2021

Distributed Control, Optimization, and State Estimation for Renewable Power System

Qiao Li
University of Denver

Distributed Control, Optimization, and State Estimation for Renewable Power System

Abstract

The traditional power systems are usually centralized systems, in which the control, operation and monitoring are performed by the centralized control center, e.g., SCADA. However, with the development of renewable energy, power systems are getting more and more distributed. So, it becomes necessary to establish the distributed power system operation methods for these power systems. In this research, the distributed techniques for the renewable power systems are proposed based on the consensus protocol technique from graph theory. These techniques cover the three important problems in power systems, i.e., economic dispatch, state estimation, and optimal power flow. First, the Distributed Economic Dispatch (DED) approach is proposed. In this part, both the PI controller and Neural Network (NN) controller are utilized to design the distributed algorithm to minimize the power system's operational cost in a distributed way. The communication-failure-tolerant DED algorithm is proposed to improve the robustness of the approach during communication failure. Also, the DED algorithm considering line loss model is proposed. On the other hand, an information propagation method is provided to develop the Distributed State Estimation (DSE) algorithm. Then, the bad data detection and measurement accuracy improvement topics in state estimation are discussed. Then, based on the proposed DED algorithm and DSE algorithm, the Distributed Optimal Power Flow (DOPF) method is developed. Finally, the AC power flow model is considered to build the distributed AC State Estmiation method and distributed AC Optimal Power Flow method. At the end, the proposed methods are verified in the MATLAB/simulation software. The 4-generator system model, IEEE 10-generator 39-bus system model, WSCC 9-Bus system model, and some specially designed power system models are employed in the tests. The results of the simulation show that the proposed methods reach the desired performance.

Document Type

Dissertation

Degree Name

Ph.D.

Department

Electrical Engineering

First Advisor

David Gao

Second Advisor

Yunbo Yi

Third Advisor

Margareta Stefanovic

Keywords

Consensus protocol, Economic dispatch, Optimal power flow, Renewable energy, State estimation

Subject Categories

Electrical and Computer Engineering | Electrical and Electronics | Engineering | Power and Energy

DISTRIBUTED CONTROL, OPTIMIZATION, AND STATE ESTIMATION FOR RENEWABLE POWER SYSTEM

_____

A Dissertation

Presented to

the Faculty of the Daniel Felix Ritchie School of Engineering and Computer Science

University of Denver

_____

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

_____

by

Qiao Li

November 2021

Advisor: Dr. David Gao

Author: Qiao Li
Title: Distributed Control, Optimization, and State Estimation for Renewable Power System
Advisor: Dr. David Gao
Degree Date: November 2021

Abstract

The traditional power systems are usually centralized systems, in which the control, operation and monitoring are performed by the centralized control center, e.g., SCADA. However, with the development of renewable energy, power systems are getting more and more distributed. So, it becomes necessary to establish the distributed power system operation methods for these power systems. In this research, the distributed techniques for the renewable power systems are proposed based on the consensus protocol technique from graph theory. These techniques cover the three important problems in power systems, i.e., economic dispatch, state estimation, and optimal power flow. First, the Distributed Economic Dispatch (DED) approach is proposed. In this part, both the PI controller and Neural Network (NN) controller are utilized to design the distributed algorithm to minimize the power system's operational cost in a distributed way. The communication-failure-tolerant DED algorithm is proposed to improve the robustness of the approach during communication failure. Also, the DED algorithm considering line loss model is proposed. On the other hand, an information propagation method is provided to develop the Distributed State Estimation (DSE) algorithm. Then, the bad data detection and measurement accuracy improvement topics in state estimation are discussed. Then, based on the proposed DED algorithm and DSE algorithm, the Distributed Optimal Power Flow (DOPF) method is developed. Finally, the AC power flow model is considered to build the distributed AC State Estmiation method and distributed AC Optimal Power Flow method.

At the end, the proposed methods are verified in the MATLAB/simulation software. The 4-generator system model, IEEE 10-generator 39-bus system model, WSCC 9-Bus system model, and some specially designed power system models are employed in the tests. The results of the simulation show that the proposed methods reach the desired performance.

Acknowledgements

Thanks to my advisors Dr. David Wenzhong Gao for the great support and professional suggestions.

Table of Contents

List of Figures

List of Tables

List of Symbols

Symbol	Definition
Constants:	
f_{rated}	The rated frequency (60Hz or 50Hz)
P_i^{max}	The upper limit of the real power output of the ith generator.
P_i^{min}	The lower limit of the real power output of the ith generator.
P_{ij}^{max}	The limit of active power flow on the transmission line between bus i and bus j.
Q_{ij}^{max}	The limit of reactive power flow on the transmission line between bus i and bus j.
P_f^{max}	The vector of line flow limits.
Indices:	
i, j	Index for bus.
n_b	Number of buses.
n_g	Number of generators.
Sets:	
G	Graph.
$\mathcal{N}_i$	Neighborhood set.
E	Edge set.
V	Vertex set.
Matrices:	
I	Incidence matrix.
A	Adjacency matrix.
D	Degree matrix.
L	Laplacian.
L_w	Weighted Laplacian.
M_w	The state transfer matrix of the consensus protocol.
$\boldsymbol{X}(k)$	The vector containing the states of all vertices at time k.
H	Observation matrix.
$\boldsymbol{\lambda}(k)$	The vector of $\lambda_i(k)$.
$P_f(k)$	The vector of line flow of all lines at time k.
B'	“B-prime" matrix in DC power flow model.
$OF(k)$	The indicator vector of overflowed power lines at time k.
$H_v(k)$	The matrix related to the lines which will violate the constraint at time k.
M_b	The basis matrix of the solution space of the homogeneous equation.
Parameters:	
$\alpha_i, \beta_i, \gamma_i$	The cost function parameters of the ith generator.

L_i	The parameters of the line losses model
w_{ij}	The weight coefficient on the link between v_i and v_j.
φ_1, φ_2	The weight coefficient in the Auto-Regressive model.
x_{ij}	The reactance between bus i and bus j.
G_{ij}	The conductance on the transmission line between bus i and bus j.
B_{ij}	The susceptance on the transmission line between bus i and bus j.
B_{ij}^{cap}	The line charging susceptance on the transmission line from bus i to bus j .

Variables:

v_i, v_j	The ith vertex in the graph.
$d(v_i)$	The degree of the vertex v_i.
λ	Lagrange multiplier.
$\lambda_i(k)$	The estimated λ by the ith generator at time k.
$\rho(M)$	The maximum eigenvalue (in absolute value) of the matrix M.
τ	The time interval between two update iterations.
e_l	The lth edge.
$f_i(k)$	The frequency measured by the ith generator at k.
$F_i(P_i)$	The cost function of the ith generator.
$G_i^*(k)$	The goal attraction function.
m	Number of edges in the graph.
n	Number of nodes in the graph.
$P_i^*(k)$	The solution of the ED problem for the ith generator at time k.
P_D	The total real power demand of the power system.
P_i^D	The real power demand on the bus i.
Q_i^D	The reactive power demand on the bus i.
P_i^G	The real power generation on the bus i.
Q_i^G	The reactive power generation on the bus i.
P_i	The ith generator's power production.
P_{ij}	The real power flow from bus i to bus j.
Q_{ij}	The reactive power flow from bus i to bus j.
P_{load}	The total real power load in the system.
P_{loss}	The total real power loss in the system.
$x_i(k)$	The state of the ith vertex at time k.
M_{ij}	The measured power flow between node i and j.
θ_i	The phase angle on the bus i.
V_i	The voltage on the bus i.
η	The noise or error in the measurement.
z_i	The reading on the ith meter.
Z_i	The estimated measurements on the ith node.
$P_{g,i}$	The power reference of generator i.

$\Delta P_{g,i}^{j}(k)$	The update of generation reference of the ith generator calculated by the jth generator.
$\Delta P_{g}^{\eta}(k)$	The particular solution part in the power reference correction at time k.
$\Delta P_{g}^{\zeta}(k)$	The general solution part in the power reference correction at time k.
$\Delta P_{g}^{\pi}(k)$	The penalty term in the power reference correction at time k.

Chapter 1. Introduction

In recent decades, renewable energy is getting more and more popular, a lot of PV panels and wind turbines are connected to the power grid, in which power generation becomes distributed in modern power systems. With the distributed energy resources (DER) being connected, it is more complex and less efficient to use the traditional centralized control center to coordinate all of these distributed facilities. Therefore, it is necessary to develop a distributed control, operation, and optimization scheme to achieve higher efficiency and better performance in distributed power systems.

There are several advantages for distributed methods. Because of the distributed characteristic, the distributed system can be more robust, more economical, and higher efficient. For example, in [1], a distributed algorithm for the electricity market is proposed to protect the privacy of load aggregators and generators and to reduce the computational complexity of the centralized approach. This is because that the connectivity of the distributed structure is usually higher than the centralized structure (i.e. star network), and the calculation can be done by a bunch of cheap devices in parallel in a distributed system instead of the expensive control center. Besides these advantages, the development of distributed control system also benefits the conventional centralized power system, since the distributed control system can be implemented as an auxiliary or backup system for the centralized control system to improve the robustness and efficiency of the power system.

In this research, three important topics for power systems will be covered and the distributed solutions for these topics will be proposed. These topics are: Economic Dispatch (ED), State Estimation (SE), and Optimal Power Flow (OPF).

Economic dispatch is an important problem in power systems. The solution of the economic dispatch ensures the power system to be operated under the most economical condition. State estimation is also an important technique for power system operation. It is usually performed by Energy Management System (EMS) [2-4] in control centers to acquire the states of the power system. The major objective of state estimation is to monitor the states of power systems, estimate the unmeasured data, improve measurement accuracy, and detect bad measurements. Optimal power flow is a technique used in power system control centers to achieve optimal operation under certain constraints. Similar to the economic dispatch problem, the OPF schemes minimize the operating cost of power systems. In addition, more power flow related constraints are considered in OPF rather than the simple cost minimization in the ED problem. In this research, the power flow limits on all power lines in the power system will be considered in the OPF problem.

Based on the consensus protocol technique [5], the distributed version of the above techniques is realized in this research. They are Distributed Economic Dispatch (DED), Distributed State Estimation (DSE), and Distributed Optimal Power Flow (DOPF). For the DSE and DOPF methods, both the DC version which is developed on the basis of the DC power flow model and the AC version which considers the AC power flow model are introduced in this work,

Chapter 2. Literature Review

Different methods have been proposed to realize the distributed algorithms for different topics in power systems. In this section, the existing methods for distributed economic dispatch, distributed state estimation, and distributed optimal power flow are introduced.

2.1. Distributed Economic Dispatch Methods

The conventional economic dispatch approaches are designed mainly for centralized power systems, in which a centralized control center is needed to calculate the power output reference for each generator. Many different methods have been developed to solve the ED problem in the conventional centralized power system. These methods include Lambda-iteration, gradient methods, Newton's method, and dynamic programming [6]. Since the ED is an optimization problem, the heuristic optimization methods, such as genetic algorithm [7], particle swarm optimization [8], and Neural Network Approach [9], are also applicable for the ED problem. Also, an advanced method like oblivious network design has been employed in [10] to develop the new oblivious routing economic dispatch (ORED) algorithm, in which the economic dispatch can be achieved while managing congestion and mitigating power losses.

Some distributed economic dispatch methods have been investigated recently as well. One of the major methods is the consensus-based DED approaches [11-17]. These methods are mainly based on the consensus protocol from the graph-theoretic method [5].

Since the necessary condition for the solution to the ED problem requires the Lagrange multipliers λ of all generators to be identical, the consensus protocol was applied to achieve all equal Lagrange multipliers in these research. However, the solution of the distributed ED problem should also meet the power balance constraint (the power supply is equal to the power consumption), which has not been solved perfectly in the previous works. As said in [18], most of these methods generally require a centralized node to ensure that the demand-supply balance is satisfied. For example, a consensus-based algorithm was proposed in [11] to solve the DED problem in Microgrids, in which the optimal solution is obtained by introducing the global supply-demand mismatch value (power imbalance). This value is estimated locally by each generator through an estimation algorithm, and the result of the estimation heavily depends on the initial value. However, a centralized information center is still required to calculate the initial value of the mismatch, which makes the method not totally distributed. Also, the estimation algorithm in this paper cannot deal with the load variations. In [12], the authors discussed a consensus-based distributed gradient algorithm for the DED problem. However, the authors assume that the total demand in the power system is known by each generator, which is hard to be obtained in a totally distributed system without a centralized node. Similarly, in [15, 16], the authors developed the distributed economic dispatch methods by assuming that the power imbalance between the total load and the total generation is known. So, these methods failed to be totally distributed due to their strong assumption on the availability of the power mismatch in distributed systems. Differently, in [17], a lambda consensus protocol with an innovation term was introduced to solve the DED problem for a distributed power

system, and the innovation term was used to ensure the power balance in the power system. However, the convergence of the protocol cannot be guaranteed by the algorithm due to the lack of power mismatch information. Thus, two decaying weight parameters α_t and β_t were applied to force the solution to converge, but the optimal solution cannot be obtained. In order to solve this problem, several communication protocols, such as the “timestamp" algorithm in [13] and graph discovery algorithm in [19], are proposed to collect the load and generation information in the power system, then calculate the power imbalance. However, this method requires the algorithm to be run on both the generators and the loads, which increases the cost and complexity of the system.

2.2. Distributed State Estimation Methods

Some efforts have been made in recent years to develop the DSE method. The most commonly used method is the distributed multi-area state estimation as in [20-25]. In this type of method, the entire power system is partitioned into several small areas. Each area has a local state estimator to estimate the states of its own buses and communicates with the nearby estimators to ensure the observability of the estimation. The state estimation in this method is realized by decomposing the centralized state estimation problem into several smaller problems, e.g., using the Alternating Direction Method of Multipliers (ADMM) method. However, since there are some requirements for the area partitioning, such as the size of the areas, the way to divide the system, and the overlapping between the areas, this method may not be compatible with many power systems. Also, this method is not fully distributed since it has a centralized structure (all sensors in an area are connected

to the local estimator) inside each area and the state estimation algorithm is partially centralized (all data in an area are assembled at the estimator for calculation).

Some other methods are also proposed to address the DSE problem, e.g. [26, 27]. However, centralized facilities, such as GPS (Global Positioning System) and SCADA (Supervisory Control and Data Acquisition), are still required in these methods. So, these methods are not totally distributed as well.

According to the above discussion, the existing DSE methods are not totally distributed. This drawback makes the distributed power system less robust since a centralized structure still exists in the system. Also, due to the same reason, these methods are not compatible with some other distributed techniques in power systems, e.g. distributed economic dispatch methods [28], so that the development of the more advanced distributed techniques in power systems is limited.

2.3. Distributed Optimal Power Flow Methods

There are some existing papers for distributed OPF problem. In [29], a distributed OPF algorithm is developed, in which the alternating direction multiplier method (ADMM) is used to decompose the optimization problem into several subproblems. However, in this method, the power system is required to be divided into several areas, and the centralized method is still used within each area. So, the algorithm is not completely distributed. On the other hand, only the local power constraint and voltage limit constraint are considered in [29], but the constraints related to multiple buses, such as line flow constraint, are not discussed. It makes the proposed method in [29] less useful in practice since many important problems in OPF are multi-bus problems, e.g., the line flow constraint and the

N-1 contingency problem. Similarly, three Distributed DC Optimal Power Flow (DDCOPF) methods based on ADMM are proposed in [30], in which the power system is also divided into small areas, and the centralized method is also needed within each area. Similarly, the ADMM-based methods are discussed in [31-33]. Since ADMM-based methods require the power system to be partitioned into sub-areas, these methods are not totally distributed methods.

Other distributed OPF methods are also discussed in [34-38]. These papers are also based on the area partitioning method which is not fully distributed. In [39], the authors proposed a Distributed Security-Constrained Unit Commitment (DSCUC) which is closely related to distributed OPF. The paper uses the analytical target cascading (ATC) method to achieve the DSCUC. However, this method also requires the system to be partitioned into several sub-areas.

In sum, the existing research always relies on the area partitioning concept to separate the global ACOPF problem into smaller and decoupled optimization problems. But, as the centralized structure still exists inside the sub-areas, these methods are not fully distributed methods in which the advantages, such as the robustness of the communication, of the distributed system cannot be fully achieved.

Chapter 3. Preliminary Knowledge

3.1. Some Basic Knowledge from Graph Theory

In a distributed power system, the communication network can be modeled by an undirected graph $G = (V, E)$, where $V = \{v_1, v_2, \dots, v_n\}a$ is the set of vertices in the graph which denotes the n buses in the power system. $E = \{e_l = \{v_i, v_j\} \mid v_i, v_j \in V\}$ is a special subset of $V \times V$ which represents the set of m edges between any two linked vertices in G. The two linked vertices are called "neighbors" to each other.

There are four important matrices from graph theory that will be useful in this research. They are A (adjacency matrix), D (degree matrix), I (incidence matrix), and L (Laplacian). The adjacency matrix $A = [a_{ij}]$ for an undirected graph is a $n \times n$ symmetric matrix which describes the adjacency relationships of the vertices in the graph. The element a_{ij} at the ith row and jth column of the matrix A is 1 if there is an edge that connects the vertex v_i to $v_j (i \neq j)$, i.e. v_j is a neighbor of v_i. Otherwise, a_{ij} is zero. Note that the diagonal elements of the adjacency matrix are all zeros since there is no edge to connect a vertex with itself.

In a graph, the set of neighbor vertices of a certain vertex v_i is defined as the neighborhood set $\mathcal{N}_i$ of vertex v_i, and the cardinality of the neighborhood set $\mathcal{N}_i$ is called the degree $d(v_i)$ of vertex v_i, which represents the number of vertices connected to vertex v_i. Matrix $D = [d(v_i)]$ is called the degree matrix, in which the elements of the matrix are

the degrees of the vertices in the graph. Contrary to the adjacency matrix, the degree matrix only has non-zero elements on its diagonal. That is $D = Diag([d(v_1), d(v_2), \ldots, d(v_n)])$.

I is the incidence matrix of a graph. It is a $n \times m$ matrix and its elements are defined as

$$[I]_{ij} = \begin{cases} -1, & if\ v_i\ is\ the\ tail\ of\ e_j \\ 1, & if\ v_i\ is\ the\ head\ of\ e_j \\ 0, & otherwise \end{cases} \quad (1)$$

The incidence matrix represents the relationship between vertices and edges.

Laplacian L is one of the most important matrices in graph theory, it is related to the dynamic characteristic of the graph. The Laplacian L can be defined by

$$L = D - A = II^T \quad (2)$$

According to equation (2), Laplacian L is symmetrical as the degree matrix D and the adjacency matrix A are both symmetrical. As discussed before, the ith row of the adjacency matrix A describes the connections of vertex v_i to its neighbors by ones. Then, the sum of the ith row of A is the number of neighbors of v_i. The row sum of the degree matrix D is also the number of neighbors since all the elements in a row are zeros except the diagonal element which is the cardinality of the neighborhood set $\mathcal{N}_i$. Therefore, the row sum of Laplacian is equal to zero, and since Laplacian is symmetrical, the column sum is also zero.

3.2. Consensus Protocol

Base on the above knowledge, a consensus protocol can be designed for a network system. The consensus protocol can be used for several important applications of power systems. For example, suppose the generators (vertices) in a power system (an undirected

graph G) need to calculate (or estimate) a quantity such as the desired incremental cost of the generators. The consensus protocol can help the generators to reach a consensus about the quantity.

The consensus protocol for vertex v_i can be designed as the update formula form in discrete-time as

$$x_i(k) = x_i(k-1) + \tau \sum_{j \in \mathcal{N}_i} w_{ij}\left[x_j(k-1) - x_i(k-1)\right] \tag{3}$$

where $x_i(k)$ is the ith vertex's state at iteration k. τ is the time interval between two update iterations. $\mathcal{N}_i$ is the neighbors of vertex i. w_{ij} is a weight coefficient on the link between v_i and v_j. It can also be expressed as w_l where l is from edge $e_l = \{v_i, v_j\}$.

Let $X(k) = [x_1(k), x_2(k), x_3(k), \dots, x_n(k)]^T$ denotes the states of vertices in the graph at iteration k. The equation (3) can be rewritten as matrix form (4) by considering all vertices in the graph.

$$\boldsymbol{X}(k) = \boldsymbol{X}(k-1) - \tau L_w \boldsymbol{X}(k-1) \tag{4}$$

where $L_w = IWI^T$ is the weighted Laplacian, and $W = Diag([w_1, w_2, \dots, w_m])$ is the weight matrix. Equation (4) follows Lemma 1 [5].

Lemma 1: The vector $\boldsymbol{X}(k)$ is convergent as:

$$\lim_{k \to \infty} X(k) = \left(\frac{1}{n}\sum_{i=1}^{n} x_i(0)\right)\mathbf{1} = \frac{1}{n}\mathbf{1}\mathbf{1}^{\mathrm{T}} X(0) \tag{5}$$

if and only if the graph is connected and

$$\rho(L_w) < \frac{2}{\tau} \tag{6}$$

where "the graph is connected" means that, for every two vertices in the graph, there is at least one path between them. In other words, there are no isolated parts in the graph. $x_i(0)$ is the initial value of x_i. $\mathbf{1}$ is a $n \times 1$ vector which all elements are 1, i.e. $\mathbf{1} = [1,1,\dots,1]^T$. $\rho(L_w)$ is the maximum eigenvalue (in absolute value) of L_w.

Therefore, the stability of the consensus protocol (4) in a connected graph can be guaranteed by tuning the weight coefficient w_k on each vertex to force the weighted Laplacian L_w to satisfy the Lemma 1.

3.3. Centralized Economic Dispatch

The objective of the ED problem is to calculate the power generation for each generator so that the entire power system works in the most efficient condition. By neglecting the transmission losses, the ED problem is an optimization problem which to minimize the following objective function,

$$\text{Minimize } F(P_G) = \sum_{i=1}^{n} F_i(P_i) \tag{7}$$

$$s.t. \quad P_D = \sum_{i=1}^{n} P_i \tag{8}$$

$$P_i^{min} \le P_i \le P_i^{max} \tag{9}$$

where P_i is the ith generator's power production. $F_i(P_i)$ is the cost function of the ith generator. P_D is the total demand of the entire power system, which includes loads and losses, i.e., $P_D = P_{loss} + P_{load}$. P_i^{min} and P_i^{max} are the lower and upper active power output limits of the generator i, respectively. Since the losses are very small compared to the loads, the losses can be neglected in some ED problems. The constraint (8) ensures the

power balance between the load and generation. Meanwhile, constraint (9) limits the output of each generator in its permitted range. Typically, the generation cost function in (7) is quadratic as

$$F_i(P_i) = \alpha_i + \beta_i P_i + \gamma_i P_i^2 \tag{10}$$

where α_i, β_i and γ_i are the coefficients of the cost function for the ith generator.

In the traditional centralized power system, the ED problem can be solved by the Lagrange multiplier method. By introducing the Lagrange multiplier, the solution of the ED problem can be obtained by solving the following equations

$$\frac{dF_i(P_i)}{dP_i} = \lambda \quad \text{, for } i = 1, 2, \dots, n \tag{11}$$

Therefore, the Lagrange multiplier equals the incremental cost of each generator. Now, considering the cost function (7) and the constraints (8) and (9), the solution of the objective function can be written as

$$P_i^*(k) = \begin{cases} \dfrac{\lambda - \beta_i}{2\gamma_i} & \text{, if } P_i^{min} < P_i^*(k) < P_i^{max} \\ P_i^{min} & \text{, if } P_i^*(k) \le P_i^{min} \\ P_i^{max} & \text{, if } P_i^*(k) \ge P_i^{max} \end{cases} \tag{12}$$

where $P_i^*(k)$ is the solution of the ED problem at time k.

If the line losses are specifically considered, the power balance constraint (8) becomes

$$\sum_{i=1}^{n} P_i = P_{load} + P_{loss} \tag{13}$$

and the necessary conditions (11) for the solution of the ED problem is given by

$$\frac{dF_i(P_i)}{dP_i} = \lambda\left(1 - \frac{\partial P_{loss}}{\partial P_i}\right), \text{ for } i = 1, 2, \ldots, n \quad (14)$$

where $\partial P_{loss}/\partial P_i$ is called incremental loss.

The above method is obvious a centralized approach since the single Lagrange multiplier λ should be computed for all generators. Also, the total power supply $\sum_{i=1}^{n} P_i$ in (8) or (13) is determined at the centralized control center by collecting the generation data from all generators.

3.4. Centralized State Estimation

The Weighted Least-Squares (WLS) is a commonly used method in power system state estimation [6]. Assuming that there is a measurement device to detect a state θ in the system, the mathematical model of the measurement is

$$z = H\theta + \eta \quad (15)$$

where z is the reading on the instrument. H is the observation matrix that represents the characteristic of the power system. η is the noise or error in the measurement. Then, the WLS method in (16) can be used to estimate the state θ with the noisy measurement z.

$$\hat{\theta} = (H^{\mathrm{T}}R^{-1}H)^{-1}H^{\mathrm{T}}R^{-1}z \quad (16)$$

where, $\hat{\theta}$ is the estimation of the state θ. R is the covariance matrix of measurement errors. For a power system, the WLS state estimation is employed to estimate the voltage magnitudes and phase angles on different buses by the noisy measurements of the real power, reactive power, current, transformer tap position, and voltage magnitude [6].

In this research, the work is focused on state estimation using DC power flow analysis. So, the measurements are the real power on each transmission line and the states are the phase angles on different buses. In addition, the estimation can be used to improve

the accuracy of the measurements. By inputting the estimated state $\hat{\theta}$ back to the measurement equation without noise as in (17), the better estimation $\hat{z}$ of the measurement values can be obtained.

$$\hat{z} = H\hat{\theta} \tag{17}$$

3.5. Distributed State Estimation over Sensor Networks

In a distributed network, numerous meters (or sensors) are installed to measure the same quantity in the system. By connecting the meters together to form a sensor network, the redundant measurement can be used to improve the accuracy of state estimation and measurement results.

In order to realize the DSE, two consensus protocols are required [5],

$$P_i(k+1) = P_i(k) + \tau \sum_{j \in \mathcal{N}_i} w_{ij}^P \left(P_j(k) - P_i(k) \right) \tag{18}$$

$$\Theta_i(k+1) = \Theta_i(k) + \tau \sum_{j \in \mathcal{N}_i} w_{ij}^\Theta \left(\Theta_j(k) - \Theta_i(k) \right) \tag{19}$$

where, $P_i(0) = H_i^T R_i^{-1} H_i$ and $\Theta_i(0) = H_i^T R_i^{-1} z_i$. w_{ij}^P and w_{ij}^Θ are the weights for graph $G(P)$ and $G(\Theta)$, respectively. According to Lemma 1, $P_i(k)$ and $\Theta_i(k)$ converge to values,

$$\lim_{k \to \infty} P_i(k) = \frac{1}{n} \sum_{i=1}^{n} H_i^T R_i^{-1} H_i \tag{20}$$

$$\lim_{k \to \infty} \Theta_i(k) = \frac{1}{n} \sum_{i=1}^{n} H_i^T R_i^{-1} z_i \tag{21}$$

Therefore, by implementing the equations (18) and (19) on each node, the same result of the centralized WLS state estimation [5] can be achieved by (22) on the ith node in a distributed fashion,

$$\hat{\theta}_i(k) = P_i(k)^{-1}\Theta_i(k) \tag{22}$$

Since $\lim_{k\to\infty} \hat{\theta}(k) = \hat{\theta}$.

3.6. Optimal Power Flow Problem

The OPF (optimal power flow) problem is an advanced version of economic dispatch in which it also aims to minimize the cost of the power system meanwhile it does not only consider the power balance between the total load and total generation in the power system but also take the power flow model in consideration to ensure the power is balanced on each bus. In addition, the OPF problem usually has more constraints such as the line flow constraints.

As mentioned above, the OPF problem is an advanced economic dispatch problem, so it has the same objective function (7) and constraints (8)-(9). In addition, the following constraints are also included in the OPF problem,

$$P_i^{D,meas} = P_i^G - \sum_{i_g=1}^{n_g} P_{ij} \tag{23}$$

$$Q_i^{D,meas} = Q_i^G - \sum_{i_g=1}^{n_g} Q_{ij} \tag{24}$$

$$\left|P_{ij}\right| \le P_{ij}^{max} \tag{25}$$

$$\left|Q_{ij}\right| \le Q_{ij}^{max} \tag{26}$$

Where (23)-(24) are node power balance constraints. Meanwhile, (25)-(26) are called line flow limit constraints; $P_i^{D,meas}$ and $Q_i^{D,meas}$ are the active and reactive demands measured on the bus i, respectively.

Chapter 4. Proposed Methods

Based on the preliminary knowledge above, the proposed distributed algorithms for economic dispatch, state estimation, and optimal power flow will be introduced. In the first two sections, the distributed economic dispatch and distributed state estimation based on the DC power flow model will be developed. Then, the distributed optimal power flow method will be established on the basis of the first two distributed methods. In the last two sections, the distributed AC state estimation and distributed ACOPF methods will be presented.

4.1. Distributed Economic Dispatch

According to the existing research discussed in Chapter 2, the challenge for the consensus-based distributed economic dispatch is to obtain the power mismatch information in a distributed system and thus satisfy the power balance constraint in the DED problem. However, this problem can be easily solved by control methods. For example, the frequency controller in the peer-to-peer control method [40] is designed to balance the active power in the power system. Based on this concept, the PI controller and Neural Network (NN) controller are employed to develop the DED method.

4.1.1. DED Control with PI controller

Based on the consensus protocol (3), the Lagrange multipliers λ's can be treated as the states of vertices in the graph, so they can be equalized by the consensus protocol. In addition, the power balance constraint (8) should be satisfied, so the peer-to-peer frequency

control method [40] can be introduced by adding a goal attraction function [41] in the consensus protocol to balance the power in the system. Therefore, a modified consensus protocol having the form (27) can be adopted to develop the DED control method.

$$\lambda_i(k) = \lambda_i(k-1) + \tau \sum_{j \in \mathcal{N}_i} w_{ij}\left[\lambda_j(k-1) - \lambda_i(k-1)\right] + G_i^*(k) \tag{27}$$

where $G_i^*(k)$ is a goal attraction function [41] to force $\lambda(k)$ to reach a certain consensus, which is the optimal solution for ED problem, rather than the naive "average value" solution as in (5). The goal attraction function is a function of the distance between the desired λ^* and the current λ, and the function becomes zero when the goal is reached.

