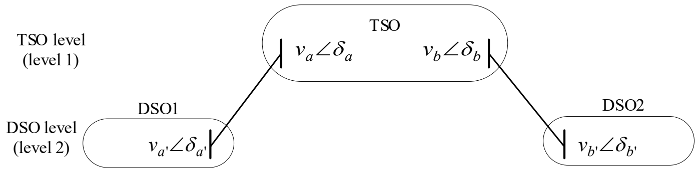


Fig.1. Interdependency of TSO and DSOs with coupling variables.

The remainder of the paper is organized as follows. The decentralized ATC-based optimal power flow is presented in Section II. The proposed decentralized decision-making framework as well as DQA and TDQA solution algorithms are presented in Section III. Numerical results are discussed in Section IV. Concluding remarks are provided in Section V.

II. DECENTRALIZED ATC-BASED OPF IMPLEMENTATION

A. Dependency of TSO and DSO

Assume that the transmission network is not connected to the active distribution grids. In this case (isolated mode), TSO and DSOs are capable of solving their local OPF problems completely independent from one another. However, this is not the case in reality since distribution grids are interconnected to transmission networks via one or more connection points. Consider the system shown in Fig. 1, which includes one TSO and two DSOs. The system has two levels. TSO is on the first level (upper-level), and DSOs are in the second level (lower-level). Control variables of buses a and a' (i.e., voltage magnitudes and angles) couple TSO and DSO1. Both TSO and DSO1 are interested in controlling these coupling variables to improve their grid performance. Likewise, TSO and DSO2 are coupled via control variables of buses b and b'. The coupling variables, i.e., $\{v_a\angle\delta_a, v_{a'}\angle\delta_{a'}, v_b\angle\delta_b, v_{b'}\angle\delta_{b'}\}$, make decisions of TSO, DSO1, and DSO2 interdependent (note that active and reactive power flows in a tie-line are by-products of the voltage magnitudes and angles of ending terminals of the tie-line). Thus, coordination of the aforementioned coupling variables is in great interest of TSO and DSOs.

B. Characterization of Analytical Target Cascading

The general concept of analytical target cascading (ATC) is similar to the auxiliary problem principle (APP) and alternating direction method of multipliers (ADMM) [20, 23, 33-35]. The ATC procedure (which is suitable for multilevel management of complex engineering systems) first decomposes the system into a multilevel hierarchical structure (as shown in Fig. 2) and recognizes parents and children. At the next step, penalty functions are introduced to model subproblems' interdependencies. Whereas in APP and ADMM, the duality concept is applied and penalty functions are introduced, and then the system is decomposed into several subproblems. As shown in Fig.1, TSO in the upper-level is hierarchically connected to DSOs in the lower-level. Thus, ATC is a suitable method to solve the collaborative TSO+DSO operation in a decentralized manner. In ATC, subproblems (also called elements or autonomous systems) in the upper-levels are parents of subproblems in the lower-levels. By the same token, subproblems in the lower-levels are children of subproblems in the upper-levels. Although a child has only one parent, a parent could have multiple children. This hierarchical interconnection means that there is no loop in the ATC structure. This further implies that subproblems at the same level do not share any connection/information with each other. If we assume the ATC structure as a graph, subproblems and tie-lines are respectively nodes and edges of the graph.

By decomposing the system into parents/children, dimensionality of each subproblem reduces. An iterative solution procedure can be applied to coordinate TSO and DSOs and determine the optimal solution of the SoS-based power system. In ATC, the coupling variables between two connected elements appear in the form of target variables and response copiers. TSO solves its OPF subproblem and propagates the target values down toward its children (i.e., DSOs). Then, DSOs use the updated target values, solve their local OPF problems, and send the updated values of the response copiers back to TSO. The responses determined by the children define how close they are to the parent's targets [33].

To enforce the decentralized optimization problem to converge, a proper coordination strategy is required. Several methods have been proposed in literature with different options to penalize the coupling variables into the objective functions. These options for selection of penalty terms and coordination strategies make ATC more flexible than ADMM and APP. Augmented Lagrangian block coordinate descent (AL-BCD) and exponential penalty function (EPF) are two popular ATC formulations that use coordination strategies with two loops, inner loop and outer loop. The penalty terms in these two methods are not separable, and thus the solution algorithm is a *sequential* procedure as shown in Fig. 3. It should be noted that if no direct link exists among the subproblems in each level i, the subproblems (only those is level i) can be solved in parallel.

The ATC structure converges to first order optimality conditions, if the problem is convex[35]. Thus, ATC provides the optimal solution for a convex problem. As shown in the literature, ATC shows good performance for non-convex problems, such as ACOPF presented in this paper [27-29, 36]. In addition, as explained in the following sections, a set of convex penalty functions, such as a quadratic function, are added to the objective function. These convex penalty functions act as local convexifiers for the subproblems and mitigate the non-convexity of the parents' and children' subproblems. The convergence and optimality of the decentralized algorithm, when applied to the studied problem, are demonstrated through several numerical simulations.

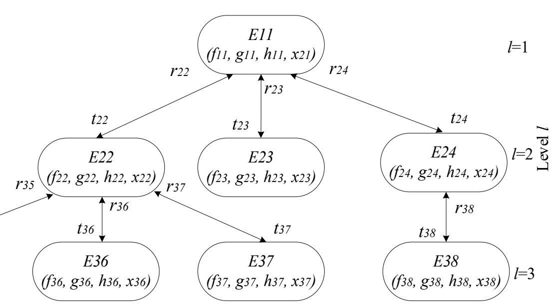


Fig. 2. Decomposing a system into a multilevel hierarchical structure.

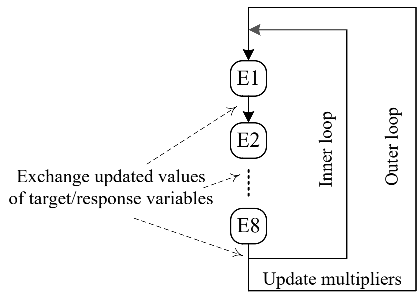


Fig. 3. Solution procedure of AL-BCD and EPF algorithms.

C. Hierarchical Two-Level TSO+DSO Operation

In this section, we formulate the collaborative TSO+DSO OPF within the ATC framework. Consider that optimization (1) expresses a centralized OPF problem for the entire power transmission and distribution systems.

$$\min_{X} F(X) \tag{1}$$
$$s.t.\ \ g(X) \leq 0, \quad h(X) = 0$$

where X denotes all variables of the entire power system, F is the overall objective function, and g and h represent all inequality and equality constraints. The power system has a two-level hierarchical structure (a simplified version of Fig. 2). Thus, within the ATC framework, we can rewrite (1) as follows:

$$\min_{(x_{ij},t_{2j})} f_{11}(x_{11}, t_{2j}) + \sum_{j=2}^{N_d+1} f_{2j}(x_{2j}, t_{2j}) \tag{2}$$
$$s.t.\ \ g_{ij}(x_{ij}, t_{2j}) \leq 0 \qquad h_{ij}(x_{ij}, t_{2j}) = 0$$

where subscript ij indicates subproblem j^{th} in level i, $X = \{x_{ij}, t_{2j}\}$, x_{ij} is local variables of subproblem j in level i, and t_{2j} represents the target variables. Note that in ATC, the shared variables that couple TSO to DSOs (i.e., voltage of border buses as discussed in Section II. A) appear in form of target variables. Parameter N_d is number of DSO in level 2, f_{11} is the objective function of TSO, f_{2j} is the objective function of DSO j in level 2, and g_{ij} and h_{ij} are compact representations of inequality and equality constraints of subproblem j in level i. If t is an empty set ($t = \{\}$), TSO and DSOs are isolated and can solve their local OPF subproblems completely separate from each other. However, if t is non-empty ($t \neq \{\}$), which is the case in the power systems, subproblems that share t (an element of t) need to achieve an agreement on its value.

