Distributed Online Voltage Control for Wind Farms Using Generalized Fast Dual Ascent

Yifei Guo, *Member, IEEE*, Houlei Gao, *Member, IEEE*, and Zhaoyu Wang, *Member, IEEE*

***Abstract*—This paper considers the voltage control problem in wind farms, with the goals of maintaining feasible voltage profile and adequate reactive power (VAR) reserves by coordinating the reactive power outputs of wind turbines (WTs) and static synchronous compensators (STATCOMs). Rather than a conventional centralized and offline fashion, we propose a distributed online voltage control scheme based on dual decomposition. First, the optimal voltage control problem is formulated based on the branch flow model. Then, thanks to the sparsity of the constraints in this problem, a distributed solution framework is developed based on the dual decomposition and its online implementation is achieved through voltage feedback. To pursue fast convergence of the distributed algorithm and consequently facilitate the online implementation, the fast gradient method is generalized and then applied to solve the dual problem. The effectiveness of the proposed voltage control scheme is numerically validated under both static and dynamic cases.**

***Index Terms*—dual ascent (DA), distributed control, reactive power (VAR), voltage control, wind turbine (WT), wind farm**

NOMENCLATURE

A. Abbreviations

ADMM	Alternating direction method of multipliers
DA	Dual ascent
DFIG	Doubly-fed induction generator
Ef-DA	Enhanced fast dual ascent
f-DA	Fast dual ascent
FSC	Full-scale converter
Gf-DA	Generalized fast dual ascent
PCC	Point of common coupling
QP	Quadratic programming
s-DA	Standard dual ascent
STATCOM	Static synchronous compensator
VAR	Reactive power
WT	Wind turbine

B. Variables

p_i, q_i	Active/reactive power injections at node i
P_{ij}, Q_{ij}	Active/reactive power flows over branch (i,j)
q_i^c	Reactive power command at node i
q_i^*	Reactive power value at node i obtained by primal update
$\boldsymbol{q}_{\min}, \boldsymbol{q}_{\max}$	Vectors of min/max VAR limits
$\boldsymbol{q}_{\mathrm{mid}}$	Vector of middle reactive power level
t	Time stamp/iteration step
$\boldsymbol{v}$	Vector of squared voltage magnitude
$\boldsymbol{v}^\star, \boldsymbol{q}^\star$	Optimal value of voltage and reactive power
v_i, v_i^m	Squared voltage magnitude at node i and its instantaneous measurement
v_i^*	Squared voltage magnitude value at node i obtained by primal update
$\boldsymbol{\lambda}_{\min}, \boldsymbol{\lambda}_{\max}$	Dual vectors associated with voltage limits
$\boldsymbol{\mu}_{\min}, \boldsymbol{\mu}_{\max}$	Dual vectors associated with VAR limits
$\boldsymbol{\omega}$	Dual vector associated with the power flow
$\boldsymbol{\omega}_{\mathcal{N}_i}$	Subvector of $\boldsymbol{\omega}$ associated with i
$\boldsymbol{\nu}$	Compact representation of all dual vectors
$\boldsymbol{\eta}_{\min}, \boldsymbol{\eta}_{\max}$	Auxiliary vectors associated with $\boldsymbol{\lambda}_{\min}, \boldsymbol{\lambda}_{\max}$
$\boldsymbol{\xi}_{\min}, \boldsymbol{\xi}_{\max}$	Auxiliary vectors associated with $\boldsymbol{\mu}_{\min}, \boldsymbol{\mu}_{\max}$
$\boldsymbol{\chi}$	Auxiliary vector associated with $\boldsymbol{\omega}$
$\boldsymbol{\nu}^\star$	Optimal dual vector

C. Parameters

$\bar{\boldsymbol{A}}$	Graph incidence matrix of network
$\boldsymbol{A}$	Reduced graph incidence matrix
g_{ij}, b_{ij}	Parallel conductance/susceptance at node i
$\boldsymbol{H}_v$	Weighting matrix associated with voltage
$\boldsymbol{H}_q$	Weighting matrix associated with VAR
$\boldsymbol{I}_N$	$N \times N$ identity matrix
$L, \boldsymbol{L}$	Lipschitz constant and its generalization
N	Number of nodes
r_{ij}, x_{ij}	Resistance/ reactance of branch (i,j)
T_γ	Time period of γ updating
$\boldsymbol{v}_{\min}, \boldsymbol{v}_{\max}$	Vector of min/max voltage limit
$\boldsymbol{v}_r$	Vector of voltage reference
$\boldsymbol{X}$	Diagonal branch reactance matrix
$\boldsymbol{\Lambda}$	Diagonally scaled step size matrix in s-DA
$\boldsymbol{0}_N$	$N \times N$ zero matrix
$\boldsymbol{1}_N$	$N \times 1$ vector with all entries being 1

D. Sets

$\mathcal{E}$	Set of branches
$\mathcal{N}$	Set of nodes
$\mathcal{N}_i$	Set of neighbors of node i in cyber layer
$\mathcal{N}_j^+$	Set of adjacent downstream nodes of node j

This work was supported by National Natural Science Foundation of China under grant 51877127. (*Corresponding author: Houlei Gao.*)

Y. Guo is with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011 USA (e-mail: yfguo.sdu@gmail.com).

H. Gao is with Key Laboratory of Power System Intelligent Dispatch and Control of Ministry of Education, Shandong University, Jinan 250061, China (e-mail:houleig@sdu.edu.cn).

Z. Wang is with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011 USA (e-mail:wzy@iastate.edu).

I. INTRODUCTION

AS modern wind farms tend to be large-scale, they could run as relatively autonomous systems including generation, collection and grid integration. In recent years, a number of wind turbine (WT) trip-off failures occurred in the large-scale wind bases in Northern and Northwestern China due to over-/under-voltage problems, suggesting that voltage control/management is an essential issue for the secure and economical operation of wind farms [1]–[2].

Generally, transmission system operators only specify the voltage/reactive power (VAR) requirements at the point of common coupling (PCC). Therefore, several simple PCC voltage/VAR control schemes of wind farms have been proposed in [3]–[6], wherein the voltage, VAR or power factor at the PCC is controlled using the proportional-integral or droop control, and the total required VAR is then proportionally (or equally) dispatched to each WT. Obviously, such methods are unable to tackle the over-/under-voltage issues at each WT terminal bus. This motivates the studies on optimization-based control, which not only optimize the voltage profile but also achieve some specific farm-level goals, e.g, reducing losses, minimizing active power curtailment and keeping VAR reserves, by optimally coordinating multiple voltage regulation devices, e.g., on-load-tap-changing transformers, static VAR compensators, static synchronous compensators (STATCOMs) as well as modern WTs [7]–[15].

For centralized voltage control, a central farm controller is usually designed as a supervisor for device-level controllers, with a relatively slow sampling rate. The decentralized (purely local) and distributed control are appealing to the wind farm control since a wind farm naturally consists of a number of WTs that act highly autonomously. Several distributed/decentralized voltage control methods for wind farms have been proposed in [16]–[19]. Most of them aim to achieve an equilibrium without pursuing specific optimality. Additionally, the decentralized methods might cause VAR oscillations and voltage flicker due to interactions among independent control loops [20]. Considering the similarity of network topology between distribution systems and wind farms, some existing voltage control algorithms for distribution systems could also work for wind farms. We refer the readers to [21]–[23] for more comprehensive surveys on that topic, especially for the distributed control and optimization algorithms with computational and robust superiorities. Additionally, recently, some reinforcement learning-based control methods, e.g., [24]–[26], were proposed to better adapt to various operating conditions, which generally rely on massive offline or online training.

In those optimal control methods, optimization problems are mostly solved in an *offline* way, namely, all the variables are iteratively updated in the cyber layer until they converge and then the final solution is applied to the physical networks [7]–[15]. Therefore, they might not be able to track fast voltage fluctuations caused by the high variability of renewables or loads. This motivates the recently increasing interest of online feedback-based optimization (see [22, Sec.IV] and references therein for related works). The so-called *online* algorithms apply each update to physical networks without waiting for convergence and are thus able to catch up with voltage fluctuations at a fast timescale.

Dual-based methods, e.g. *dual ascent* (DA) and *alternating direction method of multipliers* (ADMM) [27], are widely used to develop distributed optimization and control algorithms for power systems [22]–[23]. Especially, the authors in [28]–[30] have contributed valuable research on online voltage control based on the dual methods. A known drawback of DA with conventional gradient methods is that it often suffers from slow convergence. However, convergence speed is of great importance for online algorithms since it is closely related with the tracking capabilities. In [31], DA was improved for distributed voltage control by combining with an accelerated gradient method; however, it was designed under an offline case and the theoretic convergence analysis was not provided.

In this context, we propose a distributed online voltage control (DOnVC) scheme and address its application in wind farms. The VAR outputs of WTs and STATCOM are adjusted to maintain the voltage profile across the wind farm network while optimizing VAR reserves. A distributed solution framework is established based on dual decomposition and then its online implementation is achieved by sharing information with immediate neighbors. To pursue a fast convergence of the distributed algorithm, the fast gradient method [32] is generalized and used to efficiently solve the dual problem. Compared with the existing works, the main contributions of this paper are summarized as follows:

- First, the proposed algorithm builds on the mildest requirements for sensing, measurement and communication. This DOnVC scheme only uses local voltage measurement as feedback and exchanges it with immediate neighbors without the need of other operation information that are necessary for some offline algorithms. Besides, a central coordination is not required.
- Second, the fast gradient (projection) method is generalized and then applied to solve the dual problem, significantly improving the convergence performance of DA. Beyond an analysis of sufficient conditions for convergence, we further explore the convergence rates not only for dual function, but also for primal sequences (voltage and VAR) and primal infeasibility via the primal-dual relation. Numerical comparisons with traditional DA and ADMM are provided to demonstrate the faster convergence rate of the proposed algorithm in this problem, implying a better tracking capability.
- Third, in addition to the fast tracking capability, the online feedback strategy can compensate for the model errors due to its closed-loop nature, even though the algorithm establishes on a linearized power flow model, which implies good robustness.

The reminder of this paper is organized as follows. Section II presents the wind farm model. In Section III, the voltage control problem is formulated. Section IV presents the distributed solution framework, followed by convergence analyses in Section V. Section VI outlines the online implementation strategy of the distributed voltage control algorithm. Numerical results are given in Section VII, followed by conclusions.

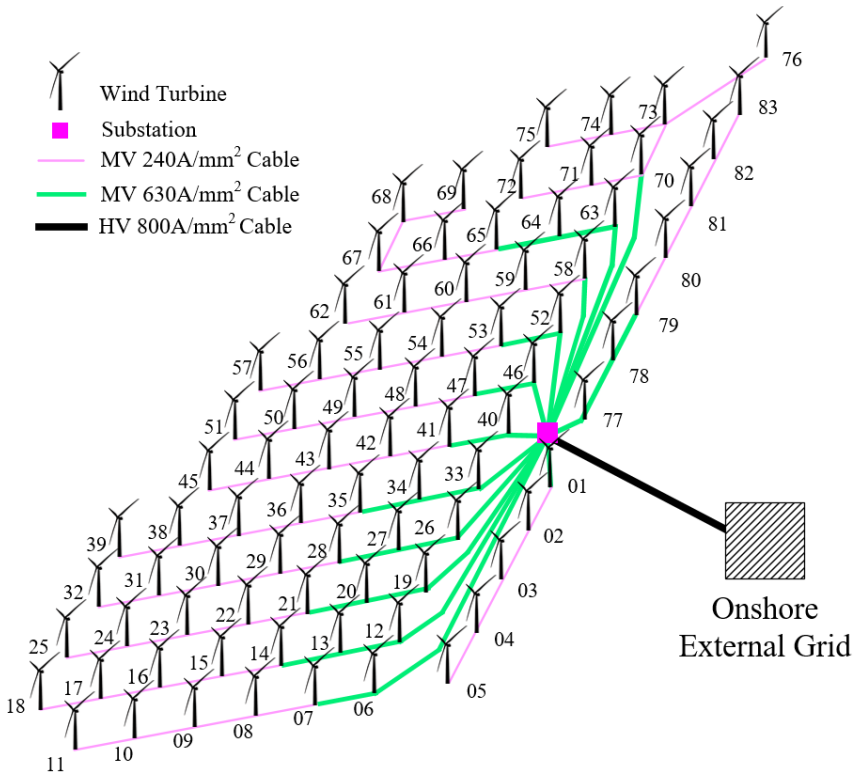


Fig. 1. Layout of French Fécamp offshore wind project [9].

