Learning-based Distributed Optimal Power Sharing and Frequency Control under Cyber Contingencies

Siyu Xie[b], Masoud H. Nazari[a,*], Le Yi Wang[a]

[a]*Department of Electrical and Computer Engineering, Wayne State University, Detroit, MI 48202, USA*
[b]*School of Aeronautics and Astronautics, University of Electronic Science and Technology of China, Chengdu 611731, China*

Abstract

This paper proposes a learning-based distributed optimization framework to enhance the resilience of distributed power system algorithms against communication failures. The autoregressive–moving-average (ARMA) model is used to estimate the missing states and control actions of neighboring agents during communication contingencies. The proposed framework allows agents to predict the future control actions of neighbors. Thus, even during complete loss of communication, agents can efficiently perform distributed optimization. We use the Distributed Optimal Frequency Control (DOFC) algorithm, which includes optimal power sharing to achieve frequency stability, as a benchmark application platform to show the effectiveness of the proposed framework. The theoretical findings are evaluated on two practical power systems. The results show that the ARMA-based DOFC algorithm can asymptotically reach the same convergence rate as the power system without communication interruptions.

Keywords: ARMA model, resilient distributed optimal frequency control, multi-agent network, prediction-based distributed optimization

1. Introduction

Power grids are moving toward a distributed architecture for decision making and control [1, 2]. Several factors are contributing to this move such as vulnerability of centralized power grid to single point of failure, insufficient flexibility

*Corresponding author

of centralized architecture for large penetration of distributed energy resources (DERs), etc. [3].

Distributed power system algorithms have been proposed as a scalable solution for real-time coordination of emerging DERs and producer-consumer (prosumers) agents [4, 5, 6, 7, 8, 9]. Reference [10] presents a comprehensive survey of distributed optimization and control algorithms for electric power systems until late 2017. The communication network performance and reliability are key elements in the overall success of distributed algorithms [11, 12]. Failures in the communication network, caused by disconnection of communication links, cyber-attacks, delay and other channel imperfections, can jeopardize the performance of distributed algorithms and lead to system-level efficiency, reliability and security problems [13]. In other words, distributed algorithms require convergence in the cyber network before the solutions can be implemented on the physical grid and the intermittent iterations are not satisfying power flow and other system constraints [14, 15, 16]. On the other hand, fully decentralized methods provide sub-optimality and can lead to inter-area oscillations among agents and sub-systems [17, 18, 19, 20, 21]

Most of the recent efforts to improve the resilience of distributed algorithms against communication failures are limited to architecture solutions [22, 23, 24]. For instance, in our previous work [25] we introduced multiple-input multiple-output (MIMO) architecture to increase the reliability of communication networks and improve the convergence rate of distributed algorithms against communication delays. Also, in [26, 27] we introduced an enhanced communication network architecture, which allows for multiple-hop communication among agents to increase the resilience of distributed algorithms. The proposed solutions require hardware improvement in the cyber layer, which can be a costly approach particularly for large-scale power systems. On the other hand, the state-of-the-art resilient distributed algorithms do not provide a systematic approach to guarantee the convergence of distributed algorithms under communication failures [28, 29, 30].

This paper addresses the gap by introducing a learning-based approach based on the autoregressive–moving-average (ARMA) model to estimate the missing data and bypass convergence in the cyber network during communication failures. The ARMA model has less computational efforts compared with deep learning methods. Thus, it can be a computationally effective approach for linearized power system models in real-time operation. The ARMA model estimates the states and control actions of neighboring agents, when there is loss of communication. Also, it can predict the future control action of neighbors, thus it can expedite the convergence of distributed algorithms. We use Distributed Optimal

Frequency Control (DOFC) algorithms based on the linearized model of the power system to showcase the effectiveness of the ARMA-based model. We show that the ARMA-based DOFC can asymptotically reach the same convergence rate as the power system without communication interruptions.

The main contributions of this paper are:

1. Developing the distributed ARMA model to estimate the states and control actions of neighboring agents during peer-to-peer communication interruptions.

2. Embedding the ARMA model into the DOFC algorithm, and laying a foundation for learning-based DOFC to mitigate the impact of communication failures.

3. Quantitatively characterizing the fundamental relationship between the estimation errors and convergence rates of the ARMA-based DOFC to develop a practical criterion for securing reliability of optimal frequency control under communication uncertainties.

4. Illustrating that the convergence rate of the ARMA-based DOFC under communication failures can asymptotically reach the Cramer–Rao (CR) lower bound. Note that the CR bound represents the lower bound on the variance of error between the optimal control actions and estimated control solutions of the algorithm [31].

The rest of the paper is organized as follows. Section 2 presents an overview of realistic communication performance and its impact on distributed power system algorithms. Section 3 proposes a distributed prediction framework based on the ARMA model to mitigate the impact of disrupted communication channels on distributed algorithms such as DOFC. The main results are established in this section, where error bounds, strong convergence, and asymptotic optimality errors are derived. The technical findings are illustrated on two realistic power systems in Sections 4 to validate the performance of the ARMA-based DOFC under different scenarios. The paper concludes with discussions of overall findings in Section 5.

2. Realistic Communication Performance

Algorithmic convergence in a distributed setting depends on the characteristics of the control network, such as its topology, but also on the quality of the communication network that serves peer-to-peer interactions, as well as on the nature of

the distributed algorithm logic. The major modes of communication failures include: Delayed, Disrupted, and Malicious Agents. Various use cases can manifest as one or more of the three fault conditions. In this paper, we develop algorithmic solutions for Delayed and Disrupted use cases.

2.1. Communication Delays

Timing is critical for supporting the functionalities of distributed algorithms [32, 33]. IEEE and the International Electrotechnical Commission (IEC) have defined rigorous standards for communication delay requirements in smart grids [34]. However, in reality the communication networks are not always able to meet the strict communication delay requirements of IEEE and IEC. For instance, experimental results on communication delays between substations, reported in [34, 33], show that in many scenarios the packet delays exceed the maximum required limit for the most critical messages.

DOFC is one of the time critical power system algorithms. NERC A1 criterion requires that frequency regulators bring area control error (ACE) to zero once every 10 minutes [35] and NERC B2 criterion requires that frequency regulators begin to return ACE to zero within 1 minute after a disturbance [36]. Prolonged communication delays slow down the distributed algorithms (iterations take more time) and increase the risk of violating the NERC reliability criteria.

2.2. Disrupted Communication Network

When all communication links connecting a sub-system (e.g. DERs, prosumers, etc) to the rest of the grid are disconnected, information isolation, denoted as "muteness", occurs. During this abnormal condition, other sub-systems need to take an algorithmic contingency-based action. Due to physical coupling, the control actions of each sub-system depend on the action of its neighbors. An isolated sub-system cannot control power flow and reroute power according to Kirchhoff's and Ohm's laws [37]. This lack of controllability can lead to system-level stability problems.

In [38, 39], we developed topological conditions under which the power system can maintain controllability under single or multiple communication failures. But the proposed solution requires communication architecture upgrade. Figure 1 illustrates the schematics of an $N-1$ communication-failure-resilient architecture for a loop power grid with eight prosumers [38].

As discussed in Section 1, this is a costly approach and may not be scalable for large-scale power grids. In this paper, we propose an algorithmic solution

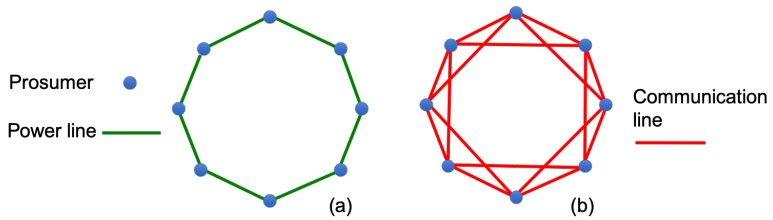

Figure 1: Architecture of a communication network resilient to a single communication failure. (a) Loop power grid with 8 prosumers (b) Communication network with two-hop information exchange capability [26]. This architecture may not be scalable for large-scale power systems.

to mitigate the impact of communication delays and disrupted communication channels without needs for upgrading the cyber infrastructure.