To separate the TSO's and DSOs' OPF subproblems as well as the variables that are governed by each subproblem, response copiers are introduced. The response copiers are duplicates of the target variables. We consider that the target variables are the shared variables (voltage of border buses) that are handled by TSO and the response copiers are the shared variables that are governed by DSOs. We can include the response variables, denoted generically by vector r, in (2) by enforcing a set of consistency constraints as:

$$\text{C:}\ \ t_{2j} - r_{2j} = 0 \tag{3}$$

One consistency constraint is required for each target-response pair. We relax the consistency constraints in the objective function using a penalty term.

$$\min_{(x_{ij},t_{2j},r_{2j})} f_{11}(x_{11}, t_{2j}) + \sum_{j=2}^{N_d+1} f_{2j}(x_{2j}, t_{2j}) + \sum_{j=2}^{N_d+1} \pi(t_{2j} - r_{2j}) \tag{4}$$

Now, we can completely separate the local OPF subproblems of TSO and DSOs. Let us represent the target variables (i.e., voltage of border buses) by their common notations in power system communities, i.e., $v\angle\delta$. Also, $\tilde{v}\angle\tilde{\delta}$ represents the response variables. The local OPF subproblem of TSO is:

$$\min_{(x_{11},\delta_{2j},v_{2j})} f_{11}(x_{11}, \delta_{2j}, v_{2j}) + \sum_{j=2}^{N_d+1} \pi(\delta_{2j} - \tilde{\delta}_{2j}) + \pi(v_{2j} - \tilde{v}_{2j}) \tag{5}$$
$$s.t.\ \ g_{11}(x_{11}, \delta_{2j}, v_{2j}) \leq 0 \qquad h_{11}(x_{11}, \delta_{2j}, v_{2j}) = 0$$

And the local OPF subproblem of DSO j is:

$$\min_{(x_{2j},\tilde{\delta}_{2j},\tilde{v}_{2j})} f_{2j}(x_{2j}, \tilde{\delta}_{2j}, \tilde{v}_{2j}) + \pi(\delta_{2j} - \tilde{\delta}_{2j}) + \pi(v_{2j} - \tilde{v}_{2j}) \tag{6}$$
$$s.t.\ \ g_{2j}(x_{2j}, \tilde{\delta}_{2j}, \tilde{v}_{2j}) \leq 0 \qquad h_{2j}(x_{2j}, \tilde{\delta}_{2j}, \tilde{v}_{2j}) = 0$$

The OPF subproblem (5) ((6)) is formulated using the local information of TSO (DSO j) as well as its shared variables with DSOs (TSO). The generation cost function of each subproblem (i.e., f) is a quadratic function as $f(p) = a + bp + cp^2$. The equality constraints h and the inequality constraints g of TSO and DSOs are as follows:

$$h\!:\begin{cases}\text{Nodal power balance equations}\\ \text{Voltage angle of the reference bus} = 0\end{cases}$$

$$g\!:\begin{cases}\text{Generation capacity limits}\\ \text{Bus voltage limits}\\ \text{Line flow limits}\end{cases}$$

While TSO is allowed to decide about its local and target variables $v\angle\delta$, each DSO j determines its local and corresponding response variables $\tilde{v}\angle\tilde{\delta}$. This is, while $v\angle\delta$ is constant in the DSOs' OPF subproblems, $\tilde{v}\angle\tilde{\delta}$ is constant in the TSO's OPF subproblem. In the ATC framework, TSO sends the target values $v\angle\delta$ down to DSOs, and each DSO sends its response values $\tilde{v}\angle\tilde{\delta}$ back upward TSO.

An iterative procedure needs to be implemented to enforce the difference between $v - \tilde{v}$ and $\delta - \tilde{\delta}$ to zero and find the optimal solution of the entire two-level power system. Depending on the choice of the penalty function $\pi(\cdot)$, the iterative solution procedure could be implemented in a sequential or a parallel fashion. An algorithm in which the TSO and DSOs OPF subproblems are sequentially and iteratively solved is called *block coordinate descent*. The convergence of the algorithm is guaranteed in [35, 37]. This is independent of the choice of the penalty function since the constraint sets of TSO and DSOs are completely independent.

In [28, 29], we have applied AL-BCD and EPF methods to model the penalty function (π). In AL-BCD, the penalty term is

$$\lambda^T(t-r)+\|w\circ(t-r)\|_2^2 \qquad t=\{v,\delta\}, r=\{\tilde{v},\tilde{\delta}\} \tag{7}$$

And in EPF, the penalty term is

$$\alpha\left(\mathrm{e}^{(t-r)}-1\right)+\beta\left(\mathrm{e}^{(r-t)}-1\right) \quad t=\{v,\delta\}, r=\{\tilde{v},\tilde{\delta}\} \tag{8}$$

where λ, w, α, and β are penalty multipliers, and "$\circ$" denotes the Hadamard product. Setting the penalty factor w to a small value enhances the accuracy of the distributed algorithm but it increases the number of iterations. A large w potentially reduces the number of iterations but it may degrade the accuracy of the results. Indeed w should be set to a large enough value (this value is problem dependent) to balance the cost function f and the penalty function and make a trade-off between the accuracy and speed [38]. The penalty multipliers λ, α, and β should be initialized close to their optimal values. A user may utilize historical data (e.g., a hot start strategy) or its experience to initialize the penalty multipliers.

Penalty functions (7) and (8) include non-separable terms. Thus, a sequential solution procedure (hierarchical and level by level similar to Fig. 3) is required to solve the collaborative ATC-based TSO+DSO OPF. That is, in each iteration k, TSO (DSOs) needs to know the response (target) values of DSOs (TSO) in that iteration, i.e., r^k (t^k). Hence, when the TSO's (DSOs') OPF subproblem is being solved, the DSOs' (TSO's) OPF subproblems should stay idle (see[28] for more details). This degrades computational efficiency of the decentralized solution procedure.

III. Diagonal Quadratic Approximation Method for Parallel Solution

It is highly desirable to solve the OPF subproblems in a parallel manner as shown in Fig. 4, especially when multiple levels of hierarchy (e.g., TSO, DSO, and microgrid levels) and many subproblems exist. In this paper, diagonal quadratic approximation (DQA) and truncated diagonal quadratic approximation (TDQA) are presented to *parallelize* the solution procedure of the collaborative ATC-based TSO+DSO OPF. In these two algorithms, a subproblem with the longest solution time determines the algorithms' solution time in each iteration. In contrast, in a sequential algorithm, such as AL-BCD, the summation of TSO's solution time and the longest solution time of DSOs determines the overall solution time of the algorithm.

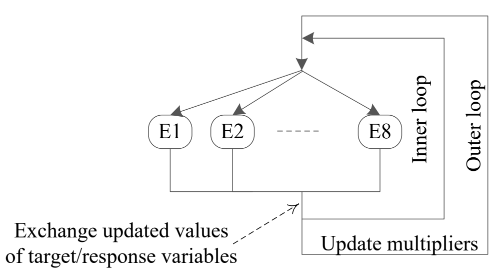


Fig. 4. Solution of OPF subproblems with parallel ATC coordination strategy.

A. Diagonal Quadratic Approximation (DQA)

The objective functions f, i.e., the generation cost functions, in TSO and DSO subproblems are convex functions. In addition, the local equality and inequality constraints of each subproblem are fully separable. In ATC, we have the flexibility to select the penalty term $\pi(\cdot)$ to relax the consistency constrains in the local objective functions. We have followed the concept of augmented Lagrangian and selected a combination of linear and quadratic penalty functions as in (7). The augmented term, i.e., the quadratic term, improves the convergence performance compared with the ordinary Lagrangian function. In addition, this penalty term acts as a local convexifier and enhances behavior of the subproblems. However, this quadratic term of the augmented Lagrangian penalty function is not separable. We apply the diagonal quadratic approximation (DQA) method to make the augmented Lagrangian terms separable[39, 40]. Consider the penalty function corresponding to the voltage angle. We expand its quadratic term as:

$$\left\|\delta_{ij}-\tilde{\delta}_{ij}\right\|_2^2=\left\|\delta_{ij}\circ\delta_{ij}+\tilde{\delta}_{ij}\circ\tilde{\delta}_{ij}-2(\delta_{ij}\circ\tilde{\delta}_{ij})\right\| \tag{9}$$

We use the first order Taylor expansion for multiple variable scalar functions to linearize the cross term $\delta_{ij}\circ\tilde{\delta}_{ij}$ at the point (δ_{ij}^{k-1}, $\tilde{\delta}_{ij}^{k-1}$) [34, 35].