II. Modelling of a Wind Farm

A. Network

A typical layout of a wind farm is illustrated in Fig. 1. Consider a radial power collector system consisting of $N+1$ nodes denoted by set $\mathcal{N}\bigcup\{0\}$, $\mathcal{N} := \{1,\dots,N\}$ and N branches denoted by $\mathcal{E}$. Node 0 denotes the connection point with external grids, thereby treated as a slack bus of the collector system. Such a radial network can be modelled by the branch flow model [33], which is as, $\forall(i,j)\in\mathcal{E}$,

$$P_{ij} = \sum_{k\in\mathcal{N}_j^+} P_{jk} - p_j + r_{ij}\frac{P_{ij}^2+Q_{ij}^2}{v_i} + g_j v_j, \tag{1a}$$

$$Q_{ij} = \sum_{k\in\mathcal{N}_j^+} Q_{jk} - q_j + x_{ij}\frac{P_{ij}^2+Q_{ij}^2}{v_i} - b_j v_j, \tag{1b}$$

$$v_i = v_j + 2\left(r_{ij}P_{ij}+x_{ij}Q_{ij}\right) - \left(r_{ij}^2+x_{ij}^2\right)\frac{P_{ij}^2+Q_{ij}^2}{v_i}. \tag{1c}$$

This yields a compact model as,

$$\boldsymbol{M}\boldsymbol{v} = \boldsymbol{q} + \boldsymbol{b} \tag{2}$$

where $\boldsymbol{q} := [q_1,\dots,q_N]$ and $\boldsymbol{v} := [v_1,\dots,v_N]$. $\boldsymbol{M} := [m_{ij}] = (1/2)\boldsymbol{A}\boldsymbol{X}^{-1}\boldsymbol{A}^T \in \mathbb{R}^{N\times N}$ is the weighted Laplacian matrix. $\boldsymbol{X} := \mathrm{diag}(\{X_l\}) \in \mathbb{R}^{N\times N}$ with $X_l = x_{ij}$ where (i,j) denotes the lth branches in $\mathcal{E}$. $\boldsymbol{A} := [a_{ij}] \in \mathbb{R}^{N\times N}$ is the reduced incidence matrix of the network graph[1]. Vector $\boldsymbol{b} \in \mathbb{R}^N$ is a function of $\boldsymbol{p}$, $\boldsymbol{q}$ and v_0, which reflects the impacts of active power injections of WTs, external grids and nonlinear nature of ac power flow relation. It can be computed with the knowledge of real-time measurements of $\boldsymbol{v}$ and $\boldsymbol{q}$.

Remark 1 (Sparsity of Matrix $\boldsymbol{M}$): The weighted Laplacian matrix $\boldsymbol{M}$ is determined by the physical electrical network. $m_{ij} = \sum_l a_{il}X_{ll}^{-1}a_{jl}$ is non-zero only for the nodes i and j which are adjacent nodes because $a_{il}a_{jl} = 0, \forall l \in \mathcal{E}$ definitely holds for nodes i and j which are not adjacent. Thus, $\boldsymbol{M}$ is inherently sparse for the radial power collector network. This key feature indicates the power flow constraint in the coming voltage control problem is sparsely coupling, so that the dual decomposition can be applied to solve the problem in a distributed fashion, only relying on information exchange between immediate neighbors.

B. VAR Capability of Wind Farm

In this work, modern WTs (Type 3 and Type 4) and STATCOMs are considered to provide VAR support. For a full-scale-converter (FSC)-WT, the VAR capability depends on the capacity limit of the GSC [34]. For a doubly-fed induction generator (DFIG)-WT, both the stator and grid-side converter (GSC) can provide VAR support [35]. The VAR limit of stator side depends on the stator and rotor current limits considering the heating due to the winding Joule losses, as well as the steady state stability limit (see [36] for details). The VAR capability of GSC is limited by the converter capacity. Further, we assume that the GSC has the higher priority in providing VAR support in this study. Once the required VAR exceeds the capability limit of the GSC, the DFIG is controlled to provide the extra VAR. Moreover, in the control loop of the GSC, the active power current has the higher priority compared with

[1]First, let $\bar{\boldsymbol{A}} \in \{0,\pm1\}^{N\times(N+1)}$ be the full incidence matrix of the network graph and then, let $\boldsymbol{a}_0^T$ be the first row of $\bar{\boldsymbol{A}}$ and thus $\bar{\boldsymbol{A}} = [\boldsymbol{a}_0, \boldsymbol{A}^T]^T$.

the VAR current, which aims to stabilize the dc voltage of the back-to-back dc system. The voltage source converter-based STATCOM has a similar structure and control loops with the GSC of WTs and therefore, its VAR limit depends on the converter capacity. Then, the VAR limit at each node can be uniformly represented by,

$$q_{\min,i} \le q_i \le q_{\max,i}, \quad \forall i \in \mathcal{N} \tag{3}$$

where $q_{\min,i}$ and $q_{\max,i}$ denote the min/max VAR limits.

In this paper, for a WT, active power is considered to have higher priority than reactive power. Thus, the VAR capabilities of WTs are closely related to their active power outputs and should be estimated in real time for online control. Without loss of generality, we assume that all the WTs operate at the maximum power point tracking mode.

C. Cyber Layer

We assume each WT and STATCOM corresponds to an intelligent agent in the cyber layer. Each agent has computation, sensing and communication capabilities and can also actuate WTs or STATCOM by setting their power output reference signals. However, the agents are only allowed to communicate with neighbors via communication networks that could be very similar to the power networks due to the special layout of wind farms. Besides, it is assumed that agent i is aware of its neighbor set $\mathcal{N}_i$.

III. Problem Formulation

We first present the static voltage control problem (primal problem (P)) formulation with the aim of minimizing the voltage deviations and maintaining adequate VAR reserves by optimally commanding VAR outputs of WTs and STATCOM, which can be compactly formulated as,

$$\underset{\boldsymbol{v},\boldsymbol{q}}{\text{minimize}} \quad f(\boldsymbol{v},\boldsymbol{q}) := \frac{1}{2}\,||\boldsymbol{v}-\boldsymbol{v}_r||^2_{\boldsymbol{H}_v} + \frac{1}{2}\,||\boldsymbol{q}-\boldsymbol{q}_{\rm mid}||^2_{\boldsymbol{H}_q} \tag{4a}$$

$$\text{subject to} \quad \boldsymbol{M}\boldsymbol{v} = \boldsymbol{q} + \boldsymbol{b} \tag{4b}$$

$$\boldsymbol{v}_{\min} \le \boldsymbol{v} \le \boldsymbol{v}_{\max} \tag{4c}$$

$$\boldsymbol{q}_{\min} \le \boldsymbol{q} \le \boldsymbol{q}_{\max} \tag{4d}$$

where the middle VAR level is defined by $\boldsymbol{q}_{\rm mid} := (\boldsymbol{q}_{\min} + \boldsymbol{q}_{\max})/2$. The diagonal weighting matrices $\boldsymbol{H}_v := \text{diag}\{h_{v,1},\ldots,h_{v,N}\}$ and $\boldsymbol{H}_q := \text{diag}\{h_{q,1},\ldots,h_{q,N}\}$ are defined to be positive definite so that f is strictly convex.

The first term in (4a) is to minimize the voltage deviations from the references. Not only the voltage at the PCC but also the WT terminal voltages are optimized since they should be kept within a secure range otherwise the WT will be tripped by its protection system. Since the protection configuration is typically set as [0.9,1.1] p.u., $\boldsymbol{v}_r = 1.0^2\mathbf{1}_N$ provides adequate downward and upward margins for possible voltage deviations. Another potential benefit of a flat voltage profile is that it could also result in network loss reduction [37]. The second term in (4a) is designed to drive the VAR output of each unit to its middle level. The reasons are twofold. First, to tackle potential large disturbance of the system, VAR reserves should be maintained. Second, $\boldsymbol{q}_{\rm mid}$ generally equals zero (for FSC-WTs and STATCOMs) or approximates to zero (for DFIG-WTs). This results in VAR reduction from WTs and STATCOM and thus reduces inner losses and network losses.

Remark 2 (Voltage Constraint Concerns): In above formulation, we give a general problem formulation including voltage constraints, which can be removed if needed. There may be a concern that the existence of voltage constraints may lead to primal infeasibility in some cases, e.g., when VAR capabilities of WTs are very small due to the high production or a voltage rise/drop suddenly occurs caused by external systems. However, interestingly, the proposed iteration-based algorithm could still asymptotically drive the system towards a better operating point instead of being invalid or even unstable.

IV. Distributed Solution Framework

In this section, the static problem (P) will be used for developing the distributed online control algorithms to adapt to time-varying system operating conditions. To facilitate algorithm design and theoretic analysis, we assume $\boldsymbol{b}$ which captures the uncontrollable disturbances is fixed but this is not required when applying our method.

A. Dual Problem

The Lagrangian function of problem (P) is given by,

$$\begin{aligned}\mathcal{L}(\boldsymbol{v},\boldsymbol{q},\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\omega}) = {} & f(\boldsymbol{v},\boldsymbol{q}) + \boldsymbol{\omega}^T(\boldsymbol{M}\boldsymbol{v}-\boldsymbol{q}-\boldsymbol{b}) \\ & + \boldsymbol{\lambda}^T_{\max}(\boldsymbol{v}-\boldsymbol{v}_{\max}) + \boldsymbol{\lambda}^T_{\min}(\boldsymbol{v}_{\min}-\boldsymbol{v}) \\ & + \boldsymbol{\mu}^T_{\max}(\boldsymbol{q}-\boldsymbol{q}_{\max}) + \boldsymbol{\mu}^T_{\min}(\boldsymbol{q}_{\min}-\boldsymbol{q})\end{aligned} \tag{5}$$

Accordingly, from $\nabla_{\{\boldsymbol{v},\boldsymbol{q}\}}\mathcal{L} = 0$, we have

$$d(\boldsymbol{\nu}) := \inf_{\boldsymbol{v},\boldsymbol{q}} \mathcal{L} = -\frac{1}{2}\boldsymbol{\nu}^T\boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{\Phi}^T\boldsymbol{\nu} - (\boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{g}+\boldsymbol{z})^T\boldsymbol{\nu} \tag{6}$$

where

$$\boldsymbol{\nu} := \begin{bmatrix}\boldsymbol{\lambda}_{\max}\\ \boldsymbol{\lambda}_{\min}\\ \boldsymbol{\mu}_{\max}\\ \boldsymbol{\mu}_{\min}\\ \boldsymbol{\omega}\end{bmatrix}, \boldsymbol{\Phi} := \begin{bmatrix}\boldsymbol{I}_N & \boldsymbol{0}_N\\ -\boldsymbol{I}_N & \boldsymbol{0}_N\\ \boldsymbol{0}_N & \boldsymbol{I}_N\\ \boldsymbol{0}_N & -\boldsymbol{I}_N\\ \boldsymbol{M} & -\boldsymbol{I}_N\end{bmatrix}, \boldsymbol{z} := \begin{bmatrix}\boldsymbol{v}_{\max}\\ -\boldsymbol{v}_{\min}\\ \boldsymbol{q}_{\max}\\ -\boldsymbol{q}_{\min}\\ \boldsymbol{b}\end{bmatrix}$$

$$\boldsymbol{H} := \begin{bmatrix}\boldsymbol{H}_v & \\ & \boldsymbol{H}_q\end{bmatrix}, \quad \boldsymbol{g} := -\begin{bmatrix}\boldsymbol{H}_v & \\ & \boldsymbol{H}_q\end{bmatrix}\begin{bmatrix}\boldsymbol{v}_r\\ \boldsymbol{q}_{\rm mid}\end{bmatrix}.$$

Then, the dual problem (D) is given by

$$\underset{\boldsymbol{\lambda}_{\max},\boldsymbol{\lambda}_{\min},\boldsymbol{\mu}_{\max},\boldsymbol{\mu}_{\min}\ge 0,\,\boldsymbol{\omega}}{\text{maximize}} \; d(\boldsymbol{\nu}). \tag{7}$$

Given that f is strictly convex and the constraint is convex, d is differentiable and its gradient is

$$\nabla d(\boldsymbol{\nu}) = \boldsymbol{\Phi}\begin{bmatrix}\boldsymbol{v}^*(\boldsymbol{\nu})\\ \boldsymbol{q}^*(\boldsymbol{\nu})\end{bmatrix} - \boldsymbol{z} \tag{8}$$

where

$$\{\boldsymbol{v}^*(\boldsymbol{\nu}),\boldsymbol{q}^*(\boldsymbol{\nu})\} := \underset{\boldsymbol{v},\boldsymbol{q}}{\arg\min} \; \mathcal{L}(\boldsymbol{v},\boldsymbol{q},\boldsymbol{\nu})$$

which has the closed-form solution as,

$$\boldsymbol{v}^*(\boldsymbol{\nu}) = \boldsymbol{v}_r - \boldsymbol{H}_v^{-1}\left(\boldsymbol{\lambda}_{\max} - \boldsymbol{\lambda}_{\min} + \boldsymbol{M}^T\boldsymbol{\omega}\right) \tag{9a}$$

$$\boldsymbol{q}^*(\boldsymbol{\nu}) = \boldsymbol{q}_{\rm mid} - \boldsymbol{H}_q^{-1}\left(\boldsymbol{\mu}_{\max} - \boldsymbol{\mu}_{\min} - \boldsymbol{\omega}\right). \tag{9b}$$

B. Generalized Fast Dual Ascent

The traditional gradient methods, e.g., steepest descent, can be used to solve problem (D), referred to as **standard dual ascent** (s-DA) [27, Ch.7]. However, as mentioned before, it suffers from slow convergence if the dual is ill-conditioned. Therefore, we propose to exploit the fast gradient method [32], provided a tight Lipschitz condition of ∇d can be found.

