

# Online Optimization for Networked Distributed Energy Resources with Time-Coupling Constraints

Shuai Fan, *Student Member, IEEE*, Guangyu He, *Senior Member, IEEE*, Xinyang Zhou, *Member, IEEE*,  
Mingjian Cui, *Senior Member, IEEE*

**Abstract**—This paper proposes a Lyapunov optimization-based online distributed (LOOD) algorithmic framework for active distribution networks with numerous photovoltaic inverters and invert air conditionings (IACs). In the proposed scheme, ADNs can track an active power setpoint reference at the substation in response to transmission-level requests while concurrently minimizing the utility loss and ensuring the security of voltages. In contrast to conventional distributed optimization methods that employ the setpoints for controllable devices only when the algorithm converges, the proposed LOOD can carry out the setpoints immediately relying on the current measurements and operation conditions. Notably, the time-coupling constraints of IACs are decoupled for online implementation with Lyapunov optimization technique. An incentive scheme is tailored to coordinate the customer-owned assets in lieu of the direct control from network operators. Optimality and convergency are characterized analytically. Finally, we corroborate the proposed method on a modified version of 33-node test feeder.

**Index Terms**—Active distribution networks, online distributed optimization, photovoltaic, inverter air conditionings.

## I. INTRODUCTION

ACTIVE distribution networks (ADNs) integrated with high penetrations of distributed energy resources (DERs) provide increasing flexibility for power systems and accommodate advanced ancillary services such as automatic generation control, fast ramping, and power reserves [1]. However, coordinating numerous DERs to achieve some objectives while considering their distinct dynamics and constraints in a time-varying environment is extremely challenging. Moreover, since the customer-owned DERs are not directly dispatched by the utilities, an incentive-based scheme instead of the direct control from network operators is required.

There has been extended studies on optimal coordination of DERs with the ADN in the literatures. Some works such as [2] design a centralized solver for the formulated optimization problems, which is valid for the small-scale application and utility-owned assets. Refs. [3]–[5] present distributed optimization frameworks, where multiple subproblems need to be solved iteratively until the convergence for each time slot. We term such scheme as solving the problem in a batch fashion [6]. In this case, the system profiles are presumed to be stationary and unchanged during the whole iterative procedure. However, if we use a small time-slot duration to track the optimal setpoints for DERs in the

fast variation environment, the batch fashion will be communicationally costly and rarely feasible to yield the convergency before the system profiles change [6]. Additionally, if the batch mode is applied to design an incentive scheme, various rounds of bargains between ADN operators and customers are required before revealing the optimal price [5], which may be user-unfriendly.

Alternatively, the online convex optimization (OCO) has emerged as a promising paradigm. Unlike conventional batch fashion, a limited number of iterations are performed at each time slot in OCO. The generated coordination signals or setpoints are applied directly without waiting for convergency. Based on this computationally affordable online method, DERs can continuously pursue the trajectory of the time-varying optimizers using a small sampling time in the fast variable environment. For instance, Enyioha *et al.* [7] propose an online decentralized algorithm for the transmission-level economic dispatch. However, it only considers the active power balancing of large generation units while DERs are not involved. Refs. [8] and [9] coordinate the networked microgrids and DERs to minimize the system cost and loss, respectively. A unified online feedback-based controller for DERs is presented in [6] to pursue a given objective. To provide auxiliary services, a primal-dual-based algorithm is proposed in [10] to realize a virtual power plant. Zhou *et al.* [11] present an incentive-enabled online optimization framework.

For online ADN optimization, the aforementioned algorithms cannot well integrate the energy storage devices with time-coupling dynamics. For instance, the inverter air conditioning (IAC) is thermal storage devices whose power setpoints can be adjusted continuously to provide control flexibility to the ADN [12]. The main barrier for integrating IACs is that they feature constraints of the states of temperatures, coupling their power setpoints within the entire operating period. OCO refs. [10] and [11] avoid DERs with the time-coupling constraints to advocate a fast online controller. Li *et al.* [13] propose an online algorithm for the optimization problems considering switching costs but only focuses on the temporal coupling between two successive time slots without considering the whole time span. Model predictive control (MPC) is leveraged to coordinate networked battery energy storages [14]. The predictive-based approach is also tailored in an OCO framework in [15]. However, MPC-based framework can only employ a limited number of time windows ahead to avoid prohibitively high computational complexity with larger predictive window sizes. Some researchers leverage stochastic gradient-based methods to transfer these time-coupling constraints [2], [16]. However, [2] is designed in a centralized manner while [16] only considers coordinating batteries at the transmission level. Furthermore, they are not formulated in an OCO framework.

This paper investigates a Lyapunov optimization-based online distributed (LOOD) algorithmic scheme to achieve an incentive-based DER coordination. In the proposed algorithm, the networked customer-owned DERs are coordinated to provide the

S. Fan and G. He are with the Department of Electrical Engineering, Shanghai Jiao-Tong University, Shanghai, 200240, China (e-mail: {fanshuai, gyhe}@sjtu.edu.cn).

X. Zhou is with National Renewable Energy Laboratory (NREL), Golden, CO 15013, USA (e-mail: Xinyang.Zhou@nrel.gov).

M. Cui is with the Department of Electrical and Computer Engineering, Southern Methodist University, Dallas, TX 75275, USA (e-mail: mingjiancui@smu.edu).

active power tracking service at the substation, while simultaneously minimizing the utility loss and maintain node voltages within an acceptable range. Compared with most existing ADN online optimization works, the main innovations of our developed method are summarized as follows:

(1) The proposed method can integrate numerous DERs with time-coupling constraints and tailor their distinct attributes to the OCO framework. Unlike the greedy decoupling methods and MPC-based scheme, the time-coupling constraints are decoupled from a long-term vision to each time slot in LOOD. It provides the most cost-efficient result since the whole time period is sighted, while also advocates the real-time deployment due to the existence of the closed-form solution for the IAC.

(2) We design a proper incentive scheme to coordinate DERs in lieu of direct controlling. A first-order filter is applied in the incentive generator to alleviate potential fluctuations of the incentives and corresponding responses to smoothen the control and convergence process.

(3) We conduct a rigorous mathematical analysis to demonstrate the impact on the optimality and convergence of our algorithm caused by the step size, weight coefficients, and initialization of virtual queues. In particular, the relaxation for the time-coupling constraints is proved to cause no violation in our setting.

The remainder of this paper is organized as follows. Section II formulates a mathematical model to coordinate networked DERs. In Section III, LOOD is proposed. The performance analysis is analytically characterized in Section IV. Case studies and benchmarks are described in Section V. Concluding remarks are summarized in Section VI.

*Notations:*

We denote the  $L2$ -norm of a vector as  $\|\cdot\|$ .  $P_{\Xi}(\cdot)$  denotes a projection operator onto set  $\Xi$ .

*Definitions:*

- 1). We define the *time-average value*  $\bar{x}$  of a variable  $x^t$  as its mean over the whole time period, such that  $\bar{x} = \lim_{T \rightarrow \infty} (1/T) \sum_{t=1}^T x^t$ .
- 2). A function  $f(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is  $l$ -Lipschitz continuous on  $\mathbb{R}^n$  if there is a constant  $l > 0$  such that

$$\|f(x_1) - f(x_2)\| \leq l \|x_1 - x_2\|, \forall x_1, x_2 \in \mathbb{R}^n. \quad (1)$$

- 3). A function  $f(\cdot)$  is  $\sigma$ -strongly convex if for all  $x_1$  and  $x_2$  in the feasible set and some constant  $\sigma > 0$ , we have

$$[\nabla f(x_1) - \nabla f(x_2)]^T (x_1 - x_2) \geq \sigma \|x_1 - x_2\|^2. \quad (2)$$

- 4). A function  $f(\cdot)$  is strongly concave if  $-f(\cdot)$  is strongly convex.

## II. SYSTEM MODELS

Consider a distribution network with high penetration of DERs. All the control actions are performed in a discrete-time manner with a time interval  $\Delta t$ . The time slots are indexed by  $t$  in  $\mathcal{T} = \{1, 2, \dots, T\}$ . Let  $\mathcal{N} = \{1, 2, \dots, N\}$  collect all the nodes in the network excluding the substation which is denoted by node 0. The sets of nodes connected with PV inverters and IACs are denoted by  $\mathcal{K} \subseteq \mathcal{N}$  and  $\mathcal{A} \subseteq \mathcal{N}$ , respectively. In this paper, each node  $\forall i \in \mathcal{N}$  in the distribution system is regarded as a customer.

### A. Node Models

Two representative types of DERs, i.e., the PV inverter and IAC, are modeled through defining their feasible sets and utility

loss functions.

#### 1) PV Inverter Model

The PV is connected to the network through an inverter. Let  $P_i^{\text{av},t}$  be the maximal available active power output of PV  $i$  at time slot  $t$  and  $S_i^{\text{PV}}$  be the rated apparent capacity of inverter  $i$ . The active power output  $p_i^{\text{PV},t}$  and reactive power output  $q_i^{\text{PV},t}$  belong to a feasible set  $\mathcal{Z}_i^{\text{PV},t}$  given by:

$$\mathcal{Z}_i^{\text{PV},t} = \{(p_i^{\text{PV},t}, q_i^{\text{PV},t}) \mid 0 \leq p_i^{\text{PV},t} \leq P_i^{\text{av},t}, (p_i^{\text{PV},t})^2 + (q_i^{\text{PV},t})^2 \leq (S_i^{\text{PV}})^2\}, \forall i, \forall t, \quad (3)$$

If  $i \notin \mathcal{K}$ , we have  $P_i^{\text{av},t} = 0$  and  $S_i^{\text{PV}} = 0$  for  $\forall t$ .

