CENTRALIZED AND DISTRIBUTED OPTIMIZATION OF LOAD DISPATCH PROBLEM INCORPORATING DISTRIBUTED GENERATION

TAN SICONG

NATIONAL UNIVERSITY OF SINGAPORE

2015

CENTRALIZED AND DISTRIBUTED OPTIMIZATION OF LOAD DISPATCH PROBLEM INCORPORATING DISTRIBUTED GENERATION

TAN SICONG

B.Eng.(Hons.), NTU

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

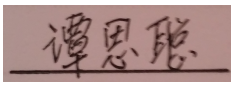
NUS GRADUATE SCHOOL FOR INTEGRATIVE SCIENCES AND ENGINEERING NATIONAL UNIVERSITY OF SINGAPORE

2015

Declaration

I hereby declare that the thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.



19 SEP 2015

Acknowledgements

I would like to express the deepest appreciation to my supervisor Professor Xu Jianxin and co-supervisor Associate Professor Sanjib Kumar Panda, who have shown the attitude and the substance of a genius: they continually and persuasively conveyed a spirit of adventure in regard to research and scholarship, and an excitement in regard to teaching. Without their supervision and constant help this thesis would not have been possible.

I would like to thank my thesis advisory committee members, Professor Wu Yihong and Associate Professor Martin Henz, for their encouragement, insightful comments, and hard questions.

I would like to thank NUS Graduate School of Integrative Sciences and Engineering for providing the scholarship. Moreover, I would like to thank all the staff in NGS office for providing support on all the administrative matters.

My sincere thanks also go to Mr. Yang Shiping for collaborating on these projects.

I thank the lab technicians Mr. Woo Ying Chee, Mr. M Chandra and Mr Seow Hung Cheng for providing IT support.

Last but not the least, I would like to thank my family: my parents Zhou

Guangping, Tan Yuji, and my wife Zhang Dan, for supporting me throughout my life.

Contents

1 Introduction **22**

1.1 Literature Review 22

1.2 Problem and Motivation 40

1.3 Objectives and Challenges 41

1.4 Contributions and Outline 42

2 Multi-objective Optimization of Economic Load Dispatch for Micro-grids Using Evolutionary Computation **45**

2.1 Introduction 45

2.2 Formulation of Load Dispatch 46

2.2.1 Operating cost and emission of the thermal generator . 49

2.2.2 Operating cost of the wind farm 50

2.2.3 Operating cost of PEM fuel cell 52

2.2.4 System structure 54

2.3 Overview of SPEA2 and NSGA-II 54

2.3.1 SPEA2 55

2.3.2 NSGA-II 58

2.4 Application Examples 60

2.4.1 Case study 1: 12-generator test system 61
2.4.2 Case study 2: 33-generator test system 63
2.4.3 Case study 3: 54-generator test system 66
2.4.4 Case study 4: 180-generator test system 68
2.5 Conclusion 70

3 Optimization of Distribution Network Incorporating Micro-grid: An Integrated Approach 72

3.1 Introduction 72
3.2 Problem Formulation 73
3.2.1 System structure 74
3.2.2 Objective and constraints 74
3.2.3 Power flow formulation 77
3.2.4 Forecasting 79
3.2.5 Models 82
3.3 Methodology 87
3.3.1 Acquiring parameters 87
3.3.2 Optimization techniques 89
3.3.3 Parametric setting 95
3.3.4 Encoding strategy 97
3.3.5 Overall workflow 98
3.4 Integrated Approach 99
3.5 Application Example 100
3.5.1 Case study 1: renewable energy at the same place . . . 102
3.5.2 Case study 2: renewable energy at different places . . . 108

3.5.3 Case study 3: fault occurrence 114
3.5.4 Case study 4: fixed structure 117
3.6 Conclusion 119

4 Consensus Based Approach for Economic Dispatch Problem in a Micro-grid 121

4.1 Introduction 121
4.2 Preliminary 123
4.2.1 Graph theory 123
4.2.2 Consensus algorithm 124
4.2.3 Analytic solution to ED problem 126
4.3 Main Results 128
4.3.1 Algorithm design without power generation constraints 128
4.3.2 Generalization to constrained case 134
4.3.3 Fully distributed implementation 136
4.4 Learning gain design 137
4.5 Application Examples 141
4.5.1 Case study 1: without generator constraints 143
4.5.2 Case study 2: with generator constraints 144
4.5.3 Case study 3: robustness of command node connections 145
4.5.4 Case study 4: plug and play test 147
4.5.5 Case study 5: time-varying demand 148
4.5.6 Case study 6: relation between convergence speed and learning gain 150
4.5.7 Case study 7: comparison with Lambda-Iteration method 152

4.5.8 Case study 8: fully Distributed Implementation with IEEE 14-Bus Test Systems 155
4.5.9 Case study 9: comparison with Kar's Work 158
4.5.10 Case study 10: Application in large micro-grid 160
4.6 Conclusion 161

5 Hierarchical Consensus Based Approach with Loss Consideration for Economic Dispatch Problem under Micro-grid Context 163

5.1 Introduction 163
5.2 Preliminary 165
5.2.1 Analytic solution to ED problem with loss calculation . 166
5.3 Main Results 168
5.3.1 Upper level: estimating the power loss 168
5.3.2 Lower level: solving economic dispatch distributively . 170
5.4 Application Examples 176
5.4.1 Case study 1: convergence test 178
5.4.2 Case study 2: robustness of command node connections 179
5.4.3 Case study 3: plug and play test 180
5.4.4 Case study 4: time-varying demand 183
5.4.5 Case study 5: relation between convergence speed and learning gain 184
5.4.6 Case study 6: IEEE 14-bus system test 186
5.4.7 Case study 7: IEEE 57-bus system test 188
5.4.8 Case study 8: Application in large micro-grid 190

5.5 Conclusion 191

6 Conclusion and Future Work **193**

A Publication **208**

A.1 Journal Publication 208

A.2 Conference Publication 208

A.3 Under Revision 209

Summary

Economic dispatch (ED) is an important problem in power system operations. The objective of this problem is to minimize the operating cost of all the generation in the power network with consideration of various system constraints. With the increasing environmental concerns, the traditional power system is facing a problem of increasing penetration of distributed energy resources (DERs). These DERs such as wind turbines and Photovoltaic panels produce electricity from renewable energy sources which can reduce the emission and help protect the environment. Nevertheless, these DERs are heavily relied on stochastic weather conditions, which results in the uncertainty of the power outputs. Furthermore, the accurate models of these generators' cost functions are highly non-linear and non-convex, which significantly increases the computational complexity. Moreover, unlike traditional grid, DERs are distributively implemented, and the users have flexibility of installing and uninstalling the generators, which requires high computational scalability. These problems have raised the difficulty in integrating DERs into the traditional grid. Thus, the integration of these DERs is a challenging and important task in modern power system development. In this thesis, four different approaches are proposed to tackle the ED problem.

First two approaches present centralized optimization schemes for solving the ED problem with the consideration of distributed generation. The last two approaches present distributed optimization schemes for solving the ED problem with distributed generation.

The thesis starts investigating the ED problem from the micro-grid islanding operations by two multi-objective optimization algorithms. This approach formulates the ED problem as a bi-objective optimization problem. This approach provides the user with various choices to support the load demand in a micro-grid islanding operation depending on the user's minimization objectives preferences.

Furthermore, the thesis investigates the ED problem of the distribution network with distributed energy resources. This approach minimizes the operating cost of the whole network with distributed generation and utilize more renewable energy into the traditional grid.

Additionally, the thesis proposed a novel consensus based algorithm to solve ED problem in a distributed manner. The distributed algorithm enables generators to collaboratively learn the mismatch between demand and total power generation in a distributed environment, and all generators collectively minimize the total cost while satisfying power balance constraint.

On top of that, the thesis improves the existing distributed ED algorithm in literature by bringing loss information into consideration. This is a hierarchical consensus based algorithm. This approach minimizes total operating cost with loss consideration while maintaining the power balance constraint under sparse communication network.

List of Tables

2.1 Mean and standard deviation for different number of generations by 12-generator system 62
2.2 Mean and standard deviation for different number of generations by 33-generator system 65
2.3 Mean and standard deviation for different number of generations by 54-generator system 67
2.4 Mean and standard deviation for different number of generations by 180-generator system 69

3.1 Case study 1: results obtained by AV-AIS providing different renewable power. 103
3.2 Case study 1: operating cost obtained by different algorithms. 106
3.3 Case study 2: results obtained by adaptive AV-AIS providing different renewable power. 110
3.4 Operating cost obtained by different algorithms. 112
3.5 Case study 3: results obtained by AV-AIS providing different renewable power. 115

3.6 Case study 4: results obtained by AV-AIS providing different renewable power. 118

4.1 Generator parameters 143

4.2 Initializations 144

4.3 IEEE 14-bus test systems generator parameters 156

5.1 Generator parameters 177

5.2 Initializations 178

5.3 IEEE 14-bus test system generator parameters 187

List of Figures

2.1 The micro-grid system structure 55

2.2 Workflow of SPEA2 56

2.3 Workflow of NSGA-II 59

2.4 Crowding distance sorting 60

2.5 Optimal fronts obtained for various number of generations by 12-generator system 64

2.6 Optimal fronts obtained for various number of generations by 33-generator system 66

2.7 Optimal fronts obtained for various number of generations by 54-generator system 68

2.8 Optimal fronts obtained for various number of generations by 180-generator system 70

3.1 System structure of distributed network with network reconfiguration and micro-grid 75

3.2 An example of power flow in a radial network 78

3.3 ϵ-insensitive band for SVR 81

3.4 An example of 2-day load forecasting with error rate 0.67% . . 83

3.5 An example of wind speed distribution information acquisition 88
3.6 An example of solar irradiance distribution information acquisition 89
3.7 Vaccine extraction of a 2 dimensional search space 93
3.8 Adaptive vaccine-AIS flowchart 96
3.9 An example of 3 feeder network 97
3.10 An example of nodes supported by DG and ESS 98
3.11 Overall system workflow 99
3.12 IEEE 33-node test system 101
3.13 Structure inside the dashed ellipse: a flexible micro-grid 101
3.14 Case study 1: dynamic experiment results 107
3.15 IEEE 33-node test system with renewable energy at different places 108
3.16 Case study 2: dynamic experiment results 113
3.17 IEEE 33-node test system with fault on 9-10 114
3.18 IEEE 33-node test system with fixed structure 117

4.1 Communication topology among generators and command vertex in the network. 143
4.2 Results obtained without generator constraints. 145
4.3 Results obtained with generator constraints. 146
4.4 Robustness test when the command vertex is connected to generators 2 and 4. 147
4.5 Communication topology with the fifth generator. 149
4.6 Results obtained with the fifth generator. 150

4.7 Results obtained with time-varying demand. 151
4.8 Relation between convergence speed and learning gain. 152
4.9 Results obtained by Lambda-Iteration method. 155
4.10 Results obtained with IEEE 14-bus test systems. 157
4.11 (a) Power output from generators and (b) λ over iterations with no generator reaching its limit and strong network connections 160
4.12 Results obtained with 100 generators. 161

5.1 Communication topology among generators and command vertex in the network. 177
5.2 Results obtained with generator constraints. 179
5.3 Robustness test when the command vertex is connected to generators 2 and 4. 181
5.4 Communication topology with the fifth generator. 182
5.5 Results obtained with the fifth generator. 183
5.6 Results obtained with time-varying demand. 185
5.7 Relation between convergence speed and learning gain. 186
5.8 Results obtained with IEEE 14-bus test system. 188
5.9 Results obtained with IEEE 57-bus test system. 189
5.10 Results obtained with 100 generators. 191

Nomenclature

$[B], B_0^T, B_{00}$ power line loss B matrix coefficients

$\alpha/c1/c2$ particle moving speed parameters

α_i/α_i^* Lagrange multipliers

ϵ width of the tube

η_i fuel cell electrical efficiency of ith generator

η_{st} hydrogen storage efficiency

$\phi()$ kernel function

ρ/β decay constant/user specified parameter

ρ_c crossover rate of GA

ρ_m mutation rate of GA

$\sigma(t)/\mu(t)$ standard deviation/mean of the distribution

σ_i^k distance to the kth nearest neighbor

$\theta_i, \beta_i, \gamma_i, \zeta_i, \kappa_i$ emission coefficients of ith generator

ξ_i/ξ_i^* lower/upper training error

a_i, b_i, c_i, e_i, f_i cost coefficients of ith generator

A_t archive at generation t

ab_r antibodies in the population

C cost of error

$c(t), k(t)$ scale factor, shape factor at a given location

$C_{F,i,H}(t)$ hydrogen storing cost

$C_{F,i,Th}(t)$ thermal load cost

$C_{F,i}(t)$ operating cost of ith fuel cell generator

C_{factor} environmental cost factor

$C_{loss}(t)$ cost of power loss on all transmission lines

$C_{mg}(t)$ cost of power generated by micro-grid

C_n price of natural gas

C_{pump} pumping cost

$C_{PV,i}(t)$ cost of ith PV generator

$C_{S,i}(t)$ fuel cost of synchronous ith generator

$C_{T,i}(t)$ operating cost of ith thermal generator

$C_{utility}(t)$ cost of power purchased from utility grid

$C_{W,i,p}(t)$ penalty cost of ith wind turbine for not using all the available wind power

$C_{W,i,r}(t)$ reserve cost due to the available wind power being less than the scheduled wind power

$C_{W,i}(t)$ operating cost of ith wind turbine

$D(i)$ distance density of individual i

D_k number of grids in kth dimension

$d_m r$ vaccine affinity

d_i wind power cost

d_{ri} antibody to antibody affinity

$Diff()$ difference between actual and forecasted peak load

$Dis(i)$ crowding distance of solution i

$E_{S,i}(t)$ emission cost of ith generator at time t

$erf(x)$ Gauss error function $erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$

$f(ab_r)$ affinity (fitness) of the antibody

$f(i)$ fitness value of individual i

$f_r(ab_r)$ relative affinity (fitness) of the antibody

F_i solutions with ith rank

f_m^{max} maximum fitness value in objective m

f_m^{min} minimum fitness value in objective m

$f_w()$ wind power probability density function

$G_{ING}(t)$ incident irradiance

G_{STC} irradiance at STC 1000 (W/m^2)