Now, the objective is to create the proper goal attraction function $G_i^*(k)$ based on the peer-to-peer frequency control method so that the desired solution of ED problem can be obtained. Typically, in a distributed system, the Lagrange multipliers $\lambda_i (i = 1,2, \dots, n_b)$ converges to an identical value in a few iterations (k_c) with the consensus protocol. Therefore, the consensus term $\tau \sum_{j \in \mathcal{N}_i} w_{ij}\left[\lambda_j(k-1) - \lambda_i(k-1)\right]$ in (27) can be neglected for time $k > k_c$, and k_c is usually a small number. Then, the lambda-iteration equation (27) becomes,

$$\lambda_i(k) = \lambda_i(k-1) + G_i^*(k) \tag{28}$$

In equation (28), the term $G_i^*(k)$ acts like the change of state $\Delta\lambda_i$ in an iteration equation. So, let $G_i^*(k) = \Delta\lambda_i$. According to equation (11) and Euler approximation, the change of state is

$$G_i^*(k) = \Delta\lambda_i(k) = \frac{d\lambda}{dP_i}\Delta P_i(k) = 2\gamma_i \Delta P_i(k) \tag{29}$$

where $\Delta P_i(k)$ is the change of active power output of the ith generator. In order to balance the load and generation in the power system, the frequency controller is used for each generator to calculate the $P_i(k)$. Since the iteration is adopted in a discrete manner, the discrete PI controller can be utilized for the frequency controller. Then, the proposed DED control scheme is,

$$\lambda_i(k) = \lambda_i(k-1) + \tau \sum_{j \in \mathcal{N}_i} w_{ij}\left[\lambda_j(k-1) - \lambda_i(k-1)\right] + \Delta\lambda_i(k) \tag{30}$$

$$\Delta\lambda_i(k) = 2\gamma_i\left[Kp_i\left(\Delta f(k) - \Delta f(k-1)\right) + \tau Ki_i \Delta f(k)\right] \tag{31}$$

$$\Delta f(k) = f_{rated} - f(k) \tag{32}$$

where $\Delta f(k)$ is the frequency deviation at iteration k. $f(k)$ is the frequency of the power system at iteration k which can be measured by the Phase-Locked Loop (PLL) on each generator. f_{rated} is the rated frequency that can be reached when the load and generation are balanced under the rated operating condition. Kp_i is the proportional gain and Ki_i is the integral gain. After $\lambda_i(k)$ is obtained, the optimal power output reference $P_i^*(k)$ for the ith generator can be calculated by letting $\lambda = \lambda_i(k)$ in equation (12).

Note that, if there is no communication failure, it is not necessary for all generators to measure the frequency $f(k)$. The control scheme is still applicable even if the frequency measurement is only available for a small subset of all generators. For the generators without PLL, called followers [42], their goal attraction terms $\lambda_i(k)$ in the control algorithm (17) can be removed. The reason is that, due to the consensus protocol, the followers are attracted by the leaders, i.e., the generators with frequency measurement ability, to reach the optimal solution.

The proposed DED control algorithm is summarized in Algorithm 1.

Algorithm 1 The Algorithm on the ith Generator

1: Initialize $\lambda_i(0)$ with an arbitrary value.
2: Set parameters w_{ij}, Kp_i and Ki_i.
3: **for** $k > 0$ **do**
4: Measure the frequency $f_i(k)$ in the grid
5: Receive the Lagrange multipliers $\{\lambda_j(k-1) \mid j \in \mathcal{N}\}$ from neighbor nodes
6: Send the local Lagrange multiplier $\lambda_i(k-1)$ to the neighbor nodes.
7: Calculate the new Lagrange multiplier $\lambda_i(k)$ by (30)-(32).
8: Calculate the generator reference $P_i(k)$ by (12).
9: Set the output reference of the generator by $P_i^*(k)$.
10: Let $k = k + 1$.
11: end for

Now, let's discuss the convergence of the proposed method. By the control objective, the Lagrange multipliers of all generators should be identical to achieve the economic dispatch. So, in the following paragraphs of this section, the convergence of the Lagrange multipliers under the proposed algorithm will be discussed, i.e., $\lambda_1(k) = \lambda_2(k) = \cdots = \lambda_{n_b}(k)$.

According to equation (4), the consensus protocol (30) can be written into the vector form as

$$\boldsymbol{\lambda}(k) = \boldsymbol{\lambda}(k-1) - \tau L_w \boldsymbol{\lambda}(k-1) + \Delta\boldsymbol{\lambda}(k) \tag{33}$$

where $\boldsymbol{\lambda}(k) = \left[\lambda_1(k), \lambda_2(k), \dots, \lambda_{n_b}(k)\right]^{\mathrm{T}}$ is the vector of Lagrange multipliers of the generators. Similarly, $\Delta\lambda(k) = \left[\Delta\lambda_1(k), \Delta\lambda_2(k), \dots, \Delta\lambda_{n_b}(k)\right]^{\mathrm{T}}$. Let $M_w = I - \tau L_w$, the equation (20) becomes

$$\boldsymbol{\lambda}(k) = M_w \boldsymbol{\lambda}(k-1) + \Delta\boldsymbol{\lambda}(k) \tag{34}$$

which can be expanded as

$$\boldsymbol{\lambda}(k) = (M_w)^k\boldsymbol{\lambda}(0) + (M_w)^{k-1}\boldsymbol{\lambda}(1) + \cdots + M_w\boldsymbol{\lambda}(k-1) + \Delta\boldsymbol{\lambda}(k) \tag{35}$$

The vectors $\Delta\boldsymbol{\lambda}(\kappa)$ for $\kappa = 1,2,\ldots,k-1$ can be written into two terms by letting $\Delta\boldsymbol{\lambda}(\kappa) = \overline{\Delta\boldsymbol{\lambda}}(\kappa) + \boldsymbol{\delta}_\lambda(\kappa)$, where $\overline{\Delta\boldsymbol{\lambda}}(\kappa) = \mathbf{1}\mathbf{1}^T\Delta\lambda(\kappa)/n_b$ is the average vector of $\Delta\boldsymbol{\lambda}(\kappa)$, in which the entries of this average vector are all the same. $\boldsymbol{\delta}_\lambda(\kappa)$ is the difference between the average vector $\overline{\Delta\boldsymbol{\lambda}}(\kappa)$ and original vector $\Delta\boldsymbol{\lambda}(\kappa)$. Then, the equation (35) becomes

$$\boldsymbol{\lambda}(k) = (M_w)^k\boldsymbol{\lambda}(0) + (M_w)^{k-1}\overline{\Delta\boldsymbol{\lambda}}(1) + \cdots + M_w\overline{\Delta\boldsymbol{\lambda}}(k-1) + (M_w)^{k-1}\boldsymbol{\delta}_\lambda(1) + \cdots + M_w\boldsymbol{\delta}_\lambda(k-1) + \Delta\boldsymbol{\lambda}(k) \tag{36}$$

Here, a lemma about the matrix M_w is introduced:

Lemma 2: If the graph is connected and $\rho(L_w) < 2/\tau$, then

$$(M_w)^\infty \coloneqq \lim_{k\to\infty}(M_w)^k = \frac{1}{n}\mathbf{1}\mathbf{1}^{\mathrm{T}} \tag{37}$$

the proof can be found in [5].

By Lemma 2, the average vector $\overline{\Delta\boldsymbol{\lambda}}(\kappa)$ can be represented as $\overline{\Delta\boldsymbol{\lambda}}(\kappa) = (1/n)\mathbf{1}\mathbf{1}^T\Delta\boldsymbol{\lambda}(\kappa) = (M_w)^\infty\Delta\boldsymbol{\lambda}(\kappa)$, and thus $(M_w)^{k'}\overline{\Delta\boldsymbol{\lambda}}(\kappa) = (M_w)^{k'}(M_w)^\infty\Delta\lambda(\kappa) = (M_w)^{k'+\infty}\Delta\lambda(\kappa) = (M_w)^\infty\Delta\lambda(\kappa) = \overline{\Delta\lambda}(\kappa)$. So, the equation (36) becomes

$$\boldsymbol{\lambda}(k) = (M_w)^k\boldsymbol{\lambda}(0) + \overline{\Delta\boldsymbol{\lambda}}(1) + \cdots + \overline{\Delta\boldsymbol{\lambda}}(k-1) + \overline{\Delta\boldsymbol{\lambda}}(k) + (M_w)^{k-1}\boldsymbol{\delta}_\lambda(1) + \cdots + M_w\boldsymbol{\delta}_\lambda(k-1) + \boldsymbol{\delta}_\lambda(k) \tag{38}$$

This equation shows that the inequality between the entries of the vector$\boldsymbol{\lambda}(k)$, i.e., the Lagrange multipliers of all nodes, is caused by the terms $(M_w)^k\boldsymbol{\lambda}(0)$ and $(M_w)^{k-\kappa}\boldsymbol{\delta}_\lambda(\kappa)$ for all κ. However, by Lemma 2, the term $(M_w)^k\boldsymbol{\lambda}(0)$ converges to $\overline{\boldsymbol{\lambda}}(\kappa) = (1/n_b)\mathbf{1}\mathbf{1}^T\boldsymbol{\lambda}(\kappa)$. So, the differences between the Lagrange multipliers is mainly

caused by the terms $(M_w)^{k-\kappa}\boldsymbol{\delta}_\lambda(\kappa)$. Therefore, to prove the convergence of the proposed algorithm, the convergence of the summation $\sum_{\kappa=1}^{k}(M_w)^{k-\kappa}\boldsymbol{\delta}_\lambda(\kappa)$ should be proven.

According to equation (31), assuming that the measured frequency is the same for all nodes, the difference between the controller coefficients of the generators is the reason to make the goal attraction term $\Delta\boldsymbol{\lambda}(k)$ for each node distinct, i.e.,$\boldsymbol{\delta}_\lambda(k) \neq 0$. Also, this difference can be amplified by the frequency variation, i.e., the terms $\Delta f(k) - \Delta f(k-1)$ and $\Delta f(k)$. Since the controller coefficients are fixed during the operation, and the frequency deviation cannot be arbitrarily large due to the dynamic of the power system, then it is reasonable to assume that the term $\boldsymbol{\delta}_\lambda(k)$ has an upper bound $\boldsymbol{\delta}_\lambda$ for some norms, i.e., $\|\boldsymbol{\delta}_\lambda\| \geq \|\boldsymbol{\delta}_\lambda(k)\|$ for $k = 1, 2, \ldots$. Assuming that the term $\boldsymbol{\delta}_\lambda(k) = \boldsymbol{\delta}_\lambda$ for all time, if the proposed algorithm can still converge under this extreme condition, it means that the algorithm can always converge, thus the convergence of the algorithm is proven. So, let $\boldsymbol{\delta}_\lambda(k) = \boldsymbol{\delta}_\lambda$ for $k = 1, 2, \ldots$ and define a new time variable $\kappa' = k - \kappa$, then the summation $\sum_{\kappa=1}^{k}(M_w)^{k-\kappa}\boldsymbol{\delta}_\lambda(\kappa)$ becomes a power series when time goes to infinity,

$$\sum_{\kappa'=0}^{\infty}(M_w)^{\kappa'}\boldsymbol{\delta}_\lambda = \boldsymbol{\delta}_\lambda + M_w\boldsymbol{\delta}_\lambda + (M_w)^2\boldsymbol{\delta}_\lambda + \cdots \tag{39}$$

In order to converge the term $\sum_{\kappa'=0}^{\infty}(M_w)^{\kappa'}$ in equation (39), there must be a matrix norm $\|*\|$ such that the numerical series $\sum_{\kappa'=0}^{\infty}\|M_w\|^{\kappa'}$ converges [43]. Since the series $\sum_{\kappa'=0}^{\infty}\|M_w\|^{\kappa'}$ is a geometric series, the condition for its convergence can be easily satisfied by tuning the weights of the weighted Laplacian L_w to make $\|M_w\|^{\kappa'} < 1$. Then, the series (39) converges to

$$\sum_{\kappa'=0}^{\infty}(M_w)^{\kappa'}\boldsymbol{\delta}_\lambda = S_w\boldsymbol{\delta}_\lambda \tag{40}$$

where $S_w = \sum_{\kappa'=0}^{\infty}(M_w)^{\kappa'}$. This result cannot ensure Lagrange multipliers to be identical since the entries of $S_w\boldsymbol{\delta}_\lambda$ may be distinct. This is reasonable since the difference vector $\boldsymbol{\delta}_\lambda$ is constantly added into the equation (38) to create a new difference between the Lagrange multipliers. But this result gives the upper bound for the difference. So, the proposed method is not divergent. In addition, since the frequency deviation in the power system can be eliminated by the frequency controller in the proposed algorithm, then the goal attraction term $\Delta\boldsymbol{\lambda}(k)$ decreases to zero, and thus the difference vector $\boldsymbol{\delta}_\lambda(k)$. So, the consensus can be achieved eventually.

4.1.2. DED Control with NN Controller

As shown in the previous section, the lambda-iteration equation (27) becomes (28) after the consensus term converges. So, control methods in discrete form can be integrated with the consensus protocol. In the previous section, the PI frequency controller is integrated with the consensus protocol to obtain the DED control method. In this section, a different controller, the neural network controller, is employed to integrate with the consensus protocol to develop a new distributed optimal control scheme. So, it can show that the proposed scheme can be widely applied for different control purposes by integrating different controllers.

In this work, a two-layer neural network is employed to integrate with the consensus protocol, and it is easy to expand the neural network with more layers if needed. The two-layer neural network includes one hidden layer and one output layer. The hidden

layer has 5 neurons, and the output layer has 1 neuron. The inputs of the neural network are the frequency deviation $\Delta f(k)$ and the second-order frequency deviation $\Delta^2 f(k) = \Delta f(k) - \Delta f(k-1)$. The output of the neural network is the change of Lagrange multiplier $\Delta\lambda_i$. The structure of the two-layer neural network is shown in Figure 4-1.

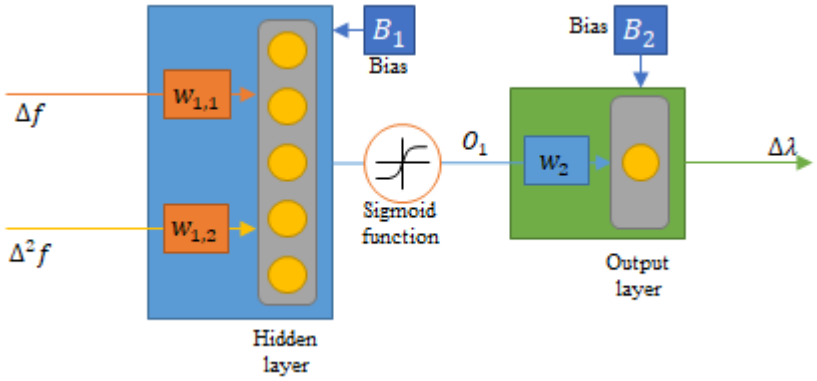


Figure 4-1 The structure of the neural network

According to the neural network structure, the proposed algorithm is the following,

$$\lambda_i(k) = \lambda_i(k-1) + \tau \sum_{j \in \mathcal{N}_i} w_{ij}\left[\lambda_j(k-1) - \lambda_i(k-1)\right] + \Delta\lambda_i(k) \tag{41}$$

$$\Delta\lambda_i(k) = \boldsymbol{O}_1^T \boldsymbol{W}_2 + B_2 \tag{42}$$

$$\boldsymbol{O}_1 = \mathrm{Sigm}\left(\Delta f(k)\boldsymbol{W}_{1,1} + \Delta^2 f(k)\boldsymbol{W}_{1,2} + \boldsymbol{B}_1\right) \tag{43}$$

$$\Delta f(k) = f_{rated} - f(k) \tag{44}$$

$$\Delta^2 f(k) = \Delta f(k) - \Delta f(k-1) \tag{45}$$

where $\boldsymbol{W}_{1,1}$, $\boldsymbol{W}_{1,2}$ and $\boldsymbol{W}_2$ are the weight vectors of the hidden layer and the output layer; $\boldsymbol{B}_1$ and B_2 are the biases of the two layers; $\boldsymbol{O}_1$ is the outputs of the hidden layer. The sigmoid function $Sigm(*)$ has the form in (33).

$$Sigm(x) = \frac{2}{1 + e^{-2x}} - 1 \tag{46}$$

This sigmoid function maps the outputs of the hidden layer $\boldsymbol{O}_1$ into the range $(-1,1)$. If the input of the sigmoid function is a vector, the function is performed componentwise.

4.1.3. Communication-failure-tolerant DED Control method

According to the consensus protocol, the communication between the controllers in the power system is required to converge their Lagrange multipliers to the consensus. However, if there is a communication failure, i.e. the communication network is disconnected into two or more parts due to faults, in the system, the consensus may not be able to be maintained for the basic DED control method proposed in the previous sections. By the equation (30), the proposed DED algorithm includes two major parts: the consensus term $\tau \sum_{j\in\mathcal{N}_i} w_{ij}\left(\lambda_j(k-1)-\lambda_i(k-1)\right)$ and the goal attraction function $\Delta\lambda_i(k)$. According to equation (31), the values of the goal attraction functions in different nodes are distinct due to the distinct parameters. Normally, these differences can be eliminated by the consensus term. However, during the communication failure, the information of the Lagrange multipliers cannot be transmitted through the consensus term, so the differences of the goal attraction functions cannot be removed. This is why the communication failure makes the generators fail to maintain consensus, i.e., the economic dispatch. So, in this section, the communication-failure-tolerant DED control method is proposed to solve this problem and make the system stay at the economic dispatch during the communication failure. The proposed communication-failure-tolerant DED approach mainly includes two parts: Controller gain design and leader-follower mode.

Controller gain design is a method to set the parameters Kp_i and Ki_i of the PI controllers so that the consensus can be maintained during the communication failure. As discussed above, the errors introduced by the goal attraction functions cause the Lagrange multipliers to fail to reach the consensus, and the errors are generated by the differences in the control parameters of the nodes. So, by carefully designing the control parameters of the controllers to make their goal attraction terms $\Delta\lambda_i(k)$ identical, the errors can be removed so that the consensus can be maintained. In equation (36), if $\Delta\boldsymbol{\lambda}(\kappa) = \overline{\Delta\boldsymbol{\lambda}}(\kappa)$ for any κ, which indicates $\boldsymbol{\delta}_\lambda(\kappa) = 0$, the equation (38) becomes

$$\boldsymbol{\lambda}(k) = (M_w)^k\boldsymbol{\lambda}(0) + \Delta\boldsymbol{\lambda}(1) + \ \ldots\ + \Delta\boldsymbol{\lambda}(k-1) + \Delta\boldsymbol{\lambda}(k) \tag{47}$$

This is also true if $\Delta\boldsymbol{\lambda}(k)$ is a vector with identical elements. In this situation, the residual term $\boldsymbol{\sigma} = \mathbf{0}$ and $\boldsymbol{\lambda}(k) = \overline{\boldsymbol{\lambda}(k)}$ for $k \geq k_c$, even if the communication network fails. So, by adjusting the coefficients Kp_i and Ki_i in equation (31) to satisfy the equation $\Delta\boldsymbol{\lambda}(k) = \overline{\Delta\boldsymbol{\lambda}}(k)$, the proposed control scheme can be more robust.

The elements in $\overline{\Delta\boldsymbol{\lambda}}(k)$ are

$$\begin{aligned}\overline{\Delta\lambda_i}(k) = \frac{2}{n_b}\left(\sum_{i=1}^{n_b}\gamma_i Kp_i\right)\big(\Delta f(k) - \Delta f(k-1)\big) \\ + \tau\frac{2}{n_b}\left(\sum_{i=1}^{n_b}\gamma_i Ki_i\right)\Delta f(k)\end{aligned} \tag{48}$$

compared with (31), for $\boldsymbol{\lambda}(k) = \overline{\boldsymbol{\lambda}(k)}$, the new coefficients Kp_i^* and Ki_i^* can be set as

$$Kp_i^* = \frac{1}{n_b\gamma_i}\sum_{i=1}^{n_b}\gamma_i Kp_i \tag{49}$$

$$Ki_i^* = \frac{1}{n_b \gamma_i} \sum_{i=1}^{n_b} \gamma_i Ki_i \tag{50}$$

With the controller gain design method, the consensus can be maintained during the communication failure.

However, if one or more nodes in the system cannot measure the frequency due to the cost or faults, then their goal attraction terms $\Delta\lambda_i(k)$ cannot be calculated. Thus, the controller gain design method cannot be directly applied. To solve this problem, the leader-follower mode is proposed. In this mode, the controllers that can measure the frequency are called "leaders", and the controllers that have no frequency measurement ability are called "followers". The idea of the leader-follower mode is to make the communication between the leaders and the followers into one-way so that the lack of goal attraction function in followers will not affect the leaders maintaining the consensus with the controller gain design method, and the followers can "trace" the leaders to reach the consensus through the consensus term. According to this idea, the algorithm of the leaders is in the equation (51),

$$\lambda_i(k) = \lambda_i(k-1) + \tau \sum_{j \in \mathcal{N}_i^L} w_{ij}\left[\lambda_j(k-1) - \lambda_i(k-1)\right] + \Delta\lambda_i(k) \tag{51}$$

where $\mathcal{N}_i^L$ is the set of the leader nodes. The goal attraction term $\Delta\lambda_i(k)$ can be calculated by the PI controller method (31) with the controller gain design or the NN controller (42) with the same weights and biases for all nodes. The difference between the algorithm of the leaders in (51) and the algorithm of the basic DED control method in (30) is that the leader nodes only communicate with other leaders but not the followers, so the error in the followers will not affect the leaders.

Then, the algorithm of the followers is in the equation (52),

$$\lambda_i(k) = \lambda_i(k-1) + \tau \sum_{j \in \mathcal{N}_i} w_{ij}\left[\lambda_j(k-1) - \lambda_i(k-1)\right] \tag{52}$$

in which the equation is the same as the equation (30) but has no goal attraction term $\Delta\lambda_i(k)$ due to the lack of frequency measurement. Since the followers can receive the information of the Lagrange multipliers from the leaders through one-way communication, they can be driven to the consensus by the consensus term.

4.1.4. Line Losses Constraint

In the ED problem, if the line loss model is considered, the line losses constraint (13) should be satisfied. The line losses of the power system can be modeled by the following function [40],

$$P_{loss} = L_1 P_1^2 + L_2 P_2^2 + \cdots + L_{n_b} P_{n_b}^2 \tag{53}$$

where $L_i (i = 1,2, \ldots n_b)$ are the parameters of the line losses model.

According to the necessary conditions with line losses constraint (14), the solution of the ED problem is

$$P_i^*(k) = \frac{\lambda_i(k) - \beta_i}{2\gamma_i + 2L_i\lambda_i(k)} \tag{54}$$

where $P_i^*(k)$ is the ED solution for the ith generator at iteration k. Considering the output limits (9) of each generator, the optimal dispatch will be

$$P_i^*(k) = \begin{cases} \dfrac{\lambda_i(k) - \beta_i}{2\gamma_i + 2L_i\lambda_i(k)} & \text{, if } P_i^{min} < P_i^*(k) < P_i^{max} \\ P_i^{min} & \text{, if } P_i^*(k) \le P_i^{min} \\ P_i^{max} & \text{, if } P_i^*(k) \ge P_i^{max} \end{cases} \tag{55}$$

This solution is obtained locally at each generator and $P_i^*(k)$ can be set as the active power output reference of each generator i at time k.

4.2. Distributed State Estimation

According to the literature review, the existing distributed state estimation methods are not fully distributed. Most of them require area partitioning on the power system. To overcome this drawback, a fully distributed state estimation method without any centralized structure or facility (e.g. SCADA, local control center, or GPS) is proposed in this section so that the state estimation can be realized with only the distributed smart meters. Also, a novel information propagation algorithm is proposed as the basis of the proposed DSE method. The information propagation algorithm can help the smart meters to broadcast their local data to the entire system with a distributed communication network. The proof of the convergence of the proposed information propagation algorithm is provided. In addition, the distributed WLS method for the normal sensor network [5] and the DC power flow model [6] are adopted to develop the DSE method.

4.2.1. Problem Statement

This work will focus on the state estimation of phase angle based on the DC power flow model, since the calculation of DC power flow is much faster than AC power flow, and the accuracy is acceptable. So, in the algorithm, the measurements are the real power on each transmission line and the states to be estimated are the phase angles on different buses.

According to the distributed state estimation method for sensor networks provided in the preliminary section, equation (22) can be used to realize the DSE. However, the

method cannot be directly applied in power systems, since power systems are different from the traditional sensor network. For example, the power system shown in Figure 4-2 has four buses, in which bus 4 is set as the reference. There are four meters (M_{12}, M_{24}, M_{31} and M_{43}) to measure the real power on the transmission lines. The meters are connected by the communication lines as the dashed lines in Figure 4-2.

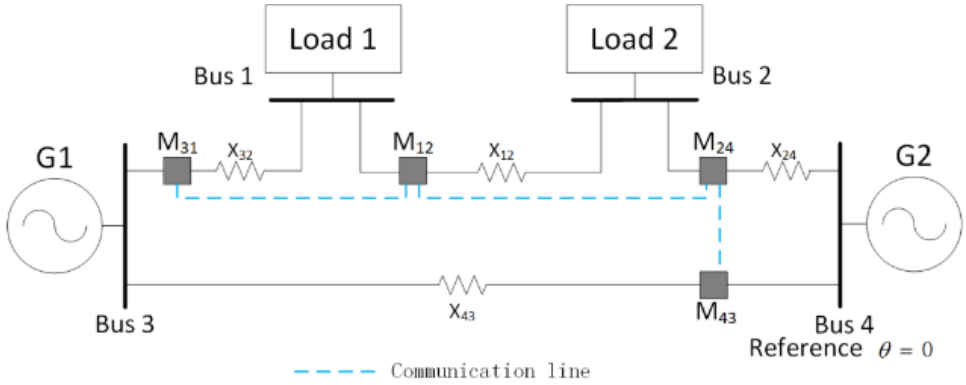


Figure 4-2 An example of a four-bus power system

According to the DC power flow, the measurement model in the centralized state estimation method is

$$\begin{bmatrix} M_{12} \\ M_{31} \\ M_{24} \\ M_{43} \end{bmatrix} = \begin{bmatrix} \frac{1}{X_{12}} & -\frac{1}{X_{12}} & 0 \\ -\frac{1}{X_{31}} & 0 & \frac{1}{X_{31}} \\ 0 & \frac{1}{X_{24}} & 0 \\ 0 & 0 & -\frac{1}{X_{43}} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} + \eta \tag{56}$$

However, in a decentralized system, since there is no centralized facility, the measurement information is incomplete for some nodes (meters). For example, in Figure 4-2, the meter M_{31} is only connected to the meter M_{12}, so the measured values M_{24} and M_{43} are unknown for M_{31}. Then, the measurement model at M_{31} is

$$\begin{bmatrix} M_{12} \\ M_{31} \end{bmatrix} = \begin{bmatrix} \frac{1}{X_{12}} & -\frac{1}{X_{12}} & 0 \\ -\frac{1}{X_{31}} & 0 & \frac{1}{X_{31}} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} + \eta \tag{57}$$

It is easy to verify that $P_{31} = H_{31}^T R_{31}^{-1} H_{31}$ in the equation (22) is singular. Thus the system is unobservable for the meter M_{31} [6], and the DSE cannot simply proceed because $P_i(k)$ is singular in (22). In order to solve this problem, an information propagation algorithm is proposed to broadcast the local measurement data of each node to the entire system, so that the system is observable for all nodes.

4.2.2. Information Propagation Algorithm

For a distributed system, in order to share information among the nodes in the system, the information propagation algorithm is proposed for the nodes to broadcast their local information (measurements) to the entire system.

Suppose that the vector $Z = [z_1; z_2; \dots; z_n]$ includes the local information z_i (the measurements on each meter in this case) of all n nodes in the system. The objective is to modify the consensus protocol (3) to estimate the vector Z, i.e. $lim_{t\to\infty} x_i(t) = Z$. Please note that this estimation is not the state estimation. So, $x_i(t) \in \mathbb{R}^n$ is an estimation vector in node i at time t, in which its jth entry $x_{i,j}(t)$ is the estimation for the jth node's measurement z_j.

The following consensus protocol serves as the proposed information algorithm,

$$\dot{x}_i(t) = I_i^0 \sum_{j\in\mathcal{N}_i} \Big(x_j(t) - x_i(t)\Big) \tag{58}$$

where the initial values $x_{i,j}(0)$ can be arbitrary if $i \neq j$, but $x_{i,i}(0) = z_i$ since the local information z_i is available from the measurement at the ith node; I_i^0 is a $n \times n$

diagonal matrix which is the same as the identity matrix but has a zero at the ith diagonal entry,

$$I_i^0 \triangleq Diag([1,1,\dots,1,0,1,\dots,1]) \tag{59}$$

With the information propagation algorithm (58), the estimation vector $x_i(t)$ converges to the local information vector Z, i.e. $lim_{t\to\infty}x_i(t) = Z$. The proof is given in the following section.

4.2.3. Proof of The Convergence of Information Propagation Algorithm

Here is a proof of the statement $lim_{t\to\infty}x_i(t) = Z$, where $x_i(t)$ is the estimation vector of the ith node, and $Z = [z_1; z_2; \dots; z_n]$ is the vector of local information of every node. The proof includes two parts: the first part proves the estimation vector $x_i(t)$ converges to an equilibrium point $x_i(\infty) \triangleq lim_{t\to\infty}x_i(t)$ under the algorithm (13), and the second part proves that the equilibrium point $x_i(\infty)$ is Z.

The estimation vector $x_i(t)$ converges to an equilibrium:

First, let us define a modified Laplacian L^* by the equation $L^* \triangleq [I_i^0]_n(L \otimes I)$, where $[I_i^0]_n \triangleq Diag([I_1^0, I_2^0, \dots, I_n^0])$, and I_1^0 is from (59). Then, the consensus protocol of the entire system can be written into the matrix form

$$\dot{X}(t) = -L^*X(t) = -[I_i^0]_n(L \otimes I)X(t) \tag{60}$$

where $\dot{X}(t)$ is the derivative of the vector $X(t)$; $X(t) = [x_1(t); x_2(t); \dots; x_n(t)]$ is a $n^2 \times 1$ vector. The symbol "$\otimes$" represents the Kronecker product.

The task is to prove the convergence of the consensus-based algorithm (60). Since (60) is a linear differential equation, the Gershgorin circle theorem [43] can be used to locate the eigenvalues of the matrix L^*.

Theorem 1 (Gershgorin theorem [43]) Suppose that there is a matrix $A = [a_{ij}] \in \mathbb{C}^{n\times n}$. Note that A is not the adjacency matrix A. Let $R_i^d = \sum_{j\neq i} |a_{ij}|$ be the sum of the absolute values of the off-diagonal entries in the ith row of A. Consider the Gershgorin discs $D(a_{ii}, R_i^d) = z \in \mathbb{C}: |z - a_{ii}| \leq R_i^d$ centered at $a_{i,i}$ with radius R_i^d. Then, all the eigenvalues of A are in the union of Gershgorin discs $G(A) = \cup_{i=1}^{n} D(a_{ii}, R_i^d)$.