$$\delta_{ij}\circ\tilde{\delta}_{ij}\cong\tilde{\delta}_{ij}^{k-1}\circ\delta_{ij}+\delta_{ij}^{k-1}\circ\tilde{\delta}_{ij}-\delta_{ij}^{k-1}\circ\tilde{\delta}_{ij}^{k-1} \tag{10}$$

where δ_{ij}^{k-1} and $\tilde{\delta}_{ij}^{k-1}$ are respectively targets and responses determined in the previous iteration $k-1$ and are constant in the current iteration k. Thus, we can approximate the quadratic penalty term (9) as:

$$\left\|\delta_{ij}-\tilde{\delta}_{ij}\right\|_2^2=\left\|\delta_{ij}^{k-1}-\tilde{\delta}_{ij}\right\|_2^2+\left\|\delta_{ij}-\tilde{\delta}_{ij}^{k-1}\right\|_2^2+\mathrm{C} \tag{11}$$

where C is a constant. The same Taylor expansion is implemented on the quadratic penalty term corresponding to voltage magnitudes. Now, the OPF problem of the entire two-level power system can reformulate as:

$$\begin{aligned}\min_{(x_{ij},\delta,v,\tilde{\delta},\tilde{v})}&\sum_{i=1}^{N_l}\sum_{j}f_{ij}\left(x_{ij},\delta_{(i+1)j},v_{(i+1)j}\right)\\&+\sum_{i\in N_{l,i\neq1}}\sum_{j}\Big(\lambda_{ij,\delta}^T\left(\delta_{ij}-\tilde{\delta}_{ij}\right)+\left\|w_{ij,\delta}\circ\left(\delta_{ij}^{k-1}-\tilde{\delta}_{ij}\right)\right\|_2^2\\&\quad+\left\|w_{ij,\delta}\circ\left(\delta_{ij}-\tilde{\delta}_{ij}^{k-1}\right)\right\|_2^2\Big)\\&+\sum_{i\in N_{l,i\neq1}}\sum_{j}\Big(\lambda_{ij,v}^T\left(v_{ij}-\tilde{v}_{ij}\right)+\left\|w_{ij,v}\circ\left(v_{ij}^{k-1}-\tilde{v}_{ij}\right)\right\|_2^2\\&\quad+\left\|w_{ij,v}\circ\left(v_{ij}-\tilde{v}_{ij}^{k-1}\right)\right\|_2^2\Big)\end{aligned} \tag{12}$$

where N_l denotes number of levels which is two in this paper. This optimization problem is subject to all-in-once constraints, i.e., all constraints of TSO and DSOs. We now decompose (12). The local OPF subproblem of TSO in iteration k of the ATC procedure is as follows:

$$\begin{aligned}\min_{(x_{11},\delta,v)}&f_{11}\left(x_{11},\delta_{22},v_{22},\delta_{23},v_{23},\dots,\delta_{2(N_d+1)},v_{2(N_d+1)}\right)\\&+\sum_{j=2}^{N_d+1}\lambda_{2j,\delta}^T\delta_{2j}+\left\|w_{2j,\delta}\circ\left(\delta_{2j}-\tilde{\delta}_{2j}^{k-1}\right)\right\|_2^2\\&+\sum_{j=2}^{N_d+1}\lambda_{2j,v}^T v_{2j}+\left\|w_{2j,v}\circ\left(v_{2j}-\tilde{v}_{2j}^{k-1}\right)\right\|_2^2\end{aligned} \tag{13}$$

subject to the local constraints of TSO (e.g., nodal power balance, line flow limits, etc.). The penalty term depends on the target variables $v_{2j}\angle\delta_{2j}$ while using the response values $\tilde{v}_{2j}^{k-1}\angle\tilde{\delta}_{2j}^{k-1}$ determined by DSOs in the previous iteration $k-1$. TSO solves its local OPF problem and find the target values. Likewise, the local OPF subproblem of each DSO j is reformulated as:

$$\min_{(x_{2j},\tilde{\delta}_{2j},\tilde{v}_{2j})} f_{2j}(x_{2j},\tilde{\delta}_{2j},\tilde{v}_{2j}) + \lambda_{2j,\delta}^T(-\tilde{\delta}_{2j}) + \lambda_{2j,v}^T(-\tilde{v}_{2j}) \\ +\left\|w_{2j,\delta}\circ(\delta_{2j}^{k-1}-\tilde{\delta}_{2j})\right\|_2^2 + \left\|w_{2j,v}\circ(v_{2j}^{k-1}-\tilde{v}_{2j})\right\|_2^2 \quad (14)$$

subject to the local constraints of DSO j. The penalty term depends on the response variables $\tilde{v}_{2j}\angle\tilde{\delta}_{2j}$ while using the target values $v_{2j}^{k-1}\angle\delta_{2j}^{k-1}$ determined by TSO in the previous iteration $k-1$. Formulations (13) and (14) allow a parallel solution of the TSO's and DSOs' OPF subproblems since each subproblem needs the target/response values determined by other subproblems in iteration $k-1$. The DQA coordination strategy is proven to converge and its convergence rate is discussed in [39] and [40].

A.1. Parallel Solution Procedure of DQA

Figure 5 illustrates the solution procedure of DQA to coordinate the OPF subproblems of TSO and DSOs. Although the problem's structure has a hierarchical two-level form, the presented coordination strategy is a parallel procedure that allows a simultaneous solution of TSO's and DSOs' subproblems. The DQA solution strategy includes two loops, inner loop and outer loop. The inner loop updates the target and response values while the penalty multipliers are fixed. This improves the linearization. The inner loop stops when the difference between each target (response) determined in two consecutive iterations are less than a threshold. Indeed, the inner loop seeks to find the best targets and responses for a given set of multipliers. If the targets and responses are determined more precise, the penalty multipliers are updated more accurate in the outer loop. If the penalty multipliers are updated more accurate, the algorithm takes less iterations to update the multipliers. Thus, although the inner loop increases the computational cost (corresponding to the inner loop iterations), it might reduce the number of outer loop iterations in which the multipliers are updated (i.e., the method of multipliers). The steps are discussed in details as follows:

Step1: Set the initial value of local variables x of each subproblem, target values $\{\delta,v\}$, response copiers $\{\tilde{\delta},\tilde{v}\}$, penalty multipliers λ and w, and parameters Γ and τ. Set the outer loop iteration index $k=1$ and the inner loop iteration index $l=0$.

Step2: Increase the inner loop iteration by one, i.e., $l=l+1$. Solve TSO's and DSOs' local OPF subproblems in parallel using targets and responses that are determined in the previous inner loop iteration ($l-1$), i.e., $\delta^{k-1,l-1}$ and $\tilde{\delta}^{k-1,l-1}$. Note that in the first iteration, the subproblems are solved using the initial values.

Step3: Check the following inner loop convergence criterion

$$\max\left(\|\delta^{k,l}-\delta^{k,l-1}\|,\left\|\tilde{\delta}^{k,l}-\tilde{\delta}^{k,l-1}\right\|,\|v^{k,l}-v^{k,l-1}\|,\|\tilde{v}^{k,l} \\ -\tilde{v}^{k,l-1}\|\right)\leq\epsilon_{inner} \quad (15)$$

where ϵ_{inner} is the stopping threshold of the inner loop. If the difference between the target (and response) values determined in the current and previous iterations are less than the acceptable threshold, then we should stop the inner loop, set $X^k = X^{k,l}$ (where $X=[x,\delta,v,\tilde{\delta},\tilde{v}]$), and go to Step 4; otherwise:

$$X^{k,l} = X^{k,l-1} + \Gamma(X^{k,l} - X^{k,l-1}) \quad (16)$$

where Γ is the step size, which determines a value among the current solution and the previous one (i.e., if $\Gamma\cong 0$, the algorithm uses with the previous solution, and if $\Gamma\cong 1$ the algorithm uses with the current solution), and then go to Step 2. Note that we update the initial values of all local and shared variables.