I sorted solutions using each objective value

$I(i)_m$ ith solution sorted in objective m sequence

$IPDF()$ irradiance probability density function

k, c weibull pdf parameters

k_{pi} penalty cost coefficient

k_{PV} temperature coefficient of power

K_{ri} reserve cost coefficient

LP_j power loss on jth branch

$MVAf_j/MVAf_j^{max}$ current/maximum power on jth branch

N population size

o, l intermediate variables

$OC(t)$ operating cost of the entire grid

P vector of all generators' outputs

$P_a(t)$ power for auxiliary devices

$P_{Dem}(t)$ power demand at time t

$P_{F,i,Th}(t)$ thermal power of ith fuel cell

$P_{F,i}(t)$ output power of ith fuel cell at time t

$P_{F,i}^{max}(t)$ maximum power output of ith fuel cell generator

$P_{F,i}^{min}(t)$ minimum power output of ith fuel cell generator

$P_H(t)$ equivalent power for hydrogen production

$P_i(t)$ ith generator's output at time t

$P_i^{max}(t)$ maximum power output of ith generator at time t

$P_i^{min}(t)$ minimum power output of ith generator at time t

$P_j(t)/Q_j(t)$ real/reactive power on jth branch

$P_{Lj}(t)$ load on jth branch

$P_{Loss}(t)$ power loss on the transmission line at time t

$P_{PV,i,ex}(t)$ expected PV output power value

$P_{PV,i}(t)$ output power of ith PV generator

$P_R(t)$ spinning reserve power at time t

$P_{S,i}(t)$ output power of ith synchronous generator

$P_{T,i}(t)$ output of ith thermal generator

$P_{T,i}^{max}(t)$ maximum power output of ith thermal generator

$P_{T,i}^{min}(t)$ minimum power output of ith thermal generator

P_t population at generation t

$P_{utility}(t)$ power purchased from utility grid

$P_{W,i,av}$ available wind power for ith wind farm

$P_{W,i,ex}(t)$ expected output of ith wind turbine

$P_{W,i,r}$ rated power of wind turbine

$P_{W,i}(t)$ power output of ith wind turbine

PLR_i part load ratio of ith generator

Q_t offspring at generation t

$R(i)$ raw fitness of individual i

r_j/x_j resistance/reactance of jth branch

$S(i)$ strength value of individual i

$S_i(t)$ status of ith generator at time t

$SOC_i(t)$ state of charge of ith battery

$SOC_{max,i}$ maximum state of charge of ith battery

$SOC_{min,i}$ minimum state of charge of ith battery

$T_c(t)$ cell temperature

T_r reference temperature

v_k^i random point in kth dimension and ith grid

$v_i(t), v_o(t), v_r(t)$ cut-in, cut-out and rated wind speed respectively

$V_j(t)$ node voltage on jth branch

$V_j^{min}(t)/V_j^{max}(t)$ minimum/maximum node voltage on jth branch

$vaccine$ generated vaccine set

w normal vector to the hyperplane

W_k grid width in kth dimension

x_k^{max} upper boundary in kth dimension

x_k^{min} lower boundary in kth dimension

Chapter 1

Introduction

1.1 Literature Review

Economic dispatch (ED) problem started from the time when there were two or more generators committed to satisfy the load demand in a power system where the capacities exceeded the load demand required in early 1920s. The problem arisen to the operator was how to divide the load demand between the two units [1, 2]. ED problem was then studied by engineers and scientists. Before 1930, various solutions such as "the base load method" and "best point loading" were developed and used in real time control to allocate the total generation among the units committed to satisfy the load. "The base load method" utilizes the most efficient generator unit to its maximum capacity first. After that, the second most efficient generator unit is loaded. The process continues until the demand is satisfied [3]. However, the fuel cost and the power output have a quadratic relationship. By maximizing the power output of one generator, the operating cost increases rapidly. Thus this

approach is not an economic solution. "Best point loading" successively loads generator units to their lowest heat rate point. The operation begins with the most efficient unit, followed by less efficient units [4]. This approach is more accurate than "The base load method". Nevertheless, with the quadratic cost-power relationship, the heat rate and efficiency vary with different power outputs. Therefore this solution is not optimal.

In early 1930, the "equal incremental method" was recognized to produce the most economic results. This method was established in [5] and [6]. In these two papers, the principle of "equal incremental method" is stated that the next increment in load demand should be balanced by the generator unit with lowest incremental cost. This is a important and fundamental principle which still applies today. It was also recognized by Steinberg [7] that this principle would result in equal incremental cost among all generator units. After that, a formal proof of "equal incremental method" was given by Steinberg and Smith in 1934 [8]. In this paper, they proved that for any required load demand ($P_{Dem}(t) = P_{S,1}(t) + P_{S,2}(t)$), the most economic solution was given by

$$\frac{d\,C_{S,1}(t)}{d\,P_1(t)} = \frac{d\,C_{S,2}(t)}{d\,P_2(t)}. \tag{1.1}$$

With the publication of Steinberg and Smith's book "Economy Loading of Power Plants and Electric Systems" [9] in 1943, the equal incremental cost criteria (EICC) was then widely accepted and applied by operators and researchers. However, it is also noted that Steinberg only considered power output variables. The transmission network and the corresponding power loss issues were ignored.

A lot of research were done under the EICC branch. During early work of the EICC branch, fuel cost curves were accurately represented with valve-point effect. In [10], the network was assumed to be lossless. The implementation was carried out by graph or by slide-rule. After that, George [11, 12], added a model for real power transmission losses to the incremental problem which led to the classical economic dispatch. The real power losses were expressed as a quadratic functions and coordination equations were developed to deal with these losses. This was a breakthrough in computation of transmission losses. Furthermore, Kron published four papers [13–16] in a series to present concise power network and loss modelling. The first two papers considered single area loss modelling. The latter two papers considered interconnected areas loss modelling. The loss models were well structured and clearly presented in these papers. Kron's major assumptions were

1. each load current remains constant ratio irrespective of load demand
2. all generators have constant VAR and WATT ratio
3. deviation of generator voltages and angles are small.

Kron proposed the well-known Kron's loss formula, which is also known as B-matrix loss formula

$$\overline{LP(t)} = P^T[B]P + B_0^T P + B_{00}, \tag{1.2}$$

where $P = [P_1(t), \ldots, P_n(t)]^T$ is the vector of all generators' outputs, $[B]$ is the square matrix, B_0^T is the vector of the same length as P and B_{00} is a constant. After that, Kirchmayer applied Kron's work in [17, 18]. With

the help of computer programs, Kirchmayer improved the loss calculation procedures. Kirchmayer also derived the classic coordination equations

$$\frac{d\,C_{S,i}(t)}{d\,P_i(t)} + \lambda \frac{\partial, P_{Loss}(t)}{\partial, P_i(t)} = \lambda; i = 1, 2.... \tag{1.3}$$

The classic coordination equations are the necessary optimization conditions for the economic dispatch Lagrangian function. From 1.3, Ward [19] further derived the transmission loss penalty factor to

$$PF_i = \frac{1}{1 - \frac{\partial P_{Loss}(t)}{\partial P_i(t)}}. \tag{1.4}$$

From 1.4, Tudor [20] found that the system constants could not be easily changed to accommodate other changes in the transmission system. Thus, he proposed easily modified approximation for the penalty factor. The errors incurred by this approximation are relatively small. Van [21] also proposed improved linear loss models based on differential information. The numerical example show that Van's model can allow more variations in the loads and reactive power at the generators. These linear models are popular as they can be easily updated in iterative algorithms. Happ [22] used a Jacobian matrix to calculate the incremental losses. His simulations on IEEE 118-bus test system show that this approach converge rapidly compared to classic economic dispatch. Shoults [23] made use of the linear relationship between real and reactive power outputs to compute the loss coefficient and demonstrated the computational advantages over classical dispatch techniques. With the increasing power system size and complexity, Wollengberg [24] realized that

more comprehensive real time control and dispatch techniques were needed to be studied. He then reviewed the criteria for the constrained optimization practical system and presented security dispatch with reduced set of variables. Therefore large system problems could be solved in real time. E-ICC was also used as a subproblem to update nonlinear information at each iteration. Deo [25] followed Steinberg's development in EICC. From EIC-C, Deo noticed that incremental cost was piece-wise linear from piece-wise linear incremental heat rate system. Then he applied linear programming to solve the economic dispatch problem with simplest expressions to avoid valve point nonlinearities. The solution is very fast because of the simplicity of linear programming technique. However, the optimal dispatch using linear programming tends to provide outputs which were near operating limits. This leads to increased transmission losses and wasted capacity. Ringlee [26] proposed dynamic programming to solve the EICC and achieved the minimum fuel cost dispatch. However, dynamic programming is computationally intensive. The results are provided by close approximation obtained from stepped incremental method.

After using EICC to solve the problem, researchers started to use linear approximations to form the search directions to tackle the problem. Those techniques are grouped as linear sub-problem branch [10]. Lots of work at that time were based on Carpentier's formulation. Carpentier's major contribution was in providing solid mathematical foundations for economic dispatch. Carpentier's work consisted of four parts [27–30]. His work led to a generalized formulation of the economic dispatch problem based on the Kuhn-Tucker conditions. The Kuhn-Tucker conditions are the first order nec-

essary conditions for a solution in nonlinear programming to be optimal if regularity conditions are satisfied [31]. Carpentier used Gauss-Seidel method to solve the problem. However, the convergence was proved to be difficult. Inspired by Carpentier's generalized reduced gradient method, Peschon [32] firstly modified Newton power-flow to provide the specified area interchange flows. After the modification, the problem was in a standard non-linear programming format. He then used Generalized Reduced Gradient and Penalty Function Method to achieve optimal power flow. Comparisons of the two methods show similar optimal numerical results. After Peschon's work, Dommel and Tinney [33] extended their load flow work on Newton's method to an optimal power flow. They solved this problem by a combination of the gradient methods. Their solution is straightforward except for the gradient updating gain. Small gain causes slow convergence, whereas too high a value causes instability. They handled the constraints by penalty functions. Cost function was penalized if there was any violation of constraints. Two other gradient based applications were also popular during that time. Wu [34] used two-stage gradient method to solve the dispatch problem. The first stage does the calculation without considering the dependent constraints. The second stage adds the violated voltage into the objective by penalty functions. This application is able to handle large system, but it often obtains infeasible values after calculation. Another application was done by Burchett [35]. He applied general purpose nonlinear programming to solve the problem. His algorithm periodically switches between conjugate directions technique and Dommel and Tinney's steepest descent technique. This algorithm is able to handle large system calculation and inequality constraints.

From 1970s, concerns about generators' emission on the environment started to attract attention from researchers. Gent [36] modeled the emission of nitrogen oxide as a polynomial function. Experiments were carried out to find the parameters of the polynomial. After that, Gent compared the results with standard economic dispatch and found that reduction of emission could be achieved by systematic scheduling. However, minimum emission dispatch resulted in higher operating cost. Delson [37] added a equality constraint to the problem. The emission was converted to price by a cost factor. The paper provides four different problem formulations according to minimum cost objective, minimum emission objective, minimum cost and emission objective, and minimum cost with constraints on two types of emissions objective. These different problem formulations can be readily adopted by system operators.

After the advancement of linear sub-problem branch, researchers moved their focus to Newton strategy methods [38]. El-Abiad [39] presented a formulation of the dispatch problem based on the Lagrangian multiplier approach. He used Newton's method to calculate the necessary conditions for load flow. All equality constraints were calculated using iterative procedure. However, the results are oscillatory. Shen [38] used a iterative indirect search method to tackle the Lagrange-Kuhn-Tucker conditions of optimality. Unlike direct search algorithm, his method starts with a initial point which may not satisfy all the constraints. After each iteration, the solution is updated by variation method. The results converge within tolerance limits. In 1973, two researchers proposed quadratic subproblems to drive the non-linear optimization. Nabon [40] firstly used the second order approximation to rep-

resent the power generation cost function with linearized sensitivity function. Then he employed quadratic programming to solve the second order problem. This technique requires less computing time compared to non-linear optimization techniques due to its simplicity and this technique avoids using penalty function to maintain the constraints. Similarly, Nicholson [41] used linear programming formulation for estimation the system constraints. He then used quadratic programming technique to solve the quadratic cost problem. Reactive power optimal allocation was done by a gradient method which minimizes the transmission loss. He also applied the optimization unit for online scheduling with the help of large scientific computer. In mid 1970s, quadratic programming based on Kuhn-Tucker conditions became popular. Dillon [42] proposed a method for calculating the sensitivity of the economic dispatch problem. The method can determine which parameters have great influence on the solutions. Thus more accurate measurement techniques or instruments may be required. The expected system performance deterioration is calculated to first order degree of accuracy. Dillon [43] then applied his sensitivity analysis in optimal power flow problem. This method allows fast re-computation of small variation conditions such as loads or constraints values. Thus less frequent re-solution of the non-linear programming technique is needed and schedules can be determined quickly.

Many researchers also applied Newton search method to solve the problem. These methods are also referred as Lagrangian-Newton methods. The projected Lagrangian programs for ED problem were proposed from late 1970s to 1980s. Biggs [44] presented constrained minimization technique based on recursive quadratic programming. Although the algorithm encoun-

ters difficulties in robustness and rapid convergence, the algorithm obtains accurate results. He also proposed to reduce the frequency of calculating Lagrangian multipliers and use Lagrangian penalty function to increase the calculation speed. Lipowski [45] improved Biggs' algorithm and proposed a modified recursive quadratic programming technique. This technique is faster and it converges in less number of iterations than Biggs' method as it has a more efficient subprogram for solving the quadratic programming. Later, Quasi-Newton methods were also introduced with penalty functions. An approximation of the Hessian is built by iterative updating formulae in these methods. Cova [46] found that the second order methods which generated quadratic programming problems had difficulties when dealing with large systems. This was caused by the non-compact formulation of the problem. He then used decomposed technique to construct and update the Hessian matrices at each iteration. This reduces the dimensions of the quadratic subproblem. The results shows that this approach is suitable for large system economic dispatch. Giras [47] used Quasi-Newton method to tackle the economic dispatch problem. The method is straightforward in incorporating power flow constraints. Furthermore, it is robust and it can converge even with infeasible starting point, and it provides fast convergence. Talukdar [48] reviewed and compared the algorithms of dispatch techniques, namely, Dommel-Tinney's methods, Generalized Reduced Gradient methods, Wu-Gross-Luini-Look methods, and Tarlukda-Giras methods. He concluded that within 1-5 minutes dispatching time frame and with large number of constraints, Quasi-Newton methods with elegant constraints handling were very good choices. Some well-known solvers were also developed. Gradient

solver [35] was developed by General Electric package. It works well when carefully tuned. However, it is slow to detect the infeasible solutions and penalty functions. Quasi-Newton solver [49] was developed in ESCA package. This solver is fast and robust. The constraints are well handled, and it can start with infeasible solution. Lagrangian-Newton solver [50] was developed in PCA package. This method is a direct simultaneous solution for all the unknowns in the Lagrange function. This method has well known quadratic convergence properties as it minimizes a quadratic approximation of the Lagrangian in each iteration.