Now, let us look at the Laplacian L. Since $L = D - A$, the center of each Gershgorin disc is the diagonal entry of the degree matrix D, i.e. the degree $d(v_i)$ of each vertex. And the radius R_i is the sum of the row entries of the adjacency matrix A. By the definition of $d(v_i)$, it is true that $R_i^d = d(v_i)$. Similarly, for the matrix $L \otimes I = D \otimes I - A \otimes I$, the centers and radiuses are determined by $D \otimes I$ and $A \otimes I$, respectively. Now, for the modified Laplacian $L^* = [I_i^0]_n(L \otimes I)$, the matrix $[I_i^0]_n$ can be considered as an operation to replace n rows of $L \otimes I$ by zeros. For these replaced n rows, the corresponding Gershgorin discs become circles at the origin with zero radii. Therefore, as an example, the Gershgorin discs of the modified Laplacian L^* have the form as in Figure 4-3.

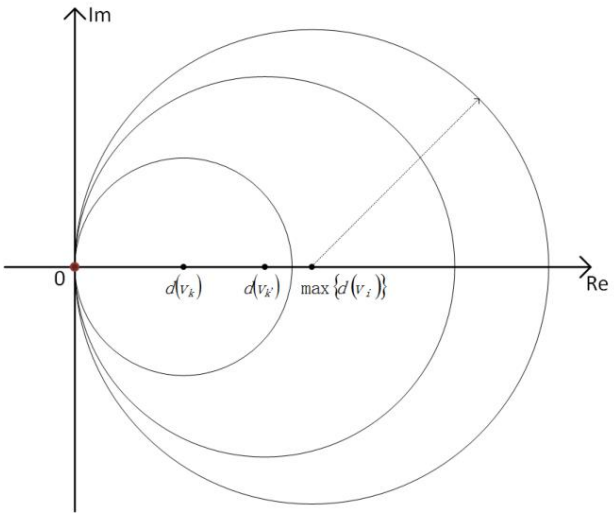


Figure 4-3 An example of the Gershgorin discs for the modified Laplacian L^*

In Figure 4-3, all the smaller Gershgorin discs (including the discs at the origin with zero radii) are contained within the largest Gershgorin discs. Thus, all the eigenvalues of L^* are located in the largest Gershgorin discs which are centered at $maxd(v_i)$ and has the radius $maxd(v_i)$, where $maxd(v_i)$ denotes the largest degree of all vertices. Because all the degrees $d(v_i) \geq 0$, the eigenvalues of L^* are either zeros or in the right half-plane. Therefore, for the differential equation (60), none of the eigenvalues of the matrix $-L^*$ are purely imaginary or has positive real part. Then, the solution of the differential equation (60) converges to a certain value as time goes to infinite.

$$X(t)|_{t\to\infty} = X(\infty) = C \tag{61}$$

where C is a constant vector.

Proof of $x_i(\infty) = Z$:

In order to complete the proof, the constant $X(\infty) = C$ must be found out. Since $\dot{X}(\infty) = 0$, $X(\infty)$ is the equilibrium point of the differential equation (60). Therefore, the objective is to find the solution $X = X(\infty)$ for the function

$$L^*X = 0 \tag{62}$$

Here, let us define a new matrix $[I]_i^0$ by the equation $[I]_i^0 \triangleq Diag([I, I, \dots, I, I_i^0, I, \dots, I])$, in which the identity matrix I has the same dimension as I_i^0, and the matrix I_i^0 only appears at the ith diagonal block. So the row (and also the column) $n(i-1)+i$ is all zeros for the particular i. The matrix $[I_i^0]_n$ can be decomposed as $[I_i^0]_n = [I]_n^0[I]_{n-1}^0 \dots [I]_1^0$. Then, the matrix L^* can be rewritten into

$$L^* = [I]_n^0[I]_{n-1}^0 \dots [I]_1^0(L \otimes I) \tag{63}$$

On the other hand, define $\mathbf{0}_i^1 \triangleq I - I_i^0$, and $P_i \triangleq \mathbf{1}^T \otimes (\mathbf{1}^T\mathbf{0}_i^1)$. $P_i \in \mathbb{R}^{1\times n^2}$ is a vector that only has ones at its $(kn+i)$th entries, for $k = 0,1,..,n-1$. According to the mixed-product property of the Kronecker product and the property $\mathbf{1}^TL = L\mathbf{1} = \mathbf{0}$, we have

$$\begin{aligned} P_i(L \otimes I) &= [\mathbf{1}^T \otimes (\mathbf{1}^T\mathbf{0}_i^1)](L \otimes I) \\ &= (\mathbf{1}^TL) \otimes (\mathbf{1}^T\mathbf{0}_i^1 I) \\ &= \mathbf{0} \otimes (\mathbf{1}^T\mathbf{0}_i^1 I) = \mathbf{0} \end{aligned} \tag{64}$$

where $\mathbf{0}$ is the zero vector. Then, let $[P_i] \in \mathbb{R}^{n^2\times n^2}$ be a matrix that has only one non-zero row vector P_i at its $\big((i-1)n+i\big)$th row, i.e., $[P_i] = [\mathbf{0}; \mathbf{0}; \dots; \mathbf{0}; P_i; \mathbf{0}; \dots; \mathbf{0}]$. By equation (64), $[P_i](L \otimes I) = \mathbf{0}$ as well. So the right-most two terms of equation (63) become

$$\begin{aligned} [I]_1^0(L \otimes I) &= [I]_1^0(L \otimes I) + [P_1](L \otimes I) \\ &= ([I]_1^0 + [P_1])(L \otimes I) \\ &= E_1(L \otimes I) \end{aligned} \tag{65}$$

where $E_1 \triangleq [I]_1^0 + [P_1]$. Since right multiplying a matrix by $[I]_1^0$ is equivalent to the column operation that eliminates the first column of the matrix, and the first column of $[P_2]$ is already all zeros, then $[P_2][I]_1^0 = [P_2]$. Thus,

$$\begin{aligned}[I]_2^0[I]_1^0(L \otimes I) &= [I]_2^0[I]_1^0(L \otimes I) + [P_2][I]_1^0(L \otimes I) \\ &= ([I]_2^0 + [P_2])[I]_1^0(L \otimes I) \\ &= E_2E_1(L \otimes I)\end{aligned} \tag{66}$$

where $E_2 \triangleq [I]_2^0 + [P_2]$. By performing the similar process as above for n times, equation (63) can be rewritten into

$$L^* = [I_i^0]_n(L \otimes I) = E_nE_{n-1} \dots E_1(L \otimes I) \tag{67}$$

where $E_i \triangleq [I]_i^0 + [P_i]$ (for $i = 1,2, \dots, n$) and it has the following form

$$E_i = \begin{bmatrix} I & | & \mathbf{0} \\ \hline & P_i & \\ \hline \mathbf{0} & | & I \end{bmatrix} \leftarrow the(n(i-1)+i)throw \tag{68}$$

Since P_i is a vector with ones at its $(kn + i)$th entries, the diagonal of the matrix E_i are all ones. So, each E_i can be decomposed as the product of a sequence of elementary matrices corresponding to row-addition transformations. According to equation (67), $[I_i^0]_n$ performs similarly as a series of row-addition transformations on the matrix $L \otimes I$. Because the row operation does not change the null space of the original matrix, the matrices L^* and $(L \otimes I)$ have the same null space, i.e., $Null(L^*) = Null(L \otimes I)$.

For a connected network, the rank of Laplacian L is $n - 1$. According to the property of Kronecker product, $rank(L \otimes I) = rank(L)rank(I) = (n-1)n$. Also, because $L \otimes I$ is a $n^2 \times n^2$ matrix and based on the rank-nullity theorem, the dimension $dim\big(Null(L \otimes I)\big) = n^2 - (n-1)n = n$. Thus, $dim\big(Null(L^*)\big) = n$.

Meanwhile, for an arbitrary n-dimensional vector x^*,

$$\begin{aligned} L^*(\mathbf{1} \otimes x^*) &= [I_i^0]_n (L \otimes I)(\mathbf{1} \otimes x^*) \\ &= [I_i^0]_n [(L\mathbf{1}) \otimes (Ix^*)] \\ &= [I_i^0]_n [\mathbf{0} \otimes (Ix^*)] = \mathbf{0} \end{aligned} \tag{69}$$

thus $\mathbf{1} \otimes x^*$ is in $Null(L^*)$. Since $span\mathbf{1} \otimes x^* \subseteq Null(L^*)$ and $dim\big(Null(L^*)\big) = dim(span\mathbf{1} \otimes x^*) = n$, then $Null(L^*) = span\mathbf{1} \otimes x^*$.

Therefore, the general solution to the equation (62) is

$$X = \mathbf{1} \otimes x^* = [x^*; x^*; \dots; x^*] \tag{70}$$

Since $X(\infty)$ is one of the solutions and $X(\infty) = [x_1(\infty); x_2(\infty); \dots; x_n(\infty)]$, then

$$x_1(\infty) = x_2(\infty) = \cdots = x_n(\infty) = x^* \tag{71}$$

This result shows that all the states x_i will converge to a consensus when the time goes to infinity (in practice, it will converge within an acceptable error in a few reasonable steps). On the other hand, according to the information propagation algorithm equation (58), the term I_i^0 sets the ith row of $\dot{x}_i(t)$ to zero, which means that the ith row of $x_i(t)$ does not change for the whole time. So, it is true that

$$x_{ii}(\infty) = x_{ii}(0) = z_i \tag{72}$$

where $x_{ii}(t)$ denotes the ith element of $x_i(t)$. Finally, all the states x_i will converge to the solution as

$$x_i(\infty) = [x_{11}(0); x_{22}(0); \dots; x_{nn}(0)] = [z_1; z_2; \dots; z_n] = Z \tag{73}$$

■

4.2.4. The Proposed Distributed State Estimation

With the information propagation algorithm, the DSE for power systems can be established. In order to integrate with the discrete form of DSE, the information propagation algorithm can be written into discrete form as

$$Z_i(k+1) = Z_i(k) + \tau I_i^0 \sum_{j \in \mathcal{N}_i} w_{ij}^Z \left(Z_j(k) - Z_i(k) \right) \tag{74}$$

where Z_i denotes the estimated measurements on the ith node. τ is the update time step. w_{ij}^Z is the weight on the edge Z_i, Z_j.

In (19), z_i in the initial value $\Theta_i(0) = H_i^T R_i^{-1} z_i$ is the measurement data of the entire power system, which is estimated by $Z_i(k)$. So, by combining the information propagation algorithm with the distributed WLS state estimation, the proposed DSE for a power system is obtained as follows,

$$Z'_i(k) = Z_i(k), where Z'_{i,i}(k) = z_{i,i}(k) \tag{75}$$

$$Z_i(k+1) = Z'_i(k) + \tau I_i^0 \sum_{j \in \mathcal{N}_i} w_{ij}^Z \left(Z'_j(k) - Z'_i(k) \right) \tag{76}$$

$$P_i(k+1) = P_i(k) + \tau \sum_{j \in \mathcal{N}_i} w_{ij}^P \left(P_j(k) - P_i(k) \right) \tag{77}$$

$$\Theta'_i(k+1) = \Theta'_i(k) + \tau \sum_{j \in \mathcal{N}_i} w_{ij}^\Theta \left(\Theta'_j(k) - \Theta'_i(k) \right) \tag{78}$$

$$\hat{\theta}_i(k+1) = \left(P_i(k+1) \right)^{-1} \Theta'_i(k+1) Z_i(k+1) \tag{79}$$

where $Z_{i,i}(k)$ is the ith entry of the vector Z_i at time k. $z_{i,i}(k)$ denotes the local measurement of the ith vertex at time k. $Z'_i(k)$ is a modification of $Z_i(k)$ that the ith entry is replaced by the actual measurement $z_{i,i}(k)$. So, equations (75) and (76) are used to

constantly broadcast the local information. The initial values are: $Z_{i,i}(0) = z_{i,i}(0)$, $Z_{i,j}(0)$ can be arbitrary initial values (e.g. zero), $P_i(0) = H_i^T R_i^{-1} H_i$, and $\Theta'_i(0) = H_i^T R_i^{-1}$. Note that $\Theta'(k)$ is different from $\Theta(k)$ in equation (19).

Different from typical sensor networks, the H_i and R_i in the proposed equations are not the observation matrix and covariance matrix of the individual sensor. However, since the power flow calculation is based on the information of the entire system, the matrices H_i and R_i are constructed based on the characteristics of the whole system as shown in section 4.2.1. So, these matrices are the same for the vertices in the same system, i.e. $H_1 = H_2 = \cdots = H$ and $R_1 = R_2 = \cdots = R$ for the constant matrices H and R of the power system. Consequently, the matrices P_i and Θ'_i are constant as well, and thus the algorithm (75)-(79) can be replaced by,

$$Z'_{i,j}(k) = Z_{i,j}(k), for j \neq i \tag{80}$$

$$Z'_{i,j}(k) = z_{i,j}(k), for j = i \tag{81}$$

$$Z_i(k+1) = Z'_i(k) + \tau I_i^0 \sum_{j \in \mathcal{N}_i} w_{ij}^Z \left(Z'_j(k) - Z'_i(k) \right) \tag{82}$$

$$\hat{\theta}_i(k+1) = (H^T R^{-1} H)^{-1} H^T R^{-1} Z_i(k+1) \tag{83}$$

4.3. Distributed Optimal Power Flow

In the proposed distributed OPF method, the power system has two networks: the power network and the communication network. The power network connects the loads and generators with the transmission lines. The communication network links all the generator controllers together and also connects all the smart meters. In addition, each generator controller connects to a smart meter (or, a smart meter can be installed as a part

of the controller) to obtain the power flow information. In the proposed scheme, the DSE algorithm runs on the smart meters, and the proposed DDCOPF algorithm is executed on the generator controllers.

4.3.1. The Idea and Framework of The DDCOPF

In the OPF problem, the optimization considers not only the economic dispatch problem with the power balance constraint (7) and the generation limit constraint (8) but also some power flow related constraints. For example, this work focuses on the DC line flow limit constraints (84) as follows [6],

$$\frac{1}{x_{ij}}\left|\theta_i - \theta_j\right| \le P_{ij}^{max}, \qquad for\ i,j = 1,2,\dots,n_b;\ i \ne j \tag{84}$$

where x_{ij} is the reactance between bus i and bus j. θ_i and θ_j are the voltage phase angles on the bus i and j, respectively. P_{ij}^{max} is the power flow limit for the transmission line between bus i and bus j. According to the DC power flow [44], the constraint (84) for the entire power system can also be written in the vector form,

$$\left|P_f\right| = |H\theta| = |H(B')^{-1}P| \le P_f^{max} \tag{85}$$

where P_f is the vector of the line flow in the power system. P_f^{max} is the vector of line flow limits. H is the matrix to convert the phase angle values to the power flow values. θ and P are the vectors of phase angles and nodal injections for buses $2,3,\dots,n_b$, respectively. B' is called the “B-prime" matrix, which transfers the phase angles into the power injection [6]. Since bus 1 is the reference bus and its phase angle is always zero, the B' matrix does not include the row and column for bus 1, so the impact of power injection

at bus 1 cannot be directly calculated with the B' matrix. Therefore, a new matrix T is used in this work to calculate the power flow. The T is constructed as

$$T = \begin{bmatrix} -B_{12} & -B_{13} & \dots & -B_{1,n_b} \\ B_{22} & -B_{23} & \dots & -B_{2,n_b} \\ -B_{32} & B_{33} & \dots & -B_{3,n_b} \\ \vdots & \vdots & \ddots & \vdots \\ -B_{n_b,2} & -B_{n_b,2} & \dots & B_{n_b,n_b} \end{bmatrix}^g \tag{86}$$

where $[\ *\ ]^g$ denotes the generalized inverse of the matrix, so T is a $(n_b - 1) \times n_b$ matrix, and

$$B_{ij} = \begin{cases} 0 & if\ buses\ i\ and\ j\ are\ not\ connected \\ \dfrac{1}{x_{ij}} & if\ buses\ i\ and\ j\ are\ connected \\ \displaystyle\sum_{k=2,k\neq i}^{n_b-1} B_{ik} & if\ i = j \end{cases} \tag{87}$$

Now, the inequality constraint (85) becomes,

$$\left|P_f\right| = |H\theta| = |HTP| \leq P_f^{max} \tag{88}$$

To solve the DDCOPF problem, the DED method [28] is employed in this research to constantly update the solution of the ED problem (7)-(9) to approach the optimal dispatch. However, the constraint (88) defines a feasible region for the solution of the DDCOPF problem. In order to keep the constraint (88) being satisfied, the DDCOPF should limit the solution P_g of the DED algorithm within the feasible region. In this work, a line flow constraint algorithm (as shown in Figure 4-4) is proposed to correct the solution of the DED method, so that the final solution of the DDCOPF is restricted in the feasible region.

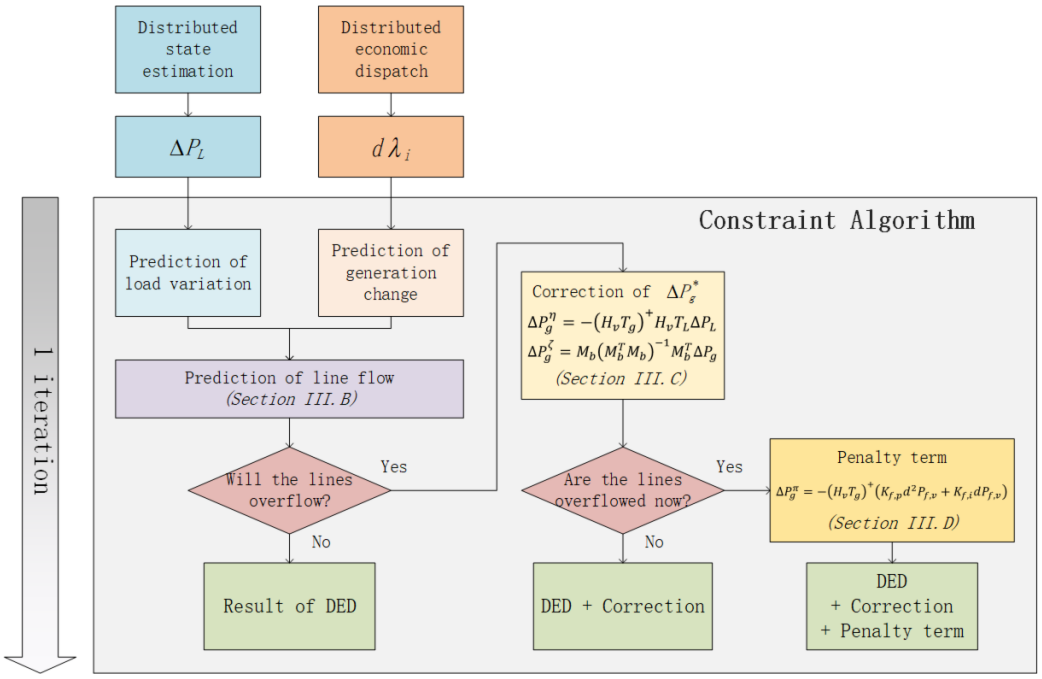


Figure 4-4 The flow chart of the proposed constraint algorithm for DDCOPF

According to Figure 4-4, at the first, the information from the DED algorithm and DSE algorithm is adopted to predict the line flows of the next step in the power system and determines whether there will be overflow. If no potential overflow in the power system, the generation references calculated by the DED algorithm are directly applied to the generators. However, If overflow will happen on any line, the correction algorithm will be performed to correct the references of the generators to prevent the overflow. In addition, if the overflow has already happened in the system, a penalty term will be added to the references to reduce the line flow. In the following sections, the different parts of the framework in Figure 4-4 will be described.

4.3.2. Line Flow Prediction and Overflow Checking

This section discusses the method to predict and check the overflow in the power system. The method is based on the DC power flow model, and the Auto-Regressive (AR) model is adopted for the load prediction.

In the DC power flow model, due to the linearity, the load buses and the generator buses in the power system can be considered separately as follows,

$$P_f = H\theta = HTP = H\left(T_g P_g + T_L P_L\right) \tag{89}$$

where P_g and P_L are the power injection vector of the generator buses and load buses, respectively. T_g is a matrix constructed according to (86) but with only the generator buses. Similarly, T_L is developed by considering only the load buses.

Before step k, the prediction for the line flow at time k can be made by (23).

$$\begin{aligned} P_f(k) &= H\left(\theta(k-1) + \Delta\theta(k)\right) \\ &= P_f(k-1) + HT_g\Delta P_g(k) + HT_L\Delta P_L(k) \end{aligned} \tag{90}$$

The power flow $P_f(k-1)$ at the previous step can be obtained from the DSE algorithm. So, the object is to predict the update of the power injection vector $\Delta P_g(k)$ and $\Delta P_L(k)$.

The update of the generation reference vector $\Delta P_g(k)$ can be obtained from the result of the DED algorithm. In (31), the update of the Lagrange multiplier $\Delta\lambda_i(k)$ is calculated. According to (12), the final result of the generation reference is computed based on the Lagrange multiplier $\lambda_i(k)$. So, for the controller of the jth generator, the update of

the generation reference $\Delta P_{g,i}^{j}(k)$ of the ith generator can be estimated by (91) with its local Lagrange multiplier $\lambda_j(k)$, since $\lambda_j(k) = \lambda_i(k)$ after the consensus protocol is converged.

$$\Delta P_{g,i}^{j}(k) = P_{g,i}^{*}(k) - P_{g,i}^{*}(k-1)$$

$$= \begin{cases} \dfrac{d\lambda_j(k)}{2\gamma_i} & ,if P_{g,i}^{min} < P_{g,i}^{*}(k) < P_{g,i}^{max} \\ 0 & ,otherwise \end{cases} \quad (91)$$

Therefore, by storing the parameters β_i and γ_i for $i = 1,2,\dots n_g$ in each generator controller, the prediction of the generation reference update vector $\Delta P_g(k) = \left[\Delta P_{g,1}^{j}(k); \Delta P_{g,2}^{j}(k); \dots; \Delta P_{g,n_g}^{j}(k)\right]$ can be obtained by the equation (91). Note that the notation $\Delta P_g(k)$ here is a vector for the jth generator. But, since the equation is the same for all controllers and for the sake of simplicity, the superscript j is omitted.

Now, let's consider the prediction of the load variation $\Delta P_L(k)$. There are many methods to predict the load in a power system [45, 46]. In this work, an Auto-Regressive (AR) model of order 2 is used to predict the load variation as follows,

$$\Delta P_L(k) = \varphi_1 \Delta P_L(k-1) + \varphi_2 \Delta P_L(k-2) \quad (92)$$

where the parameters φ_1 and φ_2 can be calculated based on historical data of $\Delta P_L(t)$ by Least Square method in (93),

$$Min \sum_{t=3}^{T} \left(\Delta P_L(t) - \varphi_1 \Delta P_L(t-1) - \varphi_2 \Delta P_L(t-2)\right)^2 \quad (93)$$

By the AR model, the prediction $\Delta P_L(k)$ can be made by (16) with the previous updates $\Delta P_L(k-1)$ and $\Delta P_L(k-2)$, whose previous updates are provided by the DSE algorithm.

After the prediction of generation reference update $\Delta P_g(k)$ and prediction of load variation $\Delta P_L(k)$ are obtained, the prediction for the line flow $P_f(k)$ can be calculated by (90). Then, the overflow in the power system can be checked, and the results are put in a column vector $OF(k)$ whose each entry is defined by,

$$OF_u(k) = \begin{cases} 1 & , P_{f,u}(k) \geq P_{f,u}^{max} \\ 0 & , otherwise \end{cases} \tag{94}$$

where $OF_u(k)$ is the uth entry of vector $OF(k)$ at time k; $u = 1,2,\dots,m_l$. $P_{f,u}(k)$ is the line flow in line u at time k. $P_{f,u}^{max}$ is the maximal line flow allowed for line u.

4.3.3. Correction of Power References

According to Figure 4-4, the power references of the generators calculated by the DED algorithm are directly outputted if the prediction shows no overflow in the power system in the next step. However, if any overflow will happen, the power references should be adjusted to avoid the violation of the line flow constraint. In this section, a method to correct the power references is proposed to make the solution of the DDCOPF stay within the feasible region.

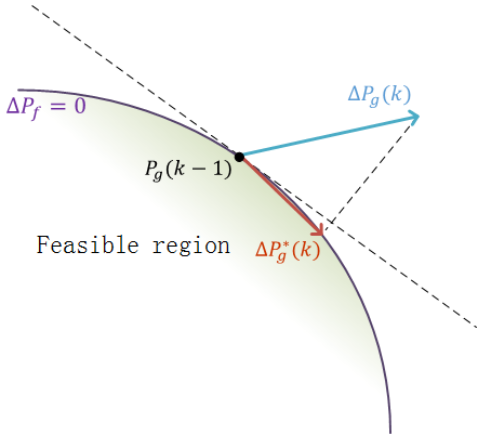


Figure 4-5 The correction of the power reference update vector

Since the power balance constraint (8) and the output limit constraint (9) are already satisfied by the DED algorithm, the correction algorithm in this section should adjust the result of the DED algorithm to meet the line flow constraint (88). Assuming that the current solution $P_g(k-1)$ of the DED algorithm is inside the feasible region but near the boundary as shown in Figure 4-5. If the line flow prediction (90) shows that the constraint (88) will be violated at time k, i.e. the solution $P_g(k)$ will be out of the feasible region, then it is necessary to adjust the update $\Delta P_g(k)$ of the DED solution to prevent the further increasing of the power flow on the lines, denoted by $P_{f,v}(k)$, which will be overflowed according to the prediction. This adjustment can be realized by letting the increment $\Delta P_{f,v}(k)$ to be zero as follows,

$$\Delta P_{f,v}(k) = H_v(k)\left(T_g \Delta P_g(k) + T_L \Delta P_L(k)\right) = 0 \tag{95}$$

where $H_v(k)$ is the matrix related to the lines which will violate the constraint, and it is defined as follows,

$$H_v(k) = Diag\left(\left[OF_1(k), OF_2(k), \dots, OF_{m_l}(k)\right]\right)H \tag{96}$$

In (95), since $\Delta P_g(k)$ is the adjusted update for the DED solution which should be calculated, the equation (95) can be rewritten into the nonhomogeneous form as follows,

$$H_v(k)T_g \Delta P_g(k) = -H_v(k)T_L \Delta P_L(k) \tag{97}$$

Because $H_v(k)T_g$ is usually not a full ranked square matrix, the solution cannot be simply computed by its inverse. So, according to linear algebra [47], the solutions of this nonhomogeneous equation are the sum of a particular solution $\Delta P_g^{\eta}(k)$ of the corresponding nonhomogeneous equation and the general solutions $\Delta P_g^{\zeta}(k)$ of the

homogeneous equation $H_v(k)T_g\Delta P_g(k) = 0$. Here, by equation (97), it is convenient to consider that the particular solution $\Delta P_g^{\eta}(k)$ is to counteract the effect of the load variation $\Delta P_L(k)$ on the line flow. Then, the solution $\Delta P_g^{\zeta}(k)$ of the homogeneous equation is the desired power reference update for the power system with constant loads.

First, let us solve the particular solution $\Delta P_g^{\eta}(k)$ of the nonhomogeneous equation. Since the particular solution is to neutralize the effect of the loads, the difference (if it is not zero) between the left-hand side (the line flow caused by generators) and the right-hand side (the line flow caused by loads) of the equation (97) is required to be minimized. In other words, the solution should be the least square solution. Also, it is better to minimize the norm of the solution $\Delta P_g^{\eta}(k)$, so that the result is more energy-efficient. Therefore, the Moore-Penrose pseudoinverse [48] is used to solve the particular solution $\Delta P_g^{\eta}(k)$ as in (98), since it gives the least square solution with minimal norm.

$$\Delta P_g^{\eta}(k) = -(H_v(k)T_g)^{+}H_v(k)T_L\Delta P_L(k) \tag{98}$$

where $\Delta P_L(k)$ comes from the prediction (92). The superscript notation $(*)^{+}$ denotes the Moore-Penrose pseudoinverse.

Next, let us consider the general solution of the homogeneous equation $H_v(k)T_g\Delta P_g(k) = 0$. All the solutions of the homogeneous equation indeed form the kernel of the matrix $H_v(k)T_g$. Since the kernel is a subspace of the $\mathbb{R}^{n_g}$, it is possible to find a basis that spans the kernel. Suppose that a basis matrix M_b is defined in which its column vectors are the basis of the kernel, and the coordinates for the vectors in the kernel are denoted by X. Because any vector in the kernel satisfies the homogeneous equation, then X can be any

value. Note that the matrices M_b and X are related to $H_v(k)$, hence they are not a constant value, i.e. $M_b(k)$ and $X(k)$. So, the general solution of the homogeneous equation is,

$$\Delta P_g^{\zeta}(k) = M_b(k)X(k) \tag{99}$$

So, the homogeneous equation becomes,

$$H_v(k)T_g M_b(k)X(k) = 0 \tag{100}$$

As discussed above, the homogeneous equation describes the power system with constant loads, thus the change of line flow is only related to the generation. Since the object is to correct the original power reference updates to meet the requirement, it is necessary to preserve as much information from the original solution $\Delta P_g(k)$ as possible, and also satisfy the homogeneous equation (100). So, it is natural to use the Least Square method to find the solution X in which the corrected vector $\Delta P_g^{\zeta}(k)$ is closest to the original vector $\Delta P_g(k)$ and is in the kernel. According to the equation (99), the Least Square solution is,

$$\Delta P_g^{\zeta}(k) = M_b(k)(M_b^T(k)M_b(k))^{-1}M_b^T(k)\Delta P_g(k) \tag{101}$$

Now, as the particular solution of the nonhomogeneous equation and the Least Square solution of the homogenous equation are obtained, the desired solution of the nonhomogeneous equation (97) is

$$\begin{aligned}\Delta P_g^*(k) &= \Delta P_g^{\eta}(k) + \Delta P_g^{\zeta}(k) \\ &= -(H_v T_g)^{+} H_v T_L \Delta P_L + M_b(M_b^T M_b)^{-1}M_b^T \Delta P_g\end{aligned} \tag{102}$$

here, to make the equation clear, the time variable k is omitted. $\Delta P_g^*(k)$ is the corrected power reference update with inputs $\Delta P_g(k)$ and $\Delta P_L(k)$.

4.3.4. Penalty Term for The Infeasible Operating Point

Ideally, the correction algorithm above can ensure the solutions of the DDCOPF method stay in the feasible region. However, in practice, some factors may cause the actual operating point (the power flow) of the power system to be infeasible.

In the correction equation (98), the load variation $\Delta P_L(k)$ is estimated by the AR model (92), thus the error of the prediction will be propagated into the final result. Consequently, the solution $\Delta P_g(k)$ may not be restricted on the boundary and drift out of the boundary due to the prediction error. Also, since the change of power flow is eliminated by the correction algorithm by letting $\Delta P_f(k) = 0$, the infeasible result cannot return to the feasible region.

In addition, the inertia of the generator causes the delay between the output reference (the solution of the DDCOPF) and the actual output. So, when the reference is bounded on the boundary of the feasible region, the actual output may not stop increasing instantly, but overshoot beyond the limit. Then, due to the equation $\Delta P_f(k) = 0$, the operating point may not go back to the feasible region.