2.3. Modeling Communication Uncertainty

The basic assumption behind most distributed algorithms is that the communication network should follow the power grid topology. This is a strong assumption and may not be practical. In this paper, we limit the communication interaction between sub-systems (prosumers). Thus, as long as the prosumer-based communication network is connected, prosumers can perform distributed optimization. During the algorithm implementation, packet loss or channel interruption may cause the cyber link to be randomly disconnected. The connectivity of each communication line is represented by a stochastic variable

$$\lambda_{ij}(k) = \begin{cases} 1, & \text{if the link } (i,j) \text{ is connected at step } k, \\ 0, & \text{if the link } (i,j) \text{ is broken at step } k, \end{cases}$$

where λ_{ij} is an indicator function. If the communication link between prosumer i and prosumer j is connected, $\lambda_{ij} = 1$; otherwise $\lambda_{ij} = 0$. The mathematical expression is provided above.

We assume that each channel's packet loss is an independent and identically distributed (i.i.d) sequence of random variables with probability Pr_k to be linked and $1 - Pr_k$ to be disconnected. This assumption is practical, because in reality most communication nodes use battery-based power supplies, providing some assurance that power outages will not result in communication outages [37]. Thus, communication nodes can independently operate and their failure rates are mutually independent. Therefore, when block erasure channels are mutually independent, $\lambda_{ij}(k)$ is an independent random variable.

3. Distributed Prediction Framework

Without loss of generality, suppose that prosumer j must receive the control action of its adjacent prosumer i (U_{ij}). Since information exchange among prosumers is subject to data loss and channel interruption $\lambda_{ij}(k)$, it is possible that at time k, $U_{ij}(k)$ is not received. How should prosumers find a remedy for this situation?

1. Solution 1: Ignore the link between prosumers i and j. For example, suppose prosumer 1 has two neighbors, prosumers 2 and 3. If the data from prosumer 2 is lost at k, only data from prosumer 3 will be used in updating the optimization iteration variable.

2. Solution 2: Use the previously received value until the next value is received. This is a zero-order-hold method. For instance, if $U_{ij}(k)$ is not available, then prosumers use:

$$\widehat{U}_{ij}(k) = U_{ij}(k-1)$$

3. Solution 3: Estimate $U_{ij}(k)$ by using previously received or estimated values. This is the focus of this paper.

Note that Solutions 1 and 2 do not require any additional models. But in these solutions, either the information on $U_{ij}(k)$ is totally lost (Solution 1) or the estimation may be subject to large errors (Solution 2). This can impact the convergence of distributed algorithms and lead to sub-optimal power sharing, or in the worst-case scenario can cause system-wide instability. The central hypothesis of this paper is that using suitable models to predict $U_{ij}(k)$ can reduce the error

$$\widehat{U}_{ij}(k) - U_{ij}(k).$$

We will show that if the prediction error is small, Solution 3 can asymptotically converge to the optimal solution under communication failures. To prove this hypothesis, we first use the ARMA model to relate the dynamics of each prosumer to the past values. This allows predicting the current value of $U_{ij}(k)$ based on the historical data. Next, we develop the prediction algorithm and establish the convergence rate of the algorithm.

Solution 3 will be implemented on Delayed and Disrupted use cases. When delay at each iteration exceeds the threshold, the missing data is considered as lost

data. In practice, 30 ms is the threshold for communication delays. Prosumers can estimate the lost data using the ARMA model.

3.1. Dynamic Model Estimation Methods

There are different types of linear dynamic model estimation methods. We first use simple examples, and then extend them to general expressions.

1. AR (Auto-regression) Model. This model involves $U_{ij}(k)$ (self recursive expression) without input. Thus, the current control action of the neighboring agents is correlated with their historical control values. For this model, we can define

$$\widehat{U}_{ij}(k) = \begin{cases} U_{ij}(k) & U_{ij}(k) \text{ received} \\ \sum_{l=1}^{q} \phi_l \hat{U}_{ij}(k-l) & \text{Otherwise,} \end{cases} \tag{1}$$

   where q is the time horizon of the historical data (e.g. q steps in the past), and ϕ_l is the coefficient of the AR model at time step $k-l$. If there is communication noise, the AR model takes the form:

$$\widehat{U}_{ij}(k) = \sum_{l=1}^{q} \phi_l \hat{U}_{ij}(k-l) + d(k), \tag{2}$$

   where $d(k)$ is the noise at time k.

2. FIR (Finite Impulse Response) Model. The FIR model involves only a finite number of the past input values

$$\widehat{U}_{ij}(k) = \sum_{l=1}^{q} \theta_l p_j(k-l). \tag{3}$$

   In other words, the control action of prosumers U_{ij} is only correlated with their historical power deviations p_j, and θ_l is the coefficient of the FIR model at time step $k-l$. If there is noise, this model becomes

$$\widehat{U}_{ij}(k) = \sum_{l=1}^{q} \theta_l p_j(k-l) + d(k).$$

3. ARMA Model with noise. This model uses both self recursive expression and past input values to estimate U_{ij}, that is

$$\widehat{U}_{ij}(k) = \sum_{l=1}^{q} \phi_l \hat{U}_{ij}(k-l) + \sum_{l=1}^{q} \theta_l p_j(k-l) + d(k), \tag{4}$$

where ϕ and θ are matrix-valued coefficients of the ARMA model. The control actions of the neighboring agents are correlated with both their historical control values and their power deviations. ϕ and θ can be obtained by the least mean squares regression, which will be discussed in Section 3.3.

The ARMA model has the advantages of both AR and FIR models. It has less computational efforts compared with deep learning methods, such as long short-term memory (LSTM). Therefore, in this paper it is used for model estimation and prediction. In the distributed framework, prosumer i has sensing of its tie-line power flows and can measure p_j. Also, under normal conditions prosumer i receives the control actions of its neighbors (U_{ij}). The ARMA model allows prosumer i to estimate U_{ij}, when the communication link between prosumers i and j is disconnected.

3.2. Prediction-based Distributed Optimal Frequency Control

In this section, we discuss how the ARMA model can be integrated with the DOFC algorithm to mitigate the impact of communication failures. DOFC involves bringing the quasi-steady state frequency deviations to zero with optimal power sharing among prosumers. The optimization aspect of the DOFC makes it different from conventional decentralized frequency control methods, such as a PI controller. Optimal power sharing will ensure that the frequency control is done in an efficient way.

As frequency deviations has linear relation with power deviations at the prosumer level [40], we can define the following performance index for the DOFC algorithm for systems with n prosumers:

$$\min_{u} \mathcal{J}(u) = \min_{u} \frac{1}{2}(p(k+1)^\top Q p(k+1) + u(k)^\top R u(k)), \tag{5a}$$

$$\text{s.t. } p(k+1) - Ap(k) - Bu(k) = 0, \tag{5b}$$

where $p = [p_1, \cdots, p_n]^\top$ is the vector of power deviations and $u = [u_1, \cdots, u_n]^\top$ is the vector of control variables. Matrices $Q = \mathrm{diag}(Q_1, Q_2, \cdots, Q_n)$ and $R =$

$\mathrm{diag}(R_1, R_2, \cdots, R_n)$ are positive definite diagonal matrices, where Q_i represents the cost of power deviation and R_i represents the cost of frequency control for prosumer i. Also, (5b) represents the linearized quasi-steady state dynamics of the power system, where A is the system matrix and B is the control matrix. Note that matrices A and B preserve the topology of the prosumer-based power grid. The A and B matrices can be obtained as

$$A = I - T_s J_{p\delta} \mathcal{S},$$

$$B = T_s J_{p\delta} \Gamma.$$

where $\mathcal{S}$ is a diagonal matrix including the droop constant of prosumers (σ_i) and Γ is a diagonal matrix where γ_i represents the internal dynamics of prosumer i. Also, $J_{p\delta}$ is the reduced Jacobean matrix and T_s is the sampling time for frequency regulation [40].