Step4: If $\max\{\|\delta^k-\tilde{\delta}^k\|,\|v^k-\tilde{v}^k\|\}\leq\epsilon_{outer}$ (i.e., if the difference between each target-response pair is less than the criterion), where ϵ_{outer} is the outer loop stopping threshold, TSO+DSO OPF has converged and the optimal values are $\bar{X}^* = \bar{X}^k$, otherwise increase the outer loop iteration index by one (i.e., $k=k+1$) and set the inner loop index to zero (i.e., $l=0$) and update the penalty multipliers as follows:

$$\lambda_\delta^k = \lambda_\delta^{k-1} + w_\delta^{k-1}\circ(\delta^{k-1}-\tilde{\delta}^{k-1}) \quad (17)$$

$$\lambda_v^k = \lambda_v^{k-1} + w_v^{k-1}\circ(v^{k-1}-\tilde{v}^{k-1}) \quad (18)$$

$$w_\delta^k = \tau_\delta w_\delta^{k-1} \quad (19)$$

$$w_v^k = \tau_v w_v^{k-1} \quad (20)$$

and then go to Step 2 (note that multipliers will be updated for every outer loop iteration). Parameter τ should be equal or large than one, i.e., $\tau\geq 1$ [41]. Depending on the optimization problem characteristics, a wide range of τ can be selected to reduce the computational cost and/or enhance the solution accuracy. Based on our experience, setting τ close to one provides an accurate solution while the computational burden is reasonable.

$\Gamma\in(0,1)$ is the step size that affects accuracy of the linearization of the second-order penalty term. A small Γ leads to more accurate results, but it decreases the convergence speed. Parameters ϵ_{outer} and ϵ_{inner} should be significantly smaller than Γ; otherwise, the obtained solution might not be optimal.

The ATC method is proven to converge to an accumulation point (i.e., the shared variables converge to a unique point) that satisfies the first-order optimality conditions of the local optimization problems. This accumulation point also satisfies the first-order optimality conditions of the original problem [35]. In addition, [39] provides the convergence proof and convergence rate of the diagonal quadratic approximation method when applied to separate subproblems of the augmented Lagrangian approach. It is worth to mention that the quadratic penalty terms act as local convexifiers and improves the performance of ATC when applied to non-convex problems.

B. Truncated Diagonal Quadratic Approximation (TDQA)

The collaborative ATC-based TSO+DSO optimal power flow converges when the optimal values of Lagrange multipliers are found. As explained in the DQA solution procedure, the multipliers are not updated in the inner loop. The inner loop helps to improve the linearization. Each iteration of the outer loop, in which the

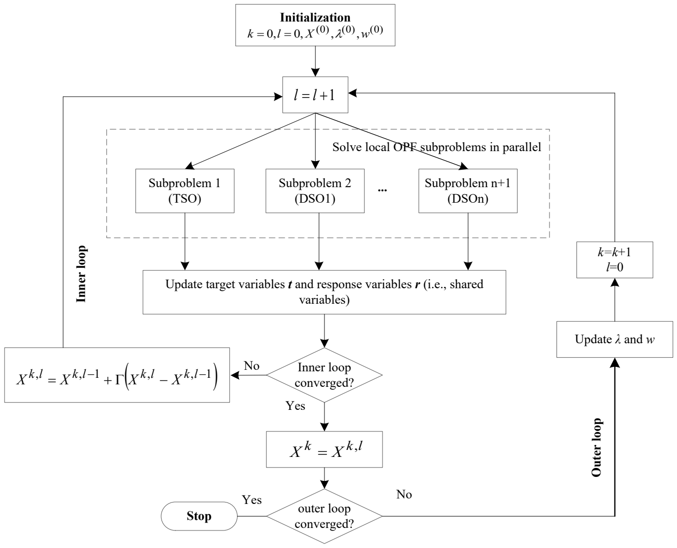


Fig. 5. DQA solution algorithm.

multipliers are updated, might take many inner loops. Thus, the inner loop increases the computational effort. Obtaining high accurate solution of the OPF subproblems in the inner loop is not necessary as the inner loop solution might not be the overall optimal solution. If we only solve the outer loop and update the Lagrange multipliers after every iteration, the multipliers move quickly toward the optimal values. Thus, we omit the inner loop and only consider the outer loop and update the Lagrange multipliers after every iteration. This single-loop coordination strategy is called truncated diagonal quadratic approximation (TDQA) [39, 40]. Note that one can consider the inner loop, but limiting its iterations with any extra criterion (in addition to DQA criterion) rather than allowing it to be able to go to infinity. This procedure is also TDQA as the inner loop is truncated compared with DQA.

The solution procedure of the TSO+DSO operation with TDQA is summarized in the following pseud code. It has a similar structure as DQA except that DQA has inner loop to decrease the gap between targets and responses and then update multipliers while TDQA performs this only by the outer loop. In the case study section, we show that TDQA provides promising results for the collaborative TSO+DSO operation.

Although inner loop enhances accuracy of the targets and responses over the course of iterations, it is not necessary for convergence. Indeed, updating penalty multipliers in the outer loop (which is based the method of multipliers) guarantees the convergence of the ATC-based algorithm to the first optimality conditions. That is, TDQA might slightly increase error; however, its convergence is still ensured.

IV. Numerical Results

We implement the DQA and TDQA coordination strategies on a 6-bus and the IEEE 118-bus test systems. The numerical simulations show efficiency and convergence of the ATC-based collaborative TSO+DSO algorithm even for the non-convex OPF problems. All computations are carried out using quadratic programming solver of Matlab on a 2.6GHz personal computer with 16GB of RAM.

Solution Algorithm of TDQA

1: **initialize** $X = [x, \delta, v, \tilde{\delta}, \tilde{v}]$, λ, w, and τ
2: **while** $\max(\|\boldsymbol{\delta}^k - \tilde{\boldsymbol{\delta}}^k\|, \|\boldsymbol{v}^k - \tilde{\boldsymbol{v}}^k\|) \leq \epsilon_{outer}$, $k = k+1$ **do**
3: Solve (13) and (14) in a parallel manner and determine X^k
4: Update X: $X^k = X^{k-1} + \Gamma(X^k - X^{k-1})$
5: Update multiplier λ^k and w^k using (17)-(20)
6: **end while**

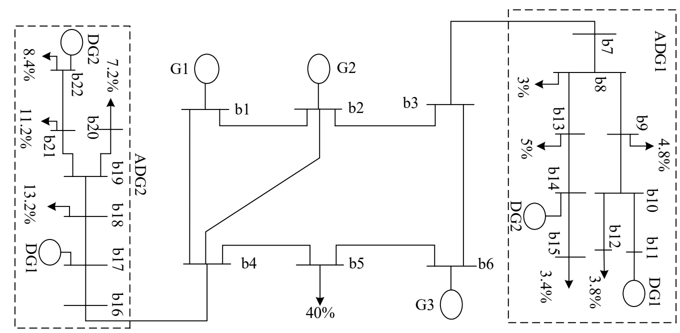


Fig. 6. Six-bus test system.

A. 6-Bus Test System

The system topology is shown in Fig. 6. The transmission system includes six buses, seven transmission lines, and three generators. Two active distribution grids are connected to the transmission system. Active distribution grid one consists of nine buses, five loads, and two DGs. Active distribution grid two includes seven buses, four loads, and two DGs. The total load is 256MW. The resistance of each distribution line is 40% of the line's reactance. The reactive power limit of each generating unit is considered to be 60% of its active power limit. The reactive power consumption of each load is assumed to be 30% of its active power demand. The rest of the information is given in [28]. We study and analyze the following four cases:

Case 1: Centralized OPF implementation considering a single operator for transmission and distribution networks

Case 2: The proposed decentralized decision-making with respect to autonomy of TSO and DSOs

Case 3: Sensitivity of the proposed collaborative TSO+DSO OPF to variation of input parameters

Cases 4: Comparison between TDQA, APP, ADMM, and ALAD

Case 1: We ignore autonomy and information privacy of TSO and DSOs and consider that the transmission and distribution networks are operated by the same operator using the centralized OPF method. Although this is not a realistic case since TSO and DSOs are autonomous, the centralized method provides the references results that can be used to evaluate the performance of the decentralized decision-making. Since the ratio of lines reactance to resistance is large in the transmission system, DC-OPF is a reasonable approximation of AC-OPF. Thus, DC-OPF is used for TSO, whereas AC-OPF is used for DSOs[1]. To model the reactive mismatch at border buses between TSO and DSOs, we consider voltage magnitude at transmission terminals equal to one, while the voltage magnitude at the distribution terminals can vary between

[1] A user may deploy a linearized model of ACOPF. Since OPF is solved continually, in each time interval, the user can deploy results of the previous interval (i.e., a hot start) to linearize OPF around the operating point. Also, the user may convexify ACOPF using techniques such as semidefinite programing or second order cone programing [42, 43]

0.95 and 1.05 (note that since the voltage at the transmission terminal is close to one, the reactive mismatch in the tie-line is not significant compare with the case that AC-OPF is considered for TSO). The total operating cost of the system is \$3,396. The operating costs of TSO, DSO1, and DSO2 are respectively \$2,375, \$351.7, and \$669.3. The voltage phasors of buses that connect TSO to DSO1 are $1\angle-0.0293$ and $0.998\angle-0.0385$, and the voltage phasors of buses that connect TSO to DSO2 are $1\angle-0.0459$and $1.0018\angle-0.0797$.