Contingency-constrained economic dispatch was emerged as a hot topic in 1980s. It was found that after major disturbance (line fault or generator outage), it was difficult to bring the system into normal state. Thus extra constraints are needed to make sure that the system is kept in normal state for a long time or it does not stay away from normal state after disturbance. The integration of contingency constraints is an important step in ED research. Monticelli [51] proposed strategies of rescheduling of control variables within short-term ramping limits. This technique is flexible to incorporate existing dispatch methods and it can include more corrective actions such as line switching, overload rotation, and multi-period rescheduling. Schnyder presented quick networking switching [52, 53] to reach $n-1$ security in the conservative sense. The technique uses corrective switching concept to allow contingencies to be treated as corrective actions. Therefore contingency conditions are transformed into inequality constraints and they can be solved by dispatching methods. These strategies take important considerations into system ED operation and alleviate the tight constraints

imposed by the system.

With the development of stochastic search algorithm, the evolutionary computation methods were used to solve the ED problem from 1990s [54]. In contrast to the conventional ED methods, evolutionary computation methods do not need the gradient information of the cost functions. Thus the researchers do not need to differentiate the cost functions and constraints. These methods use probabilistic selection rules to choose solutions which approach closer to the global optima. These solutions are then used to reproduce better solutions in an iterative manner. The end of the iterative process generated a set of solutions which are at or in the neighborhood of the global optima. Wong [55] used the simulated annealing (SA) to solve the ED problem. Simulated annealing is the technique of heating up the metal and cool down the metal in order to increase the size of the crystals and reduce the number of defects. Researchers develop this technique to a probabilistic method for global optimization. Nevertheless, the SA based algorithm is difficult for tuning the parameters of the annealing process and the computation time required is high. Walter [56] proposed a genetic algorithm (GA) to solve highly non-linear ED problems considering the valve-point loading effect and other non-convex generators' cost functions. Genetic algorithm is inspired by the natural evolution. It simulates the inheritance, mutation, crossover, and selection of the species to find the best solution for a optimization problem. The results show that the algorithm is a powerful tool to solve the economic dispatch problem. It is able to handle non-linearities and provide solution very near optimal point. However, selecting the right coding strategy and variation strategies are based on heuristics. Repetitive evalu-

ation of the fitness function can be computationally expensive. Gaing [57] proposed a particle swarm optimization (PSO) for solving the ED problem. Particle swarm optimization comes from the group behavior of bird flocks and fish schools. Their movement mechanism is studied and simplified to perform optimization tasks. In this paper, Gaing used PSO to perform the economic dispatch with ramp rate limits and valve-point zones. The results are compared to GA. It is shown that PSO is more computationally efficient and it provides higher quality solutions.

At the same time, multi-objective ED problem was introduced in literature. Wong [58] formulated the bi-objective problem by using a weighted sum of the two objectives. He then applied simulated annealing to solve the problem. The results are represented by a trade-off curve. However, the trade-off curve is based on different weighting factor. Thus it is not a complete multi-objective solution. Das [59] used a heuristic method which combined genetic algorithm and simulated annealing to perform a multi-objective optimization of the economic dispatch problem. The formulation of the problem is truly multi-objective. The results are represented by pareto-front. The method is fast in determining near-optimal solution and it was applicable for large systems. The pareto-front better represent the cost-emission relationship than weighted trade-off curve. Chandrasekaran [60] presented an optimal deviation based firefly algorithm tuned fuzzy membership function to deal with economic and reliability problem. This paper is important as it considered reliability as one dispatch objective. However, in his paper, the reliability is transformed into cost by a penalty factor. Two objectives are reduced into one. Thus this is not a truly multi-objective problem. Jubril [61] used a

Semidefinite Programming method for solving ED problem with fuel cost and transmission losses. It is interesting that he incorporate transmission losses as another objective. Nevertheless, he used weighted sum to combine the two objectives. The weighted sum can not provide a uniformly distributed solutions as the pareto-front. Thus this results are not truly multi-objective.

Recently, with the increasing environmental awareness, renewable energy sources such as wind and solar energy played a more and more important role in electricity generation. To make better use of the renewable energy resources, distributed generation devices like wind turbines and photovoltaic (PV) panels are used. These devices are grouped to form micro-grids running in the distribution networks [62]. Thus researchers shifted their focus to the ED problem of the micro-grids. Hetzer [63] developed a model to incorporate wind energy into economic dispatch problem. In addition to the operating cost of wind energy, the overestimation and underestimation of wind energy are also modeled. With the help of Weibull probability density function, the dispatch problem is solved. The results demonstrate the relationship between economic solutions and the penalty/reserve cost factors. Hernandez-Aramburo [64] proposed a cost optimization scheme for a micro-grid consisting of two reciprocating gas engines, a combined heat and power plant, a photovoltaic array and a wind generator. The micro-grid system behaves differently than the traditional grid because a penalty over excess heat generated is imposed and minimum amount of reserve power can be easily achieved. The results show that a communication infrastructure is needed to minimize the fuel cost as the power sharing technique between all generators are set for all units explicitly. Chen [65] developed a smart energy manage-

ment system consisting of forecasting module, energy storage module, and optimization module. A matrix real-coded genetic algorithm is adopted to achieve optimal results under three different operating policies. The results show that the management system can reduce the daily costs significantly.

To further assist the power system economic operation and stability, Network reconfiguration, load forecasting and weather forecasting were also studied by researchers. Network reconfiguration is to minimize the energy loss, maintain power balance and isolate faults by changing the states of sectionalizing and tie switches. Many algorithms have been developed to solve the reconfiguration problem. Civanlar [66] presented a scheme to utilize feeder reconfiguration to reduce losses by re-structuring the primary feeders. This paper simplifies the load flow analysis to avoid repetitive calculation. The computation requirements are significantly reduced. This load flow based technique is also flexible to be implemented into the existing feeder reconfiguration strategy. Rao [67] used harmonic search algorithm to find the best configuration of the radial network. The method improves voltage profile and minimizes the real loss. The results are from simulations on 119-bus test system and comparisons with other genetic algorithms are carried out. It is shown that this method converges faster with higher accuracy. Thus this method can be applied for large systems. Huang [68] presented a enhanced genetic fuzzy multi-objective algorithm for reconfiguring the network. The objectives are maximizing fuzzy satisfaction of minimizing power loss, violation of constraints, and the number of switchings. The results are compared to simulated annealing and simple genetic algorithm. Test results demonstrate that the enhanced genetic algorithm could efficiently search

the optimal and near-optimal space better than the other two algorithms. The performance shows that this approach is suitable for actual distribution system. Wu [69] presented an integer-coded particle swarm optimization to solve the distribution network reconfiguration. The advantage of the integer-coded technique is that it reduces the number of infeasible solutions significantly. Thus the search space is narrowed down to feasible solutions' space. This method is compared with genetic algorithm, discrete particle swarm optimization, and modified binary particle swarm optimization. The results show that this method outperforms other three methods in terms of speed and accuracy.

Moreover, Gupta [70] presented a stochastic procedure to forecast monthly peak load up to three years ahead. He improved the well-known stationary stochastic time series to address those non-stationary stochastic time series that could be reduced to stochastic time series by linear transformation. The results show that the forecasting technique catches almost all turning point except for a few cases. The amplitude swings are close. Nevertheless, spring and fall months are a bit off. Overall, the results are quite good for two years forecasting. Taylor [71] incorporated weather ensemble prediction in artificial neural network for load forecasting of 10 days lead times. He used 51 ensemble members for temperature, wind speed, and cloud cover to create 51 load scenarios in the artificial neural network. The 51 scenarios are important as they are equivalent of taking the expectation of load probability density function. The results show that this method can obtain better accuracy than those without weather data and those with traditional weather forecasts. Elattar [72] proposed a modified version of support vector regres-

sion for forecasting of electricity demand. The support vector regression's risk function is modified by the locally weighted regression and weighting function's bandwidth is improved by the weighted distance algorithm based on the Mahalanobis distance algorithm. The results are compared with local support regression and locally weighted regression. The proposed method outperforms the other two methods by providing more accurate predictions. Srinivasan [73] combined fuzzy logic and fuzzy set theory with artificial neural network to forecast the load demand of weekdays as well as weekends. The fuzzy logic is used with special emphasis on weekends and public holidays. This approach reduces the complexity of the mathematical calculation and improves the accuracy of prediction. The results are impressive with average error of 0.62% on weekdays, 0.83% on Saturdays and 1.17% on Sundays and public holidays.

These papers which have discussed so far are all centralized techniques. These techniques required centralized computation and communication. Nevertheless, with the development of distributed energy resources (DERs), the power system structures were becoming more distributed [62]. The users might plug-in or disconnect their DERs such as PV panels and wind turbines without prior notice to the system operators. Additionally, it might be infeasible for the system operators to have all the generators' committed information as well as communication channels. Therefore, more and more researchers have look into the distributed optimization of the ED problem.

The initial distributed optimization of the ED problem was inspired by the consensus problem. Consensus problem has been studied for more than two decades. A huge volume of results were available in the literature [74].

The main issue in a consensus problem was to achieve agreement regarding certain quantities of interest associated with agents in the multi-agent systems (MAS) by utilizing the local information exchange. The traditional consensus algorithm was a very simple local coordination rule, which resulted in agreement at the group level, and no centralized task planner or global information were required by the algorithm. Due to its distributed implementation, robustness, and scalability, consensus algorithms have been widely applied in many coordination problems, such as formation, flocking, rendezvous, and synchronization.

Researchers were interested in applying consensus algorithm in micro-grid related problems. A few of noticeable works appear in the literature. In [75], quadratic convex cost function is assumed. To meet the equal incremental cost optimization criterion, the incremental cost of each generator was chosen as the consensus variable. The consensus algorithm was applied to drive all incremental cost to a common value. To satisfy the demand constraint, the mismatch between demand and total power generated is fed back to the consensus algorithm such that the incremental cost converged to the optimal value. The communication among generators were undirected, which meant the information exchange was bidirectional. This was a restrictive assumption since the communication might not be symmetric in practical situations. In addition, the algorithm was not completely distributed because a leader agent had to be deployed to collect current power generated by each generator in order to calculate the total mismatch. [76] took a different approach. The authors first noticed the results in [77], which showed that the total power generated by all generators was a linear piecewise continuous function of in-

cremental cost. If the ED problem was solvable, the demand must lie in one of the linear segments. Once the linear segment was identified, the optimal incremental cost could be easily obtained by solving an algebraic equation. The generator constraints are considered in the algorithm. To make the algorithm distributed, a *ratio consensus* algorithm is proposed in [76]. The ratio consensus algorithm was applied to learn all the generators constraints and cost functions parameters, specifically, each generator instantiated local estimations of other generators' parameters, and when consensus was reached, the local estimations exactly equaled to the true values of other generators' parameters. Therefore, each generator could obtain all the information needed by the algorithm in [77] in a distributed fashion. Furthermore, a strongly connected communication graph was adequate for this algorithm to work. In [78], motivated by the distributed stochastic approximation theory, a consensus + innovation approach was proposed to solve the ED problem. The innovation term was synthesized by the difference of local power generation and a fixed local reference, where the local reference was either the total demand (estimated or actual), or the total load at the current bus. With the help of vanishing and persistent excitation gains, the algorithm returned the optimal dispatch asymptotically. The latest advancement in distributed ED problem was done by Du [79] in 2014. He proposed two learning algorithms with guaranteed convergence to Nash equilibria and/or optima to solve the ED problem distributively. This method converted inequality constraints into feasible action sets. After that, it incorporated equality constraints by penalty functions, which simplified the ED formulation. It also provided guaranteed convergence.

1.2 Problem and Motivation

With the increasing environmental concerns, more and more distributed generators are connected to the traditional grid. These DERs such as wind turbines and PV panels are affected by stochastic weather conditions. Thus the outputs of these DERs are fluctuating. Furthermore, the electricity cost of distributed generators is higher than the electricity generated by the large scale generators from traditional grid. This also raises the difficulty in integrating distributed generators into the grid. Moreover, distributed generators are supposed to be implemented as "plug-and-play" resources, and they are also supposed to be implemented in a large scale. This significantly increases the computation and communication burden of the centralized power systems. Therefore, the integration of distributed generators is a difficult and challenging task. It must be carefully studied and investigated.

This thesis is motivated by the increasing economical and environmental concerns of power system operation. The distributed generators connected to the grid are very different from the traditional large scale generators. These DERs are small in scale and fluctuating due to stochastic weather conditions. Thus the electricity cost of DERs is higher than the electricity generated by large scale generators. However, these DERs can be installed by normal households, which means the number of the DERs in a network can be large and the capacities of the DERs can take up a substantial portion of the network. These DERs such wind turbines and PV panels make use of renewable energy which can help protect the environment. Therefore it is very important to study the environmental and economical impacts of these

DERs. Especially, we need to investigate in-depth about how to integrate these DERs into the traditional grid and let these DERs work efficiently with traditional generators. Moreover, due to the flexibility of installing DERs and the number of DER units, traditional centralized management and control scheme needs to be revised. A more suitable management and control paradigm must be developed to cater the more distributed network structure.

1.3 Objectives and Challenges

The objective of this thesis is to investigate the economic impacts as well as the environmental impacts of incorporating distributed energy resources into traditional power grid. Economic impacts consist of cost of distributed power generation cost, penalty and reserve cost, cost of purchasing from utility grid, and cost of power loss on transmission lines. Environmental impacts are mainly caused by the emission of thermal generators.

Many challenges have been encountered. DERs are heavily affected by fluctuating weather conditions, which results in uncertainties in power outputs. The generators' cost models are highly non-linear and non-convex, which raises the difficulty in calculation. Moreover, all network users must be considered instead of micro-grid users. This increases the scope of the dispatch problem. Furthermore, unfixed network structure and large scale implementation increase the computation burden and complexity of centralized dispatch approaches. Thus, this problem is a challenging and important task for future power system development.