A quadratic function with coefficient  $c_i^p$  is designed to penalize the active power curtailment of PV. As injecting reactive power is not economic for customers, we also penalize the reactive power generation/absorbing as a quadratic function with a coefficient  $c_i^q$  as follows:

$$U_i^{\text{PV},t}(p_i^{\text{PV},t}, q_i^{\text{PV},t}) := c_i^p (p_i^{\text{PV},t} - P_i^{\text{av},t})^2 + c_i^q (q_i^{\text{PV},t})^2, \forall i, \forall t. \quad (4)$$

#### 2) IAC Model

Unlike the conventional fixed speed ACs, power consumptions of IACs can be continuously adjusted by regulating the operating frequency of the compressors. Compared with fixed speed air conditioners, IACs are more flexible and widespread. We consider various IACs may connect to one node, and each IAC is installed in an independent room. So, we index the IACs connected to node  $i \in \mathcal{A}$  as  $a \in \mathcal{A}_i = \{1, 2, \dots, A_i\}$ , where the cardinality  $A_i$  represents the number of connected IACs at  $i$ . Note that as the operating power is a linear function of its frequency [12], the operating power  $s_{i,a}^t$  of IAC  $a \in \mathcal{A}_i$  at time slot  $t$  is regarded as the optimization decision variable. The operating power  $s_{i,a}^t$  is confined in a box set  $s_{i,a}^{\min} \leq s_{i,a}^t \leq s_{i,a}^{\max}$ . We further define  $\mathbf{s}_i^t = [s_{i,1}^t, s_{i,2}^t, \dots, s_{i,A_i}^t]^T$ . Then with the lower and upper bound vectors  $\mathbf{s}_i^{\min}$  and  $\mathbf{s}_i^{\max}$  the feasible set for IACs is formulated as the following compact form:

$$\mathcal{Z}_i^{\text{AC},t} = \{\mathbf{s}_i^t \mid \mathbf{s}_i^{\min} \leq \mathbf{s}_i^t \leq \mathbf{s}_i^{\max}\}, \forall i, \forall t. \quad (5)$$

We consider IAC working in the cool mode. The IAC features the indoor temperature  $T_{i,a}^t$  following a given dynamic, which can be depicted by the simplified equivalent thermal parameters (ETP) model [12]. Let  $C_{i,a}$  denote the equivalent thermal capacity ( $\text{J}/^\circ\text{C}$ ),  $W_{i,a}$  be the equivalent thermal resistance ( $^\circ\text{C}/\text{W}$ ), and  $Q_{i,a}$  be the cooling rate (W) of the IAC. Since only one IAC is installed in an independent room, the indoor temperature evolves as follows:

$$T_{i,a}^{t+1} = T_{i,a}^t - \eta_{i,a}(T_{i,a}^t - T_{\text{amb}}^t) - Q_{i,a}W_{i,a}(1 - \eta_{i,a}), \forall a, \forall i, \forall t, \quad (6)$$

where  $\eta_{i,a} = e^{-\Delta t/(W_{i,a}C_{i,a})}$  is a constant,  $T_{\text{amb}}^t$  is the ambient temperature at time slot  $t$ . The cooling rate  $Q_{i,a}$  can be modeled as the following linear function of the operating power  $s_{i,a}^t$  [12]:

$$Q_{i,a}^t = k_{i,a}s_{i,a}^t + f_{i,a}, \quad \forall a, \forall i, \forall t, \quad (7)$$

where  $k_{i,a}$  and  $f_{i,a}$  are constant coefficients for a given IAC  $a$ .

For further discussion, the ETP model can be equivalently formulated as:

$$\begin{aligned} T_{i,a}^{t+1} &= T_{i,a}^t + \xi_{i,a}^{+,t} - \xi_{i,a}^{-,t}(s_{i,a}^t), \quad \forall a, \forall i, \forall t, \\ \left\{ \begin{aligned} \xi_{i,a}^{+,t} &= (T_{\text{amb}}^t - T_{i,a}^t)(1 - \eta_{i,a}), \\ \xi_{i,a}^{-,t}(s_{i,a}^t) &= k_{i,a}W_{i,a}(1 - \eta_{i,a})s_{i,a}^t + f_{i,a}W_{i,a}(1 - \eta_{i,a}), \end{aligned} \right. \quad (8) \end{aligned}$$

where  $\xi_{i,a}^{+,t}$  and  $\xi_{i,a}^{-,t}$  denote the temperature increase part and decrease part from time slot  $t$  to  $t+1$ , respectively. The increase part

is caused by the ambient heat radiation, which can be estimated at each time slot. The decrease part is a function of the operating power  $s_{i,a}^t$  to be optimized.

To satisfy the temperature requirements,  $T_{i,a}^t$  is restricted by:

$$T_{i,a}^L \leq T_{i,a}^t \leq T_{i,a}^H, \quad \forall a, \forall i, \forall t, \quad (9)$$

where  $T_{i,a}^L$  and  $T_{i,a}^H$  are the lower and upper temperature limits. Note that constraints (9) are temporarily coupled. Therefore, making decisions at time slot  $t$  will impact the future temperatures and decisions. To cope with the time-coupling constraints, we will tailor the Lyapunov optimization approach to an online adaptation in Section III.

To better control temperature, we define a utility loss function of all IAC of node  $i$  at time  $t$  as a quadratic function penalizing the deviation of the actual temperature from the setpoint:

$$U_i^{AC}(s_i^t) := \sum_{a \in \mathcal{A}_i} c_{i,a}^{AC} (T_{i,a}^t - T_{i,a}^{\text{set}})^2, \quad \forall i, \forall t. \quad (10)$$

where  $c_{i,a}^{AC}$  is a positive cost coefficient and  $T_{i,a}^{\text{set}}$  is the temperature setpoint defined as the median value of  $T_{i,a}^L$  and  $T_{i,a}^H$ .

Here we focus on cool mode of the IAC. Nevertheless, heat mode can be modeled similarly and is omitted here.

### 3) Aggregate Model for Nodes

Let  $\mathbf{z}_i^t := [p_i^{PV,t}, q_i^{PV,t}, (s_i^t)^T]^T$  collect all the decision variables at node  $i \in \mathcal{N}$ . We can present the feasible set of  $\mathbf{z}_i^t$  as:

$$\mathcal{Z}_i^t = \{\mathbf{z}_i^t \mid (p_i^{PV,t}, q_i^{PV,t}) \in \mathcal{Z}_i^{PV,t}, s_i^t \in \mathcal{Z}_i^{AC,t}\}, \quad \forall i, \forall t. \quad (11)$$

Note that  $\mathcal{Z}_i^t$  only comprises some simple constraints excluding the time-coupling temperature constraints.

Then, we define the total utility loss of the customer  $i$  as the accumulation of that of each DER:

$$\hat{U}_i^t(\mathbf{z}_i^t) := U_i^{PV,t}(p_i^{PV,t}, q_i^{PV,t}) + U_i^{AC}(s_i^t), \quad \forall i, \forall t. \quad (12)$$

### B. Network Model

For each node in the ADN, the aggregate active power injection  $p_i^t$  and reactive power injection  $q_i^t$  can be calculated by:

$$p_i^t = p_i^{\text{In},t} - p_i^{\text{PV},t} + \sum_{a \in \mathcal{A}_i} s_{i,a}^t, \quad \forall i, \forall t, \quad (13a)$$

$$q_i^t = q_i^{\text{In},t} - q_i^{\text{PV},t}, \quad \forall i, \forall t, \quad (13b)$$

with  $p_i^{\text{In},t}$  and  $q_i^{\text{In},t}$  being the active and reactive power consumption of the inelastic loads of node  $i$ , respectively, which are presumed to be accurately predictable in real-time operations. For notational simplicity, we collect all the node active and reactive power by vector  $\mathbf{p}^t$  and  $\mathbf{q}^t$ , respectively. Let  $v_i^t$  denote the voltage magnitude of node  $i$  at time  $t$ , and collect the voltage magnitudes of all nodes at time  $t$  by a vector  $\mathbf{v}^t = [v_1^t, v_2^t, \dots, v_N^t]^T$ . The active power at the substation at time slot  $t$  is denoted by  $p_0^t$ .

To develop a computationally affordable controller to advocate the online implementation, power flow linearization is leveraged to model the AC power flow equations, given by:

$$\mathbf{v}^t \approx \mathbf{R}\mathbf{p}^t + \mathbf{X}\mathbf{q}^t + \mathbf{V}, \quad (14a)$$

$$p_0^t \approx \mathbf{M}\mathbf{p}^t + \mathbf{N}\mathbf{q}^t + o, \quad (14b)$$

where  $\mathbf{R} \in \mathbb{R}^{N \times N}$ ,  $\mathbf{X} \in \mathbb{R}^{N \times N}$ ,  $\mathbf{V} \in \mathbb{R}^N$ ,  $\mathbf{M} \in \mathbb{R}^{1 \times N}$ ,  $\mathbf{N} \in \mathbb{R}^{1 \times N}$ , and  $o \in \mathbb{R}^1$  can be obtained by numerous linearization methods with high accuracy, such as approaches in [8] -- [10]. To cope with the inaccuracy caused by the approximation, the real-time measurements of voltage and active power at the substation, denoted by  $\mathbf{v}^{m,t}$  and  $p_0^{m,t}$ , respectively, are leveraged as feedback to reduce modeling errors, as will be shown later.

### C. Problem Formulation

The optimization problem can be formulated as a time-average utility loss minimizing problem  $\mathcal{P}_1$  as follows:

$$\bar{\Phi}^* := \min_{\{\mathbf{z}_i^t\}_{i=1}^N} \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t \in T} \Phi^t, \quad (15a)$$

$$\mathbf{v}^L \leq \mathbf{v}^t \leq \mathbf{v}^H, \quad \forall t, \quad (15b)$$

$$E^t \mid p_0^t - p_{0,\text{set}}^t \leq \varepsilon p_{0,\text{set}}^t, \quad \forall t, \quad (15c)$$

$$\mathbf{z}_i^t \in \mathcal{Z}_i^t, \quad \forall i \in \mathcal{N}, \forall t, \quad (15d)$$

$$T_{i,a}^L \leq T_{i,a}^t \leq T_{i,a}^H, \quad \forall a \in \mathcal{A}_i, \forall i \in \mathcal{N}, \forall t, \quad (15e)$$

$$(8), (13a), (13b), (14a), (14b), \quad (15f)$$

where the objective function is the time-average value of  $\Phi^t$  with  $\Phi^t = \mathbb{E}(\sum_{i \in \mathcal{N}} \hat{U}_i^t)$  being the expectation of the summarized utility losses of all customers at time slot  $t$ . We define a set  $\Theta^t = \{\{P_i^{\text{av},t}, p_i^{\text{In},t}, q_i^{\text{In},t}\}_{i=1}^N, p_{0,\text{set}}^t, T_{\text{amb}}^t\}$  collecting all random variables in (15). In practice, even though  $\Theta^t$  can be accurately estimated or obtained in real time, its realization is unknown in the  $\mathcal{P}_1$  since  $\mathcal{P}_1$  is formulated and solved from a long-term view. Thus, the expectation  $\mathbb{E}$  is taken over the vector  $\Theta^t$  for  $\Phi^t$ . Constraint (15b) confines the voltage to an acceptable range. Constraint (15c) tracks the power setpoint reference at the substation  $p_{0,\text{set}}^t$  given by the transmission-level operator in real time with a permitted tracking error  $\varepsilon$ .  $E^t \in \{0,1\}$  is a binary indicator to switch on the tracking service when it is required.

## III. LOOD ALGORITHM

### A. Virtual Queue-Based Reformulation

The online implementation requires solving  $\mathcal{P}_1$  at each time slot. To decouple the time-coupling constraints from a long-term time horizon, the technique of virtual queues (see e.g., [2], [17]) is leveraged to reformulate  $\mathcal{P}_1$ .

