This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2937513, IEEE Transactions on Industrial Informatics

A Distributed Dual Consensus ADMM Based on Partition for DC-DOPF with Carbon Emission Trading

Linfeng Yang, *Member, IEEE*, Jiangyao Luo, Yan Xu, *Senior Member, IEEE*, Zhenrong Zhang, Zhaoyang Dong, *Fellow, IEEE*

***Abstract*—This paper presents a distributed alternating direction method of multipliers (ADMM) approach for solving the direct current dynamic optimal power flow with carbon emission trading (DC-DOPF-CET) problem. Generally, the ADMM-based distributed approaches disclose boundary buses and branches information among adjacent subsystems. As opposed to these methods, the proposed method (DC-ADMM-P) adopts a novel strategy which uses consensus ADMM to solve the dual of DC-DOPF-CET while only disclose boundary branches information among adjacent subsystems. Moreover, the convergence performance of DC-ADMM-P is improved by reducing the number of dual multipliers and employing an improved update step of the multiplier. DC-ADMM-P is tested on cases ranging from 6 to 1062 buses, with comparison with other distributed/decentralized methods. The simulation results verify the high efficiency of DC-ADMM-P in solving the DC-DOPF problem with complex (nonlinear) factors which can be formulated as convex separable functions. Meanwhile, it also shows the improvement of convergence performance by reducing the number of dual multipliers and employing a new update strategy for the multiplier.**

***Index Terms*—ADMM, dual, partition, distributed DC-DOPF, carbon emission trading.**

NOMENCLATURE

Operator:

$[\cdot]^-$	$\min(0,\cdot)$.
$[\cdot]^+$	$\max(0,\cdot)$.
$\|\cdot\|$	Number of elements in "·", if "·" is a set.

Indices:

i	Index for unit/bus.
t	Index for time period.

Constants:

S^{B}	Set of buses, $\|S^{\mathrm{B}}\| = N$ and $S^{\mathrm{B}} = \{1, \dots, N\}$.
S^{G}	Set of buses with units. $\|S^{\mathrm{G}}\| = G$ and $S^{\mathrm{G}} \subseteq S^{\mathrm{B}}$.
S^{TP}	Set of time periods, and $\|S^{\mathrm{TP}}\| = T$.
S^{L}	Set of lines, and $\|S^{\mathrm{L}}\| = L$.
S^{ref}	Set of reference bus.
$\alpha_i, \beta_i, \gamma_i$	Coefficients of the quadratic production cost function of unit i.
a_i, b_i, c_i	Coefficients of the quadratic emission function of unit i.
$\pi_{\mathrm{b}}, \pi_{\mathrm{s}}$	Price of emission allowances bought and sold on the market.
$F_{ij}^{\max}$	Maximum power flow limit on line ij.
X_{ij}	Reactance of line ij.
$\underline{P}_i^{\mathrm{G}}$	Minimum power output of unit i.
$\bar{P}_i^{\mathrm{G}}$	Maximum power output of unit i.
$P_{i,t}^{\mathrm{D}}$	System load demand on bus i in time period t.
$P_i^{\mathrm{G,up}}$	Ramp up limit of unit i.
$P_i^{\mathrm{G,down}}$	Ramp down limit of unit i.
$\boldsymbol{e}$	Vector with all components one.
$\mathbf{0}$	Matrix with all components zero.
$\boldsymbol{B}$	Network admittance matrix $(\Re^{N\times N})$.
B_{ij}	Element of $\boldsymbol{B}$ in the i^{th} row and j^{th} col.
E_0	The initial emission quota of CO_2.
$\Delta E_{\mathrm{b}}^{\max}$	The maximum emission allowances which can be bought on the market.
$\Delta E_{\mathrm{s}}^{\max}$	The maximum emission allowances which can be sold on the market.
μ	Coefficients of the update step for multiplier.

Variables:

$P_{i,t}^{\mathrm{G}}$	Power output of unit on bus i in time period t.
$\theta_{i,t}$	Phase angle on bus i in time period t.
$\Delta E_{\mathrm{b}}, \Delta E_{\mathrm{s}}$	Capacity of emission allowances bought and sold on the market.

I. INTRODUCTION

Optimal power flow (OPF) which aims to minimize the generation cost subject to some constraints, e.g., bus voltage limits, unit constraints and bus power limits, is one of the fundamental problems in power systems. The idea of using a unified model to resolve the aforementioned concerns was introduced by Carpentier in 1962 [1]. Under some assumptions such as small voltage phase angles, alternating current optimal power flow (AC-OPF) is analogous to direct current optimal power flow (DC-OPF) for obtaining the optimal dispatch solution of the entire power system.

A vast amount of researches have been conducted on OPF solution methods in the power system with one conventional centralized control entity [2]-[4]. For the centralized OPF, all data must be storaged and processed in the central controller. In actuality, the distribution region in a power network may belong to different owners, who might refuse to share some data due to security concerns. The distributed OPF (D-OPF) exactly guarantees the privacy of information, since adjacent subsystems disclose only a limited amount of information relating to boundary branches or boundary buses [5], [6]. Thus, individual subsystems do not need to disclose their confidential information/parameters to other subsystems [7]. Although the dis-

This work was supported by the Natural Science Foundation of China (51767003, 11771383, 61661004).

L.F. Yang and Z.R. Zhang are with the School of Computer Electronics and Information, Guangxi University, Nanning 530004, China, and also with the Guangxi Key Laboratory of Multimedia Communication and Network Technology, Guangxi University. (e-mail: ylf@gxu.edu.cn, zzr76@gxu.edu.cn).

J.Y. Luo is with the School of Computer Electronics and Information, Guangxi University, Nanning 530004, China. (e-mail: landiljy@163.com)

Y. Xu is with School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 (e-mail: yan.xu@sydney.edu.au).

Z.Y. Dong is with the School of Electrical Engineering and Tele-communications, The University of NSW, Sydney, NSW 2052, Australia. (e-mail: zydong@ieee.org).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2937513, IEEE Transactions on Industrial Informatics

tributed approach is often computationally more expensive than the centralized approach, but the distributed approach enables the independent operation of each interconnected subsystem with achieving the global optimum. In addition, some distributed OPFs can be solved asynchronously via individual local controllers [8].

Over the past few years, a vast amount of researches have been explored in D-OPF methods [5]-[21]. These methods can be broadly classified into the following categories [6]: analytical target cascading (ATC) [9], auxiliary problem principle (APP) [10], optimality condition decomposition (OCD) [11], consensus+innovations (C+I) [12], methods based on alternating direction method of multipliers (ADMM) [5], [7], [13]-[19], and other methods that do not fall into any of the aforementioned categories [20]. ATC, ADMM, and APP are well-suited for area-based OPF rather than nodal OPF [6], and they have similar implementation and performance. However, ATC is a bit more complex than ADMM in implementation due to the hierarchical decision-making in ATC. For APP, it needs a large number of iterations to converge in solving the nodal OPF. For OCD, the convergence conditions of OCD are usually hard to check before solving the optimization problem. The C+I requires all subproblems to reach consensus on a common variable. Hence, compared to Lagrangian-based methods such as ADMM and ATC, C+I takes more iterations.

The ADMM is well suited to distributed convex optimization, especially for the large-scale problems such as OPF [22], [23]. This method was developed in the 1970s, with roots in the 1950s [24]. It can be viewed as an attempt to blend the benefits of dual decomposition and augmented Lagrangian methods for constrained optimization. It is interesting to note that with no tuning, ADMM can be competitive with the best-known methods for some problems. In [13], the general form of ADMM was introduced and extended form the 2-block ADMM to an *N*-block ADMM. Moreover, this paper focused on two distributed parallel ADMM-based optimization algorithms: Consensus ADMM (C-ADMM) and Jacobian Proximal ADMM (JP-ADMM). The numerical results of C-ADMM and JP-ADMM in DC-OPF show that ADMM is effective as distributed parallel algorithm for power networks. A fully distributed consensus-based ADMM without the central controller is proposed in [7] for DC-OPF with Demand Response (DR). This method employs a new communication strategy, which is coordinated via global consensus variables, i.e., phase angles on boundary buses of adjacent subsystem. In [14], a D-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. An ADMM-based algorithm has been applied to OPF in [15]. 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 exchanged between neighboring nodes and devices. References [5] and [16] present decentralized OPF algorithms using ADMM, where subproblems directly communicate with each other without the central coordinator.

In general, dividing the whole OPF problem to region-based or component-based sub-problems is a common scheme for ADMM-based distributed methods. At each iteration, each sub-problem (i.e., the corresponding decomposed augmented function for each subsystem) is solved only with respect to a subset of primal variables over which it is decomposed. Besides the information (such as admittances of buses and branches) belonging to its own subsystem, this sub-problem always needs some information from its adjacent subsystems. This additional information is regard as "public information". As far as knows, most of the distributed methods including the existing ADMM-based OPF methods, use boundary bus and branch information among all adjacent subsystems from public information.

As opposed to other popular ADMM-based methods, DC-ADMM-P can reduce the amount of public information by employing ADMM to solve the dual problem. In addition, operating at an absolute minimal cost is no longer the only condition for electric power generation due to the pressing public demand for cleaner air [25]. In [26] it has been summed that there are two main categories of emission functions. One is the second order polynomial function which is the type used in our paper [27]. According to the Kyoto Protocol, the emissions rights of greenhouse gas have the property of commodities, which can be traded in commodity trade. Therefore, we have introduced a simple carbon emission trading (CET), which facilitates the reduction in greenhouse gas emission while optimizing the daily operation cost of electric power systems [28]. Theoretically, our method can effectively solve DC-DOPF problem with complex constraints (i.e., CET, DR), as long as these constraints are convex and separable.

The major contributions that differentiate our method from other existing methods are the following:

1) The DC-ADMM-P only needs public boundary branches information disclosed among adjacent subsystems while most other ADMM-based methods also need private boundary buses information disclosed among adjacent subsystems. Therefore, the volume of public information of DC-ADMM-P is much smaller.

2) In DC-ADMM-P, the number of elements in the consensus variable is only determined by the number of "real" coupling constraints in the DC-DOPF-CET problem. Moreover, the convergence speed of the DC-ADMM-P is improved by identifying the uncoupled buses and reducing the number of dual multipliers.

3) A new update strategy for a multiplier is employed to reduce the iteration number of our method further without increasing communication time.

4) Theoretically, DC-ADMM-P can arbitrarily divide DC-DOPF-CET into multiple subsystems or even a series of single buses by setting the partition set. In addition, the right-hand side (RHS, i.e., b_{E}, b_{BR}, and b_{PF}) of all "real" coupling constraints aren't necessarily have to be assigned to its corresponding subsystems. They can be assigned to any subsystem according to arbitrary partitioning strategies.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2937513, IEEE Transactions on Industrial Informatics

The remainder of this paper is arranged as follows: Section II presents the mathematical formulation of the DC-DOPF-CET. The DC-ADMM-P and the presentation of reducing the number of dual multipliers, public information, and the improved update strategy of multipliers are given in Section III. In Section IV, we apply the DC-ADMM-P on a various of numerical case studies, and compare their key features and simulation results with the methods in [6] and [7]. Finally, a summary is provided in Section V.

