Data-driven Coordination of Distributed Energy Resources for Active Power Provision
Hanchen Xu, Student Member, IEEE, Alejandro D. Domínguez-García, Member, IEEE and Peter W. Sauer, Life Fellow, IEEE

Abstract—In this paper, we propose a framework for coordinating distributed energy resources (DERs) connected to a lossy power distribution system, the model of which is unknown, so that they collectively provide a specified amount of active power to the bulk power system as quantified by the power exchange between both systems at the bus interconnecting them. The proposed framework consists of (i) an input-output (IO) system model that represents the relation between the DER active power injections (inputs), and the total active power exchanged between the distribution and bulk power systems (output); (ii) an estimator that aims to estimate the IO model parameters, and (iii) a controller that determines the DER active power injections so that the power exchanged between both systems equals to the specified amount. We formulate the estimation problem and the control problem as quadratic programs with box constraints and solve them using the projected gradient descent algorithm, which we show under some mild assumptions, the estimated parameters converge to their true values, and the total active power exchanged between both systems converges to the required amount. The effectiveness of the framework is validated via numerical simulations using the IEEE 123-bus test feeder.

I. INTRODUCTION

In the modernization that electric power systems are currently undergoing, one goal is to massively integrate distributed energy resources (DERs) into power distribution systems [1]. These DERs, which include distributed generation resources, energy storage, demand response resources, and typically have small capacities, may be coordinated so as to collectively provide grid support services, e.g., reactive power support for voltage control [2]–[4], active power control for frequency regulation [5], [6], and energy management [7].

In this paper, we focus on the problem of coordinating the response of a set of DERs in a lossy power distribution system so that they collectively provide some amount of active power to the bulk power system. Specifically, the DERs will be requested to collectively provide—in real time—a certain amount of active power at the bus where the power distribution system is interconnected with the bulk power system. In order for the DERs to fulfill such request, it is necessary to develop appropriate schemes that explicitly take into consideration the losses incurred. One approach to include the losses in the problem formulation is to utilize a power-flow-like model of the system obtained offline. However, such a model for the power distribution system may not be available or if so, it may not be accurate.

As an alternative to the aforementioned model-based approach, data-driven approaches have been demonstrated to be very effective in situations where models are not readily available [8]–[15]. The fundamental idea behind data-driven approaches is to describe the system behavior by a linear time-varying (LTV) input-output (IO) model, and estimate the parameters of this model via regression using measurements of pertinent variables [8], [9]. Many previous works have applied data-driven approaches to power system problems, both in a steady-state setting [10]–[13], and a dynamical setting [14], [15]. For example, in [10], the authors developed a data-driven framework to estimate linear sensitivity distribution factors such as injection shifting factors [11]; they further proposed a data efficient sparse representation to estimate these sensitivities [11]. This framework was later tailored to the problem of estimating the power flow Jacobian [12]. In [13], the authors used the estimation framework proposed in [10] to solve the security constrained economic dispatch problem. Data-driven approaches have also been used to develop power systems stabilizers [14], and damping controls [15]. We refer interested readers to [16], [17] for a nice review of data-driven approaches and their applications in a variety of other areas.

Yet, due to the collinearity issue [18], [19], the regression problem may be ill-conditioned, thus resulting in large error in the estimation. Moreover, when estimating the parameters, the regression problem needs to be solved in a centralized way. This centralized approach relies heavily on information exchange between the centralized estimator/controller and every individual DER, which will impose a heavy burden on the communication network, and may subject to failure when any single communication link fails. Therefore, distributed control schemes that only requires information exchange between DERs and their neighbors are more desirable [20]–[22].

In this paper, we pursue the data-driven approach to develop a framework for coordinating the response of a set of DERs. The proposed framework consists of three components, namely (i) a model of the system describing the relation between the variables of interest to the problem, i.e., DER active power injections and power exchanged between the distribution and bulk power systems, (ii) an estimator, which provides estimates of the parameters that populate the model in (i); and (iii) a controller that uses the model in (i) with the parameters estimated via (ii) to determine the active power injection set-points of the DERs. Specifically, an LTV IO model is adopted as the system model to capture the relation between the DER active power injections (inputs), and the total active power exchange (output). The parameters in this...
model are estimated by the estimator via the solution of a box-constrained quadratic program, obtained by using the projected gradient descent algorithm. The controller then uses the estimated parameter to update the DER active power injections via the solution of another box-constrained quadratic program, also obtained by using the projected gradient descent algorithm. It is important to note that the projected gradient descent algorithm, together with a consensus-type algorithm, allows fully distributed implementation of the DER coordination scheme. In order to overcome the problem of collinearity in the measurements, we introduce random perturbations in the update rule used by the the controller. We show that the estimation and control algorithms converge almost surely (a.s.) under some conditions, i.e., the estimated parameters converge to the true parameters and the total provided active power converges to the required amount. The major contributions of this paper include the data-driven coordination framework, the algorithms to solve the control and estimation problems, as well as the proof of the algorithm convergence.

The remainder of this paper is organized as follows. Section II describes the power distribution system model and the DER coordination problem of interest. Section III describes the components of the data-driven DER coordination framework. A description of the algorithm used in the framework, as well as its convergence and distributed implementation, is provided in Section IV. The proposed framework is illustrated and validated via numerical simulations on a IEEE 123-bus test feeder in Section V. Concluding remarks are presented in Section VI.

II. PRELIMINARIES

In this section, we introduce the power distribution system model adopted in this work, which consists of a physical layer and a cyber layer. We then discuss the DER coordination problem of interest.

A. Power Distribution System Model

1) Physical Layer: Consider a power distribution network that represented by a directed graph \( G = (N, L) \), which consists of a set of buses indexed by the elements in the set \( N = \{0, 1, \ldots, N\} \), and a set of distribution lines indexed by the elements in some set \( L \subseteq N \times N \). Assume bus 0 corresponds to a substation bus, which is the only connection of the distribution system to the bulk power system. Further, assume that bus 0 is an ideal voltage source. Without loss of generality, assume there is at most one DER and/or load at each bus. For simplicity, we refer to the DER/load at bus \( i \) as DER/load \( i \). Let \( N^g \), where \( N^g \subseteq N := N \setminus \{0\} \), denote the DER index set; and let \( N^d \), where \( N^d \subseteq N \), denote the load index set. Assume \( |N^g| = n \), where \( |\cdot| \) denotes the cardinality of a set.

Let \( p_i^d \) and \( q_i^d \) respectively denote the active and reactive power loads at bus \( i \), \( i \in N^d \), and define \( p^d = [p_i^d]^T \), and \( q^d = [q_i^d]^T \). Let \( p_i^g \) and \( q_i^g \) respectively denote the active and reactive power injections from DER \( i \), \( i \in N^g \), and define \( p^g = [p_i^g]^T \), and \( q^g = [q_i^g]^T \). Let \( p^g \) and \( q^g \) respectively denote the minimum and maximum active power that can be provided by DER \( i \), \( i \in N^g \), and define \( p_i^g = [p_i^g]^T \), and \( q_i^g = [q_i^g]^T \). Similarly, let \( q_i^g \) and \( q_i^g \) respectively denote the minimum and maximum reactive power that can be provided by DER \( i \), \( i \in N^g \), and define \( q_i^g = [q_i^g]^T \), and \( q_i^g = [q_i^g]^T \).