Due to the above reasons, if the operating point moves out of the feasible region, a penalty function is needed in the constraint algorithm to pull the infeasible operating point back to the feasible region. In other words, the power flow of the overflowed lines should be decreased to satisfy the line flow constraint (88). According to the equation (95), the power flow increment on the overflowed lines $\Delta P_{f,v}$ includes two parts such that $\Delta P_{f,v} = H_v T_g \Delta P_g + H_v T_L \Delta P_L$. Due to the correction algorithm, the loads related term $H_v T_L \Delta P_L$ is eliminated by the solution $\Delta P_g^{\eta}(k)$ of the nonhomogeneous equation (97). Also, instead of

fixing the operating point by letting its increment $\Delta P_{f,v} = 0$ in the correction algorithm, the penalty term should decrease the power flow by making the increment $\Delta P_{f,v}$ to be negative as follows,

$$\Delta P_{f,v} = H_v(k) T_g \Delta P_g^{\pi}(k) = -F_s(k) \tag{103}$$

where $F_s(k)$ is a vector with positive entries which represents the step size of the power flow decreasing. $\Delta P_g^{\pi}(k)$ is the change added on the generation reference to produce the penalty. Due to the inertia of the power system, the change $\Delta P_g^{\pi}(k)$ on the generation reference cannot affect the power flow instantly, so it is impossible to reduce the power flow $P_{f,v}$ to the limit $P_{f,v}^{max}$ in one iteration by directly letting $\Delta P_{f,v} = -\left(P_{f,v} - P_{f,v}^{max}\right)$. Therefore, the concept of PI controller is employed to calculate the step vector F_s to gradually decrease the power flow to the feasible region.

$$F_s(k) = K_{f,p} d^2 P_{f,v}(k) + K_{f,i} dP_{f,v}(k) \tag{104}$$

where $dP_{f,v}(k) = P_{f,v}(k) - P_{f,v}^{max}$ and $d^2 P_{f,v}(k) = dP_{f,v}(k) - dP_{f,v}(k-1)$. $K_{f,p}$ and $K_{f,i}$ are the proportional gain and integral gain, respectively. Therefore, according to the equations (103) and (104), the change of generation reference to produce the penalty is obtained with the pseudoinverse as follows,

$$\Delta P_g^{\pi} = -(H_v T_g)^{+} (K_{f,p} d^2 P_{f,v} + K_{f,i} dP_{f,v}) \tag{105}$$

In the power system, when the operating point is out of the feasible region, this penalty term $\Delta P_g^{\pi}(k)$ can be added to the result from the correction algorithm, so that the final result becomes $\Delta P_g^{*}(k) = \Delta P_g^{\eta}(k) + \Delta P_g^{\zeta}(k) + \Delta P_g^{\pi}(k)$.

4.4. Distributed AC State Estimation and ACOPF

In the previous sections, the distributed state estimation and optimal power flow schemes are introduced. However, these methods are developed based on the DC power flow model which is a linear method. Because power systems in real life are nonlinear systems, the linear methods may not be able to perform very precisely. For example, the DC state estimation may not give a precise estimation when the system is highly nonlinear. To solve this problem, the AC state estimation and ACOPF (AC Optimal Power Flow) methods are presented in this section. These methods are developed based on the AC power flow model which is a nonlinear model that better describes the power system.

4.4.1. AC Power Flow Model

The AC power flow model for the active and reactive power flow from the bus i to the bus j are summarized as follows,

$$P_{ij}(x) = V_i^2 G_{ij} - V_i V_j [\cos(\theta_i - \theta_j)\, G_{ij} + \sin(\theta_i - \theta_j)\, B_{ij}] \quad (106)$$

$$Q_{ij}(x) = -V_i^2 (B_{ij}^{cap} + B_{ij}) - V_i V_j [\sin(\theta_i - \theta_j)\, G_{ij} - \cos(\theta_i - \theta_j)\, B_{ij}] \quad (107)$$

where $x = [V_{ref}, V_1, V_2, \dots, V_{n_b}, \theta_{ref}, \theta_1, \theta_2, \dots, \theta_{n_b}]^T$ is the state vector which contains the voltage and phase angle variables (to be optimized) of all buses.

4.4.2. Distributed AC State Estimation

Similar to the proposed distributed DC state estimation, the Distributed AC state estimation is also based on the information propagation algorithm introduced in the previous section. According to Newton's method [6] for the centralized AC state estimation, the state estimation problem is solved by updating the estimated states over and over again

to approach the optimal solution. The updating step $\Delta\hat{x}^s$ for the AC state estimation is the following,

$$\Delta\hat{x}^s = (H^{\mathrm{T}}R^{-1}H)^{-1}H^T R^{-1}\left(Z - h(\hat{x}^s)\right) \tag{108}$$

where H is the Jacobian of the observation model $h(\hat{x}^s)$; R is the covariance matrix of measurement errors; Z is the vector of measurements from the sensors; $\hat{x}^s$ is the estimated state of the power system.

This iteration can be combined with the iterations in the information propagation algorithm to update the states of the power system $\hat{x}^s$ and the information state x_i at the same time. That is, for the node i between buses n and m, the iterative equations of the proposed AC DSE algorithm at the step k are listed as follows,

$$\Delta Z_i^P = \tau I_i^0 \sum_{j\in\mathcal{N}_i} w_{ij}^P \left(Z_j^P(k) - Z_i^P(k)\right) \tag{109}$$

$$\Delta Z_i^Q = \tau I_i^0 \sum_{j\in\mathcal{N}_i} w_{ij}^Q \left(Z_j^Q(k) - Z_i^Q(k)\right) \tag{110}$$

$$Z_i^P(k+1) = Z_i^P(k) + \Delta Z_i^P(k) \tag{111}$$

$$Z_i^Q(k+1) = Z_i^Q(k) + \Delta Z_i^Q(k) \tag{112}$$

$$Z_i(k+1) = \left[Z_i^P(k+1);\ Z_i^Q(k+1)\right] \tag{113}$$

$$\Delta Z_i(k+1) = Z_i(k+1) - h\left(\hat{x}^s(k)\right) \tag{114}$$

$$\Delta\hat{x}^s(k+1) = (H^{\mathrm{T}}R^{-1}H)^{-1}H^T R^{-1}\Delta Z_i(k+1) \tag{115}$$

$$\hat{x}^s(k+1) = \hat{x}^s(k) + \Delta\hat{x}^s(k+1) \tag{116}$$

The procedure of the algorithm can be summarized as in Algorithm 2 below,

Algorithm 2 Distributed AC state estimation

1: Initialize $Z_i^P(0)$ and $Z_i^Q(0)$ with random values or zeros.

2: Set $Z_{ii}^{P}(0)$ to the measured real power from the sensor on node i.
3: Set $Z_{ii}^{Q}(0)$ to the measured reactive power from the sensor on node i.
4: **while** $\|\Delta\hat{x}^{s}(k)\|_{\infty} >$ tolerance **do:**
5: Update the estimation vectors Z^{P} and Z^{Q} by the information propagation equations (109)-(110).
6: Combine Z^{P} and Z^{Q} in one vector as in (113).
7: Update the estimation of the state vector $\hat{x}^{s}$ by equations (114)-(116).
8: end while

where $\|\cdot\|_{\infty}$ is the infinity norm which gives the maximal value in the vector. So, the norm $\|\Delta\hat{x}^{s}(k)\|_{\infty}$ represents the state estimation error and the algorithm stops when it is smaller than a given tolerance value.

4.4.3. The Framework of Distributed ACOPF

In section 394.3, the distributed DCOPF method was introduced. In this section, the distributed ACOPF method will be established by formulating the objective function and constraints given in section 3.6 by the AC power flow model (108) and (109).

In the node power balance constraint (23)-(24) in the OPF problem, the active and reactive demand measurements $P_i^{D,meas}$ and $Q_i^{D,meas}$ are usually not directly measured since it is not economical and reliable to install sensors on all loads. Typically, these measurements are obtained from state estimation. In the state estimation methods, instead of reading data from all loads, it only requires a few meters to measure the power flow on transmission lines. Also, the state estimation methods can improve the accuracy of the measurements and rule out bad readings. Since this is a distributed ACOPF method, the distributed AC state estimation introduced in section 4.4.2 is used to provide the active and reactive demand measurements $P_i^{D,meas}$ and $Q_i^{D,meas}$ for the distributed ACOPF method. In this way, the entire control and measurement system for the power system is totally

distributed and AC power flow based. According to this idea, the proposed distributed ACOPF scheme contains 2 layers as shown in Figure 4-6,

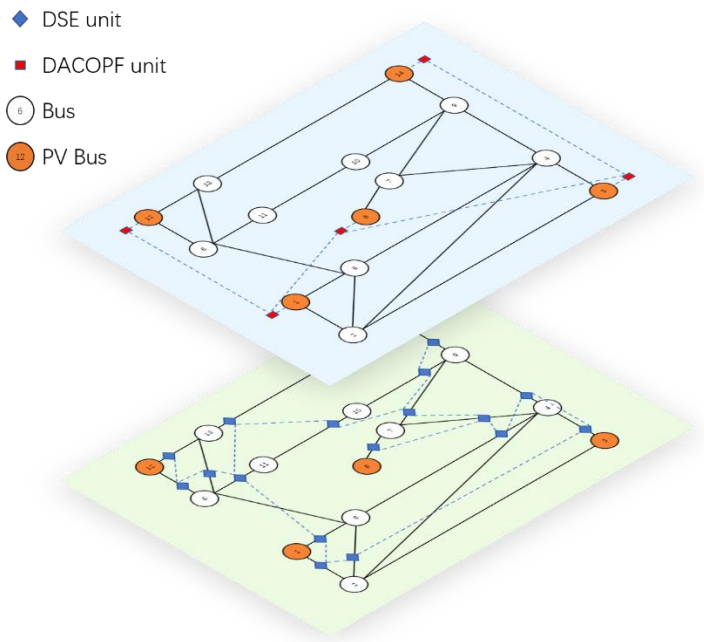


Figure 4-6 The structure of the proposed ACOPF scheme

The first layer in the ACOPF scheme is the DSE (distributed state estimation) layer which performs the distributed state estimation. In this layer, the DSE units are installed on the transmission lines in the power system and they measure the active and reactive power flow. The communication lines connect these DSE units together (typically, each DSE unit only links to the one or two nearest DSE units). The DSE units run the Distributed AC State Estimation program in Algorithm 2 to estimate the states (voltage and phase angle) of the power system.

4.4.4. Linearization of ACOPF Problem

In this research, the ACOPF (AC Optimal Power Flow) problem is solved by iteratively solving the linearized ACOPF problem. This is because, compared with other

nonlinear optimization methods, the linear programming method handles the inequality constraints better [6].

In order to linearize the ACOPF problem, the nonlinear functions such as the cost function (10) for the i_gth generator in the OPF problem should be linearized around a point $x_0 = \left[V_{0ref}, V_{01}, V_{02}, \dots, V_{0n_b}, \theta_{0ref}, \theta_{01}, \theta_{02}, \dots, \theta_{0n_b}\right]$ as follows,

$$\tilde{F}_{i_g}\left(P_{i_g}(x)\right) = F_{i_g}\left(P_{i_g}(x_0)\right) + \left(\beta_{i_g} + 2\gamma_{i_g}P_{i_g}(x_0)\right)\left.\frac{dP_{i_g}}{dx}\right|_{x_0}\Delta x \qquad (117)$$

Where $\tilde{F}_{i_g}$ is the linearized function and $\frac{dP_{i_g}}{dx}$ is

$$\frac{dP_{i_g}}{dx} = \sum_{j=1}^{n_b}\frac{dP_{ij}}{dx} \qquad (118)$$

Also, the AC power flow model (106)-(107) can be linearized at the point x_0 as

$$\tilde{P}_{ij}(x) = P_{ij}(x_0) + \left.\frac{\partial P_{ij}}{\partial V_i}\right|_{x_0}\Delta V_i + \left.\frac{\partial P_{ij}}{\partial V_j}\right|_{x_0}\Delta V_j + \left.\frac{\partial P_{ij}}{\partial \theta_i}\right|_{x_0}\Delta\theta_i + \left.\frac{\partial P_{ij}}{\partial \theta_j}\right|_{x_0}\Delta\theta_j \qquad (119)$$

$$\tilde{Q}_{ij}(x) = Q_{ij}(x_0) + \left.\frac{\partial Q_{ij}}{\partial V_i}\right|_{x_0}\Delta V_i + \left.\frac{\partial Q_{ij}}{\partial V_j}\right|_{x_0}\Delta V_j + \left.\frac{\partial Q_{ij}}{\partial \theta_i}\right|_{x_0}\Delta\theta_i + \left.\frac{\partial Q_{ij}}{\partial \theta_j}\right|_{x_0}\Delta\theta_j \qquad (120)$$

where

$$\left.\frac{\partial P_{ij}}{\partial V_i}\right|_{x_0} = 2V_{0i}G_{ij} - V_{0j}\left[\cos\left(\theta_{0i} - \theta_{0j}\right)G_{ij} + \sin\left(\theta_{0i} - \theta_{0j}\right)B_{ij}\right] \qquad (121)$$

$$\left.\frac{\partial P_{ij}}{\partial V_j}\right|_{x_0} = -V_{0i}\left[\cos\left(\theta_{0i} - \theta_{0j}\right)G_{ij} + \sin\left(\theta_{0i} - \theta_{0j}\right)B_{ij}\right] \qquad (122)$$

$$\left.\frac{\partial P_{ij}}{\partial \theta_i}\right|_{x_0} = V_{0i}V_{0j}\left[\sin\left(\theta_{0i} - \theta_{0j}\right)G_{ij} - \cos\left(\theta_{0i} - \theta_{0j}\right)B_{ij}\right] \qquad (123)$$

$$\left.\frac{\partial P_{ij}}{\partial \theta_j}\right|_{x_0} = -V_{0i}V_{0j}\left[\sin\left(\theta_{0i}-\theta_{0j}\right)G_{ij} - \cos\left(\theta_{0i}-\theta_{0j}\right)B_{ij}\right] \tag{124}$$

and

$$\begin{aligned}\left.\frac{\partial Q_{ij}}{\partial V_i}\right|_{x_0} &= -2V_{0i}\left(B_{ij}^{cap}+B_{ij}\right) \\ &\quad - V_{0j}\left[\sin\left(\theta_{0i}-\theta_{0j}\right)G_{ij} - \cos\left(\theta_{0i}-\theta_{0j}\right)B_{ij}\right]\end{aligned} \tag{125}$$

$$\left.\frac{\partial Q_{ij}}{\partial V_j}\right|_{x_0} = -V_{0i}\left[\sin\left(\theta_{0i}-\theta_{0j}\right)G_{ij} + \cos\left(\theta_{0i}-\theta_{0j}\right)B_{ij}\right] \tag{126}$$

$$\left.\frac{\partial Q_{ij}}{\partial \theta_i}\right|_{x_0} = -V_{0i}V_{0j}\left[\cos\left(\theta_{0i}-\theta_{0j}\right)G_{ij} + \sin\left(\theta_{0i}-\theta_{0j}\right)B_{ij}\right] \tag{127}$$

$$\left.\frac{\partial Q_{ij}}{\partial \theta_j}\right|_{x_0} = V_{0i}V_{0j}\left[\cos\left(\theta_{0i}-\theta_{0j}\right)G_{ij} + \sin\left(\theta_{0i}-\theta_{0j}\right)B_{ij}\right] \tag{128}$$

4.4.5. The Proposed ACOPF algorithm

As discussed in section 4.4.3, the DSE units run the distributed state estimation algorithm to estimate the states of the power system, then the results are sent to the ACOPF units for optimization. The ACOPF units perform the distributed ACOPF algorithm to solve the optimal power flow problem shown in section 3.6. The values $P_i^{D,meas}$ and $Q_i^{D,meas}$ in the constraints (23) and (24) of the optimzation are obtained by computing the AC power flow model with the power system states, i.e., phase angle and voltage, from the DSE algorithm. In order to get the precise power system states, the ACOPF algorithm does not run at the beginning of the entire algorithm. Meanwhile, the DSE algorithm is running solely to obtain the accurate power system states. After the DSE algorithm converges

(when $|\Delta\hat{x}^s|$ is less than a threshold), the ACOPF algorithm starts to run with the latest $P_i^{D,meas}$ and $Q_i^{D,meas}$. The distributed ACOPF algorithm iteratively computes the linearized ACOPF optimization problem given in section 4.4.4. In every iteration, the ACOPF algorithm obtains the power system states from the nearby DSE unit to update the values of $P_i^{D,meas}$ and $Q_i^{D,meas}$ in the constraints (23) and (24). Then, a linear programming solver is used to optimize the linearized ACOPF problem.

By running the ACOPF algorithm given in above, the ACOPF units can solve the ACOPF problem and reach the optimal power flow. However, the iterative linear program method used in the ACOPF is not a global optimization method, so it is possible that some ACOPF units get the local optimal result near the global optimal point. This may cause inconsistent generation assignments over all generators in the system, in which a part of generators are assigned with the power references from one optimal solution and another part of generators have generation references from a distinct local optimal solution. Since these local optimal solutions are different, then they cannot be consistent. So, it is likely that the system will not be balanced, which the total generation is not equal to the total loads plus the losses. This will cause the system to unstable. To resolve this problem, the consensus protocol is used to synchronize the ACOPF results in the ACOPF units.

With the ACOPF optimization algorithm and the consensus protocol, the proposed distributed ACOPF discussed above is summarized in Algorithm 3 below.

Algorithm 3 Distributed AC Optimal Power Flow

1: **while** $|\Delta\hat{x}^s|$ <tolerance **do:**
2: Read the estimated state $\hat{x}^s$ from the connected DSE unit (Algorithm 2).
3: Calculate the loads $P_i^{D,meas}$ and $Q_i^{D,meas}$ on all PQ buses by AC power flow model (106)-(107) with the estimated state $\hat{x}^s$.

4: Perform one iteration of the consensus protocol over the latest ACOPF results $x_i(k)$ on all ACOPF units. The result of the consensus protocol on the ith ACOPF unit is denoted by $x_i'(k)$.
5: Optimize the linearized ACOPF problem at the initial point $x_i'(k)$ and get the new optimal result $x_i(k+1)$.
8: end while

Chapter 5. Numerical Simulations

5.1. Simulation for Distributed Economic Dispatch

In this section, the proposed DED control scheme is simulated in several different power systems, which are a four-generator power system, an IEEE 39-bus system, and other four power systems with different sizes. The simulations are conducted in Matlab Simulink software. In the first two cases, the normal economic dispatch and frequency control problems are studied. Then, the communication-failure-tolerant consensus protocol is tested in a power system with communication failure. After that, the consensus protocol considering the line loss model is simulated. Finally, a sensitivity analysis for the system size is provided to show the applicability of the proposed method in different power systems.

5.1.1. Four-Generator System

In the first case, the simulation is done in a four-generator power system. The structure of this system is shown in Figure 5-1 Power system with four generators. There are four distributed generators (DGs) and three loads in the system. Each generator has a controller which can communicate with the neighbor controllers, and the distributed ED algorithm has been implemented in each controller. The parameters of the generators are presented in Table 5-1 and the output of each generator is limited to $70\% \sim 105\%$ of their

rated capacity P_{rated}. Meanwhile, the loads can be changed to imitate the load variation in a real power system. The model of the entire system is built in Simulink.

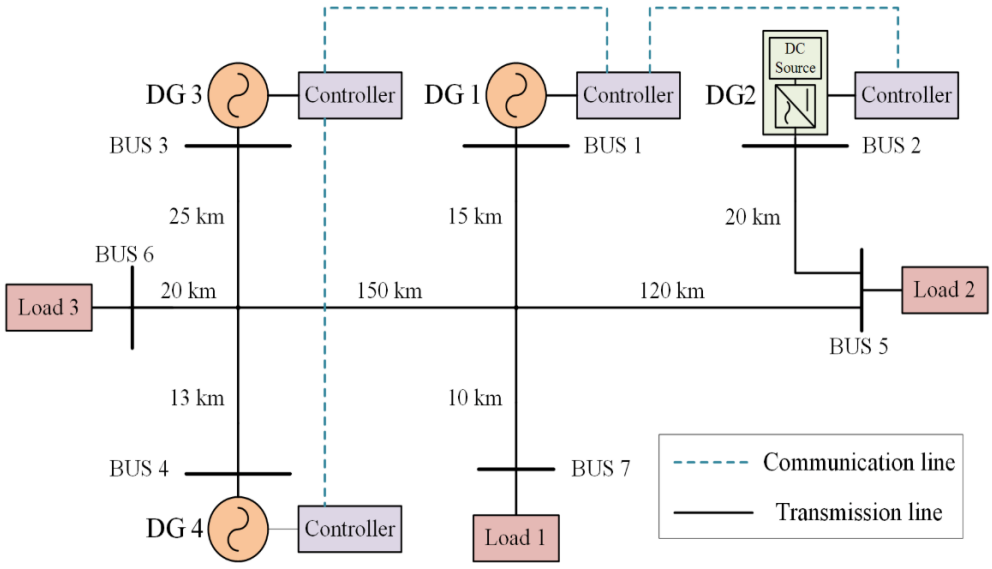


Figure 5-1 Power system with four generators

Table 5-1 The Parameters of Generators

	α_i	β_i	γ_i	P_{rated} (MW)
DG1	561	8.08	0.00118	150
DG2	310	7.8	0.00346	80
DG3	278	7.85	0.00322	80
DG4	453	8	0.00184	110

First, the simulation is conducted for the distributed economic dispatch with PI controller from section 4.1.1. To simulate the load variation, the loads are changed at different times. At 10s, Load 1 is changed from 120MW to 140MW. Then, Load 2 is decreased from 75MW to 60MW at 25s. Finally, at 40s, Load 3 is increased from 150MW to 170MW. The simulation results are shown in Figure 5-2.

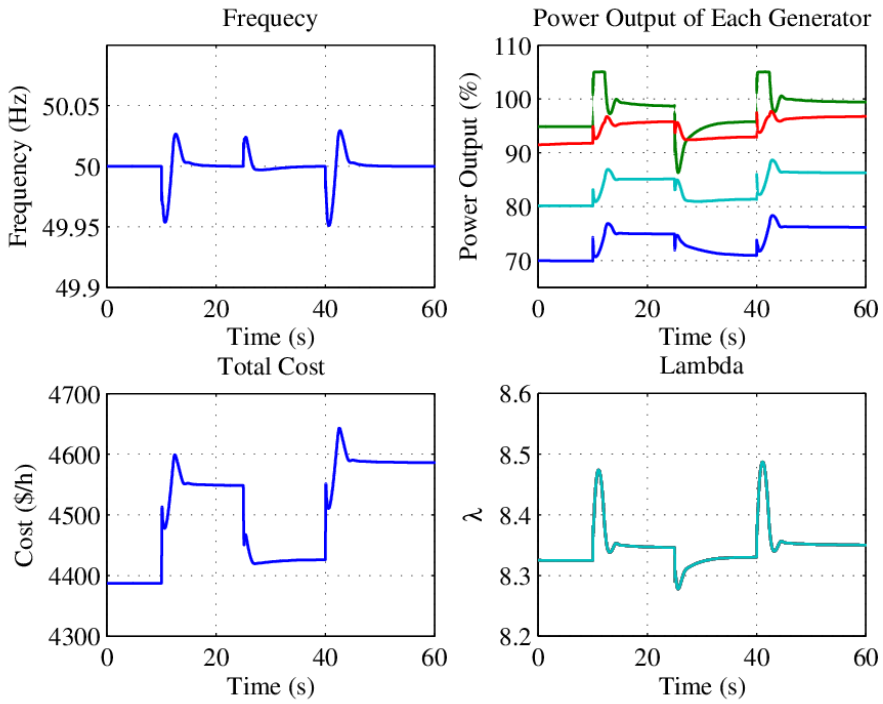


Figure 5-2 The simulation results of the DED algorithm with PI controller

In the results, the frequency of the system remains stable at 50Hz at the steady-state and the deviation is lower than 0.05Hz when the loads change. It means that the control scheme is stable and the power balance condition in equation (8) has been satisfied. The profiles of all the λ_i are displayed in the fourth subplot in Figure 5-2, and this result tells that all the Lagrange multipliers are identical and thus the optimal operation is achieved. Figure 5-3 shows that the λ's for all generators can reache the consensus in a very short time (less than 0.01s), which supports the assumption that the consensus protocol ensures the convergency of the states $\lambda_i(k)$ very fast.

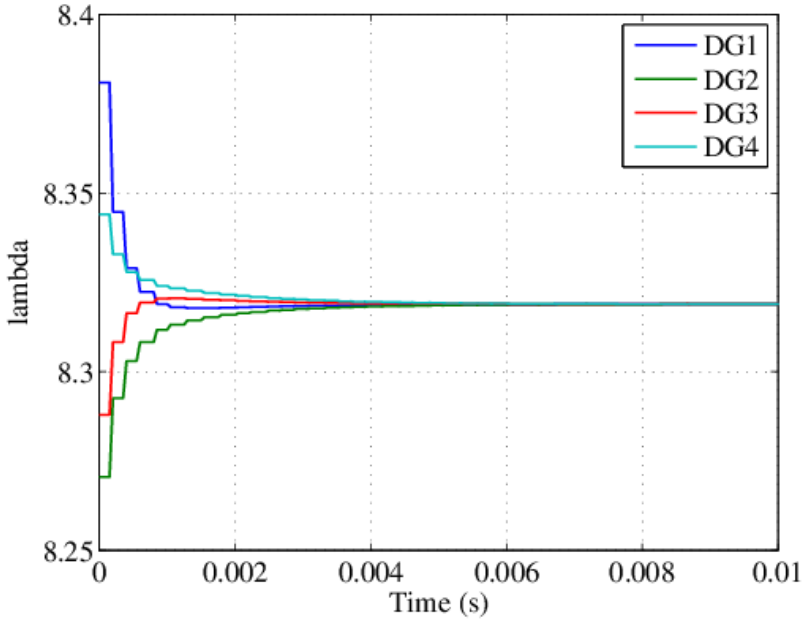


Figure 5-3 The convergence of lambda

Second, the DED control method with an NN controller is simulated in this power system. To train the neural network, the input and output data of the PI controller are collected as the training dataset for the neural network. In the training process, 70% of data from the entire dataset is the training data, 15% of data is for validation, and the remaining 15% of data is for testing. The neural network is trained by the Levenberg-Marquardt method [49, 50]. After the training, the information of the trained weights and biases of the neural network is applied in the algorithm (41)-(45). Then, the algorithm is simulated in the power system model. The results of the simulation are shown in Figure 5-4.

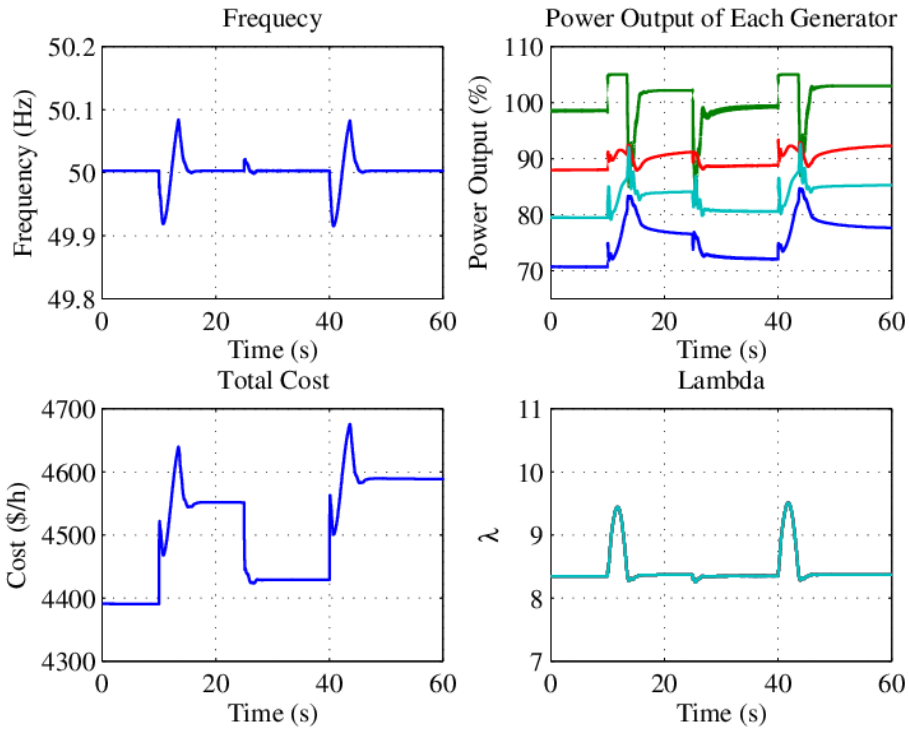


Figure 5-4 The simulation results of the DED algorithm with NN controller

A similar conclusion of the results of the DED with PI controller can be obtained from this result. The proposed distributed control method with an NN controller keeps the system stable and economic during the load variations.

Finally, the proposed distributed method is compared with the centralized method. However, it is hard to compare the two methods directly. This is because the conventional centralized economic dispatch is performed with a 5-min or longer interval, but the proposed distributed economic dispatch method is operated in real-time. So, as the simulation results are shown above, the optimal dispatch can be instantly reached under the load variations with the proposed distributed method, but this time scale is too small for the centralized method to react. Thus, to verify the performance of the proposed distributed method, the algorithm of the centralized ED method is adopted to calculate the

optimal solution of the system, and the calculation is made at every instant based on the data collected from the simulation. The costs of the power system under the two different methods are shown in Figure 5-5. According to the result, the distributed method obtains almost the same results as the centralized method. It means that the economic dispatch is reached under the distributed method. However, since the conventional centralized ED method in a real power system cannot be performed in real-time, the proposed distributed method has a better performance by tracking the optimal solution in a transient state.

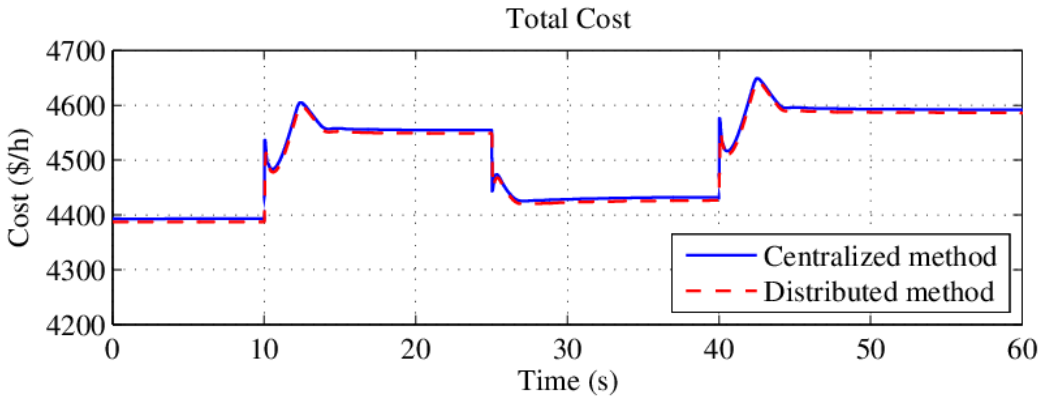


Figure 5-5 The comparison between the centralized method and distributed method in the 4-generator system

5.1.2. IEEE 10-Generator 39-Bus System

In addition to the four-generator system, the IEEE 10-generator 39-bus System, as shown in Figure 5-6, is built in Simulink to test the proposed method. The 39-bus System is operated at 60Hz, and 19 loads are installed in the system. The communication network in the system connects all the generators as the dash lines in Figure 5-6, so there is no centralized control center. The parameters of the generators in the power system are listed in Table 5-2.