Under normal condition, several distributed optimization methods can be used to solve the DOFC problem [41, 42]. In this paper, we use the alternating direction method of multipliers (ADMM). Thus, (5) can be decomposed as follows [27, 40]:

$$\begin{aligned} \min_{U_{1\cdot}, \cdots, U_{n\cdot}} \quad & \sum_{i=1}^{n} \big(Q_i \left[A_{i\cdot}^{\top} P^i(k) + B_{i\cdot}^{\top} U_{i\cdot}(k) \right]^2 + R_i U_{ii}^2(k) \big), \\ \text{s.t.} \quad & U_{ij}(k) = U_{jj}(k), \ \forall i \in N, j \in \mathcal{N}_i, \end{aligned} \tag{6}$$

where $A_{i\cdot}$ and $B_{i\cdot}$ are the ith rows of A and B respectively, and $\mathcal{N}_i$ is the set of prosumer i's neighbors. Also, $U_{i\cdot}$ is the ith rows of the matrix of decision variables $U = [U_{ij}]$, which include the control action of prosumers $U_{ii} = u_i$ and the perception of prosumers from the control actions of their neighbors U_{ij} $\forall j \in \mathcal{N}_i$. In addition, P^i is a column vector, which includes p_j, $j \in \mathcal{N}_i \cup \{i\}$. Constraints in (6) guarantee that prosumer i has a correct perception of the control strategies of its neighbors. At each iteration h, the average control strategy of prosumer j is computed using (7) and this information is shard with neighboring prosumers to ensure that the control actions and perceptions of the control actions converge to the same values:

$$\bar{U}_{jj}^h(k) := \frac{\sum\limits_{i \in \mathcal{N}_j \cup \{j\}} {U^h}_{ij}(k)}{|\mathcal{N}_j| + 1}. \tag{7}$$

To ensure that the solution is optimal, prosumer i solves the following aug-

mented Lagrangian function at each iteration:

$$\begin{aligned}\mathcal{L}_{\rho,i}(U_{i\cdot}, \bar{U}^h_{i\cdot}, \lambda^h_{i\cdot}) = Q_i \left[A_{i\cdot}^\top P^i + B_{i\cdot}^\top U_{i\cdot}\right]^2 + R_i U_{ii}^2 \\ + \lambda_{i\cdot}^{h^\top}(U_{i\cdot} - \bar{U}^h_{i\cdot}) + \frac{\rho}{2}\|U_{i\cdot} - \bar{U}^h_{i\cdot}\|_2^2,\end{aligned} \tag{8}$$

where $\rho > 0$ is a given penalty factor, and $\bar{U}^h_{i\cdot}$ is a column vector, which includes the average control strategy of prosumer i and that of its neighbors.

Prosumer i updates the primal and dual residuals $\alpha^h_{i\cdot}$ and $\beta^h_{i\cdot}$ as follows:

$$\alpha^h_{i\cdot} = U^h_{i\cdot} - \bar{U}^h_{i\cdot}, \text{ and } \beta^h_{i\cdot} = \rho(\bar{U}^h_{i\cdot} - \bar{U}^{h-1}_{i\cdot}), \forall i \in N. \tag{9}$$

The dual variables are also updated as:

$$\lambda^h_{i\cdot} = \lambda^{h-1}_{i\cdot} + \rho\alpha^h_{i\cdot}, \ \forall i \in N. \tag{10}$$

Note that at each iteration only the computed control variables will be shared. The matrices A and B remain the same and the knowledge about these matrices will be shared during the planning phase. This facilitates the distributed framework and minimizes information exchange among prosumers at each iteration.

When communication channels between prosumer i and its neighbor j is disrupted, prosumer i needs to predict the control action of prosumer j using the following equation

$$\widehat{U}_{ij}(k) = \begin{cases} U_{ij}(k) & U_{ij}(k) \text{ received} \\ \sum_{i=1}^q \phi_i \hat{U}_{i\cdot}(k-i) + \sum_{i=1}^q \theta_i P_{i\cdot}(k-i) + d(k) & \text{Otherwise} \end{cases} \tag{11}$$

In other words, prosumer i uses the previous states and control variables of prosumer j and applies the ARMA model to predict its control action at time k. As discussed in Sub-section 3.3, coefficients θ and ϕ will be learned from the historical data and will be updated frequently to keep the ARMA model up to date.

Algorithm for implementing ARMA-based DOFC

Step 0. *Prosumer i initializes its control variable (U^0_{ii}), shares its power deviations (p_i) with neighbors, and set $h = 0$.*

Step 1. *Prosumer i sends U^h_{ij} to its neighboring prosumers j.*

Step 2. *Prosumer j computes $\bar{U}^h_{jj}$ using* (7) *and sends it back to its neighbors.*

Step 3. *If prosumer i does not receive $\bar{U}^h_{ij}$, i.e., $\lambda_{ij}(k) = 0$, it estimates $\widehat{\bar{U}}^h_{ij}$*

from (12)*:*

$$\widehat{\widehat{U}}^{h}_{ij}(k) = \sum_{l=1}^{q} \phi_l \widehat{\widehat{U}}_{ij}(k-l) + \sum_{l=1}^{q} \theta_l p_j(k-l) + d(k). \tag{12}$$

Step 4. *Prosumer i updates its primal and dual residuals from* (9)*, and dual variable from* (10).

Step 5. *For given primal and dual tolerances $\epsilon^{Pri} > 0$ and $\epsilon^{Dual} > 0$, if $h > 0$, $\sum_{i \in N} \|\alpha^h_{i\cdot}\|_2^2 \leq \epsilon^{Pri}$ and $\sum_{i \in N} \|\beta^h_{i\cdot}\|_2^2 \leq \epsilon^{Dual}$ STOP and output $\bar{U}^h_{jj}$ as optimal control strategy; otherwise go to Step 5.*

Step 6. *Prosumer i updates $U^{h+1}_{i\cdot}$ by solving its cost minimization problem* (13)*. Set $h \leftarrow h+1$ and go to step 1.*

$$U^{h+1}_{i} = (2Q_i B_{i\cdot} B^{\top}_{i\cdot} + E^i)^{-1} (\rho \hat{\bar{U}}^h_{i\cdot} - \lambda^h_{i\cdot} - 2Q_i B_{i\cdot} A^{\top}_{i\cdot} P^i) \tag{13}$$

where E^i is a diagonal matrix of size $|\mathcal{N}_i| + 1$, $E^i_{ii} = 2R_i + \rho$ and $E^i_{jj} = \rho$, $\forall j \in \mathcal{N}_i$.

Note that control actions and power deviations have historical patterns. Thus, when there is delay in communication (missing data) or disruption in peer-to-peer communication (channel disconnection), prosumers can estimate the lost data using the ARMA model. Therefore, prosumers can still continue regulating frequency even during communication failures.

The ARMA model is more effective when the power system represents a linear behavior. When the power deviations are large, the system has a nonlinear behavior and therefore deep machine learning methods, such as LSTM, may be more effective. This will be the topic of our future research endeavor.

3.3. Least Mean Square System Identification

The Least Mean Square (LMS) algorithm is used to identify the coefficients of the ARMA model (θ and ϕ) based on the historical power flow and frequency control data of prosumers. This model allows estimating the missing data due to communication failures. Also, prosumers can predict the next control action of neighboring prosumers.