Let \( y \) denote the active power exchanged between the distribution and bulk power systems via bus 0, defined to be positive if the flow is from the substation to the bulk power system. Conceptually, \( y \) can be represented as a function of \( p_i^d \), \( q_i^d \), \( p_i^g \), \( q_i^g \). Note also \( q_i^g \) is typically set according to some specific reactive power control rules to achieve certain objectives such as constant voltage magnitude or constant power factor [6], and thus is a function of \( p_i^d \), \( p_i^g \), \( q_i^d \). Therefore, \( y \) can be written as a function of \( p_i^d \), \( p_i^g \), \( q_i^d \) as follows:

\[
y = f(p_i^d, p_i^g, q_i^d).
\]

For notational simplicity, define \( u := p_i^g, \ u := p_i^d, \ \pi := p_i^d, q_i^d \). Also, define \( \pi := [(p_i^g)^T, \ (q_i^g)^T]^T \); then, \( y \) can be written as:

\[
y = f(u, \pi).
\]

The explicit form of \( f \) is difficult to obtain; however, we can make the following assumptions about \( f \), which are reasonable in practice as discussed below.

Assumption 1. Within a short time horizon \( u \) changes, \( \pi \) remains constant, thus, changes in \( y \) only depend on changes in \( u \).

Assumption 2. The function \( f \) is differentiable and its first order partial derivatives with respect to \( u \) belong to \([b_1, b_2]\] \), where \( b_1, b_2 \) are some known constants. In addition, \( \frac{\partial f}{\partial u} \) is a Lipschitz function, i.e., there exists \( b_2 > 0 \) such that

\[
\left| \frac{\partial f}{\partial u}_a - \frac{\partial f}{\partial u}_b \right| \leq b_2|a - b|,
\]

where \( a, b \in [u, \pi] \), and \( ||\cdot|| \) denotes the \( L_2 \)-norm.

Assumption 1 basically implies that active power and reactive power loads remain constant over the time horizon during which the DER power injections change. Assumption 2 implies that, for fixed \( \pi \) (as in Assumption 1), the rate of change in \( y \) is bounded for bounded changes in the DER active power injections. In addition, the total active power provided to the bulk power system will increase when more active power is injected in the power distribution system. Both assumptions are reasonable in a real power system.

2) Cyber Layer: The communication network that enables the information exchange between DERs can be described by a strongly connected directed graph \( G = (V, E) \), where \( V = \{1, \ldots, n\} \) is the vertex set (each vertex/node corresponds to a DER), and \( E \subseteq V \times V \) is the set of edges, where \((j, i) \in E \) if node \( j \) can receive information from node \( i \). Without loss of generality, assume node 1 can communicate directly with the substation, which is at bus 0 in the physical layer. We allow each node to receive information from itself, i.e., \((j, j) \in E \) for any \( j \in V \). All nodes that can transmit information to and receive information from node \( j \) are referred to as the in-neighbors and the out-neighbors of node \( j \), respectively. The set of in-neighbors of node \( j \) (including itself) is denoted by \( V_j^I \), whereas the set of out-neighbors is denoted by \( V_j^O \).
The in-degree and the out-degree of node $j$ are, respectively, $D_j^− = |V_j^−|$ and $D_j^+ = |V_j^+|$. Let $d(G^c)$ denote the diameter of a strongly connected graph $G^c$, i.e., the length of the longest among all shortest paths connecting any pair of nodes.

### B. DER Coordination Problem

The DERs in the distribution system can collectively provide active power to the bulk power system as quantified by the power exchange between both systems at the substation bus. For example, the DERs can provide demand response vide active power to the bulk power system as quantified by $y_j$. The objective of the DER coordination problem is to drive $y$ to $y^*$ by adjusting the DER active power injections, $p^j$. This is difficult, however, when the model describing the power exchange with the bulk power system, as captured by $f$, is unknown. In this paper, we will resort to a data-driven approach to tackle this problem.

### C. Ratio Consensus Algorithm

Assume that the communication network enabling the information exchange between DERs conforms to a strongly connected graph $G^c = \{V, E\}$ as described in Section II-A2. In ratio consensus, each node $j$ maintains two internal states, denoted by $\mu_j$ and $\nu_j$, and updates them iteratively. Let $\mu_j[l]$ and $\nu_j[l]$ denote the respective values of $\mu_j$ and $\nu_j$ at iteration $l$, $l \in \mathbb{N}$, which are updated as follows:

\[
\begin{align*}
\mu_j[l + 1] &= \sum_{i \in V_j} \frac{1}{D_j} \mu_i[l], \\
\nu_j[l + 1] &= \sum_{i \in V_j} \frac{1}{D_j} \nu_i[l].
\end{align*}
\]

Assuming $\nu_j[l] > 0$, $\forall j \in V$, node $j$ computes

\[
\gamma_j[l] = \frac{\mu_j[l]}{\nu_j[l]}
\]

Then as shown in (20), it follows:

\[
\gamma^* := \lim_{l \to \infty} \gamma_j[l] = \frac{1}{\sum_{i=1}^n \nu_i[0]} \quad \forall j \in V.
\]

We will show later that ratio consensus is a primitive to solve the DER coordination problem in a distributed fashion.

### III. DER Coordination Framework

In this section, we describe our proposed framework that consists of an LTV IO model, an estimator, and a controller.

#### A. Input-Output System Model

Unless otherwise noted, throughout this paper, $x[k]$ denotes the value that some variable $x$ takes at time instant $k$. We note that $y[k]$ is a result of $u[k - 1]$; thus, it follows from (2) and Assumption 2 that $y[k - 1] = f(u[k - 2], \pi)$ and $y[k] = f(u[k - 1], \pi)$. Then, by the Mean Value Theorem, there exists $a_k \in [0, 1]$ and $\hat{u}[k] = a_k u[k - 1] + (1 - a_k) u[k - 2]$ such that

\[
y[k] - y[k - 1] = f(u[k - 1], \pi) - f(u[k - 2], \pi) = \phi^T[k](u[k - 1] - u[k - 2]),
\]

where $\phi^T[k] := [\phi_i[k]] = \frac{\partial f}{\partial u} |_{\hat{u}[k]}$. $\phi$ is referred to as the sensitivity vector at instant $k$. It follows from Assumption 1 that $\phi_i[k] \in [b_i, \overline{b_i}], i = 1, \ldots, n$. Therefore, at any time instant $k$, (6) can be transformed into the following equivalent LTV IO model:

\[
y[k] = y[k - 1] + \phi^T[k](u[k - 1] - u[k - 2]).
\]

#### B. Estimator

At time instant $k$, the objective of the estimator is to obtain an estimate of $\phi[k]$, denoted by $\hat{\phi}[k]$, using measurements collected up to time instant $k$; we formulate this estimation problem as follows:

\[
\hat{\phi}[k] = \arg \min_{\phi \in \Omega} J^e(\hat{\phi}) = \frac{1}{2} (y[k] - \hat{y}[k])^2,
\]

subject to

\[
y[k] = y[k - 1] + \hat{\phi}^T[k](u[k - 1] - u[k - 2]),
\]

where $\Omega = [b, \overline{b}]^n$, $J^e(\cdot)$ is the cost function of the estimator, and $\hat{y}[k]$ is the value of $y[k]$ estimated by the IO model at time instant $k$. Essentially, $\hat{\phi}[k]$ aims to find $\hat{\phi}$ that minimizes the squared error between the estimated value and the true value of $y$. Then, $\hat{\phi}[k]$ is used in the controller to determine the control for the upcoming time interval.