By (6), we have

$$\left\|-\nabla d(\boldsymbol{\nu}^{(1)})+\nabla d(\boldsymbol{\nu}^{(2)})\right\|_2 \le \left\|\boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{\Phi}^T\right\|_2 \left\|\boldsymbol{\nu}^{(1)}-\boldsymbol{\nu}^{(2)}\right\|_2 \tag{10}$$

which holds for any $\boldsymbol{\nu}^{(1)},\boldsymbol{\nu}^{(2)}$. This implies $-d$ has a Lipschitz continuous gradient with constant $L=\|\boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{\Phi}^T\|_2$, which motivates us to use the fast gradient (projection) method [32] for solving problem (D), and is referred to as **fast dual ascent** (f-DA). As per the descent lemma [27, Prop. B.3], this yields a quadratic lower bound of d with the scalar L. The basic idea of fast gradient method is maximizing this bound in each update instead of directly maximizing d. However, to better capture the shape of d for faster convergence, a tighter bound in this problem can be obtained by generalizing the Lipschitz constant $\|\boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{\Phi}^T\|_2$ to $\boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{\Phi}^T$, which yields

$$\begin{aligned} d(\boldsymbol{\nu}^{(1)}) \ge & d(\boldsymbol{\nu}^{(2)}) \\ & + \nabla d(\boldsymbol{\nu}^{(2)})^T(\boldsymbol{\nu}^{(1)}-\boldsymbol{\nu}^{(2)}) - \frac{1}{2}\left\|\boldsymbol{\nu}^{(1)}-\boldsymbol{\nu}^{(2)}\right\|_{\boldsymbol{L}}^2 \end{aligned} \tag{11}$$

holds for any $\boldsymbol{\nu}^{(1)},\boldsymbol{\nu}^{(2)}$ and any $\boldsymbol{L}\succeq\boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{\Phi}^T\in\mathbb{R}^{5N\times 5N}$. The right-hand side provides such a (generalized) lower bound for d. The fast gradient method based on this bound is used to solve problem (D), and is referred to as **generalized fast dual ascent** (Gf-DA), of which the update rule is presented in Algorithm 1 with a compact expression.

C. Distributed Implementation

Observe that in Gf-DA, (S3) and (S4) are naturally decomposable. Hence, the challenges would be how to establish a distributed implementation of (S1) and (S2) and how to choose matrix $\boldsymbol{L}$, which will be detailed as follows.

First, rewrite the Lagrangian as,

$$\begin{aligned} \mathcal{L}(\boldsymbol{v},\boldsymbol{q},\boldsymbol{\nu}) &= \sum_{i=1}^N \mathcal{L}_i(v_i,q_i,\lambda_{\max,i},\lambda_{\min,i},\mu_{\max,i},\mu_{\min,i},\boldsymbol{\omega}_{\mathcal{N}_i}) \\ &= \sum_{i=1}^N \Big\{ f_i(v_i,q_i)+\boldsymbol{\omega}_{\mathcal{N}_i}^T\boldsymbol{M}_{\mathcal{N}_i}v_i-\omega_i q_i-\omega_i b_i \\ &\quad +\lambda_{\max,i}(v_i-v_{\max,i})+\lambda_{\max,i}(v_{\min,i}-v_i) \\ &\quad +\mu_{\max,i}(q_i-q_{\max,i})+\mu_{\min,i}(q_{\min,i}-q_i)\Big\} \end{aligned} \tag{12}$$

where $\boldsymbol{\omega}_{\mathcal{N}_i}$ is the subvector of $\boldsymbol{\omega}$ consisting of $\omega_j, j\in\mathcal{N}_i\bigcup\{i\}$ and $\boldsymbol{M}_{\mathcal{N}_i}$ denotes the submatrix of $\boldsymbol{M}$ consisting of $m_{ji}, j\in\mathcal{N}_i\bigcup\{i\}$. Observe that, the operation in (S1) can be performed in parallel by minimizing the local Lagrangian function $\mathcal{L}_i$, if $\omega_j, j\in\mathcal{N}_i$, is known. Then, the gradient in (S2) can be rewritten in a separable form (c.f. (8)),

$$\begin{aligned} \nabla_{\lambda_{\max,i}} d(\boldsymbol{\nu}) &= v_i^*(\lambda_{\max,i},\lambda_{\min,i},\mu_{\max,i},\mu_{\min,i},\boldsymbol{\omega}_{\mathcal{N}_i}) - v_{\max,i} \\ \nabla_{\lambda_{\min,i}} d(\boldsymbol{\nu}) &= v_{\min,i} - v_i^*(\lambda_{\max,i},\lambda_{\min,i},\mu_{\max,i},\mu_{\min,i},\boldsymbol{\omega}_{\mathcal{N}_i}) \end{aligned}$$

Algorithm 1 Generalized Fast Dual Ascent for Problem (D)

Initialization: Set $\boldsymbol{\lambda}(1)=\boldsymbol{\xi}(0),\boldsymbol{\mu}(1)=\boldsymbol{\eta}(0)\ge 0$ and $\gamma(1)=1$.

For $t\ge 1$ Alternately update primal and dual variables by following steps (S1)–(S4):

$$\text{S1}: \{\boldsymbol{v}(t+1),\boldsymbol{q}(t+1)\} \leftarrow \arg\min_{\boldsymbol{v},\boldsymbol{q}} \ \mathcal{L}(\boldsymbol{v},\boldsymbol{q},\boldsymbol{\nu}(t))$$

$$\text{S2}: \begin{cases} \boldsymbol{\eta}_{\max}(t) \leftarrow \left[\boldsymbol{\lambda}_{\max}(t)+[\boldsymbol{L}^{-1}]_{\boldsymbol{\lambda}_{\max}}\cdot\nabla_{\boldsymbol{\lambda}_{\max}} d(\boldsymbol{\nu}(t))\right]_0^{\infty} \\ \boldsymbol{\eta}_{\min}(t) \leftarrow \left[\boldsymbol{\lambda}_{\min}(t)+[\boldsymbol{L}^{-1}]_{\boldsymbol{\lambda}_{\min}}\cdot\nabla_{\boldsymbol{\lambda}_{\min}} d(\boldsymbol{\nu}(t))\right]_0^{\infty} \\ \boldsymbol{\xi}_{\max}(t) \leftarrow \left[\boldsymbol{\mu}_{\max}(t)+[\boldsymbol{L}^{-1}]_{\boldsymbol{\mu}_{\max}}\cdot\nabla_{\boldsymbol{\mu}_{\max}} d(\boldsymbol{\nu}(t))\right]_0^{\infty} \\ \boldsymbol{\xi}_{\min}(t) \leftarrow \left[\boldsymbol{\mu}_{\min}(t)+[\boldsymbol{L}^{-1}]_{\boldsymbol{\mu}_{\min}}\cdot\nabla_{\boldsymbol{\mu}_{\min}} d(\boldsymbol{\nu}(t))\right]_0^{\infty} \\ \boldsymbol{\chi}(t) \leftarrow \boldsymbol{\omega}(t)+[\boldsymbol{L}^{-1}]_{\boldsymbol{\omega}}\cdot\nabla_{\boldsymbol{\omega}} d(\boldsymbol{\nu}(t)) \end{cases}$$

$$\text{S3}: \gamma(t+1) \leftarrow \frac{1+\sqrt{1+4\gamma(t)^2}}{2}, \text{ and } \varepsilon(t)=\frac{\gamma(t)-1}{\gamma(t+1)}$$

$$\text{S4}: \begin{cases} \boldsymbol{\lambda}_{\max}(t+1) \leftarrow \boldsymbol{\eta}_{\max}(t)+\varepsilon(t)(\boldsymbol{\eta}_{\max}(t)-\boldsymbol{\eta}_{\max}(t-1)) \\ \boldsymbol{\lambda}_{\min}(t+1) \leftarrow \boldsymbol{\eta}_{\min}(t)+\varepsilon(t)(\boldsymbol{\eta}_{\min}(t)-\boldsymbol{\eta}_{\min}(t-1)) \\ \boldsymbol{\mu}_{\max}(t+1) \leftarrow \boldsymbol{\xi}_{\max}(t)+\varepsilon(t)(\boldsymbol{\xi}_{\max}(t)-\boldsymbol{\xi}_{\max}(t-1)) \\ \boldsymbol{\mu}_{\min}(t+1) \leftarrow \boldsymbol{\xi}_{\min}(t)+\varepsilon(t)(\boldsymbol{\xi}_{\min}(t)-\boldsymbol{\xi}_{\min}(t-1)) \\ \boldsymbol{\omega}(t+1) \leftarrow \boldsymbol{\chi}(t)+\varepsilon(t)(\boldsymbol{\chi}(t)-\boldsymbol{\chi}(t-1)) \end{cases}$$

where $[*]_{\boldsymbol{a}}^{\boldsymbol{b}}$ denotes the projection operation onto the constraint set $[\boldsymbol{a},\boldsymbol{b}]$; $[\boldsymbol{L}^{-1}]_{\boldsymbol{\lambda}_{\max}}$, $[\boldsymbol{L}^{-1}]_{\boldsymbol{\lambda}_{\min}}$, $[\boldsymbol{L}^{-1}]_{\boldsymbol{\mu}_{\max}}$, $[\boldsymbol{L}^{-1}]_{\boldsymbol{\mu}_{\min}}$ and $[\boldsymbol{L}^{-1}]_{\boldsymbol{\omega}}$ are the submatrices of $\boldsymbol{L}^{-1}$ consisting of the rows corresponding to each dual vector.

$$\begin{aligned} \nabla_{\mu_{\max,i}} d(\boldsymbol{\nu}) &= q_i^*(\lambda_{\max,i},\lambda_{\min,i},\mu_{\max,i},\mu_{\min,i},\boldsymbol{\omega}_{\mathcal{N}_i}) - q_{\max,i} \\ \nabla_{\mu_{\min,i}} d(\boldsymbol{\nu}) &= q_{\min,i} - q_i^*(\lambda_{\max,i},\lambda_{\min,i},\mu_{\max,i},\mu_{\min,i},\boldsymbol{\omega}_{\mathcal{N}_i}) \\ \nabla_{\omega_i} d(\boldsymbol{\nu}) &= \sum_{j\in\mathcal{N}_i\bigcup\{i\}} m_{ij} v_j^*(\lambda_{\max,i},\lambda_{\min,i},\mu_{\max,i},\mu_{\min,i},\boldsymbol{\omega}_{\mathcal{N}_i}) \\ &\quad - q_i^*(\lambda_{\max,i},\lambda_{\min,i},\mu_{\max,i},\mu_{\min,i},\boldsymbol{\omega}_{\mathcal{N}_i}) - b_i \end{aligned}$$

which can be computed in distributed manner as long as v_i^*, q_i^* and $v_j^*, j\in\mathcal{N}_i$ are available. Notice that, the above analysis has been sufficient to achieve a distributed implementation of s-DA.