1.4 Contributions and Outline

This thesis focuses on centralized and distributed optimization of economic dispatch problem incorporating distributed generation. The proposed formulations and optimization techniques are to facilitate the economic operations of the power system. Application examples are conducted on various IEEE test systems. In this section, the contributions of this thesis are briefly summarized as below:

- In Chapter 1, detailed literature review, problem statement, and contribution of this thesis are introduced.

- In Chapter 2, investigation of both economic and environment impacts of thermal generators, wind turbines and PEM fuel cells are conducted. Penalty and reserve functions are introduced to encourage the usage of renewable energy. The goal of minimizing operating cost as well as emission is achieved. Two state-of-the-art multi-objective algorithm are also investigated in the study. The pareto fronts are obtained from 3 different systems to provide operators a variety of generation choices. The results show that SPEA2 has a faster convergence when generation number is small and NSGA-II can perform slightly better for large number of generations. NSGA-II provides more diverse solutions than SPEA2.

- In Chapter 3, an integrated technique of network reconfiguration and economic dispatch of power system with distributed generation is proposed. The stochastic nature of wind, PV and load demand is taken

into consideration by stochastic forecasting modules. Four bio-inspired optimization techniques are adopted to investigate the problem. The economic and environment benefit of the whole network is maximized by the problem formulation and the techniques. Despite higher price of renewable energy, the integrated approach can incorporate more renewable energy as well as minimizing the operating cost. Thus, distributed generation resources are more effectively connected to the grid.

- In Chapter 4, a novel consensus based algorithm is proposed to solve ED problem in a distributed manner. Quadratic cost functions are used to model generators' cost functions. The convergence of our algorithm is proved by eigenvalue perturbation. Sparse communication is used in the application examples to demonstrate the effectiveness of the algorithm. All the generators can collaboratively minimize the operating cost as well as keeping power balance constraint. This algorithm can be extended to large scale power networks which allow "plug-and-play" of distributed generators. Furthermore, comparisons with centralized Lambda-iteration method and distributed "consensus + innovation" method are carried out to demonstrate the advantages of our algorithm.

- In Chapter 5, a hierarchical consensus algorithm is developed to solve ED problem with loss consideration. This is the first work in distributed ED with loss consideration. Quadratic cost functions are adopted. The loss information is represented by B matrix and is handled by the upper layer of the algorithm. The lower layer of the algorithm makes sure that

the dispatch is divided economically while satisfying power balance constraint. This chapter improves on previous studies in literature by first bringing loss information into distributed ED problem.

- Finally, conclusion and future work are stated in Chapter 6.

Chapter 2

Multi-objective Optimization of Economic Load Dispatch for Micro-grids Using Evolutionary Computation

2.1 Introduction

With the increase of electricity demand and the environmental awareness, renewable energy sources such as wind and photovoltaic (PV) play a more and more important role in electricity generation. The wind and PV power generation are highly affected by fluctuating weather conditions [63].

Such fluctuations result in fluctuating generation outputs. These outputs have brought difficulties in keeping the power generations and load demand balance. Furthermore, the electricity demand varies from time to time. Be-

sides meeting the demand and the operating cost, environmental impacts like emission and storage of excess power must be taken into consideration [80–85].

In this chapter, the micro-grids are modeled using thermal generators, wind turbines, and PEM fuel cells. The penalty and reserve functions of wind turbines are used for encouraging the use of renewable energy and reducing emission. The fluctuations of wind turbines' power outputs are balanced by other controllable generators such as thermal generators and fuel cells. A set of feasible non-dominated solutions are generated by the multi-objective algorithms. The user can adjust the generators' outputs flexibly in response to the fluctuating weather conditions. In other words, more thermal and fuel energy are used when the system is short of wind energy so that the power balance is maintained in the system.

This chapter is organized into five sections. In Section 2.1, the background knowledge of economic load dispatch problem is introduced. In Section 2.2, the emission and operating cost of three types of generators are formulated. In Section 2.3, the algorithms of SPEA2 and NSGA-II are introduced. The simulation and results are presented in Section 2.4. Finally, the conclusions are made in Section 2.5.

2.2 Formulation of Load Dispatch

The objective is to minimize the operating cost of all the generators in the micro-grid $OC(t)$ $(P_{T,i}(t), P_{W,i}(t), P_{F,i}(t))$ as well as the emission from thermal generators $E_{S,i}(t)$ subject to system constraints.

As shown in [80, 81], the total operating cost is the summation of all the thermal generation costs, wind generation with penalty and reserve costs, and PEM fuel cell generation with thermal load and hydrogen storage costs

$$\begin{aligned} OC(t) &= \sum C_{T,i}(t) + \sum C_{W,i}(t) + \sum C_{W,i,p}(t) + \sum C_{W,i,r}(t) \\ &\quad + \sum C_{F,i}(t) + \sum C_{F,i,Th}(t) + \sum C_{F,i,H}(t). \end{aligned} \tag{2.1}$$

The system constraints are expressed in (2.2-2.4). (2.2-2.4) are the physical constraints of this optimization task. All generators' outputs are bounded by their manufactured operation ranges to prevent damaging of the equipments. (2.2) states that the thermal generator must operate within its minimum and maximum value. (2.3) ensures that the wind turbine must operate within 0 and its rated value. (2.4) states that the fuel cell unit must operate within its minimum and maximum value. (2.5) is the power balance equation, which ensures that the total power generated is equal to the power demand and the power losses [81–83].

$$P_{T,i}^{min}(t) \leq P_{T,i}(t) \leq P_{T,i}^{max}(t), \forall i, \tag{2.2}$$

$$0 \leq P_{W,i}(t) \leq P_{W,i,r}, \forall i, \tag{2.3}$$

$$P_{F,i}^{min}(t) \leq P_{F,i}(t) \leq P_{F,i}^{max}(t), \forall i, \tag{2.4}$$

$$\sum P_{T,i}(t) + \sum P_{W,i}(t) + \sum P_{F,i}(t) = P_{Loss}(t) + P_{Dem}(t), \forall i. \tag{2.5}$$

The total loss on all the transmission lines is computed using B coefficients [86]. B coefficients, also known as Kron's loss formula, was first introduced in the early 1950s as a practical method for loss and incremental loss calculation. B coefficients are one of the most important methods for approximating the total transmission loss as a function of generators' power. It models the total transmission loss as a second order function. At that time, automatic dispatching was performed by analog computers and the loss formula was stored in the analog computers by setting potentiometers. This method provides reasonably accurate calculation of loss coordination in ED problem. The matrix form is expressed as:

$$\overline{LP(t)} = P^T[B]P + B_0^T P + B_{00}. \tag{2.6}$$

where $P = [P_1(t), \ldots, P_n(t)]^T$ is the vector of all generators' outputs, $[B]$ is the square matrix, B_0^T is the vector of the same length as P and B_{00} is a constant.

2.2.1 Operating cost and emission of the thermal generator

2.2.1.1 Fuel cost

Thermal generator has a number of valves that are opened according to the power output. As the output requirement increases, the generator opens one more valve to allow more steam to come out. However, when a valve is opened, the incremental heat rate rises rapidly. This valve-point effect results in non-smooth, non-convex input-output relationship. A recurring rectified sinusoid modeling the valve-point effect is added to the traditional quadratic cost function as shown in (2.7) [85, 87–90]

$$C_{T,i}(t) = \frac{a_i}{2} P_{T,i}^2(t) + b_i P_{T,i}(t) + c_i + |e_i \sin(f_i(P_{T,i}^{min}(t) - P_{T,i}(t)))|. \quad (2.7)$$

2.2.1.2 Environmental cost

Nitrogen-Oxide (NOx) emission represents the environmental impact of thermal generators. The amount of NOx emission is proportional to generator output as shown by (2.8) [84]

$$E_{T,i}(t) = \theta_i + \beta_i P_{T,i}(t) + \gamma_i P_{T,i}^2(t) + \zeta_i exp(\kappa_i P_{T,i}(t)). \quad (2.8)$$

2.2.2 Operating cost of the wind farm

2.2.2.1 Cost function

For wind turbine, when the wind farm is not owned by the system operator, the system operator needs to pay a constant price for using the power from wind farm. The cost is linearly proportional to the power output [81, 85] as shown in (2.9)

$$C_{W,i}(t) = d_i P_{W,i}(t). \tag{2.9}$$

2.2.2.2 Penalty cost

In case the available wind power is more than the amount being generated. The energy is wasted. A penalty cost related to the amount of energy being wasted is added. The penalty cost function is shown by (2.10) [81]

$$\begin{aligned} C_{W,i,p}(t) &= k_{pi}(P_{W,i,av} - P_{W,i}(t)) \\ &= k_{pi} \int_{P_{W,i}(t)}^{P_{W,i,r}(t)} (w - P_{W,i}(t)) f_w(w) dw \\ &= k_{pi} \left[0.5 P_{W,i,av}^2 \sqrt{\pi} erf(u) - (P_{W,i,av}^2 u - P_{W,i}(t) \right. \\ &\quad \left. P_{W,i,av} + 0.5 P_{W,i,av}^2) e^{-u^2} \right]_{u=P_{W,i}(t)}^{u=P_{W,i,av}}. \end{aligned} \tag{2.10}$$

2.2.2.3 Reserve cost

In case the available wind power is less than the amount required, the load must be shed or the energy need to be bought from somewhere else. A re-

serve cost is imposed as shown by (2.11) [81]

$$\begin{aligned} C_{W,i,r}(t) &= k_{ri}(P_{W,i}(t) - P_{W,i,av}) \\ &= k_{ri}\int_0^{w_i}(P_{W,i}(t) - w)f_w(w)dw \\ &= k_{ri}\Big[(P_{W,i,av}^2 u - P_{W,i}(t)P_{W,i,av} + 0.5P_{W,i,av}^2)e^{-u^2} \\ &\quad -0.5P_{W,i,av}^2\sqrt{\pi}erf(u)\Big]_{u=0.5}^{u=0.5+\frac{P_{W,i}(t)}{P_{W,i,av}}} . \end{aligned} \tag{2.11}$$

Wind speed distribution is modeled as Weibull probability density function. The probability density function of wind power output can also be represented by [85]

$$f_w(w) = \frac{klv_i}{c}(\frac{(1+ol)v_i}{c})^{k-1}exp(-(\frac{(1+ol)v_i}{c})^k), \tag{2.12}$$

for $0 < w < P_{W,i,av}$, and

$$f_w(0) = 1 - exp(-(\frac{v_i}{c})^k) + exp(-(\frac{v_0}{c})^k), \tag{2.13}$$

$$f_w(w_r) = exp(-(\frac{v_r}{c})^k) - exp(-(\frac{v_0}{c})^k), \tag{2.14}$$

where $o = \frac{w}{P_{W,i,av}}$ and $l = \frac{v_r - v_i}{v_i}$.

2.2.3 Operating cost of PEM fuel cell

2.2.3.1 Fuel cost

For PEM fuel cell, the fuel cost is linearly proportional to the sum of the power generated, the power consumed by auxiliary devices and the power for hydrogen production as shown by (2.15) [82]

$$C_{F,i}(t) = C_n(\frac{P_{F,i}(t) + P_a(t) + P_H(t)}{\eta_i}). \tag{2.15}$$

2.2.3.2 Recovered Thermal Energy Calculation

Not only PEM fuel cell produces electrical energy, but also it produces thermal energy. The thermal energy is used to satisfy the thermal load. Thermal load is caused due to space heating and hot water, thus is a part of the load along with electric load. The thermal load is satisfied by utilizing the recovered thermal energy. The electricity-fuel conversion efficiency is related to the thermal load ratio as shown in (2.16), (2.17) and [82].

PLR means part load ratio. If the part load ratio is small ($PLR_i < 0.05$), the efficiency and the thermal energy to electrical energy ratio can be expressed in simple terms. If the PLR is not small ($PLR_i \geqslant 0.05$), the efficiency and the thermal energy to electrical energy ratio need to be expressed by polynomial functions, which is more complicated. Non-linear functions increase the difficulty of computation for traditional gradient methods.

For $PLR_i < 0.05$, the efficiency is a simple factor given by [82]

$$\eta_i = 0.2716. \tag{2.16}$$

The thermal energy to electrical energy ratio is given by [82]

$$\gamma_{TE,i} = 0.6801. \tag{2.17}$$

For $PLR_i \geqslant 0.05$, the efficiency is represented by a fifth-order polynomial [82]

$$\begin{aligned} \eta_i &= 0.9033PLR_i^5 - 2.9996PLR_i^4 + 3.6503PLR_i^3 \\ &\quad -2.0704PLR_i^2 + 0.4623PLR_i + 0.3747. \end{aligned} \tag{2.18}$$

The thermal energy to electrical energy ratio is given by a fourth order polynomial [82]

$$\begin{aligned} \gamma_{TE,i} &= 1.0785PLR_i^4 - 1.9739PLR_i^3 \\ &\quad +1.5005PLR_i^2 - 0.2817PLR_i + 0.6838. \end{aligned} \tag{2.19}$$

The thermal power recovered from fuel cell is given by [82]

$$P_{F,i,Th}(t) = \gamma_{TE,i}(P_{F,i}(t) + P_a(t) + P_H(t)). \tag{2.20}$$

The cost due to thermal load is given by [82]

$$C_{F,i,Th}(t) = C_{n2}max(P_{L,i,Th}(t) - P_{F,i,Th}(t), 0). \tag{2.21}$$

2.2.3.3 Hydrogen Management Cost

The hydrogen storage cost is proportional to the hydrogen pumping cost. The hydrogen reservoir is assumed to have 95% storage efficiency [82]. The mathematical expression is

$$C_{F,i,H}(t) = C_{pump} P_H(t) \eta_{st}. \tag{2.22}$$

2.2.4 System structure

The micro-grid system for simulation consists of thermal generators, wind turbines and PEM fuel cells. The whole structure is shown in Fig. 2.1. The wind turbines are connected to the common ac bus through AC/DC and DC/AC inverters. The thermal generators are connected to the common ac bus through AC/DC and DC/AC inverters. The fuel cells are connected to the common ac bus through DC/AC inverters. The power generated by these generators is then transmitted to balance the distributed load.