Table 5-2 The Parameters of Generators in the 39-bus System

	α_i	β_i	γ_i	P_{rated} (MW)
DG1	561	8.08	0.00118	1000
DG2	310	7.8	0.00346	1000
DG3	278	7.85	0.00322	1000
DG4	453	8	0.00184	1000
DG5	453	8.1	0.00248	1000
DG6	524	7.95	0.00385	1000
DG7	384	7.86	0.00268	1000
DG8	368	7.75	0.00362	1000
DG9	572	8.12	0.00262	1000
DG10	426	8.03	0.00368	1000

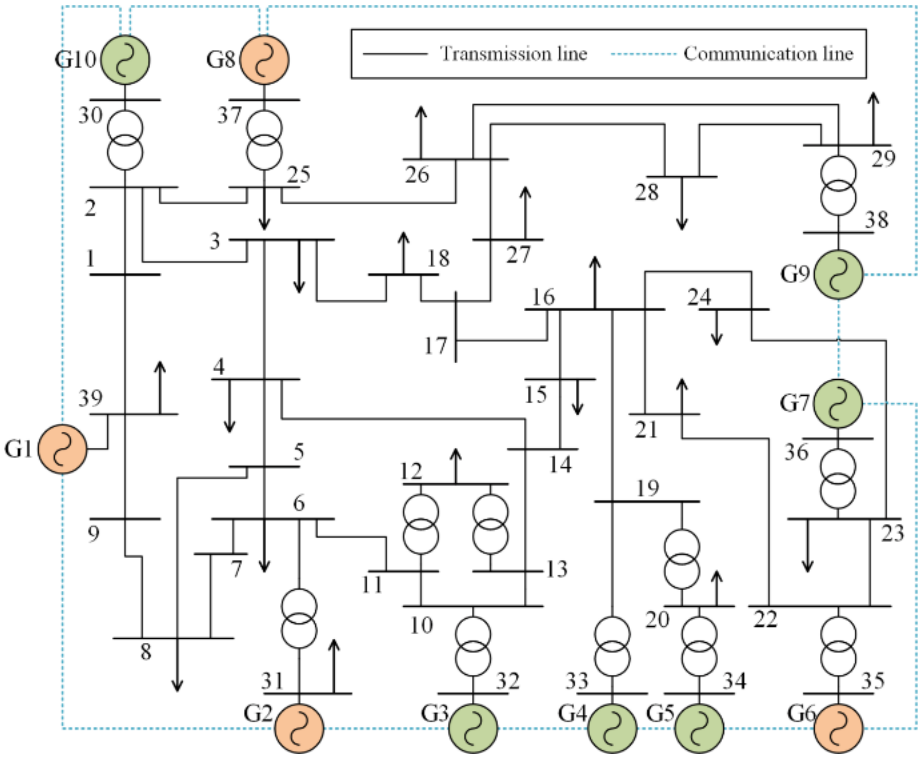


Figure 5-6 A 39-bus Power System with 10 generators

In the simulation, the load (158MW, 30MVAR) on bus 18 is turned off at 10s, then a load (100MW, -50MVAR) is added on bus 24 at 20s. This is to simulate the load

variations in a real-world power system and test the stability of the power system with the proposed control scheme. Since these load changes are much bigger than the normal load variations in a real power system, the performance of the proposed control scheme can be exhibited. The results are shown in Figure 5-7.

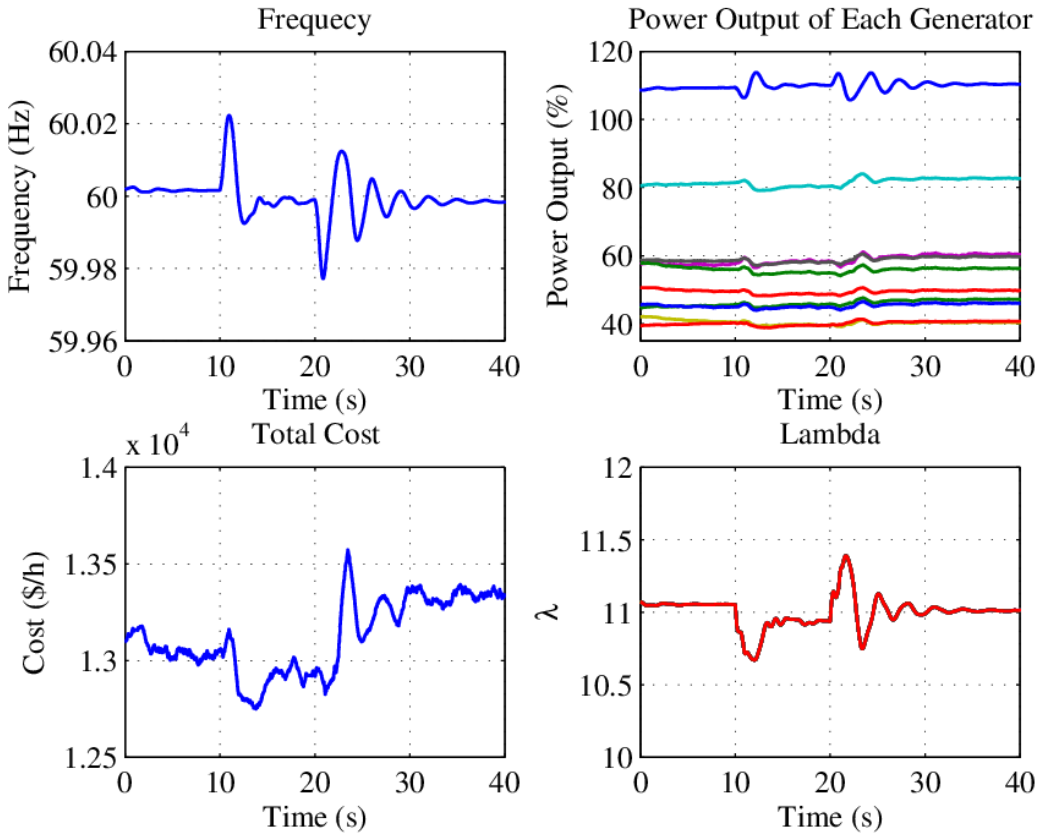


Figure 5-7 The results of the distributed control for the 39-bus Power System

In the results, the first sub-figure (upper-left) of Figure 5-7 shows the frequency profile of the power system during the simulation. Two fluctuations occurred when the loads changed at 10s and 20s, but the system is stabilized quickly by the distributed controllers. The second sub-figure (upper-right) shows the active power output in the percentage of each generator. The real-time calculated cost of the system is shown in the

third sub-figure (lower-left). The last sub-figure (lower-right) represents the Lagrange multipliers λ's of all generators. The profiles of all the λ's are overlapped in the figure, which means the Lagrange multipliers of all generators are identical as expected, and the economic dispatch is achieved continuously.

The results show that the frequency control (stabilize the frequency) and economic dispatch (equalize the incremental costs of generators) are achieved by the proposed DED control method in a distributed manner. The results of this 39-bus system are similar to the results in the four-generator system in the last case, so the proposed control scheme is a universal method.

In Figure 5-8, the costs of the distributed method and the centralized method are presented. The two methods get the same result, which shows that the economic dispatch can be achieved by the proposed distributed method.

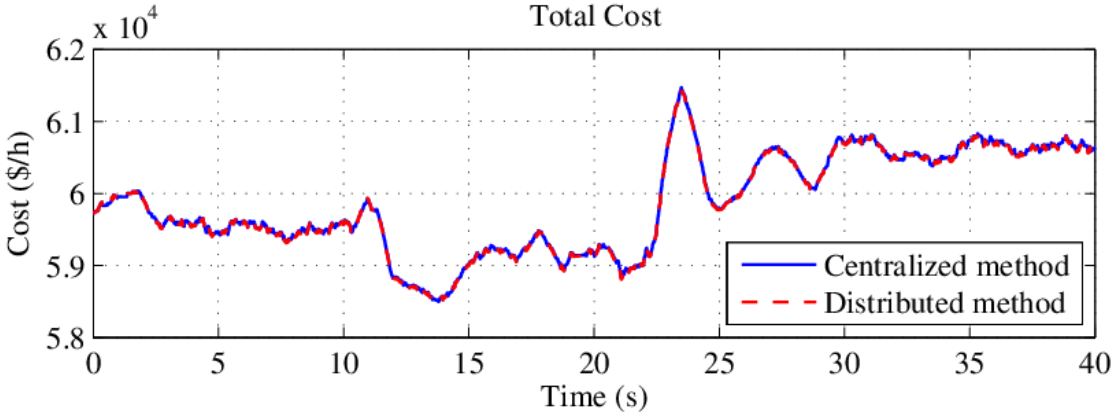


Figure 5-8 The comparison between the centralized method and distributed method in the 39-bus system

5.1.3. Communication Failure

In this section, the communication failure is introduced in the four-generator system to test the controller gain design scheme in the communication-failure-tolerant

DED control method. Then, the leader-follower mode is also tested in the 39-bus system under the link faults.

To simulate the communication failure in the four-generator system in Figure *5-1*, the link between DG1 and DG3 is disconnected at 1s. Then, a 15MW of Load 1 on Bus-7 is turned off at 5s. The results of the DED control method and the communication-failure-tolerant DED control method with controller gain design are shown in Figure 5-9 and Figure 5-10, respectively.

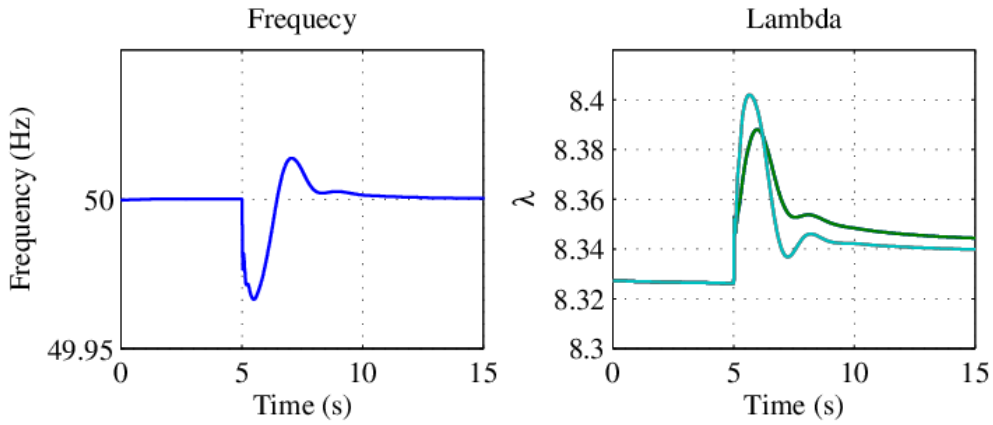


Figure 5-9 Communication failure for the basic consensus protocol

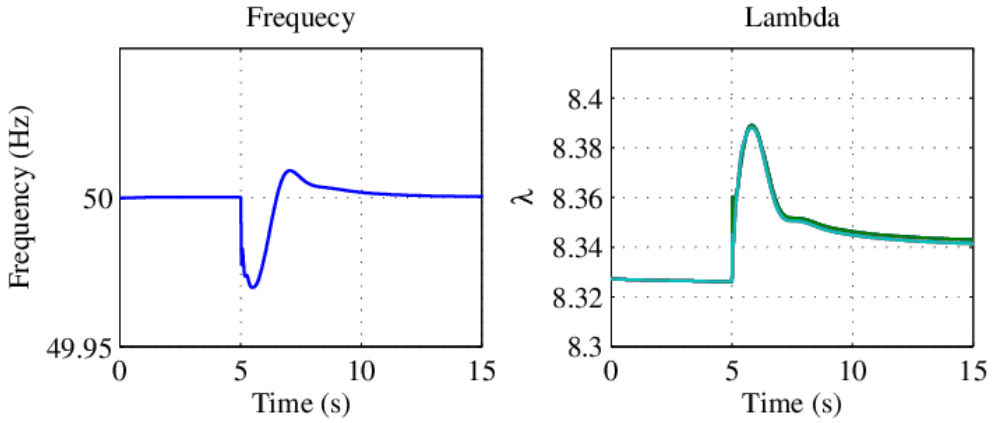


Figure 5-10 Communication failure for the communication-failure-tolerant consensus protocol

According to the results, in both cases, the system operation is not affected when the communication failure just happens (1s~5s). This is because, in the steady-state, the consensus protocol acts according to equation (28) and does not rely on the information from neighbors. But, after the loads change at 5s, these two types of consensus protocols perform differently. In the system with the basic consensus protocol applied, the Lagrange multipliers λ_i cannot reach the consensus after 5s. While the communication-failure-tolerant consensus protocol shows a better result. However, the λ's in the system with communication-failure-tolerant consensus protocol do not perfectly remain identical. This is because the design of the communication-failure-tolerant controller gain is based on the assumption that the measured frequency is the same at all buses in the power system. But in a large power system, this assumption may not be satisfied due to the measurement error and the difference of the dynamic between generators, especially during the transient state. Figure 5-11 shows that the frequency measured at different buses is different during the transient period. To reduce the influence of this problem, it is better to set the integral coefficient Kp_i in the equation (31) as small as possible, so that the error culminates slower.

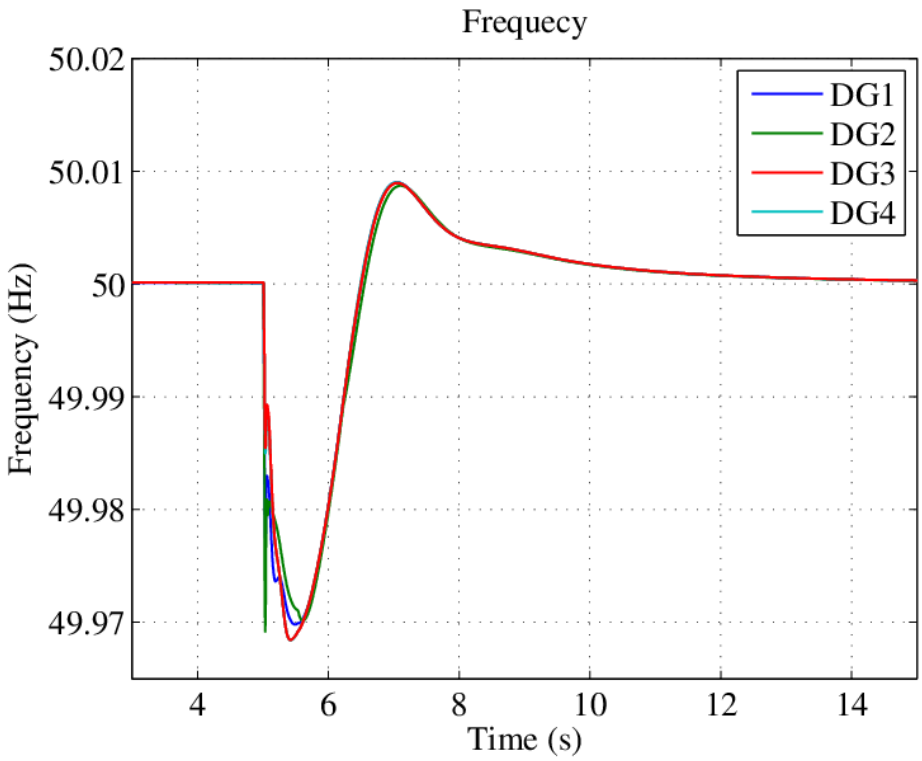


Figure 5-11 The measured frequency on each generator

On the other hand, two link faults are injected in the 39-bus system in Figure *5-6* to test the leader-follower mode in the communication-failure-tolerant DED control method. The G1, G2, G6, and G8 in the system are the leaders, and the rest generators are the followers which cannot measure the frequency. The two faults occur at the link between G1 and G10 and the link between G9 and G7 at 5s, in which the entire communication network is divided into two areas: Area 1 includes G10, G8, and G9; area 2 contains G1-G7. Then, at 10s and 20s, two loads variation are performed. The result is shown in Figure 5-12. As the result, the Lagrange multipliers stay identical after the communication failure. So it shows that the leader-follower method improves the robustness of the DED control method during the faults.

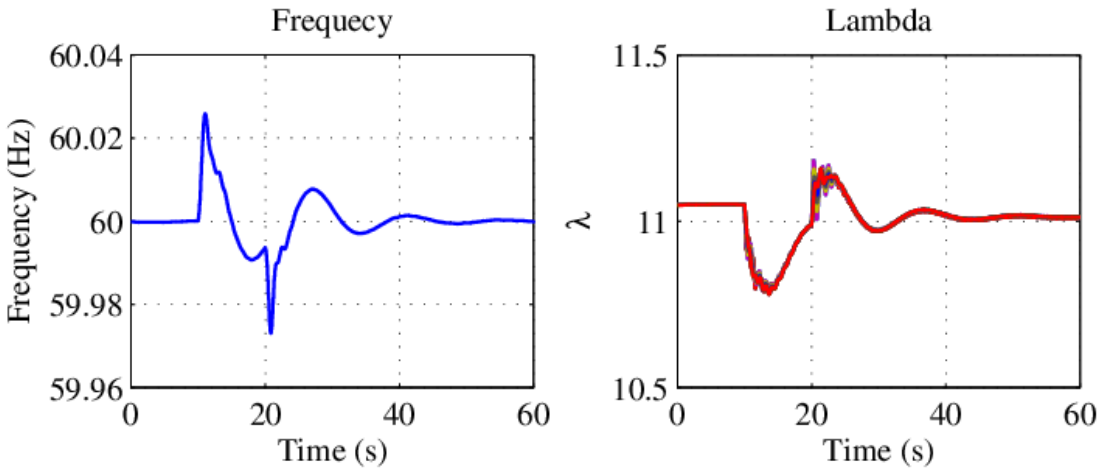


Figure 5-12 The communication failure in the 39-bus system with the leader-follower mode

In addition, the robustness of the proposed method during the communication failure is also influenced by the topology of the communication network. According to the graph theory, the cycle graph has higher edge-connectivity than the path graph, which means that the cycle network is more robust than the path network. The 39-bus system shown in the previous case study has a cycle network as shown in Figure *5-6*. To show the impact of communication failure on the cycle network, a simulation is made in the 39-bus system to disconnect the communication between G1 and G10 at 5s. The result is shown in Figure 5-13. According to the result, the system stays stable after the communication failure at 5s. Also, by the profile in the second subplot, the Lagrange multipliers stay together during the load variation at 10s and 20s, which means that the system stays at the optimal as desired. The result shows that the proposed distributed control method is more robust in this cyclical communication network, and the control objective can be still reached even though one link in the network is disconnected. Furthermore, if a topology

with higher edge-connectivity is chosen to construct the communication network, the system can be more robust.

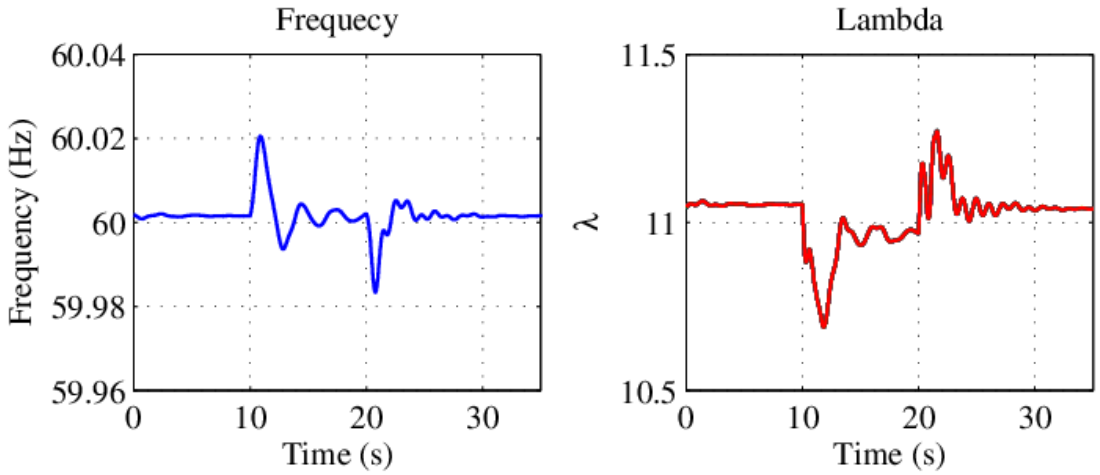


Figure 5-13 The communication failure in the 39-bus system

5.1.4. Transmission Loss

In the last case, the line losses model (53) is employed in the DED solution. The parameters of the line losses model are: $L_1 = 0.00003367$, $L_2 = 0.00003416$, $L_3 = 0.00006402$ and $L_4 = 0.00002715$. According to the analysis in section 4.1.4, the solution of the ED problem becomes (55). In the simulation, Load 1 is increased by 15MW at 5s. The results are compared with the basic consensus approach in Figure *5-14*, Figure *5-15*, and Figure *5-16*.

In Figure *5-14*, the two methods show almost the same results in terms of the total cost of the power system. The reason is that, with the traditional economic dispatch, the total demand P_D in the power balance constraint (8) is obtained by collecting the load data and calculating the line loss by the model (53). Thus, the power balance constraint cannot be actually satisfied if there is no line loss model. However, in the proposed distributed ED

method, the frequency information is used to satisfy the power balance constraint, in which the frequency deviation is related to the power imbalance. So, no matter whether the line loss model is included, the constraint can be satisfied. Then, the cost should be the same, though the ED solution of the optimization may be different.

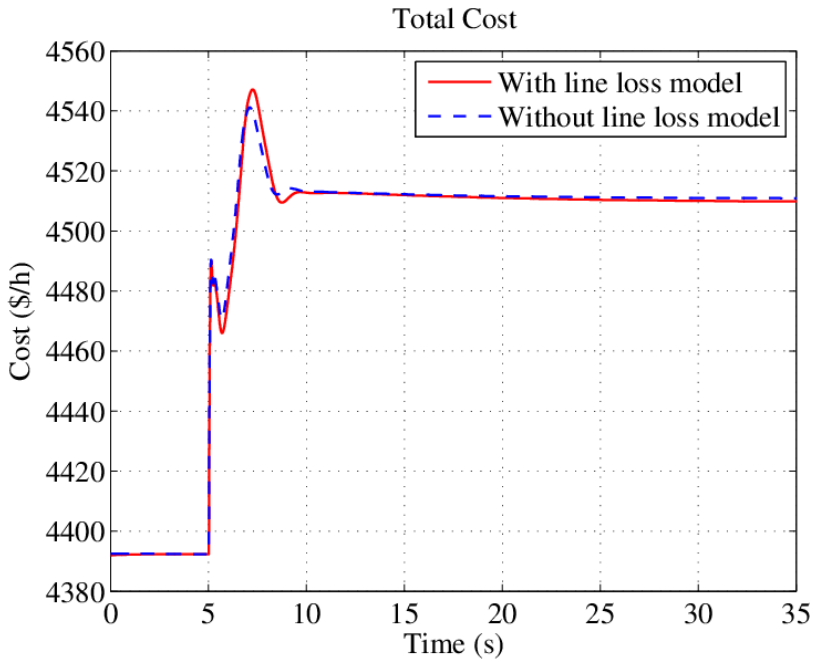


Figure 5-14 Total cost of the power system

In Figure *5-15*, the algorithm with the line loss model considered has the lower line loss. According to the line loss model (53), the parameters reflect the “weight" of each generator in the line loss. So, the bigger the L_i is, the more losses the ith generator causes. On the other hand, based on equation (54), the generation reference $P_i^*(k)$ will become smaller as the L_i gets bigger, thus the line losses decrease.

In this case study, the parameter L_4 is the smallest and L_3 is the largest one. To decrease the line losses, according to the line loss model (53), the DG4 should generate more power and DG3 should output less, which are supported by the results in Figure *5-16*.

Therefore, the modified algorithm is useful to minimize line losses while maintaining the lowest cost.

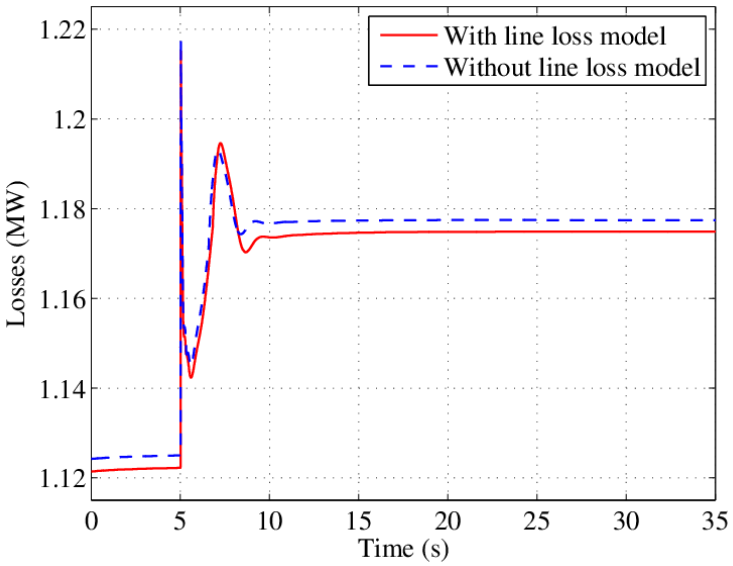


Figure 5-15 The line losses of the power system

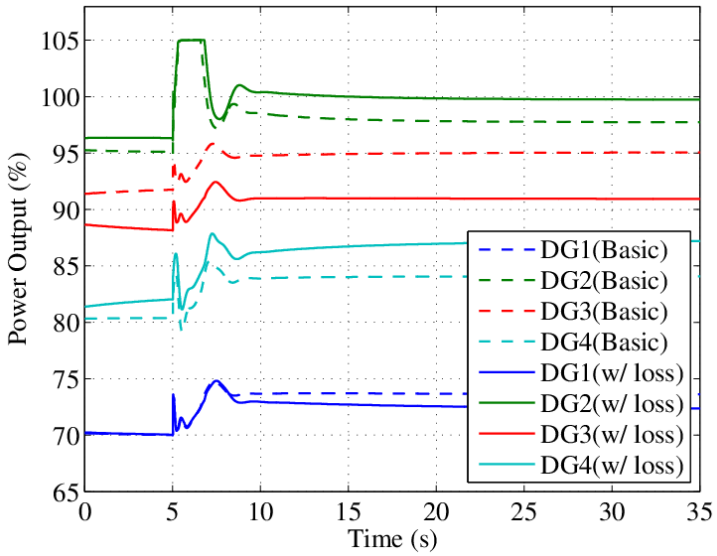


Figure 5-16 The real power outputs of generators (DGi (Basic) represents the generator performing the basic consensus protocol; DGi (w/ loss) loss denotes the generator with modified consensus protocol)

5.1.5. Sensitivity Analysis

To verify the performance of the proposed method in different power systems, a sensitivity analysis of the control scheme for the power system size is provided in this section. In this test, four power systems with different sizes are adopted. The power systems are 4-generator 7-bus system, 8-generator 15-bus system, 16-generator 31-bus system, and 32-generator 63-bus system. The 7-bus system is built at the first, including 4 generators (one 200MW generator and three 100MW generators) and 3 loads (100MW each). In the simulation, a 15MW load will be connected to the system at 5s. For the other three bigger systems, they are created by adding the "load-generation pairs" in the small system as shown in Figure *5-17*. A load-generation pair contains two buses: a PV bus with a 100MW generator and a PQ bus with an 80MW load. In addition, a 4MW load will be added on the PQ bus at 5s. For example, the 15-bus system is obtained by adding four load-generation pairs in the 7-bus system. By creating the power systems with those "load-generation pair" modules, the load-generation ratios of these power systems all remain 0.8, and the increased loads at 5s are always 5% of the total loads. Also, the controller settings of these four power systems are the same. Therefore, the only difference, except for the generator parameters, between these power systems is the size, and the simulations of these power systems are comparable.

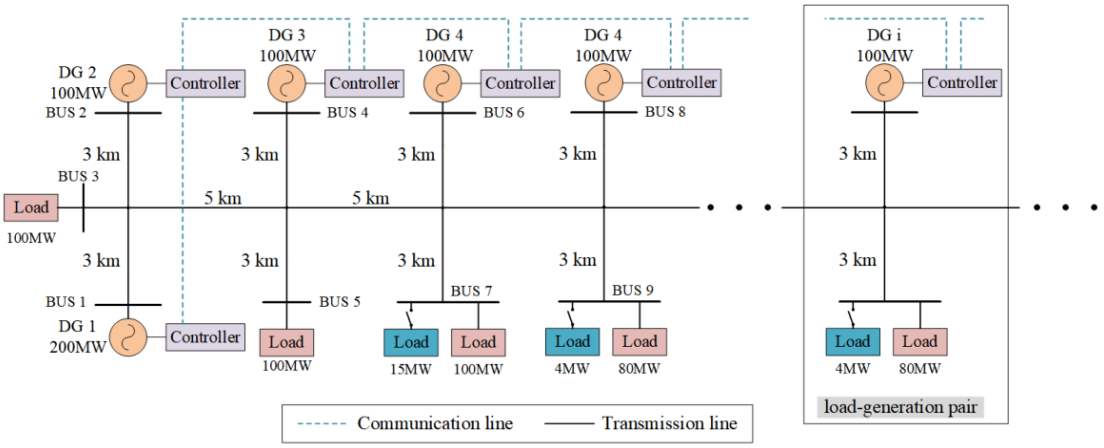


Figure 5-17 The structure of the power systems for sensitive analysis

The frequency profiles of the four different power systems are shown in Figure *5-18*. According to the results, the proposed control scheme can stabilize the power systems with different sizes under load variation. In the figure, the frequency fluctuation is slightly increased with the power system size. This is because the large power system has bigger inertia, so the system fluctuation is higher under the same control settings.