II. Mathematical Formulation of DC-DOPF-CET

In smart electric power grids, the system's physical structure is becoming more distributed. In such infrastructures, different entities, which are physically interconnected via transmission lines, take the control responsibility of different parts of the system. This means that, in entire transmission network, the entities need to collaborate with each other to achieve effective and reliable operation of the entire grid. The DC-DOPF-CET problem can be used to optimize the day-ahead hourly schedule of generators and trade of emission quota for minimizing the total operation cost.

Consider a power network containing several generators ($|S^{\mathrm{G}}|$ generators in this paper) and known loads. The network has finite numbers of buses belonging to S^{B} and finite numbers of branches (edges) belonging to S^{L}. Then, the numbers of nodes and branches are $|S^{\mathrm{B}}|$ and $|S^{\mathrm{L}}|$ respectively. A typical formulation of DC-DOPF-CET can be presented below.

A. Objective Function

The objective of operational planning for DC-DOPF-CET is to minimize the total operation cost F_{TC}

$$F_{\mathrm{TC}} = F_{\mathrm{TH}} + F_{\mathrm{CET}} \tag{1}$$

where F_{TH} is the fuel cost of thermal units and F_{CET} is the cost of CET.

1) Fuel cost of thermal units (F_{TH})

The fuel cost of thermal units is shown as

$$F_{\mathrm{TH}} = \sum_{i\in S^{\mathrm{G}}}\sum_{t\in S^{\mathrm{TP}}} f_i(P_{i,t}^{\mathrm{G}}) \tag{2}$$

where S^{TP} is the set of time periods and S^{G} is the set of buses with units; $P_{i,t}^{\mathrm{G}}$ is the power output on bus i in time period t. The production cost for unit i can be represented as a quadratic function $f_i(\cdot)$, shown as below

$$f_i(P_{i,t}^{\mathrm{G}}) = \alpha_i + \beta_i P_{i,t}^{\mathrm{G}} + \gamma_i \left(P_{i,t}^{\mathrm{G}}\right)^2 \tag{3}$$

where α_i, β_i, and γ_i are the coefficients of the quadratic production cost function of unit i.

2) Cost of CET (F_{CET})

A CET system is a cap-and-trade market mechanism that has been introduced in the European Union aiming to facilitate emissions management. In this scheme, a certain amount of allowances, which is also called quota, is usually given in advance. This quota is usually used to cover emissions produced during energy generation [34]. Under CET, emission quota is thought as goods, then the power producer can buy and sell emission quotas on the market [35], [36]. In this paper, it is assumed that the emission quota for DC-DOPF-CET is E_0. The actual total emission may be higher or lower than the cap because of the quota trades. Referring to [34], the cost caused in this trading is

$$F_{\mathrm{CET}} = \pi_{\mathrm{b}}\Delta E_{\mathrm{b}} - \pi_{\mathrm{s}}\Delta E_{\mathrm{s}} \tag{4}$$

where π_{b} and π_{s} are the price of emission allowances bought and sold on the market. ΔE_{b} and ΔE_{s} are the capacity of emission allowances bought and sold on the market.

B. Constraints

The constraints of DC-DOPF-CET include operation constraints of thermal units, network constraints, and CET constraints, etc. The detailed constraints are presented as follows.

1) Nodal power balance restriction

$$P_{i,t}^{\mathrm{G}} - P_{i,t}^{\mathrm{D}} = \sum_{j\in S^{\mathrm{B}}} B_{ij}\theta_{j,t}, \forall t \in S^{\mathrm{TP}}, \forall i \in S^{\mathrm{B}} \tag{5}$$

where $P_{i,t}^{\mathrm{D}}$ is the system load demand on bus i in time period t; B_{ij} is the element of $\boldsymbol{B}$ (network admittance matrix) in the i^{th} row and j^{th} column; $\theta_{j,t}$ is the phase angle on bus j in time period t; S^{B} is the set of buses.

2) Capacity limits of transmission lines

$$\left|\frac{\theta_{i,t}-\theta_{j,t}}{X_{ij}}\right| \le F_{ij}^{\max}, \forall t \in S^{\mathrm{TP}}, \forall ij \in S^{\mathrm{L}}, \forall i,j \in S^{\mathrm{B}} \tag{6}$$

where X_{ij} is the reactance of line ij; $F_{ij}^{\max}$ is the maximum power flow limit on line ij; S^{L} is the set of lines.

3) Reference bus

$$\theta_{i,t} = 0, \forall i \in S^{\mathrm{ref}}, \forall t \in S^{\mathrm{TP}} \tag{7}$$

where S^{ref} is the set of reference bus.

4) Buses without generators

$$P_{i,t}^{\mathrm{G}} = 0, \forall i \in S^{\mathrm{B}} - S^{\mathrm{G}}, \forall t \in S^{\mathrm{TP}} \tag{8}$$

5) Capacity limits of generators

$$\underline{P_i}^{\mathrm{G}} \le P_{i,t}^{\mathrm{G}} \le \bar{P}_i^{\mathrm{G}}, \forall i \in S^{\mathrm{G}}, \forall t \in S^{\mathrm{TP}} \tag{9}$$

where $\underline{P_i}^{\mathrm{G}}$ and $\bar{P}_i^{\mathrm{G}}$ are the minimum/maximum power output of unit i.

6) Ramp up limits of generators

$$P_{i,t}^{\mathrm{G}} - P_{i,t-1}^{\mathrm{G}} \le P_i^{\mathrm{G,up}}, \forall i \in S^{\mathrm{G}}, \forall t \in S^{\mathrm{TP}} \tag{10}$$

where $P_i^{\mathrm{G,up}}$ is the ramp up limit of unit i.

7) Ramp down limits of generators

$$P_{i,t-1}^{\mathrm{G}} - P_{i,t}^{\mathrm{G}} \le P_i^{\mathrm{G,down}}, \forall i \in S^{\mathrm{G}}, \forall t \in S^{\mathrm{TP}} \tag{11}$$

where $P_i^{\mathrm{G,down}}$ is the ramp down limit of unit i.

8) Capacity limit of carbon emission [28]

$$\sum_{i\in S^{\mathrm{G}}}\sum_{t\in S^{\mathrm{TP}}} E_i(P_{i,t}^{\mathrm{G}}) \le E_0 + \Delta E_{\mathrm{b}} - \Delta E_{\mathrm{s}} \tag{12}$$

where $E_i\left(P_{i,t}^{\mathrm{G}}\right) = a_i + b_i P_{i,t}^{\mathrm{G}} + c_i\left(P_{i,t}^{\mathrm{G}}\right)^2$ is the emission function; E_0 is the initial emission quota of CO_2; a_i, b_i, and c_i are the coefficients of the quadratic emission function of unit i.

9) Maximum emission allowance can be bought

$$0 \le \Delta E_{\mathrm{b}} \le \Delta E_{\mathrm{b}}^{\max} \tag{13}$$

where $\Delta E_{\mathrm{b}}^{\max}$ is the maximum emission allowances that can be bought on the market.

10) Maximum emission allowance can be sold

$$0 \le \Delta E_{\mathrm{s}} \le \Delta E_{\mathrm{s}}^{\max} \tag{14}$$

where $\Delta E_{\mathrm{s}}^{\max}$ is the maximum emission allowances that can be sold on the market.

III. Dual Consensus ADMM Based on Partition

As mentioned in the introduction of this paper, the distributed approach enables the independent operation of each interconnected subsystem, meanwhile the global optimum is

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2937513, IEEE Transactions on Industrial Informatics

achieved. In this section, the DC-DOPF-CET problem is re-written as a concise separable formulation primarily. Based on this formulation, DC-ADMM-P is introduced in following.

A. Description of the Model for DC-ADMM-P

In this paper, the following notations are used for DC-DOPF-CET problem. For $i \in S^{\mathrm{B}}$, let $x_i = [\boldsymbol{P}_i^{\mathrm{G}}; \boldsymbol{\theta}_i]$. Moreover, we have $\boldsymbol{P}_i^{\mathrm{G}} = \left[P_{i,1}^{\mathrm{G}}; P_{i,2}^{\mathrm{G}}; \dots; P_{i,T}^{\mathrm{G}}\right]$, $\boldsymbol{P}_i^{\mathrm{D}} = \left[P_{i,1}^{\mathrm{D}}; P_{i,2}^{\mathrm{D}}; \dots; P_{i,T}^{\mathrm{D}}\right]$, $\boldsymbol{P}^{\mathrm{D}} = [\boldsymbol{P}_1^{\mathrm{D}}; \boldsymbol{P}_2^{\mathrm{D}}; \dots; \boldsymbol{P}_N^{\mathrm{D}}]$, $\boldsymbol{\theta}_i = \left[\theta_{i,1}; \theta_{i,2}; \dots; \theta_{i,T}\right]$, and $x_0 = [\Delta E_{\mathrm{b}}; \Delta E_{\mathrm{s}}]$. As the dual consensus ADMM is a distributed algorithm with a central coordinator, $i = 0$ is introduced as a central coordinator in the following section, which only knows the initial emission quota of CO_2.

Then, the (5) and the (6) can be reformulated as

$$\boldsymbol{P}_i^{\mathrm{G}} - \boldsymbol{P}_i^{\mathrm{D}} = \sum_{j \in S^{\mathrm{B}}} B_{ij} \boldsymbol{\theta}_j, \forall i \in S^{\mathrm{B}} \tag{15}$$

$$-F_{ij}^{\max} \boldsymbol{e} \le \frac{\boldsymbol{\theta}_i - \boldsymbol{\theta}_j}{x_{ij}} \le F_{ij}^{\max} \boldsymbol{e}, \forall ij \in S^{\mathrm{L}},\ \forall i, j \in S^{\mathrm{B}}. \tag{16}$$

Let $b_{\mathrm{PF}} = \boldsymbol{P}^{\mathrm{D}}$, $\tilde{A}_i = [\boldsymbol{0}, -B_{1i}\boldsymbol{I}; \cdots; \boldsymbol{I}, -B_{ii}\boldsymbol{I}; \cdots;\ \boldsymbol{0}, -B_{Ni}\boldsymbol{I}]$, and $\tilde{A}_0 = \boldsymbol{0}$, (15) can be reformulated as

$$\sum_{i \in S^{\mathrm{B}} \cup \{0\}} \tilde{A}_i x_i = b_{\mathrm{PF}}. \tag{17}$$

Similarly, let $\widetilde{M}_0 = \boldsymbol{0}$, (16) can be reformulated as

$$\sum_{i \in S^{\mathrm{B}} \cup \{0\}} \widetilde{M}_i x_i \le b_{\mathrm{BR}} \tag{18}$$

where $\widetilde{M}_i$ is a coefficient matrix for x_i, and b_{BR} is a constant vector.

Let $\tilde{h}_0(x_0) = [-1, 1]x_0$, (12) can be reformulated as

$$\sum_{i \in S^{\mathrm{B}} \cup \{0\}} \tilde{h}_i(x_i) \le b_{\mathrm{E}} \tag{19}$$

where $b_{\mathrm{E}} = E_0$; $\tilde{h}_i(x_i) = \sum_{t \in S^{\mathrm{TP}}} E_i\left(P_{i,t}^{\mathrm{G}}\right)$, $i \in S^{\mathrm{G}}$; $\tilde{h}_i(x_i) = 0$, $i \in S^{\mathrm{B}} - S^{\mathrm{G}}$.