2.3 Overview of SPEA2 and NSGA-II

Economic load dispatch is a multi-objective optimization problem because two different objectives: operating cost and emission, are involved [84]. When multiple objectives are conflicting, we may not be able to find a dominated solution, that is, there does not exist a solution that is strictly better than other solutions in terms of all objectives. Instead, we may have a set of non-dominated solutions in which a solution cannot be better than other

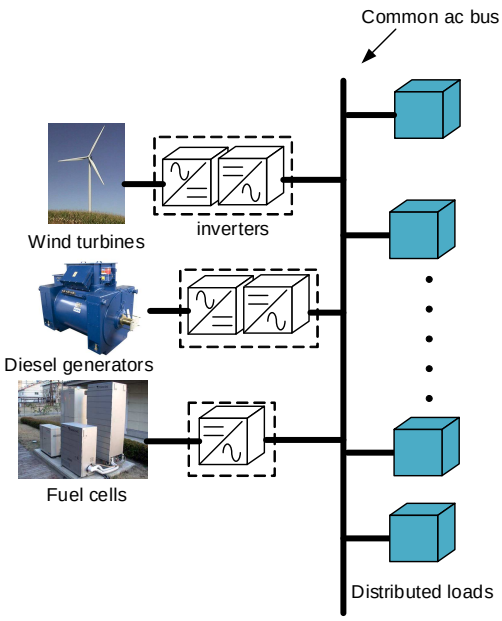


Figure 2.1: The micro-grid system structure

solutions in all objectives. When all solutions are non-dominated, we seek a pareto front that consists of optimal non-dominated solutions.

2.3.1 SPEA2

Strength pareto evolutionary algorithm 2 (SPEA2) is the state-of-the-art technique for finding the pareto-optimal solutions for multi-objective optimization problems [91]. Fig. 2.2 shows the workflow of SPEA2.

2.3.1.1 Initialization

Randomly generate N number of solutions as initial population P_0 and an empty archive A_0

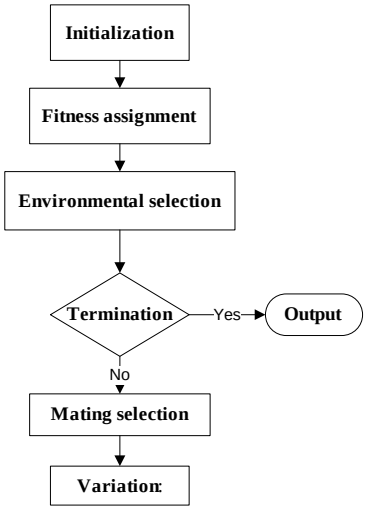


Figure 2.2: Workflow of SPEA2

2.3.1.2 Fitness assignment

Each individual i in the population P_t and archive A_t is assigned a strength value $S(i)$. The value is equal to the number of individuals dominated by i

$$S(i) = |\{j | j \in P_t \cup A_t \wedge i \succ j\}|, \tag{2.23}$$

where $|\bullet|$ denotes the cardinality of the set. Based on the S values, the raw fitness $R(i)$ of an individual i is

$$R(i) = \sum_{j \in P_t \cup A_t, j \succ i} S(j), \tag{2.24}$$

where $R(i)$ represents the sum of the individual i's dominators' strength values. The raw fitness is to be minimized. Additionally, the density information is introduced. The density is calculated as a function of kth nearest neighbor

$$D(i) = \frac{1}{\sigma_i^k + 2}, \tag{2.25}$$

where $k = \sqrt{N + N_A}$. The 2 in the denominator is added to prevent division by 0 and to keep $D(i) < 1$.

Thus the fitness of individual i is

$$f(i) = R(i) + D(i). \tag{2.26}$$

2.3.1.3 Environmental selection

Firstly, all non-dominated individuals are copied to the archive of next generation

$$A_{t+1} = \{i | i \in P_t \cup A_t \wedge f(i) < 1\}. \tag{2.27}$$

If the number of non-dominated individuals fits exactly into the archive, namely $|A_{t+1}| = N$, then the environmental selection is finished. If $|A_{t+1}| > N$, the archive truncation procedure is called. The individual i is chosen for removal for which $i \leq_d j$ for all $j \in A_{t+1}$ with

$$i \leq_d j \Leftrightarrow B \vee (C \wedge D), \tag{2.28}$$

where $B : (\sigma_i^k = \sigma_j^k, \forall 0 < k < |A_{t+1}|)$, $C : (\sigma_i^l = \sigma_j^l, \forall 0 < l < k, \exists 0 < k < |A_{t+1}|)$ and $D : (\sigma_i^k < \sigma_j^k, \exists 0 < k < |A_{t+1}|)$.

In other words, the individual i with minimum distance to another individual is chosen for removal. In the case of $|A_{t+1}| < N$, sort the $P_t + A_t$

individuals and copy the best $N - |A_{t+1}|$ individuals into A_{t+1}.

2.3.1.4 Termination

If maximum number of generations is reached or other criteria is met, output the set of non-dominated solutions in A_{t+1}. Otherwise continue the process.

2.3.1.5 Mating selection

Binary tournament selection is performed on A_{t+1} to fill the mating pool.

2.3.1.6 Variation

Use recombination and mutation operators to the mating pool and let A_{t+1} equal to the resulting population and go to Fitness assignment.

2.3.2 NSGA-II

Non-dominated sorting genetic algorithm (NSGA-II) with elitism is another state-of-the-art technique in multi-objective optimization using genetic algorithm [92]. Fig. 2.3 shows the workflow of NSGA-II.

2.3.2.1 Initialization

A parent population P_0 is randomly generated and each individual is assigned a fitness equal to its dominance rank.

2.3.2.2 Variation

Recombination, mutation and binary tournament selection is performed to produce a offspring Q_0 of size N. The total population is $P_0 \cup Q_0$. Thus

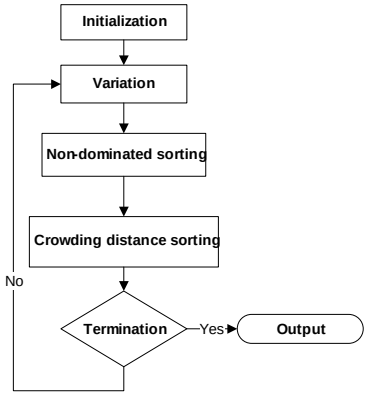


Figure 2.3: Workflow of NSGA-II

elitism is introduced.

2.3.2.3 Non-dominated sorting

Sort the population $P_0 \cup Q_0$ according to its dominance rank in ascending order.

2.3.2.4 Crowding distance sorting

After the Non-dominated sorting, all the individuals are ranked according to their dominance rank. The best individuals F_1, also known as the non-dominated individuals, must be emphasized. If the number of F_1 individuals is less than the population size, all F_1 individuals are selected for reproduction. Additionally, the second best individuals F_2 are also considered. The process repeats until N best individuals are selected.

As shown by Fig. 2.4, F_1 and F_2 are copied to the next generation population P_{t+1} because $(|F_1| + |F_2|) < N$. However, F_3 cannot be directly copied into P_{t+1} due to its large number of individuals. Thus crowding distance sorting is performed:

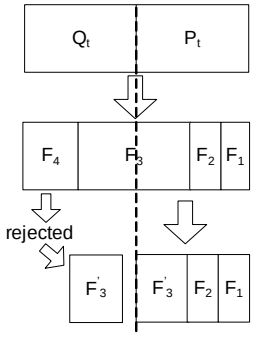


Figure 2.4: Crowding distance sorting

$$Dis(i) = Dis(i) + \left(\frac{I(i+1)_m - I(i-1)_m}{f_m^{max} - f_m^{min}} \right), \forall i \in I, i \neq 1, l, \qquad (2.29)$$

where $Dis(1) = Dis(l) = \infty$. After the Crowding distance sorting, the best $(N - |F_1| - |F_2|)$ individuals from F_3 are copied.

2.3.2.5 Termination

If maximum number of generations is satisfied or other criteria is met, output the set of non-dominated solutions in P_{t+1}. Otherwise continue the process.

2.4 Application Examples

This section presents the results of 12-generator test system, 33-generator test system, 54-generator, and 180-generator test system. These case studies are carried out on an Intel Core2 Duo E8400 with 4 GB memory machine. Matlab 2009 is used for all the case studies.

The optimal pareto-fronts are formed by running both SPEA2 and NSGA-

II 20 times. 2000 generations are performed for each time. The 40 sets of solutions are combined and the non-dominated solutions are selected as the optimal pareto-front.

Simulations are conducted with various number of generations to investigate the performance of the algorithms. Both algorithms are run 20 times for each number of generations. The hyper-volume indicator is calculated for every run. A hyper-volume indicator measures the the front obtained with respect to the optimal front reference point. The larger the hyper-volume indicator is, the closer the front converges towards the optimal front. The average and standard deviation values of the hyper-volume indicator for each number of generations are recorded in respective tables.

2.4.1 Case study 1: 12-generator test system

The optimization technique is performed on a 12-generator test system. The average and standard deviation values of the hyper-volume indicator for each number of generations are recorded in Table 2.1. As shown in Table 2.1, when the number of generation is 100, the mean hyper-volume indicator of SPEA2 is 0.0749 whereas that of the NSGA-II is 0.0594. SPEA2 has faster convergence performance initially. As the number of generations increases to 700, the mean hyper-volume indicator of SPEA2 is 0.0752 whereas that of NSGA-II is 0.0761. NSGA-II reaches closely to SPEA2. As the number of generations increases to 1500 and further, the mean hyper-volume indicator of SPEA2 is 0.0754 whereas that of NSGA-II is 0.0800. NSGA-II obtains better results.

Table 2.1: Mean and standard deviation for different number of generations by 12-generator system

No. of generations	Mean (SPEA2)	Deviation (SPEA2)	Mean (NSGA-II)	Deviation (NSGA-II)
100	0.0749	0.0096	0.0594	0.0104
300	0.0730	0.0075	0.0678	0.0141
500	0.0734	0.0057	0.0757	0.0087
700	0.0752	0.0098	0.0761	0.0103
900	0.0764	0.0066	0.0783	0.0106
1100	0.0750	0.0070	0.0757	0.0089
1300	0.0756	0.0108	0.0783	0.0111
1500	0.0771	0.0083	0.0790	0.0103
1700	0.0754	0.0077	0.0805	0.0068
1900	0.0754	0.0092	0.0800	0.0099

Figure 2.5 shows four examples of the optimal front and the two fronts obtained by the both algorithms. It is shown that SPEA2 converges faster than NSGA-II initially. However, NSGA-II catches up and obtains similarly good results when the number of generations increases. As shown by the plots, NSGA-II provides more diverse solutions than SPEA2.

Fig 2.5 are the fronts obtained by SPEA2 and NSGA-II algorithms. The x-axis of sub-figures in Fig 2.5 is the operating cost of all the generators. The y-axis of sub-figures in Fig 2.5 is the amount of emission by the generators. Since this problem is a two-objective problem. We cannot find a solution that is best in both objectives. Instead, we can only find solutions which are only better in one objective. These solutions form the "pareto front". The solutions on "pareto front" are all feasible and good solutions. It is up to the

user to decide which solution he will choose. From Fig 2.5, we can see that if we want to reduce the operating cost of the system, we need to move towards left-hand-side of the x-axis. However, if we do that, the amount of emission will go up. This can also be explained by practical case. In reality, renewable energy is more expensive. If we want to reduce the operating cost, we need to use more thermal energy instead of renewable energy. Nevertheless, using more thermal energy means burning more fuel. This results in more emission into the atmosphere. On the other hand, if we want to reduce emission, we need to use more renewable energy, which will incur higher operating cost. The environmental and economical impacts are shown by the "pareto front". Therefore, we do not provide a single dominating answer. Instead, we obtain a set of non-dominated solutions "pareto front", which contains minimization of both environmental and economical information, to the user. The user will then choose the solution based on his preferences.

Both algorithms take similar amount of time to finish the same number of generations. It takes 5 seconds for both algorithms to complete 100 generations. The time increases to 40 seconds for the 700 generations. The algorithms need one minute to finish 1100 generations. For 1500 generations, the algorithms need to run one and a half minute.

2.4.2 Case study 2: 33-generator test system

The optimization technique is performed on a 33-generator test system. The average and standard deviation values of the hyper-volume indicator for each number of generations are recorded in Table 2.2. As shown in Table 2.2,

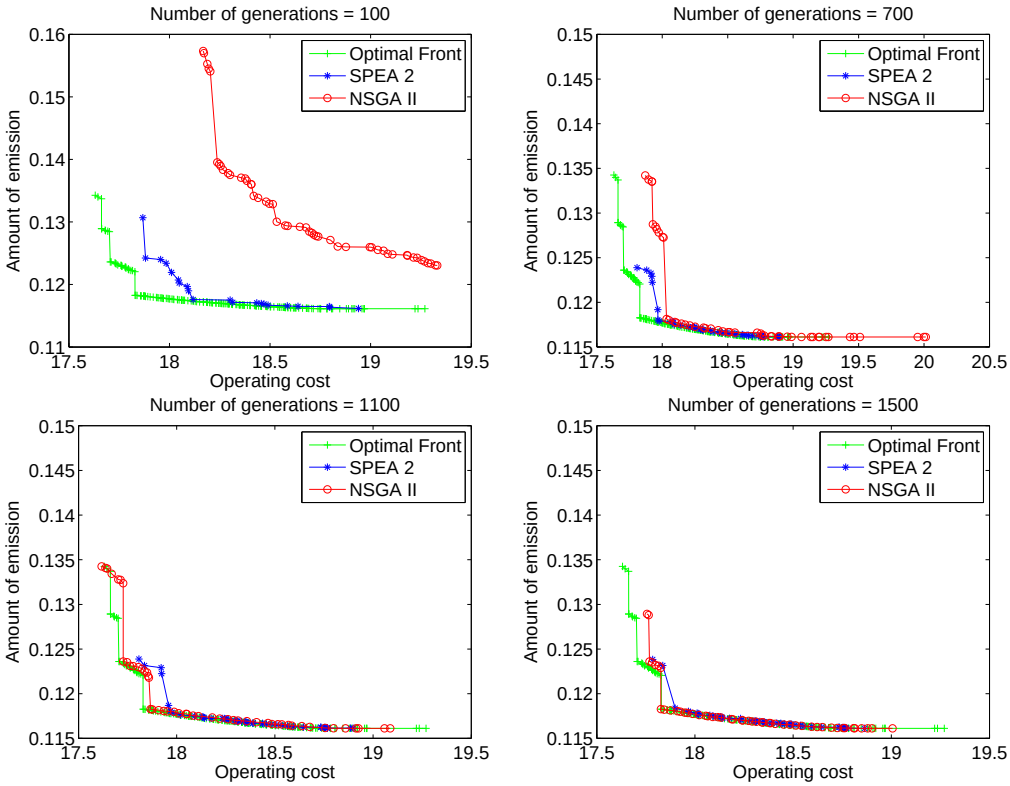


Figure 2.5: Optimal fronts obtained for various number of generations by 12-generator system

when the number of generation is 100, the mean hyper-volume indicator of SPEA2 is 0.0722 whereas that of the NSGA-II is 0.0367. SPEA2 has faster convergence performance initially. As the number of generations increases to 1300, the mean hyper-volume indicator of SPEA2 is 0.0822 whereas that of NSGA-II is 0.0803. NSGA-II reaches closely to SPEA2. As the number of generations increases to 1500 and further, the mean hyper-volume indicator of SPEA2 is 0.0800 whereas that of NSGA-II is 0.0868. NSGA-II overtakes SPEA2 by obtaining better results.