Case 2: In this case, autonomy and information privacy of the three systems (i.e., TSO, DSO1, and DSO2) are taken into account, and each system is operated by an independent operator. We have considered the same operation horizon for TSO and DSOs, e.g., 5-minute interval. The OPF problems are run for one snapshot, and it is assumed that the entities start solving their subproblems simultaneously. Note that even if the operation intervals of TSO and DSOs are not the same, to allow a collaborative operation, we can consider the operation horizon equal to the longest interval. TSO is the parent, and its children are DSOs 1 and 2. A tie-line connects the border bus b3 of the transmission system to the border bus b7 of ADG1, and another tie-line links the border bus b4 of the transmission system to the border bus b16 of ADG2. Thus, voltage of buses b3 and b7 are the shared variables between TSO and DSO1, and voltage of buses b4 and b16 are the shared variables between TSO and DSO2. TSO includes four target variables, and each DSO has two response variables. We analyze cold start and hot start conditions.

Cold start: The initial values for targets/responses are set to zero, and the initial values of penalty multipliers/parameters are λ^0=1000, w_δ^0=1500, $w_{v7,ADG1}^0=30$, $w_{v16,ADG2}^0=10$, Γ=0. 9, and $\tau=1$. For DQA, the inner and outer loops' convergence thresholds are $\epsilon_{inner}=0.004$ and $\epsilon_{outer}=1.4\times10^{-4}$, respectively. Note that TDQA has only the outer loop with the convergence threshold of $\epsilon_{outer}=1.4\times10^{-4}$. The DQA coordination strategy converges after 33 outer loop iterations and the total number of executed inner loops is 40, whereas TDQA converges after 34 iterations. Figure 7(a) shows updating process of the target-response pair corresponding to the voltage angles of bus 3 of TSO and bus 1 of DSO1 over the course of iterations (note that for DQA, we show the updating process over the course of overall iterations, i.e., inner and outer loops). The difference between each pair of target-response becomes smaller and it is less that the convergence threshold in iteration 40 (34) of DQA (TDQA). Although in several iterations (e.g., iteration 18 in TDQA) a target and its corresponding response value might differ less than the stopping threshold, the algorithms stop when the differences between every pair of target-response become less than the stopping threshold. Although TDQA takes more (outer loop) iterations than DQA, it does not need the inner loop. In overall, DQA needs 40 iterations (sum of inner and outer loops iterations) which is 6 iterations more than that in TDQA. Although DQA needs more iterations than TDQA, it finds the solution more precisely especially when lower ϵ is chosen. Figure 7(b) shows the average difference between the target-response values over the course of iterations. Note that for DQA, the updating process is shown over the course of overall iterations. The error decays faster in TDQA than DQA because while DQA tries to enhance the solution by repeating the inner loop with the fixed multipliers, TDQA seeks to improve the solution by updating the multipliers. This reduces the overall number of function evaluations and the computation time of TDQA (DQA and TDQA take respectively 3.97 and 3.62 seconds to converge); however, it might slightly increase the overall error. Table I shows the generation dispatch for the three systems. Since the stopping threshold is not zero, the dispatch results obtained the centralized and decentralized algorithms are slightly different. However, as shown in Table II, the operating costs determined by the decentralized and decentralized algorithms are similar. The operating costs of TSO, DSO1, and DSO2 determined by DQA are \$2,379.8, \$ 350.8, and \$669.3, and they are \$2,378.3, \$351.7, and \$ 669.3 when using TDQA. The total power system operating costs determined by DQA and TDQA are \$3,399.9 and \$3,399.3 that are almost the same as the cost obtained by the centralized OPF (i.e., \$3,396). Note that the sensitivity of the solver and solution to changes in power generated by TSO's and DSOs' generators might be difference since the cost functions and geographical locations of the units are different. Thus, although power dispatch of DSOs' and TSO's units are slightly different from the centralized results, both algorithms yield almost the same operating cost.

To evaluate the performance of the proposed ATC-based TSO+DSO OPF in more details, we formulate two convergence indices. The first index is Euclidean norm of mismatch between the power flow in tie-lines connecting transmission system to active distribution grids (P_l^{ATC}) and the optimal value obtained by the centralized OPF (P_l^*):

$$P_{mis}=\left\|\frac{P_l^*-P_l^{ATC}}{P_l^*}\right\| \tag{19}$$

The second index is the relative distance of the total cost determined by ATC (f^{ATC}) from the optimal value determined by the centralized OPF (f^*):

$$rel=\frac{|f^*-f^{ATC}|}{f^*} \tag{20}$$

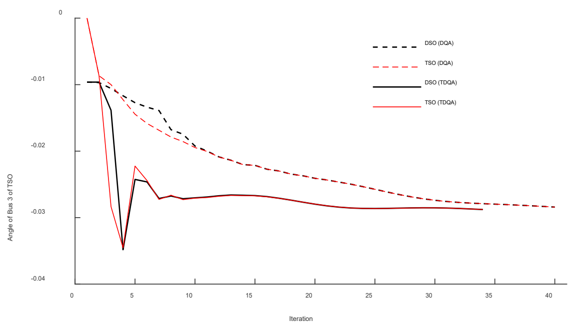
The values of the two convergence measures are zero at the optimal point. Hence, the closer these convergence measures get to zero, the better solution is obtained. Figure 8 shows P_{mis} and rel values over the course of iterations. The values of the convergence measures decrease when more iterations are carried out. They are small enough upon the algorithm convergence.

TABLE I. POWER OUTPUT OF GENERATING UNITS

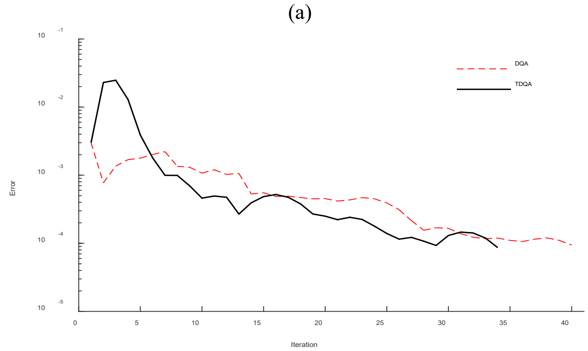
Algorithm	TSO			DSO1		DSO2	
	G1	G2	G3	DG1	DG 2	DG 1	DG2
Centralize	129.41	34.03	25	15	18	25	13.17
Dec. (DQA)	126.65	37.07	25	14.91	18	25	13.17
Dec. (TDQA)	127.93	35.72	25	15	18	25	13.17

TABLE II. OPERATING COST OBTAINED BY DIFFERENT ALGORITHMS

Algorithm	TSO	DSO1	DSO2	Total cost
Centralize	\$2,375	\$351.7	\$669.3	\$3,396
Dec. (DQA)	\$2,379.8	\$350.8	\$669.3	\$3,399.9
Dec. (TDQA)	\$2,378.3	\$351.7	\$ 669.3	\$3,399.3



(a)



(b)

Fig. 7. a) Target and response values corresponding to b3 of TSO and b7 of DSO1 and b) the average difference between targets and responses.

Hot start: In practice, we usually have good initial values for the variables and penalty multipliers. For example, when we solve the OPF problem for interval ω, we have the optimal results of interval $\omega-1$. We know that, in the most cases, the OPF input parameters, e.g., power demand, vary slightly from interval $\omega-1$ to interval ω. Thus, the solution of interval $\omega-1$ can be utilized to initialized the problem in interval ω. This is called hot start. We can solve the problem faster and more precise by selecting appropriate initial values. We assume that the load changes 5% between intervals $\omega-1$ and ω and use the solution obtained in interval $\omega-1$ to initialized the variables and penalty multipliers in interval ω. Figure 8 shows the convergence measure for the hot start and cold start cases. While DQA and TDQA take 40 and 34 iterations using the cold start, they take 16 and 17 iterations using the hot start, respectively. The rel index of cold start is 9×10^{-4}, whereas it is 1.06×10^{-4} using the hot start. Note that since we have good initial conditions (i.e., good guesses for target/response and penalty multipliers), DQA and TDQA behaviors are similar.