For our Gf-DA, the last remaining work is the selection of matrix $\boldsymbol{L}$. According to (11), $\boldsymbol{L}=\boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{\Phi}^T$ provides the tightest bound, which guarantees the best convergence. Unfortunately, this selection is not allowed since $\boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{\Phi}^T$ is not invertible[2]. Hence, we first restrict $\boldsymbol{L}$ to be a diagonal matrix, i.e., $\boldsymbol{L} := \text{blkdiag}(\boldsymbol{L}_{\lambda_{\max}},\boldsymbol{L}_{\lambda_{\min}},\boldsymbol{L}_{\mu_{\max}},\boldsymbol{L}_{\mu_{\min}},\boldsymbol{L}_{\omega})$ where $\boldsymbol{L}_{\lambda_{\max}},\boldsymbol{L}_{\lambda_{\min}},\boldsymbol{L}_{\mu_{\max}},\boldsymbol{L}_{\mu_{\min}},\boldsymbol{L}_{\omega}\in\mathbb{R}^{N\times N}$ represent the diagonal submatrices associated with $\boldsymbol{\lambda}_{\max}$, $\boldsymbol{\lambda}_{\min}$, $\boldsymbol{\mu}_{\max}$, $\boldsymbol{\mu}_{\min}$ and $\boldsymbol{\omega}$, respectively. Then, $\boldsymbol{L}$ can be computed by solving the following semi-definite programming problem,

$$\underset{\boldsymbol{L}}{\text{minimize}} \ \text{Trace}(\boldsymbol{L}) \tag{13a}$$

$$\text{subject to} \ \ \boldsymbol{L}\succeq\boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{\Phi}^T. \tag{13b}$$

Then, the elements of $\boldsymbol{L}$ associated with unit i are selected to

[2]Obviously, there must be linear dependent rows in $\boldsymbol{\Phi}$. So, we have $\boldsymbol{\Phi}=\left[\bar{\boldsymbol{\Phi}}^T,\boldsymbol{0}^T\right]^T$ via row transformations where matrix $\boldsymbol{P}$ is invertible. Then,

$$\boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{\Phi}^T=\boldsymbol{P}\begin{bmatrix}\bar{\boldsymbol{\Phi}}\boldsymbol{H}^{-1}\bar{\boldsymbol{\Phi}}^T & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\boldsymbol{P}^T.$$

This indicates it is not invertible.

TABLE I
CONVERGENCE RATES COMPARISON

Proposition	Description	s-DA	Gf-DA
1	Dual Function	$d(\boldsymbol{\nu}^\star) - d(\boldsymbol{\nu}(t)) \leq \frac{C_2}{2\rho(\boldsymbol{\Lambda})t}$	$d(\boldsymbol{\nu}^\star) - d(\boldsymbol{\nu}(t)) \leq \frac{2C_L}{(t+1)^2}$
2	Primal Sequence	$\left\| \begin{bmatrix} \boldsymbol{v}^*(t) \\ \boldsymbol{q}^*(t) \end{bmatrix} - \begin{bmatrix} \boldsymbol{v}^\star \\ \boldsymbol{q}^\star \end{bmatrix} \right\|_2 \leq \frac{\sqrt{C_2}}{\sqrt{\sigma_{\min}(\boldsymbol{H})\rho(\boldsymbol{\Lambda})}\sqrt{t}}$	$\left\| \begin{bmatrix} \boldsymbol{v}^*(t) \\ \boldsymbol{q}^*(t) \end{bmatrix} - \begin{bmatrix} \boldsymbol{v}^\star \\ \boldsymbol{q}^\star \end{bmatrix} \right\|_2 \leq \frac{2\sqrt{C_L}}{\sqrt{\sigma_{\min}(\boldsymbol{H})}(t+1)}$
3	Primal Infeasibility	$\|\boldsymbol{M}\boldsymbol{v}^*(t) - \boldsymbol{q}^*(t) - \boldsymbol{b}\|_\infty \leq \frac{\left\|[\boldsymbol{M}, -\boldsymbol{I}_N]\right\|_\infty \sqrt{C_2}}{\sqrt{\sigma_{\min}(\boldsymbol{H})\rho(\boldsymbol{\Lambda})}\sqrt{t}}$	$\|\boldsymbol{M}\boldsymbol{v}^*(t) - \boldsymbol{q}^*(t) - \boldsymbol{b}\|_\infty \leq \frac{2\left\|[\boldsymbol{M}, -\boldsymbol{I}_N]\right\|_\infty \sqrt{C_L}}{\sqrt{\sigma_{\min}(\boldsymbol{H})}(t+1)}$
		$\left[\begin{bmatrix} \boldsymbol{v}^*(t) - \boldsymbol{v}_{\max} \\ \boldsymbol{v}_{\min} - \boldsymbol{v}^*(t) \\ \boldsymbol{q}^*(t) - \boldsymbol{q}_{\max} \\ \boldsymbol{q}_{\min} - \boldsymbol{q}^*(t) \end{bmatrix}\right]_0^\infty \leq \left(\frac{2\sqrt{C_2}}{\sqrt{\sigma_{\min}(\boldsymbol{H})\rho(\boldsymbol{\Lambda})}\sqrt{t}}\right)\mathbf{1}_{4N}$	$\left[\begin{bmatrix} \boldsymbol{v}^*(t) - \boldsymbol{v}_{\max} \\ \boldsymbol{v}_{\min} - \boldsymbol{v}^*(t) \\ \boldsymbol{q}^*(t) - \boldsymbol{q}_{\max} \\ \boldsymbol{q}_{\min} - \boldsymbol{q}^*(t) \end{bmatrix}\right]_0^\infty \leq \left(\frac{2\sqrt{C_L}}{\sqrt{\sigma_{\min}(\boldsymbol{H})}(t+1)}\right)\mathbf{1}_{4N}$

establish $\boldsymbol{L}_i := \text{diag}(L_{\lambda_{\max,i}}, L_{\lambda_{\min,i}}, L_{\mu_{\max,i}}, L_{\mu_{\min,i}}, L_{\omega_i})$, which should be stored by unit i. Since $\boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{\Phi}^T$ is only system-dependent, $\boldsymbol{L}$ can be offline computed using (13).[3]

A simple example is provided in Appendix A for better understanding and following the distributed Gf-DA algorithm.

D. Simplified and Enhanced Variant

In Gf-DA, all the constraints are relaxed so that a smooth case is obtained with explicit d, which makes it easier to conduct some analyses. However, such relaxation requires updating and storing many dual variables. Inspired by the advanced studies of gradient methods on non-smooth cases [39]–[40], we further propose a simplified but enhanced version of Gf-DA based on the equivalent problem of (4) with an extended-value objective function (eliminating the inequality constraints) and power flow constraint, referred to as Ef-DA. Then, the step of updating primal variables $\boldsymbol{v}$ and $\boldsymbol{q}$ in (S1) will be replaced by the constrained minimization with the constraints of VAR and voltage as in (4c) and (4d). In this way, the multipliers corresponding to voltage and VAR constraints, i.e., $\boldsymbol{\lambda}_{\max}, \boldsymbol{\lambda}_{\min}, \boldsymbol{\mu}_{\max}$ and $\boldsymbol{\mu}_{\min}$ are eliminated. The benefits of Ef-DA are: i) the updates of dual variables are significantly simplified with less computation and storage burdens and ii) thanks to fewer Lagrange multipliers and the hard bounds on $\boldsymbol{v}$ and $\boldsymbol{q}$ in each update, a faster convergence rate is expected.

V. CONVERGENCE ANALYSIS

For a generalized s-DA using a diagonally scaled step size matrix $\boldsymbol{\Lambda} \succ 0 \in \mathbb{R}^{5N\times 5N}$, a sufficient condition for convergence on problem (D) is $\rho(\boldsymbol{\Lambda}) \leq 2/\rho(\boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{\Phi}^T)$. This can be proven by generalizing the descent lemma [27, Ch.3] to a weighted case. As discussed before, a prerequisite for using Gf-DA to solve problem (D) is that such a quadratic upper bound exists for any $\boldsymbol{\nu}$, indicating that for any $\boldsymbol{L} \succeq \boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{\Phi}^T$, Gf-DA will converge when applied to solve problem (D).

Beyond the discussions on sufficient conditions, we further compare the convergence rates of s-DA and fast-gradient-based DA on this voltage control problem, including the convergence rates not only on dual function, but also on primal variables $\boldsymbol{v}$ and $\boldsymbol{q}$ which we are more concerned with.

Let $\{\boldsymbol{v}^\star, \boldsymbol{q}^\star\}$ and $\boldsymbol{\nu}^\star$ be the optimal solutions of primal problem (P) and dual problem (D), respectively. Let $\{\boldsymbol{v}^*(t), \boldsymbol{q}^*(t)\}$ be the primal sequence generated by the s-DA or Gf-DA. The (weighted) initial gaps are given by,

$$C_2 = \|\boldsymbol{\nu}(0) - \boldsymbol{\nu}^\star\|_2^2\,,\; C_L = \|\boldsymbol{\nu}(0) - \boldsymbol{\nu}^\star\|_{\boldsymbol{L}}^2\,.$$

Suppose, for s-DA, we have $\rho(\boldsymbol{\Lambda}) \leq 1/\rho(\boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{\Phi}^T)$, and for Gf-DA, we have $\boldsymbol{L} \succeq \boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{\Phi}^T$. Then, Propositions 1–3 listed in Table I detail the convergence rates of s-DA and Gf-DA on dual function, primal sequence and primal infeasibility, respectively. The rigorous proofs are provided in Appendix B.

From Propositions 1–3, it can be concluded that i) Gf-DA solves problem (D) with a convergence rate no worse than $O(1/t^2)$, however, s-DA only achieves a rate of $O(1/t)$; ii) With the primal-dual relation, we further know that the primal sequence $\{\boldsymbol{v}^*(t), \boldsymbol{q}^*(t)\}$ converges to the optimal solution with rates of $O(1/t)$ and $O(1/\sqrt{t})$, respectively; iii) Last but not least, the primal infeasibility converges to zero with rates of $O(1/t)$ and $O(1/\sqrt{t})$, respectively. These analyses might shed light on why the proposed method significantly improves the convergence of dual ascent, which is helpful in tracking system voltage variations. Besides, this will also reduce the computation and communication burdens when applied offline.

Remark 3 (Convergence of f-DA and Ef-DA): The convergence analysis for Gf-DA is applicable to the standard f-DA by applying the l_2-norm. Interestingly, the relationship in (11) also holds for the dual function in Ef-DA, given that the resulting equivalent primal problem of (4) satisfies several specific conditions [40]. Thus, the convergence analysis for Gf-DA also works for Ef-DA and the selection method of $\boldsymbol{L}$ can be used for Ef-DA. Besides, given that $\boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{\Phi}^T = [\boldsymbol{M}, -\boldsymbol{I}_N]\boldsymbol{H}^{-1}[\boldsymbol{M}, -\boldsymbol{I}_N]^T = \boldsymbol{M}\boldsymbol{H}_v^{-1}\boldsymbol{M}^T + \boldsymbol{H}_q^{-1}$ is positive definite, the tightest upper bound for d can be achieved by selecting $\boldsymbol{L} = \boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{\Phi}^T$, which is exactly the Hessian matrix of $-d$ and consequently, it can be found that the update rule in (S2) tends to be a constrained version of the pure form of Newton's method [27, Ch. 2]. Hence, if the centralized coordination is allowed to perform the computation of $\boldsymbol{\chi}$, a superlinear convergence is expected.

[3] $\boldsymbol{L}$ can be also online computed in a distributed way based on the distributed version of the generalized Lipschitz condition.

VI. Online Implementation

A. Estimation of b_i

The update of ω_i requires b_i, which reflects the system operation states. For an online and distributed implementation, based on (2), b_i can be estimated online by measuring v_i^m and q_i and sharing the instantaneous voltage measurements v_i^m with neighbors, which is as follows:

$$b_i(t) = \sum_{j\in\mathcal{N}_i\bigcup\{i\}} m_{ij}v_j^m(t) - q_i(t), \quad \forall i \in \mathcal{N}. \tag{14}$$

B. Estimation of $q_{\max,i}$ and $q_{\min,i}$

VAR limits ($q_{\max,i}$ and $q_{\min,i}$) of each WT are locally updated online according to their VAR characteristics. As aforementioned, they are quite different for different types of WTs. For DFIG-WTs, the stator and rotor-side information are required to compute the VAR limits but for FSC-WT, only the active power output is needed. The online update of VAR limits is helpful to guarantee WTs operate within a secure range, especially for preventing converters from overload.