Figure 2.6 shows four examples of the optimal front and the two fronts obtained by the both algorithms. It is shown that SPEA2 converges faster

Table 2.2: Mean and standard deviation for different number of generations by 33-generator system

No. of generations	Mean (SPEA2)	Deviation (SPEA2)	Mean (NSGA-II)	Deviation (NSGA-II)
100	0.0722	0.0094	0.0367	0.0140
300	0.0774	0.0098	0.0603	0.0108
500	0.0769	0.0081	0.0685	0.0137
700	0.0826	0.0080	0.0728	0.0128
900	0.0793	0.0058	0.0691	0.0222
1100	0.0818	0.0061	0.0756	0.0203
1300	0.0822	0.0086	0.0803	0.0198
1500	0.0800	0.0080	0.0840	0.0083
1700	0.0825	0.0079	0.0868	0.0101
1900	0.0827	0.0076	0.0874	0.0116

than NSGA-II initially. However, NSGA-II catches up and obtains similarly good results when the number of generations increases. As shown by the plots, NSGA-II provides more diverse solutions than SPEA2.

Both algorithms take similar amount of time to finish the same number of generations. It takes 15 seconds to completes 100 generations. The time increases to 100 seconds for the 700 generations. The algorithms need three minutes to finish 1100 generations. For 1500 generations, the algorithms need to run 5 minutes.

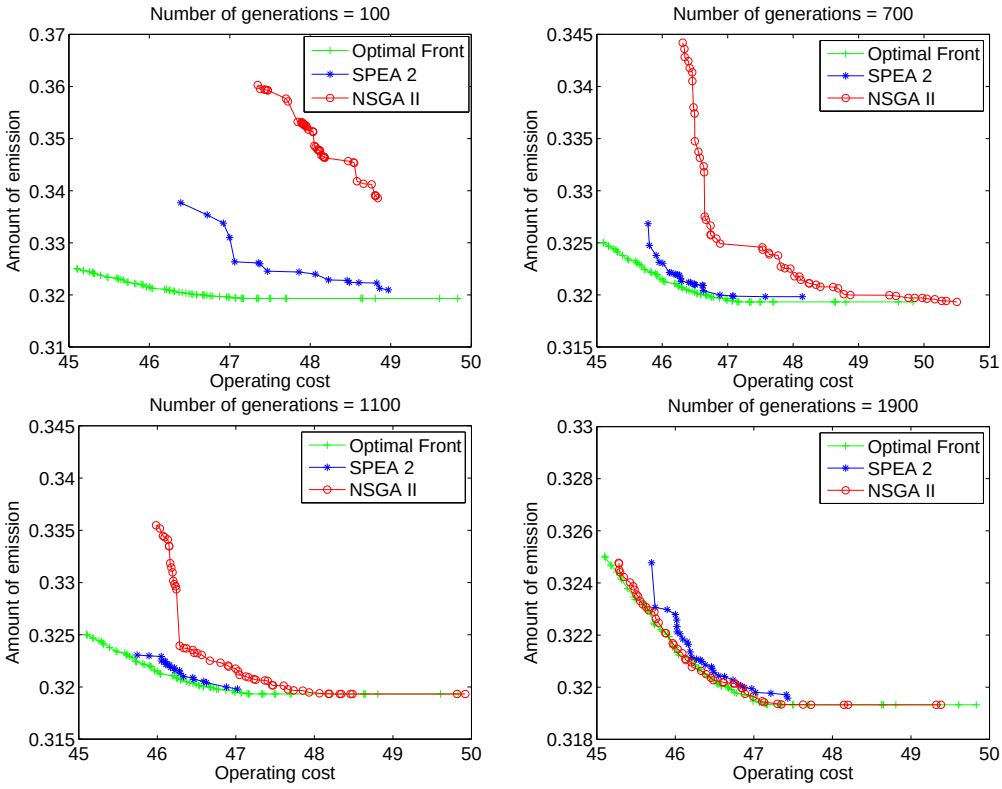


Figure 2.6: Optimal fronts obtained for various number of generations by 33-generator system

2.4.3 Case study 3: 54-generator test system

The optimization technique is performed on a 54-generator test system. The average and standard deviation values of the hyper-volume indicator for each number of generations are recorded in Table 2.3. As shown in Table 2.3, when the number of generation is 100, the mean hyper-volume indicator of SPEA2 is 0.0700 whereas that of the NSGA-II is 0.0430. SPEA2 has faster convergence performance initially. As the number of generations increases to 900, the mean hyper-volume indicator of SPEA2 is 0.0850 whereas that of NSGA-II is 0.0800. NSGA-II reaches closely to SPEA2. As the number of generations increases to 1500 and further, the mean hyper-volume indicator

of SPEA2 is 0.0841 whereas that of NSGA-II is 0.0875. NSGA-II outperforms SPEA2 by obtaining closer results.

Table 2.3: Mean and standard deviation for different number of generations by 54-generator system

No. of generations	Mean (SPEA2)	Deviation (SPEA2)	Mean (NSGA-II)	Deviation (NSGA-II)
100	0.0700	0.0086	0.0430	0.0108
300	0.0802	0.0075	0.0608	0.0095
500	0.0763	0.0111	0.0666	0.0101
700	0.0833	0.0111	0.0657	0.0104
900	0.0850	0.0086	0.0800	0.0082
1100	0.0841	0.0082	0.0847	0.0090
1300	0.0869	0.0089	0.0871	0.0117
1500	0.0841	0.0094	0.0875	0.0110
1700	0.0857	0.0078	0.0882	0.0100
1900	0.0859	0.0072	0.0891	0.0086

Figure 2.7 shows four examples of the optimal front and the two fronts obtained by the both algorithms. It is shown that SPEA2 converges faster than NSGA-II initially. However, NSGA-II catches up and obtains better results when the number of generations increases. As shown by the plots, NSGA-II provides more diverse solutions than SPEA2.

Both algorithms take similar amount of time to finish the same number of generations. It takes 25 seconds to completes 100 generations. The time increases to three minites for the 700 generations. The algorithms need five minutes to finish 1100 generations. For 1500 generations, the algorithms need to run six minutes.

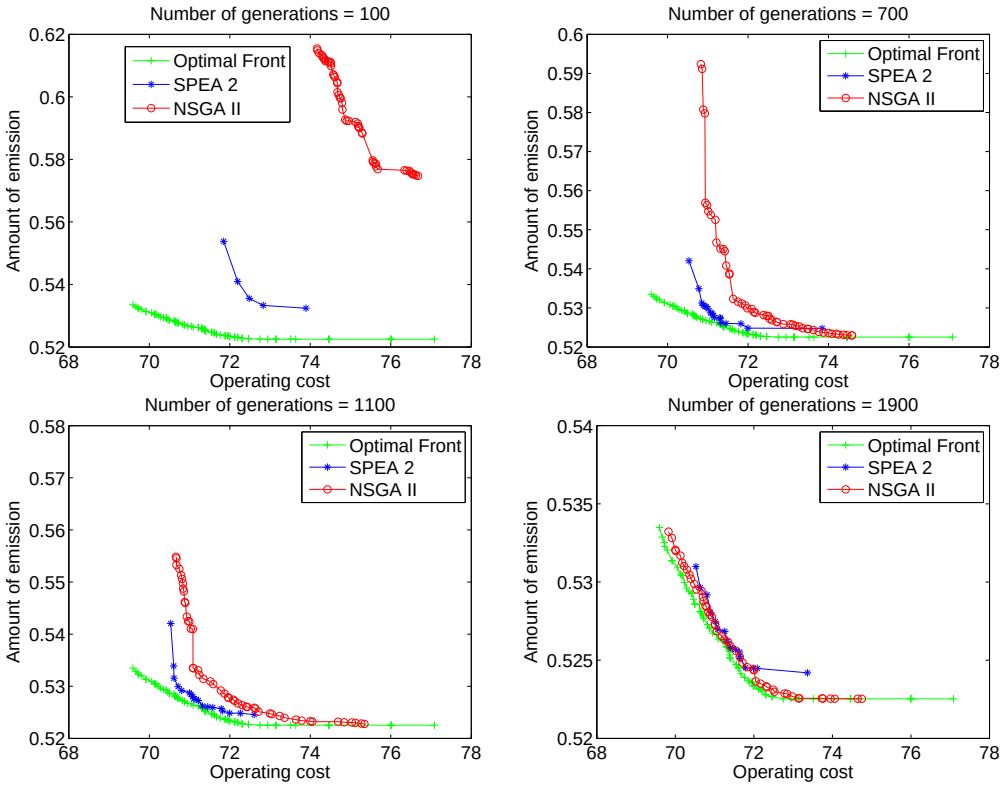


Figure 2.7: Optimal fronts obtained for various number of generations by 54-generator system

2.4.4 Case study 4: 180-generator test system

The optimization technique is performed on a 180-generator test system. The average and standard deviation values of the hyper-volume indicator for each number of generations are recorded in Table 2.4. As shown in Table 2.4, when the number of generation is 100, the mean hyper-volume indicator of SPEA2 is 0.0268 whereas that of the NSGA-II is 0.0128. SPEA2 has faster convergence performance initially. As the number of generations increases to 1500, the mean hyper-volume indicator of SPEA2 is 0.0632 whereas that of NSGA-II is 0.0634. NSGA-II reaches closely to SPEA2. As the number of generations increases to 1900, the mean hyper-volume indicator of SPEA2 is

0.0648 whereas that of NSGA-II is 0.0654. NSGA-II outperforms SPEA2 by obtaining closer results.

Table 2.4: Mean and standard deviation for different number of generations by 180-generator system

No. of generations	Mean (SPEA2)	Deviation (SPEA2)	Mean (NSGA-II)	Deviation (NSGA-II)
100	0.0268	0.0048	0.0128	0.0093
300	0.0479	0.0059	0.0332	0.0137
500	0.0539	0.0069	0.0666	0.0227
700	0.0594	0.0086	0.0249	0.0253
900	0.0570	0.0066	0.0443	0.0130
1100	0.0624	0.0084	0.0560	0.0214
1300	0.0618	0.0078	0.0578	0.0088
1500	0.0632	0.0075	0.0634	0.0097
1700	0.0641	0.0076	0.0645	0.0230
1900	0.0648	0.0100	0.0654	0.0319

Figure 2.8 shows four examples of the optimal front and the two fronts obtained by the both algorithms. It is shown that SPEA2 converges faster than NSGA-II initially. However, NSGA-II catches up and obtains better results when the number of generations increases. As shown by the plots, NSGA-II provides more diverse solutions than SPEA2.

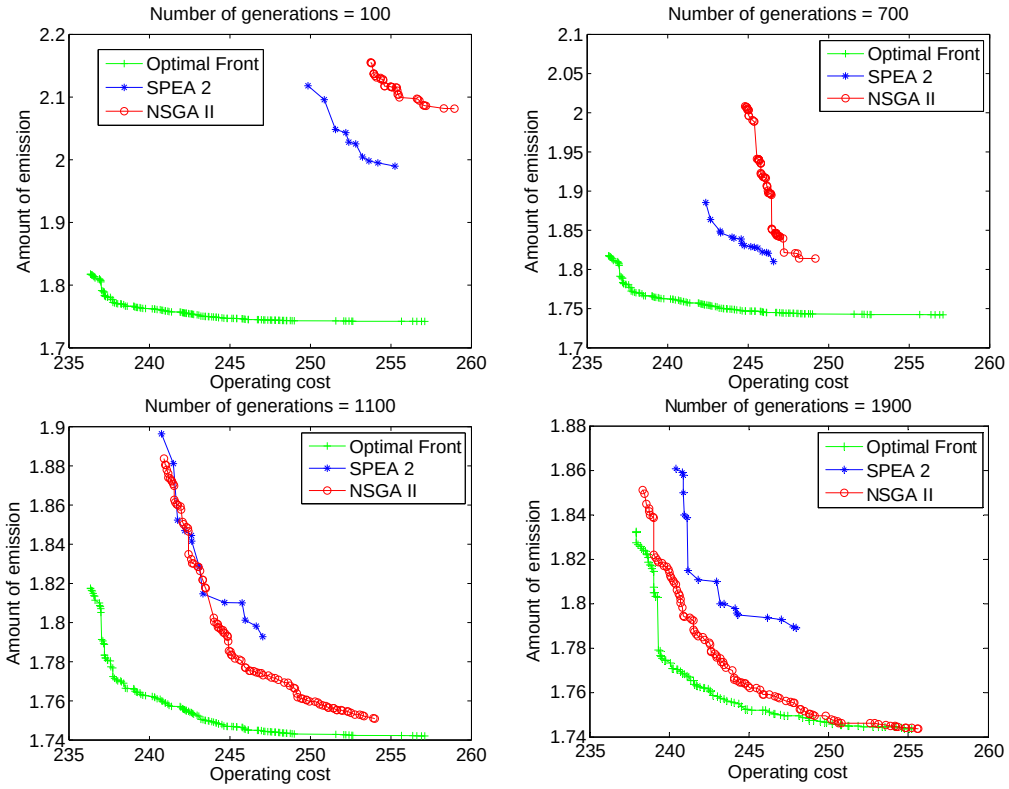


Figure 2.8: Optimal fronts obtained for various number of generations by 180-generator system

Both algorithms take similar amount of time to finish the same number of generations. It takes 33 seconds to completes 100 generations. The time increases to four minutes for the 700 generations. The algorithms need seven minutes to finish 1100 generations. For 1500 generations, the algorithms need to run eleven minutes.