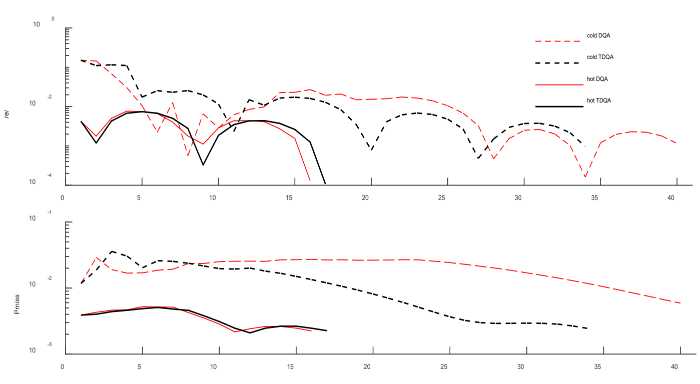


Fig. 8. Convergence property of DQA- and TDQA-based collaborative OPF.

A Full ACOPF: We have tested the proposed algorithm on a full ACOPF (i.e., ACOPF for TSO and ACOPF for DSOs). The initial values for targets/responses and multipliers are the same as the cold start strategy. We stop the algorithm after 60 iterations. The rel indices obtained from both approaches are shown in Fig. 9. The decentralized algorithm provides acceptable rel indices. Note that the combination of DCOPF and ACOPF provides good results in the first few iterations; however, the rel index gradually goes down for the case of the full ACOPF (a user may run the algorithm more than 60 iterations to get a smaller rel index). The DCOPF approximation for TSO slightly increases the error (because of linearization of ACOPF) but it enhances the performance of the decentralized optimization algorithm. The user may prefer to use such an approximation to get faster results from the decentralized algorithm. Note that using DCOPF for TSO and ACOPF for DSOs is aligned with the power industry.

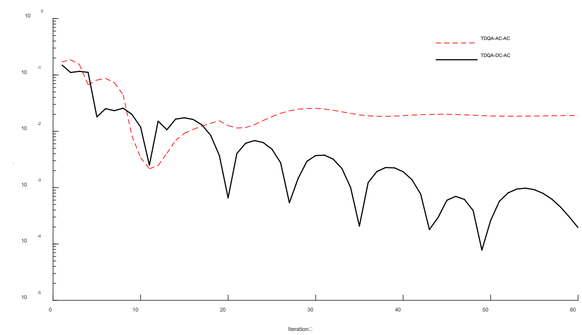
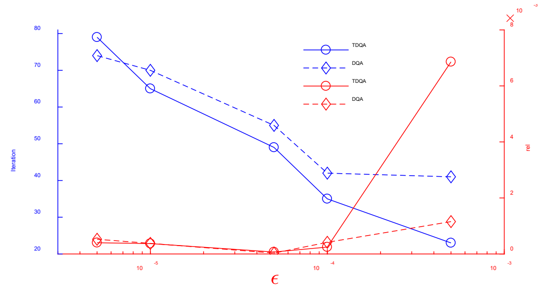


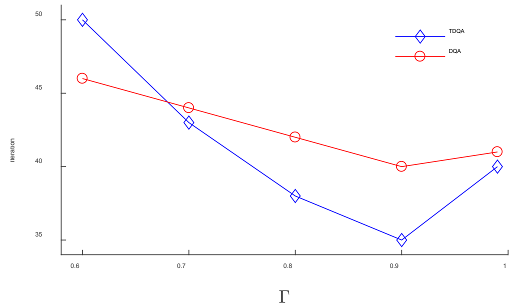
Fig. 9. The rel index for DCOPF+ACOPF and a full ACOPF.

Case 3: To evaluate the convergence behavior of the proposed collaborative TSO+DSO OPF with respect to variations of DQA/TDQA parameters, we perform multiple sensitivity analysis. This provides a user with insights on how to initialize the algorithms' parameters. We initialize λ^0=1000, w_δ^0=1500, $w_{v7,ADG1}^0=30$, $w_{v16,ADG2}^0=10$ and Γ=0.9. We select various stopping thresholds as 5×10^{-4}, 1×10^{-4}, 5×10^{-5}, 1×10^{-5} and 5×10^{-6} and demonstrate the relative error and number of iterations in Fig. 10(a). By decreasing the stopping criteria, the relative error decreases generally, but the number of iterations increases. One can select a small enough stopping threshold to make a trade-off between the stepped and error. When the stopping threshold is large, DQA's error is slightly smaller than that for TDQA. However, for the small thresholds, the relative errors of TDQA and DQA are almost the same. In overall, comparing the error and number of iterations shows that TDQA has better performance than DQA.

We set $\epsilon_{outer}=1.4\times10^{-4}$, λ^0=1000, w_δ^0=1500, $w_{v7,ADG1}^0=30$, $w_{v16,ADG2}^0=10$, and evaluate the convergence behaviors with respect to the step size Γ. Parameter Γ reflects the level of dependency of the target-response variables in each iteration to their values obtained in the previous iteration. We vary Γ in the range of {0.6, 0.7 …, 0.8, 0.99}. Figure 10(b) shows that, in general, increasing Γ decreases the number of iterations and computation time. For DQA and TDQA, the least number of iterations is obtained by setting Γ= 0.9. DQA has more stability to variation of Γ. This is because of the existence of the inner loop in which the algorithm seeks to reduce the error between the target and response values without updating the penalty multipliers.



(a)



(b)

Fig. 10. a) Iteration and *rel* index vs. stopping criterion, and b) iterations vs. Γ.

Setting different initial values for the penalty multipliers changes the speed of the algorithms and accuracy of the obtained results. We select different initial values for multipliers λ and w and the step size Γ and calculate the rel index. Figure 11 shows a contour plot of the rel index versus variations of initial values of λ, w, and Γ. If λ^0=1000 and w_δ^0=1000, setting the step size Γ to 0.6 provides the least error ($rel = 2.48 \times 10^{-4}$) after and 51 iterations. If the user selects the parameter badly (e.g., λ^0=2500, w_δ^0=2500), Γ equal to 0.5 yields the relative error of 0.0054 within 75iterations. Note that although we get the least rel with $\Gamma = 0.6$, λ^0=1000, and w_δ^0=1000, it takes a relatively long time to converge.

Case 4: We consider a full ACOPF and implement the proposed algorithm and three other methods, namely ADMM[20], APP[22], and ALAD (that is based on ATC)[33]. We compare the TDQA-based TSO+DSO optimal power flow to OPF solved by the other three methods [3]. Although all these four methods are based on the augmented Lagrangian relaxation, TDQA and APP solve the problem in a fully parallel manner while ADMM and ALAD are sequential solution algorithms. This means that in iteration k of ADMM and ALAD, DSOs cannot solve their subproblems without having the TSO's shared variable values in iteration k, whereas in TDQA and APP, TSO and DSOs solve their subproblems in parallel as they need their neighbors' shared variable values determined in iteration $k-1$. Each iteration of TDQA and APP takes around 0.1 seconds while it is 0.16 seconds for ADMM and ALAD. That is, each iteration of TDQA (and APP) is generally faster than that in ALAD and ADMM. Figure 12 shows the rel index after 100 iterations. All four methods converge to acceptable rel indices. APP takes many iterations for the rel index to go below an acceptable threshold, whereas the other three methods reach to an acceptable threshold after the first few iterations. Since ALAD and ADMM are sequential algorithms and TDQA is a parallel one, the solution time of TDQA, in each iteration, is faster than that for ALAD and ADMM. In the simulations, we have tried to provide reasonable and fair conditions to compare the four algorithms. Although the TDQA algorithm shows a good performance for the considered TSO+DSO OPF in this paper, we cannot make a solid conclusion that which algorithm has a better overall performance. The performance of the algorithms depends on the type of the problem and initial values of variables and penalty multipliers/factors (refer to [3] for more details).