#### 1) Virtual Queue Definition

By summarizing (8) over time from 1 to  $T$  and taking the expectation of each term, we get:

$$\mathbb{E}(T_{i,a}^{T+1}) = \mathbb{E}(T_{i,a}^{t=1}) + \sum_{t \in T} \mathbb{E}[\xi_{i,a}^{+,t} - \xi_{i,a}^{-,t}(s_{i,a}^t)], \quad \forall a, \forall i \quad (16)$$

Divide both sides of (16) by  $T$  and take  $T \rightarrow \infty$  to get:

$$\lim_{T \rightarrow \infty} (\frac{1}{T}) \sum_{t \in T} \mathbb{E}[\xi_{i,a}^{+,t} - \xi_{i,a}^{-,t}(s_{i,a}^t)] = 0, \quad \forall a, \forall i, \quad (17)$$

as  $\mathbb{E}(T_{i,a}^{t=1})$  and  $\mathbb{E}(T_{i,a}^{T+1})$  are both bounded by  $[T_{i,a}^L, T_{i,a}^H]$ .

**Remark 1:** Constraint Eq. (17) is a relaxed version of (15e). If we replace Eq. (15e) in  $\mathcal{P}_1$  by (17) and denote the pertinent optimizer as  $\Phi_r^*$ , we must have  $\Phi_r^* < \Phi^*$ .

To address the relaxed temperature constraints in (17), we can define a virtual queue  $H_{i,a}^t$  for  $\forall a \in \mathcal{A}_i$  and  $\forall i \in \mathcal{N}$  as:

$$H_{i,a}^{t+1} = H_{i,a}^t + [\xi_{i,a}^{+,t} - \xi_{i,a}^{-,t}(s_{i,a}^t)], \quad \forall a, \forall i, \forall t. \quad (18)$$

The arrival rate of the queue is the injected temperature while the departure rate is the cooling temperature at time slot  $t$ . Following the rate stability theorem [17], we place (17) with:

$$\lim_{t \rightarrow \infty} (\mathbb{E}[\|H_{i,a}^t\|/t]) = 0, \quad \forall a, \forall i. \quad (19)$$

In most existing researches such as [2], the initial value of the virtual queue  $H_{i,a}^1$  is set to zero because it handles the time-averaging constraints without any relaxation. Differently, the original constraints must be satisfied at each time slot in  $\mathcal{P}_1$ . To avoid the violations of constraints due to the relaxation, a hot start approach of the virtual queues will be illustrated in the Section

### III-C.

#### 2) Lyapunov Optimization

The constraint (19) still hinders the online deployment as it is coupled over a long-term time span. Consequently, the Lyapunov optimization presented in [17] is leveraged to transfer them to a penalty term attached to the objective function at each time slot based on the observation of the current states.

Let  $\mathbf{H}_i^t \in \mathbb{R}^A$  collect all the virtual queues defined for  $\forall a \in \mathcal{A}_i$  of node  $i$ . Then, we define a Lyapunov function  $L(\mathbf{H}_i^t) : \mathbb{R}^A \rightarrow \mathbb{R}^1$  to measure the size of the queues:

$$L(\mathbf{H}_i^t) := \frac{1}{2} \sum_{a \in \mathcal{A}_i} (H_{i,a}^t)^2. \quad (20)$$

Then, the conditional one-slot Lyapunov drift can be defined as follows to measure the expected queue size growth under observation of current state  $\mathbf{H}_i^t$ :

$$\Delta(\mathbf{H}_i^t) := \mathbb{E}(L(\mathbf{H}_i^{t+1}) - L(\mathbf{H}_i^t) | \mathbf{H}_i^t). \quad (21)$$

To satisfy constraints (19), we minimize the Lyapunov drift to push the queues toward a less congested state. Following the *minimizing drift-plus-penalty* method in [17], we minimize the weighted sum of the drift and cost at each time slot to obtain  $\Phi_r^*$  as follows:

$$\Phi_r^t := (1/V) \sum_{i \in \mathcal{N}} \Delta(\mathbf{H}_i^t) + \mathbb{E}(\sum_{i \in \mathcal{N}} \hat{U}_i^t | \mathbf{H}_i^t). \quad (22)$$

where  $V$  is a positive coefficient to achieve a tradeoff between the stability of queues and original costs.

**Lemma 1:** The drift-plus-penalty function is upper bounded at each time slot  $t$  by:

$$\Phi_r^t \leq B/V + (1/V) \mathbb{E}[\sum_{i \in \mathcal{N}} (U_i^{AC\Delta,t}(s_i^t) | \mathbf{H}_i^t)] + \mathbb{E}(\sum_{i \in \mathcal{N}} \hat{U}_i^t), \quad (23)$$

where

$$B = (1/2) \sum_{i \in \mathcal{N}} \sum_{a \in \mathcal{A}_i} (\xi_{i,a}^{+, \max} - \xi_{i,a}^{-, \min})^2; \quad (24a)$$

$$U_i^{AC\Delta,t}(s_i^t) = \sum_{a \in \mathcal{A}_i} H_{i,a}^t (\xi_{i,a}^{+, t} - \xi_{i,a}^{-, t}). \quad (24b)$$

with  $\xi_{i,a}^{+, \max} = \max_{t \in \mathcal{T}} (\xi_{i,a}^{+, t})$  and  $\xi_{i,a}^{-, \max} = \max_{t \in \mathcal{T}} (\xi_{i,a}^{-, t})$ .

*Proof:* See Appendix A.

Based on Lemma 1, instead of optimizing the drift-plus-penalty function, we will minimize its upper bound alternatively. Following the theorem on opportunistically minimizing an expectation in [17] (c.f. 1.8 in [17]), the policy for the optimization is to observe the current state  $\mathbf{H}_i^t$  and then select the minimizer of  $\sum_{i \in \mathcal{N}} (U_i^{AC\Delta,t}(s_i^t) + \hat{U}_i^t)$ .

In practice,  $U_i^{AC\Delta,t}(s_i^t) : \mathbb{R}^A \rightarrow \mathbb{R}^1$  can be interpreted as an additional utility loss function of the aggregated IACs that carries more temporal knowledge than (10). Hence, we reformulate the utility loss function of each node as:

$$U_i^t(z_i^t) = \hat{U}_i^t + U_i^{AC\Delta,t}(s_i^t), \quad \forall i, \forall t. \quad (25)$$

**Assumption 1:** The utility loss function  $U_i^t$  is  $\sigma$ -strongly convex and  $L$ -Lipschitz continuous for  $\forall i, \forall t$ .

Hereafter, the long-term optimization problem  $\mathcal{P}_1$  comprising the time-coupling constraints can be reformulated as a simple real-time problem to be executed at each time slot without reliance on high-complex solvers. The new problem  $\mathcal{P}_2^t$  is given by:

$$\Phi_1^{t,*} := \min_{\{z_i^t\}_{i=1}^N} \sum_{i \in \mathcal{N}} U_i^t(z_i^t), \quad (26a)$$

$$\mathbf{R}\mathbf{p}^t + \mathbf{X}\mathbf{q}^t + \mathbf{V} - \mathbf{v}^H \leq 0, \quad : \mathbf{u}^{H,t} \quad \forall t \in \mathcal{T}, \quad (26b)$$

$$\mathbf{v}^L - \mathbf{R}\mathbf{p}^t - \mathbf{X}\mathbf{q}^t - \mathbf{V} \leq 0, \quad : \mathbf{u}^{L,t} \quad \forall t \in \mathcal{T}, \quad (26c)$$

$$E^t[\mathbf{M}\mathbf{p}^t + \mathbf{N}\mathbf{q}^t + o - (1 + \varepsilon)p_{0,\text{set}}^t] \leq 0, \quad : \lambda^{H,t} \quad \forall t \in \mathcal{T}, \quad (26d)$$

$$E^t[(1 - \varepsilon)p_{0,\text{set}}^t - \mathbf{M}\mathbf{p}^t - \mathbf{N}\mathbf{q}^t - o] \leq 0, \quad : \lambda^{L,t} \quad \forall t \in \mathcal{T}, \quad (26e)$$

$$\mathbf{z}_i^t \in \mathcal{Z}_i^t, \quad \forall i \in \mathcal{N}, \forall t \in \mathcal{T}, \quad (26f)$$

where the update of virtual queues  $H_{i,a}^t$  follows Eq. (18). Let  $\mathbf{u}^{H,t} = [u_1^{H,t}, u_2^{H,t}, \dots, u_N^{H,t}]^T$  and  $\mathbf{u}^{L,t} = [u_1^{L,t}, u_2^{L,t}, \dots, u_N^{L,t}]^T$  collect dual variables associated with constraints (26b) and (26c), respectively, while  $\lambda^{H,t}$  and  $\lambda^{L,t}$  be dual variables for constraint (26d) and (26e), respectively. Note that all the dual variables are non-negative. For notational simplicity, we denote the objective function of  $\mathcal{P}_2^t$  by  $U^t(\mathbf{z}^t) = \sum_{i \in \mathcal{N}} U_i^t(\mathbf{z}_i^t)$  and the functional constraints (26b) -- (26e) by a compact stacked form,  $g^t(\mathbf{z}^t) \leq 0$  with  $\mathbf{z}^t = \{\mathbf{z}_i^t\}_{i=1}^N$  being all the decision variables. Due to the strong convexity of  $U_i^t$ , the next result follows naturally.

**Lemma 2:** The objective function  $U^t(\mathbf{z}^t)$  is  $\sigma_u$ -strongly convex.

Further, because  $g^t(\mathbf{z}^t)$  is a set of linear constraints, the Jacobian of  $g^t(\mathbf{z}^t)$  is bounded by a positive constant  $\sigma_g$  over the feasible set of  $\mathbf{z}^t$ , such that  $\|g^t(\mathbf{z}^t)\|_F \leq \sigma_g$ . Notice that  $\sigma_g$  can be characterized according the parameter matrices  $\mathbf{R}, \mathbf{X}, \mathbf{M}$ , and  $\mathbf{N}$ .

**Theorem 1:** The difference between time-average value of  $\Phi_1^{t,*}$  denoted as  $\bar{\Phi}_1^*$ , and the optimizer of  $\mathcal{P}_1$ , i.e.,  $\bar{\Phi}^*$  is bounded, such that  $\bar{\Phi}_1^* \leq \bar{\Phi}^* + B/V$ .

*Proof:* See Appendix B.

**Remark 2:**  $\mathcal{P}_2^t$  provides a time decouple reformulation within  $\mathcal{O}(1/V)$  of the optimal results of original  $\mathcal{P}_1$  together with  $\mathcal{O}(V)$  tradeoff in the time-coupling constraints. A large  $V$  can decrease the optimality gap but also bring about constraint's violation. In Section III-C, the upper limit for  $V$  that ensures the constraints is demonstrated.