2.5 Conclusion

In this chapter, we investigate the environmental and economical impacts of the thermal, wind turbine and PEM fuel cell generators, achieving the goal of

minimizing operating cost as well as minimizing emission. Furthermore, we investigate the performance of two state-of-the-art multi-objective optimization techniques on this problem. The techniques are performed on 3 different test systems. The results show that SPEA2 has a faster convergence when generation number is small and NSGA-II can perform better for large number of generations. NSGA-II provides more diverse solutions than SPEA2. It is suggested that SPEA2 is recommended if time is the most important concern. However, if the accuracy of the results is top priority, NSGA-II is preferred.

Chapter 3

Optimization of Distribution Network Incorporating Micro-grid: An Integrated Approach

3.1 Introduction

In 2003, the severe black out that took place in Eastern United States and Canada affected 50 million people. On September 23 of 2003, the network failure affected 2.4 million people in Eastern Denmark and Southern Sweden [93]. These severe blackouts have shown that the reliability of the power transmission is important in distribution networks. Furthermore, with the increase of power demand and environmental awareness, the power loss during transmission and the use of renewable energy resources have drawn much

attention. To address these issues, network reconfiguration and forming of micro-grids are typically used. Thus the integration between the distribution network and the distributed generators becomes a significant and complex problem.

The novelty of this work is that it proposes an integrated solution that takes care of both micro-grid load dispatch and network reconfiguration. The proposed scheme makes use of the power flow technique to minimize the total operating cost of a distribution network with multiple micro-grids. Despite higher renewable energy cost, the proposed integrated approach can still incorporate more renewable energy into the network by largely reducing the power losses on transmission lines. The scheme is also able to alter the network structure in order to handle faults occurred on the distribution feeders.

This chapter is organized into five sections. In Section 3.1, the background information and the literature reviews are introduced. In Section 3.2, the network reconfiguration, forecasting, system modeling and system constraints are presented. In Section 3.3, the optimization technique and implementation are discussed. The application examples are shown in Section 3.5. Finally, conclusions are made in Section 3.6.

3.2 Problem Formulation

This section discusses about the overall system structure, objective functions and constraints. After that power flow formulation and forecasting methods are introduced. Various models used for different types of generators are also

discussed.

3.2.1 System structure

The overall system structure is shown in Fig. 3.1. It describes the interactions between different modules. For distributed generators, the weather history data and generator data are input to the optimization and support vector regression to acquire parameters and forecast. After optimization, the system determines how many nodes are supported by the distributed generators and how much power each generator provides. For energy storage system, the state of charge information is the input to optimization module. After optimization, the system decides how much power the energy storage system charges or discharges. For utility grid, the market electricity price is the input to the optimization module. The optimization module decides how much power the network imports from utility grid. For power demand, history data are used to do forecasting. The optimization reconfigures the network and supplies power to meet the demand.

3.2.2 Objective and constraints

The objective is to minimize the operating cost $OC(t)$ of the whole distribution network. The $OC(t)$ consists of cost of micro-grid generators, cost of power losses on transmission lines, and cost of energy from utility grid as shown by (3.1)

$$OC(t) = C_{loss}(t) + C_{mg}(t) + C_{utility}(t), \tag{3.1}$$

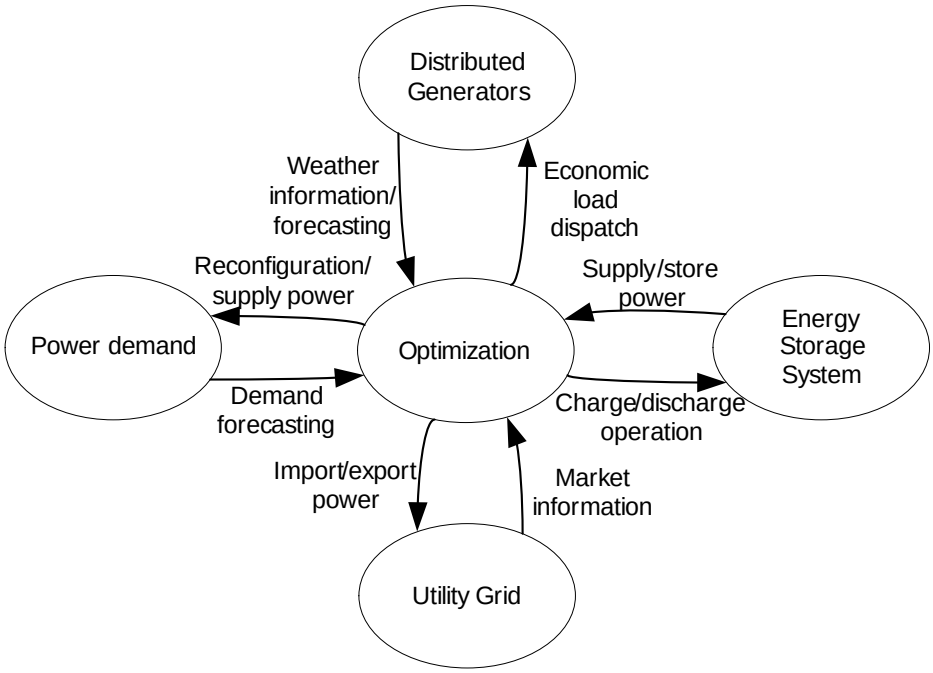


Figure 3.1: System structure of distributed network with network reconfiguration and micro-grid

The cost of micro-grid consists of synchronous generator's fuel and emission cost, solar energy operating cost, wind turbine operating cost as well as fuel cell energy cost and hydrogen management cost

$$C_{mg}(t) = \sum(C_{S,i}(t) + C_{factor} * E_{S,i}(t)) + \sum C_{PV,i}(t) + \sum C_{w,i}(t) + \sum(C_{F,i}(t) + C_{F,i,H}(t)), \tag{3.2}$$

subject to the following constrains

3.2.2.1 Power balance equation

The sum of the power purchased from utility grid and the total power generated by the synchronous generators, wind turbines, PV panels and fuel cells

in the micro-grid must be balanced by the local demand and the power loss on the transmission lines

$$\sum P_{S,i}(t) + \sum P_{W,i,ex}(t) + \sum P_{F,i}(t) + \sum P_{PV,i,ex}(t) + P_{utility}(t) = P_{Loss}(t) + P_{Dem}(t). \tag{3.3}$$

3.2.2.2 Spinning reserve

Spinning reserve is the reserved energy generation capacity during operation. The reserved capacities provide extra security and reliability for the system. It can be made available through transmission lines within a certain amount of time. the spinning reserve is considered because this is a large network consisting micro-grid. It is more difficult to deal with load or generators' variations after the reconfiguration of the network structure. Some of the nodes may not be reachable from utility grid after reconfiguration. Furthermore, the load demand is high compared to the generators' capacities. Thus the spinning reserve is needed to ensure that the promised load is supported constantly even with stochastic weather conditions

$$\sum S_i(t) P_i^{max}(t) \geqslant P_{Dem}(t) + P_R(t). \tag{3.4}$$

3.2.2.3 Generation limit

Every generator has its lowest and highest generation limit due to its physical constraints. The generator cannot operate beyond the physical limit,

otherwise the generator would be severely damaged

$$P_i^{min}(t) \leqslant P_i(t) \leqslant P_i^{max}(t). \tag{3.5}$$

3.2.2.4 Voltage constraint

The nodal voltage cannot be too high or too low. Otherwise it would cause serious problems to end users. It would cause power apparatus damage or instability in the power system. Moreover, it would cause unavailability of power for end user

$$V_j^{min}(t) \leq V_j(t) \leq V_j^{max}(t). \tag{3.6}$$

3.2.2.5 Power limit

The maximum power magnitude of a branch is limited due to the material of the transmission line. Excessive power on the line would damage the transmission element and result in disconnection

$$MVAf_j \leq MVAf_j^{max}. \tag{3.7}$$

3.2.3 Power flow formulation

The reconfiguration of a distribution network is a process that modifies the states of the sectionalizing switches and tie switches to isolate a fault in the network or to meet given optimal requirements such as minimizing power loss of the network, maintaining the power balance equation and reducing the load of the transformers.

Since we are using network reconfiguration techniques to re-structure the network, after reconfiguration, the network structure is in radial structure. Thus radial power flow formulation is used.

Considering a radial network in Fig. 3.2, a set of recursive equations is used to model the power flow in the radial network. The equations are shown as follows [94]

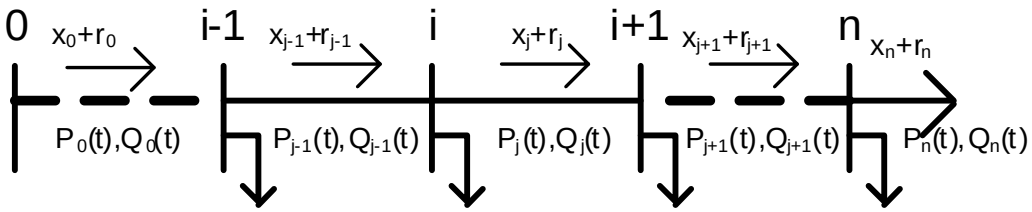


Figure 3.2: An example of power flow in a radial network

The power drawn by branch $j+1$ is equal to the power transmitted from branch j minus the sum of line loss on branch j and the load on branch $j+1$

$$P_{j+1}(t) = P_j(t) - r_j \frac{P_j(t)^2 + Q_j(t)^2}{V_j(t)^2} - P_{Lj+1}(t), \tag{3.8}$$

$$Q_{j+1}(t) = Q_j(t) - x_j \frac{P_j(t)^2 + Q_j(t)^2}{V_j(t)^2} - Q_{Lj+1}(t), \tag{3.9}$$

$$\begin{aligned} V_{j+1}(t)^2 &= V_j(t)^2 - 2(r_j P_j(t) + x_j Q_j(t)) + \\ &\quad (r_j^2 + x_j^2) \frac{P_j(t)^2 + Q_j(t)^2}{V_j(t)^2}. \end{aligned} \tag{3.10}$$

It is noticed that the quadratic terms in the (3.8-3.10) are much smaller than the branch power $P_j(t)$ and $Q_j(t)$. Therefore the equations can be

simplified by dropping quadratic terms

$$P_{j+1}(t) = P_j(t) - P_{Lj+1}(t), \tag{3.11}$$

$$Q_{j+1}(t) = Q_j(t) - Q_{Lj+1}(t), \tag{3.12}$$

$$V_{j+1}(t)^2 = V_j(t)^2 - 2(r_j P_j(t) + x_j Q_j(t)). \tag{3.13}$$

For the radial network, the branch power can be obtained by the following terms

$$P_{j+1}(t) = \sum_{k=j+2}^{n} P_{Lk}(t), \tag{3.14}$$

$$Q_{j+1}(t) = \sum_{k=j+2}^{n} Q_{Lk}(t), \tag{3.15}$$

$$V_{j+1}(t)^2 = V_j(t)^2 - 2(r_j P_j(t) + x_j Q_j(t)). \tag{3.16}$$

The power loss on a branch is calculated using

$$LP_j(t) = r_j \frac{P_j(t)^2 + Q_j(t)^2}{V_j(t)^2}. \tag{3.17}$$

Thus, the total power loss on the network is calculated as

$$P_{Loss}(t) = \sum_{j=0}^{n-1} r_j \frac{P_j(t)^2 + Q_j(t)^2}{V_j(t)^2}. \tag{3.18}$$

3.2.4 Forecasting

Due to the stochastic nature of the weather and load demand, forecasting unit is needed for the system in order to forecast the parameters and load

values for the next time unit. For the load, with previous eight mean and standard deviation values, it is able to forecast the mean and standard deviation values for the next time unit. With the forecasted values, it is possible to calculate the difference between the forecasted and actual load based on the probabilistic model. Together with the forecasted load value, it is able to obtain the interval that the actual load will fall in. Similarly, it is possible to forecast the mean and standard deviation parameters for PV generator as well as shape factors and scale factors for wind turbines. Then it is able to obtain the expected value of wind and PV generators power outputs based on the probabilistic models.

Support vector regression is a well-proved machine learning approach for time series prediction. It is based on statistical learning theory which allows it to generalize well for the unseen data. Given the training data $(x_1, y_1) \ldots, (x_i, y_i) \ldots (x_l, y_l)$, where x_i are the input vector and y_i are the corresponding output value, the support vector regression solves following minimization problem [95]

$$min \;\; \frac{1}{2} w^T w + C \sum_{i=1}^{l} (\xi_i + \xi_i^*) \tag{3.19}$$

subject to

$$\begin{aligned} y_i - (w^T \phi(x_i) + b) &\leq \epsilon + \xi_i^*, \\ (w^T \phi(x_i) + b) - y_i &\leq \epsilon + \xi_i, \\ \xi_i^*, \xi_i &\geq 0, i = 1, \ldots, l \end{aligned}$$

where $\frac{1}{2}w^T w$ is the regularization term and (ξ_i^*, ξ_i) are the upper and lower training errors subject to the ϵ-insensitive tube $|y - (w^T\phi(x_i) + b)| \leq \epsilon$ as shown in Fig. 3.3. C is the cost of error that controls the regression quality. ϵ is the width of the tube and ϕ is the kernel mapping function.

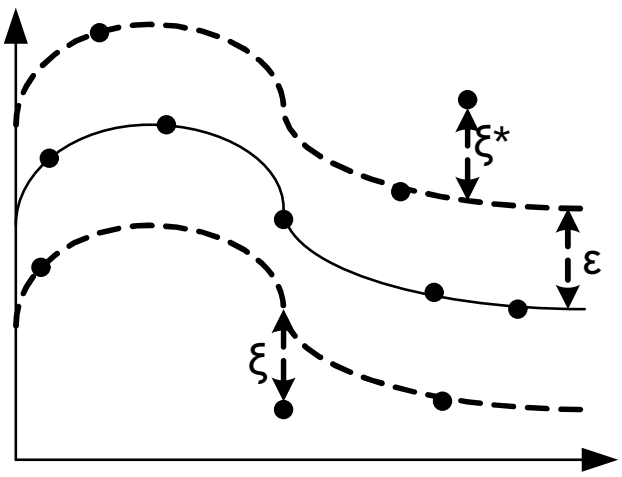


Figure 3.3: ϵ-insensitive band for SVR

The constraints of the objective equation (3.19) indicate that most data, x_i, are put in the tube $|y - (w^T\phi(x_i) + b)| \leq \epsilon$. If x_i is not in the tube, there are upper/lower training errors (ξ_i^*, ξ_i) which are going to be minimized in the objective function. Overfitting and underfitting are mitigated by minimizing the regularization term $\frac{1}{2}w^T w$.