#### C. Controller

At time instant $k$, the objective of the controller is to determine the control $u[k]$ such that $y[k + 1] = y^*$. Thus, by using the model in (6), replacing $\hat{\phi}[k]$ with $\hat{\phi}[k]$, we can formulate the control problem as follows:

\[
u[k] = \arg \min_{u} J^c(u) = \frac{1}{2} (y^* - \hat{y}[k + 1])^2,
\]

subject to

\[
y[k + 1] = y[k] + \hat{\phi}^T[k](u[k] - u[k - 1]),
\]

where $u = \{u \in \mathbb{R}^n : u \leq \overline{u} \leq \underline{u}\}$, and $J^c(\cdot)$ is the cost function of the controller. Note that $\hat{\phi}[k]$ is used in (6) to predict the value of $y[k + 1]$ for a given control $u$.

### IV. Algorithm and Its Convergence

In this section, we propose a projected gradient descent algorithm to solve the estimation/control problem. We then prove the convergence of the proposed algorithm. We show in the end how the projected gradient descent algorithm can be implemented in a distributed fashion leveraging ratio consensus.

1 We adopt the convention that the partial derivative of a scalar function with respect to a vector is a row vector.
A. Algorithm

We first describe the basic workflow of the proposed algorithm. At time instant $k$, $y[k]$ is measured and becomes available to the estimator, which uses it to update the estimate of the sensitivity vector. The updated estimate of the sensitivity vector, $\hat{\phi}[k]$, is then used in the controller to determine the control, $u[k]$. Then, the DERs are instructed to change their active power set-points based on $u[k]$. At time instance $k+1$, the estimation and control update steps are repeated once $y[k+1]$ becomes available. The sequential process described above is illustrated as follows:

\[
\cdots u[k-1] \rightarrow y[k] \rightarrow \hat{\phi}[k] \rightarrow u[k] \rightarrow y[k+1] \rightarrow \hat{\phi}[k+1] \cdots
\]

Problems (7) and (8) can be solved using the projected gradient descent method. Let $P_{V_1 \rightarrow V_2}$ denote the projection operator from a vector space $V_1$ to its (arbitrary) subspace $V_2$, i.e.,

\[
P_{V_1 \rightarrow V_2}(v_1) = \arg\min_{v_2 \in V_2} \|v_2 - v_1\|,
\]

where $v_1 \in V_1$. For ease of notation, when the vector space to which $v_1$ belongs is unambiguous, we simply write $P_{V_2}(v_1)$ instead of $P_{V_1 \rightarrow V_2}(v_1)$.

Define the tracking error at time instant $k$ as $e[k] = y[k] - y^*$. In addition, define $\Delta y[k] = y[k] - y[k-1]$ and $\Delta u[k] = u[k] - u[k-1]$. The partial derivative vector of $J^*(\phi)$ with respect to $\phi$ is

\[
\frac{\partial J^*(\phi)}{\partial \phi} = \Delta u[k-1][\Delta u^T[k-1]\phi - \Delta y[k]],
\]

and that of $J^*(u)$ with respect to $u$ is

\[
\frac{\partial J^*(u)}{\partial u} = \phi[k](\phi^T[k](u - u[k-1]) + e[k]).
\]

Instead of solving both (7) and (8) to completion, we iterate the projected gradient descent algorithm that would solve them for one step at each time instant. Specifically, at time instant $k$, we evaluate the new gradient at $\hat{\phi}[k-1]$ and $u[k-1]$ and iterate once. Thus, by using (10) and (9), the update rules for $\hat{\phi}$ and $u$, respectively, are

\[
\hat{\phi}[k] = P_{\Delta \phi} \left( \hat{\phi}[k-1] - \alpha_k \Delta u[k-1] \right),
\]

\[
u[k] = P_U \left( u[k-1] - \beta_k e[k] \hat{\phi}[k] \right),
\]

where $\alpha_k > 0$ and $\beta_k > 0$ are the step sizes at time instant $k$.

In order to obtain richer information about the system, we introduce random perturbations in the control update rule. Define $W[k] = \text{diag}(w_1[k], \ldots, w_n[k])$, where $w_i[k]$’s are independent random variables that follow a Bernoulli distribution with parameter $p = 0.5$. Then, the control update rule in (12) is modified, resulting in:

\[
u[k] = P_U \left( u[k-1] - \beta_k e[k] W[k] \hat{\phi}[k] \right).
\]

Intuitively, this means that at each time instant, the control of each DER is updated with a probability of $p = 0.5$.

The random perturbation in the control is key to establish convergence of the parameter estimation process. The DER coordination algorithm, along with its initialization, is summarized in Algorithm 1.

Algorithm 1: DER Coordination Algorithm

Input: $y$, $y^*$, $\delta > 0$
Output: $u$, $\phi$
Initialization: set $y[0] = 0$, $\hat{\phi}[0] = [1, \ldots, 1]^T$, $u[0] = P_U(-\beta_0 W[0] \phi[0] e[0])$, $u[1] = [0, \ldots, 0]^T$ set $k = 1$
while $|e[k]| > \delta$
do
obtain new measurement of $y[k]$
compute $\Delta e[k] = y[k] - y^*$
compute $\Delta y[k] = y[k] - y[k-1]$
compute $\Delta u[k-1] = u[k-1] - u[k-2]$
update the sensitivity vector estimate, $\hat{\phi}$, according to

\[
\hat{\phi}[k] = P_{\Delta \phi} \left( \hat{\phi}[k-1] - \alpha_k \Delta u[k-1] \right)
\]

\[
(\Delta u^T[k-1] \hat{\phi}[k-1] - \Delta y[k])
\]

update the control vector, $u$, according to

\[
u[k] = P_U \left( u[k-1] - \beta_k e[k] W[k] \hat{\phi}[k] \right)
\]
change DER active power injections to $\|u[k]\|$
set $k = k + 1$
end

B. Convergence of Control Update Rule

The main convergence result for the control update rule is stated next.

Theorem 1. Using the estimation update rule in (11) and the control update rule in (13) with $\beta_k \in (\frac{\alpha_k}{2}, \frac{1}{\beta_0})$, where $0 < \epsilon < \frac{\alpha_k^2}{\beta_0}$ is a given parameter, the system attains one of the following equilibria: 1) $e[k]$ converges to 0 a.s.; 2) $e[k]$ converges to some positive constant and $u[k]$ stays at $\bar{u}$; 3) $e[k]$ converges to some negative constant and $u[k]$ stays at $\bar{u}$. In all cases, $\lim_{k \to \infty} \Delta u[k] = 0_n$, where $0_n \in \mathbb{R}^n$ is an all-zeros vector.

Theorem 1 shows something intuitive, i.e., the tracking error will be positive (negative) if the requested active power is less (more) than the minimum (maximum) amount of active power the DERs can provide; otherwise, the tracking error goes to zero a.s.

We note that $\epsilon$ has a direct impact on the convergence rate of the control algorithm. This is more obvious in a deterministic setting, when the control update rule in (13) is used instead of the one in (12). A result on the convergence rate is given in the following corollary.

Corollary 1. Assume $u[k] \neq \bar{u}$ and $u[k] \neq \bar{u}$, $\forall k \in \mathbb{N}$. Using the estimation update rule in (11) and the control update rule in (13) with $\beta_k \in (\frac{\alpha_k}{2}, \frac{1}{\beta_0})$, where $\epsilon > 0$ is a given parameter, $e[k]$ converges to 0 at a rate smaller that $1 - \epsilon$. 


We refer the readers to Appendix A for detailed proofs of the convergence results.