C. Projection of q_i

Since the VAR limit is relaxed in updates, to allow for an online implementation, the reactive power command q_i^c should be projected to the feasible range at each control point, which is as follows:

$$q_i^c(t) = [q_i^*(t)]_{q_{\min,i}(t)}^{q_{\max,i}(t)}, \quad \forall i \in \mathcal{N}. \tag{15}$$

D. Reset γ

Generally, the offline fast gradient method starts with $\gamma = 1$ and then γ is updated in each iteration. For an online implementation, we propose to reset $\gamma \leftarrow 1$ every T_γ time steps to guarantee the tracking capability.

E. Algorithm Design

The detailed rules of DOnVC are presented in Algorithm 2. The schematic diagram of the distributed controller for each DFIG-WT is shown in Fig. 2 where the measurements of the active power output of stator $p_{S,i}^m$, stator voltage $V_{S,i}^m$ and active power output of GSC $p_{C,i}^m$ are locally obtained. The VAR commands of stator $q_{S,i}$ and GSC $q_{C,i}$ are dispatched to the corresponding controllers, respectively. For FSC-WTs, the VAR estimation method is different because of different VAR characteristics and the inner VAR dispatch block does not exist. For STATCOMs, both the VAR estimation and inner VAR dispatch are removed.

Remark 4 (Model Errors): Though we establish the algorithms and also the convergence analyses under a fixed-point condition (linear model), the closed-loop nature of online algorithms with an up-to-date knowledge of $\boldsymbol{b}$ could asymptotically mitigate the model errors caused by the linear approximation. This will be numerically validated later.

Algorithm 2 Distributed Online Voltage Control Algorithm

For any agent i at time t

- **Reset γ:** If t mod $T_\gamma = 0$ then reset $\gamma(t) \leftarrow 1$.
- **Estimate VAR Limit:** Locally update $q_{\max,i}$, $q_{\min,i}$.
- **Share v_i^m:** Receive $v_j^m(t)$ and send $v_i^m(t)$ to $j \in \mathcal{N}_i$.
- **Estimate Operation Point:** Update b_i as in (14).
- **Share ω_i:** Receive ω_j from $j \in \mathcal{N}_i$ and send ω_i to $j \in \mathcal{N}_i$.
- **Update Primal Variables:** Update v_i and q_i by,

$$v_i^*(t+1) \leftarrow v_{r,i} - \frac{\sum_{j\in\mathcal{N}_i\bigcup\{i\}} m_{ji}\omega_j(t) + \lambda_{\max,i}(t) - \lambda_{\min,i}(t)}{h_{v,i}}$$

$$q_i^*(t+1) \leftarrow q_{\mathrm{mid},i}(t) + \frac{\omega_i(t) - \mu_{\max,i}(t) + \mu_{\min,i}(t)}{h_{q,i}}.$$

- **Share v_i^*:** Receive $v_j^*(t+1)$ from $j \in \mathcal{N}_i$ and send $v_i^*(t+1)$ to $j \in \mathcal{N}_i$.
- **Update Dual Variables:** Update $\eta_{\max,i}$, $\eta_{\min,i}$, $\xi_{\max,i}$, $\xi_{\min,i}$ and χ_i as

$$\eta_{\max,i}(t) \leftarrow \left[\lambda_{\max,i}(t) + \frac{v_i^*(t+1) - v_{\max,i}}{L_{\lambda_{\max,i}}}\right]_0^\infty$$

$$\eta_{\min,i}(t) \leftarrow \left[\lambda_{\min,i}(t) + \frac{v_{\min,i} - v_i^*(t+1)}{L_{\lambda_{\min,i}}}\right]_0^\infty$$

$$\xi_{\max,i}(t) \leftarrow \left[\mu_{\max,i}(t) + \frac{q_i^*(t+1) - q_{\max,i}(t)}{L_{\mu_{\max,i}}}\right]_0^\infty$$

$$\xi_{\min,i}(t) \leftarrow \left[\mu_{\min,i}(t) + \frac{q_{\min,i}(t) - q_i^*(t+1)}{L_{\mu_{\min,i}}}\right]_0^\infty$$

$$\chi_i(t) \leftarrow \omega_i(t) + \frac{\sum_{j\in\mathcal{N}_i\bigcup\{i\}} m_{ij}v_j^*(t+1) - q_i^*(t+1) - b_i(t)}{L_{\omega_i}}$$

  and then update $\lambda_{\max,i}$, $\lambda_{\min,i}$, $\mu_{\max,i}$, $\mu_{\min,i}$ and ω_i as in (S4) in Algorithm 1.
- **Implement:** Update the reactive power set-point as in (15).

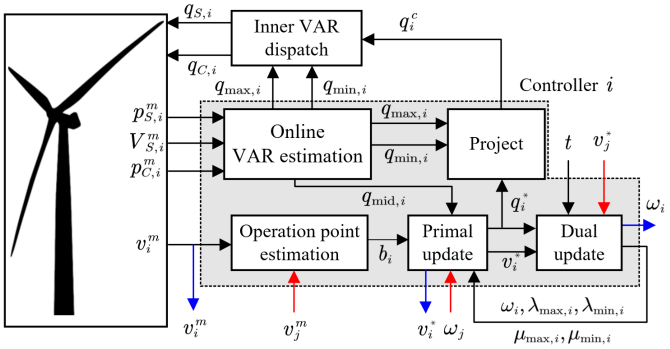


Fig. 2. Schematic diagram of the proposed DOnVC with a DFIG-WT case. (The red arrows represent the received information, the blue arrows represent the sent information and the black arrows represent the local information.)

VII. Numerical Results

In this section, numerical simulations are performed to demonstrate the proposed voltage control schemes. The static performance is tested on the modified French Fécamp offshore wind farm [9] with 83×6 MW FSC-WTs and 1×±30 MVar STATCOM (see Fig. 1). The dynamic simulation is performed on a wind farm with 20×5 MW DFIG-WTs and 1×±10 MVar

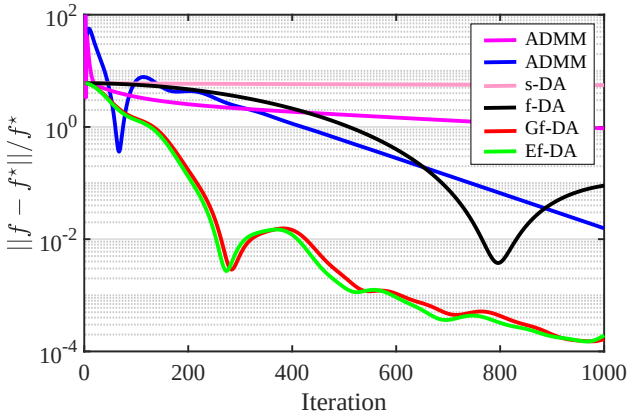


Fig. 3. Convergence performance of different algorithms. ($f^\star$ is the optimal objective value computed using the IPM-based solver IPOPT; For s-DA, the step size is $2/\rho(\boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{\Phi}^T)$; For f-DA, $\boldsymbol{L}=\rho(\boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{\Phi}^T)\boldsymbol{I}_N$; The first ADMM case (rose red) is with $\rho=1$ and the second (blue) is with $\rho=10$).

TABLE II
OBJECTIVE VALUES COMPUTED BY DIFFERENT ALGORITHMS

	Problem	Algorithm	Value	Solver
A1	NLP (P)	IPM	**0.0015216**	IPOPT 3.12
A2	NLP (P)	SQP	0.0015216	FMINCON
A3	NLP (P)	Active Set	0.0015217	FMINCON
A4	QP	IPM	0.0464197	CPLEX 12.9
A5	SOCP	IPM	0.2284767	SeDuMi 1.3
A6	NLP (P)	ADMM (feedback)	0.0015454	[29]
A7	NLP (P)	Gf-DA (feedback)	**0.0015218**	—
A8	NLP (P)	Ef-DA (feedback)	**0.0015219**	—
A9	NLP (P)	Ef-DA (tightest $\boldsymbol{L}$)	**0.0015216**	—

STATCOM [19]. Note that the power flow is computed using the full ac power flow model instead of a linear approximation.

A. Static Performance

1) Convergence Rate: Let $\boldsymbol{\Phi}_v=\boldsymbol{I}_N$, $\boldsymbol{\Phi}_q=5\boldsymbol{I}_N$, $\boldsymbol{v}_{\max}=1.05^2\mathbf{1}_N$ and $\boldsymbol{v}_{\min}=0.95^2\mathbf{1}_N$ the convergence performances (i.e., relative error to optimal objective value) of f-DA, Gf-DA and Ef-DA are compared with s-DA [28] and ADMM [29] (see Fig. 3). Gf-DA and Ef-DA enjoy faster convergence than others on this problem and a slight superiority of Ef-DA is observed compared with Gf-DA. Particularly, s-DA shows very slow convergence because the conventional gradient method (steepest descent) often suffers from very slow convergence when solving ill-conditioned problems. The Gf-DA and Ef-DA can converge to the solutions within 10^{-2} error bound after about 400 iterations. However, ADMM and s-DA require more than 1000 iterations. This implies the proposed algorithm will save communication bandwidth in real-life implementations when achieving the similar control performance. Besides, according to the results, Ef-DA with the tightest bound (i.e., $\boldsymbol{L}=\boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{\Phi}^T$) can compute a solution with a relative error of 10^{-5} using only 8 iterations, validating our previous analysis. Seen from Fig. 4, a larger penalty weighting for VAR output results in a more stable but slower convergence speed. So, there will be a trade-off between stability and convergence rate in practical implementation.

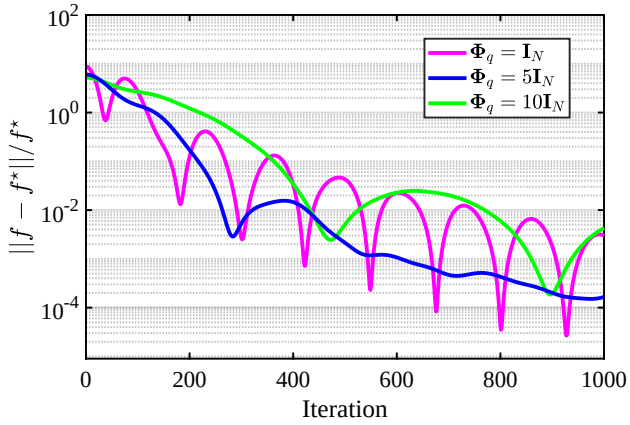
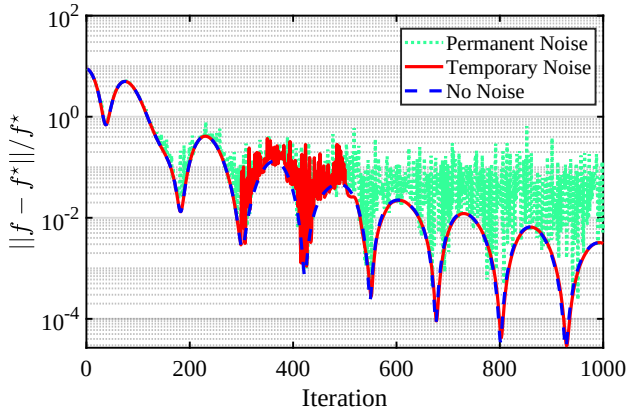
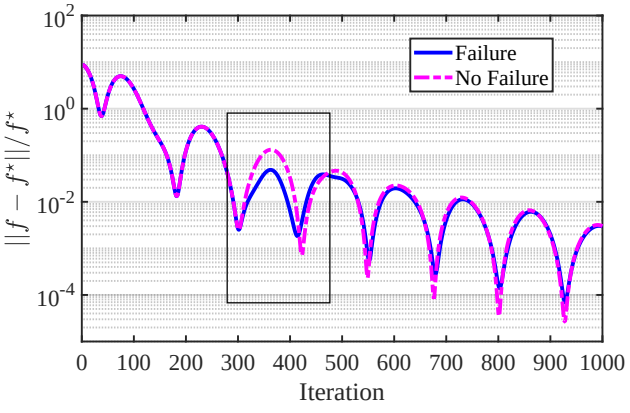


Fig. 4. Convergence performance of Gf-DA with different weighting settings.



(a)



(b)

Fig. 5. Robust performance under a) measurement noise (A random noise following a normal distribution of $N(0,0.01^2)$ is considered on terminal voltage of WT16 and the temporary noise only occurs during 300th–500th iteration.) and b) communication link failure (A temporary communication link failure between WT16 and WT17 is considered during 300th–500th iteration.).