The LMS algorithm [43] is a stochastic gradient algorithm that uses a fixed step-size parameter to control the correction applied to each weight as the iterations proceed. In this paper, LMS uses a filtering process, which applies an input vector to calculate the error. The error is the difference between desired response

and the output of the filter. Denote

$$\begin{aligned} Y(k+1) =& U_{ij}(k+1), \\ U(k) =& [U_{ij}(k), U_{ij}(k-1), \ldots, U_{ij}(k-q+1), \\ & p_j(k), p_j(k-1), \ldots, p_j(k-q+1)]^\top, \\ W =& [\phi_1, \ldots, \phi_q, \theta_1, \ldots, \theta_q]^\top. \end{aligned} \tag{14}$$

Algorithm for LMS ARMA

Step 0. *Initialization: Select* $e_0 > 0$ *as the threshold error, given the initial values for* $W(0)$ *and* $U(0)$*, and let* $k = 0$*.*

Step 1. *Actual data:* $Y(k+1) = U(k)^\top W(k)$*.*

Step 2. *Output prediction:* $\widehat{Y}(k+1) = U(k)^\top W(k)$*.*

Step 3. *Calculating the error:* $e(k+1) = Y(k+1) - \widehat{Y}(k+1)$*.*

Step 4. *Calculating the updated weight:* $W(k+1) = W(k) + \eta U(k) e(k+1)$*.*

Step 5. *Termination condition: If* $|e(k+1)| \leq e_0$*, end the loop. Otherwise, let* $k = k+1$ *and go to Step 1.*

The step size parameter is η, which should be selected appropriately to guarantee fast convergence with reasonable steady-state error. Typically, a larger step size η can generate faster rate of initial convergence when the estimation error is large. But it will result in a larger persistent estimation due to observation noise. In addition, if the underlying system dynamics of the power grid changes with time, a large step size can have better tracking capability. When a constant step size is used, this tradeoff is fundamental and must be resolved in practical usage of LMS algorithms. In this paper, the step size is tuned in our case studies to achieve a reasonable tradeoff in the specific system.

3.4. Convergence Analysis

Next, we prove that under some mild assumptions, the ARMA-based DOFC algorithm can improve the convergence rate under communication failures. In this sub-section, we consider the gradient-based algorithm, which is easier to be analyzed. In a certain sense, ADMM algorithm is also a gradient-based algorithm. Using the performance index in (5), the distributed gradient of the performance index can be defined as $\nabla_u \mathcal{J}(u) = Gu + B^\top QAp$, where $G = R + B^\top QB$. Thus, the optimal solution can be obtained in a closed form as $u^* = -G^{-1}B^\top QAp$. This solution cannot be obtained in a distributed framework, since it requires calculating the inverse of matrix $R + B^\top QB$.

In order to find a distributed solution, we denote that for prosumer i the following performance index can be obtained:

$$\begin{aligned}\mathcal{J}_i(u) &= \frac{1}{2}(r_i u_i^2 + q_i(A_{i\cdot}p + B_{i\cdot}u)^2) \\ &= \frac{1}{2}\left[r_i u_i^2 + q_i\left(\sum_{j\in\mathcal{N}_i\cup\{i\}} A_{ij}p_j + \sum_{j\in\mathcal{N}_i\cup\{i\}} B_{ij}u_j\right)^2\right] \end{aligned} \tag{15}$$

Then, we have $\mathcal{J}(u) = \sum_{i=1}^n \mathcal{J}_i(u)$. Note that

$$\nabla_u \mathcal{J}(u) = \begin{pmatrix} \nabla_{u_1}\mathcal{J}(u) \\ \nabla_{u_2}\mathcal{J}(u) \\ \vdots \\ \nabla_{u_n}\mathcal{J}(u) \end{pmatrix} \in \mathbb{R}^n, \tag{16}$$

where

$$\nabla_{u_i}\mathcal{J}(u) = \sum_{\ell\in\mathcal{N}_i\cup\{i\}} \nabla_{u_i}\mathcal{J}_\ell(u), \tag{17}$$

and

$$\nabla_{u_i}\mathcal{J}_\ell(u) = \begin{cases} r_\ell u_\ell + q_\ell B_{\ell\ell}\left(\sum\limits_{j\in\mathcal{N}_\ell\cup\{\ell\}} A_{\ell j}p_j + \sum\limits_{j\in\mathcal{N}_\ell\cup\{\ell\}} B_{\ell j}u_j\right) & \text{if } \ell = i, \\ q_\ell B_{\ell i}\left(\sum\limits_{j\in\mathcal{N}_\ell\cup\{\ell\}} A_{\ell j}p_j + \sum\limits_{j\in\mathcal{N}_\ell\cup\{\ell\}} B_{\ell j}u_j\right) & \text{if } \ell \in i. \end{cases}$$

Thus, prosumer i can adopt the following gradient algorithm to track the optimal solution:

$$\begin{aligned} u_i(k+1) &= u_i(k) - \mu(k)\nabla_{u_i}\mathcal{J}(u(k)) \\ &= u_i(k) - \mu(k)\sum_{\ell\in\mathcal{N}_i\cup\{i\}} \nabla_{u_i}\mathcal{J}_\ell(u(k)), \end{aligned} \tag{18}$$

where the step size $\mu(k) = 1/k^\gamma$, $1/2 < \gamma < 1$. This algorithm is strictly distributed, since for prosumer i it only requires the gradient information from its neighbors, i.e., $\nabla_{u_i}\mathcal{J}_\ell(u(k))$ where $\ell \in \mathcal{N}_i \cup \{i\}$. Since p_j can be measured via

tie-line flows and is the physical power imbalance, only u_j needs to be shared with neighbors. If prosumer i does not receive $u_j(k)$, then the ARMA-based estimate $\widehat{u}_{ji}(k)$ will be used.

Thus, the update algorithm with the prediction error and gradient error can be written in the following vector form:

$$u(k+1) = u(k) - \mu(k)(Gu(k) + B^\top QAp + d_1(k) + d_2(k)), \tag{19}$$

where $d_1(k) \in \mathbb{R}^n$ denotes the prediction error and $d_2(k) \in \mathbb{R}^n$ denotes the gradient error. Next, we make the following basic assumptions for the theoretical analysis.

Assumption 3.1. *1. power graph is connected.*

2. *The prediction error $\{d_1(k) \in \mathbb{R}^{n\times 1}\}$ and the gradient error $\{d_2(k) \in \mathbb{R}^{n\times 1}\}$ are two mutually independent sequences of i.i.d. random variables such that $\mathbb{E}[d_1(k)] = \mathbb{E}[d_2(k)] = \mathbf{0}_{n\times 1} \in \mathbb{R}^{n\times 1}$, $\mathbb{E}[d_1(k)d_1^\top(k)] = \Sigma_{d_1} \in \mathbb{R}^{n\times n}$, $\mathbb{E}[d_2(k)d_2^\top(k)] = \Sigma_{d_2} \in \mathbb{R}^{n\times n}$, where Σ_{d_1} and Σ_{d_2} are symmetric positive definite.*

Using the ordinary differential equation (ODE) method in stochastic approximation [44], the following strong convergence result can be obtained for (19).

Theorem 3.1. *Under* Assumption 3.1*, the iterates $\{u(k)\}$ generated by the DOFC algorithm* (19) *converge to the optimal solution $u(k) \to u^*$ with probability one (w.p.1) as $k \to \infty$.*

For simplicity, we omit the verbatim proof and refer the reader to [44, Chapters 5 and 6]. While the actual proof of Theorem 3.1 will be skipped, the main ideas can be summarized as follows. Define $\xi(k) = \sum_{j=0}^{k-1} \mu(j)$, $\varpi^\xi = \max\{k : \xi(k) \le \xi\}$, the piecewise constant interpolation $u_0^\xi = u(k)$ for $\xi \in [\xi(k), \xi(k+1))$, and the shift sequence $u_k^\xi = u_0^{\xi+\xi(k)}$. Since the step size $\mu(k) = 1/k^\gamma$ $(1/2 < \gamma < 1)$, the interpolated sequence $\{u_t^{(\cdot)}\}$ is uniformly bounded and equicontinuous. By Ascoli-Arzéla's theorem, we can extract a subsequence $\{u_{t_\ell}^{(\cdot)}\}$, which converges to $u(\cdot)$ on any compact intervals w.p.1 such that $u(\cdot)$ is a solution. The ODE has a unique equilibrium point, which is the optimal solution. Now, by using the Lyapunov method, the equilibrium point u^* is an asymptotically stable point, since $-G$ is stable. This theoretical result leads to the desired property for (19).