C. Convergence of Estimation Update Rule

Next, we establish the convergence of the estimation update rule. Define the estimation error vector at time instant $k$ as $\varepsilon[k] = \hat{\phi}[k] - \phi[k]$. Since both $\hat{\phi}[k]$ and $\phi[k]$ are bounded, $\varepsilon[k]$ is also bounded. Define $\Delta \phi[k] = \phi[k] - \phi[k-1]$. The convergence result for the estimation update rule is stated next.

**Theorem 2.** Using the estimation update rule in (11) and the control update rule (13), with $\alpha_{k+1} = \frac{2}{\|\Delta u[k]\|^2}$, $\beta_k \in (\frac{\epsilon}{\|n\|^2}, \frac{1}{\|n\|^2})$, where $0 < \epsilon < \frac{\mu^2}{\beta_1}$ is a given parameter, if $u[k] \in (\mu, \overline{\mu})$ and $e[k] \neq 0$, $\forall k \in \mathbb{N}$, then $\|\varepsilon[k]\|$ converges to 0 a.s.

The intuition is that the estimation error goes to zero if the system can be continuously excited (guaranteed by the condition $u[k] \in (\mu, \overline{\mu})$ and $e[k] \neq 0$, $\forall k \in \mathbb{N}$).

We refer the readers to Appendix B for detailed proofs of the convergence results.

D. Distributed Implementation

At each iteration, node $i$, $i \in \mathcal{V}$, needs to update $\hat{\phi}_i$ and $u_i[k]$ according to the following rules from (11) and (13). Once each node learns the values of $y[k]^* - y^*$, $\Delta u^T[k-1] \hat{\phi}[k-1] - \Delta y[k]$, and $\alpha_k$, both updates can be done locally. The learning of $y[k]^* - y^*$ can be done through simple message passing and it can be learned by all nodes after at most $d(G^c)$ iterations. The learning of the $\Delta u^T[k-1] \hat{\phi}[k-1] - \Delta y[k]$ can also be done distributedly through ratio consensus as follows. Each node $i$ executes one copy of the numerator iteration with

$$\mu_i[k] = \begin{cases} \Delta u_i[k-1] \hat{\phi}_i[k-1] - \Delta y_i[k], & \text{if } i = 1, \\ \Delta u_i[k-1] \hat{\phi}_i[k-1], & \text{if } i \in \mathcal{V} \backslash \{1\}, \end{cases}$$

and one copy of the denominator iteration with $\nu_i[k] = \frac{1}{\nu_i[k-1]}$. Updating $\mu_i[k]$ and $\nu_i[k]$ according to (14), by (15), each node $i$ will asymptotically learn $\Delta u^T[k-1] \hat{\phi}[k-1] - \Delta y[k]$, as follows:

$$\lim_{l \to \infty} \frac{\mu_i[l]}{\nu_i[l]} = \Delta u^T[k-1] \hat{\phi}[k-1] - \Delta y[k].$$

The learning of $\alpha_k$ can be done similarly by initializing as follows: $\mu_i[0] = \frac{2}{\alpha_0}$, and $\nu_i[0] = \Delta u^2_i[k-1]$.

The ratio consensus iteration happens between time instant $k$ and time instant $k+1$ on a faster time-scale than that of the control/estimation. In practice, the consensus algorithm needs to be stop based on some criterion, for example, when maximum iteration is reached or using finite time consensus algorithms (see, e.g., [23]). Leveraging ratio-consensus, the data-driven DER coordination is now fully distributed.

V. NUMERICAL SIMULATION

In this section, we illustrate the application of the proposed DER coordination framework and validate the theoretical results presented earlier. A modified single-phase IEEE 123-bus distribution test feeder from [24] (see Fig. 1 for the one-line diagram) is used for all numerical simulations. This balanced test feeder has a total active power load of 3000 kW, and a total reactive power load of 1575 kVAR. DERs are added at buses 19, 26, 38, 49, 56, 64, 78, 89, 99, as indicated by empty circles at the corresponding buses in the one-line diagram of Fig. 1. We assume each DER can output active power from 0 kW to 40 kW. To illustrate the impacts of reactive power control, assume all DERs operate at unity power factor except DERs 78 and 89, which will inject reactive power to maintain a constant voltage magnitude of 0.95 p.u. The underlying nonlinear power flow problem is solved using Matpower [25].

The cyber layer, represented by a directed communication graph, is shown in Fig. 2. Figure 3 shows the process of the nodes reaching consensus on the value of 1 via the ratio consensus algorithm. In all subsequent simulations, we set $\beta_1 = 0.8$, $\overline{\mu}_1 = 1.2$, which are reasonable values for real power systems. We emphasize that a time instant is essentially a control/estimation step, which depends on specific applications and may correspond to different time durations in real systems. Thus, we do not associate the time instant with any units.

A. Tracking Performance

Under this simulation setting, as given in Theorem 1, the upper bound of the control update rule size is $\frac{1}{\max(\mu_1)} \approx 0.0694$. For $y^* = 100$ kW and a constant step size $\beta_k = 0.02$, the DER active power injections are shown in Fig. 4. The
non-smoothness in the active power profiles is caused by the random perturbation imposed on the control update rule. Also as shown in Fig. 5, the convergence rate of the tracking error is not affected by the tracking target, i.e., the total active power required from the bulk power system. The tracking error $e[k]$ under various constant control update rule sizes is shown in Fig. 6. As expected, a larger step size will reduce the tracking error faster than a small step size.

B. Estimation Accuracy

With $\beta_k = 0.02$, true values of the sensitivities, $\phi$, and estimated values of the sensitivities, $\hat{\phi}$, are plotted in Figs. 7 and 8 respectively. While the estimator update rule step size, $\alpha_k$, in the proposed algorithm is adaptive, we also investigate the results when $\alpha_k$ is chosen to be constant. Fig. 9 shows the mean absolute error (MAE) of estimation, i.e., $\frac{\sum_{i=1}^{n} |e_i[k]|}{N}$, under various constant estimation step sizes. As can be seen from Fig. 9, the MAE of estimation will converge to some non-zero constant under constant estimation step sizes.

The impact of the control step sizes on the estimation accuracy is also investigated. Figure 10 shows the MAE of estimation under various control step sizes. With a large control update rule step size, the tracking error $e[k]$ converges to 0 quickly, leading to a situation where the system cannot get sufficient excitation and consequently, the estimation errors cannot be further reduced.

C. Time-varying Tracking Targets

We next illustrate the application of the proposed algorithm when the tracking target is time-varying, i.e., the DERs are

---

Fig. 3. Consensus process in the cyber layer.
Fig. 4. DER active power injections for $\beta_k = 0.02$ and $y^* = 100$ kW. (Legends indicate DER indexes.)
Fig. 5. Tracking error for $\beta_k = 0.02$ and various tracking targets. (Legends indicate values of $y^*$.)
Fig. 6. Tracking error for $y^* = 100$ kW and various constant control step sizes. (Legends indicate values of $\beta_k$.)
Fig. 7. True values of the sensitivities $\phi$. (Legends indicate DER indexes.)
Fig. 8. Estimated values of the sensitivities $\hat{\phi}$ with $\beta_k = 0.02$. (Legends indicate DER indexes.)
required to provide different amounts of active power at different time intervals. Specifically, assume the DERs are required to provide 100 kW, 300 kW, and 600 kW during time instants 0 to 200, 201 to 400, and 401 to 600, respectively. The tracking error for this case is shown in Fig. 11 and the MAE of estimation is shown in Fig. 12. As is shown in these two figures, the DERs can track the target quickly after the target changes, and in the meantime, the estimation error remains small, which indicates the good performance of the parameter estimators.