It can be seen that the constraints (7) to (11) are fully decoupled for each unit. So we define set $\chi_i = \{x_i | (7) - (11), i \in S^{\mathrm{B}}\}$, and denote the fuel cost of thermal units by $g_i(x_i) = 0, i \in S^{\mathrm{B}} - S^{\mathrm{G}}$ and $g_i(x_i) = \sum_{t \in S^{\mathrm{TP}}} f_i\left(P_{i,t}^{\mathrm{G}}\right), i \in S^{\mathrm{G}}$. The CET can be rewritten as $g_0(x_0) = [\pi_{\mathrm{b}}; -\pi_{\mathrm{s}}]x_0$ and $\chi_0 = \{x_0 | (13)\ (14)\}$.

In order to further simplification, let $\psi_0(x_0) = \left[\tilde{h}_0(x_0) - b_{\mathrm{E}}; \widetilde{M}_0\right]$ and $A_0(x_0) = \tilde{A}_0$. For $i \in S^{\mathrm{B}}$, let $A_i(x_i) = \left[\tilde{A}_i x_i - b_{\mathrm{PF}}^i\right]$ and $\psi_i(x_i) = \left[\tilde{h}_i(x_i); \widetilde{M}_i x_i - b_{\mathrm{BR}}^i\right]$, where $\sum_{i \in S^{\mathrm{B}}} b_{\mathrm{BR}}^i = b_{\mathrm{BR}}$ and $\sum_{i \in S^{\mathrm{B}}} b_{\mathrm{PF}}^i = b_{\mathrm{PF}}$.

Then, the model of DC-DOPF-CET can be reformulated as

$$\min \sum_{i \in S^{\mathrm{B}} \cup \{0\}} g_i(x_i)$$
$$\text{s.t.} \begin{cases} \sum_{i \in S^{\mathrm{B}} \cup \{0\}} \psi_i(x_i) \le 0 \\ \sum_{i \in S^{\mathrm{B}} \cup \{0\}} A_i(x_i) = 0 \\ x_i \in \chi_i, i \in \{0\} \cup S^{\mathrm{B}}. \end{cases} \tag{20}$$

Different with the traditional DC-DOPF, which is a convex quadratic programming (QP) problem, the DC-DOPF-CET is a convex quadratically constrained quadratic programming (QCQP) problem. Theoretically, all RHS (i.e., b_{E}, b_{BR}, and b_{PF}) can be arbitrary assigned to one or more subsystems in $\{0\} \cup S^{\mathrm{B}}$ for privacy protection. Furthermore, any subsystem in S^{B} can be viewed as a central coordinator. In this case, $\{0\}$ will not be included in problem (20).

B. Methodology

Let $|\pi| = n + 1$ and $\pi = \{\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n\}$ is a partition of the set $S^{\mathrm{B}} \cup \{0\}$, i.e., $\emptyset \notin \pi$, $\cup_{\mathcal{A}_j \in \pi, j=0,\cdots,n} \mathcal{A}_j = S^{\mathrm{B}} \cup \{0\}$, and $(\forall \mathcal{A}, \mathcal{B} \in \pi)\mathcal{A} \ne \mathcal{B} \Rightarrow \mathcal{A} \cap \mathcal{B} = \emptyset$.

Because problem (20) is a convex QCQP problem, according to Slater constraint specification [23], the duality gap between (20) and its dual problem is zero. So we can obtain the solutions of the problem (20) by solving its dual problem.

Then, the partition-based Lagrangian dual function of problem (20) is

$$L(\lambda; v) = \sum_{j=0}^{n} L_j(\lambda; v) \tag{21}$$

where $L_j(\lambda; v) = \inf_{(x_i \in \chi_i, i \in \mathcal{A}_j)} \left\{\sum_{i \in \mathcal{A}_j} g_i(x_i) + \left\langle \sum_{i \in \mathcal{A}_j} \psi_i(x_i), \lambda \right\rangle + \left\langle \sum_{i \in \mathcal{A}_j} A_i(x_i), v \right\rangle\right\}$, $\lambda \ge 0$, and v are Lagrange multiplier vectors corresponding to $\sum_{i \in \mathcal{A}_j} \psi_i(x_i)$ and $\sum_{i \in \mathcal{A}_j} A_i(x_i)$, and "$\langle \cdot, \cdot \rangle$" denotes the inner product.

Let $y = (\lambda; v)$. Then, the dual problem of problem (20) is

$$\max L(y). \tag{22}$$

This problem can be rewritten as a minimization problem with local variables z_i and a common consensus variable y [37]:

$$\min\left\{-\sum_{j=0}^{n} L_j(z_j)\right\}$$
$$\text{s.t.}\ y - z_j = 0, j = 0, \dots, n. \tag{23}$$

Based on partitioning the multiplier vector $p = (p_0, \cdots, p_n)$ and giving penalty parameter $\rho > 0$, solving problem (23) employing ADMM consists of the following iterations [22]:

$$y^{k+1} := \underset{y}{\operatorname{argmin}} \left\{\left\langle \sum_{j=0}^{n} p_j^k, y \right\rangle + \frac{\rho}{2} \sum_{j=0}^{n} \left\| y - z_j^k \right\|^2\right\} \tag{24}$$

$$z_j^{k+1} := \underset{z_j}{\operatorname{argmin}} \left\{-L_j(z_j) - \left\langle p_j^k, z_j \right\rangle + \frac{\rho}{2} \left\| y^{k+1} - z_j \right\|^2\right\} \tag{25}$$

$$p_j^{k+1} := p_j^k + \rho\left(y^{k+1} - z_j^{k+1}\right). \tag{26}$$

The y-update problem in problem (24) is minimizing a convex quadratic function of y, then we can obtain the minimum of y from the optimality condition, i.e.,

$$y^{k+1} = \frac{1}{n+1} \sum_{j=0}^{n} z_j^k - \frac{1}{(n+1)\rho} \sum_{j=0}^{n} p_j^k. \tag{27}$$

By substituting the function on the right-hand side of $L_j(\lambda; v)$ in (21) into (25) to eliminate the $L_j(z_j)$, we can obtain:

$$\inf_{z_j = [\lambda_{z_j}; v_{z_j}], \lambda_{z_j} \ge 0} \left\{-\inf_{(x_i \in \chi_i, i \in \mathcal{A}_j)} \left\{\sum_{i \in \mathcal{A}_j} g_i(x_i) + \left\langle \lambda_{z_j}, \sum_{i \in \mathcal{A}_j} \psi_i(x_i) \right\rangle + \left\langle v_{z_j}, \sum_{i \in \mathcal{A}_j} A_i(x_i) \right\rangle\right\} - \left\langle p_j^k, z_j \right\rangle + \frac{\rho}{2} \left\| y^{k+1} - z_j \right\|^2\right\} =$$

$$\sup_{z_j = [\lambda_{z_j}; v_{z_j}], \lambda_{z_j} \ge 0} \inf_{(x_i \in \chi_i, i \in \mathcal{A}_j)} \left\{\sum_{i \in \mathcal{A}_j} g_i(x_i) + \left\langle z_j, \left(\sum_{i \in \mathcal{A}_j} \psi_i(x_i); \sum_{i \in \mathcal{A}_j} A_i(x_i)\right) + p_j^k \right\rangle - \frac{\rho}{2} \left\| y^{k+1} - z_j \right\|^2\right\} =$$

$$\inf_{(x_i \in \chi_i, i \in \mathcal{A}_j)} \sup_{z_j = [\lambda_{z_j}; v_{z_j}], \lambda_{z_j} \ge 0} \left\{\sum_{i \in \mathcal{A}_j} g_i(x_i) + \left\langle z_j, \left(\sum_{i \in \mathcal{A}_j} \psi_i(x_i); \sum_{i \in \mathcal{A}_j} A_i(x_i)\right) + p_j^k \right\rangle - \frac{\rho}{2} \left\| y^{k+1} - z_j \right\|^2\right\} \tag{28}$$

where λ_{z_j} is the sub vector including the elements that corresponds to λ in z_j; v_{z_j} can be defined similarly.

For any fixed $\{x_i \in \chi_i, i \in \mathcal{A}_j\}$, the minimum on the "sup.$\{\cdot\}$" in the right-hand side of problem (28) is uniquely attained by

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2937513, IEEE Transactions on Industrial Informatics

$$z_j = \begin{bmatrix} \lambda_{z_j} \\ v_{z_j} \end{bmatrix} = \begin{bmatrix} \max\left\{0, \lambda_{y^{k+1}} + \frac{1}{\rho}\left(\sum_{i\in\mathcal{A}_j}\psi_i(x_i) + \lambda_{p_j^k}\right)\right\} \\ v_{y^{k+1}} + \frac{1}{\rho}\left(\sum_{i\in\mathcal{A}_j} A_i(x_i) + v_{p_j^k}\right) \end{bmatrix} \quad (29)$$

where $\lambda_{p_j^k}$ and $v_{p_j^k}$ have the similar definitions as λ_{z_j}.

We may thus substitute (29) into the function on the right-hand side of (28) to eliminate the variables z_j. Then, we can obtain

$$\left\{x_i^*, i\in\mathcal{A}_j\right\} \coloneqq \operatorname{argmin}_{(x_i\in\chi_i, i\in\mathcal{A}_j)} \left\{ \sum_{i\in\mathcal{A}_j} g_i(x_i) + \frac{\rho}{2}\left\| \begin{bmatrix} \max\left\{0, \lambda_{y^{k+1}} + \frac{1}{\rho}\left(\sum_{i\in\mathcal{A}_j}\psi_i(x_i) + \lambda_{p_j^k}\right)\right\} \\ v_{y^{k+1}} + \frac{1}{\rho}\left(\sum_{i\in\mathcal{A}_j} A_i(x_i) + v_{p_j^k}\right) \end{bmatrix} \right\|^2 \right\}. \quad (30)$$

Therefore, $\left\{x_i^*, i\in\mathcal{A}_j\right\}$ is a solution of the following problem:

$$\min \sum_{i\in\mathcal{A}_j} g_i(x_i) + \frac{\rho}{2}\left\| \begin{bmatrix} \tau \\ v_{y^{k+1}} + \frac{1}{\rho}\left(\sum_{i\in\mathcal{A}_j} A_i(x_i) + v_{p_j^k}\right) \end{bmatrix} \right\|^2$$
$$\text{s.t.} \begin{cases} \tau \geq 0 \\ \tau \geq \left[\lambda_{y^{k+1}}\right] + \frac{1}{\rho}\left(\sum_{i\in\mathcal{A}_j}\psi_i(x_i) + \lambda_{p_j^k}\right) \\ x_i \in \chi_i, i\in\mathcal{A}_j \end{cases} \quad (31)$$

where τ is the auxiliary variable.

It is not difficult to verify that problem (31) is strictly convex; then, the unique solution $\left\{x_i^*, i\in\mathcal{A}_j\right\}$ can be found. Therefore, z_j can be determined by (29).