Next, we demonstrate the convergence rate of the DOFC algorithm. Here, we consider the following averaging form:

$$\bar{u}(k) = \sum_{j=0}^{k-1} u^j/k.$$

Theorem 3.2. *Under* Assumption 3.1, $\sqrt{k}(\bar{u}(k) - u^*)$ *converges weakly to a normal random variable with zero mean and asymptotic covariance*

$$\Sigma^* = G^{-1}(\Sigma_{d_1} + \Sigma_{d_2})G^{-1}. \tag{20}$$

Remark 3.1. *For a coverage on the CR lower bound, we refer the reader to [45] (pp. 300-302). Since the proof of* Theorem 3.2 *is similar to Chapter 11 of [44], here we omit it. Note that $\bar{u}(k) - u^*$ is asymptotically normal (Gaussian distributed) with zero mean and covariance Σ^*/k. Thus, $\bar{u}(k)$ will converge to its limit at a convergence rate that approaches asymptotically the corresponding CR lower bound [46]. As a result, we use the error covariance matrix Σ^* to evaluate how fast convergence to the optimal solution can be achieved. It can be observed that by increasing the accuracy of the estimate $\widehat{u}_{ji}(k)$, the prediction error variance Σ_{d_1} decreases, thus the algorithm converges faster to the optimal solution.*

By [47], if all the nodes have a uniform packet delivery probability $\rho \in (0, 1]$, the asymptotic covariance will be $\Sigma^*_\rho = \rho^{-1}G^{-1}\Sigma_{d_2}G^{-1}$. Note that $\Sigma^*_\rho - \Sigma^* = G^{-1}[(\rho^{-1}-1)\Sigma_{d_2} - \Sigma_{d_1}]G^{-1}$. If $(1-\rho)\Sigma_{d_2} \geq \rho\Sigma_{d_1}$, we know that $\Sigma^*_\rho \geq \Sigma^*$ holds, i.e. if ρ or Σ_{d_1} is small enough, the ARMA-based DOFC algorithm will converge to the optimal solution faster than the DOFC algorithm with packet loss.

4. Simulation Results

In this section, we evaluate the performance of the ARMA-based DOFC on two practical power systems. The first system is the electric power grid on Flores Island and the second system is the electric power grid on Sao Miguel Island.

4.1. Flores Island

Flores Island is one of the smaller islands of the Azores Archipelago. Figure 2 shows the electrical network of Flores, which consists of a 15 kV radial distribution network with 45 nodes and 44 branches. The total demand of the island is around 2MW. Three small power plants supply the electrical demand. More than 50% of the electricity is provided by four diesel generators whose total capacity

is 2.5MW. Around 35% of the demand is supplied by four hydro power plants with an overall capacity of 1.65MW. Two wind power plants with a total capacity of 0.65MW provide the rest of demand (15%) [48]. We cluster the island into three prosumers, where each prosumer has power generation units and local loads. Figure 3 illustrates the equivalent prosumer-based structure of Flores Island. We assume that the physical communication network is the same as the topology of the prosumer grid.

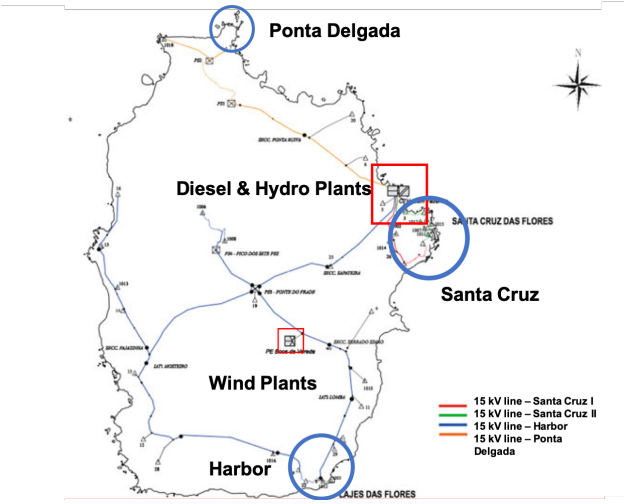

Figure 2: The electrical power grid on Flores Island.

We use the one-month historical power flow data of the island, which consists of 4464 data points, to train and test the LMS model and identify the coefficients of the ARMA model (θ and ϕ). 80% of the data was used for training and 20% was used to test the LMS model. Table IV shows the estimated ARMA coefficients of Flores Island. Also, Fig. 8 shows the accuracy of the LMS prediction model.

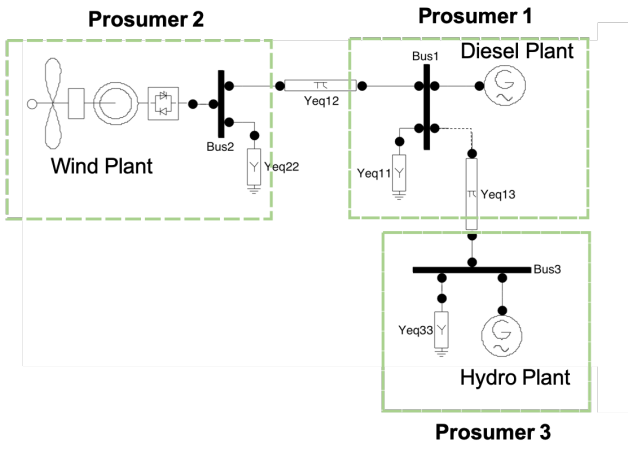

Figure 3: The equivalent prosumer-based power grid on Flores Island.

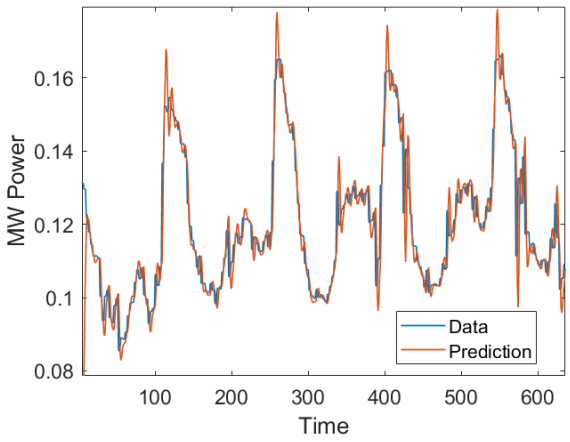

Figure 4: Comparing the actual data and prediction data of Diesel Prosumer for Flores Island.

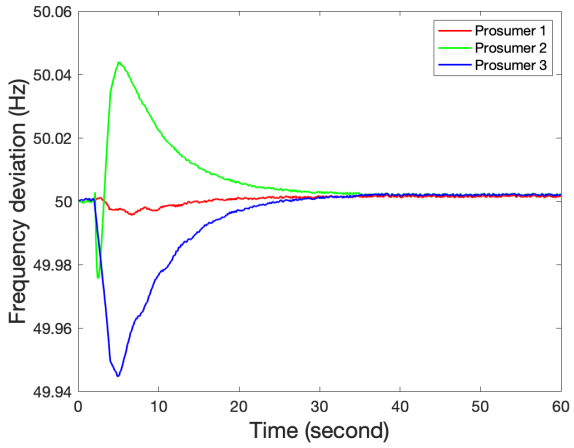

Figure 5: Frequency deviation of prosumers when there is no communication failures. Note that the ARMA-based DOFC performance (third scenario) is almost identical to this scenario.

Table 1: ARMA coefficients for Prosumers on Flores Island.