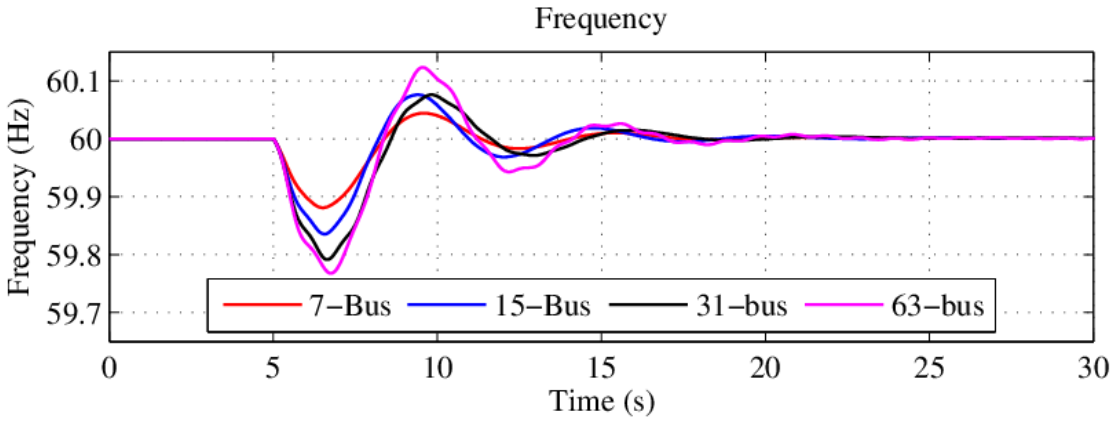


Figure 5-18 The frequency fluctuation of the four power systems

In addition, the convergence of the consensus algorithm in the four power systems is tested. In this test, the initial values of the Lagrange multipliers of the generators are randomly chosen. The results are presented in Figure *5-19*. The results show that the

consensus algorithm can converge the Lagrange multipliers to the identical value as desired in power systems with different sizes. However, the convergence of the large power systems is slower than the small power systems, since there are too many nodes. But this can be overcome by increasing the weight coefficients between the nodes or reducing the time interval 'τ' in the equations 17 and 18 to accelerate the convergence. Also, the topology of the communication network influences the convergence speed of the algorithm. So the convergence speed can be improved by adding cycles in the communication network. In addition, according to the characteristics of the consensus algorithm, once the Lagrange multipliers are converged, they will remain converged unless there is a communication failure.

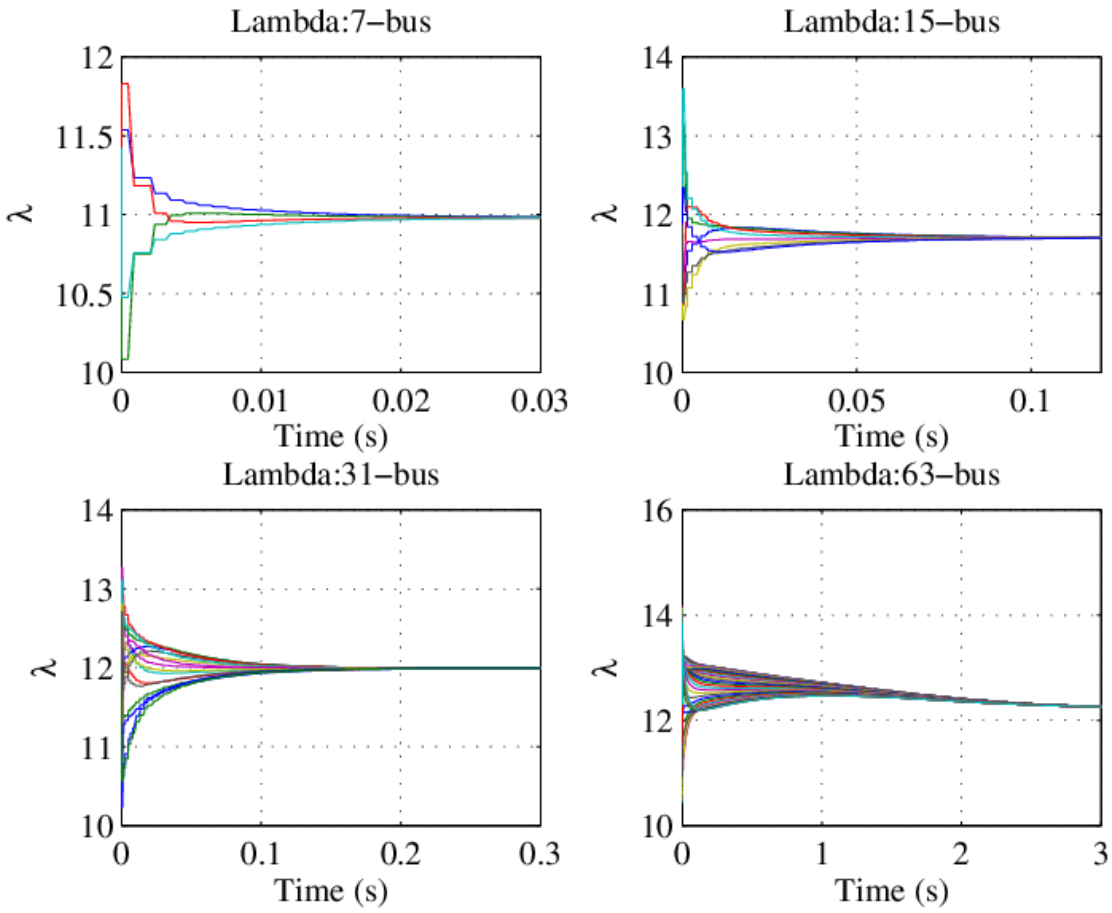
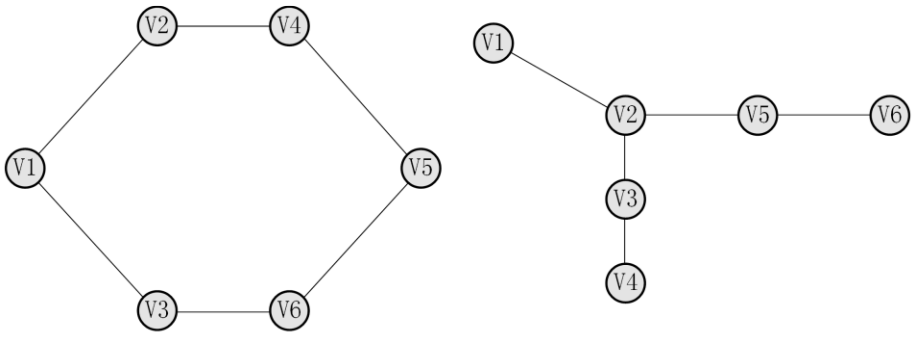


Figure 5-19 The convergency of the Lagrange multipliers in the power systems

5.2. Simulation for Distributed State Estimation

5.2.1. Information Propagation Algorithm in Example Graphs

In order to verify that the proposed information propagation algorithm can share the local value of each vertex to the others, the simulations are conducted in the communication networks with two different graphs as shown in Figure *5-20*, which are the cycle graph (a) and the tree graph (b), respectively [51].



(a) Cycle graph (b) Tree

Figure 5-20 The two different graphs: (a) Cycle graph and (b) Tree

In the two case studies, the vertices $V_1, V_2, \ldots, V_6$ have initial local values 8, 2, 25, 12, 13, and 3, respectively. Then, at time $t = 0.5s$, the local values are turned into 4, 6, 14, 9, 21, and 11. By applying the information propagation algorithm on the graphs, the results are shown in Figure *5-21* and Figure *5-22*.

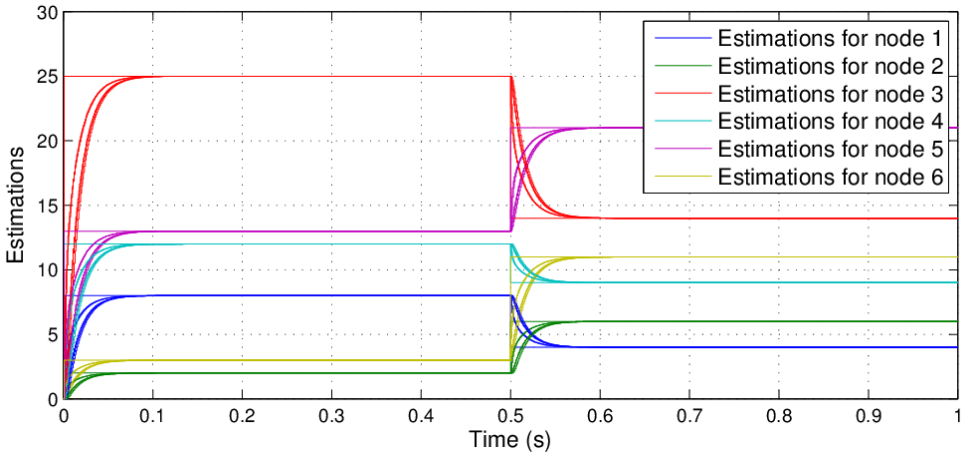


Figure 5-21 The information propagation algorithm in the graph (a)

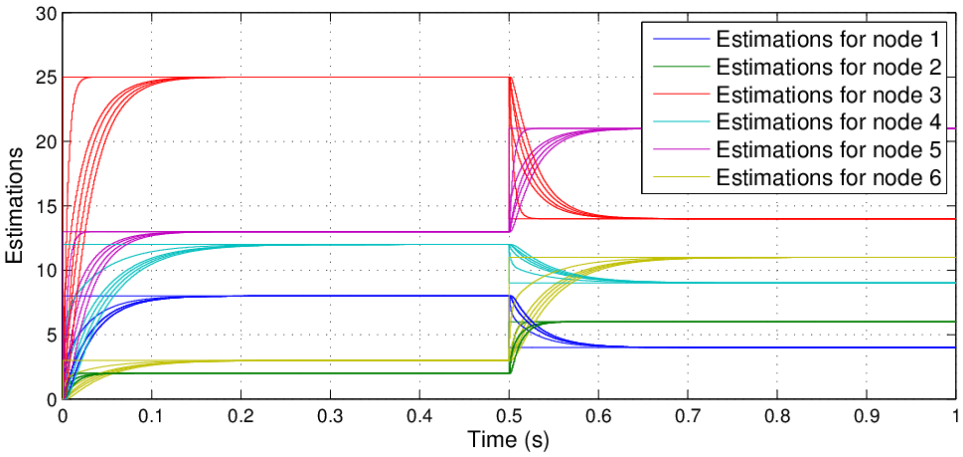


Figure 5-22 The information propagation algorithm in the graph (b)

The results show that local data can be successfully propagated in the graphs by the information propagation algorithm. According to Figure *5-21* and Figure *5-22*, the estimations can converge to the correct values in about $0.15s$.

5.2.2. State Estimation in WSCC 9-Bus Power System

In this work, the WSCC 3-Generator 9-Bus system [52] is adopted to test the proposed DSE method. WSCC 9-Bus system is an approximated model of the Western System Coordinating Council (WSCC) power system at 60Hz, as shown in Figure *5-23*.

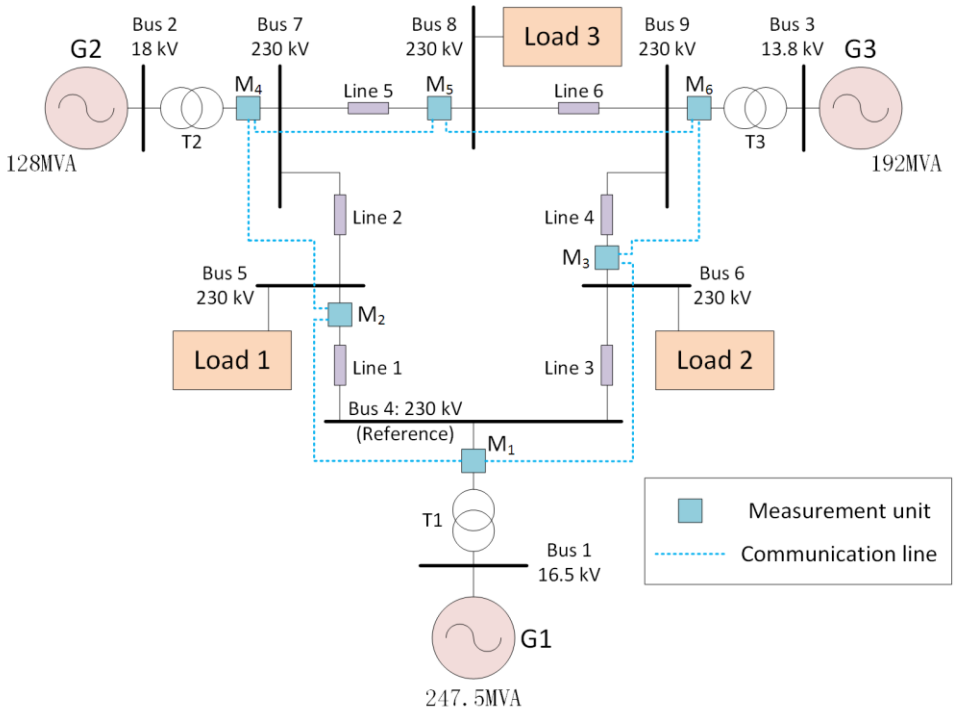


Figure 5-23 WSCC 9-Bus system

The generator G1 and G3 in the power system are Hydro plant and thermal plant, respectively. Considering the rapid development of renewable energy, the generator G2 is modeled as an inverter-based generator. The three transformers T1, T2, and T3 step up the generators' voltages to 230kV for the transmission lines. The transmission lines are Line 1 $(0.01 + j0.085)$, Line 2 $(0.032 + j0.161)$, Line 3 $(0.017 + j0.092)$, Line 4 $(0.039 + j0.17)$, Line 5 $(0.0085 + j0.072)$ and Line 6 $(0.0119 + j0.1008)$ in p.u.. There are three loads, load 1 (187.5MVA), load 2 (153MVA), and load 3 (170MVA), on bus-5, bus-6, and bus-8, respectively. In addition, six meters, M_1, M_2, ..., M_6, are installed to measure the real power on the transmission lines in the system as shown in Figure *5-23*. Each meter is connected with its neighbor meters.

In a 25s simulation, load 1 is increased by 30MW at 5s, while load 3 is decreased by 20MW at 15s. A Gaussian noise with zero mean and $\sigma = 0.01$ are added on the meters. The proposed DSE algorithm (80)-(83) is used to estimate the phase angle on each bus. Figure *5-24* shows the phase angle estimation results by the distributed algorithm executed at node M1. The result is compared with the centralized state estimation and the actual value. The result shows the DSE has less noise than the centralized method. This is because the consensus protocol is a recursive algorithm that has the inherent filtering effect, which reduces the noise.

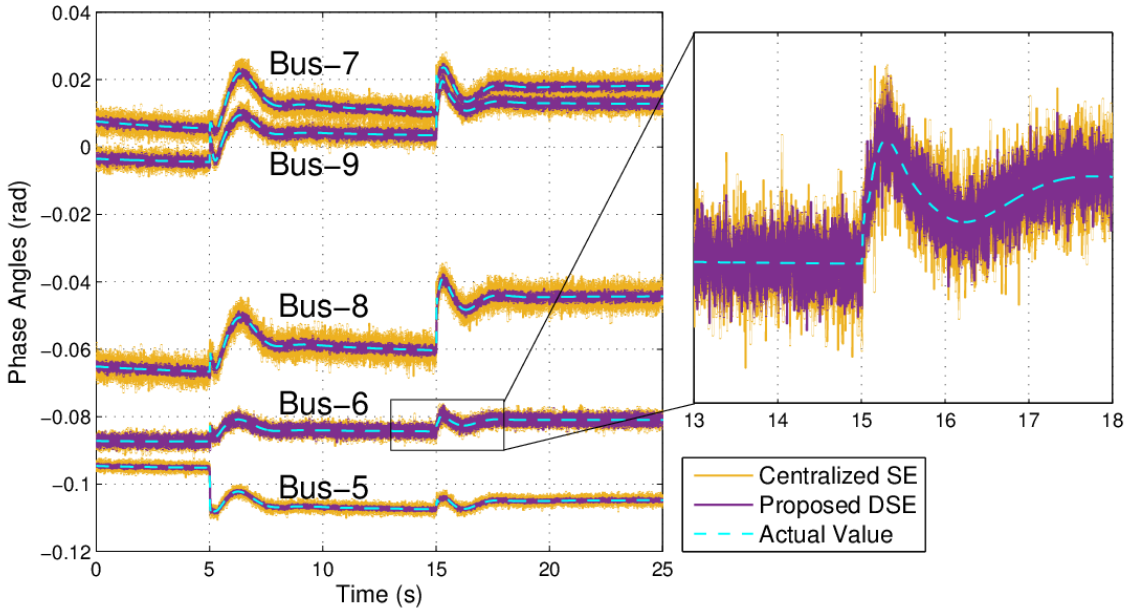


Figure 5-24 Comparison between the actual value, the estimations by the centralized method and by the proposed distributed method at node M1

In order to evaluate the performance of the proposed method, the Mean Squared Error (MSE) is adopted to compute the accuracy of the estimation. The MSE is defined by $MSE = \frac{1}{T}\sum_{t=0}^{T}\left(x(t) - \hat{x}(t)\right)^2$, where $x(t)$ is the actual value; $\hat{x}(t)$ is the estimated value; T is the total time.

The results are shown in Figure *5-25*. In the bar chart, the bars are categorized into five groups, in which each group represents the results for each bus in the power system. Meanwhile, inside each group, the bars show the estimation errors of different meters and methods. Since the proposed DSE algorithm can be performed on every meter, each meter has its own state estimation results. Therefore, the left six bars in each group are the results of the proposed DSE executed on the six different meters, and the last bar represents the result of the centralized method performed by the control center. According to the results, the error of our proposed method is much smaller than the centralized method.

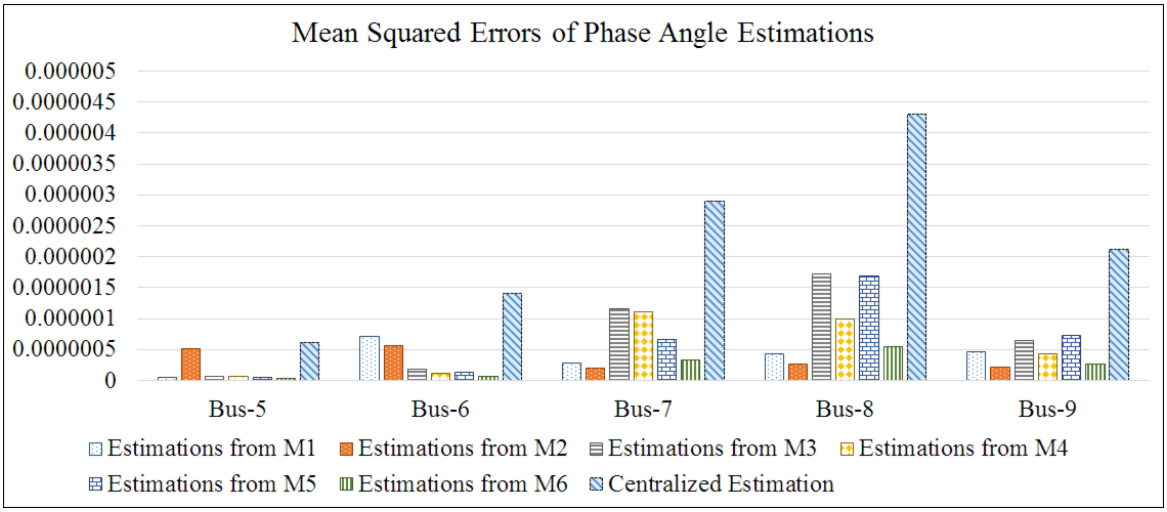


Figure 5-25 Comparison of the MSE between different methods

5.2.3. State Estimation in IEEE 39-Bus Power System

IEEE 39-Bus system is a power system model with 10 generators and 39 buses. Figure *5-26* [53] shows the diagram of the 39-Bus power system. In the system, 37 smart meters running the proposed DSE are installed on different lines. The communication network connects these smart meters as shown in the diagram. The results of the simulation

are shown in Figure *5-27*. According to the results, the proposed DSE is more accurate than the traditional centralized method.

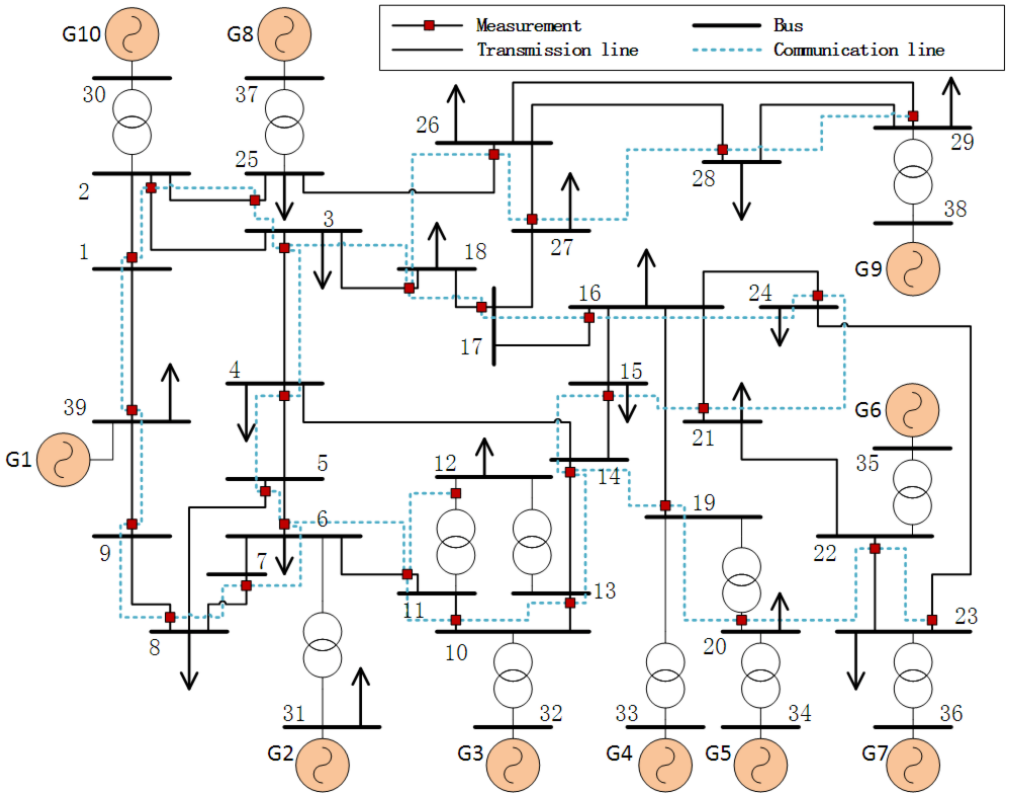


Figure 5-26 IEEE 39-Bus Power System for State Estimation

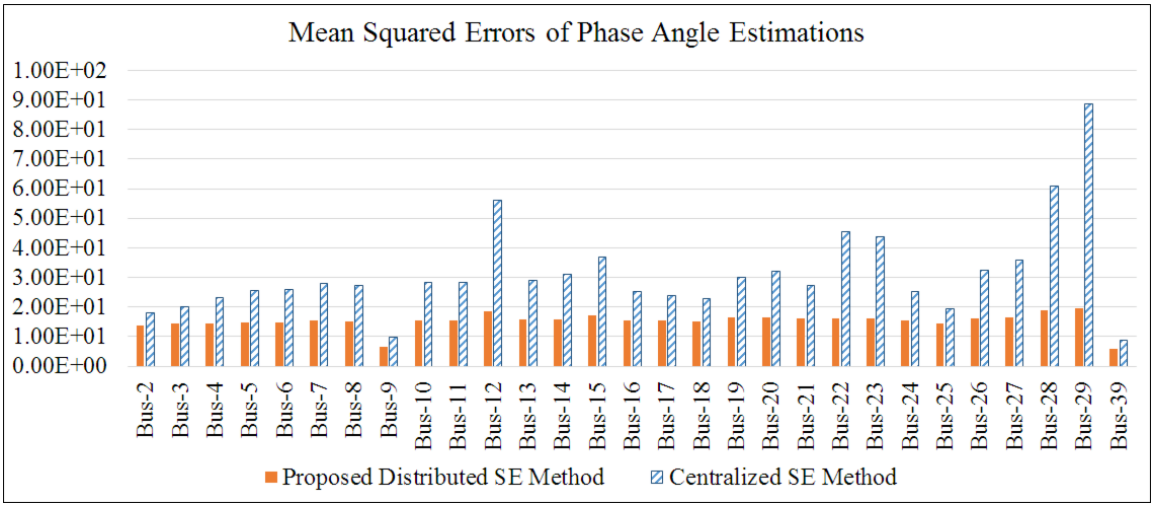


Figure 5-27 MSE of different methods for IEEE 39-Bus system

5.2.4. Improving the Measured Data

On the other hand, the estimated states can be used to improve the accuracy of the measurements by the equation $\hat{z} = H\hat{\theta}$, where $\hat{\theta}$ is the estimated phase angle obtained from the proposed DSE method, and $\hat{z}$ is the estimation of the power flow. By this method, the noise in the raw measurements of the power flow from smart meters can be reduced, thus the measurements are more precise. The case study is done on the 39-Bus power system model. The results of the estimation from the meter M1 in the IEEE 39-Bus system, as an example, are presented in Figure *5-28*.

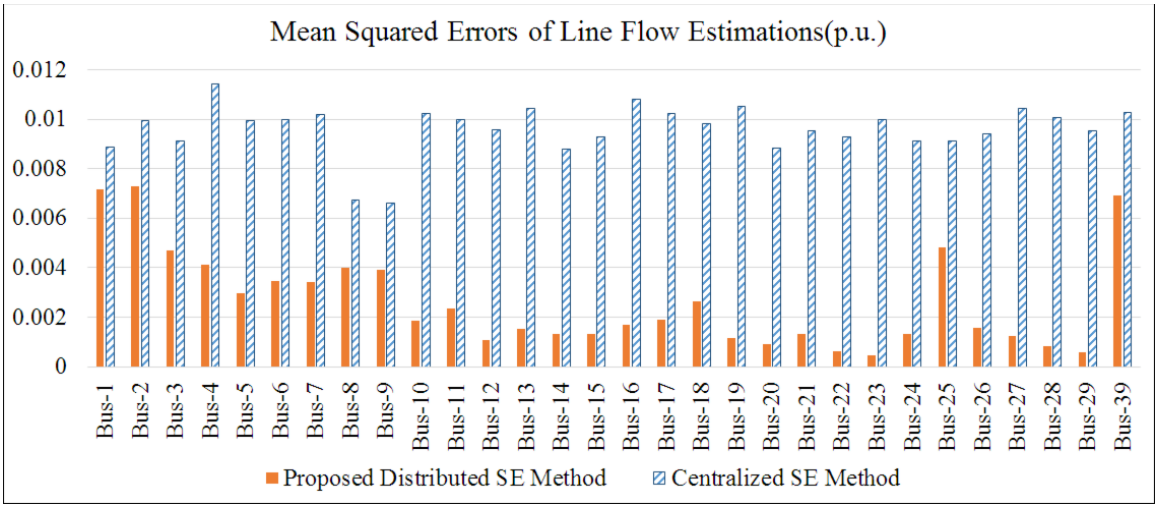


Figure 5-28 MSE of the power flow estimation in M1 in IEEE 39-Bus system

The results of the IEEE 39-Bus system in Figure *5-28* compare the MSEs of the power flow estimation between the proposed DSE method and the traditional centralized SE method. According to the results, the proposed method has a lower MSE for the estimation of each bus, which means that it is much more accurate than the centralized method. The same is true for other meters.

5.2.5. Bad Measurement Detection

Bad measurement detection is one of the most important topics in power system estimation. In a large sensor network, the hardware malfunction or cyber-attack may cause the sensors to provide incorrect readings. So, it is very important to identify and eliminate bad measurements. In this case study, the system runs for 8s and two bad measurements (by increasing the raw measurement reading by 1.0) are injected between 2s-3s and 5s-6s on M1 and M14, respectively. Based on the proposed DSE method, the bad measurement can be detected by the ith meter by calculating the measurement residual $J_i(\hat{\theta}_i)$ [6],

$$J_i(\hat{\theta}_i) = \sum_{j=1}^{n} \frac{(z_{i,j} - \hat{z}_{i,j})^2}{\sigma_{i,j}^2} \tag{129}$$

where $z_{i,j}$ is the jth raw measurements and $\hat{z}_i = H\hat{\theta}_i$ is the estimation of the measurements. $\sigma_{i,j}$ is the standard deviation of the measurement and it is 0.01 in this system. In this 39-Bus system, the number of estimated states $\hat{\theta}_i$ is 29 and the number of measurements is 37, hence the degrees of freedom for the state estimation is $37 - 29 = 8$. The measurement residual $J_i(\hat{\theta}_i)$ satisfies the χ^2 distribution. Then, according to the PDF (Probability Distribution Function) of χ^2 distribution, by choosing 3σ confidence, the threshold value of $J_i(\hat{\theta}_i)$ for the bad measurement detection is 23.3.

In the proposed method, bad data detection can be performed by any meter in the system. In Figure 5-29, the residuals $J_i(\hat{\theta}_i)$ calculated by all meters are presented. The results show that the residuals on all meters exceed the threshold during 2s-3s and 5s-6s. It means that the bad data injection on M1 and M14 are detected.

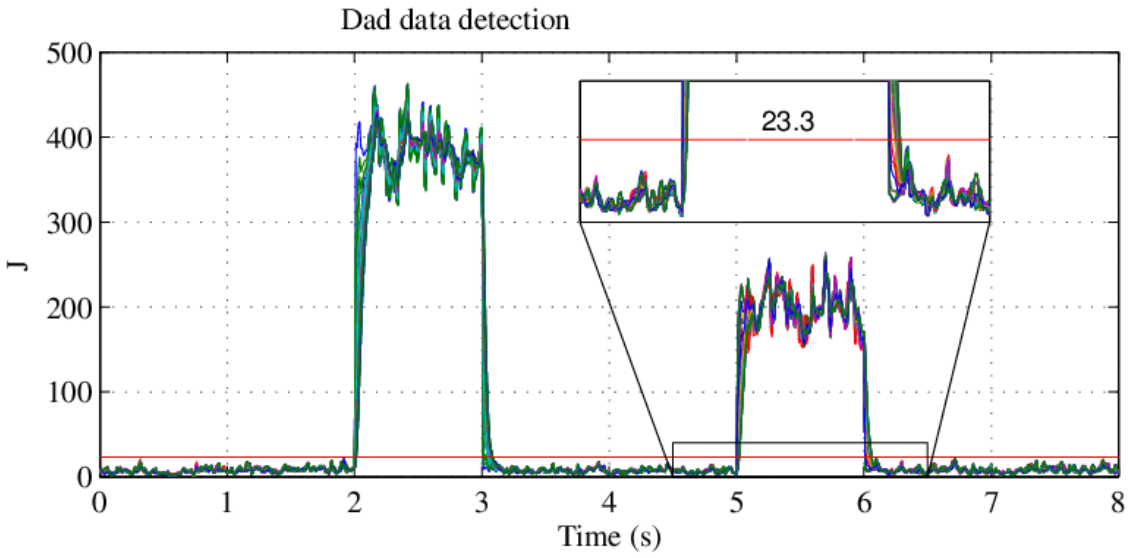


Figure 5-29 The values of the objective function J for bad data detection

5.3. Simulation for Distributed Optimal Power Flow

An IEEE 39-Bus power system model with 10 generators and 39 buses is built in Matlab/Simulink as shown in Figure 5-30 [53]. The 39-Bus system is operated at 60Hz with 19 loads on different buses. The communication network in the power system includes two parts, one network connects all generator controllers together while another network links all smart meters in the system. In addition, each generator connects to a nearby meter to get the power flow information. As shown in Figure 5-30, there is no centralized control center in the power system, so the system is totally distributed. The parameters of the generators are listed in Table 5-2.