VI. CONCLUDING REMARKS

In this paper, we proposed a data-driven coordination framework for DERs in a lossy power distribution system with unknown model to collectively provide some pre-specified amount of active power to a bulk power system. We showed that using the distributed coordination algorithm proposed in the framework, the total provided active power converges to the required amount, and the estimated parameters converge to the true parameters a.s. The data-driven nature of this framework makes it adaptive to various system conditions. We validated the effectiveness of the proposed framework through numerical simulations on a modified version of the IEEE 123-bus test feeder.

While the proposed framework is capable to coordinate the DERs to provide the requested amount of active power accurately without knowing the system model, this may not be achieved in a way that is optimal in some sense, e.g., in a way that minimizes the total cost. We leave the extension of this framework to situations where optimal dispatch of the DERs are required for future work.

APPENDIX

A. Proof of Theorem 7

The convergence analysis of the control update rule relies on the following two lemmas.

Lemma 1. There exists \( \hat{\phi}[k] \) satisfying \( 0_n \leq \hat{\phi}[k] \leq \hat{\phi}[k] \), such that (A.3) is equivalent to

\[
\bar{u}[k] = \bar{u}[k-1] - \beta_k e[k] W[k] \hat{\phi}[k].
\]

Also, \( \hat{\phi}[k] = 0_n \) if and only if \( \bar{u}[k] = \bar{u} \) or \( \bar{u}[k] = \bar{u} \).

Furthermore, if \( \bar{u}[k] \neq \bar{u} \) and \( \bar{u}[k] \neq \bar{u} \), there exists \( i \in V := \{1, 2, \ldots, n\} \) such that \( \bar{\phi}_i[k] = \bar{\phi}_i[k] = [\bar{u}_i[k] \mid \bar{u}_i[k] \mid 0_n] \).

Proof. If \( \bar{u}[k-1] - \beta_k e[k] W[k] \hat{\phi}[k] \in \mathcal{U} \), then we simply set \( \hat{\phi}[k] = \hat{\phi}[k] \). Without loss of generality, first consider the case where the following holds for some \( i \in V \):

\[
u_i[k-1] - \beta_k e[k] w_i[k] \hat{\phi}_i[k] = \bar{u}_i.
\]

Then, \( e[k] < 0 \) and \( w_i[k] > 0 \) since otherwise (A.1) cannot not hold. Therefore,

\[
u_i[k] = P_{\mathcal{U}}(\nu_i[k-1] - \beta_k e[k] w_i[k] \hat{\phi}_i[k]) = \bar{u}_i.
\]

Let \( \bar{\phi}_i[k] = \frac{\nu_i[k-1] - \bar{u}_i}{\beta_k e[k] w_i[k]} \); by definition, \( \bar{\phi}_i[k] = 0 \) if and only if \( \nu_i[k-1] = \bar{u}_i \). Then, we have that:

\[
u_i[k] = \nu_i[k-1] - \beta_k e[k] w_i[k] \hat{\phi}_i[k].
\]
If follows then from (A.1), (A.2), and (A.3) that
\[ \beta_k e[k] \phi_i[k] w_i[k] < \beta_k e[k] \phi_i[k] w_i[k], \] (A.4)
which leads to \( 0 \leq \phi_i[k] < \hat{\phi}_i[k] \). A similar argument can be used for the case where \( u_i[k-1] - \beta_k e[k] w_i[k] \phi_i[k] < u_i \) and for some \( i \in V \).

If \( u[k] \neq \underline{u} \) and \( u[k] \neq \overline{u} \), then there exists \( i \in V \) such that \( u_i < u_i[k] < \overline{u}_i \), which implies
\[ u_i[k] = u_i[k-1] - \beta_k e[k] w_i[k] \phi_i[k]. \] (A.5)
Therefore, \( \phi_i[k] = \hat{\phi}_i[k] \). Consequently, \( \phi_i[k] = \hat{\phi}_i[k] \in [\underline{u}_i, \overline{u}_i] \). It can be easily seen that if \( U \) is sufficiently large and no DER hits its capacity limits, then \( \phi[k] = \hat{\phi}[k] \).

**Lemma 2.** Let \( X_k, k = 1, 2, \cdots \), be independently identically distributed (i.i.d.) random variables. Assume \( X_k \geq 0 \) and \( \mathbb{E}[X_k] \in (0, 1) \), where \( \mathbb{E} \) denotes expectation. Let \( Y_k = \prod_{i=1}^k X_i \). Then, \( \lim_{k \to \infty} Y_k = 0 \) a.s.

**Proof.** Note that \( Y_k = \exp \left\{ \sum_{i=1}^k \log X_i \right\} \). By the Strong Law of Large Numbers (see Proposition 2.15 in [20]), we have that
\[ \lim_{k \to \infty} \sum_{i=1}^k \frac{1}{k} \log X_i = \mathbb{E} \log X_i , \text{ a.s.} \] (A.6)
By Jensen’s inequality (see Theorem 2.18 in [20]), we have that
\[ \mathbb{E} \log X_i \leq \log \mathbb{E} [X_i] < 0. \] (A.7)
Therefore,
\[ \lim_{k \to \infty} \sum_{i=1}^k \frac{1}{k} \log X_i = -\infty, \text{ a.s.}, \] (A.8)
which leads to
\[ \lim_{k \to \infty} Y_k = \lim_{k \to \infty} \exp \left\{ \sum_{i=1}^k \frac{1}{k} \log X_i \right\} = \exp \left\{ \lim_{k \to \infty} \sum_{i=1}^k \frac{1}{k} \log X_i \right\} = 0, \text{ a.s.}; \] (A.9)
this completes the proof.

Using Lemma 1 and Lemma 2, we can prove the Theorem 1 as follows:

**Proof.** By (6), we have that
\[ e[k+1] - e[k] = \phi^\top[k] \Delta u[k]. \] (A.10)
By Lemma 1 we have that
\[ \Delta u[k] = -\beta_k e[k] W[k] \phi[k], \] (A.11)
where \( 0_n \leq \phi[k] \leq \hat{\phi}[k] \). Substituting (A.11) into (A.10) leads to
\[ e[k+1] = (1 - \beta_k \phi^\top[k] W[k] \phi[k]) e[k]. \] (A.12)
Define \( \rho_k = 1 - \beta_k \phi^\top[k] W[k] \phi[k] \), then
\[ e[k+1] = e[0] \prod_{i=0}^{k} \rho_i . \] (A.13)
By Assumption 1 \( 0 < \underline{b}_1 \leq \phi_i[k] \leq \overline{b}_1 \). In addition, it follows from Lemma 7 that \( 0 \leq \phi_i[k] \leq \hat{\phi}_i[k] \leq \overline{b}_1 \). Therefore, \( \phi^\top[k] W[k] \phi[k] \) can be bounded as follows:
\[ 0 \leq \phi^\top[k] W[k] \phi[k] = \sum_{i=1}^n w_i[k] \phi_i[k] \phi_i[k] \leq n \overline{b}_1^2 \] (A.14)
Since \( \beta_k < \frac{1}{n \overline{b}_1^2} \), then all \( e[k] \) has the same sign for all \( k \) (positive if \( e[0] > 0 \), and negative otherwise). As a result, the entries of \( \Delta u[k] \) always have the same sign by (A.11).
(a) If \( e[k] = 0 \) for some \( k \in \mathbb{N} \), then \( \Delta u[k] = 0_n \). The control and estimation algorithms will stop updating according to (11) and (15). In this case, \( u[k] \) may equal to \( \underline{u} \) or \( \overline{u} \) or neither, and the system attains an equilibrium.
(b) Now suppose \( e[k] \neq 0, \forall k \in \mathbb{N} \). Since the increments of \( u \) always have the same sign, the entries of \( u \) cannot hit their bounds in different directions, i.e., some hit their lower bounds while others hit their upper bounds.
(b.1) If \( u[k-1] = \underline{u} \) for some time instant \( k \), then \( e[k] > 0, \forall k \in \mathbb{N} \). By (13), we have that
\[ u[k] = \mathbb{P}_U (u - \beta_k e[k] W[k] \phi[k]) = u \] (A.15)
Thus, \( \Delta u[k] = 0_n \), which leads to \( e[k+1] = e[k] \) by (A.10). Therefore, \( u \) will equal to \( \underline{u} \) and \( e[k'] = e[k] > 0 \) for all \( k' > k \).
Similarly, when \( u[k] = \overline{u} \), \( u \) will be equal to \( \overline{u} \), and \( e \) will be equal to \( e[k] < 0 \) in all future time intervals. The system attains an equilibrium in both cases.
(b.2) If \( u[k] \neq \underline{u} \) and \( u[k] \neq \overline{u} \), \( \forall k \in \mathbb{N} \), by Lemma 1 there exists \( i \in V \) such that \( \phi_i[k] \in [\overline{b}_1, \overline{b}_1] \). Then,
\[ \phi^\top[k] W[k] \phi[k] = \sum_{i=1}^n \phi_i[k] \phi_i[k] w_i[k] \geq 2 \overline{b}_1^2 w_i[k]. \] (A.16)
Thus, by using (A.14) and (A.16), it follows that \( \rho_k \in [1 - \beta_k n \overline{b}_1^2, 1 - \beta_k b_1^2 w_i[k] \phi_i[k]]) \). Define \( \overline{p}_k = 1 - e w_i[k] \phi_i[k] \), then \( \overline{p}_k \) equals to \( 1 - \epsilon \) or \( 1 \), each with probability 0.5, and \( \mathbb{E} [\overline{p}_k] = 1 - \frac{\epsilon}{2} \in (0, 1) \). Note that \( 0 < \epsilon < \frac{\beta_k}{n \overline{b}_1^2} \) implies \( \overline{p}_k > 0 \). By Lemma 2
\[ \lim_{k \to \infty} \prod_{i=0}^{k} \overline{p}_i = 0, \text{ a.s.} \] (A.17)
When \( \beta_k \in \left( \frac{1}{2 \overline{b}_1^2}, \frac{1}{n \overline{b}_1^2} \right] \), \( 0 \leq \rho_k \leq \overline{p}_k \). Then, in an a.s. sense,
\[ \lim_{k \to \infty} |e[k+1]| = \lim_{k \to \infty} |e[0]| \lim_{k \to \infty} \prod_{i=0}^{k} \rho_i \leq \lim_{k \to \infty} |e[0]| \lim_{k \to \infty} \prod_{i=0}^{k} \overline{p}_i = 0. \] (A.18)
Since \( |e[k+1]| \geq 0 \), \( \lim_{k \to \infty} |e[k+1]| = 0 \) a.s. In addition, by (A.11), \( \lim_{k \to \infty} \Delta u[k] = 0_n \) a.s.

**Remark 1.** If \( U \) is sufficiently large and no DER hits the capacity limits, then \( \phi[k] = \hat{\phi}[k] \) and \( \phi^\top[k] W[k] \phi[k] \geq \frac{\beta_k^2}{n \overline{b}_1^2} \sum_{i=1}^n w_i[k] \). Following a similar argument as in part (b.2)
in the proof of Theorem 1 we can show $e[k]$ converges to 0 a.s. when $\beta_k \in (\frac{\lambda_n}{2}, \frac{1}{2\lambda_n})$, where $0 < \epsilon < \frac{\lambda_n}{2}$.  

Following a similar argument, Corollary 1 can be proved as follows:

**Proof.** When the control update rule in (12) is used instead of the one in (13),

$$e[k + 1] = (1 - \beta_k \phi^T[k] \phi[k])e[k].$$  

(19)

If $u[k] \neq u$ and $u[k] \neq \mathbf{w}$, $\phi^T[k] \phi[k] = \phi_i[k] \phi_i[k] \geq \lambda_i^2$. Define $\rho_k = 1 - \beta_k \phi^T[k] \phi[k]$. When $\beta_k \in \left(\frac{\lambda_n}{2}, \frac{1}{2\lambda_n}\right)$, $\rho_k < 1 - \epsilon$. Therefore, $\lim_{k \to \infty} \frac{\rho_k}{\rho_k} = \rho_k < 1 - \epsilon$. □

**B. Proof of Theorem 2**

The convergence analysis of the estimation update rule uses some convergence results for $\Delta \phi[k]$, which are presented next.

**Lemma 3.** Let $X_k$, $k = 1, 2, \cdots$, be i.i.d. random variables that take value 1 with probability 0.5, or some constant $x \in (0, 1)$, also with probability 0.5. Let $Y_k = \prod_{i=1}^k X_i$ and $Z = \sum_{i=1}^k Y_i$. Then, $Z$ is bounded a.s.

**Proof.** Let $M$ denote the maximum number of 1’s that appears continuously in the sequence $\{X_k\}$; then, the sequence $\{Y_k\}$ will have a new (smaller) value at most after $M + 1$ steps. We claim $Z$ is unbounded only if $M$ is infinite. Suppose $X_j = x$, and $X_k = 1$ for $k = j + 1, \cdots, j + m$, then $Y_j = Y_{j+1} = \cdots = Y_{j+m}$ and $\sum_{i=j}^{j+m} Y_j = (m+1)Y_j \leq (M+1)Y_j$. Therefore, $Z = \sum_{i=1}^k Y_i \leq (M+1) \sum_{i=0}^\infty x^i = \frac{M+1}{1-x}$. (20)

It follows from (A.20) that $Z$ is unbounded only if $M$ is infinite. However, $\mathbb{P}\{M = \infty\} = \mathbb{P}\{X_{i+1} = \cdots = X_{i+M} = 1, \text{ for some } i\} = \frac{1}{2^\infty} = 0$. Thus, $Z$ is bounded a.s. □

**Lemma 4.** Using estimation update rule (11) and control update rule (13), with $\beta_k \in (\frac{\lambda_n}{2}, \frac{1}{2\lambda_n})$, where $0 < \epsilon < \frac{\lambda_n}{2}$ is a given parameter, then

$$\lim_{k \to \infty} \|\Delta \phi[k]\| = 0, \text{ a.s.}$$  

(21)

and

$$\sum_{k=1}^\infty \|\Delta \phi[k]\| < \infty, \text{ a.s.}$$  

(22)

**Proof.** If follows from the proof of Theorem 1 that the entries of $\Delta u[k]$ always have the same sign. First consider the case where $\Delta u[k] \geq 0_n$ for all $k \in \mathbb{N}$. Note that $\phi^T[k] = \frac{\partial f}{\partial u} \hat{u}[k]$, where $\hat{u}[k] = a_k u[k - 1] + (1 - a_k) u[k - 2]$ with $a_k \in [0, 1]$, i.e., $u[k - 2] \leq \hat{u}[k] \leq u[k - 1]$. Similarly, $\phi^T[k - 1] = \frac{\partial f}{\partial u} \hat{u}[k-1]$, where $u[k - 3] \leq \hat{u}[k - 1] \leq u[k - 2]$. Thus, by Assumption 1 we have that

$$\|\Delta \phi[k]\| \leq b_2 \|\Delta u[k - 1]\|$$  

$$\leq b_2 \|u[k - 1] - u[k - 2]\|$$  

$$= b_2 \|\Delta u[k - 1] + \Delta u[k - 2]\|$$  

$$\leq b_2 \|\Delta u[k - 1]\| + \|\Delta u[k - 2]\|.$$  

(A.23)