It should be noted that the considered OPF problem in this paper has two levels. For problems with several levels of hierarchy, for instance, if TSO (level 1), DSOs (level 2), and microgrids (level 3) want to solve a collaborative OPF, the parallel solution algorithm, such as TDQA, is expected to be faster than the sequential one..

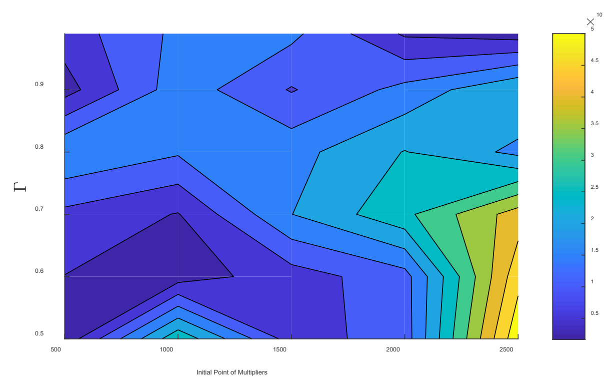


Fig. 11. Convergence measure rel versus initial values of multipliers and Γ.

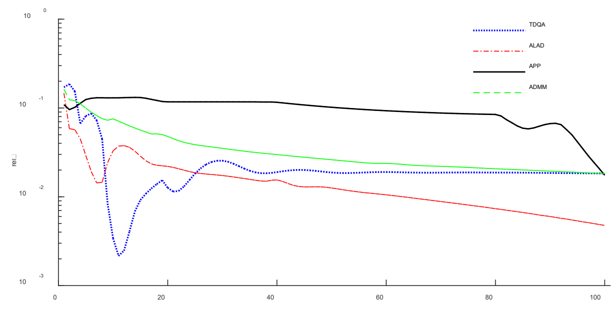


Fig. 12. Comparison between TDQA, APP, ADMM, and ALAD.

B. *IEEE 118-bus system*

Cold start: This test system comprises 31 autonomous systems: one TSO and 30 DSOs. The transmission system includes 118 buses, 186 lines, and 54 generators. Each active distribution grid consists of two generating units. We use a cold start, i.e., the initial values of the local variables of TSO and DSOs as well as targets/responses (i.e., shared variables) are set to zero. The initial values for penalty multipliers are selected as λ^0 =100, w^0=100, Γ=0.6, and the stopping thresholds are $\epsilon_{inner} = 0.2$ and $\epsilon_{outer} =$ 0.002. As shown in Table III, the collaborative ATC-based TSO+DSO operation with the DQA coordination strategy converges after 66 outer loop and 122 inner loop iterations (around 2 seconds). If we use the TDQA coordination strategy, the algorithm takes 99 iterations (only outer loop) to converge after less than 2 seconds). The overall operating cost of the system determined by DQA and DTQA is \$5,115.6 and \$5,120.6, respectively. In order to evaluate accuracy of the results, we ignore autonomy and information privacy of TSO and DSOs and solve the centralized OPF. The overall operating cost is \$5,098.6. Although error of DQA and TDQA is negligible, DQA is slightly more accurate.

Hot start: Since in practice we can use a hot start scenario, we are potentially capable of speeding up the convergence process. We choose appropriate initial points and run a hot start scenario. The results show that TDQA and DQA respectively take 11 and nine iterations to converge, which is much faster (almost ten times) than the cold start scenario. Figure 13 shows the rel index over the course of iterations.

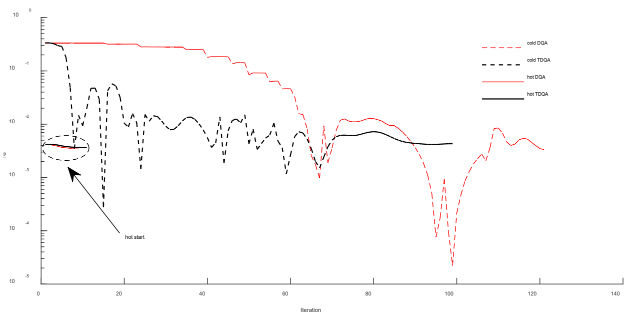


Fig. 13. rel index for the IEEE 118-bus system (cold and hot start scenarios).

TABLE III. TSO+DSO OPF FOR THE IEEE 118-BUS

Algorithm		Total cost	Iteration
Centralized TSO+DSO OPF		\$5,098.6	-
Decentralized TSO+DSO OPF *cold start*	DQA	\$5,116.5	92
	TDQA	\$5,120.6	98
Decentralized TSO+DSO OPF *hot start*	DQA	\$5,117.4	9
	TDQA	\$5,116.8	11

V. CONCLUSION

Power systems are being transformed to distributed energy infrastructures in which electricity is generated in both transmission and distribution levels. This paper presents a decentralized OPF algorithm for collaborative management of transmission and distribution systems. The proposed algorithm is based on analytical target cascading (ATC) for multilevel hierarchical optimization. The OPF problem associated with TSO is formulated in the upper level of the hierarchy while OPFs of DSOs are assigned to the lower level. As TSO and DSOs are autonomous systems, the information privacy plays a critical role in their joint management decisions. In the proposed framework, TSO and DSOs exchange only limited information namely target and response (shared) variables, and they are not required to reveal their commercially sensitive information to other parties. Two coordination strategies namely DQA and TDQA are presented to coordinate TSO and DSOs in a decentralized manner. DQA and TDQA allow parallel execution of local OPF problems (associated with TSO and DSOs).

The numerical tests on a 6-bus and the IEEE 118-bus systems show the accuracy and convergence performance of our proposed ATC-based collaborative TSO+DSO OPF method. Although both DQA and TDQA coordination strategies provide promising results, the collaborative TSO+DSO operation with TDQA usually determines optimal results with almost the same accuracy as DQA while usually taking less iterations. Using a hot start scenario to initialize the target/response pairs and multipliers significantly enhances the solution speed of the proposed TSO+DSO operation algorithm, especially for the large-scale system.

REFERENCE

[1] S. Chowdhury and P. Crossley, *Microgrids and active distribution networks*: The Institution of Engineering and Technology, 2009.

[2] A. Khodaei and M. Shahidehpour, "Microgrid-based co-optimization of generation and transmission planning in power systems," *IEEE transactions on power systems,* vol. 28, pp. 1582-1590, 2013.

[3] A. Kargarian, J. Mohammadi, J. Guo, S. Chakrabarti, M. Barati, G. Hug*, et al.*, "Toward distributed/decentralized DC optimal power flow implementation in future electric power systems," *IEEE Transactions on Smart Grid,* 2017.

[4] Y. Wang, S. Wang, and L. Wu, "Distributed optimization approaches for emerging power systems operation: A review," *Electric Power Systems Research,* vol. 144, pp. 127-135, 2017.

[5] D. K. Molzahn, F. Dörfler, H. Sandberg, S. H. Low, S. Chakrabarti, R. Baldick*, et al.*, "A survey of distributed optimization and control algorithms for electric power systems," *IEEE Transactions on Smart Grid,* 2017.

[6] S. Kar, G. Hug, J. Mohammadi, and J. M. Moura, "Distributed State Estimation and Energy Management in Smart Grids: A Consensus ${+} $ Innovations Approach," *IEEE Journal of Selected Topics in Signal Processing,* vol. 8, pp. 1022-1038, 2014.

[7] L. Kristov, P. De Martini, and J. D. Taft, "A Tale of Two Visions: Designing a Decentralized Transactive Electric System," *IEEE Power and Energy Magazine,* vol. 14, pp. 63-69, 2016.

[8] A. Ferrante, N. Constantinescu, and J. A. Jackson, "Lines of Convergence: R&D for Transmission and Distribution: Coordination and the Regulatory Challenge," *IEEE Power and Energy Magazine,* vol. 13, pp. 52-59, 2015.

[9] D. K. Molzahn, F. Dörfler, H. Sandberg, S. H. Low, S. Chakrabarti, R. Baldick*, et al.*, "A survey of distributed optimization and control algorithms for electric power systems," *IEEE Transactions on Smart Grid,* vol. 8, pp. 2941-2962, 2017.

[10] B. H. Kim and R. Baldick, "A comparison of distributed optimal power flow algorithms," *IEEE Transactions on Power Systems,* vol. 15, pp. 599-604, 2000.