In summary, in this subsection, we apply the ADMM to solve the dual problem of DC-DOPF-CET based on partition. Mathematically, our method actually obtains the solutions for dual variables (i.e., Lagrange multipliers) and the optimal scheduling schemes (including unit power output $P_{i,t}^{\mathrm{G}}$, bus phase angle $\theta_{i,t}$, and line flow $\frac{\theta_{i,t}-\theta_{j,t}}{X_{ij}}$, etc.) of DC-DOPF-CET only through the products provided by solving problem (31). Thus, our method can achieve the locational marginal prices (LMPs) in distributed manner. Moreover, our method is also applicable when considering other complicated factors, including emission, demand response, etc., as long as these factors can be formulated as convex separable (nonlinear) functions [7][38][39]. However, this method is not applicable in AC-OPF problem resulting from non-convexity of AC power flow constraint. Although the AC power flow constraint can be relaxed to be convex semi-definite (SD) constraint or second-order cone (SOC) constraint [38], our method is also not applicable because the resulted SD and SOC constraints are not separable. The consensus-based ADMM with duplicating the coupling variables (C-ADMM-D) [7], ADMM with gaussian back substitution (ADMM-G) [30], traditional ADMM (T-ADMM) [22], and JP-ADMM [31] can't directly solve the problem with nonlinear constraints similar to CET. Then, (12) can be rewritten as

$$\sum_{t\in S^{\mathrm{TP}}} E_i\left(P_{i,t}^{\mathrm{G}}\right) \leq D_i, i\in S^{\mathrm{G}} \quad (32)$$
$$\sum_{i\in S^{\mathrm{G}}} D_i - \Delta E_{\mathrm{b}} + \Delta E_{\mathrm{s}} + H = E_0 \quad (33)$$
$$H \geq 0 \quad (34)$$

where D_i is the auxiliary variable, H is the relaxation variable.

After introducing many auxiliary and relaxation variables, lots of constraints, (32)-(34) is a suitable form for the T-ADMM, ADMM-G, and JP-ADMM, although it is clear that these methods will not be very efficient, especially when $|S^{\mathrm{G}}|$ is large. However, C-ADMM-D still cannot decouple (33) by duplicating the coupling variables in different subsystems. Moreover, ADMM-G ensure the convergence by adding the Gaussian back substitution, which causes the blocks to be updated one after another. Therefore, ADMM-G is not amenable for parallelization. JP-ADMM adds a proximal term $\frac{1}{2}\left\|x_i - x_i^k\right\|_{P_i}^2$ for each x_i-subproblem and an additional damping parameter $\delta > 0$ for the update of dual variable.

As the separability of problem (23), the updates of (29), and (26) can be performed in parallel for our partition $\pi = (\mathcal{A}_j, j = 0,1,\dots,n)$. Then, the details of the proposed preliminary DC-ADMM-P is shown as following.

Let M be the predefined maximum number of iterations. $\varepsilon^{\mathrm{feasible}}$ be a positive tolerance parameter. ρ be a predefined positive parameter. Let z_j^0 and p_j^0 be the initial values of z_j and p_j.

Algorithm 1: DC-ADMM-P

Initialization: $z_j^0, p_j^0, M > 0, \rho > 0, \varepsilon^{\mathrm{feasible}} > 0$;
for $k = 0,\dots,M$
  Obtain y^{k+1} according to (27);
  for each sub set of units $j = 0,\dots,n$: (in parallel)
    Obtain $\left\{x_i^*, i\in\mathcal{A}_j\right\}$ by solving (31);
    Obtain z_j^{k+1} according to (29);
    Obtain p_j^{k+1} according to (26);
  end
  if $\varepsilon \coloneqq \|y^{k+1} - y^k\|_\infty < \varepsilon^{\mathrm{feasible}}$ **break**;
end.

As shown in Fig. 1, $\mathcal{A}_0$ is viewed as a central coordinator. Each subsystem only communicates with the central coordinator and there are no direct communication links between each other. The information to be exchanged between the central coordinator and subsystems is y^{k+1}, p_j^{k+1}, and z_j^{k+1}. In this structure, p_j^{k+1} and z_j^{k+1} are sent from subsystems to the central coordinator while y^{k+1} is sent from the central coordinator to subsystems.

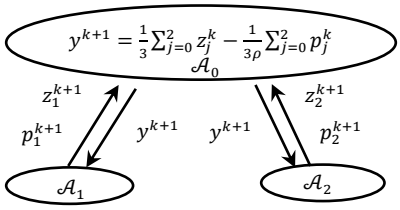

Fig. 1. COMMUNICATION STRATEGY FOR DC-ADMM-P.

In addition, the computing efficiency of algorithm 1 can be improved by reducing the number of dual multipliers and employing the new update step of multiplier. The improvements and discussion about the advantages of our method in disclosing less public information are shown in the following subsections.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2937513, IEEE Transactions on Industrial Informatics

C. Reducing the Number of Dual Multipliers

In the process of transforming the problem (20) into its dual problem, all constraints include $\sum_{i\in S^{B}\cup\{0\}}\psi_i(x_i)\le 0$ and $\sum_{i\in S^{B}\cup\{0\}}A_i(x_i)=0$ introduce their corresponding dual multipliers. However, only the dual multipliers introduced by "real" coupling constraints formed by the boundary buses and branches information among adjacent subsystems are necessary. Therefore, with the aim of obtaining better convergence in DC-ADMM-P, the number of dual multipliers should be minimum. Normally, the number of dual multipliers is the same as the number of elements in a consensus variable y of the problem (23).

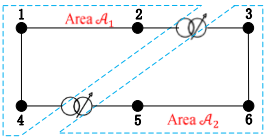

Fig. 2. THE 6-BUS SYSTEM AS AN ILLUSTRATIVE EXAMPLE.

A 6-bus system shown in Fig. 2 is used as an example to illustrate the procedure of identifying "real" coupling constraints. The 6-bus system is divided into three subsystems, i.e., $\pi=\{\{\mathcal{A}_0=\{0\},\mathcal{A}_1=\{1,2,4\},\mathcal{A}_2=\{3,5,6\}\}$. Buses 2 and 4 are called boundary buses of area $\mathcal{A}_1$, and bus 1 is the inner bus of area $\mathcal{A}_1$. The transmission line (2,3) connecting the bus 2 and bus 3 is called boundary branch between area $\mathcal{A}_1$ and $\mathcal{A}_2$. Similarly, transmission lines (1,2) and (1,4) are inner branches of $\mathcal{A}_1$. Then, the boundary branch information between area $\mathcal{A}_1$ and $\mathcal{A}_2$ is not private for area $\mathcal{A}_1$ and $\mathcal{A}_2$. But the boundary bus information is the private information to its corresponding area. Hence, we use "public boundary branches information" and "private boundary buses information" in the paper.

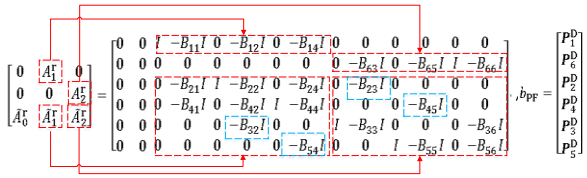

Fig. 3. $\sum_{i\in S^{B}\cup\{0\}}\tilde{\boldsymbol{A}}_i\boldsymbol{x}_i=\boldsymbol{b}_{\mathbf{PF}}$ FOR THE 6-BUS SYSTEM.

The Fig. 3 gives the coefficient matrix and RHS for $\sum_{i\in S^{B}\cup\{0\}}\tilde{A}_i x_i=b_{\mathrm{PF}}$. It can be seen that there is $v\in R^{6T}$ in consensus variable y. In other words, before identifying "real" coupling constraints, the coupling constraints for the DC nodal power balance equation are $[\mathbf{0},A_1^{\mathrm{r}},\mathbf{0};\mathbf{0},\mathbf{0},A_2^{\mathrm{r}};\tilde{A}_0^r,\tilde{A}_1^r,\tilde{A}_2^r][x_0;x_1;x_2;x_4;x_3;x_5;x_6]-[\boldsymbol{P}_1^{\mathrm{D}};\boldsymbol{P}_6^{\mathrm{D}};\boldsymbol{P}_2^{\mathrm{D}};\boldsymbol{P}_4^{\mathrm{D}};\boldsymbol{P}_3^{\mathrm{D}};\boldsymbol{P}_5^{\mathrm{D}}]=\mathbf{0}$. However, bus 1 in subsystem $\mathcal{A}_1$ has no association with subsystem $\mathcal{A}_2$, and bus 6 in subsystem $\mathcal{A}_2$ has not associated with subsystem $\mathcal{A}_1$. So $A_1^{\mathrm{r}}\cdot[x_1;x_2;x_4]-\boldsymbol{P}_1^{\mathrm{D}}=\mathbf{0}$ can be moved from constrains $\sum_{i\in S^{B}}\tilde{A}_i x_i=b_{\mathrm{PF}}$ to decouple constraint sets χ_1. Similarly, $A_2^{\mathrm{r}}\cdot[x_3;x_5;x_6]-\boldsymbol{P}_6^{\mathrm{D}}=\mathbf{0}$ can be moved from constraints $\sum_{i\in S^{B}}\tilde{A}_i x_i=b_{\mathrm{PF}}$ to decouple constraint sets χ_2. Then, the "real" coupling constraints for the DC nodal power balance equation are reduced to $[\tilde{A}_0^r,\tilde{A}_1^r,\tilde{A}_2^r][x_0;x_1;x_2;x_4;x_3;x_5;x_6]-[\boldsymbol{P}_2^{\mathrm{D}};\boldsymbol{P}_4^{\mathrm{D}};\boldsymbol{P}_3^{\mathrm{D}};\boldsymbol{P}_5^{\mathrm{D}}]=\mathbf{0}$. Now, there is $v\in R^{4T}$ in consensus variable y. The reduction of elements in the consensus variable accelerates the convergence of the DC-ADMM-P (As seen in Fig. 7 of section IV).

D. Analysis of Public Information

As discussed in Section I, all the information that does not belong to the current subsystem is regarded as public information while solving the subproblem corresponding with a subsystem. In simple expression, all information disclosed between adjacent subsystems is called public information. A 6-bus system shown in Fig. 2 is used as an example to illustrate the difference of public information number between our paper and literature [7].

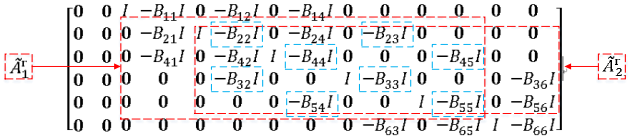

Fig. 4. PUBLIC INFORMATION FOR THE METHOD IN [7].

As shown in Fig. 3 and Fig. 4, the number of the "real" coupling constraints in [7] is the same as our paper. However, the coefficient matrix of "real" coupling constraints in our paper (corresponds to $\tilde{A}_1^r$ and $\tilde{A}_2^r$ in Fig. 3) is different from the coefficient matrix of "real" coupling constraints in [7] (corresponds to $\tilde{A}_1^r$ and $\tilde{A}_2^r$ in Fig. 4. This is the reason why DC-ADMM-P only needs public boundary branches information disclosed among adjacent subsystems while the method in [7] also needs private boundary buses information disclosed among adjacent subsystems.