#### B. Online Distributed Dual Ascent Algorithm

To design an online distributed solver for  $\mathcal{P}_2^t$ , we consider its regularized Lagrangian function defined as follows:

$$\begin{aligned} \mathcal{L}^t(\mathbf{z}^t, \mathbf{d}^t) = & \sum_{i \in \mathcal{N}} U_i^t + (\mathbf{u}^{H,t})^T (\mathbf{R}\mathbf{p}^t + \mathbf{X}\mathbf{q}^t + \mathbf{V} - \mathbf{v}^H) \\ & + (\mathbf{u}^{L,t})^T (\mathbf{v}^L - \mathbf{R}\mathbf{p}^t - \mathbf{X}\mathbf{q}^t - \mathbf{V}) \\ & + \lambda^{H,t} [\mathbf{M}\mathbf{p}^t + \mathbf{N}\mathbf{q}^t + o - (1 + \varepsilon)p_{0,\text{set}}^t] \\ & + \lambda^{L,t} [(1 - \varepsilon)p_{0,\text{set}}^t - \mathbf{M}\mathbf{p}^t - \mathbf{N}\mathbf{q}^t - o] - \frac{\varphi}{2} \|\mathbf{d}^t\|^2, \end{aligned} \quad (27)$$

where  $\mathbf{d}^t = [(\mathbf{u}^{H,t})^T, (\mathbf{u}^{L,t})^T, \lambda^{H,t}, \lambda^{L,t}]^T$  collects all the dual variables, and  $\frac{\varphi}{2} \|\mathbf{d}^t\|^2$  is a regularization term to ensure the concavity of the dual function with a predefined parameter  $\varphi \geq 0$ . Such regulation is widely used in OCO, such as [9]--[11]. The bounded gap between the saddle point of the regularized Lagrangian function and the original one can be found in [19].

We next propose a dual ascent algorithm to find the saddle point of (27). To that end, consider the following dual problem:

$$\max_{\mathbf{d}^t \in \mathbb{R}_+^{2N+2}} D^t(\mathbf{d}^t), \quad (28)$$

where  $D^t(\mathbf{d}^t)$  is the dual function calculated from:

$$D^t(\mathbf{d}^t) = \min_{\mathbf{z}^t \in \mathcal{Z}^t := \prod_{i \in \mathcal{N}} \mathcal{Z}_i^t} U^t(\mathbf{z}^t) + (\mathbf{d}^t)^T g^t(\mathbf{z}^t) - \frac{\varphi}{2} \|\mathbf{d}^t\|^2. \quad (29)$$

**Assumption 2:**  $\mathcal{P}_2^t$  is strictly feasible for  $\forall t$ , i.e., it satisfies the Slater's condition.

The strong duality of  $\mathcal{P}_2^t$  holds based on Assumption 2 [20]. Thus, if  $\mathbf{d}^{t,*}$  is the solver to (28),  $\mathbf{z}^{t,*} = \arg \min D^t(\mathbf{d}^{t,*})$  is the optimal solution to  $\mathcal{P}_2^t$ .

We continue to investigate how to solve the problem in a distributed manner based on incentives. Given the optimal dual vari-

ables, the primal problem is ready to be divided and equivalently solved through  $N$  subproblems. Particularly, each subproblem only requires the local information and a couple of coordination signals composed of the dual variables. The local subproblem denoted as  $\mathcal{P}_{L,i}^t$  for  $\forall i$  and  $\forall t$ :

$$\min_{\mathbf{z}_i^t} U_i^t(\mathbf{z}_i^t) + \alpha_i^{t,*} (\sum_{a \in \mathcal{A}_i} s_{i,a}^t - p_i^{PV}) + \beta_i^{t,*} (-q_i^{PV}), \quad (30a)$$

$$\mathbf{z}_i^t \in \mathcal{Z}_i^t. \quad (30b)$$

In  $\mathcal{P}_{L,i}^t$ ,  $\alpha_i^{t,*}$  and  $\beta_i^{t,*}$  are the coordination signals with the vector forms being  $\alpha^{t,*}$  and  $\beta^{t,*}$ , respectively. By [11], [15], we design the signals as follows:

$$\alpha^{t,*} = \underbrace{\mathbf{R}(\mathbf{u}^{H,t*} - \mathbf{u}^{L,t*})}_{\text{incentive for voltage regulation}} + \underbrace{[(\lambda^{H,t*} - \lambda^{L,t*})\mathbf{M}]^T}_{\text{incentive for power tracking}} \quad (31a)$$

$$\beta^{t,*} = \underbrace{\mathbf{X}(\mathbf{u}^{H,t*} - \mathbf{u}^{L,t*})}_{\text{incentive for voltage regulation}} + \underbrace{[(\lambda^{H,t*} - \lambda^{L,t*})\mathbf{N}]^T}_{\text{incentive for power tracking}} \quad (31b)$$

Note that according to Theorem 2 in [11], the design of the coordination signals ensures an exact distributed reformulation, i.e., the optimal solutions of  $N$  subproblems  $\mathcal{P}_{L,i}^t$  coincide with the centralized optimization results of  $\mathcal{P}_i^t$ .

In practice, the coordination signal  $\alpha_i^{t,*}$  and  $\beta_i^{t,*}$  play roles as the monetary incentive for the active and reactive power for  $\forall i$ , respectively. The incentive signals comprise two components. The first term is the price for voltage regulation to ensure an acceptable range. The second term is used to induce customers to regulate the DERs for a power tracking service.

We further define  $c(\mathbf{d}^t): \mathbb{R}_+^{2N+2} \rightarrow \mathbb{R}^{2N}$  to compactly denote Eq.(31a) and Eq.(31b). From the structure of  $c(\mathbf{d}^t)$ , we have the following results.

**Remark 3:**  $c(\mathbf{d}^t)$  is  $L_c$ -Lipschitz continuous.

**Assumption 3:** The optimal dual variables between two successive time slots are bounded by a positive constant  $\Delta_d$ , such that  $\|\mathbf{d}^{t,*} - \mathbf{d}^{t+1,*}\| \leq \Delta_d$ .

This assumption is standard in the domain of OCO, such as [15].

### 3) Online Algorithm

According to our setting,  $\mathcal{P}_2^t$  is required to be online implemented. Concurrently, to reduce the fluctuations of incentives, a first-order filter is also applied in this algorithm. To recap, the proposed LOOD algorithm is illustrated as:

---

#### LOOD Algorithm

---

##### Initialization:

Customers choose an initial value for virtual queue  $\mathbf{H}_i^{t=0}$  according to the hot start policy range. See (35a) and (35b).

ADN operator sets initial dual variables  $\mathbf{d}^{t=0} = 0$ ; sets  $\delta \in (0, \delta^{\max}]$  and  $V \in (0, V^{\max}]$ . (See (35c) and Theorem 2).

**for**  $t = 0, 1, 2, \dots, T$

[P1] ADN operator sends  $(\alpha_i^t, \beta_i^t)$  to each customer.

[P2] Customer  $i$  receives  $(\alpha_i^t, \beta_i^t)$  and solves  $\mathcal{P}_{L,i}^t$  locally.

[P3] ADN operator measures the node voltage and active power at the substation, i.e.,  $\mathbf{v}^{m,t}$  and  $p_0^{m,t}$ , respectively.

[P4] ADN operator updates dual variables according to (32a)–(32d).

[P5] ADN operator updates  $(\alpha_i^{t+1}, \beta_i^{t+1})$  based on (32e)–(32h).

**end for**

$$\mathbf{u}^{H,t+1} = P_{\mathbb{R}^+}[(1-\varphi)\mathbf{u}^{H,t} + \delta(\mathbf{v}^{t,m} - \mathbf{v}^H)], \quad (32a)$$

$$\mathbf{u}^{L,t+1} = P_{\mathbb{R}^+}[(1-\varphi)\mathbf{u}^{L,t} + \delta(\mathbf{v}^L - \mathbf{v}^{t,m})], \quad (32b)$$

$$\lambda^{H,t+1} = P_{\mathbb{R}^+}[(1-\varphi)\lambda^{H,t} + \delta(p_0^{t,m} - (1+\varepsilon)p_{0,\text{set}}^t)], \quad (32c)$$

$$\lambda^{L,t+1} = P_{\mathbb{R}^+}[(1-\varphi)\lambda^{L,t} + \delta((1-\varepsilon)p_{0,\text{set}}^t - p_0^{t,m})], \quad (32d)$$

$$\widehat{\alpha}^{t+1} = \mathbf{R}(\mathbf{u}^{H,t+1} - \mathbf{u}^{L,t+1}) + [(\lambda^{H,t+1} - \lambda^{L,t+1})\mathbf{M}]^T, \quad (32e)$$

$$\widehat{\beta}^{t+1} = \mathbf{X}(\mathbf{u}^{H,t+1} - \mathbf{u}^{L,t+1}) + [(\lambda^{H,t+1} - \lambda^{L,t+1})\mathbf{N}]^T, \quad (32f)$$

$$\alpha^{t+1} = (1-\phi)\widehat{\alpha}^{t+1} + \phi\alpha^t, \quad (32g)$$

$$\beta^{t+1} = (1-\phi)\widehat{\beta}^{t+1} + \phi\beta^t. \quad (32h)$$

In our algorithm, [P1] and [P3] to [P5] are processed by the ADN operator based on measurements and the received setpoint reference from transmission-level operators. [P2] is solved by customers locally depending on the private information and the incentives. It is hard to get the optimal incentives in (31) in a time varying conditions since the optimal dual variables can be revealed only after various iterations between the ADN operator and customers. However, the environments and customers' responses have changed during the multiple rounds' bargains. Alternatively, based on the OCO framework, we update the incentives only relying on current measurements and conditions. Convergence and optimality gaps of this online algorithm will be characterized analytically in Section IV.

In this algorithm, the actual incentive signal is tuned by a first-order filter, as shown in (32g) and (32h). The motivation of this filter is to smooth the incentive signal. In practice, fast fluctuations of the monetary incentive are not user-friendly. More importantly, the fluctuations of incentive signals will be reflected in the node power and voltage finally. As the violations are essential to power systems' stability, we smoothen the incentive signals to reduce the violation of node power and voltage accordingly. We also characterize the discrepancy on the solver of the optimization problems after filtering the incentive signals.

Note that we leverage measurements feedback to cope with the accurate AC power flow in the update of dual variables with the following assumption to bound the discrepancy between the linearization power flow model and the actual measurements.