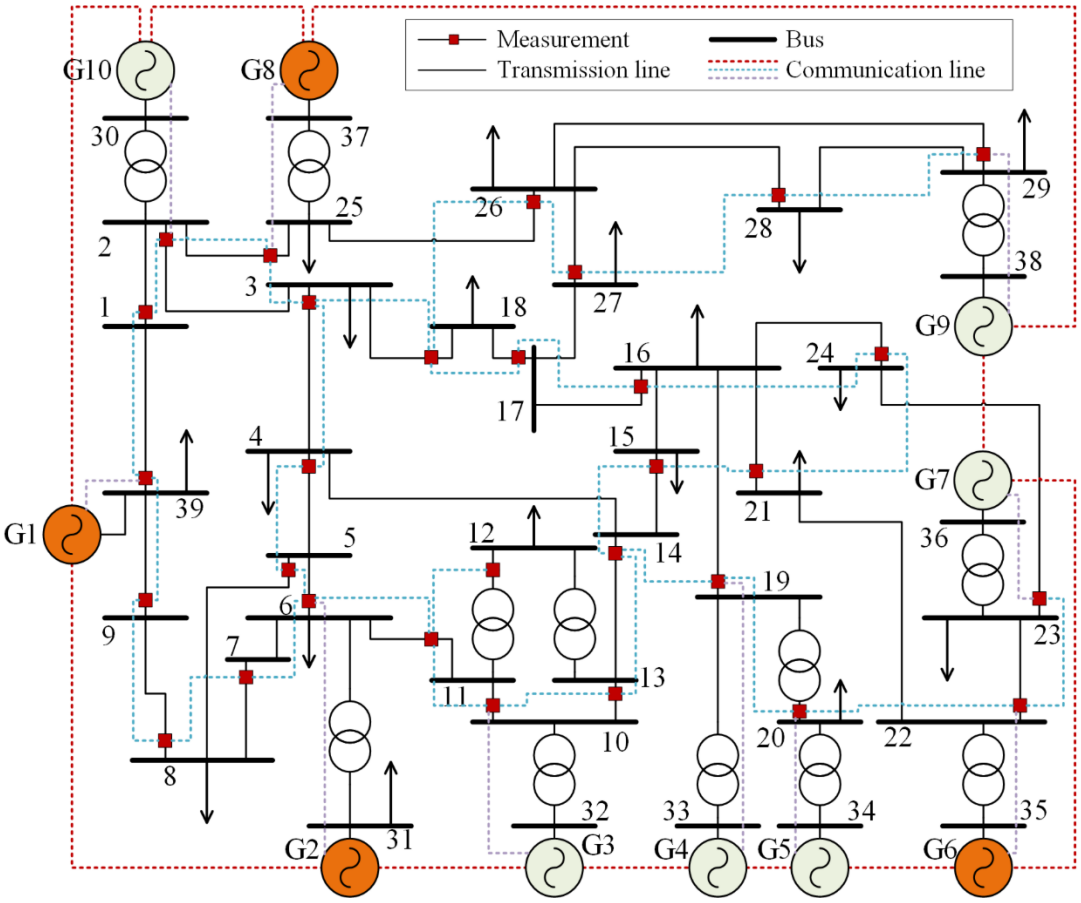


Figure 5-30 IEEE 39-Bus Power System for Optimal Power Flow

In the simulations, two case studies are provided to show the performance of the proposed method.

5.3.1. Overflow on One Transmission Line

In the first case, the limit of the line flow on line 24 is set to 0.8 p.u. To simulate the load variation and overflow, the load on bus 24 increases by 100MW from 5s to 7s, so that line 24 will overflow if no action is taken. The results are shown in Figure 5-31 and Figure 5-32.

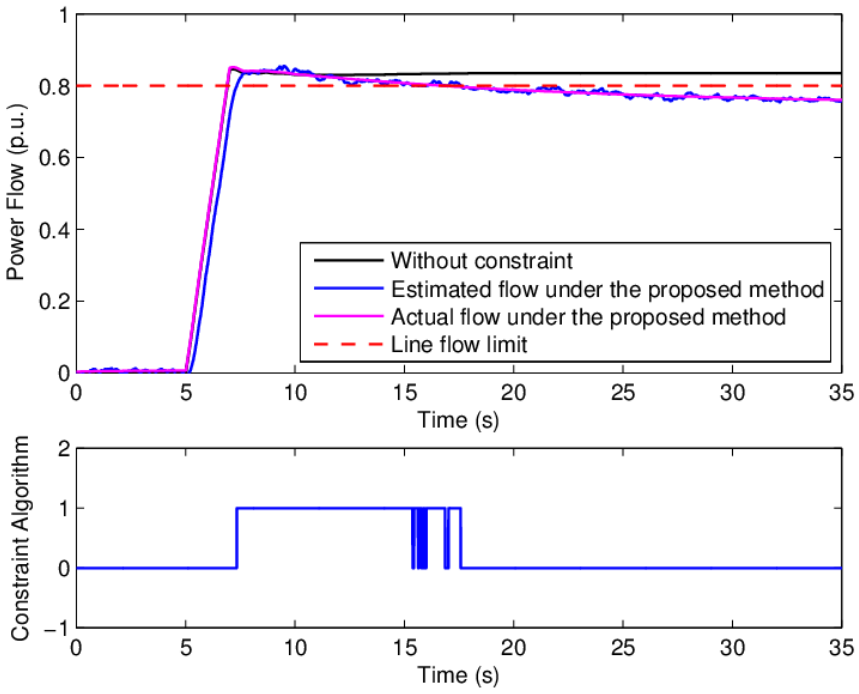


Figure 5-31 Case 1: The line flow on line 24 with 0.8 p.u. limit.

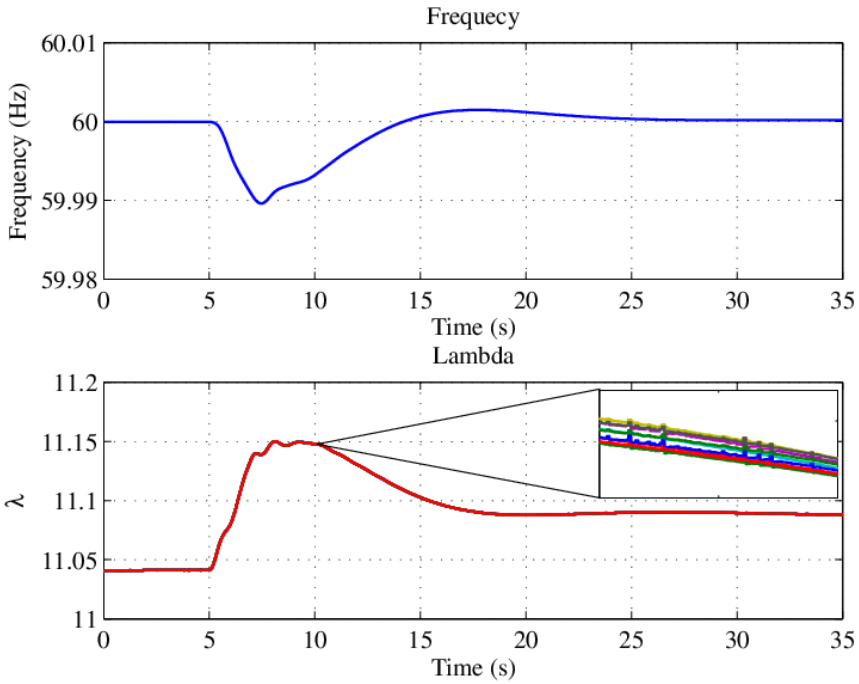


Figure 5-32 Case 1: The Lagrange multiplier and frequency of the system

The top figure in Figure 5-31 shows that, without applying the proposed DDCOPF method, the line flow on line 24 will go beyond the limit and stay at about 0.82. But, with the proposed method, the line flow decreases down below the limit. In the figure, the estimated flow is from the smart meter with the DSE method. The estimation is almost the same as the actual power flow on the line, which means that the DSE method is pretty accurate. The sub-figure at the bottom of Figure 5-31 shows when the constraint algorithm is activated. In Figure 5-32, the frequency of the system returns to 60Hz after the load variation and the Lagrange multipliers (Lambdas) are identical for all generators, which means that the optimal solution of the ED problem with power balance constraint as described in equation (7)-(9) is achieved.

5.3.2. Overflow on Two Transmission Lines

In another case study, the scenario in which two lines in the power system are overflowed at the same time is simulated. In this case, line 24 and line 27 exceed their maximal line flow, i.e., 0.8 p.u. for line 24 and 1.4 p.u. for line 27, after the load increases on bus 24. The simulation results are shown in Figure 5-33, Figure 5-34, and Figure 5-35. Similar to the previous case, the line flow on both line 24 and line 27 decreases to the feasible region. Also, the frequency and Lagrange multiplier show that the system operates in the optimal condition.

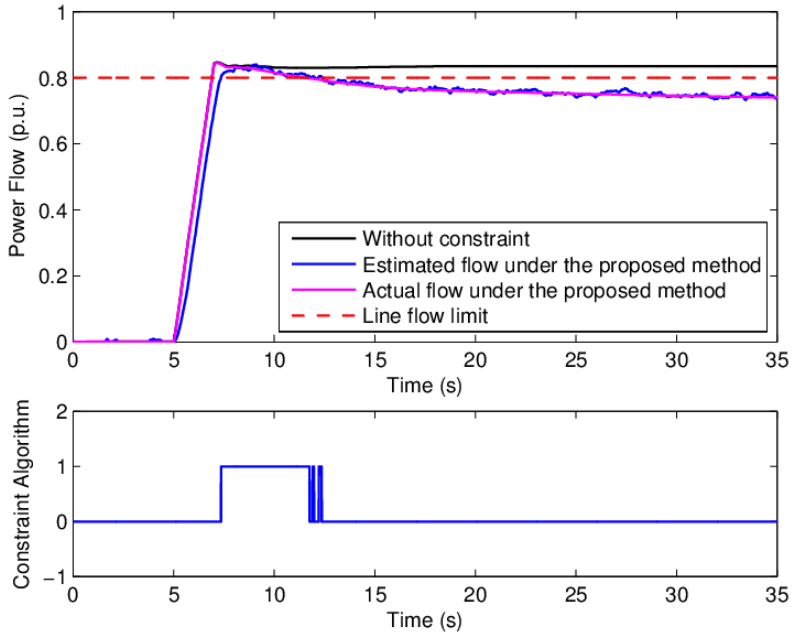


Figure 5-33 Case 2: The line flow on line 24 with 0.8 p.u. limit.

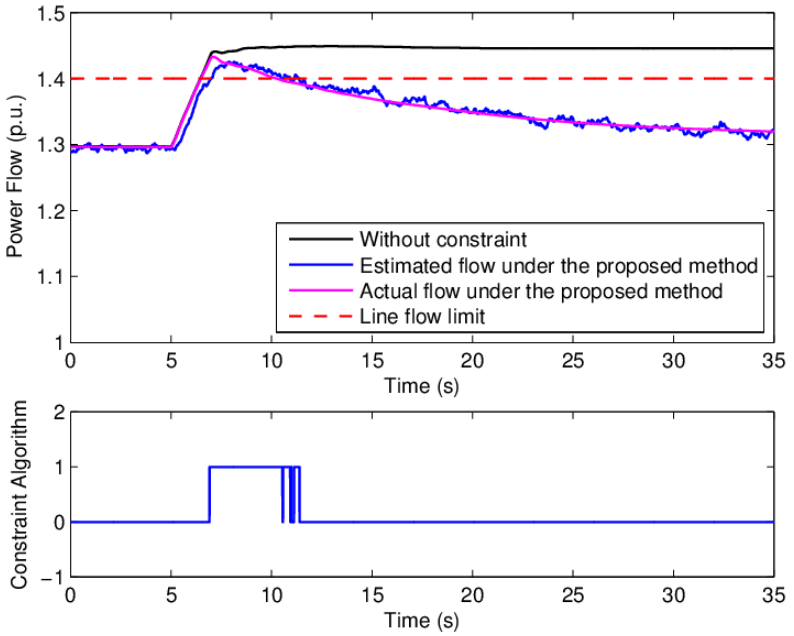


Figure 5-34 Case 2: The line flow on line 27 with 1.4 p.u. limit.

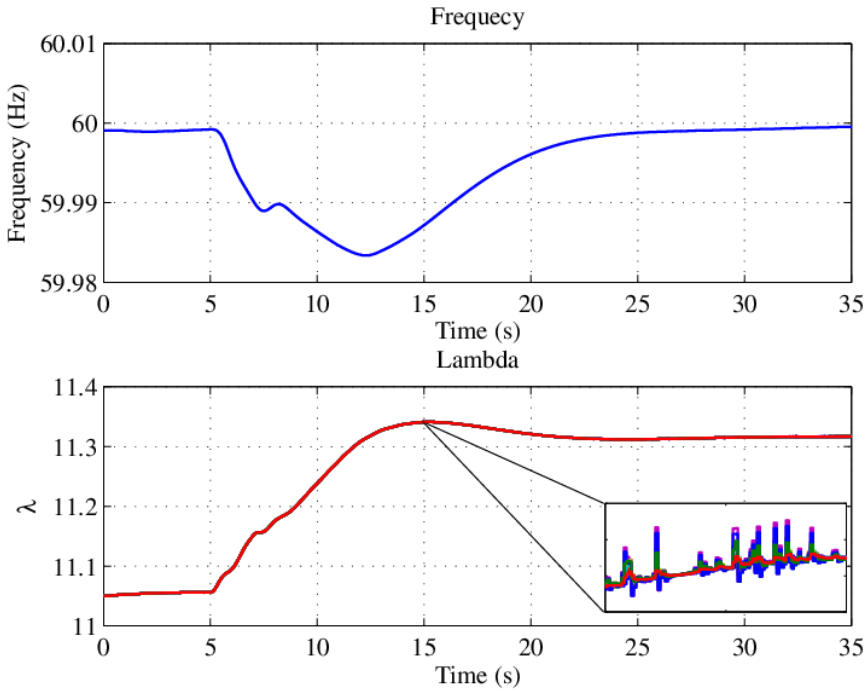


Figure 5-35 Case 2: The Lagrange multiplier and frequency of the system

The simulation runs on a PC with Intel i7-7700HQ 2.8GHz CPU and 8GB memory. The time consumed by each iteration of the algorithm is less than 0.000002s. So, the algorithm does not require a very powerful computation device and is valid for real-time running.

5.4. Simulation for Distributed AC State Estimation

In this section, the proposed Distributed AC state estimation in section 4.4.2 is tested in MATLAB on IEEE 14-bus system, IEEE 39-bus system, IEEE 118-bus system, and IEEE 300-bus system. In addition, a case study over IEEE 118-bus system with different communication network topology is provided to show the impact of the network structure on the performance of the AC DSE.

5.4.1. Case study on IEEE 14 Bus System

The structure of the IEEE 14 bus system is shown in Figure 5-36. In this system, 17 DSE units are installed on the power lines. The dash lines in the figure are the communication lines that connect the DSE units.

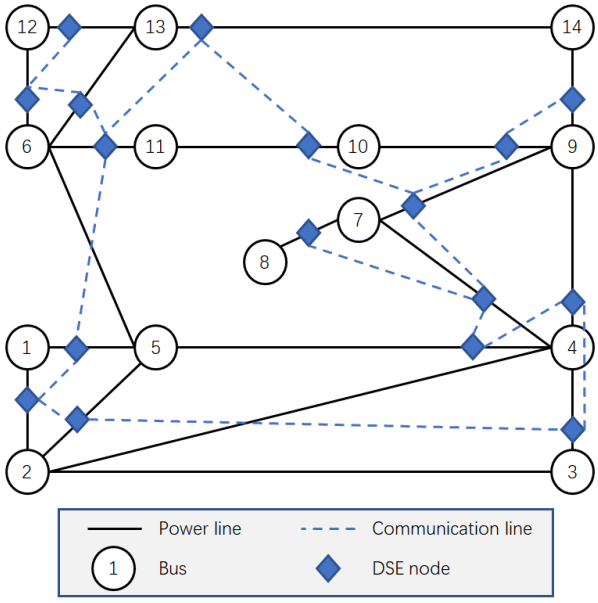


Figure 5-36 The IEEE 14 bus system

The trajectory of the estimated states (the phase angles and voltage magnitudes of the 14 buses) by the proposed AC DSE method is shown in Figure 5-37. The figure shows that the estimated states reach the actual values in about 200 iterations. For the simulation on a PC with Intel Core i7 2.8Ghz CPU and 24GB memory, each iteration takes about 0.001s, so the 200 iterations take about 0.2s which is acceptable compare with the 15 minutes interval between two state estimation in a centralized state estimation method.

The proposed AC DSE method is also compared with the traditional centralized state estimation method [6]. The estimated phase angles and voltage magnitudes from node 1 by the proposed AC DSE are compared with the results from the centralized method as

shown in Figure 5-38. In the results, the phase angle is in radian and the voltage magnitude is in p.u.. The MSE (Mean squared error) and MAE (Mean absolute error) of each method are listed in Table 5-3.

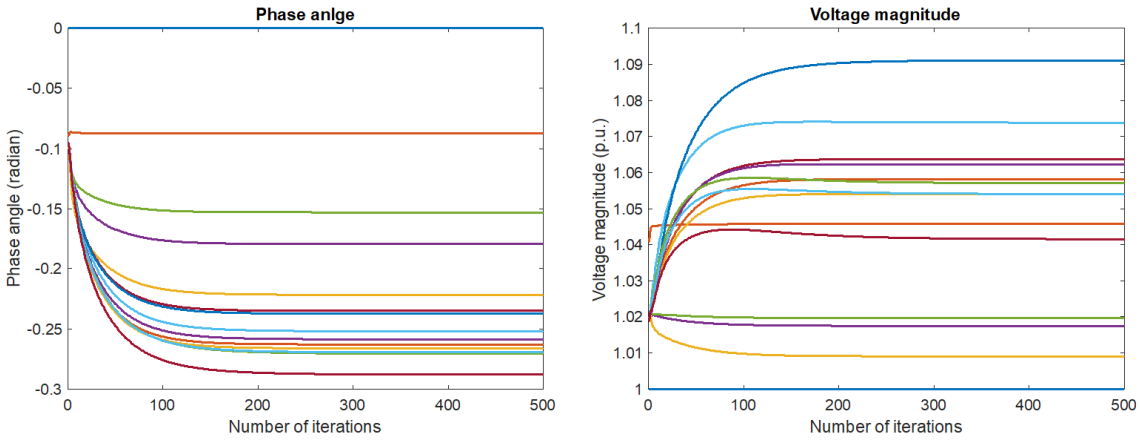


Figure 5-37 The convergence of the AC DSE on IEEE 14 bus system

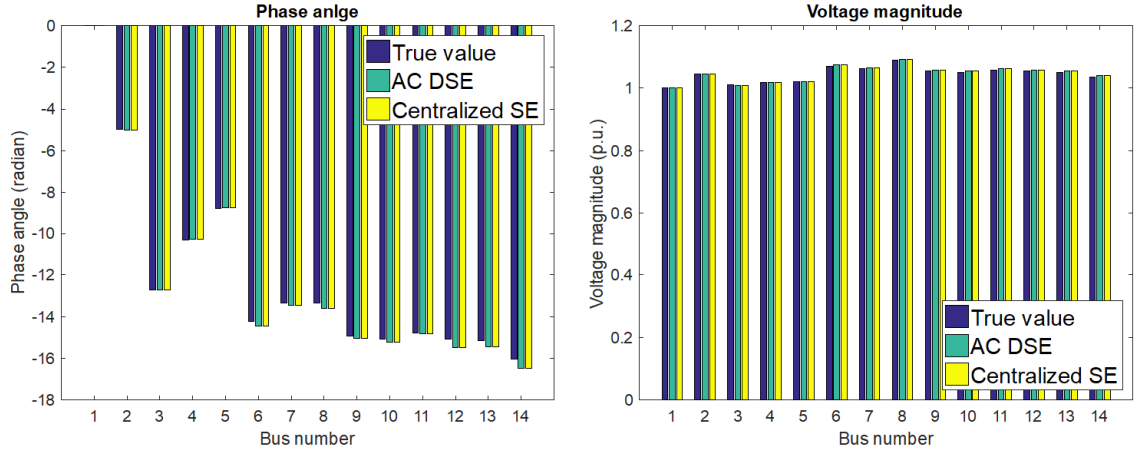


Figure 5-38 Comparison of centralized SE and AC DSE on IEEE 14 bus system

Table 5-3 The Errors of State Estimation Methods on IEEE 14 Bus System

Methods	Centralized SE	Proposed AC DSE	Difference
Phase angle MAE (rad)	0.0016514	0.0016535	0.12716%
Voltage MAE (p.u.)	0.0007402	0.00074041	0.028371%
Phase angle MSE (rad)	4.6457×10^{-6}	4.6579×10^{-6}	0.26261%
Voltage MSE (p.u.)	1.065×10^{-5}	1.0618×10^{-5}	−0.30047%

According to the above results, the accuracy of the proposed AC DSE method is similar to the centralized state estimation method in the IEEE 14-bus system.

5.4.2. Case Study on IEEE 39 Bus System

The IEEE 39-bus system as shown in Figure 5-39 is adopted in the simulation. There are 46 DSE nodes in the system and the communication network is represented as the dash lines in the figure. The results of this simulation case are listed in Table 5-4.

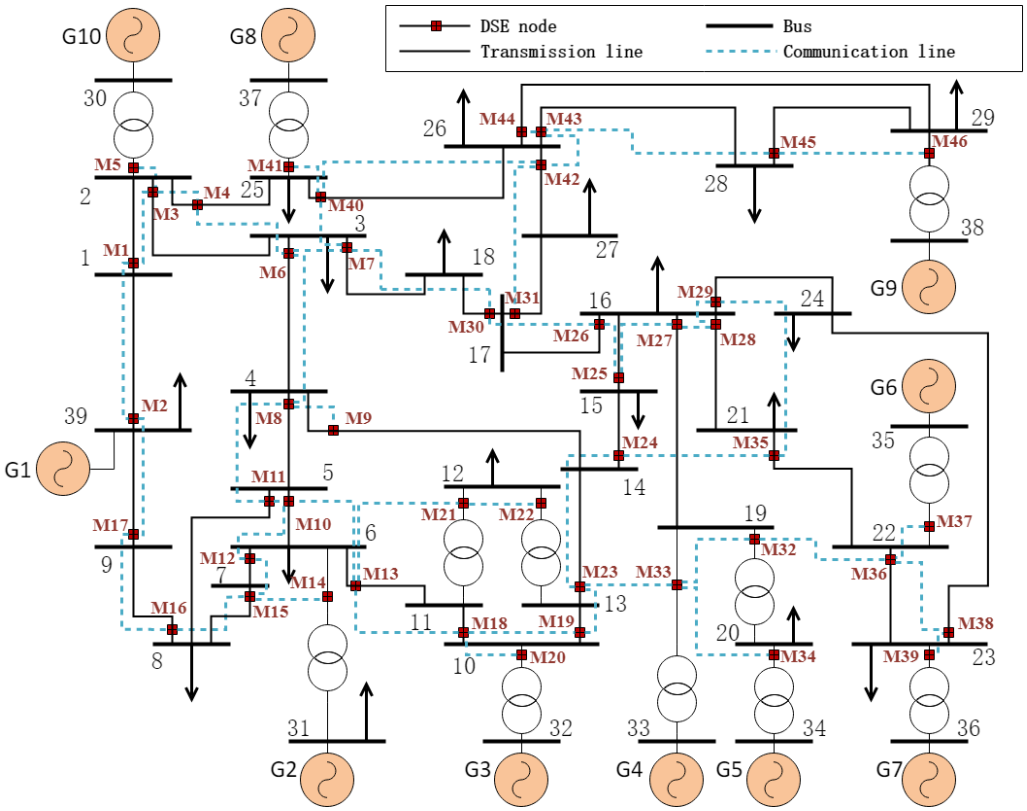


Figure 5-39 The IEEE 39 bus system

Table 5-4 The Errors of State Estimation Methods on IEEE 39 Bus System

Methods	Centralized SE	Proposed AC DSE	Difference
Phase angle MAE (rad)	0.00018428	0.00018409	−0.1031%
Voltage MAE (p.u.)	0.00019664	0.00019667	0.015256%
Phase angle MSE (rad)	5.5541×10^{-8}	5.5518×10^{-8}	−0.041411%
Voltage MSE (p.u.)	5.0927×10^{-8}	5.0954×10^{-8}	0.053017%

5.4.3. Case Study on IEEE 118 Bus System

To verify the performance of the proposed method in a big power system, the proposed method is also simulated in the IEEE 118-bus power system model [54]. There are 179 DSE nodes in the system. The results are shown in Table III.

Table 5-5 The Errors of State Estimation Methods on IEEE 118 Bus System

Methods	Centralized SE	Proposed AC DSE	Difference
Phase angle MAE (rad)	0.0010439	0.0010585	1.3986%
Voltage MAE (p.u.)	0.0007395	0.00073666	−0.38404%
Phase angle MSE (rad)	1.6417×10^{-6}	1.6991×10^{-6}	3.4964%
Voltage MSE (p.u.)	8.2586×10^{-7}	8.2088×10^{-7}	−0.60301%

5.4.4. Case Study on IEEE 300 Bus System

Finally, the proposed method is tested in the IEEE 300-Bus system [55] and compared with the centralized state estimation method. There are 409 DSE nodes in the system. The results are shown in Table 5-6. The errors of the DSE method are the same as the centralized method in this case.

Table 5-6 The Errors of State Estimation Methods on IEEE 118 Bus System

Methods	Centralized SE	Proposed AC DSE	Difference
Phase angle MAE (rad)	0.0032567	0.0032567	≈ 0%
Voltage MAE (p.u.)	0.002517	0.002517	≈ 0%
Phase angle MSE (rad)	0.00013898	0.00013898	≈ 0%
Voltage MSE (p.u.)	0.0001109	0.0001109	≈ 0%

According to the results from the different cases above, the errors of the centralized state estimation and the proposed AC DSE are very close. This result confirms that the proposed AC DSE method has similar accuracy as the centralized method.

5.4.5. The influence of the communication network structure

The performance of the DSE method is usually related to the topology of the communication network. In this case study, the IEEE 118-bus system is used to demonstrate the impact of the communication network structure on the convergence of the estimation. Two different communication network structures are investigated in this study: One is chain connection and another is fully connected. The chain connection is that each node connects to its previous node if it is not the first node and connects to the next node if it is not the last node. For example, for the 179 DSE nodes in the IEEE 118-bus system, node 1 connects with node 2, node 2 connects to node 1 and node 3, node 3 connects to node 2 and node 4, and goes on until node 179 only connects to the node 178. On the other hand, the full connection means that each node connects to all other nodes in the system. The chain connection structure has the fewest communication lines and the full connection structure has the most communication lines.

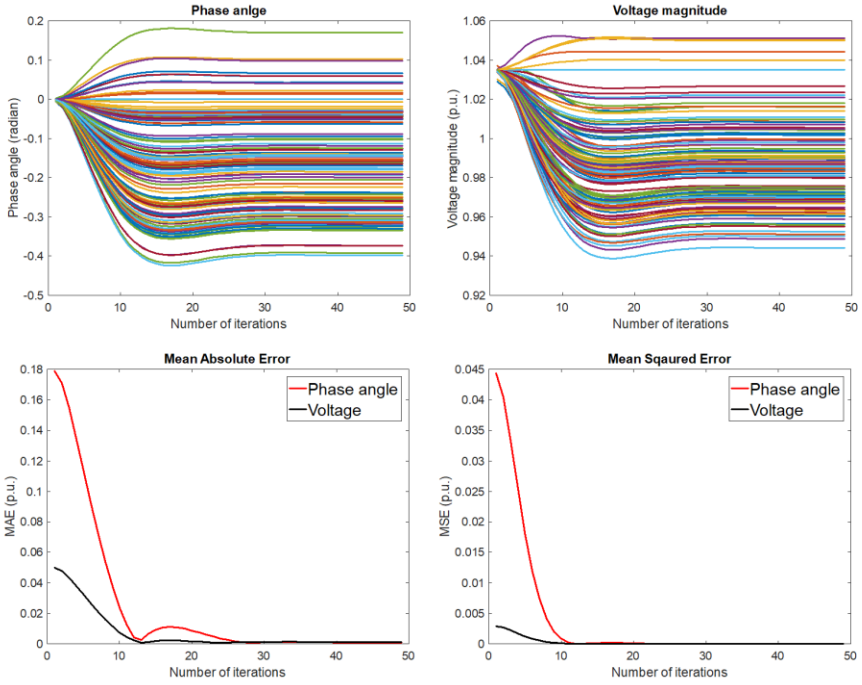


Figure 5-40 The results of fully connected communication network

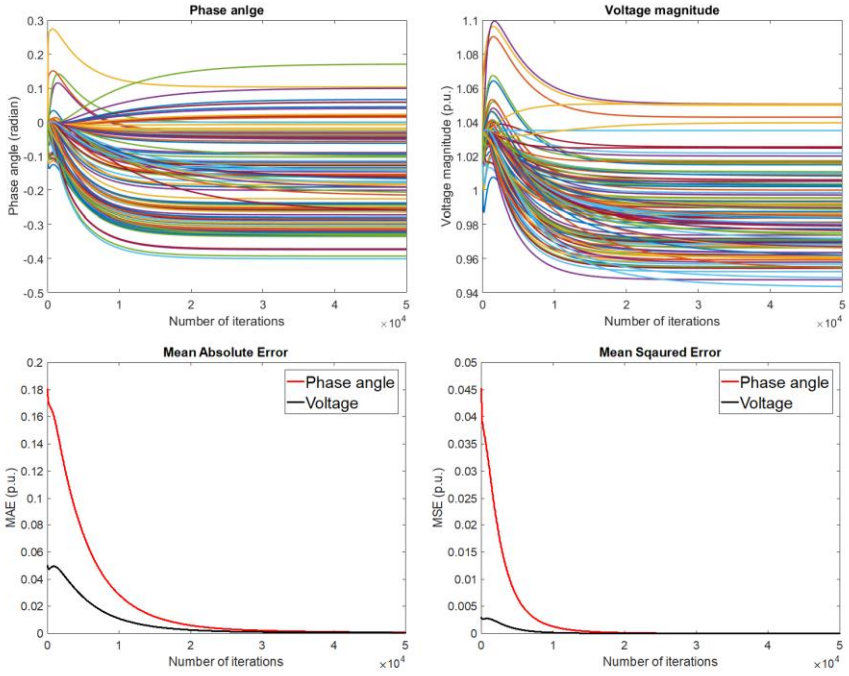


Figure 5-41 The results of chain connected communication network

The simulation results of the two communication network structures in the IEEE 118-bus system are shown in Figure 5-40 and Figure 5-41. The figures show the trajectory of the estimated states and the errors of the results. According to the results, the state estimation in the fully connected communication structure is much fast than in the chain-connected system. The chain-connected system spends about 40000 iterations to reach an accurate result, but the fully connected system only takes 30 iterations. Since the chain-connected structure and the fully connected structure are the extreme cases, other communication structures will have the speed between these two results. In addition, although the chain-connected system takes 40000 iterations to reach the acceptable result, the speed of the algorithm can be improved in the real world by properly setting the initial values of the estimated states with the previously estimated results.