In Fig. 4, the subsystems partition is the same as in Fig. 3 ($\pi=\{\mathcal{A}_0=\{0\},\mathcal{A}_1=\{1,2,4\},\mathcal{A}_2=\{3,5,6\}\}$). As shown in Fig. 3, B_{32} and B_{23} are the admittance information of the same boundary branch between subsystem $\mathcal{A}_1$ and subsystem $\mathcal{A}_2$. Similarly, B_{54} and B_{45} are the admittance information of the same boundary branch. Admittance B_{23} and B_{45} are not private information for subsystem $\mathcal{A}_1$ and subsystem $\mathcal{A}_2$. Then, admittance B_{23} and B_{45} are called public boundary branches information. For $\sum_{i\in S^{B}\cup\{0\}}\tilde{A}_i x_i=b_{\mathrm{PF}}$, the number of public information for the 6-bus system using DC-ADMM-P is 2.

In Fig. 4, with the known the admittance B_{23}, B_{32}, B_{54}, and B_{45}, subsystem $\mathcal{A}_1$ (corresponds to $\tilde{A}_1^r$ in Fig. 4) still needs to know the admittances B_{33} and B_{55} which are private information for subsystem $\mathcal{A}_2$. Then, admittance B_{33} and B_{55} are called private boundary buses information. Similarly, subsystem $\mathcal{A}_2$ (corresponds to $\tilde{A}_2^r$ in Fig. 4) needs to know the admittances B_{23}, B_{45}, B_{22}, and B_{44}. For $\sum_{i\in S^{B}\cup\{0\}}\tilde{A}_i x_i=b_{\mathrm{PF}}$, the number of public information for the 6-bus system using the method in [7] is 6.

Table I
THE COMPARISON OF PUBLIC INFORMATION FOR 6-BUS SYSTEM BETWEEN DC-ADMM-P AND THE METHOD IN [7]

Constraints	The public information of the method in [7]	The public information of DC-ADMM-P
$\sum_{i\in S^{B}\cup\{0\}}\tilde{A}_i x_i=b_{\mathrm{PF}}$	B_{23}, B_{45}, B_{22}, B_{44}, B_{33}, B_{55}	B_{23} , B_{45}
$\sum_{i\in S^{B}\cup\{0\}}\widetilde{M}_i x_i\le b_{\mathrm{BR}}$	θ_2, θ_3, θ_4, θ_5, x_{23}, x_{45}	x_{23} , x_{45}

For the sake of discussion, Table I shows that the comparison of public information between DC-ADMM-P and the method in [7]. From Table I, we know that there is less information disclosed between adjacent subsystems in DC-ADMM-P comparing with the method in [7]. Therefore, the privacy pro-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2937513, IEEE Transactions on Industrial Informatics

tection of our method is better.

E. Improving the Update Step of the Multiplier

As analyzed in [40], ADMM is the application of the Douglas-Rachford splitting method (DRSM) in the dual problem of (20). The Peaceman-Rachford splitting method (PRSM) scheme differs from ADMM "only through the addition of the intermediate multiplier update (i.e., (35)); it thus offers the same set of advantages."

$$p_j^{k+1/2} := p_j^k + \rho\left(y^{k+1} - z_j^k\right) \tag{35}$$

Compared with the DRSM, PRSM requires more restrictive assumptions to ensure its convergence, while it is always faster whenever it is convergent. The reason for this difference is that the iterative sequence generated by DRSM is strictly contractive, but the iterative sequence generated by PRSM is only contractive with respect to the solution set of the model. In [29], an underdetermined relaxation factor with PRSM to guarantee the strict contraction of its iterative sequence is attached to propose a strictly contractive PRSM. Numerical results in [29] show that the convergence rate of the PRSM is faster than ADMM when it is convergent. Similar to [29], we can obtain the new multiplier update step without increasing communication time of the Algorithm 1.

$$p_j^{k+1/2} := p_j^k + \mu\rho\left(y^{k+1} - z_j^k\right) \tag{36}$$

$$z_j^{k+1} = \begin{bmatrix} \max\left\{0, \lambda_{y^{k+1}} + \frac{1}{\rho}\left(\sum_{i\in\mathcal{A}_j}\psi_i(x_i) + \lambda_{p_j^{k+1/2}}\right)\right\} \\ v_{y^{k+1}} + \frac{1}{\rho}\left(\sum_{i\in\mathcal{A}_j}A_i(x_i) + v_{p_j^{k+1/2}}\right) \end{bmatrix} \tag{37}$$

$$p_j^{k+1} := p_j^{k+\frac{1}{2}} + \mu\rho\left(y^{k+1} - z_j^{k+1}\right) \tag{38}$$

where $\mu \in (0, 1)$ is the underdetermined relaxation factor.

After adopting this new multiplier update step, the improved DC-ADMM-P (P-DC-ADMM-P) is denoted as Algorithm 2.

Algorithm 2: P-DC-ADMM-P

Initialization: z_j^0, p_j^0, $M > 0$, $\rho > 0$, $\varepsilon^{\text{feasible}} > 0$;
for $k = 0, \ldots, M$
  Obtain y^{k+1} according to (27);
  for each sub set of units $j = 0, \ldots, n$: (in parallel)
    Obtain $\{x_i^*, i \in \mathcal{A}_j\}$ by solving (31);
    Obtain $p_j^{k+1/2}$ according to (36);
    Obtain z_j^{k+1} according to (37);
    Obtain p_j^{k+1} according to (38);
  end
  if $\varepsilon := \|y^{k+1} - y^k\|_\infty < \varepsilon^{\text{feasible}}$ **break**;
end.

IV. Numerical Case Studies

DC-ADMM-P is implemented using MATLAB on a desktop with an Intel Core i7-8700K CPU and 8GB of RAM. Problem (31) is solved by Cplex 12.6.2. All the codes and cases of the simulations for this paper can be freely downloaded from https://github.com/linfengYang/Dual-Consensus-ADMM-for-DC-OPF-CET.

In all numerical case studies, $T = 24$, $\pi_b = 12$, $\pi_s = 10$, and $M = 4000$. ρ is a predefined positive parameter. $r_k = [y^k - z_1^k; \ldots; y - z_n^k]$ is the primal residual at iteration k. $s_k = -\rho[z_1^k - z_1^{k-1}; \ldots; z_n^k - z_n^{k-1}]$ is the dual residual. $\|.\|_2$ represents the ℓ_2-norm of a vector. The "F_{TC}" is the calculation result of objective function. The "$F_{\text{TC}}^{\text{Cplex}}$" is the calculation result of objective function obtained by solving problem (20) with Cplex. The "RG" ($=\left|F_{\text{TC}} - F_{\text{TC}}^{\text{Cplex}}\right|/F_{\text{TC}}^{\text{Cplex}}$) is the relative gap between the result obtained by Cplex and the one obtained by our method. And this gap can be used to evaluate the solution quality of the algorithm. The "N_{it}" represents the total number of iterations required to solve the current optimization. Stopping criterion of S_C 1 is $\varepsilon := \|y^{k+1} - y^k\|_\infty < \varepsilon^{\text{feasible}}$, and stopping criterion of S_C 2 is $\|r_k\|_2 < \varepsilon^{\text{pri}}$ and $\|s_k\|_2 < \varepsilon^{\text{dual}}$. Unless otherwise specified, the calculation results of this section are all obtained after reducing the number of dual multipliers.

A. 6-Bus System

The detailed data of the 6-bus system that includes 3 generators and 3 loads can be found in [41]. As shown in Fig. 2, the 6-bus system is divided into seven subsystems, i.e., $= \{\mathcal{A}_j = \{j\}, j = 0, \cdots 6\}$. Then, the number of elements in the consensus variable is $20T$ in Case 1. $E_0 = 600$, $\Delta E_b^{\max} = 10000$, $\Delta E_s^{\max} = 10000$, and $\varepsilon^{\text{feasible}}$ =1e-5.

Table II
THE RESULT FOR THE 6-BUS SYSTEM

Case	ρ	F_{TC}	$\|r_k\|_2$	$\|s_k\|_2$	RG	N_{it}
1	0.14	1.430133e+5	1.95e-4	6.43e-6	4.62E-05	3132

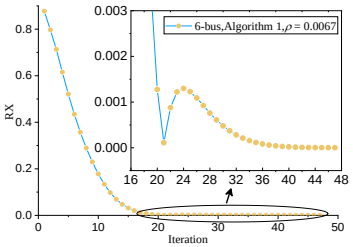

Fig. 5. ITERATIVE VALUES OF RX

Table III
THE RESULTS OF THE METHOD IN [7]

ρ	2	4	8	10	20	50	100
# of iterations in Case 1	285	139	**49**	59	103	160	209
# of iterations in Case 2	3096	1536	740	597	**278**	566	963

For this 6-bus system, the result obtained by Cplex is $F_{\text{TC}}^{\text{Cplex}}$ =1.430057e+5. The computation result of Algorithm 1 for the 6-bus system are listed in Table II. Table II shows that Algorithm 1 obtains almost the same solutions and the solution is nearly equal to the solution obtained by Cplex. Furthermore, the number of iterations in Case 1 of Table II is even larger than that in Table XIV for the 1062-bus system. It indicates that the convergence performance in terms of the number of iterations does not necessarily depend on the scale of power systems.

The 6-bus system in [7] is utilized as an example to compare DC-ADMM-P and the method in [7]. The results of DC-OPF without considering CET are shown in Table III and Table IV. In Table IV, the $\varepsilon^{\text{feasible}}$ for our method is set as 1e-5. The iterative values of RX for DC-ADMM-P is shown in Fig. 5, where RX $= \sqrt{\sum_{i\in S^G}[(P_i^* - P_i)/P_i^*]^2}$. The definition of RX in

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2937513, IEEE Transactions on Industrial Informatics

our paper is the same as the definition of RE in [7]. The iterative values curve of RX in Fig. 5 is smoother than that in Fig. 6 of [7].

Table IV
THE RESULTS OF DC-ADMM-P

ρ	0.002	0.0067	0.0085	0.025	0.06	0.09
# of iterations in Case 1	129	**47**	143	828	614	920
# of iterations in Case 2	2506	1174	961	**237**	770	1139

B. *30-Bus System*

Table V
THE COEFFICIENTS OF THE 30-BUS SYSTEM

Gen	α	β	γ	a	b	c
1	10	0.02	0.0002	11.50486	-0.15586	0.00194
2	10	0.015	0.00024	11.50486	-0.15586	0.00194
3	20	0.018	0.00008	13.86	0.33	0.0042
4	10	0.01	0.00012	13.3056	0.3168	0.00403
5	20	0.018	0.00008	13.86	0.33	0.0042
6	10	0.015	0.0002	11.50486	-0.15586	0.00194

The 30-bus system including 6 generators and 21 loads is given in Fig. 6. The coefficients of production cost function and coefficients of emission function are listed in Table V. The data of emission coefficients are from [42] and reduced in proportion. In this subsection, $E_0 = 600$, $\Delta E_{\mathrm{b}}^{\max} = 10000$, $\Delta E_{\mathrm{s}}^{\max} = 600$, and $\varepsilon^{\text{feasible}} = 1\mathrm{e}-5$.