**Assumption 4:** There exists a positive constant  $e$  such that:

$$\|g^t(\mathbf{z}^t) - g^t(\mathbf{z}^{t,m})\| \leq e. \quad (33)$$

### C. Solving Local Problems

In the LOOD algorithm, customers need to solve  $\mathcal{P}_{L,i}^t$  locally. According to the form of  $\mathcal{P}_{L,i}^t$ , it is a standard quadratic problem that can be solved efficiently. Although there may be hundreds or thousands of IACs connecting to one node, they operate independently. We can get a closed-form solver for each IAC at time slot  $t$  as follows:

$$s_{i,a}^{t,*} = P_{\mathbb{R}^+}^{AC,t} \left[ \frac{(T_{i,a}^{t-1} - T_{i,a}^{\text{set}} + \xi_{i,a}^{+,t})}{\Omega_{i,a}} + \frac{H_{i,a}^t}{c_{i,a}^A \Omega_{i,a} V} - \frac{\alpha_i^t}{c_{i,a}^A \Omega_{i,a}^2} - f_{i,a} \right], \quad (34)$$

where  $\Omega_{i,a} = k_{i,a} W_{i,a} (1 - \eta_{i,a})$  is a time-invariant constant.

**Lemma 3:** The relaxation of the original time-coupled constraints in (15e) will not bring about violations if the initialization of the virtual queue, denoted by  $h_{i,a}$  and weight coefficient  $V$  are chosen from  $[h_{i,a}^{\min}, h_{i,a}^{\max}]$  and  $(0, V^{\max}]$ , respectively.

The boundaries of  $h_{i,a}$  and  $V$  are given by:

$$h_{i,a}^{\min} = (\alpha_i^{\max} / \Omega_{i,a} + c_{i,a}^A T_{i,a}^{\text{set}} - T_{i,a}^H) V - T_{i,a}^H + T_{i,a}^{t=0}, \forall a, \forall i, \quad (35a)$$

$$h_{i,a}^{\max} = (\alpha_i^{\min} / \Omega_{i,a} + c_{i,a}^A T_{i,a}^{\text{set}} - T_{i,a}^L) V - T_{i,a}^L + T_{i,a}^{t=0}, \forall a, \forall i, \quad (35b)$$

$$V_{\max} = \min_{i \in \mathcal{N}} \left[ \frac{\Omega_{i,a} (T_{i,a}^L - T_{i,a}^H)}{c_{i,a}^A \Omega_{i,a} (T_{i,a}^L - T_{i,a}^H) + \alpha_i^{\min} - \alpha_i^{\max}} \right] \quad (35c)$$

where  $\alpha_i^{\max} = \max(\alpha_i^t)$  and  $\alpha_i^{\min} = \min(\alpha_i^t)$  for all  $t \in \mathcal{T}$ . We assume  $\alpha_i^{\max}$  and  $\alpha_i^{\min}$  can be estimated based on the history data.

The proof of Lemma 3 is that if  $T_{i,a}^{t-1} + \xi_{i,a}^{+,t} - \xi_{i,a}^{-,t} \geq T_{i,a}^H$ , a solver  $s_{i,a}^{t,*}$  that makes sure  $\xi_{i,a}^{+,t} - \xi_{i,a}^{-,t} \leq 0$  will be obtained based on (34). Similarly, if  $T_{i,a}^{t-1} + \xi_{i,a}^{+,t} - \xi_{i,a}^{-,t} \leq T_{i,a}^L$ , we will have  $\xi_{i,a}^{+,t} - \xi_{i,a}^{-,t} \geq 0$ . To recap, the generated solver for the IAC will cool the room if the indoor temperature is going to exceed the upper limit, while stop cooling the room if the indoor temperature outrides the lower limit.

**Remark 4:** We only need to estimate the upper and lower limits of  $\alpha_i$ . Thus, the accuracy requirements are relatively low. In practice, the ADN operator can also confine the incentive in a predefined range beforehand like many current demand response programs, which better advocate the estimation.

#### IV. PERFORMANCE ANALYSIS

In this section we analytically characterize the performance of the LOOD method. We first introduce the following useful lemmas.

**Lemma 4:** Under Assumption 1, the inverse function of  $\nabla U_i^t$  denoted by  $(\nabla U_i^t)^{-1}$  exists and is  $(1/\sigma)$ -Lipschitz continuous.

*Proof:* See Appendix C.

**Lemma 5:** The dual function, i.e.,  $D^t(\cdot)$ , has an  $L_D$ -Lipschitz continuous gradient, where  $L_D = (\sigma_g)^2 / \sigma_u + \varphi$ , based on Lemma 2 and the bounded Jacobian of  $g^t(\mathbf{z}^t)$ . Furthermore,  $D^t(\cdot)$  is  $\sigma_D$ -strongly concave where  $\sigma_D \leq L_D$ .

The proof can be found in Lemma 2 of [15].

**Theorem 2:** (Convergence of dual variables) If the step size in LOOD is chosen according to  $0 < \delta \leq 1/(\sigma_D + L_D)$ , the discrepancy between optimal dual variables of  $\mathcal{P}_2$  and the dual variables generated by LOOD are bounded by:

$$\limsup_{t \rightarrow \infty} \|\mathbf{d}^{t+1} - \mathbf{d}^{t+1,*}\| = \ell, \quad (36)$$

where  $\ell = (\Delta_d + \delta e) / (1 - \kappa)$  and  $\kappa = \sqrt{2 - (4\delta\sigma_D L_D) / (\sigma_D + L_D)}$ .

*Proof:* See Appendix D.

**Corollary 1:** (Convergence of the incentives) To characterize the convergence of the incentives, we first collect the incentives at each node by  $\mathbf{c}_i^t = [\alpha_i^t, \beta_i^t]^T$  with a vector form  $\mathbf{c}^t$  compactly denoting the incentives for all the nodes  $i \in \mathcal{N}$ . Then, the discrepancy between optimal incentives  $\mathbf{c}^{t+1,*}$  generated by (31) and the online created ones  $\mathbf{c}^{t+1}$  in LOOD is bounded by:

$$\limsup_{t \rightarrow \infty} \|\mathbf{c}^{t+1} - \mathbf{c}^{t+1,*}\| = L_c(\ell + \phi\Delta_d). \quad (37)$$

*Proof:* See Appendix E.

**Corollary 2:** (Convergence of the primal variables) The discrepancy between optimal solvers of  $\mathcal{P}_2$  and the solvers generated by LOOD are bounded by:

$$\limsup_{t \rightarrow \infty} \|\mathbf{z}_i^{t+1} - \mathbf{z}_i^{t+1,*}\| = Y_i, \quad \forall i. \quad (38)$$

where  $Y_i = L_c(1 + A_i) / [\ell + (1 + \phi)\Delta_d]$ .

*Proof:* See Appendix F.

**Theorem 3:** (Main results) If LOOD is used to solve the original ADN optimization problem  $\mathcal{P}_1$ , the difference between the

LOOD-based time-average optimizer, denoted as  $\Phi_{OD}^*$ , and the original optimizer  $\Phi^*$  is upper bounded as:

$$\overline{\Phi_{OD}^*} \leq \overline{\Phi^*} + B/V + NLY, \quad (39)$$

where  $Y = \max(Y_i)_{i \in \mathcal{N}}$ .

*Proof:* See Appendix G.

#### V. CASE STUDY

##### A. Simulation Setup

Consider a modified version of 33-node test feeder [21]. We run our method from 8:00 to 19:00, while it is divided into 660 time slots, i.e.,  $\Delta t = 1 \text{ min}$ . It is assumed that PV systems with a 500kVA rating inverter are located at node 2, 3, 10, 12, 16, and 18. The PV systems with a 750 kVA rating inverter are connected to node 5, 6, 7, 8, 9, 20, 24, 26, 29, 30, and 32. The available active power generation profiles of these PV are obtained from Pecanstreet [22]. We set  $c_i^p = 3$  and  $c_i^q = 2$  in the utility loss functions for  $\forall i$ . The inelastic load profile comes from Open Energy Information [23]. All the data are pretreated to have a guaranty of 1 min sampling rate. As for the IACs, we assume that nodes 2, 9, 10, 12, 14, 15, and 30 are connected by 300 IACs, while nodes 3, 6, 7, 17, 21, 25, 28, 31 and 32 are installed with 500 IACs. The maximal operating power of IAC is selected in the range of [500, 800]W. The minimal power is set as 10% of the maximal power. The equivalent thermal capacity of the environment is selected in the range of [2000, 3000] J/°C, The equivalent heat rate is selected in a range of [0.05, 0.08] °C/W. The ambient temperature is simulated by a function [24], given by:

$$T_{amb}^t = T_{amb}^{\text{set}} | \sin(2\pi t / 2T) | + T_{amb}^{\min}, \quad \forall t, \quad (40)$$

where  $T_{amb}^{\text{set}} = 4^\circ\text{C}$  and  $T_{amb}^{\min} = 35^\circ\text{C}$ . The temperature setpoint is set as  $T_{i,a}^{\text{set}} = 23^\circ\text{C}$  with a bandwidth  $\Delta T = 2^\circ\text{C}$  for all the IACs. Without loss of generality, we set  $k_{i,a} = 1.2$ ,  $f_{i,a} = 0$  and  $c_{i,a}^{AC} = 10^{-5}$  for all the IACs. As for our algorithm, we set the step size to 0.1. Some predefined parameters are set as  $V = 0.9V_{\max}$ ,  $\varphi = 10^{-4}$ ,  $\phi = 0.1$ , and  $h_{i,a} = (h_{i,a}^{\max} + h_{i,a}^{\min}) / 2$ .

##### B. Benchmarks

We use four different strategies to compare with the proposed method. Strategy 1 (S1) operates the ADN without any control. PV systems maintain the maximal active power output and IACs operate based on the gap between the current indoor temperature and the temperature setpoint to maintain a comfortable indoor temperature, given by:

$$s_{i,a}^t = \begin{cases} s_{i,a}^{\max} & T_{i,a}^t \geq T_{i,a}^H, \\ k_{i,a}^{\text{droop}} (T_{i,a}^t - T_{i,a}^{\text{set}}) + s_{i,a}^{\text{base}}, & \forall a, \forall i, \forall t, \\ s_{i,a}^{\min} & T_{i,a}^t \leq T_{i,a}^L, \end{cases} \quad (41)$$

where  $k_{i,a}^{\text{droop}}$  is the droop coefficient and  $s_{i,a}^{\text{base}}$  is the base operating power of a given IAC.

Strategy 2 (S2) operates the ADN based on a modified droop control scheme. In S2, PV inverters use a linear Q-V droop control scheme to decide their VAR outputs with a slope coefficient  $k_{\text{droop}}$ . Note that  $k_{\text{droop}}$  may significantly impact the voltage regulation performance. To strictly testify our proposed method,  $k_{\text{droop}}$  is manually adjusted to well perform in this case study. When there is no tracking requirement at the substation, i.e.,  $E^t = 0$ , PV inverters keep the maximal power outputs in their feasible regions and IACs operate according to (41). In the presence of power tracking

requirements, i.e.,  $E^t = 1$ , the active power at each node connected with PV inverters or IACs is regulated with a predefined droop coefficient  $\gamma$ :

$$p_i^{t+1} = P_{Z_i^{t+1}}[p_i^t - \gamma(p_{0,\text{set}}^t - p_0^{m,t})] \quad \forall i, \forall t. \quad (42)$$

In practice, a proper  $\gamma$  is crucial to the performance of power tracking. However, it is rarely practical to get an optimal  $\gamma$ . Multiple values of  $\gamma$  will be deployed in the following case studies for a clear comparison.