2) Optimality: Table II reports the comparison of solution quality computed by different methods. A1–A3 directly solves the possibly non-convex original problem with full ac power flow constraints using interior point method (IPM), sequential quadratic programming method (SQP) and active-set method, respectively. In A4, the approximate QP problem with linear power flow constraints without voltage feedback is considered. While it is much easier to solve, however, it fails to provide

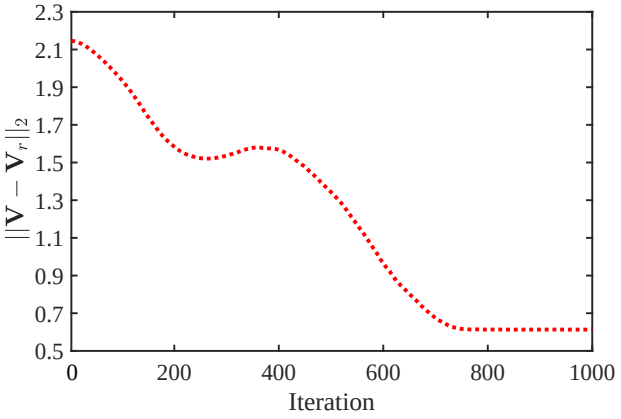


Fig. 6. Voltage control performance under a primal infeasible case.

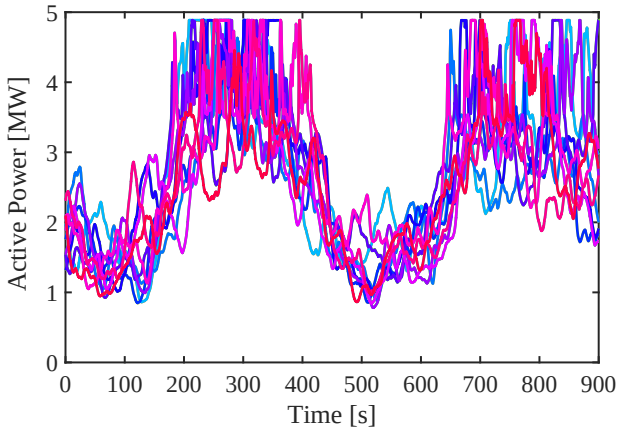


Fig. 7. Active power production of 20 DFIG-WTs.

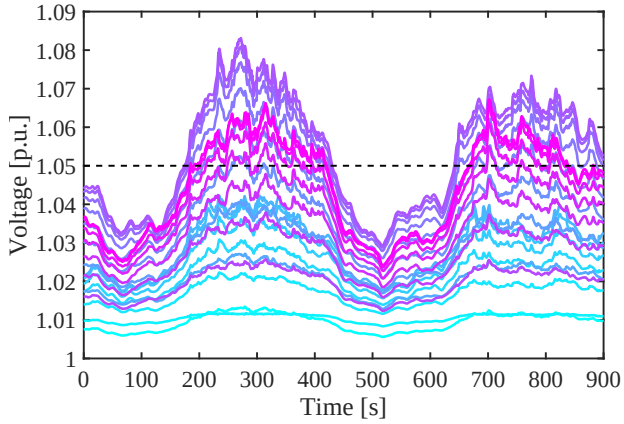


Fig. 8. Voltage profile of the wind farm network without control.

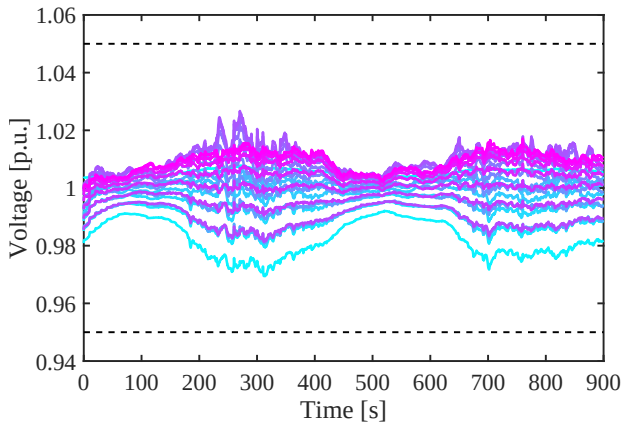


Fig. 9. Voltage profile of the wind farm network with Gf-DA based control.

a high-quality solution due to its open-loop nature. In comparison, ADMM (A6), Gf-DA (A7) and Ef-DA (A8, A9) with online feedback can achieve the very close optimality as the centralized algorithms A1–A3 due to the fact that they are able to asymptotically compensate the model errors via the feedback mechanism. For the SOCP relaxation-based method [41], it is found that the gaps $||l_{ij} - (P_{ij}^2 + Q_{ij}^2)/v_i||$ for most branches are unacceptable, indicating the relaxation is inexact and thus it fails to get a high-quality solution in this problem.

3) Robustness: As shown in Fig. 5, the robustness of the proposed algorithm is tested. In Fig. 5(a), the temporary noise only causes oscillation during that period. The permanent noise will lead to an increasing amplitude oscillation and then the system become unstable due to the error accumulation. That is, for real-life implementation, the measurement errors should be guaranteed to fall within a tiny level. As for communication link failures (see Fig. 5(b)), a "freeze" strategy [28]–[29] is used, i.e., the exchanged information remains unchanged until the new value comes. It can be observed that the proposed algorithm is slightly affected during the failure period but the convergence can still be guaranteed.

4) Primal Infeasibility: An extreme case is designed, in which the total wind power production is 457 MW ($> 90\%$) and an over-voltage condition is created. The centralized solver IPOPT reports that it fails to give an optimal solution. Hence, problem (P) is believed to be primal-infeasible. Seen from Fig. 6, under such case, Gf-DA can still drive system voltages towards a secure operation range and it finally converges to an stable operation point where all bus voltages are corrected within 1.02–1.04 p.u. (acceptable in practical operation). This validates the previous remark regarding voltage constraints.

B. Dynamic Performance

The dynamic test is performed with Matlab/SIMULINK environment. The controllers are implemented using a function block with a packaging triggered subsystem. The wind field modeling considering turbulence and wake effects for the wind farm was generated from SimWindFarm Toolbox [38]. The total simulation time is 15 min. Fig. 7 shows the active power production of 20 WTs. Fig. 8 shows the voltage profile across the network without any additional voltage/VAR control. It can be seen that, some WT terminal voltages exceed 1.05 p.u. and even reach 1.08 p.u., resulting in a high risk of being tripped by the protection devices. In comparison, as shown in Fig. 9, the proposed Gf-DA based DOnVC scheme is able to regulate the voltages within the predefined range [0.95,1.05] p.u. by coordinating the VAR outputs of WTs and STATCOM, which is illustrated in Fig. 10.

The overall voltage control performance is presented in Fig. 11. The control periods of all these methods are set as 0.1 s. It

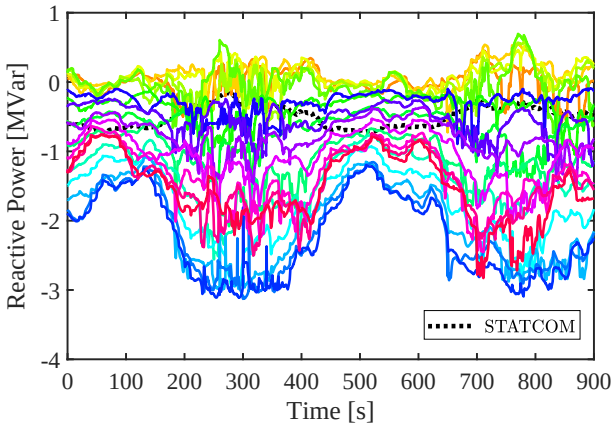


Fig. 10. Reactive power outputs of STATCOM and 20 DFIG-WTs.

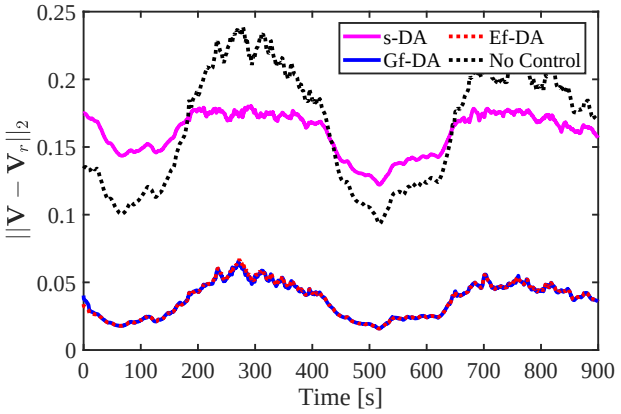


Fig. 11. Overall voltage performance with different control strategies. (For s-DA, the step size is $2/\rho(\boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{\Phi}^T)$; For Ef-DA, the tightest bound $\boldsymbol{L}=\boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{\Phi}^T$ is used; The weighting matrices are designed as $\boldsymbol{\Phi}_v=\boldsymbol{I}_N$ and $\boldsymbol{\Phi}_q=\boldsymbol{I}_N$ for all cases.)

can be found that Gf-DA-based and Ef-DA (with centralized coordination)-based voltage control schemes show much better voltage control performance than the s-DA-based method, though s-DA can alleviate the over-voltage issue in some degree. This is due to the fact that Gf-DA and Ef-DA hold much faster convergence than s-DA, indicating a good tracking capability, which is consistent with the previous discussions on convergence. This further demonstrates the convergence rate of algorithms is of great significance in online voltage control.

VIII. Conclusion

This paper has proposed a fast dual ascent-based distributed online voltage control algorithm and addressed its application in wind farms. The voltage control problem is established based on a linearized branch flow model with sparse couplings among nodes so that the dual decomposition method can be used to develop a fully distributed solution algorithm. To improve the convergence of dual ascent, the fast gradient method is generalized and then applied to solve the dual problem. Theoretical convergence analyses on both sides of dual and primal have been also provided in details. Besides, an online implementation scheme via feedback was designed in order to achieve better tracking performance. Finally, numerical results have validated the effectiveness of the proposed DOnVC methods under both static and dynamic cases.

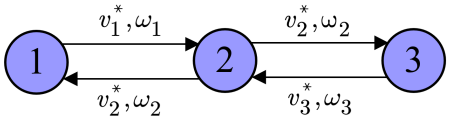


Fig. 12. An example with 3 units. ($\boldsymbol{\omega}_{\mathcal{N}_1}=[\omega_1\,\omega_2]^T$, $\boldsymbol{\omega}_{\mathcal{N}_2}=[\omega_1\,\omega_2\,\omega_3]^T$ and $\boldsymbol{\omega}_{\mathcal{N}_3}=[\omega_2\,\omega_3]^T$; $\boldsymbol{M}_{\mathcal{N}_1}=[m_{11}\,m_{21}]^T$, $\boldsymbol{M}_{\mathcal{N}_2}=[m_{12}\,m_{22}\,m_{32}]^T$ and $\boldsymbol{M}_{\mathcal{N}_3}=[m_{23}\,m_{33}]^T$)

Appendix A
A Simple Example

As shown in Fig. 12, a simple example with 3 units labelled by agents ①, ② and ③, respectively, is provided to clarify some details of the proposed distributed Gf-DA algorithm. The distributed Gf-DA algorithm is achieved by alternately carrying out the following steps.

Step 1 (Primal Update): With the latest updated dual variables $\lambda_{\max,i}$, $\lambda_{\min,i}$, $\mu_{\max,i}$, $\mu_{\min,i}$, and $\omega_{\mathcal{N}_i}$, each distributed controller $i\in\{1,2,3\}$ solves v_i^* and q_i^*, respectively, by

$$\{v_1^*,q_1^*\}:=\arg\min_{v_1,q_1}\ \mathcal{L}_1(v_1,q_1,:) \tag{16}$$

$$\{v_2^*,q_2^*\}:=\arg\min_{v_2,q_2}\ \mathcal{L}_2(v_2,q_2,:) \tag{17}$$

$$\{v_3^*,q_3^*\}:=\arg\min_{v_3,q_3}\ \mathcal{L}_3(v_3,q_3,:) \tag{18}$$

where the closed-form solutions are solved by $\nabla\mathcal{L}_i(\cdot)=0$.