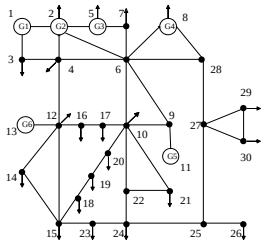

Fig. 6. THE 30-BUS SYSTEM

Table VI
COMPARISON OF PUBLIC INFORMATION FOR THE 30-BUS SYSTEM

Case	Subsystems $\mathcal{A}_n$	# of public information for the method in [7]	# of public information for DC-ADMM-P	# of elements in consensus variable
1	$\mathcal{A}_1 = \{1{:}8,25{:}30\}$ $\mathcal{A}_2 = \{9{:}24\}$	22	**8**	15T
2	$\mathcal{A}_1 = \{1{:}8,25{:}30\}$ $\mathcal{A}_2 = \{9{:}20\}$ $\mathcal{A}_3 = \{21{:}24\}$	36	**14**	25T
3	$\mathcal{A}_1 = \{1{:}10\}$ $\mathcal{A}_2 = \{11{:}20\}$ $\mathcal{A}_3 = \{21{:}30\}$	46	**18**	32T

The comparison of public information between our method and the method in [7] is shown in Table VI. Additionally, the numbers of elements in consensus variable are also shown in this table. It is clear that the amount of public information of DC-ADMM-P is significantly less than that in [7]. In addition, it is easier to calculate the amount of the public information in DC-ADMM-P. Moreover, DC-ADMM-P only needs the number of the boundary branches between adjacent subsystems to calculate the amount of the public information while the method in [7] also need to consider the partition strategy on the basis of knowing the number of boundary branches.

The computation results of Algorithm 1 for three case studies of the 30-bus system are listed in Table VII. The result obtained using Cplex for this system is 5.174874e+4. Table VI and Table VII show that there are often more iterations with more elements in the consensus variable. Thus, a useful guidance for dividing the entire system into subsystems is the number of elements in the consensus vector.

Table VII
RESULTS OF THREE CASES FOR THE 30-BUS SYSTEM

Case	ρ	F_{TC}	$\|r_k\|_2$	$\|s_k\|_2$	RG	N_{it}
1	0.49	5.174889e+4	2.24e-4	2.03e-5	2.90e-6	380
2	0.39	5.174896e+4	7.40e-4	1.86e-5	4.25e-6	622
3	0.38	5.174887e+4	5.45e-4	2.13e-5	2.51e-6	775

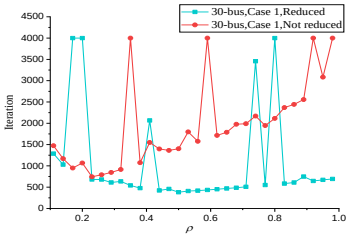

Fig. 7. COMPARISON OF DUAL MULTIPLIERS BEFORE AND AFTER REDUCTION

Furthermore, for Case 1, the iteration numbers of Algorithm 1 with dual multiplier reduction (labeled "Reduced") and without reduction (labeled "Not reduced") were given in Fig. 7. This figure shows that reducing the number of dual multipliers can indeed reduce the number of iterations for most values of ρ.

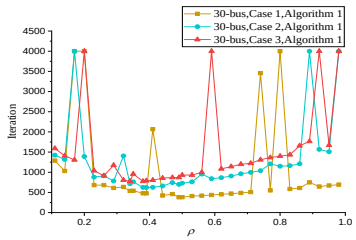

Fig. 8. IMPACT OF ρ ON THE CONVERGENCE RATE

For Case 1-3 of 30-bus system, Fig. 8 gives the number of iterations using Algorithm 1 with different values of ρ. This figure shows that, for most values of ρ, the number of elements in the consensus variable and the number of iterations are positively correlated. In Fig. 8, it also shows that an improper ρ could lead to a slow convergence. In addition, in order to obtain the best convergence performance for different case studies, ρ often are set as different, as shown in Table VII and Fig. 8. Thus, it cannot give a unique value for ρ that would be best for all practical problems [43]. Our experiences show that setting ρ in the range of 0.3-0.55 is usually acceptable for different partitions of the 30-bus system.

In Fig. 8, it can be seen that the Algorithm 1 cannot reach the preset accuracy with 4000 iterations for some values of ρ. References [22] and [29] show that ADMM can be very slow to converge to high accuracy and the numerical performance of the ADMM is highly dependent on the penalty parameter. For these values of ρ, Algorithm 1 always can rapidly converge to acceptable accuracies, which are modest accuracies a little lower than our preset higher accuracy.

Let $\mathrm{RE} = \log\|\mathrm{RG}\|_2$. Then, "RE" of Algorithm 1 (labeled "30-bus, Case 1, Algorithm 1, $\rho = 0.49$") and Algorithm 2 with different ρ and μ values (labeled "30-bus, Case 1, Algo-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2937513, IEEE Transactions on Industrial Informatics

rithm 2, …") is shown in Fig. 9. In this figure, 0.49 is the best choice for ρ in Algorithm 1. $\rho = 0.6$ and $\mu = 0.84$ are the best choices for Algorithm 2. This figure shows that with proper ρ and μ, the new update step of multiplier can improve the convergence performance of Algorithm 1. However, an improper choice of μ may instead exacerbate the convergence performance.

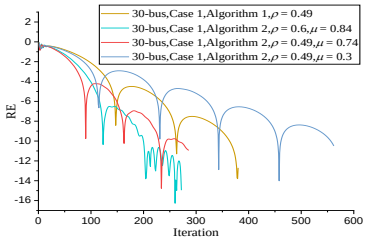

Fig. 9. TWO TYPES OF UPDATE STEPS FOR THE MULTIPLIER

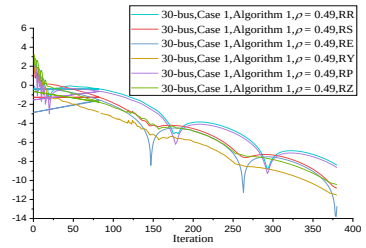

Fig. 10. SEVERAL ITERATIVE LOGARITHMIC VALUES OF 30-BUS SYSTEM

Let $\mathrm{RR} = \log\|r_k\|_2$, $\mathrm{RS} = \log\|s_k\|_2$, $\mathrm{RY} = \log\|y^{k+1} - y^k\|_\infty$, $\mathrm{RP} = \log\|p^{k+1} - p^k\|_\infty$, and $\mathrm{RZ} = \log\|z^{k+1} - z^k\|_\infty$, where $p^{k+1} = [p_1^{k+1}, \ldots, p_n^{k+1}]$ and $z^{k+1} = [z_1^{k+1}, \ldots, z_n^{k+1}]$. As shown in Fig. 10, RR and RS are similar to the RY, which indicates that the primal residual and dual residual achieve an acceptable accuracy when the algorithm stops.

Table VIII
RESULTS OF THREE CASES FOR A 30-BUS SYSTEM AFTER IMPROVING THE UPDATE STEP OF THE MULTIPLIER

Case	ρ	μ	F_{TC}	$\|r_k\|_2$	$\|s_k\|_2$	RG	N_{it}
1	0.6	0.84	5.174875e+4	8.46e-4	3.27e-5	1.93e-7	272
2	0.3	0.70	5.174903e+4	7.61e-4	1.01e-5	5.60e-6	559
3	0.39	0.75	5.174884e+4	3.13e-4	3.58e-5	1.93e-6	580

The computation results of Algorithm 2 are listed in Table VIII. Furthermore, Table VII and Table VIII show that the number of iterations is significantly reduced in all cases when the new update step of multiplier is employed. Moreover, the values of ρ leading to the best performance for the algorithm changes. The calculation time of Algorithm 2 for solving three cases are 23.5s, 59.7s, and 67.8s, respectively.

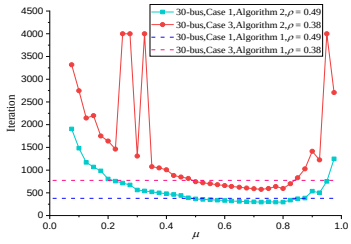

Fig. 11. RESULTS FOR THE NEW UPDATE STEP FOR THE MULTIPLIER UNDER DIFFERENT μ

For different values of μ, Fig. 11 gives the number of iterations for Algorithm 2 (labeled "…, Algorithm 2, …"). In this figure, the dotted line represents the fewest iterations for Algorithm 1 (labeled "…, Algorithm 1, …"). Our results show that μ in the range of 0.55-0.85 is usually acceptable for the 30-bus system.

For Case 1, Fig. 12 gives the number of iterations of our method with $\mu = 0.68$ and different values of ρ. In this figure, the lines with label "30-bus, Case 1, Algorithm 1" and "30-bus, Case 1, Algorithm 2, $\mu = 0.68$" report the results for Algorithm 1 and Algorithm 2 respectively. This figure shows that a good choice of μ can often improve the performance of the algorithm for most values of ρ.

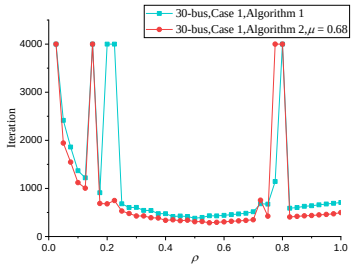

Fig. 12. THE NUMBER OF ITERATIONS WITH μ =0.68 AND DIFFERENT VALUES OF ρ (CASE 1)

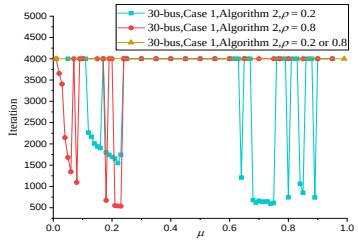

Fig. 13. COMPARISON OF UNIMPROVED AND IMPROVED MULTIPLIER UPDATE STEP WITH ρ OF 0.2 AND 0.8 (CASE 1)

When ρ is 0.2 or 0.8, Algorithm 1 does not converge while solving Case 1 as shown in Fig. 8 and Fig. 12. [22] presents that ADMM is very slow to converge to high accuracy. However, it is often the case that ADMM converges to modest accuracy which is sufficient for many applications. Furthermore, when ρ is 0.2 or 0.8, the primal residual and dual residual can reach 1e-4, which is sufficient for ADMM. For this situation, Fig. 13 gives the numbers of iterations for Algorithm 1 (labeled "30-bus, Case 1, Algorithm 1, ρ=0.2 or 0.8") and Algorithm 2 (labeled "30-bus, Case 1, Algorithm 2, …"). The figure shows that the number of iterations is reduced significantly after using the improved update step of multiplier. In addition, for all values of ρ, DC-ADMM-P can converge after improving the update step of multiplier. Furthermore, there are many other methods to improve the convergence performance of algorithm [31]-[33].

C. RTS-48 bus system

Literature [6] applies six distributed/decentralized algorithms on an OPF test system, and discuss their key features and simulation results. The six distributed/decentralized algorithms are ATC, OCD, ADMM, APP, C+I, and PMP. The general visions of ATC and ADMM are distributed algorithms. Also, the general visions of PMP, APP, OCD, and C+I are decentralized algorithms. In order to verify the effectiveness of our method, we compare the simulation results of our method with

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2937513, IEEE Transactions on Industrial Informatics

those obtained by other methods. The test system in [6] is the IEEE 48-bus system which can obtain from [44]. The generating units capacity, coefficients, and ramp limits are given in [45]. The computation time largely depend on how the implementation code is written. In addition, the choice of the initial values of the variables and starting point highly influence the algorithms' performances. Hence, we focus on the effort that it takes to carry out one iteration and the solution quality rather than the total computation times.