Strategy 3 (S3) is based on the greedy optimization, as shown in (43). The main discrepancy between S3 and our proposed method is that we consider the temperature constraints for IACs in a long-term form. However, S3 directly decouples  $P_i$  to  $T$  sub-problems. The greedy algorithm is shortsighted as it optimizes the cost at each time slot without considering the future. To compare with our proposed method, the greedy optimization model is run in an online distributed form.

$$\min_{\{z_i^t\}_{i=1}^N, \sum_{i \in \mathcal{N}} \hat{U}_i(z_i^t)}, \quad (43)$$

(13a),(13b),(14a),(14b),(15b)–(15e).

Strategy 4 (S4) runs the online distributed optimization without considering IACs. Alternatively, IACs operate according to their own policy, i.e., (41). This benchmark is used to verify the performance of integrating IACs into the ADN coordination.

### C. Results

First of all, we test the node voltage under the control of S1, S2, and the proposed LOOD method. We illustrate the voltage magnitude at node 10, 12, and 15 in Fig. 1. Following S1, voltage magnitudes between 10:00 – 16:00 exceed the upper limit due to lacking any control. S2 is effective for maintaining voltage security to some extent, but one oscillation of voltage occurs at 12:00. In practice,  $k_{\text{droop}}$  and  $\gamma$  both impact the performance of S2. We set  $\gamma = 0.03$  in this test since it performs relatively better than any other settings. Too large coefficients could bring about oscillations and even overshoots, while too small ones cannot effectively and quickly regulate both the voltage and power. To get an optimal coefficient usually calls for the global and detailed information of these customer-owned devices. In contrast, the proposed LOOD method outperforms both S1 and S2 as the voltage magnitudes maintain in an acceptable and relatively smooth region. Notably, the flat voltage profiles near to the upper limit are obtained before the tracking requirements come (before 12:00) since it is economically efficient. In the presence of tracking requirements (after 12:00), the curtailment of PV occurs to support the tracking performance. It can be seen that the voltages decrease accordingly.

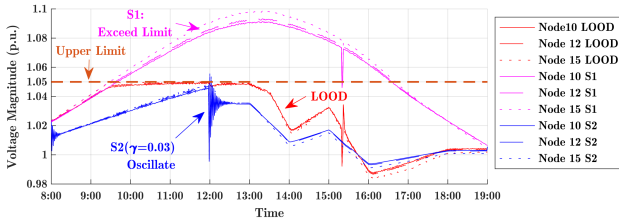


Fig. 1. Node voltage magnitude.

Fig. 2 shows the active power at the substation. To test the power tracking ability of these strategies, a setpoint reference from 12:00 to 19:00 is applied, including a sudden increasing (12:00 – 13:00), fast ramping up or down (13:00 – 18:00) and keeping flat

output (12:00 – 13:00, 18:00 – 19:00). When there is no tracking requirement (8:00 – 12:00), the power using the LOOD method is larger than that using the S2. This is because LOOD can consider the whole-time span and raise the power in advance to better track the sudden change at 12:00. From 12:00 to 19:00, LOOD can guarantee an effective power tracking except for a disturbance that is caused by the sudden change of available PV outputs. In practice, it is hard to reveal this disturbance beforehand, and the sudden change will not cause a huge impact to the whole system. Conversely, the tracking ability of S2 depends on a proper  $\gamma$ . As illustrated in Fig. 2, when  $\gamma = 0.03$ , the tracking performs well. However, undershoot ( $\gamma = 0.003$ ), oscillation ( $\gamma = 0.06$ ), and overshoot ( $\gamma = 0.09$ ) occur if  $\gamma$  is not well set.

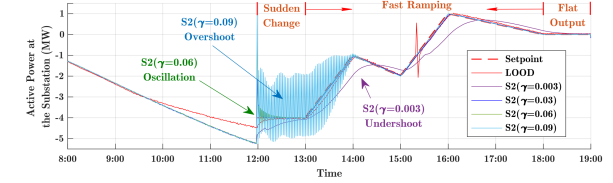


Fig. 2. Active power at the substation.

As for minimizing the utility loss, we compare the performances of LOOD, S2 and S3, as shown in Fig. 3. Intuitively, S2 brings about the largest utility loss. Although the utility loss gets smaller when  $\gamma = 0.003$ , the undershoot will occur. S3 is a variant of our method, thus it also well assures the voltage constraints and power tracking requirement. However, the utility loss caused by S3 is obviously larger than the time-average utility loss caused by LOOD. Actually, although S3 takes into account the flexibility of IAC, the greedy optimization is shortsighted and potential to employ the flexibility of IACs excessively. When the tracking signal keeps for a long time, IACs cannot respond to it sustainably.

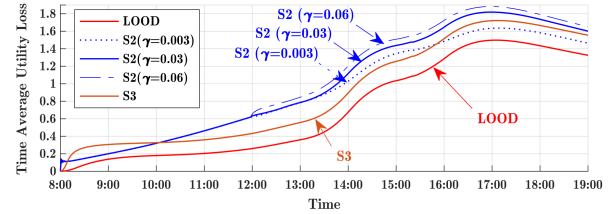


Fig. 3. Time-average utility loss.

S4 is another variant of LOOD, where IACs are not coordinated but operate according to their own local controllers. As illustrated in Fig. 4, LOOD can reduce the curtailment and reactive power absorbing of PV inverters obviously than S4. In LOOD, numerous IACs are well coordinated with PV inverters, while all the regulation relies on PV inverters in S4. If we increase the comfortable temperature bandwidth ( $\Delta T$ ) and IACs can get more flexibility accordingly, the performance of LOOD can be enhanced. Note that residents cannot accept too large temperature bandwidths, so we need to control the tradeoff in the practical deployment.

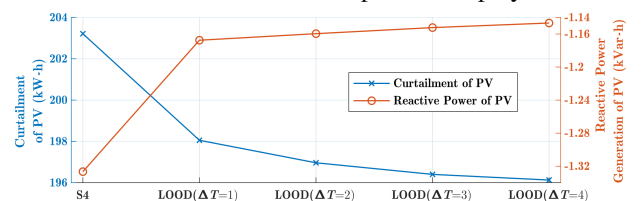


Fig. 4. Curtailment and reactive power generation of PV inverters.

Fig. 5 shows the dynamics of indoor temperature at node 2 (300



IACs) and node 32 (500 IACs). Although the temperature constraints are relaxed to a time-average form, it can obey the original strict temperature constraints as we set the coefficients  $V$  and  $h_{i,a}$  according to the demonstrated ranges, i.e., (35). Intuitively, the temperature dynamics match the regulation requirements. For instance, IACs will raise the operating power to cool the rooms from 12:00~14:00, where the setpoint reference at substation requires the ADN to increase the actual power consumption.

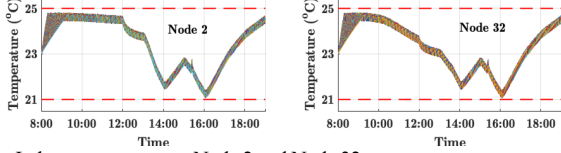


Fig. 5. Indoor temperature at Node 2 and Node 32.

Finally, we discuss the power fluctuation at the substation under different  $\phi$ . We define:

$$\vartheta = \left(\frac{1}{t_2 - t_1 - 1}\right) \sum_{\tau=t_1}^{t_2} (p_0^{m,\tau+1} - p_0^{m,\tau})^2 \quad (44)$$

to quantify the so-termed power fluctuation from time slot  $t_1$  to  $t_2$ . We divide the day into 7 fragments according to the power tracking signals. As shown in Table I,  $\vartheta$  is obviously smaller when  $\phi = 0.1$  than  $\phi = 0$  because incentive signals are smoothed by a filter and the DERs will not respond to the signals too drastically. When we increase  $\phi$  to 0.2 and 0.4,  $\vartheta$  continues to decrease since the enhanced filtering performances. Nevertheless, Corollary 2 has illustrated that  $\phi$  will bring a fix discrepancy between actual solvers and optimal solvers. Thus, a relatively small  $\phi$  can not only help smooth the power at the substation and also ensure the economic efficiency of the algorithm.

TABLE I  
POWER FLUCTUATION ( $\vartheta$ ) AT SUBSTATION

	8~12	12~13	13~14	14~15	15~16	16~18	18~19
$\phi = 0$	0.073	0.396	0.438	0.285	16.047	0.166	0.072
$\phi = 0.1$	0.069	0.148	0.405	0.245	13.825	0.068	0.036
$\phi = 0.2$	0.063	0.105	0.366	0.227	12.215	0.062	0.035
$\phi = 0.4$	0.058	0.102	0.319	0.113	11.151	0.060	0.019

## VI. CONCLUSION

This paper proposes an online distributed optimization algorithmic framework for ADNs to track a setpoint reference at the substation while concurrently minimizing the utility loss and assuring the security of voltages. Unlike most existing optimization methods for ADNs, the proposed LOOD algorithm can generate the setpoints for PV inverters and IACs immediately only relying on current measurements and environment conditions. Notably, the time-coupling constraints for IACs are considered and tackled in an online optimization framework by the Lyapunov optimization technique. Moreover, an incentive scheme is tailored in the proposed method to coordinate customer-owned DERs instead of dispatching them directly from ADN operators. Our incentive generator also considers a first-order filter to alleviate high fluctuations of incentives and corresponding responses. The theoretical analysis demonstrates the bounded gap between the optimizer from the proposed algorithm and the ideal global optimizer. Numerical results show the tracking ability, voltage control performance, and utility loss in our algorithm outperform those in benchmarks.

In this work, the ADNs are assumed to be three-phase balances. Thus, future work will extend our algorithm to a three-phase unbalanced case. Besides, the parameters of IACs model are pre-

sumed to be estimated beforehand and time-invariant in the operation. An online modeling-embedded optimization framework for networked DERs will be further studied.