Step 2 (Information Exchange—v_i^*): Each agent exchanges the latest updated v_i^* with neighbors:

$$① \Longrightarrow ② : v_1^* \quad ① \Longleftarrow ② : v_2^*$$
$$② \Longrightarrow ③ : v_2^* \quad ② \Longleftarrow ③ : v_3^*$$

Step 3 (Dual Update): After receiving the latest updated primal variables from neighbors, each agent updates the associated dual variables and auxiliary variables as in the step "*Update Dual Variables*" in Algorithm 2. Then, we obtain the newly updated $\lambda_{\max,i}$, $\lambda_{\min,i}$, $\mu_{\max,i}$, $\mu_{\min,i}$, and ω_i.

Step 4 (Information Exchange—ω_i): Each agent exchanges the updated ω_i with neighbors:

$$① \Longrightarrow ② : \omega_1 \quad ① \Longleftarrow ② : \omega_2$$
$$② \Longrightarrow ③ : \omega_2 \quad ② \Longleftarrow ③ : \omega_3$$

Appendix B
Proofs of Propositions 1–3

A. Proof of Proposition 1

Proof: The first part can be proven by generalizing the descent lemma with Lipschitz constant $\rho(\boldsymbol{\Phi}\boldsymbol{H}^{-1}\boldsymbol{\Phi}^T)$. The second part can be obtained by generalizing the proof in [39], i.e., replacing constant L with $\boldsymbol{L}$-norm. We should also remark that fast gradient methods-based DA does not guarantee a monotone increase of d. ■

B. Proof of Proposition 2

Proof: For any given $\boldsymbol{\nu}$, define

$$g(\boldsymbol{v},\boldsymbol{q};\boldsymbol{\nu}) := \mathcal{L}(\boldsymbol{v},\boldsymbol{q},\boldsymbol{\nu}) \tag{19}$$

which implies $\boldsymbol{\nu}$ is parametric in g. Since f is strongly convex with $\boldsymbol{H} \succ 0$, so is the function g. Therefore, we have

$$g(\boldsymbol{v}^*(\boldsymbol{\nu}),\boldsymbol{q}^*(\boldsymbol{\nu});\boldsymbol{\nu}) - g(\boldsymbol{v},\boldsymbol{q};\boldsymbol{\nu}) \geq \frac{1}{2}\left\|\begin{bmatrix}\boldsymbol{v}^*(\boldsymbol{\nu})\\ \boldsymbol{q}^*(\boldsymbol{\nu})\end{bmatrix} - \begin{bmatrix}\boldsymbol{v}^\star\\ \boldsymbol{q}^\star\end{bmatrix}\right\|_{\boldsymbol{H}}^2 \tag{20}$$

with $\{\boldsymbol{v}^*(\boldsymbol{\nu}),\boldsymbol{q}^*(\boldsymbol{\nu})\} := \arg\min_{\boldsymbol{v},\boldsymbol{q}} g(\boldsymbol{v},\boldsymbol{q};\boldsymbol{\nu})$. Then, by the definition of g, for any $\boldsymbol{\lambda}_{\max},\boldsymbol{\lambda}_{\min},\boldsymbol{\mu}_{\max},\boldsymbol{\mu}_{\min} \geq 0$ and any $\boldsymbol{\omega}$, we have,

$$\begin{aligned}
&g(\boldsymbol{v}^\star,\boldsymbol{q}^\star;\boldsymbol{\nu}) - g(\boldsymbol{v}^*(\boldsymbol{\nu}),\boldsymbol{q}^*(\boldsymbol{\nu});\boldsymbol{\nu})\\
&\quad= f(\boldsymbol{v}^\star,\boldsymbol{q}^\star) + \boldsymbol{\omega}^T(\boldsymbol{M}\boldsymbol{v}^\star - \boldsymbol{q}^\star - \boldsymbol{b})\\
&\qquad+ \boldsymbol{\lambda}_{\max}^T(\boldsymbol{v}^\star - \boldsymbol{v}_{\max}) + \boldsymbol{\lambda}_{\min}^T(\boldsymbol{v}_{\min} - \boldsymbol{v}^\star)\\
&\qquad+ \boldsymbol{\mu}_{\max}^T(\boldsymbol{q}^\star - \boldsymbol{q}_{\max}) + \boldsymbol{\mu}_{\min}^T(\boldsymbol{q}_{\min} - \boldsymbol{q}^\star)\\
&\qquad- f(\boldsymbol{v}^*(\boldsymbol{\nu}),\boldsymbol{q}^*(\boldsymbol{\nu})) - \boldsymbol{\omega}^T(\boldsymbol{M}\boldsymbol{v}^*(\boldsymbol{\nu}) - \boldsymbol{q}^*(\boldsymbol{\nu}) - \boldsymbol{b})\\
&\qquad- \boldsymbol{\lambda}_{\max}^T(\boldsymbol{v}^*(\boldsymbol{\nu}) - \boldsymbol{v}_{\max}) - \boldsymbol{\lambda}_{\min}^T(\boldsymbol{v}_{\min} - \boldsymbol{v}^*(\boldsymbol{\nu}))\\
&\qquad- \boldsymbol{\mu}_{\max}^T(\boldsymbol{q}^*(\boldsymbol{\nu}) - \boldsymbol{q}_{\max}) - \boldsymbol{\mu}_{\min}^T(\boldsymbol{q}_{\min} - \boldsymbol{q}^*(\boldsymbol{\nu}))\\
&\quad= d(\boldsymbol{\nu}^\star) - d(\boldsymbol{\nu}) + \boldsymbol{\omega}^T(\boldsymbol{M}\boldsymbol{v}^\star - \boldsymbol{q}^\star - \boldsymbol{b})\\
&\qquad+ \boldsymbol{\lambda}_{\max}^T(\boldsymbol{v}^\star - \boldsymbol{v}_{\max}) + \boldsymbol{\lambda}_{\min}^T(\boldsymbol{v}_{\max} - \boldsymbol{v}^\star)\\
&\qquad+ \boldsymbol{\mu}_{\max}^T(\boldsymbol{q}^\star - \boldsymbol{q}_{\max}) + \boldsymbol{\mu}_{\min}^T(\boldsymbol{q}_{\min} - \boldsymbol{q}^\star)\\
&\quad\leq d(\boldsymbol{\nu}^\star) - d(\boldsymbol{\nu})
\end{aligned} \tag{21}$$

where the second equality holds thanks to the strong duality, namely, $d(\boldsymbol{\lambda}^\star,\boldsymbol{\mu}^\star) = f(\boldsymbol{v}^\star,\boldsymbol{q}^\star)$, given that the Slater's condition holds for problem (P). The counterpart of Ef-DA can be easily conducted by removing the terms associated with VAR and voltage limits in g, which establishes a strict equality relation. Then, with (20), we can obtain,

$$\frac{1}{2}\left\|\begin{bmatrix}\boldsymbol{v}^*(\boldsymbol{\nu})\\ \boldsymbol{q}^*(\boldsymbol{\nu})\end{bmatrix} - \begin{bmatrix}\boldsymbol{v}^\star\\ \boldsymbol{q}^\star\end{bmatrix}\right\|_{\boldsymbol{H}}^2 \leq d(\boldsymbol{\nu}^\star) - d(\boldsymbol{\nu}) \tag{22}$$

Then, we have,

$$\left\|\begin{bmatrix}\boldsymbol{v}^*(\boldsymbol{\nu})\\ \boldsymbol{q}^*(\boldsymbol{\nu})\end{bmatrix} - \begin{bmatrix}\boldsymbol{v}^\star\\ \boldsymbol{q}^\star\end{bmatrix}\right\|_2^2 \leq \frac{2}{\sigma_{\min}(\boldsymbol{H})}\big(d(\boldsymbol{\nu}^\star) - d(\boldsymbol{\nu})\big) \tag{23}$$

where $\sigma_{\min}(\cdot)$ denotes the smallest eigenvalue. Further with Proposition 1, we finish the proof. ∎

C. Proof of Proposition 3

Proof: For Gf-DA or Ef-DA, we have

$$\begin{aligned}
&\|\boldsymbol{M}\boldsymbol{v}^*(t) - \boldsymbol{q}^*(t) - \boldsymbol{b} - \boldsymbol{M}\boldsymbol{v}^\star + \boldsymbol{q}^\star + \boldsymbol{b}\|_\infty\\
&\quad\leq \big\|[\boldsymbol{M},-\boldsymbol{I}_N]\big\|_\infty \left\|\begin{bmatrix}\boldsymbol{v}^*(\boldsymbol{\nu})\\ \boldsymbol{q}^*(\boldsymbol{\nu})\end{bmatrix} - \begin{bmatrix}\boldsymbol{v}^\star\\ \boldsymbol{q}^\star\end{bmatrix}\right\|_\infty\\
&\quad\leq \big\|[\boldsymbol{M},-\boldsymbol{I}_N]\big\|_\infty \left\|\begin{bmatrix}\boldsymbol{v}^*(\boldsymbol{\nu})\\ \boldsymbol{q}^*(\boldsymbol{\nu})\end{bmatrix} - \begin{bmatrix}\boldsymbol{v}^\star\\ \boldsymbol{q}^\star\end{bmatrix}\right\|_2\\
&\quad= \max_{i\in\mathcal{N}}\left\{\sum_{j\in\mathcal{N}_i}(1+|m_{ij}|)\right\}\left\|\begin{bmatrix}v^*(\boldsymbol{\nu})\\ \boldsymbol{q}^*(\boldsymbol{\nu})\end{bmatrix} - \begin{bmatrix}\boldsymbol{v}^\star\\ \boldsymbol{q}^\star\end{bmatrix}\right\|_2.
\end{aligned} \tag{24}$$

Then with Proposition 2, we can obtain the first row in Proposition 3, which corresponds to the power flow constraint.

Further, only for Gf-DA, we have

$$\begin{aligned}
&\left\|\begin{bmatrix}\boldsymbol{I}_N & \boldsymbol{0}_N\\ -\boldsymbol{I}_N & \boldsymbol{0}_N\\ \boldsymbol{0}_N & \boldsymbol{I}_N\\ \boldsymbol{0}_N & -\boldsymbol{I}_N\end{bmatrix}\begin{bmatrix}\boldsymbol{v}^*(t)\\ \boldsymbol{q}^*(t)\end{bmatrix} - \begin{bmatrix}\boldsymbol{I}_N & \boldsymbol{0}_N\\ -\boldsymbol{I}_N & \boldsymbol{0}_N\\ \boldsymbol{0}_N & \boldsymbol{I}_N\\ \boldsymbol{0}_N & -\boldsymbol{I}_N\end{bmatrix}\begin{bmatrix}\boldsymbol{v}^\star\\ \boldsymbol{q}^\star\end{bmatrix}\right\|_\infty\\
&\quad\leq \left\|\begin{bmatrix}\boldsymbol{I}_N & \boldsymbol{0}_N\\ -\boldsymbol{I}_N & \boldsymbol{0}_N\\ \boldsymbol{0}_N & \boldsymbol{I}_N\\ \boldsymbol{0}_N & -\boldsymbol{I}_N\end{bmatrix}\right\|_\infty \left\|\begin{bmatrix}\boldsymbol{v}^*(t)\\ \boldsymbol{q}^*(t)\end{bmatrix} - \begin{bmatrix}\boldsymbol{v}^\star\\ \boldsymbol{q}^\star\end{bmatrix}\right\|_\infty\\
&\quad\leq \left\|\begin{bmatrix}\boldsymbol{v}^*(t)\\ \boldsymbol{q}^*(t)\end{bmatrix} - \begin{bmatrix}\boldsymbol{v}^\star\\ \boldsymbol{q}^\star\end{bmatrix}\right\|_2.
\end{aligned} \tag{25}$$

Accordingly, we can obtain,

$$\begin{bmatrix}\boldsymbol{v}^*(t) - \boldsymbol{v}_{\max}\\ \boldsymbol{v}_{\min} - \boldsymbol{v}^*(t)\\ \boldsymbol{q}^*(t) - \boldsymbol{q}_{\max}\\ \boldsymbol{q}_{\min} - \boldsymbol{q}^*(t)\end{bmatrix} - \begin{bmatrix}\boldsymbol{v}^\star - \boldsymbol{v}_{\max}\\ \boldsymbol{v}_{\min} - \boldsymbol{v}^\star\\ \boldsymbol{q}^\star - \boldsymbol{q}_{\max}\\ \boldsymbol{q}_{\min} - \boldsymbol{q}^\star\end{bmatrix} \leq \left\|\begin{bmatrix}\boldsymbol{v}^*(t)\\ \boldsymbol{q}^*(t)\end{bmatrix} - \begin{bmatrix}\boldsymbol{v}^\star\\ \boldsymbol{q}^\star\end{bmatrix}\right\|_2 \mathbf{1}_{4N} \tag{26}$$

which indicates,

$$\left[\begin{bmatrix}\boldsymbol{v}^*(t) - \boldsymbol{v}_{\max}\\ \boldsymbol{v}_{\min} - \boldsymbol{v}^*(t)\\ \boldsymbol{q}^*(t) - \boldsymbol{q}_{\max}\\ \boldsymbol{q}_{\min} - \boldsymbol{q}^*(t)\end{bmatrix}\right]^\infty \leq \left\|\begin{bmatrix}\boldsymbol{v}^*(t)\\ \boldsymbol{q}^*(t)\end{bmatrix} - \begin{bmatrix}\boldsymbol{v}^\star\\ \boldsymbol{q}^\star\end{bmatrix}\right\|_2 \cdot \mathbf{1}_{4N}. \tag{27}$$

Then with Proposition 2, we finish the proof. ∎

REFERENCES

[1] E. Sáiz-Marín, E. Lobato, and I. Egido, "New challenges to wind energy voltage control. Survey of recent practice and literature review," *IET Renew. Power Gener.*, vol. 12, no. 3, pp. 267-278, Feb. 2018.