Table IX
SUMMARY OF THE RESULTS FOR RTS48-BUS SYSTEM IN [6]

Algorithm	ATC	PMP	APP	OCD	C+I
Iteration	96	24157	104	76(34)	250
Average computation time per iteration	0.08	0.02	0.05	<0.01(>0.01)	0.01

Table X
SUMMARY OF THE RESULTS FOR T-ADMM AND OUR METHODS

Algorithm	Classical ADMM in [6]	Algorithm 1		Algorithm 2	
		S_C 1	S_C 2	S_C 1	S_C 2
Parameter	$\rho = 5000$	$\rho = 0.53$	$\rho = 0.53$	$\rho = 0.53$, $\mu = 0.73$	$\rho = 0.53$, $\mu = 0.73$
Iteration	250	300	253	**207**	**164**
Average computation time per iteration	0.08	**0.057**	**0.057**	**0.057**	**0.057**

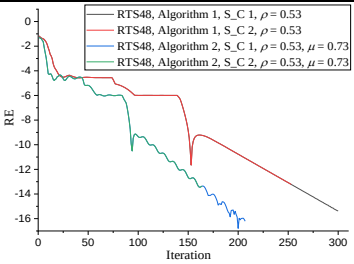

Fig. 14. THE CONVERGENCE PROPERTY OF ALGORITHM 1 AND 2

Table IX and Table X give a summary over the number of iterations and the average computation time per iteration of each algorithm. In Table X, Algorithm 1 and Algorithm 2 adopt two different types of stopping criterion. $\varepsilon^{\text{feasible}}$ is set to 1e-5. ε^{pri} and $\varepsilon^{\text{dual}}$ are both set to 1e-3.

Table XI
THE RESULTS OF ALGORITHM 1 AND 2 FOR RTS48-BUS SYSTEM

Algorithm	ρ	μ	S_C	F_{TC}	$\|r_k\|_2$	$\|s_k\|_2$	RG
1	0.53	None	1	1.8128228812e+6	1.14e-4	1.47e-5	2.11e-7
1	0.53	None	2	1.8128257889e+6	9.93e-4	1.28e-4	1.82e-6
2	0.53	0.73	1	1.8128226639e+6	6.24e-5	1.54e-5	9.16e-8
2	0.53	0.73	2	1.8128253628e+6	9.41e-4	2.82e-4	1.58e-6

In [6], it shows detailed comparison of six popular decomposition coordination algorithms: ATC, PMP, classical ADMM, APP, OCD, and C+I. The comparison of classical ADMM in [6], Algorithm 1 and Algorithm 2 are given in this paper. The classical ADMM can't solve the problem with nonlinear constraints (like CET) directly. The average computation time per iteration of Algorithm 1 and Algorithm 2 which solve the dual of DC-OPF is shorter than the classical ADMM in [6]. Algorithm 1 and Algorithm 2 have similar or even fewer iterations than the classical ADMM. The convergence property of classical ADMM, Algorithm 1 and Algorithm 2 is compared. The convergence property of Algorithm 1 and Algorithm 2 are shown in Fig. 14 and the convergence property of classical ADMM is shown in [6]. In addition, the results of Algorithm 1 and Algorithm 2 for RTS48-bus system are given in Table XI. For the RTS48-bus system, the result obtained by Cplex is 1.8128224979e+6.

D. 118-bus system

With the aim of sufficient comparison, the proposed Algorithm 1 and Algorithm 2 is utilized to solve the DC-OPF problem with DR in literature [7]. The detailed data for the IEEE 118-bus system in [7] is from [46]. As in [7], an eight-piece piece-wise segments is utilized to approximate the price-demand curves. The reference point for the demand response curve is (80,25) [47]. The price elasticity is set to -1 [48]. $\varepsilon^{\text{feasible}}$ is set to 1e-5. The values of ε^{pri} and $\varepsilon^{\text{dual}}$ are both set to 0.01, as in [7]. The definition of RE in [7] is the same as the definition of RX in our article. The IEEE 118-bus system is tested for a 24-hour period via three cases.

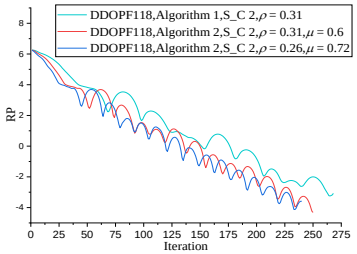

Fig. 15. ITERATIVE LOGARITHMIC VALUES OF RX IN CASE 1

Table XII
COMPARISON OF THE METHODS IN [7] AND OUR ALGORITHMS

Case	1				2				3			
Algorithm	2 in [7]	3 in [7]	1	2	2 in [7]	3 in [7]	1	2	2 in [7]	3 in [7]	1	2
# of iteration	488	307	**268**	**250**	2126	1957	**560**	**370**	792	621	**267**	**248**
# of exchanged information unit per iteration	40	102	105	105	64	142	186	186	56	134	141	141

The iterative logarithmic values of RX in Case 1 are shown in Fig. 15. It indicates that the convergence rate of Algorithm 2 is accelerated as compared to Algorithm 1.

To compare communication efficiency of the methods in [7] and our proposed algorithms, number of information units exchanged in the distributed optimization procedure are compared. The definition of one information unit in [7] is transforming 24 values of one variable for the 24 hours from one subsystem to another. Note that information units exchanged in our proposed Algorithm 1 and Algorithm 2 include p_j^{k+1}, z_j^{k+1}, and y^{k+1}. p_j^{k+1} and z_j^{k+1} are sent from subsystems to the central coordinator. y^{k+1} is sent from the central coordinator to subsystems. Table XII shows the number of iterations and the number exchanged information unit per iteration of the methods in [7] and our proposed algorithms. The stopping criterion used in Table XII is S_C 2 ($\|r_k\|_2 \le \varepsilon^{\text{pri}}$, $\|s_k\|_2 \le \varepsilon^{\text{dual}}$). Although it shows that number of information units exchanged per iteration in our proposed algorithms is larger than those of Algorithm 2 and Algorithm 3 in [7]. However, considering the iterations of our methods are notable fewer, the total number of

information units exchanged in our methods are fewer than of comparable with the methods in [7]. And when the scale of problem is large, the computation times highly depend on the computation time in each iteration rather than the communication time.

The two major factors that impact the computational performance of the Algorithm 1 and Algorithm 2 are the number of iterations and the computational time of individual subsystems. However, the experimental results show that the total computational time of DC-OPF problem heavily depend on the number of iterations. As for the IEEE 118-bus system, total computational time of Cases 1-3 for Algorithm 1 are 53.10s, 168.40s, 61.00s, respectively. And the total computational time of Case 1-3 for Algorithm 2 are 45.28s, 90.87s, 58.32s, respectively. It is confirmed that the computational performance of the distributed algorithms is highly impacted by the number of iterations.

In addition, when the stopping criterion is S_C 1, the number of iterations of Case 1-3 for Algorithm 1 are 425, 871, 513, respectively, and the number of iterations of Case 1-3 for Algorithm 2 are 391, 613, 420, respectively. The number of iterations of our proposed Algorithm 1 and Algorithm 2 are also fewer than the number of iterations of Algorithm 2 and Algorithm 3 in [7].

E. 1062-bus system

As shown in Fig. 16, the 1062-bus system is tested to explore whether our algorithm would generalize to larger systems The 1062-bus system includes 9 subsystems and each subsystem is a 118-bus system which can be retrieved from http://motor.ece.iit.edu/Data/. $E_0 = 60000$, $\Delta E_{\mathrm{s}}^{\max} = 60000$ and $\Delta E_{\mathrm{b}}^{\max} = 1000000$.

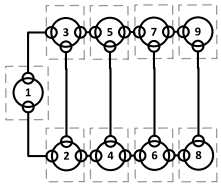

Fig. 16. 1062-Bus System

Table XIII
COMPARISON OF PUBLIC INFORMATION

Subsystems $\mathcal{A}_n$	# of public information for the method in [7]	# of public information for our method	# of the element in consensus vector
$\mathcal{A}_j = \{118 * (j-1)+1 : 118 * j\}$, $j = 1, \cdots, 9$	72	**24**	*48T*

Table XIV
COMPARISON OF ALGORITHM 1 AND 2 FOR A 1062-BUS SYSTEM

Algorithm	ρ	μ	F_{TC}	$\|\|r_k\|\|_2$	$\|\|s_k\|\|_2$	RG	N_{it}
1	4.4	None	1.834461926e+7	2.12e-4	2.56e-4	4.73e-5	2344
2	4.4	0.61	1.834461447e+7	2.20e-4	2.87e-4	4.70e-5	2120

As shown in Table XIII, the 1062-bus system is divided into ten subsystems ($\mathcal{A}_0 - \mathcal{A}_9$), and it contains 12 boundary branches. The calculated result of Cplex for the 1062-bus system is 1.834375218e+07.

It gives the comparison of public information between our methods and the method of [7] in Table XIII. The computational results of Algorithm 1 and Algorithm 2 for 1062-bus system are summarized in Table XIV. The values of $\varepsilon^{\mathrm{feasible}}$ for Algorithm 1 and Algorithm 2 are both 1e-5. From Table XIV, it shows the number of iterations is reduced by Algorithm 2. Moreover, it is proven that Algorithm 1 and Algorithm 2 are efficient for large-scale power systems.

V. CONCLUSION

This paper proposes a distributed ADMM algorithm for solving the DC-DOPF-CET. Since the DC-DOPF-CET problem is a convex programming with linear and nonlinear constraints, DC-ADMM-P guarantees the global optimal solutions. Essentially, DC-ADMM-P need significantly less public information than other ADMM-based methods, and is more suitable for a distributed optimization to large-scale power systems whose privacy protection is important.

In this paper, we focus on the impact of parameter ρ and μ, the number of elements in the consensus variable, the number of subsystems and the scale of power systems on the convergence performance. In order to verify the performance of our proposed algorithms and the effective of improved, we apply them on a various of numerical case studies, and compare their key features and simulation results with the method in [6] and [7]. The numerical case studies include a 6-bus system, a 30-bus system, a RTS48-bus system in [6], a 118-bus system in [7] and a 1062-bus system. The results show that the convergence performance largely depends on the number of elements in the consensus variable rather than the number of subsystems or the scale of power systems. Numerical results also confirm that the computation efficiency of DC-ADMM-P is improved obviously, especially when applied to large-scale power systems. Therefore, an excellent subsystem partition strategy could enhance the convergence performance, since it has less global variables.