Since $\lim_{k \to \infty} \|\Delta u[k]\| = 0$ a.s. by Theorem 1 as a result, $\lim_{k \to \infty} (\|\Delta u[k - 1]\| + \|\Delta u[k - 2]\|) = 0$ a.s., which gives

$$\lim_{k \to \infty} \|\Delta \phi[k]\| = 0, \text{ a.s.}$$  

(A.24)

Assume $u[k] = 0_n$ for all $k < 0$, then we have that

$$\sum_{k=1}^\infty \|\Delta \phi[k]\| \leq \sum_{k=1}^\infty b_2 (\|\Delta u[k - 1]\| + \|\Delta u[k - 2]\|)$$  

$$= 2b_2 \sum_{k=1}^\infty \|\Delta u[k]\|$$  

$$\leq 2b_2 \sum_{k=1}^\infty \beta_k W[k] \phi[k] e[k]$$  

$$\leq \frac{2b_2}{\sqrt{n_0}} \beta_k \sum_{k=0}^\infty \|e[k]\|$$  

$$= \frac{2b_2}{\sqrt{n_0}} \sum_{k=0}^\infty \|e[k]\|. $$  

(A.25)

Recall that $\bar{\phi}_k$ equals to $1 - \epsilon$ or 1, each with probability 0.5, where $\bar{\phi}_k$ is defined in the proof of Theorem 1. Therefore, by Lemma 3, $\sum_{k=1}^\infty \|\Delta \phi[k]\|$ is bounded a.s. When $\beta_k \in \left(\frac{\lambda_n}{2}, \frac{1}{2\lambda_n}\right)$, $0 \leq \rho_k \leq \rho_k$,

$$\sum_{k=0}^\infty \|e[k]\| = \sum_{k=1}^\infty \|e[0]\| (1 + \prod_{i=1}^{k-1} \rho_i) \leq \|e[0]\| \sum_{k=1}^\infty (1 + \prod_{i=0}^{k-1} \bar{\phi}_k).$$  

(A.26)

As a result, $\sum_{k=0}^\infty \|e[k]\|$ is bounded a.s. The case where $\Delta u[k] \leq 0_n$ for all $k \in \mathbb{N}$ can be proved similarly. □

The convergence analysis of the estimation update rule also relies on the following lemma (see Theorem 1 in [27]).

**Lemma 5.** Let $X_k, Y_k, Z_k, k = 1, 2, \cdots$, be non-negative variables in $\mathbb{R}$ such that $\sum_{k=0}^\infty Y_k < \infty$, and $X_{k+1} \leq X_k + Y_k - Z_k$, then $X_k$ converges and $\sum_{k=0}^\infty Z_k < \infty$.

Using Lemma 4 and Lemma 5, Theorem 2 can then be proved as follows:

**Proof.** Consider an arbitrary sample path. Without loss of generality, assume $e[k] < 0$, it follows from Theorem 1 that $e[k] < 0, \forall k \in \mathbb{N}$. Since $u[k] \in (\mathbf{u}, \mathbf{w})$, $\forall k \in \mathbb{N}$, (13) becomes

$$\Delta u[k] = -\beta_k e[k] W[k] \phi[k].$$  

(27)

It follows from (6) and (11) that

$$\hat{\phi}[k + 1] = \mathbb{P}_Q(\hat{\phi}[k] - \alpha_{k+1} \Delta u[k] \Delta u^T[k] e[k]).$$  

(28)

By definition, the estimation error at time instant $k$ is

$$e[k + 1] = \mathbb{P}_Q(\hat{\phi}[k] - \alpha_{k+1} \Delta u[k] \Delta u^T[k] e[k] - \phi[k + 1]).$$  

(29)

Since $\phi[k + 1] = \mathbb{P}_Q(\phi[k + 1])$, by the non-expansiveness of the projection operation (see Proposition 1.1.9 in [28]), then

$$\|e[k + 1]\| \leq \|e[k] - \alpha_{k+1} \Delta u[k] \Delta u^T[k] e[k] - \Delta \phi[k + 1]\|$$  

$$\leq \|e[k] - \alpha_{k+1} \Delta u[k] \Delta u^T[k] e[k] - \Delta \phi[k + 1]\| + \|\Delta \phi[k + 1]\|.$$  

(A.30)

Let $f(\alpha_{k+1}) = \|e[k] - \alpha_{k+1} \Delta u[k] \Delta u^T[k] e[k]\|^2$; then, $f$ attains its minimum at $\alpha_{k+1} = \frac{\|\Delta u[k]\|^2 e[k]}{\|\Delta u[k]\|^2}$, which is

$$\|e[k]\|^2 - (\varepsilon T[k] \Delta u[k])^2 = \|e[k]\|^2 - (\varepsilon T[k] W[k] \phi[k])^2.$$  

(A.31)

Define $\cos \theta_k = \frac{\varepsilon T[k] W[k] \phi[k]}{\|e[k]\|^2}$. Consequently, $f(\alpha_{k+1}) = (1 - \sin^2 \theta_k) \|e[k]\|^2$, and

$$\|e[k + 1]\| \leq \|e[k]\| + \|\Delta \phi[k + 1]\|.$$  

(A.32)
Let $X_k = [\|e[k]\|, Y_k = [\Delta \phi[k+1]|, Z_k = (1 - |\sin \theta[k]|)\|e[k]\|]$. Then, $X_{k+1} = X_k + Y_k - Z_k$. Also, $\sum_{k=0}^{\infty} Y_k = \sum_{k=1}^{\infty} \|\Delta \phi[k]\| < \infty$ by Lemma 4. Therefore, by Lemma 5 $\|e[k]\|$ converges, and $\sum_{k=1}^{\infty} (1 - |\sin \theta[k]|)\|e[k]\|$ < \infty, which further implies $\lim_{k \to \infty} (1 - |\sin \theta[k]|)\|e[k]\| = 0$.

Let $\varepsilon^*$ denote the limit of $\|e[k]\|$: then,

$$\lim_{k \to \infty} |\sin \theta[k]|\|e[k]\| = \lim_{k \to \infty} (|\sin \theta[k] - 1|)\|e[k]\| + \lim_{k \to \infty} \|e[k]\| = \varepsilon^*. \quad (A.33)$$

Next, we show $\varepsilon^* = 0$ by contradiction. Assume $\varepsilon^* > 0$. Since both $\|e[k]\|$ and $|\sin \theta[k]|\|e[k]\|$ converges to $\varepsilon^*$,

$$\lim_{k \to \infty} |\sin \theta[k]|\|e[k]\| = \lim_{k \to \infty} (|\sin \theta[k] - 1|)\|e[k]\| + \lim_{k \to \infty} \|e[k]\| = 1, \quad (A.34)$$

which implies $|\cos \theta[k]|$ converges to 0. Since $\|e[k]\|$ and $\|W[k]\|\|\phi[k]\|$ are bounded, then $|\varepsilon[k]W[k]\|\|\phi[k]\|$ converges to 0. Define $E_k = \{w_j[k] = 1 \text{ if } j = i, w_j[k] = 0 \text{ otherwise}\}$; then $P\{E_k\} = \frac{1}{2}$. Consequently, $\sum_{k=1}^{\infty} P\{E_k\} = \infty$. Also note that $E_k, k \in \mathbb{N}$, are independent.