[2] A. K. Pathak, M. P. Sharma, M. Bundele, "A critical review of voltage and reactive power management of wind farms," *Renew. Sustain. Energy Reviews*, vol. 51, pp. 460-471 ,2015.

[3] B. R. Karthikeya and R. J. Schutt, "Overview of wind park control strategies," *IEEE Trans. Sustain. Energy*, vol. 5, no. 2, pp. 416-422, Apr. 2014.

[4] Y. Li, Z. Xu, J. Zhang, and K. Meng, "Variable droop voltage control for wind farm," *IEEE Trans. Sustain. Energy*, vol. 9, no. 1, pp. 491-493, Jan. 2018.

[5] J. Martínez, P. C. Kjær, P. Rodriguez, and R. Teodorescu, "Comparison of two voltage control strategies for a wind power plant," *in Proc. IEEE/PES Power Syst. Conf. Expo. (PSCE)*, pp. 1-9, 2011.

[6] S. Asadollash, R. Zhu, and M. Liserre M, "Analysis of voltage control strategies for wind farms," *IEEE Trans. Sustain. Energy*, in press.

[7] B. Zhang, P. Hou, W. Hu, M. Soltani, C. Chen, and Z. Chen, "A reactive power dispatch strategy with loss minimization for a DFIG-based wind farm," *IEEE Trans. Sustain. Energy*, vol. 7, no. 3, pp. 914-923, Jul. 2016.

[8] M. Montilla-DJesus, D. Santos-Martin, S. Arnaltes, and E. D. Castronuovo, "Optimal reactive power allocation in an offshore wind farms with LCC-HVdc link connection," *Renew. Energy*, vol. 40, no. 1, pp. 157-166, Apr. 2012.

[9] K. Schönleber, S. Ratés-Palau, M. De-Prada-Gil, and O. Gomis-Bellmunt, "Reactive power optimization in HVDC-connected wind power plants considering wake effects," *14th Wind Integr. Work, Energynautics GmbH, Brussels*, 2015.

[10] K. Schönleber, S. Ratés-Palau, and O. Gomis-Bellmunt, "Analysis of reactive power strategies in HVDC-connected wind power plant clusters," *Wind Energy*, vol. 20, no. 12, pp. 1971-1982, Dec. 2017.

[11] S. Jung and G. Jang. "A loss minimization method on a reactive power supply process for wind farm," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3060-3068, Jul. 2017.

[12] Q. Guo, H. Sun, B. Wang, B. Zhang, W. Wu, and L. Tang, "Hierarchical automatic voltage control for integration of large-scale wind power: design and implementation," *Electr. Power Syst. Res.*, vol. 120, pp. 234-241, 2015.

[13] H. Zhao, Q. Wu, Q. Guo, H. Sun, S. Huang, and Y. Xue, "Coordinated voltage control of a wind farm based on model predictive control," *IEEE Trans. Sustain. Energy*, vol. 7, no. 4, pp. 1440-1451, Oct. 2016.

[14] H. Zhao, Q. Wu, J. Wang, Z. Liu, M. Shahidepour, and Y. Xue, "Combined active and reactive power control of wind farms based on model predictive control," *IEEE Trans. Energy Convers.*, vol. 32, no. 3, pp. 1177-1187, Sep. 2017.

[15] Y. Guo, H. Gao, Q. Wu, H. Zhao, J. Østergaard, and M. Shahidehpour, "Enhanced voltage control of VSC-HVDC connected offshore wind farms based on model predictive control," *IEEE Trans. Sustain. Energy*, vol. 9, no. 1, pp. 474-487, Jan. 2018.

[16] S. Asadollah, R. Zhu, M. Liserre, and C. Vournas, "Decentralized reactive power and voltage control of wind farms with type-4 generators," *in PowerTech 2017 IEEE*, Manchester, pp. 1-6, Jun. 2017.

[17] Y. Guo, H. Gao, H. Xing, Q. Wu, and Z. Lin, "Decentralized coordinated voltage control for VSC-HVDC connected wind farms based on ADMM," *IEEE Trans. Sustain. Energy*, vol. 10, no. 2, pp. 800-810, Apr. 2019.

[18] Y. Zhao, J. Chai, and X. Sun, "Relative voltage control of the wind farms based on the local reactive power regulation," *Energies*, vol. 10, no. 3, pp. 1-13, Feb. 2017.

[19] Y. Guo, H. Gao, and Q. Wu, "Distributed cooperative voltage control of wind farms based on consensus protocol," *Int. J. Electr. Power Energy Syst.*, vol. 104, pp. 593-602, Jan. 2019.

[20] M. P. S. Gryning , Q. Wu, Ł. Kocewiak, et al, "Stability boundaries for offshore wind park distributed voltage control," *IEEE Trans. Control Syst. Technol.*, vol. 25, no. 4, pp. 1496-1504, 2017.

[21] H. Sun, et al, "Review of challenges and research opportunities for voltage control in smart grids," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 2790-2801, Jul. 2019.

[22] D. K. Molzahn, F. Dörfler, H. Sandberg, et al, "A survey of distributed optimization and control algorithms for electric power systems," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2941-2962, Nov. 2017.

[23] K. E. Antoniadou-Plytaria, I. N. Kouveliotis-Lysikatos, P. S. Georgilakis, and N. D. Hatziargyriou, "Distributed and decentralized voltage control of smart distribution networks: Models, methods, and future research," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2999-3008, Nov. 2017.

[24] R. Diao, Z. Wang, D. Shi, Q. Chang, J. Duan, and X. Zhang, "Autonomous voltage control for grid operation using deep reinforcement learning," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Atlanta, GA, USA, Aug. 2019, pp. 1-5.

[25] J. Duan, D. Shi, R. Diao, H. Li, Z. Wang, B. Zhang, D. Bian, Z. Yi, "Deep-reinforcement-learning-based autonomous voltage control for power grid operations," *IEEE Trans. Power Syst.*, in press.

[26] Q. Yang, G. Wang, A. Sadeghi, G. B. Giannakis, and J. Sun, "Two timescale voltage control in distribution grids using deep reinforcement learning," *IEEE Trans. Smart Grid*, in press.

[27] D. P. Bertsekas, "Nonlinear programming," 3rd ed. Belmont, MA, USA: Athena scientific, 2016.

[28] J. Li, Z. Xu, J. Zhao J, and C. Zhang, "Distributed online voltage control in active distribution networks considering PV curtailment," *IEEE Trans. Ind. Informat.*, in press, 2019.

[29] H. J. Liu , W. Shi, and H. Zhu,"Distributed voltage control in distribution networks: online and robust implementations," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 6106-6117, 2018.

[30] G. Qu and N. Li, "Optimal distributed feedback voltage control under limited reactive power," *IEEE Trans. Power Syst.*, vol. 35, no. 1, pp. 315-331, Jan. 2018.

[31] Z. Tang, D. J. Hill, and T. Liu, "Fast distributed reactive power control for voltage regulation in distribution networks," *IEEE Trans. Power Syst.*, vol. 34, no. 1, pp. 802-805, Jan. 2019.

[32] Y. Nesterov, "A method for solving the convex programming problem with convergence rate $O(1/k^2)$," *Dokl. Akad. Nauk SSSR*, vol. 269, no. 3, pp. 543-547, 1983.

[33] M. E. Baran and F. F. Wu, "Optimal capacitor placement on radial distribution systems," *IEEE Trans. Power Del.*, vol. 4, no. 1, pp. 725-734, Jan. 1989.

[34] J. N. Sakamuri, Z. H. Rather, J. Rime, M. Altin, Ö. Göksu, and N. A. Cutululis, "Coordinated voltage control in offshore HVDC connected cluster of wind power plants," *IEEE Trans. Sustain. Energy*, vol. 7, no. 4, pp. 1592-1601, Oct. 2016.

[35] O. Noureldeen and I. Hamdan, "Design of robust intelligent protection technique for large-scale grid-connected wind farm," *Prot. Control Modern Power Syst.*, vol. 3, no. 3, pp. 169-182, Jun. 2018.

[36] D. Santos-Martin, S. Arnaltes, and J. R. Amenedo, "Reactive power capability of doubly fed asynchronous generators," *Electr. Power Syst. Res.*, vol. 78, no. 11, pp. 1837-1840, Nov. 2008.

[37] A. Maknouninejad and Z. Qu, "Realizing unified microgrid voltage profile and loss minimization: a cooperative distributed optimization and control approach," *IEEE Trans. Smart Grid*, vol. 5, no. 4, pp. 1621-1630, Jul. 2014.

[38] J. D. Grunnet, M. Soltani, T. Knudsen, M. Kragelund, and T. Bak, "Aeolus toolbox for dynamic wind farm model, simulation and control," *in Proc. Eur. Wind Energy Conf. Exhib.*, Warsaw, 2010, pp. 3119-3129.

[39] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183-202, Mar. 2009.

[40] P. Giselsson, "Improved dual decomposition for distributed model predictive control," *IFAC Proc.*, vol. 47, no. 3, pp. 1203-1209, Aug. 2014.

[41] C. Feng, Z. Li, M. Shahidehpour, F. Wen, W. Liu, and X. Wang, "Decentralized short-term voltage control in active power distribution systems," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4566-4576, Sep. 2018.

Yifei Guo (M'19) received the B.E. and Ph. D. degrees in electrical engineering from Shandong University, Jinan, China, in 2014 and 2019, respectively. Currently, he is a Postdoctoral Research Associate with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA, USA. He was a visiting student with the Department of Electrical Engineering, Technical University of Denmark, Lyngby, Denmark, in 2017–2018.

His research interests include voltage/var control, wind farm control, renewable energy integration, and distribution system optimization and control.

Houlei Gao (M'98) was born in Shandong, China, in 1963. He received the B.Sc. and M.Sc. degrees in electrical power engineering from Shandong University, Jinan, China, in 1983 and 1988, respectively, and the Ph.D. degree from Tianjin University, Tianjin, China, in 1997. From 2004 to 2005, he was with the School of Electrical and Electronic Engineering, Queen's University Belfast, Belfast, U.K. He is a Professor at the School of Electrical Engineering, Shandong University.

His research interests include power system protection, feeder automation, distributed generation, and digital substation.

Zhaoyu Wang (S'13–M'15) is the Harpole-Pentair Assistant Professor with Iowa State University. He received the B.S. and M.S. degrees in electrical engineering from Shanghai Jiaotong University in 2009 and 2012, respectively, and the M.S. and Ph.D. degrees in electrical and computer engineering from Georgia Institute of Technology in 2012 and 2015, respectively. His research interests include power distribution systems and microgrids, particularly on their data analytics and optimization. He is the Principal Investigator for a multitude of projects focused on these topics and funded by the National Science Foundation, the Department of Energy, National Laboratories, PSERC, and Iowa Energy Center.

Dr. Wang is the Secretary of IEEE Power and Energy Society (PES) Award Subcommittee, Co-Vice Chair of PES Distribution System Operation and Planning Subcommittee, and Vice Chair of PES Task Force on Advances in Natural Disaster Mitigation Methods. He is an editor of IEEE Transactions on Power Systems, IEEE Transactions on Smart Grid, IEEE PES Letters and IEEE Open Access Journal of Power and Energy, and an associate editor of IET Smart Grid.