REFERENCES

[1] J. Carpentier, "Contribution à l'étude du dispatching économique," *Bull. Soc. Françai se D'Electricité*., 1962.
[2] J. A. Momoh and J. Z. Zhu, "Improved interior point method for OPF problems," *IEEE Trans. Power Syst.*, vol. 14, no. 3, pp. 1114-1120, Aug. 1999.
[3] 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.
[4] 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.
[5] T. Erseghe, "Distributed optimal power flow using admm," *IEEE Trans. Power Syst.*, vol. 29, no. 5, pp. 2370–2380, Sept. 2014.
[6] A. Kargarian, J. Mohammadi, J. Guo, S. Chakrabarti, M. Barati, G. Hug, S. Kar and R. Baldick, "Toward Distributed/Decentralized DC Optimal Power Flow Implementation in Future Electric Power Systems," *IEEE Trans. Smart Grid*, 2016.
[7] Y. Wang, L. Wu and S. Wang, "A Fully-Decentralized Consensus-Based ADMM Approach for DC-OPF With Demand Response," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2637-2647, Nov. 2017.
[8] Z. Liang, S. Lin and M. Liu, "Distributed Optimal Power Flow of AC/DC Interconnected Power Grid Using Synchronous ADMM," *IOP Conference Series: Materials Science and Engineering*, vol. 199, no. 1, 2017.
[9] A. Kargarian, 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.
[10] A. Ahmadi-Khatir, A. Conejo, and R. Cherkaoui, "Multi-area unit scheduling and reserve allocation under wind power uncertainty," *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1701–1710, July. 2014.
[11] 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.
[12] S. Yang, S. Tan, and J. Xu, "Consensus based approach for economic dispatch

problem in a smart grid," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4416–4426, Nov. 2013.
[13] M. Ma, L. Fan and Z. Miao, "Consensus ADMM and Proximal ADMM for economic dispatch and AC OPF with SOCP relaxation," *North American Power Symposium*, 2016.
[14] S. Magnusson, P. Weeraddana, and C. Fischione, "A distributed approach for the optimal power flow problem based on admm and sequential convex approximations," *IEEE Trans. Control of Net. Syst.*, vol. 2, no. 3, pp. 238–253, Sept. 2015.
[15] M. Kraning, E. Chu, J. Lavaei, and S. Boyd, "Dynamic network energy management via proximal message passing," *Foundations and Trends in Optimization*, vol. 1, no. 2, pp. 70–122, 2013.
[16] A. Sun, D. Phan, and S. Ghosh, "Fully decentralized AC optimal power flow algorithms," *IEEE Power and Energy Society General Meeting*, July. 2013.
[17] Y. Xu, J. Hu, W. Gu, W. Su and W. Liu, "Real-Time Distributed Control of Battery Energy Storage Systems for Security Constrained DC-OPF," *IEEE Trans. on Smart Grid*, vol. 9, no. 3, pp. 1580-1589, May. 2018.
[18] G. Chen, C. Li and Z. Y. Dong, "Parallel and Distributed Computation for Dynamical Economic Dispatch," *IEEE Trans. Smart Grid*, vol. 8, no. 2, pp. 1026-1027, Mar. 2017.
[19] G. Chen and J. Y. Li, "A Fully Distributed ADMM-Based Dispatch Approach for Virtual Power Plant Problems," *Applied Mathematical Modelling*, vol. 58, pp. 300-312, June. 2018.
[20] Z. Li, Q. Guo, H. Sun, and J. Wang, "Coordinated economic dispatch of coupled transmission and distribution systems using heterogeneous decomposition," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4817-4830, Nov. 2016.
[21] E. Dall'Anese, H. Zhu and G. B. Giannakis, "Distributed Optimal Power Flow for Smart Microgrids," *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1464-1475, Sept. 2013.
[22] S. Boyd, N. Parikh, E. Chu, B. Peleato and J. Eckstein, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," *Now Foundations and Trends*, vol. 3, no. 1, pp. 1-128, 2011.
[23] S. Boyd; L. Vandenberghe, "Convex Optimization," *Cambridge University Press*, 2004.
[24] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximation," *Computers & Mathematics with Applicaions*, vol. 2, no. 1, pp. 17–40, 1976.
[25] (Feb. 5, 2013). [Online]. Available: https://www.weforum.org/reports/energy-vision-2013-energy-transitions-past-and-future.
[26] J. H. Talaq, F. EI-Hawary and M. E. EI-Hawary, "A Summary of Environmental/Eco nomic Dispatch Algorithms," *IEEE Trans. Power Syst.*, vol. 9, no. 3, pp. 1508-1516, 1994.
[27] A. Abdollahi, M. P. Moghaddam, M. Rashidinejad and M. K. Sheikh-EI-Eslami, "Inv estigation of Economic and Environmental-Driven Demand Response Measures Inco rporating UC," *IEEE Trans. Smart Grid*, vol. 3, no. 1, pp. 12-25, 2012.
[28] J. Xie, J. Zhong, Z. Li and D. Gan, "Environmental-economic unit commitment using mixed-integer linear programming," *Euro. Trans. Electr. Power*, vol.21, pp. 772-786, 2011
[29] B. S. He, H. Liu, Z. Wang and X. Yuan, "A Strictly Contractive Peaceman-Rachford Splitting Method For Convex Programming," *Siam Journal on Optimization A Publication of the Society for Industrial and Applied Mathematics*, vol. 24, no. 3, pp. 1011-1040, 2014.
[30] B. S. He, M. Tao, and X. M. Yuan, "Alternating direction method with Gaussian back substitution for separable convex programming," *SIAM J. Optim.*, vol. 22, no. 2, pp. 313-340, Apr. 2012.
[31] W. Deng, M. J. Lai, Z. Peng, *et al*. "Parallel Multi-Block ADMM with o(1/k) Convergence," *Journal of Scientific Computing*, vol.71, no. 2, pp. 712-736, 2013.
[32] J. J. Wang and W. Song. "An algorithm twisted from generalized ADMM for multi-block separable convex minimization models," vol. 309, pp. 342-358, 2016.
[33] C. J. Li, X. H. Yu, W. W. Yu, G. Chen and J. H. Wang. "Efficient Computation for Sparse Load Shifting in Demand Side Management," *IEEE Trans. Smart Grid*, vol. 8, no. 1, pp. 250-261, 2016.
[34] S. Bisanovic, M. Hajro and M. Dlakic, "Unit commitment problem in deregulated environment," *International Journal of Electrical Power and Energy Systems*, vol. 42, no. 1, pp. 150-157, Nov. 2012.
[35] (Apr. 22, 2016) [Online]. Available: https://ec.europa.eu/clima/policies/international /negotiations/paris_en.
[36] I. Kockar, A. J. Conejo and J. R. Mcdonald, "Influence of the Emissions Trading Scheme on generation scheduling," *International Journal of Electrical Power and Energy Systems*, vol. 31, no. 9, pp. 465-473, 2009.
[37] M. Fukushima, "Application of the alternating direction method of multipliers to separable convex programming problems," *Computational Optimization and Applications*, vol. 1, no. 1, pp. 93-111, Jan. 1992.
[38] Y. L. Xu, J. F. Hu, W. Gu, W. C. Su and W. X. Liu, "Real-Time Distributed Control of Battery Energy Storage Systems for Security Constrained DC-OPF," *IEEE Trans. on Smart Grid*, vol. 9, no. 3, pp. 1580-1589, May, 2018.
[39] C. J. Li, Y. Xu, X. H. Yu, C. Ryan and T. W. Huang, "Risk-Averse Energy Trading in Multienergy Microgrids: A Two-Stage Stochastic Game Approach," *IEEE Trans. on Industrial Informatics*, vol. 13, no. 5, pp. 2620-2630, Oct. 2017.
[40] D. Gabay, "Applications of the method of multipliers to variational inequalities," *Elsevier B.V.*, pp.299-331, 1983
[41] X. Bai and H. Wei, "Semi-definite programming-based method for security-constrained unit commitment with operational and optimal power flow constraints," *Iet Generation Transmission and Distribution*, vol. 3, no. 2, pp. 182-197, Feb. 2009.
[42] U. Güvenç, Y. Sönmez, S. Duman and N Yörükeren, "Combined economic and emission dispatch solution using gravitational search algorithm," *Scientia Iranica*, vol. 19, no. 6, pp. 1754-1762, 2012.
[43] E. Ghadimi, A. Teixeira, I. Shames, M. Johansson, "Optimal Parameter Selection for the Alternating Direction Method of Multipliers (ADMM): Quadratic Problems," *IEEE Trans. Automatic Control*, vol. 60, no. 3, pp. 644-658, Mar. 2015.
[44] A. W. Schneider. "The IEEE Reliability Test System – 1996," *IEEE Trans. Power Syst.*, vol. 14, no. 3, pp. 1019-1020, 1999.
[45] C. Wang and S. M. Shahidehpour. "Effects of ramp-rate limits on unit commitment and economic dispatch," *IEEE Trans. Power Syst.*, vol. 8, no. 3, pp. 1341-1350, 1993.
[46] [Online]. Available: http://people.clarkson.edu/~lwu/data/IEEE118_ADMM/, accessed Jan. 2016.
[47] C. Zhao, J. H. Wang, J.-P. Watson, and Y. P. Guan, "Multi-stage robust unit commitment considering wind and demand response uncertainties," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2708-2717, Aug. 2013.
[48] P. Thimmapuram, J.Kim, A. Botterud, and Y. Nam, "Modeling and simulation of price elasticity of demand using an agent-based model," *2010 Innovative Smart Grid Technologies(ISGT)*, pp.1-8, 2010.

Linfeng Yang (M'18) received the B. S. degree in computer science from Guilin University of Technology, Guilin, China, in 2002, the M.S. degree from Shanghai University, Shanghai, China, in 2005, and the Ph.D. degree in electrical engineering from Guangxi University, Nanning, China, in 2012. Now he is a professor in Guangxi University. His research interests include parallel and distribute computing and optimization in power systems.

Jiangyao Luo received the B. S. degree in Ningbo University of Technology, Ningbo, China, in 2016, the M. S. degree in the School of Computer Electronics and Information from Guangxi University, Nanning, China, in 2019. Now he is pursuing his Ph.D degree in Beijing University of Posts and Telecommunications.

Yan Xu (S'10-M'13-SM'19) received the B.E. and M.E degrees from South China University of Technology, Guangzhou, China in 2008 and 2011, respectively, and the Ph.D. degree from The University of Newcastle, Australia, in 2013. He is now the Nanyang Assistant Professor at School of Electrical and Electronic Engineering, Nanyang Technological University (NTU), and a Cluster Director at Energy Research Institute @ NTU (ERI@N), Singapore. Previously, he held The University of Sydney Postdoctoral Fellowship in Australia. His research interests include power system stability and control, microgrid, and data-analytics for smart grid applications. Dr Xu is an Editor for IEEE TRANSACTIONS ON SMART GRID, CSEE Journal of Power and Energy Systems, and an Associate Editor for IET Generation, Transmission & Distribution.

Zhenrong Zhang received the B.S. and M.S. degrees from the Electronics Department, Peking University, China, in 1998 and 2001, respectively, and the Ph.D. degree from Nanyang Technological University, Singapore, in 2006. He is currently a Postdoctoral Research Associate with National Laboratory on Local Optic-fiber Communication Networks, Peking University, and also a Professor with the School of Computer, Electronic and Information, Guangxi University. (The corresponding author)

ZhaoYang Dong (M'99–SM'06-F'17) obtained Ph.D. degree from the University of Sydney, Australia in 1999. He is currently the SHARP professor, Director of UNSW Digital Grid Futures Institute, with the University of New South Wales, Australia. He is also Director of ARC Research Hub for Integrated Energy Storage Solutions. He was previously Ausgrid Chair and Director of the Centre for Intelligent Electricity Networks, the University of Newcastle, Australia. He also held industrial positions with Transend Net-works (now TAS Networks), Australia. His research interest includes Smart Grid, power system planning, power system security, renewable energy systems, electricity market, load modelling, and computational intelligence and its application in power engineering. He is an editor of a number of IEEE transactions and IET journals. He is a Fellow of IEEE.