[11] A. G. Bakirtzis and P. N. Biskas, "A decentralized solution to the DC-OPF of interconnected power systems," *IEEE Transactions on Power Systems,* vol. 18, pp. 1007-1013, 2003.

[12] M. D. Ilic, L. Xie, and J.-Y. Joo, "Efficient coordination of wind power and price-responsive demand—Part I: Theoretical foundations," *IEEE Transactions on Power Systems,* vol. 26, pp. 1875-1884, 2011.

[13] J. Mohammadi, G. Hug, and S. Kar, "Role of communication on the convergence rate of fully distributed DC optimal power flow," in *Smart Grid Communications (SmartGridComm), 2014 IEEE International Conference on*, 2014, pp. 43-48.

[14] G. Hug-Glanzmann and G. Andersson, "Decentralized optimal power flow control for overlapping areas in power systems," *IEEE Transactions on Power Systems,* vol. 24, pp. 327-336, 2009.

[15] M. H. Amini, B. Nabi, and M.-R. Haghifam, "Load management using multi-agent systems in smart distribution network," in *Power and Energy Society General Meeting (PES), 2013 IEEE*, 2013, pp. 1-5.

[16] S. D. Manshadi and M. E. Khodayar, "A hierarchical electricity market structure for the smart grid paradigm," *IEEE Transactions on Smart Grid,* vol. 7, pp. 1866-1875, 2016.

[17] A. R. Malekpour, A. Pahwa, and B. Natarajan, "Hierarchical architecture for integration of rooftop PV in smart distribution systems," *IEEE Transactions on Smart Grid,* 2016.

[18] A. Kargarian, M. Mehrtash, and B. Falahati, "Decentralized Implementation of Unit Commitment with Analytical Target Cascading: A Parallel Approach," *IEEE Transactions on Power Systems,* 2017.

[19] Z. Yuan and M. R. Hesamzadeh, "Hierarchical coordination of TSO-DSO economic dispatch considering large-scale integration of distributed energy resources," *Applied Energy,* vol. 195, pp. 600-615, 2017.

[20] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning,* vol. 3, pp. 1-122, 2011.

[21] Z. Li, Q. Guo, H. Sun, and J. Wang, "A New LMP-Sensitivity-Based Heterogeneous Decomposition for Transmission and Distribution Coordinated Economic Dispatch," *IEEE Transactions on Smart Grid,* 2017.

[22] B. H. Kim and R. Baldick, "Coarse-grained distributed optimal power flow," *IEEE Transactions on Power Systems,* vol. 12, pp. 932-939, 1997.

[23] G. Cohen, "Auxiliary problem principle and decomposition of optimization problems," *Journal of optimization Theory and Applications,* vol. 32, pp. 277-305, 1980.

[24] J. Guo, G. Hug, and O. K. Tonguz, "Intelligent partitioning in distributed optimization of electric power systems," *IEEE Transactions on Smart Grid,* vol. 7, pp. 1249-1258, 2016.

[25] J. Mohammadi, S. Kar, and G. Hug, "Distributed approach for DC optimal power flow calculations," *arXiv preprint arXiv:1410.4236,* 2014.

[26] S. Chakrabarti, M. Kraning, E. Chu, R. Baldick, and S. Boyd, "Security constrained optimal power flow via proximal message passing," in *Power Systems Conference (PSC), 2014 Clemson University*, 2014, pp. 1-8.

[27] A. Kargarian, Y. Fu, and H. Wu, "Chance-Constrained System of Systems Based Operation of Power Systems," *IEEE Transactions on Power Systems,* vol. 31, pp. 3404-3413, 2016.

[28] A. Kargarian and Y. Fu, "System of systems based security-constrained unit commitment incorporating active distribution grids," *IEEE Transactions on Power Systems,* vol. 29, pp. 2489-2498, 2014.

[29] A. K. Marvasti, Y. Fu, S. DorMohammadi, and M. Rais-Rohani, "Optimal operation of active distribution grids: A system of systems framework," *IEEE Transactions on Smart Grid,* vol. 5, pp. 1228-1237, 2014.

[30] H. Sun, Q. Guo, B. Zhang, Y. Guo, Z. Li, and J. Wang, "Master–slave-splitting based distributed global power flow method for integrated transmission and distribution analysis," *IEEE Transactions on Smart Grid,* vol. 6, pp. 1484-1492, 2015.

[31] Z. Li, Q. Guo, H. Sun, and J. Wang, "Coordinated transmission and distribution AC optimal power flow," *IEEE Transactions on Smart Grid,* 2016.

[32] Z. Li, Q. Guo, H. Sun, and J. Wang, "Coordinated economic dispatch of coupled transmission and distribution systems using heterogeneous decomposition," *IEEE Transactions on Power Systems,* vol. 31, pp. 4817-4830, 2016.

[33] S. Tosserams, L. Etman, P. Papalambros, and J. Rooda, "An augmented Lagrangian relaxation for analytical target cascading using the alternating direction method of multipliers," *Structural and multidisciplinary optimization,* vol. 31, pp. 176-189, 2006.

[34] S. DorMohammadi and M. Rais-Rohani, "Analytical target cascading framework using the exponential method of multipliers," in *12th AIAA conference on aviation technology, integration and operations*, 2012.

[35] N. Michelena, H. Park, and P. Y. Papalambros, "Convergence properties of analytical target cascading," *AIAA journal,* vol. 41, pp. 897-905, 2003.

[36] S. DorMohammadi and M. Rais-Rohani, "Exponential penalty function formulation for multilevel optimization using the analytical target cascading framework," *Structural and Multidisciplinary Optimization,* vol. 47, pp. 599-612, 2013.

[37] D. P. Bertsekas, "Dynamic programming and suboptimal control: A survey from ADP to MPC," *European Journal of Control,* vol. 11, pp. 310-334, 2005.

[38] D. P. Bertsekas, *Nonlinear programming*: Athena scientific Belmont, 1999.

[39] A. Ruszczyński, "On convergence of an augmented Lagrangian decomposition method for sparse convex optimization," *Mathematics of Operations Research,* vol. 20, pp. 634-656, 1995.

[40] Y. Li, Z. Lu, and J. J. Michalek, "Diagonal quadratic approximation for parallelization of analytical target cascading," *Journal of Mechanical Design,* vol. 130, p. 051402, 2008.

[41] S. DorMohammadi and M. Rais-Rohani, "Comparison of alternative strategies for multilevel optimization of hierarchical systems," *Applied Mathematics,* vol. 3, pp. 1448-1462, 2012.

[42] J. Lavaei and S. H. Low, "Zero duality gap in optimal power flow problem," *IEEE Transactions on Power Systems,* vol. 27, pp. 92-107, 2012.

[43] B. Kocuk, S. S. Dey, and X. A. Sun, "Strong SOCP relaxations for the optimal power flow problem," *Operations Research,* vol. 64, pp. 1177-1196, 2016.

Ali Mohammadi (S'17) received his BS degree in Electrical Engineering (Telecommunication) from Shahid Bahonar university of Kerman, Kerman, Iran, 2012, and his MS degree in Electrical Engineering (Telecommunication) from Shiraz University of Technology, Shiraz, Iran, 2014. He is currently pursuing his Ph.D. degree at the Department of Electrical and Computer Engineering, Louisiana State University, LA, USA. His research interests include power system optimization, machine learning, game theory, and graph theory.

Mahdi Mehrtash (S'16) received the B.Sc. degree from Shahid Chamran University of Ahvaz, Ahvaz, Iran, in 2012, and the M.Sc. degree from Shiraz University, Shiraz, Iran, in 2014, both in electrical engineering (power) with first-rank honors. He is currently working toward the Ph.D. degree at the Department of Electrical and Computer Engineering, Louisiana State University, Baton Rouge, LA, USA. His research interests include power system optimization, stochastic optimization, and electricity market.

Amin Kargarian (S'10–M'14) received the B.S. degree from the University of Isfahan, Isfahan, Iran, in 2007, the M.S. degree from Shiraz University, Shiraz, Iran, in 2010, and the Ph.D. degree fromMississippi State University, Starkville, MS, USA, in 2014, all in electrical and computer engineering. In August 2014, he joined the ECE Department, Carnegie Mellon University, as a Postdoctoral Research Associate. He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, Louisiana State University, Baton Rouge, LA, USA. His research interests include power system optimization and economics.