5.5. Simulation for Distributed AC Optimal Power Flow

In this section, the proposed distributed ACOPF method is tested on IEEE 9 bus system and IEEE 39 bus system.

5.5.1. Case Study on IEEE 9 Bus System

The first case is the ACOPF over the IEEE 9-bus system. There are 9 DSE units and 3 ACOPF units in the system. The 9 DSE units are installed on the transmission lines between buses.

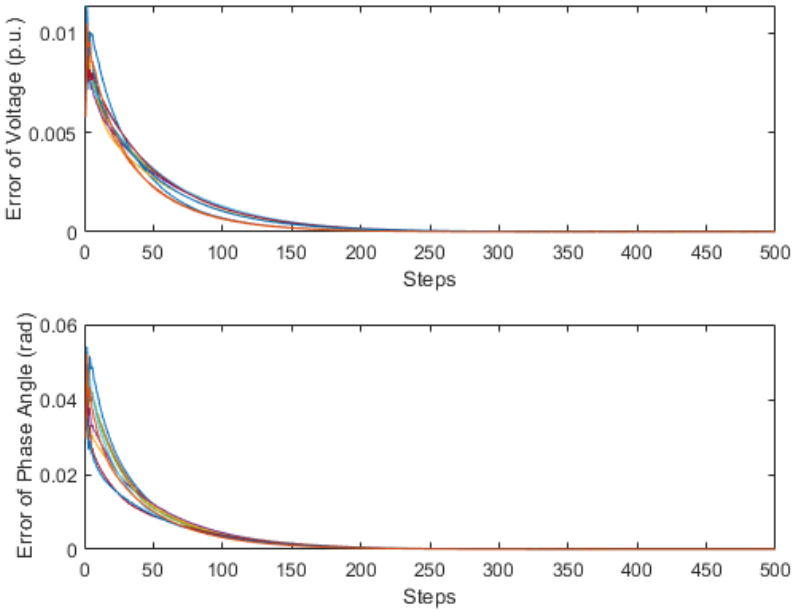


Figure 5-42 The errors of DSE in the IEEE 9 Bus System

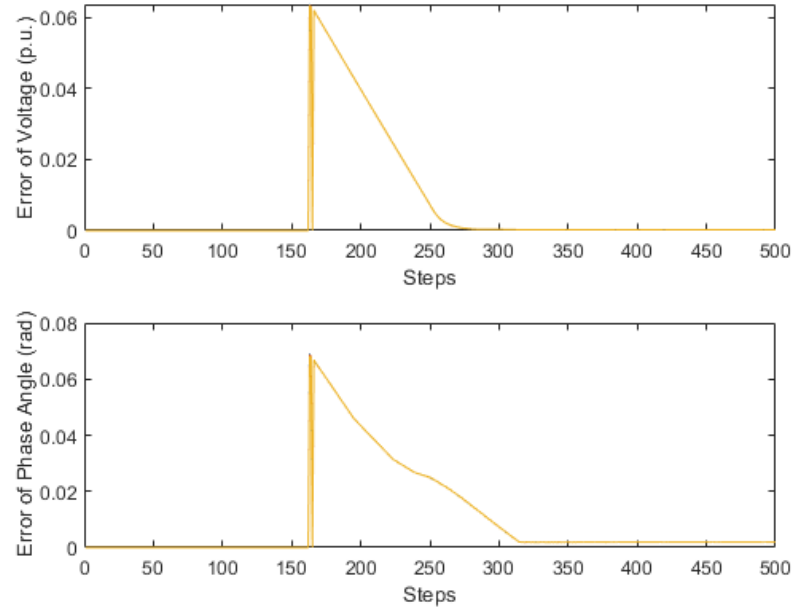


Figure 5-43 The errors of the ACOPF in the IEEE 9 Bus System

The result of the distributed ACOPF over the IEEE 9 bus system is shown in Figure 5-42 and Figure 5-43. The error of distributed state estimation over the IEEE 9 bus system

is shown in Figure 5-42. As shown in the results, the estimated power system states from the DSE convergences to the real value at around 200 steps. Meanwhile, The error of distributed ACOPF over the IEEE 9 bus system is shown in Figure Figure 5-43. It shows that the ACOPF result converges at around 300 steps. The errors between the results of the proposed distributed ACOPF algorithm and the results of the traditional centralized ACOPF algorithm are as shown in Table 5-7. It shows that the proposed distributed method has a similar performance as the traditional centralized method.

Table 5-7 The errors of ACOPF results

	MSE (p.u.)
Phase angle (rad)	1.933×10^{-3}
Voltage (p.u.)	2.385×10^{-4}

In addition, Figure 5-44 shows the line flows on transmission line 4 computed by the proposed distributed ACOPF algorithm. For the result without constraint in the figure, it means the power flow on the transmission line is not limited. The result with constraint shows the power flow of the transmission line 4 when there is a 1.4 p.u. limit applied on it. The results show that the proposed distributed ACOPF algorithm can limit the line flow of transmission line 4 to the upper limit effectively.

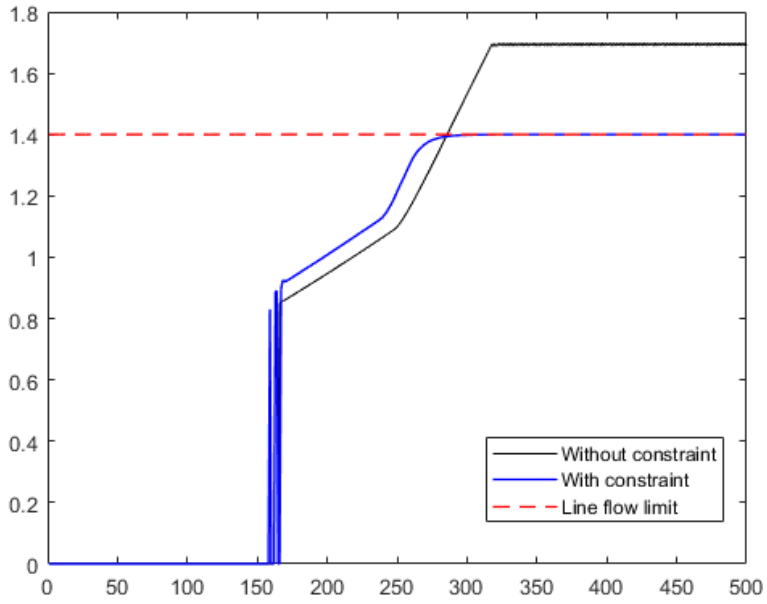


Figure 5-44 The line flows on transmission line 4 with and without the constraint

5.5.2. Case Study on IEEE 39 Bus System

In the IEEE 39-bus system, there are 46 transmission lines in total. On each transmission line, a DSE unit is installed. Each DSE unit connects to three other DSE units in the network. Meanwhile, each generator in the system is equipped with an ACOPF unit. These ACOPF units are connected with the nearby DSE units to obtain the state estimation data.

The results of the AC DSE on the IEEE 39 bus system are shown in Figure 5-45. The errors in this plot are the average absolute error between the actual states and the estimated states on each DSE unit. The actual states are obtained by running the centralized state estimation on the system with error-free data input. The plot shows that the error of the estimated states is continuously decreasing.

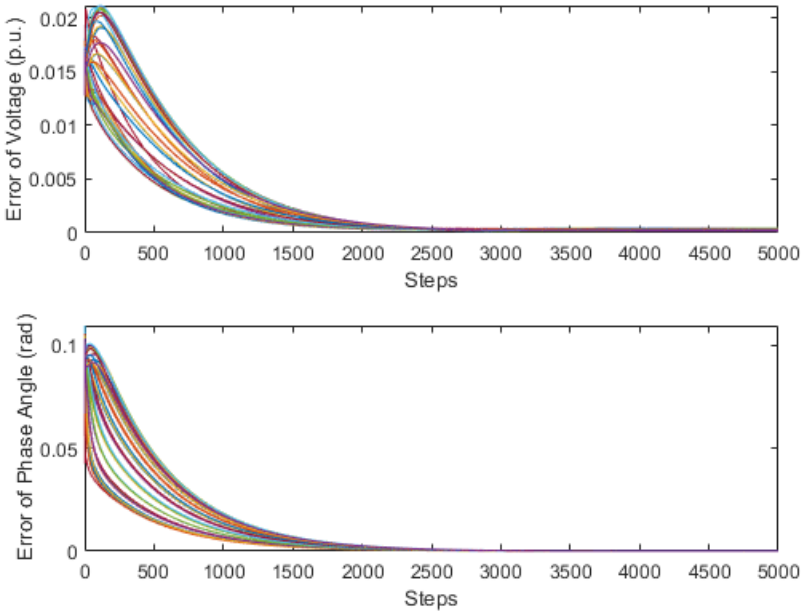


Figure 5-45 The errors of DSE in the IEEE 39-bus system

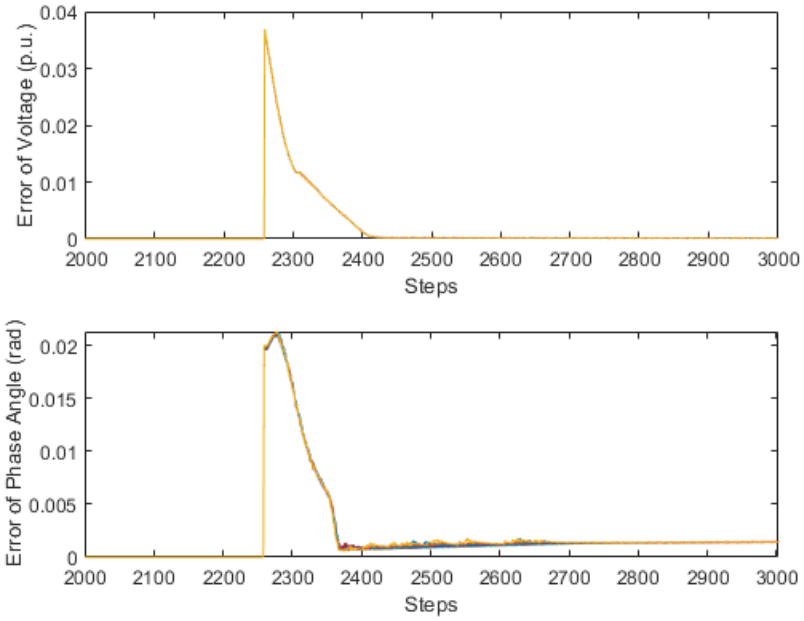


Figure 5-46 The errors of the distributed ACOPF in the IEEE 39-bus system

The results of the ACOPF are shown in Figure 5-46, in which the result shows the average error of optimized voltage and the average error of the optimized phase angle on

the buses of the system. The errors are the difference between the results (in p.u.) of the centralized ACOPF method and the proposed distributed ACOPF method. It shows that the error of the proposed distributed ACOPF gets very small after 2400 steps. As shown in section 5.4.5, if there are more connections between the DSE units, the state estimation will be much faster.

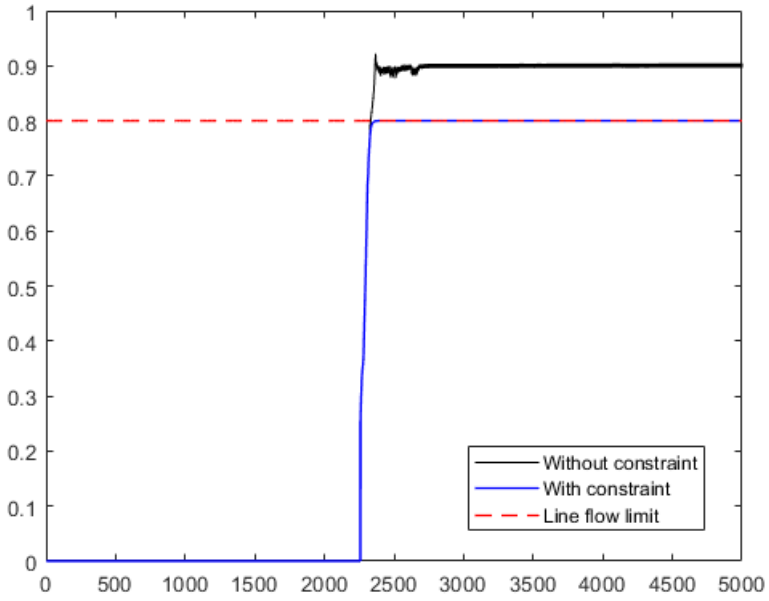


Figure 5-47 The line flows on transmission line 31 with and without constraint

The case study of the line flow limit constraint is also given in Figure 5-47. The power flow on transmission line 31 is limited to 0.8 p.u.. The result shows that the proposed distributed ACOPF algorithm can ensure the line flow limit constraint is satisfied. Meanwhile, compared with the results in section 5.3, the distributed ACOPF algorithm handles the line flow limit much better than the distributed DCOPF algorithm.

Finally, the results of these case studies show that the errors between the proposed distributed ACOPF method and the centralized ACOPF method are very close, which

means that the proposed distributed ACOPF method has a similar performance as the centralized method. In addition, since the distributed method is more robust than the centralized method under communication disconnection, the proposed distributed ACOPF is more suitable for power systems with high renewable DER penetration.

Chapter 6. Research Conclusion and Future Work

6.1. Conclusion

This research focuses on the three important topics in power systems, i.e., economic dispatch, state estimation, and optimal power flow. In order to handle the distributed characteristics of renewable power systems, the distributed economic dispatch, distributed state estimation and distributed optimal power flow methods are developed on the basis of consensus protocol technique.

For the distributed economic dispatch, the distributed algorithm is developed to minimize the power system's operational cost without the centralized control center. To prove the versatility of the proposed approach, both the PI controller and NN controller are utilized to design the distributed algorithm. In addition, a communication-failure-tolerant distributed economic dispatch is provided to ensure the optimal dispatch during serious communication failure. Also, the line loss in the power system is considered to achieve a more detailed optimization. In the simulation, the proposed methods are tested in a 4-generator system and the IEEE 10-generator 39-bus system. Also, a sensitivity analysis is provided to show that the proposed method works well in power systems with different configurations.

For the state estimation problem, an information propagation algorithm is developed to share the measured data in the distributed system. With the information propagation algorithm, the distributed state estimation is established. The proposed method

is also tested in a WSCC 9-Bus system and the IEEE 39-bus system in simulation. The bad data detection and the measurement accuracy improvement features of state estimation are also verified in the simulations. Also, the distributed AC state estimation method is developed based on the AC power flow model to get a more accurate result in the estimation. The proposed method has been tested in several IEEE standard power system models to demonstrate its performance in power systems with different sizes and structures.

In the end, the distributed optimal power flow is developed on the basis of the proposed distributed economic dispatch and distributed state estimation methods. The line flow constraint is satisfied by the specially designed update term and penalty term in the optimization algorithm. The proposed method is tested in the IEEE 39-bus system with two simulation cases. In addition, the distributed AC optimal power flow scheme is provided to achieve a better optimization result. The simulations show that the proposed distributed AC optimal power flow achieves similar performance as the centralized AC optimal power flow method but is more robust due to its decentralized structure.

6.2. Future research

The next step of the reaserch is to improve the distributed ACOPF scheme. The current algorithm is a combination of two algorithms, i.e., distributed state estimation (DSE) and distributed ACOPF. However, the two algorithms are running separatedly, that is the DSE algorithm runs first and the distributed ACOPF starts to run after the DSE converges. So, the algorithm is not most efficient since the ACOPF algorithm is not optimizing before DSE converges. To improve the efficiency, the next version of the distributed ACOPF scheme will integrate the two algorithms together, in which the ACOPF algorithm starts to

run from the very beginning, so it will not waste the time to wait until the DSE to converge. To achieve the integration of DSE and distributed ACOPF, the framework of the entire distributed ACOPF scheme should be re-designed to ensure the algorithms converges without stuck at the local minimum.

In addition, more distributed optimization methods for power systems will be studied. Since the existing distributed optimization methods are majorly based on area partitioning methods, e.g., the ADMM technique. Then, they are not fully distributed. As shown in this paper, consensus-based methods can realize fully distributed optimization. So, consensus-based methods will play an important role in distributed power systems in the future. The author will continue to work on the research of distributed optimization for power systems with different objective functions and constraints. Also, the author will study the interesting features of the distributed control and management system, e.g., plug-and-play of the nodes and parallel computing over the nodes.

On the other hand, as mentioned in the introduction, the major motivation of the research of distributed control and management system is due to the advancement of renewable energy technology. So, the distributed control and optimization for renewable sources will be researched in the future. In a published paper by the author, an optimal controller for wind turbines is proposed. It shows better performance than the traditional methods. This control method will be integrated with the proposed distributed economic dispatch and distributed optimal power flow methods in the future. Also, to optimize the renewable power system, the different possible operational scenarios of the renewable generators should be considered. However, due to the massive number of possible

scenarios, the scenario set reduction is usually required. The author is currently working on a paper to address the scenario reduction problem for PV power with a deep learning technique. At the next step, the new scenario reduction methods will be integrated with the distributed optimal power flow to achieve a better result in the renewable power system.

List of Publications

Qiao Li, David Wenzhong Gao, Huaguang Zhang, Ziping Wu, Fei-Yue Wang, "Consensus-Based Distributed Economic Dispatch Control Method in Power Systems", IEEE Transactions on Smart Grid, vol. PP, no. 99, pp. 1-1. 2017.

Qiao Li, Lin Cheng, Wei Gao, and David Wenzhong Gao, "Fully Distributed State Estimation for Power System with Information Propagation Algorithm," in Journal of Modern Power Systems and Clean Energy, vol. 8, no. 4, pp. 627-635, July 2020.

Qiao Li, Tianqi Gao, David Gao, Xiao Wang, "Adaptive LQR Control with Kalman Filter for the Variable-Speed Wind Turbine in Region II", 2017 North American Power Symposium (NAPS), Morgantown, WV, 2017, pp. 1-6.

David Wenzhong Gao, Qiang Wang, Fang Zhang, Xiaojing Yang, Zhigang Huang, Shiqian Ma, **Qiao Li**, Xiaoyan Gong, and Fei-Yue Wang, "Application of AI techniques in monitoring and operation of power systems." Frontiers in Energy 13.1 (2019): 71-85.

Qiang Wang, Xiaojing Yang, Zhigang Huang, Shiqian Ma, **Qiao Li**, David Wenzhong Gao, Fei-Yue Wang, "A novel design framework for smart operating robot in power system," in IEEE/CAA Journal of Automatica Sinica, vol. 5, no. 2, pp. 531-538, Mar. 2018.

Qiao Li, David Wenzhong Gao, etc., "Fully Distributed DC Optimal Power Flow Based on Distributed Economic Dispatch and Distributed State Estimation," *arXiv preprint*. arXiv:1903.01128

Qiao Li, David Wenzhong Gao, "Fast Scenario Reduction for Power Systems by Deep Learning," *arXiv preprint*. arXiv:1908.11486

Qiao Li, David Wenzhong Gao, "Fully Distributed AC Power System State Estimation with Information Propagation Algorithm," *(to be submitted to Journal of Modern Power Systems and Clean Energy)*.

Qiao Li, David Wenzhong Gao, "Distributed AC Optimal Power Flow with Line Flow Limits Based on Consensus Protocol," *(to be submitted to IEEE Transactions on Smart Grid)*.

References

[1] S. Bahrami, and M. H. Amini, "A decentralized trading algorithm for an electricity market with generation uncertainty," *Applied energy,* vol. 218, pp. 520-532, 2018.

[2] F. F. Wu, "Power system state estimation: a survey," *International Journal of Electrical Power & Energy Systems,* vol. 12, no. 2, pp. 80-87, 1990.

[3] F. F. Wu, K. Moslehi, and A. Bose, "Power system control centers: Past, present, and future," *Proceedings of the IEEE,* vol. 93, no. 11, pp. 1890-1908, 2005.

[4] H. Li, A. T. Eseye, J. Zhang, and D. Zheng, "Optimal energy management for industrial microgrids with high-penetration renewables," *Protection and Control of Modern Power Systems,* vol. 2, no. 1, pp. 12, 2017.

[5] M. Mesbahi, and M. Egerstedt, *Graph theoretic methods in multiagent networks*: Princeton University Press, 2010.

[6] A. J. Wood, and B. Wollenberg, "Power generation operation and control—2nd edition," in Fuel and Energy Abstracts, 1996, pp. 195.

[7] A. Bakirtzis, V. Petridis, and S. Kazarlis, "Genetic algorithm solution to the economic dispatch problem," *IEE proceedings-generation, transmission and distribution,* vol. 141, no. 4, pp. 377-382, 1994.

[8] Z.-L. Gaing, "Particle swarm optimization to solving the economic dispatch considering the generator constraints," *IEEE transactions on power systems,* vol. 18, no. 3, pp. 1187-1195, 2003.

[9] A. Mohammadi, M. H. Varahram, and I. Kheirizad, "Online solving of economic dispatch problem using neural network approach and comparing it with classical method," in 2006 International Conference on Emerging Technologies, 2006, pp. 581-586.

[10] K. G. Boroojeni, M. H. Amini, S. Iyengar, M. Rahmani, and P. M. Pardalos, "An economic dispatch algorithm for congestion management of smart power networks," *Energy Systems,* vol. 8, no. 3, pp. 643-667, 2017.

[11] Y. Xu, and Z. Li, "Distributed optimal resource management based on the consensus algorithm in a microgrid," *IEEE Transactions on Industrial Electronics,* vol. 62, no. 4, pp. 2584-2592, 2015.

[12] W. Zhang, W. Liu, X. Wang, L. Liu, and F. Ferrese, "Online optimal generation control based on constrained distributed gradient algorithm," *IEEE Transactions on Power Systems,* vol. 30, no. 1, pp. 35-45, 2015.

[13] G. Binetti, A. Davoudi, F. L. Lewis, D. Naso, and B. Turchiano, "Distributed consensus-based economic dispatch with transmission losses," *IEEE Transactions on Power Systems,* vol. 29, no. 4, pp. 1711-1720, 2014.

[14] R. Mudumbai, S. Dasgupta, and B. B. Cho, "Distributed control for optimal economic dispatch of a network of heterogeneous power generators," *IEEE Transactions on Power Systems,* vol. 27, no. 4, pp. 1750-1760, 2012.

[15] Z. Zhang, and M.-Y. Chow, "Convergence analysis of the incremental cost consensus algorithm under different communication network topologies in a smart grid," *IEEE Transactions on Power Systems,* vol. 27, no. 4, pp. 1761-1768, 2012.

[16] S. Yang, S. Tan, and J.-X. Xu, "Consensus based approach for economic dispatch problem in a smart grid," *IEEE Transactions on Power Systems,* vol. 28, no. 4, pp. 4416-4426, 2013.
[17] S. Kar, and G. Hug, "Distributed robust economic dispatch in power systems: A consensus+ innovations approach," in 2012 IEEE Power and Energy Society General Meeting, 2012, pp. 1-8.
[18] S. Kar, G. Hug, J. Mohammadi, and J. M. Moura, "Distributed State Estimation and Energy Management in Smart Grids: A Consensus ${+} $ Innovations Approach," *IEEE Journal of Selected Topics in Signal Processing,* vol. 8, no. 6, pp. 1022-1038, 2014.
[19] F. Guo, C. Wen, J. Mao, and Y.-D. Song, "Distributed economic dispatch for smart grids with random wind power," *IEEE Transactions on Smart Grid,* vol. 7, no. 3, pp. 1572-1583, 2016.
[20] D. Du, X. Li, W. Li, R. Chen, M. Fei, and L. Wu, "ADMM-based distributed state estimation of smart grid under data deception and denial of service attacks," *IEEE Transactions on Systems, Man, and Cybernetics: Systems,* vol. 49, no. 8, pp. 1698-1711, 2019.
[21] A. Gómez-Expósito, A. de la Villa Jaén, C. Gómez-Quiles, P. Rousseaux, and T. Van Cutsem, "A taxonomy of multi-area state estimation methods," *Electric Power Systems Research,* vol. 81, no. 4, pp. 1060-1069, 2011.
[22] V. Kekatos, and G. B. Giannakis, "Distributed robust power system state estimation," *IEEE Transactions on Power Systems,* vol. 28, no. 2, pp. 1617-1626, 2013.
[23] G. N. Korres, "A distributed multiarea state estimation," *IEEE Transactions on Power Systems,* vol. 26, no. 1, pp. 73-84, 2011.
[24] G. N. Korres, A. Tzavellas, and E. Galinas, "A distributed implementation of multi-area power system state estimation on a cluster of computers," *Electric Power Systems Research,* vol. 102, pp. 20-32, 2013.
[25] X. Li, and A. Scaglione, "Robust decentralized state estimation and tracking for power systems via network gossiping," *IEEE Journal on Selected Areas in Communications,* vol. 31, no. 7, pp. 1184-1194, 2013.
[26] A. K. Singh, and B. C. Pal, "Decentralized dynamic state estimation in power systems using unscented transformation," *IEEE Transactions on Power Systems,* vol. 29, no. 2, pp. 794-804, 2014.
[27] J. Du, S. Ma, Y.-C. Wu, and H. V. Poor, "Distributed hybrid power state estimation under PMU sampling phase errors," *IEEE Transactions on Signal Processing,* vol. 62, no. 16, pp. 4052-4063, 2014.
[28] Q. Li, D. W. Gao, H. Zhang, Z. Wu, and F.-Y. Wang, "Consensus-based distributed economic dispatch control method in power systems," *IEEE Transactions on Smart Grid,* 2017.
[29] T. Erseghe, "Distributed optimal power flow using ADMM," *IEEE transactions on power systems,* vol. 29, no. 5, pp. 2370-2380, 2014.

[30] Y. Wang, L. Wu, and S. Wang, “A fully-decentralized consensus-based ADMM approach for DC-OPF with demand response,” *IEEE Transactions on Smart Grid,* vol. 8, no. 6, pp. 2637-2647, 2017.
[31] W. Lu, M. Liu, S. Lin, and L. Li, “Incremental-oriented ADMM for distributed optimal power flow with discrete variables in distribution networks,” *IEEE Transactions on Smart Grid,* vol. 10, no. 6, pp. 6320-6331, 2019.
[32] N. Meyer-Huebner, M. Suriyah, and T. Leibfried, “Distributed optimal power flow in hybrid AC–DC grids,” *IEEE Transactions on Power Systems,* vol. 34, no. 4, pp. 2937-2946, 2019.
[33] S. Zhang, and K. W. Hedman, "A Distributed Optimal Power Flow Algorithm that Incorporates A Priori Congestion Information through Bounded Dual Updates." pp. 1-6.
[34] G. Hug-Glanzmann, and G. Andersson, “Decentralized optimal power flow control for overlapping areas in power systems,” *IEEE Transactions on Power Systems,* vol. 24, no. 1, pp. 327-336, 2009.
[35] F. J. Nogales, F. J. Prieto, and A. J. Conejo, “A decomposition methodology applied to the multi-area optimal power flow problem,” *Annals of operations research,* vol. 120, no. 1-4, pp. 99-116, 2003.
[36] E. Dall'Anese, H. Zhu, and G. B. Giannakis, “Distributed optimal power flow for smart microgrids,” *IEEE Transactions on Smart Grid,* vol. 4, no. 3, pp. 1464-1475, 2013.
[37] A. G. Bakirtzis, and P. N. Biskas, “A decentralized solution to the DC-OPF of interconnected power systems,” *IEEE Transactions on Power Systems,* vol. 18, no. 3, pp. 1007-1013, 2003.
[38] P. Biskas, and A. Bakirtzis, “A decentralized solution to the security constrained DC-OPF problem of multi-area power systems,” in 2005 IEEE Russia Power Tech, 2005, pp. 1-7.
[39] A. Kargarian, Y. Fu, and Z. Li, “Distributed security-constrained unit commitment for large-scale power systems,” *IEEE Transactions on Power Systems,* vol. 30, no. 4, pp. 1925-1936, 2015.
[40] R. H. Lasseter, and P. Piagi, “Microgrid: A conceptual solution,” in IEEE Power Electronics Specialists Conference, 2004, pp. 4285-4291.
[41] T. Gustavi, D. V. Dimarogonas, M. Egerstedt, and X. Hu, “Sufficient conditions for connectivity maintenance and rendezvous in leader–follower networks,” *Automatica,* vol. 46, no. 1, pp. 133-139, 2010.
[42] M. Ji, A. Muhammad, and M. Egerstedt, “Leader-based multi-agent coordination: Controllability and optimal control,” in 2006 American Control Conference, 2006, pp. 6 pp.
[43] R. A. Horn, R. A. Horn, and C. R. Johnson, *Matrix analysis*: Cambridge university press, 1990.
[44] J. D. Glover, M. S. Sarma, and T. Overbye, *Power System Analysis & Design, SI Version*: Cengage Learning, 2012.

[45] S. Fan, Y.-K. Wu, W.-J. Lee, and C.-Y. Lee, "Comparative study on load forecasting technologies for different geographical distributed loads," in 2011 IEEE Power and Energy Society General Meeting, 2011, pp. 1-8.
[46] A. K. Singh, S. Khatoon, M. Muazzam, and D. Chaturvedi, "Load forecasting techniques and methodologies: A review," in 2012 2nd International Conference on Power, Control and Embedded Systems, 2012, pp. 1-10.
[47] M. Artin, *Algebra*, 2nd Edition ed.: Pearson Prentice Hall, 2011.
[48] S. H. F. A. J. I. L. E. Spence, *Linear Algebra*: Pearson Education, 2003.
[49] K. Levenberg, "A method for the solution of certain non-linear problems in least squares," *Quarterly of applied mathematics*, vol. 2, no. 2, pp. 164-168, 1944.
[50] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431-441, 1963.
[51] R. Balakrishnan, and K. Ranganathan, *A textbook of graph theory*: Springer Science & Business Media, 2012.
[52] P. Anderson, A. Fouad, and H. Happ, "Power system control and stability," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 2, pp. 103-103, 1979.
[53] M. Pai, *Energy function analysis for power system stability*: Springer Science & Business Media, 2012.
[54] "118 bus power flow test case," 10/16, 2021; https://labs.ece.uw.edu/pstca/pf118/pg_tca118bus.htm.
[55] "300 bus power flow test case," 10/16, 2021; http://labs.ece.uw.edu/pstca/pf300/pg_tca300bus.htm.