Number	Type	ϕ_1	ϕ_2	θ_1	θ_2
Prosumer 1	Diesel	-0.1760	-0.1581	$+0.4340$	$+0.8952$
Prosumer 2	Hydro	-0.2389	-0.3480	$+0.4721$	$+1.0978$
Prosumer 3	Wind	-0.2218	-0.2794	$+0.4609$	$+1.0350$

Next, we compare the performance of DOFC after a disturbance in three scenarios. The first scenario is the normal condition, in which there is no communication disconnection. In the second scenario, one communication node is disconnected (prosumer 3) and prosumers use the zero-order-hold method (Solution 2). In the third scenario, the communication links to prosumer 3 is disconnected, but the ARMA model is applied to predict the states of prosumer 3 (Solution 3).

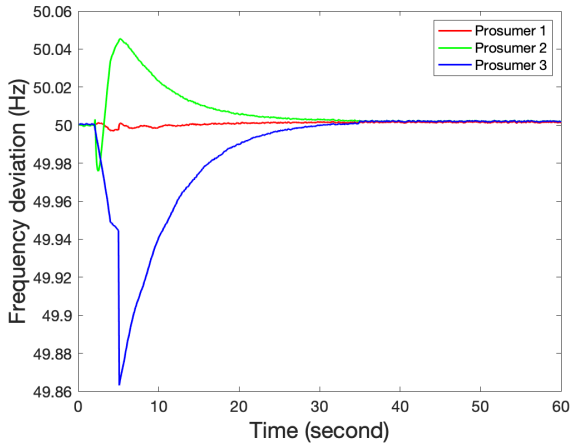

Figure 6: Frequency deviation of prosumers when prosumer 3 is disconnected (second scenario).

Fig. 5 shows the frequency dynamics of prosumers for the first scenario (no communication loss), and Fig. 6 presents the frequency dynamics of prosumers for the second scenario. Also, the third scenario has a similar frequency response to the normal condition (first scenario).

Fig. 7 compares the dynamic response of the center of inertia COI for all three scenarios. As shown in the figure, the zero-order-hold method has poor performance. On the other hand, the ARMA-based DOFC is effective in overcoming the negative impact of communication packet loss and achieving enhanced performance that is very close to the ideal condition of no communication loss.

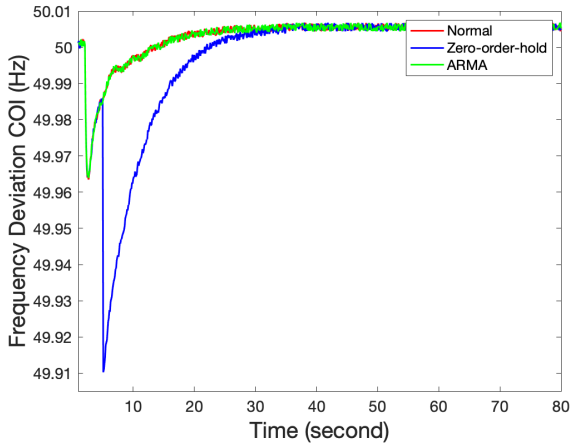

Figure 7: Frequency deviation of COI when prosumer 3 is disconnected (second scenario).

4.2. Sao Miguel Island

The next case study is Sao Miguel Island, which consists of a 60-kV transmission network connecting the large power plants to large loads. The average demand of the island is 70 MW and fifteen generators are serving the island. There are two large diesel generators located in the middle of the island. Also, there are two large geothermal plants and seven hydro plants. The schematics of the electrical method on Sao Miguel Island is provided in [25]. We assume that the island is clustered into fifteen prosumers, where each prosumer is a balancing area for frequency control. The equivalent prosumer-based network of Sao Miguel Island is shown in Figure 8. Similar to the previous case study, one-month historical power flow data of the island is used for the training and testing the LMS model. Table III shows the coefficients of the LMS model for prosumers on Sao Miguel Island.

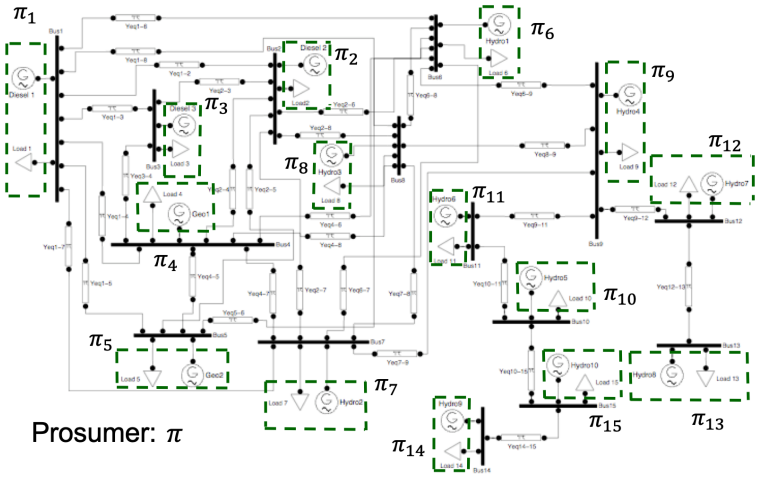

Figure 8: The equivalent prosumer-based power grid on Sao Miguel Island.

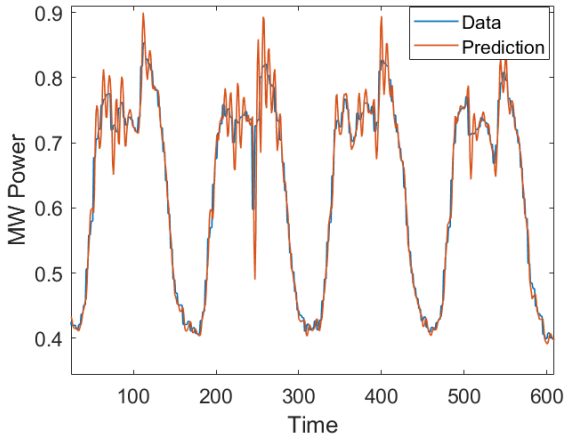

Figure 9: Comparing the actual data and LMS-based prediction data for Sao Miguel Island.

Table 2: Prediction Results of ARMA for Sao Miguel Island.

Prosumer type	ϕ_1	ϕ_2	θ_1	θ_2
Hydro1	$+0.2056$	$+0.1732$	-0.2524	-0.2110
Hydro2	$+0.0666$	$+0.0468$	-0.1765	-0.1019
Hydro3	$+0.0507$	$+0.0463$	-0.1424	-0.1186
Hydro4	$+0.0545$	$+0.1290$	-0.3256	-0.2825
Hydro5	$+0.2025$	$+0.2473$	$+0.3846$	$+0.3552$
Hydro6	-0.0591	-0.0174	-0.0280	$+0.0062$
Hydro7	$+0.1062$	$+0.1108$	$+0.3056$	$+0.2893$
Hydro8	$+0.0727$	$+0.0673$	$+0.2067$	$+0.1933$
Hydro9	-0.0246	$+0.0744$	$+0.1333$	$+0.1705$
Hydro10	$+0.1265$	$+0.3655$	$+0.3166$	$+0.3168$
Geothermal1	$+0.2056$	$+0.1732$	-0.2524	-0.2110
Geothermal2	$+0.0666$	$+0.0468$	-0.1765	-0.1019
Diesel1	$+0.0507$	$+0.0463$	-0.1424	-0.1186
Diesel2	$+0.0545$	$+0.1290$	-0.3256	-0.2825
Wind	$+0.2025$	$+0.2473$	$+0.3846$	$+0.3552$

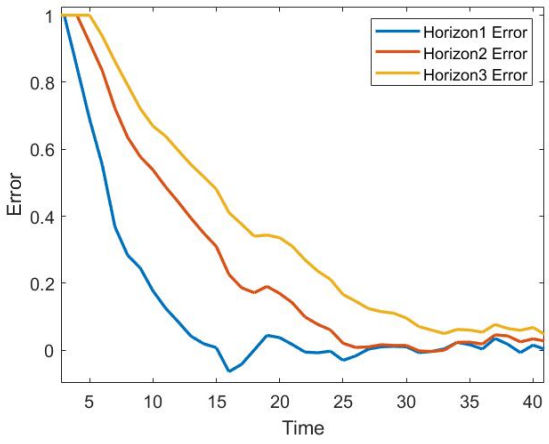

Figure 10: Performance of the LMS prediction under different time-step prediction horizons.