By the Borel-Cantelli Lemma (see Lemma 1.3 in [26]), $P\{E_k \text{ infinitely often}\} = 1$; therefore, there are infinitely many time instances that $w_j[k] = 1$ and $w_j[k] = 0$ for all $j \neq i$. Let $K_i$ denote the set of such time instances. Then $\varepsilon[k]W[k]\|\phi[k]\| = \varepsilon[k]\|\phi_i[k]\|$, for $k \in K_i$. The sequence $\{\varepsilon[k]\|\phi_i[k]\|, k \in K_i\}$ is a subsequence of $\{\varepsilon[k]W[k]\|\phi[k]\|, k \in K_i\}$; therefore, it also converges to 0. Note that $\|\phi_i[k]\| > 0$; thus, $\varepsilon[k]$ converges to 0. Since $k$ is arbitrary, we conclude that $\|e[k]\|$ converges to 0, which implies $\varepsilon^* = 0$, contradiction. Since this result holds for all sample paths, then we conclude that $\|e[k]\|$ converges to 0 a.s.

REFERENCES

[1] J. Driesen and F. Katiraei, “Design for distributed energy resources,” IEEE Power Energy Mag., vol. 6, no. 3, pp. 30–40, May 2008.

[2] H. Xu, A. D. Domínguez-García, and P. W. Sauer, “A data-driven voltage control framework for power distribution systems,” in Proc. of IEEE PES General Meeting, Portland, OR, Aug. 2018, pp. 1–5.

[3] A. Kulmala, S. Repo, and P. Jarventausa, “Coordinated voltage control in distribution networks including several distributed energy resources,” IEEE Trans. Smart Grid, vol. 5, no. 4, pp. 2010–2020, July 2014.

[4] H. Li, F. Li, Y. Xu, D. T. Rizy, and J. D. Kueck, “Adaptive voltage control with distributed energy resources: Algorithm, theoretical analysis, simulation, and field test verification,” IEEE Trans. Power Syst., vol. 25, no. 3, pp. 1638–1647, Aug. 2010.

[5] E. Dall’Anese, 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, March 2018.

[6] H. Xu, A. D. Domínguez-García, and P. W. Sauer, “Adaptive coordination of distributed energy resources in lossy power distribution systems,” in Proc. of IEEE PES General Meeting, Portland, OR, Aug. 2018, pp. 1–5.

[7] M. A. A. Pedrasa, T. D. Spooner, and I. F. MacGill, “Coordinated scheduling of residential distributed energy resources to optimize smart home energy services,” IEEE Trans. Smart Grid, vol. 1, no. 2, pp. 134–143, Sept. 2010.

[8] Z. Hou and S. Jin, “Data-driven model-free adaptive control for a class of mimo nonlinear discrete-time systems,” IEEE Trans. Neural Netw., vol. 22, no. 12, pp. 2173–2188, Dec. 2011.

[9] J. Zhang, X. Zheng, Z. Wang, L. Guan, and C. Y. Chung, “Power system sensitivity identification – inherent system properties and data quality,” IEEE Trans. Power Syst., vol. 32, no. 4, pp. 2756–2766, July 2017.

[10] Y. C. Chen, A. D. Domínguez-García, and P. W. Sauer, “Measurement-based estimation of linear sensitivity distribution factors and applications,” IEEE Trans. Power Syst., vol. 29, no. 3, pp. 1372–1382, May 2014.

[11] Y. C. Chen, A. D. Domínguez-García, and P. W. Sauer, “A sparse representation approach to online estimation of power system distribution factors,” IEEE Transactions on Power Systems, vol. 30, no. 4, pp. 1727–1738, July 2015.

[12] Y. C. Chen, J. Wang, A. D. Domínguez-García, and P. W. Sauer, “Measurement-based estimation of the power flow jacobian matrix,” IEEE Trans. Smart Grid, vol. 7, no. 5, pp. 2507–2519, Sept. 2016.

[13] K. E. V. Horn, A. D. Domínguez-García, and P. W. Sauer, “Measurement-based real-time security-constrained economic dispatch,” IEEE Trans. Power Syst., vol. 31, no. 5, pp. 3548–3560, Sept. 2016.

[14] C. Lu, Y. Zhao, K. Men, L. Tu, and Y. Han, “Wide-area power system stabiliser based on model-free adaptive control,” IET Control Theory Appl., vol. 9, no. 13, pp. 1996–2007, 2015.

[15] J. Zhang, C. Y. Chung, and Y. Han, “Online damping ratio prediction using locally weighted linear regression,” IEEE Trans. Power Syst., vol. 31, no. 3, pp. 1954–1962, May 2016.

[16] Z.-S. Hou and Z. Wang, “From model-based control to data-driven control: Survey, classification and perspective,” Information Sciences, vol. 235, pp. 3–35, June 2013.

[17] Z. Hou, R. Chi, and H. Gao, “An overview of dynamic-linearization-based data-driven control and applications,” IEEE Trans. Ind. Electron., vol. 64, no. 4, pp. 4076–4090, May 2017.

[18] J. Zhang, Z. Wang, X. Zheng, L. Guan, and C. Y. Chung, “Locally weighted ridge regression for power system online sensitivity identification considering data collinearity,” IEEE Trans. Power Syst., vol. PP, no. 99, pp. 1–17, 2017.

[19] J. Zhang, C. Y. Chung, and L. Guan, “Noise effect and noise-assisted ensemble regression in power system online sensitivity identification,” IEEE Trans. Ind. Informat., vol. 13, no. 5, pp. 2302–2310, Oct. 2017.

[20] A. D. Domínguez-García and C. N. Hadjicostis, “Coordination and control of distributed energy resources for provision of ancillary services,” in Proc. of IEEE International Conference on Smart Grid Communications, Oct. 2010, pp. 537–542.

[21] H. Xu, S. C. Utomi, A. D. Domínguez-García, and P. W. Sauer, “Coordination of distributed energy resources in lossy networks for providing frequency regulation,” in Proc. of X Bulk Power Systems Dynamics and Control Symposium, Aug. 2017, pp. 1–7.

[22] A. D. Domínguez-García, C. N. Hadjicostis, and N. H. Vaidya, “Reilient networked control of distributed energy resources,” IEEE Journal on Selected Areas in Communications, vol. 30, no. 6, pp. 1137–1148, July 2012.

[23] S. T. Cady, A. D. Domínguez-García, and C. N. Hadjicostis, “Finite-time approximate consensus and its application to distributed frequency regulation in islanded ac microgrids,” in Proc. of Hawaii International Conference on System Sciences, Jan. 2015, pp. 2664–2670.

[24] IEEE distribution test feeders. [Online]. Available: https://ewh.ieee.org/soc/pes/dsom/testfeeders/.

[25] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, “Matpower: Steady-state operations, planning, and analysis tools for power systems research and education,” IEEE Trans. Power Syst., vol. 26, no. 1, pp. 12–19, Feb. 2011.

[26] B. Hajek, Random processes for engineers. Cambridge University Press, 2015.

[27] H. Robbins and D. Siegmund, “A convergence theorem for non negative almost supermartingales and some applications,” in Herbert Robbins Selected Papers. Springer, 1985, pp. 111–135.

[28] D. P. Bertsekas, Convex optimization theory. Athena Scientific Belmont, 2009.