## REFERENCE

- [1] Y. Jiang, C. Wan, J. Wang, Y. Song, and Z. Y. Dong, "Stochastic receding horizon control of active distribution networks with distributed renewables," *IEEE Trans. Power Syst.*, vol. 34, no. 2, pp. 1325-1341, Mar. 2019.
- [2] W. Shi, N. Li, C. Chu, and R. Gadh, "Real-time energy management in microgrids," *IEEE Trans. Smart Grid*, vol. 8, no. 1, pp. 228-238, Jan. 2017.
- [3] T. Zhao and Z. Ding, "Cooperative optimal control of battery energy storage system under wind uncertainties in a microgrid," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 2292-2300, Mar. 2018.
- [4] T. Zhao and Z. Ding, "Distributed agent consensus-based optimal resource management for microgrids," *IEEE Trans. Sustain. Energy*, vol. 9, no. 1, pp. 443-452, Jan. 2018.
- [5] J. Li, C. Zhang, Z. Xu, J. Wang, J. Zhao, and Y. A. Zhang, "Distributed transactive energy trading framework in distribution networks," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 7215-7227, Nov. 2018.
- [6] A. Bernstein and E. Dall'Anese, "Real-time feedback-based optimization of distribution grids: a unified approach," *IEEE Trans. Control Netw. Syst.*, vol. 6, no. 3, pp. 1197-1209, Sept. 2019.
- [7] C. Enyioha, S. Magnússon, K. Heal, N. Li, C. Fischione, and V. Tarokh, "On variability of renewable energy and online power allocation," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 451-462, Jan. 2018.
- [8] W. Ma, J. Wang, V. Gupta, and C. Chen, "Distributed energy management for networked microgrids using online ADMM with regret," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 847-856, Mar. 2018.
- [9] J. Li, Z. Xu, J. Zhao, and C. Zhang, "Distributed online voltage control in active distribution networks considering PV curtailment," *IEEE Trans. Ind. Inform.*, in press, 2019.
- [10] E. Dall'Anese, S. S. Guggilam, A. Simonetto, Y. C. Chen, and S. V. Dhople, "Optimal regulation of virtual power plants," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1868-1881, Mar. 2018.
- [11] X. Zhou, E. Dall'Anese, L. Chen, and A. Simonetto, "An incentive-based online optimization framework for distribution grids," *IEEE Trans. Autom. Control*, vol. 63, no. 7, pp. 2019-2031, July 2018.
- [12] M. Song, C. Gao, H. Yan, and J. Yang, "Thermal battery modeling of inverter air conditioning for demand response," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 5522-5534, Nov. 2018.
- [13] Y. Li, G. Qu and N. Li, "Using predictions in online optimization with switching costs: a fast algorithm and a fundamental limit," in *2018 Annual American Control Conference (ACC)*, Milwaukee, WI, 2018, pp. 3008-3013.
- [14] A. Di Giorgio, F. Liberati, A. Lanna, A. Pietrabissa and F. D. Priscoli, "Model predictive control of energy storage systems for power tracking and shaving in distribution grids," *IEEE Trans. Sustain. Energy*, vol. 8, no. 2, pp. 496-504, Apr. 2017.
- [15] X. Zhou, E. Dall'Anese, and L. Chen, "Online stochastic optimization of networked distributed energy resources," *IEEE Trans. Autom. Control*, in press, 2019.
- [16] B. Li, T. Chen, X. Wang and, G. B. Giannakis, "Real-time energy management in microgrids with reduced battery capacity requirements," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1928-1938, Mar. 2019.
- [17] M. J. Neely, "Stochastic network optimization with application to communication and queueing systems," San Rafael, CA, USA: Morgan and Claypool, 2010.
- [18] S. Fan, G. He, X. Zhou, and M. Cui, "Online optimization for networked distributed energy resources with time-coupling constraints (Extended version)," [Online].
- [19] J. Koshal, A. Nedic, and U. Y. Shanbhag, "Multiuser optimization: Distributed algorithms and error analysis," *SIAM J. Optim.*, vol. 21, no. 3, pp. 1046-1081, 2011.
- [20] S. Boyd and L. Vandenberghe, "Convex Optimization," New York, NY, USA: Cambridge Univ. Press, 2004.
- [21] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Trans. Power Deliv.*, vol. 4, no. 2, pp. 1401-1407, April 1989.
- [22] Dataport, <https://www.pecanstreet.org/dataport/>.
- [23] OpenEI, [https://openei.org/wiki/Main\\_Page](https://openei.org/wiki/Main_Page).
- [24] Lesage-Landry, Antoine, and Joshua A. Taylor, "Learning to shift thermostatically controlled loads," in *Pro. 50th Hawaii International Conference on System Sciences*. 2017.



## APPENDIX A PROOF OF LEMMA 1

*Proof:* According to the definition of the virtual queue and Lyapunov drift function, we have:

$$\begin{aligned}
& (1/V) \sum_{i \in \mathcal{N}} \Delta(\mathbf{H}_i^t) + \mathbb{E}(\sum_{i \in \mathcal{N}} \widehat{U}_i^t) = \\
& (1/2V) \sum_{i \in \mathcal{N}} \{\mathbb{E}[(\mathbf{H}_i^{t+1})^2 - (\mathbf{H}_i^t)^2 | \mathbf{H}_i^t]\} + \mathbb{E}(\sum_{i \in \mathcal{N}} \widehat{U}_i^t) \\
& = (1/2V) \sum_{i \in \mathcal{N}} \{\mathbb{E}[\sum_{a \in \mathcal{A}_i} ((\xi_{i,a}^{+,t} - \xi_{i,a}^{-,t})^2 + 2H_{i,a}^t (\xi_{i,a}^{+,t} - \xi_{i,a}^{-,t})) | \mathbf{H}_i^t]\} \\
& \quad + \mathbb{E}(\sum_{i \in \mathcal{N}} \widehat{U}_i^t) \\
& \leq (1/V) (1/2) \sum_{i \in \mathcal{N}} \sum_{a \in \mathcal{A}_i} (\xi_{i,a}^{+,t} - \xi_{i,a}^{-,t})^2 + \\
& \quad (1/V) \mathbb{E}[\sum_{i \in \mathcal{N}} \sum_{a \in \mathcal{A}_i} H_{i,a}^t (\xi_{i,a}^{+,t} - \xi_{i,a}^{-,t}) | \mathbf{H}_i^t] + \mathbb{E}(\sum_{i \in \mathcal{N}} \widehat{U}_i^t). \tag{45}
\end{aligned}$$

Then, we define  $B = \frac{1}{2} \sum_{i \in \mathcal{N}} \sum_{a \in \mathcal{A}_i} (\xi_{i,a}^{+, \max} - \xi_{i,a}^{+, \min})^2$ . Thus, we can get Lemma 1.

## APPENDIX B PROOF OF THEOREM 1

With Lemma 1 and theorem 4.8 in [17], we derive  $\overline{\Phi}_i^* \leq \overline{\Phi}_i^* + B/V$  if all the random variables in  $\{\Theta\}^t$  are independent identically distributed (i.i.d.). The result is robust to non-i.i.d., nonergodic situations as proved in theorem 4.13 in [17]. Concurrently, we have  $\overline{\Phi}_i^* \leq \overline{\Phi}^*$  as illustrated in Remark 1. Thus, we can obtain  $\overline{\Phi}_i^* \leq \overline{\Phi}^* + B/V$ .

## APPENDIX C PROOF OF LEMMA 4

First of all, we recall the definition of the utility loss function  $U_i^t: \mathbb{R}^{A_i+2} \rightarrow \mathbb{R}^1$ . The gradient function of  $U_i^t$  is denoted by  $\nabla U_i^t: \mathbb{R}^{A_i+2} \rightarrow \mathbb{R}^{A_i+2}$ . Then, we denote the inverse function of  $\nabla U_i^t$  by  $(\nabla U_i^t)^{-1}: \mathbb{R}^{A_i+2} \rightarrow \mathbb{R}^{A_i+2}$ .

Assume  $\mathbf{x}_1$  and  $\mathbf{x}_2$  are defined in the feasible set of the function  $\nabla U_i^t$ . Let  $\mathbf{y}_1 = \nabla U_i^t(\mathbf{x}_1)$  and  $\mathbf{y}_2 = \nabla U_i^t(\mathbf{x}_2)$ , we have  $\mathbf{x}_1 = (\nabla U_i^t)^{-1}(\mathbf{y}_1)$  and  $\mathbf{x}_2 = (\nabla U_i^t)^{-1}(\mathbf{y}_2)$ , where  $(\nabla U_i^t)^{-1}$  is the inverse function of  $\nabla U_i^t$ . Since  $U_i^t$  is  $\sigma$ -strongly convex (see Assumption 1), we have:

$$(\mathbf{x}_1 - \mathbf{x}_2)^T [\nabla U_i^t(\mathbf{x}_1) - \nabla U_i^t(\mathbf{x}_2)] \geq \sigma \|\mathbf{x}_1 - \mathbf{x}_2\|^2 \tag{46}$$

By substituting  $\mathbf{y}_1$  and  $\mathbf{y}_2$  into (46), we will get:

$$\begin{aligned}
& (\mathbf{y}_1 - \mathbf{y}_2)^T [(\nabla U_i^t)^{-1}(\mathbf{y}_1) - (\nabla U_i^t)^{-1}(\mathbf{y}_2)] \\
& \geq \sigma \|(\nabla U_i^t)^{-1}(\mathbf{y}_1) - (\nabla U_i^t)^{-1}(\mathbf{y}_2)\|^2
\end{aligned} \tag{47}$$

Following the Cauchy-Schwartz inequality, (47) is transferred to:

$$[(\nabla U_i^t)^{-1}(\mathbf{y}_1) - (\nabla U_i^t)^{-1}(\mathbf{y}_2)] \leq \frac{1}{\sigma} \|\mathbf{y}_1 - \mathbf{y}_2\| \tag{48}$$

Thus,  $(\nabla U_i^t)^{-1}$  is  $\frac{1}{\sigma}$ -Lipschitz continuous.