Fig. 9 also shows the performance of the LMS prediction model for one of the prosumers. As shown in the figure, the LMS model accurately predicts the next-

step power deviation ($k+1$). By increasing the prediction horizon, the accuracy of the LMS model decreases. This is shown in Fig. 10 for two-step ($k+2$) and three-step ($k+3$) prediction horizons.

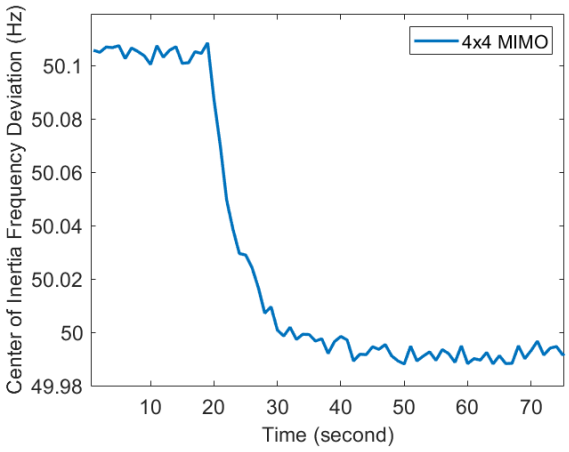

Figure 11: ARMA frequency deviation for Sao Miguel Island Island.

Next, we simulate the performance of DOFC after a power imbalance under three scenarios:

- Normal communication system.
- Communication disconnection at node 15 (prosumer 15). Prosumers use zero-order-hold method to mitigate the communication loss.
- Communication disconnection at node 15, but using ARMA model to predict the states of neighbors of prosumer 15.

Fig. 11 shows the performance of COI for the ARMA-based DOFC scenario. Also, Fig. 12 compares the quasi-steady state dynamics of COI for the three scenarios. As shown in the figure, the ARMA-based DOFC has very similar performance to the normal condition. In other words, the ARMA model allows the disconnected prosumer to accurately predict the states of its neighbors and continue performing DOFC.

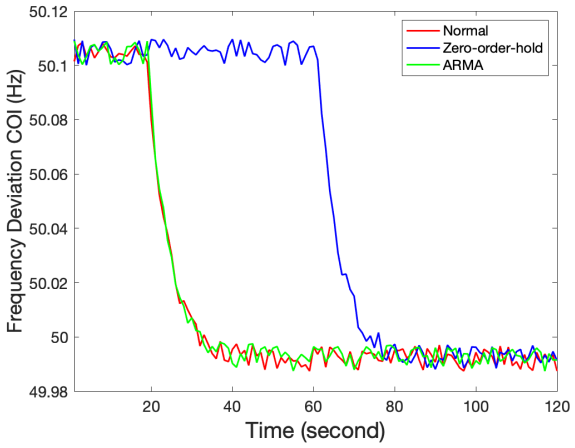

Figure 12: Frequency deviation of COI on Sao Miguel Island when three scenarios.

5. Conclusions

This paper proposed a learning framework based on the ARMA model to increase the resilience of distributed power algorithms, such as DOFC, against communication failures. The ARMA model estimates the missing states and control actions of neighboring agents, when there is loss of communication. Also, the ARMA model allows prosumers to predict the future control action of neighbors. Thus, even during communication failures, affected prosumers could continue distributed optimization and contribute to the system-level performance, such as optimal power sharing for frequency control. The performance of the proposed method was evaluated on two practical power systems. The results showed that the ARMA-based DOFC algorithm could asymptotically reach the same convergence rate of the power systems without communication interruptions.

The future research endeavored is to extend the model to more general nonlinear models, such as AC Optimal Power Flow. To this end, machine learning methods, such as neural network can be used to handle system non-linearities.

Acknowledgment

This work is supported by NSF DMS-2229109 grant.

References

[1] S. Grijalva, M. Costley, N. Ainsworth, Prosumer-based control architecture for the future electricity grid, in: Control Applications (CCA), 2011 IEEE International Conference on, IEEE, 2011, pp. 43–48.

[2] T. Ramachandran, Z. Costello, P. Kingston, S. Grijalva, M. Egerstedt, Distributed power allocation in prosumer networks, in: IFAC Necsys, 2012.

[3] Z. Wang, F. Liu, J. Z. F. Pang, S. H. Low, S. Mei, Distributed optimal frequency control considering a nonlinear network-preserving model, IEEE Transactions on Power Systems 34 (1) (2019) 76–86. doi:10.1109/TPWRS.2018.2861941.

[4] A. Abazari, H. Monsef, B. Wu, Coordination strategies of distributed energy resources including fess, deg, fc and wtg in load frequency control (lfc) scheme of hybrid isolated micro-grid, International Journal of Electrical Power & Energy Systems 109 (2019) 535–547. doi:https://doi.org/10.1016/j.ijepes.2019.02.029.
URL https://www.sciencedirect.com/science/article/pii/S0142061518333787

[5] C. Zhao, E. Mallada, S. H. Low, J. Bialek, Distributed plug-and-play optimal generator and load control for power system frequency regulation, International Journal of Electrical Power & Energy Systems 101 (2018) 1–12. doi:https://doi.org/10.1016/j.ijepes.2018.03.014.
URL https://www.sciencedirect.com/science/article/pii/S0142061517324171

[6] M. Velay, M. Vinyals, Y. Besanger, N. Retiere, Fully distributed security constrained optimal power flow with primary frequency control, International Journal of Electrical Power & Energy Systems 110 (2019) 536–547. doi:https://doi.org/10.1016/j.ijepes.2019.03.028.
URL https://www.sciencedirect.com/science/article/pii/S0142061518316260

[7] N.-L. Mo, Z.-H. Guan, D.-X. Zhang, X.-M. Cheng, Z.-W. Liu, T. Li, Data-driven based optimal distributed frequency control for islanded ac microgrids, International Journal of Electrical Power & Energy Systems 119 (2020) 105904. doi:https://doi.org/10.1016/j.ijepes.2020.105904.

URL https://www.sciencedirect.com/science/article/pii/S0142061519331059

[8] A. Kargarian, J. Mohammadi, J. Guo, S. Chakrabarti, M. Barati, G. Hug, S. Kar, R. Baldick, Toward distributed/decentralized dc optimal power flow implementation in future electric power systems, IEEE Transactions on Smart Grid 9 (4) (2016) 2574–2594.

[9] X.-K. Liu, Y.-W. Wang, P. Lin, P. Wang, Distributed supervisory secondary control for a dc microgrid, IEEE Transactions on Energy Conversion 35 (4) (2020) 1736–1746. doi:10.1109/TEC.2020.2994251.

[10] D. K. Molzahn, F. Dörfler, H. Sandberg, S. H. Low, S. Chakrabarti, R. Baldick, J. Lavaei, A survey of distributed optimization and control algorithms for electric power systems, IEEE Transactions on Smart Grid 8 (6) (2017) 2941–2962.

[11] T. Ramachandran, M. H. Nazari, S. Grijalva, M. Egerstedt, Overcoming communication delays in distributed frequency regulation, IEEE Transactions on Power Systems 31 (4) (2016) 2965–2973.

[12] C. Lenzen, R. Wattenhofer, Distributed algorithms for sensor networks, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 370 (1958) (2012) 11–26.

[13] F. Dorfler, S. Grammatico, Gather-and-broadcast frequency control in power systems, Automatica 79 (2017) 296 – 305. doi:https://doi.org/10.1016/j.automatica.2017.02.003.
URL http://www.sciencedirect.com/science/article/pii/S0005109817300572

[14] M. Kraning, E. Chu, S. Boyd, J. Lavaei, Dynamic Network Energy Management Via Proximal Message Passing, Foundations and Trends in Optimization Series, Now Publishers, 2013.

[15] A. Abboud, R. Couillet, M. Debbah, H. Siguerdidjane, Asynchronous alternating direction method of multipliers applied to the direct-current optimal power flow problem, in: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2014, pp. 7764–7768.

[16] S. Chakrabarti, M. Kraning, E. Chu, R. Baldick, S. Boyd, Security constrained optimal power flow via proximal message passing, in: 2014 Clemson University Power Systems Conference, 2014, pp. 1–8.

[17] M. Ilic, H. Allen, W. Chapman, C. King, J. Lang, E. Litvinov, Preventing Future Blackouts by Means of Enhanced Electric Power Systems Control: From Complexity to Order, Proceedings of The IEEE 93 (2005) 1920–1941.

[18] M. Ilic, X. Liu, A modeling and control framework for operating large-scale electric power systems under present and newly evolving competitive industry structures, Mathematical Problems in Engineering 1 (4) (1995) 317–340.

[19] D. Ke, C. Chung, X. Yusheng, An eigenstructure-based performance index and its application to control design for damping inter-area oscillations in power systems, Power Systems, IEEE Transactions on 26 (4) (2011) 2371–2380.

[20] D. Rai, S. Faried, G. Ramakrishna, A. Edris, Damping inter-area oscillations using phase imbalanced series compensation schemes, Power Systems, IEEE Transactions on 26 (3) (2011) 1753–1761.

[21] M. H. Nazari, M. Ilic, Dynamic modelling and control of distribution energy systems: Comparison with transmission power systems, IET Generation, Transmission and Distribution (2013).

[22] W. Zeng, M.-Y. Chow, Resilient distributed control in the presence of misbehaving agents in networked control systems, IEEE Transactions on Cybernetics 44 (11) (2014) 2038–2049. doi:10.1109/TCYB.2014.2301434.

[23] Q. Zhou, M. Shahidehpour, A. Alabdulwahab, A. Abusorrah, A cyber-attack resilient distributed control strategy in islanded microgrids, IEEE Transactions on Smart Grid 11 (5) (2020) 3690–3701. doi:10.1109/TSG.2020.2979160.

[24] L. Chen, Y. Wang, X. Lu, T. Zheng, J. Wang, S. Mei, Resilient active power sharing in autonomous microgrids using pinning-consensus-based distributed control, IEEE Transactions on Smart Grid 10 (6) (2019) 6802–6811. doi:10.1109/TSG.2019.2911344.

[25] M. H. Nazari, S. Xie, L. Yi Wang, Communication-delay-tolerant mimo architecture for distributed frequency regulation in multi-agent smart grids, International Journal of Electrical Power & Energy Systems (In Press).

[26] M. H. Nazari, L. Y. Wang, S. Grijalva, M. Egerstedt, Communication-failure-resilient distributed frequency control in smart grids: Part i: Architecture and distributed algorithms, IEEE Transactions on Power Systems 35 (2) (2020) 1317–1326. doi:10.1109/TPWRS.2019.2943820.

[27] M. H. Nazari, L. Y. Wang, S. Grijalva, M. Egerstedt, Communication-failure-resilient distributed frequency control in smart grids: Part ii: Algorithmic implementation and system simulations, IEEE Transactions on Power Systems 35 (4) (2020) 3192–3202. doi:10.1109/TPWRS.2020.2968581.

[28] J. Coulson, J. Lygeros, F. Dörfler, Distributionally robust chance constrained data-enabled predictive controlSubmitted. Available at https://arxiv.org/abs/2006.01702. (2020).
URL https://arxiv.org/abs/2006.01702

[29] Guest editorial: Special issue on security and privacy of distributed algorithms and network systems, IEEE Transactions on Automatic Control 65 (9) (2020) 3725–3727. doi:10.1109/TAC.2020.3004329.

[30] J. S. Gómez, D. Sáez, J. W. Simpson-Porco, R. Cárdenas, Distributed predictive control for frequency and voltage regulation in microgrids, IEEE Transactions on Smart Grid 11 (2) (2020) 1319–1329. doi:10.1109/TSG.2019.2935977.

[31] T. Anderson, An introduction to multivariate statistical analysis, John Wiley & sons, New York, USA, 1984.

[32] W. Wang, Y. Xu, M. Khanna, A survey on the communication architectures in smart grid, Computer Networks 55 (15) (2011) 3604 – 3629.

[33] I. 61850-5, Communication networks and systems in substations–part 5: Communication requirements for functions and device models (2003).

[34] V. Skendzic, A. Guzma, Enhancing power system automation through the use of real-time ethernet, in: Power Systems Conference: Advanced Metering, Protection, Control, Communication, and Distributed Resources, 2006. PS'06, IEEE, 2006, pp. 480–495.

[35] North American Reliability Corporation, BAL-001-2 –Real Power Balancing Control Performance Standard Background Document, www.nerc.com (February 2013).

[36] H. F. Illian, Frequency control performance measurement and requirements, Lawrence Berkeley National Laboratory (2011).

[37] M. Korkali, J. Veneman, B. Tivnan, et al., Reducing cascading failure risk by increasing infrastructure network interdependence, Sci Rep 7 (2017) 44499. URL https://doi.org/10.1038/srep44499

[38] M. H. Nazari, L. Y. Wang, S. Grijalva, M. Egerstedt, Communication-failure-resilient distributed frequency control in smart grids: Part i: Architecture and distributed algorithms, IEEE Transactions on Power Systems 35 (2) (2020) 1317–1326. doi:10.1109/TPWRS.2019.2943820.

[39] M. H. Nazari, L. Y. Wang, S. Grijalva, M. Egerstedt, Communication-failure-resilient distributed frequency control in smart grids: Part ii: Algorithmic implementation and system simulations, IEEE Transactions on Power Systems 35 (4) (2020) 3192–3202. doi:10.1109/TPWRS.2020.2968581.

[40] M. H. Nazari, Z. Costello, M. Feizollahi, S. Grijalva, M. Egerstedt, Distributed frequency control of prosumer-based electric energy systems, IEEE Transactions on Power Systems 29 (6) (2014) 2934–2942.

[41] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, Foundations and Trends in Machine Learning 3 (1) (2011) 1–124.

[42] M. J. Feizollahi, M. Costley, S. Ahmed, S. Grijalva, Large-scale decentralized unit commitment, International Journal of Electrical Power & Energy Systems 73 (2015) 97–106.

[43] B. Widrow, J. M. McCool, M. G. Larimore, C. R. Johnson, Stationary and nonstationary learning characteristic of the lms adaptive filters, Proceedings of the IEEE 64 (1976) 1151–1162. doi:10.1007/978-94-010-1223-2_23.

[44] H. Kushner, G. Yin, Stochastic approximation and recursive algorithms and applications, 2nd ed. New York, NY, USA: Springer-Verlag, 2003.

[45] J. A. Rice, Mathematical statistics and data analysis, Brooks/Cole, 2007.

[46] G. Yin, On extensions of polyak's averaging approach to stochastic approximation, Stochastics: An International Journal of Probability and Stochastic Processes 36 (3-4) (1991) 245–264.

[47] M. H. Nazari, S. Xie, L. Yi Wang, G. Yin, W. Chen, Impact of communication packet delivery ratio on reliability of optimal load tracking and allocation in dc microgrids, IEEE Transactions on Smart Grid 12 (4) (2021) 2812–2821. doi:10.1109/TSG.2021.3062024.

[48] M. H. Nazari, Electrical Networks of the Azores Archipelago, in Chapter 3 of Engineering IT-Enabled Sustainable Electricity Services, Springer, 2013.