## APPENDIX D PROOF OF THEOREM 2

Before proving the (36), we have the following reasoning:

$$\begin{aligned}
& \|\mathbf{d}^{t,*} - \mathbf{d}^{t+1}\|^2 \\
& \stackrel{(a)}{=} \|P_{\mathbb{R}_+}[\mathbf{d}^{t,*} + \delta \nabla D^t(\mathbf{d}^{t,*})] - P_{\mathbb{R}_+}[\mathbf{d}^t + \delta \nabla D^{t,m}(\mathbf{d}^t)]\|^2 \\
& \stackrel{(b)}{\leq} \|\mathbf{d}^{t,*} - \mathbf{d}^t + \delta [\nabla D^t(\mathbf{d}^{t,*}) - \nabla D^t(\mathbf{d}^t) + \\
& \quad \nabla D^t(\mathbf{d}^t) - \nabla D^{t,m}(\mathbf{d}^t)]\|^2 \\
& \stackrel{(c)}{\leq} 2 \|\mathbf{d}^{t,*} - \mathbf{d}^t\|^2 + 4\delta (\mathbf{d}^t - \mathbf{d}^{t,*})^T [\nabla D^t(\mathbf{d}^t) - \nabla D^t(\mathbf{d}^{t,*})] \\
& \quad + 2\delta^2 \|\nabla D^t(\mathbf{d}^t) - \nabla D^t(\mathbf{d}^{t,*})\|^2 + 2(\delta e)^2 \\
& \stackrel{(d)}{\leq} (2 - \frac{4\delta \sigma_D L_D}{\sigma_D + L_D}) \|\mathbf{d}^{t,*} - \mathbf{d}^t\|^2 + 2(\delta e)^2 + \\
& \quad (2\delta^2 - \frac{4\delta}{\sigma_D + L_D}) \|\nabla D^t(\mathbf{d}^t) - \nabla D^t(\mathbf{d}^{t,*})\|^2.
\end{aligned} \tag{49}$$

where  $\nabla D^{t,m}(\mathbf{d}^t) = g^t(\mathbf{z}^{t,m})$  is the measurement-enabled gradient that is used in LOOD. The equality (a) comes from the dual variables update policy; (b) is due to the non-expansiveness property of the projection operator; (c) considers Assumption 4. For inequality (d), We consider the fact that:

$$\begin{aligned}
& [\nabla D^t(\mathbf{d}^t) - \nabla D^t(\mathbf{d}^{t,*})]^T (\mathbf{d}^t - \mathbf{d}^{t,*}) \leq -(\frac{\sigma_D L_D}{\sigma_D + L_D} \|\mathbf{d}^t - \mathbf{d}^{t,*}\|^2 \\
& \quad - \mathbf{d}^{t,*})^2 + \frac{1}{\sigma_D + L_D} \|\nabla D^t(\mathbf{d}^t) - \nabla D^t(\mathbf{d}^{t,*})\|^2.
\end{aligned} \tag{50}$$

Eq. (50) holds because  $D^t(\cdot)$  is  $\sigma_D$ -strongly concave and its gradient function is  $L_D$ -Lipschitz continuous (see Lemma 5). The proof of Eq. (50) can be referred to Theorem 2.1.12 in [20].

Then, if  $0 < \delta \leq 1/(\sigma_D + L_D)$  holds, the last term of (49) must be non-positive. So, we can obtain:

$$\begin{aligned}
& \|\mathbf{d}^{t,*} - \mathbf{d}^{t+1}\| \leq \sqrt{\kappa^2 \|\mathbf{d}^{t,*} - \mathbf{d}^t\|^2 + (\delta e)^2} \\
& \leq \kappa \|\mathbf{d}^{t,*} - \mathbf{d}^t\| + \delta e,
\end{aligned} \tag{51}$$

where  $\kappa = 2 - 4\delta[\sigma_D L_D/(\sigma_D + L_D)]$ . It is readily to obtain that:

$$\begin{aligned}
& \kappa = 2 - 4\delta[\sigma_D L_D/(\sigma_D + L_D)] \\
& \leq 2 - \frac{4\sigma_D L_D}{(\sigma_D + L_D)^2} \leq 2 - \frac{4\sigma_D L_D}{2\sigma_D L_D} \leq 1.
\end{aligned} \tag{52}$$

Then, we have:

$$\begin{aligned}
& \|\mathbf{d}^{t+1} - \mathbf{d}^{t+1,*}\| \\
& \stackrel{(a)}{\leq} \|\mathbf{d}^{t+1} - \mathbf{d}^{t,*}\| + \|\mathbf{d}^{t,*} - \mathbf{d}^{t+1,*}\| \\
& \stackrel{(b)}{\leq} \kappa \|\mathbf{d}^{t,*} - \mathbf{d}^t\| + \delta e + \Delta_d \\
& \stackrel{(c)}{\leq} \kappa^t \|\mathbf{d}^{t=0} - \mathbf{d}^{t=0,*}\| + \sum_{\tau=0}^t \kappa^\tau (\Delta_d + \delta e) \\
& = \kappa^t \|\mathbf{d}^{t=0} - \mathbf{d}^{t=0,*}\| + \frac{1 - \kappa^t}{1 - \kappa} (\Delta_d + \delta e) \\
& = \underbrace{(\kappa)^t (\|\mathbf{d}^{t=0} - \mathbf{d}^{t=0,*}\| - \frac{\Delta_d}{1 - \kappa})}_{:= \bar{h}^t \text{ (transient term)}} + \underbrace{\frac{\Delta_d + \delta e}{1 - \kappa}}_{:= \ell \text{ (fixed term)}}.
\end{aligned} \tag{53}$$

In (53), the inequality (a) comes from the dual variables update policy in LOOD and the triangle inequality; (b) is based on (51) and Assumption 3; (c) is resulted from using triangle inequality repeatedly. According to (53) the first term of the reasoning result, denoted by  $\bar{h}^t$ , is transient that will vanish to be 0 when  $t \rightarrow \infty$  due to the fact that  $0 \leq \kappa \leq 1$  (as shown in Theorem 2). The second term denoted by  $\ell$  is fixed.

APPENDIX E  
PROOF OF COROLLARY 1

$$\begin{aligned}
& \| \mathbf{c}^{t+1} - \mathbf{c}^{t+1,*} \| = \| (1-\phi)c(\mathbf{d}^{t+1}) + \phi c(\mathbf{d}^t) - c(\mathbf{d}^{t+1,*}) \| \\
& \stackrel{(a)}{\leq} (1-\phi) \| c(\mathbf{d}^{t+1}) - c(\mathbf{d}^{t+1,*}) \| + \phi \| c(\mathbf{d}^t) - c(\mathbf{d}^{t+1,*}) \| \\
& \stackrel{(b)}{\leq} (1-\phi)L_c(\bar{h}^t + \ell) + \phi L_c \| \mathbf{d}^t - \mathbf{d}^{t,*} + \mathbf{d}^{t,*} - \mathbf{d}^{t+1,*} \| \quad (54) \\
& \stackrel{(c)}{\leq} (1-\phi)L_c(\bar{h}^t + \ell) + \phi L_c[(\bar{h}^{t-1} + \ell) + \Delta_d] \\
& = \underbrace{(1-\phi)L_c\bar{h}^t + \phi L_c\bar{h}^{t-1}}_{\text{transient term}} + \underbrace{L_c(\ell + \phi\Delta_d)}_{\text{fixed term}}.
\end{aligned}$$

The inequality (a) uses triangle inequality; (b) is according to Lipschitz continuity of the function  $c(\cdot)$  (see Remark 3) and Theorem 2; (c) comes from Assumption 3 and Theorem 2. Similar to arguments in Theorem 2, the resultant two terms comprise a transient term that vanishes when  $t \rightarrow \infty$ . Thus, we obtain the result in (37).

APPENDIX F  
PROOF OF COROLLARY 2

*Proof:*

Before the proof of this corollary, we rewrite the incentive vector at node  $i$  i.e.,  $\mathbf{c}_i^t = [\alpha_i^t, \beta_i^t]^\top$  as  $\mathbf{c}_i^{D,t} = \underbrace{[\alpha_i^t, \beta_i^t, \alpha_i^t, \dots, \alpha_i^t]}_{(2+A_i)\text{-dimensions}}$ .

Then, we have:

$$\begin{aligned}
& \| \mathbf{z}_i^{t+1} - \mathbf{z}_i^{t+1,*} \| \\
& \stackrel{(a)}{=} \| P_{\mathcal{Z}_i^{t+1}} \{ [\nabla U_i^{t+1}]^{-1}(\mathbf{c}_i^{D,t+1}) \} - P_{\mathcal{Z}_i^{t+1}} \{ [\nabla U_i^{t+1}]^{-1}(\mathbf{c}_i^{D,t+1,*}) \} \| \\
& \stackrel{(b)}{\leq} \| [\nabla U_i^{t+1}]^{-1}(\mathbf{c}_i^{D,t+1}) - [\nabla U_i^{t+1}]^{-1}(\mathbf{c}_i^{D,t,*}) \| + \\
& \quad \| [\nabla U_i^{t+1}]^{-1}(\mathbf{c}_i^{D,t,*}) - [\nabla U_i^{t+1}]^{-1}(\mathbf{c}_i^{D,t+1,*}) \| \\
& \stackrel{(c)}{\leq} \frac{1}{\sigma} (\| \mathbf{c}_i^{D,t+1} - \mathbf{c}_i^{D,t,*} \| + \| \mathbf{c}_i^{D,t,*} - \mathbf{c}_i^{D,t+1,*} \|) \\
& \stackrel{(d)}{\leq} \frac{(1+A_i)}{\sigma} (\| \mathbf{c}_i^{t+1} - \mathbf{c}_i^{t+1,*} \| + \| \mathbf{c}_i^{t,*} - \mathbf{c}_i^{t+1,*} \|) \\
& \stackrel{(e)}{\leq} \underbrace{\frac{(1+A_i)}{\sigma} (1-\phi)L_c\bar{h}^t + \phi L_c\bar{h}^{t-1}}_{:=\psi_i \text{ transient term}} + \underbrace{\frac{L_c(1+A_i)}{\sigma} [\ell + (1+\phi)\Delta_d]}_{:=Y_i \text{ fixed term}}. \quad (55)
\end{aligned}$$

In (55), the equality (a) is according to the optimal condition of the local problem  $\mathcal{P}_{L,i}^t$ ; (b) comes from triangle inequality and (c) considers Lemma 4; (e) uses the results Corollary 1. The resultant two terms also comprise a transient term that will vanish when  $t \rightarrow \infty$ .

APPENDIX G  
PROOF OF THEOREM 3

*Proof:*

First of all, we can characterize the difference between the two optimizers generated by LOOD and  $\mathcal{P}_2$  at each time slot:

$$\begin{aligned}
& \| \Phi_{OD}^{t,*} - \Phi_i^{t,*} \| = \| \sum_{i \in \mathcal{N}} U_i^t(\mathbf{z}_i^{t,*}) - \sum_{i \in \mathcal{N}} U_i^t(\mathbf{z}_i^t) \| \\
& \leq \sum_{i \in \mathcal{N}} \| U_i^t(\mathbf{z}_i^{t,*}) - U_i^t(\mathbf{z}_i^t) \| \leq NL \| \mathbf{z}_i^{t,*} - \mathbf{z}_i^t \|. \quad (56)
\end{aligned}$$

Thus,

$$\begin{aligned}
& \overline{\Phi_{OD}^{t,*}} - \overline{\Phi_i^{t,*}} \leq \\
& \stackrel{(a)}{=} \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t \in \mathcal{T}} (\Phi_{OD}^{t,*} - \Phi_i^{t,*}) \\
& \stackrel{(b)}{\leq} NL (\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t \in \mathcal{T}} \| \mathbf{z}_i^{t,*} - \mathbf{z}_i^t \|) \\
& \stackrel{(c)}{\leq} NLY_i \leq NLY, \quad (57)
\end{aligned}$$

where the equality (a) is based on the definition of time-average of a variable; inequality (b) and (c) come from Eq. (56) and Corollary 2, respectively. Then, due to Theorem 1 that  $\overline{\Phi_i^{t,*}} \leq \overline{\Phi^*} + \frac{B}{V}$ , we have

$$\overline{\Phi_{OD}^{t,*}} \leq \overline{\Phi^*} + B/V + NLY. \quad (58)$$