Since x_i was transformed using kernel ϕ into higher dimensional feature space, it is more convenient to deal with the dual problem 3.20

$$min \;\; \frac{1}{2}(\alpha - \alpha^*)^T Q(\alpha - \alpha^*) + \epsilon \sum_{i=1}^{l}(\alpha_i - \alpha_i^*) + \sum_{i=1}^{l} y_i(\alpha_i - \alpha_i^*) \qquad (3.20)$$

subject to

$$\sum_{i=1}^{l}(\alpha_i - \alpha_i^*) = 0,$$
$$0 \leq \alpha_i, \alpha_i^* \leq C, i = 1, \ldots, l$$

where $Q_{ij} = \phi(x_i)^T \phi(x_j)$.

Fig. 3.4 shows an example of 2-day load demand forecasting of Singapore electricity market by support vector regression using data from Energy Market Company Pte Ltd (EMC). The load demand of 14 Mar 2011 and 15 Mar 2011 are used in this example. The data are measured every half an hour. Support vector regression uses previous 8 actual load data to forecast for the following time unit. Thus the actual data ranges from 1 to 96 whereas the forecasted data ranges from 9 to 96. The spline kernel function is used. The mean average percentage error (MAPE) rate of this forecasting is 0.67%.

3.2.5 Models

This sub-section introduces load models and various types of generator models. Instead of deterministic models, we use stochastic models to represent the load demand and outputs of the different types of generators.

3.2.5.1 Load

The actual peak load and the forecasted peak load will be different due to its stochastic nature. Load forecasting uncertainty thus is an important parameter in economic load dispatch. The difference is modeled as a normal

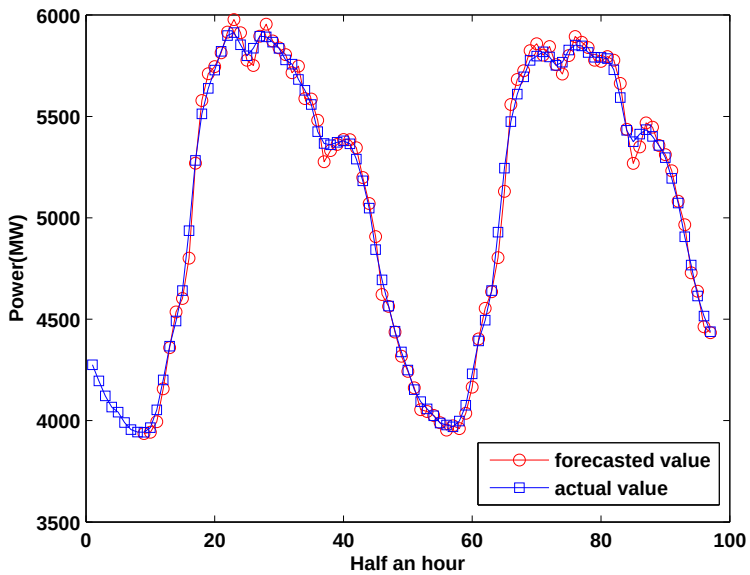


Figure 3.4: An example of 2-day load forecasting with error rate 0.67%

distribution function [96]

$$Diff(P_{Dem}(t)) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(P_{Dem}(t)-\mu)^2/2\sigma^2} \tag{3.21}$$

3.2.5.2 PV

The output power of PV generator proportional to solar irradiance and it is also affected by cell temperature. It is modeled as [97]

$$P_{PV,i}(t) = P_{STC}\frac{G_{ING}(t)}{G_{STC}}[1 + k_{PV}(T_c(t) - T_r)]. \tag{3.22}$$

Due to the stochastic nature of the weather, the irradiance probability

density function is modeled using Log-normal distribution

$$IPDF(G_{ING}(t)) = \frac{1}{\sqrt{2\pi ln(1+\frac{\sigma(t)^2}{\mu(t)^2})G_{ING}(t)}}$$
$$\cdot exp\left(-\frac{1}{2}\left(\frac{ln(G_{ING}(t)) - ln(\mu(t)) + \frac{1}{2}ln(1+\frac{\sigma(t)^2}{\mu(t)^2})}{\sqrt{ln(1+\frac{\sigma(t)^2}{\mu(t)^2})}}\right)^2\right). \quad (3.23)$$

The expected output power value can be calculated as

$$P_{PV,i,ex}(t+1) = \int_0^\infty P_{PV,i}(G_{ING}(t+1)) \cdot$$
$$IPDF(G_{ING}(t+1))dG_{ING}(t+1). \quad (3.24)$$

3.2.5.3 Wind turbine

Wind speed profile at a given location is modeled using Weibull density function. The wind speed distribution is further transformed into a wind power distribution as shown below [81]

$$f_w(P_{W,i}(t),t) = \frac{k(t)lv_i(t)}{c(t)}\left(\frac{(1+\rho l)v_i(t)}{c(t)}\right)^{k(t)-1}$$
$$\cdot exp\left(-\left(\frac{(1+\rho l)v_i(t)}{c(t)}\right)^{k(t)}\right), \quad (3.25)$$

for $0 < P_{W,i}(t) < P_{W,i,r}$, and

$$f_w(0,t) = 1 - exp\left(-\left(\frac{v_i(t)}{c(t)}\right)^{k(t)}\right) + exp\left(-\left(\frac{v_0(t)}{c(t)}\right)^{k(t)}\right), \quad (3.26)$$

$$f_w(P_{W,i,r},t) = exp\left(-\left(\frac{v_r(t)}{c(t)}\right)^{k(t)}\right) - exp\left(-\left(\frac{v_0(t)}{c(t)}\right)^{k(t)}\right), \tag{3.27}$$

where $\rho = \frac{P_{W,i}(t)}{P_{W,i,av}}$ and $l = \frac{v_r(t)-v_i(t)}{v_i(t)}$.

The expected value of power output is

$$\begin{aligned} P_{W,i,ex}(t+1) &= \int_0^{P_{W,i,r}} P_{W,i}(t+1) f_w(P_{W,i}(t+1), \\ &\quad t+1) \cdot dP_{W,i}(t+1) + P_{W,i,r} \cdot \\ &\quad f_w(P_{W,i,r}, t+1). \end{aligned} \tag{3.28}$$

3.2.5.4 Fuel cell

3.2.5.4.1 Fuel cost For PEM fuel cell, the fuel cost is linearly proportional to the sum of the power generated and the power consumed by the power for hydrogen production as shown by [82]

$$C_{F,i}(t) = C_n(\frac{P_{F,i}(t) + P_H}{\eta_i}). \tag{3.29}$$

3.2.5.4.2 Hydrogen Management Cost The hydrogen storage cost is proportional to the hydrogen pumping cost. The hydrogen reservoir is assumed to have 95% storage efficiency. The mathematical expression is [82]

$$C_{F,i,H}(t) = C_{pump} P_H \eta_{st}. \tag{3.30}$$

3.2.5.5 Synchronous generator

3.2.5.5.1 Fuel cost Synchronous generator has a number of valves that are opened according to the power output. As the output requirement increases, the generator opens one more valve to allow more steam to come out. However, when a valve is opened, the incremental heat rate rises rapidly. This valve-point effect results in non-smooth, non-convex input-output relationship. A recurring rectified sinusoid modeling the valve-point effect is added to the traditional quadratic cost function as shown in (3.31) [98]

$$C_{S,i}(t) = \frac{a_i}{2} P_{S,i}(t)^2 + b_i P_{S,i}(t) + c_i + |e_i \sin(f_i(P_{S,i}^{min} - P_{S,i}(t)))|. \quad (3.31)$$

3.2.5.5.2 Environmental cost Nitrogen-Oxide (NOx) emission represents the environmental impact of thermal generators. The amount of NOx emission is proportional to generator output as shown by [84]

$$E_{S,i}(t) = \theta_i + \beta_i P_{S,i}(t) + \gamma_i P_{S,i}(t)^2 + \zeta_i exp(\kappa_i P_{S,i}(t)). \quad (3.32)$$

Thus the cost of the emission would be the product of cost factor and the amount of emission ($C_{factor} * E_{S,i}(t)$).

3.2.5.6 Battery

Battery is a storage device which can reserve additional energy. Whenever the wind turbine and PV cannot meet the demand, the battery will be

discharged. The state of charge (SOC) of a battery is limited by

$$SOC_{min,i} \leqslant SOC_i(t) \leqslant SOC_{max,i}. \tag{3.33}$$

The cost of the battery is modeled as a constant $C_{B,i}$ per kWh.

3.3 Methodology

This section discusses the parameters acquiring process and optimization techniques. After that, the parameter settings and encoding strategy is introduced. Then it discusses the overall problem solving workflow for the readers to have a better understanding of the whole process.

3.3.1 Acquiring parameters

Since only wind speed and solar irradiance can be measured at a time, it is necessary to acquire the mean and standard deviation parameters from these data in order to obtain the distribution information and the expected values. In this chapter, the Vaccine-enhanced optimization technique is used to search for the parameters. Fig. 3.5 and Fig. 3.6 show the acquired wind speed and solar irradiance distribution information from the actual data. The data are provided by Solar Energy Research Institute of Singapore (SERIS). The irradiance and wind speed are measured every minute. The data of every half an hour containing 30 records are used to acquire the parameters for the probabilistic models. The blocks are the probability histograms of the actual data. The line with circles represents the acquired probability density

function. The parameter settings are the same as the optimization module which is shown in Section 3.5.

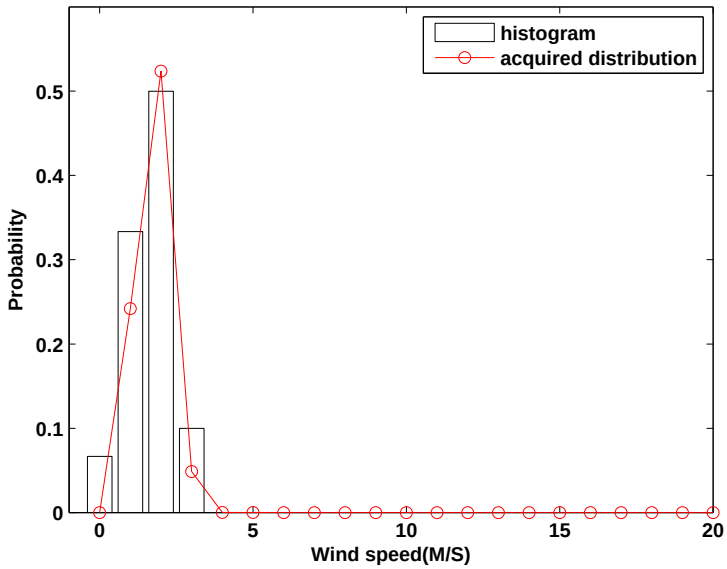


Figure 3.5: An example of wind speed distribution information acquisition

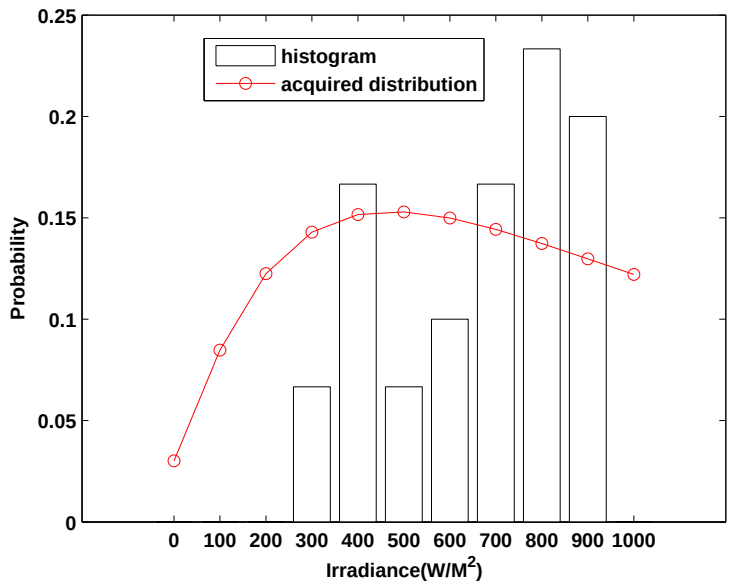


Figure 3.6: An example of solar irradiance distribution information acquisition

3.3.2 Optimization techniques

The optimization task of network reconfiguration together with economic load dispatch is a non-convex and highly nonlinear problem. Evolutionary computation techniques are suitable for solving this problem [69]. This chapter has investigated four different types of evolution computation techniques. They are Genetic Algorithm (GA) [99], Particle Swarm Optimization (PSO) [100], Vaccine-enhanced Artificial Immune System (V-AIS) [101], and Adaptive Vaccine-enhanced Artificial Immune System (AV-AIS). The first three algorithms are well-known in literature, but the fourth algorithm is proposed in this chapter. Thus this sub-section only introduces the adaptive Vaccine-enhanced Artificial Immune System.

The biological immune system has evolved over millions of years. The im-

mune system defends the body by using multilevel defense (either in parallel or sequential manner). The invader of the body will either be neutralized or destroyed by the immune system.

Two white blood cells are the most important cells in this immune process. They are T-cells and B-cells. They are both originated from bone marrow. T-cells are classified into helper T-cells, killer T-cells and suppressor T-cells. Helper T-cells are to activate B-cells. Killer T-cells are to inject poisonous chemicals to neutralize or kill the antigens. Suppressor T-cells are responsible for inhibiting the effects of other immune cells. Thus they can prevent allegoric reactions. B-cells are mainly in charge of production and secretion of antibodies.

3.3.2.1 Immune Network Theory

The immune network theory was proposed by Jern [102]. The theory stated that the immune system forms a idiotypic network of interconnected B-cells to recognize antigens. These B-cells both stimulate and suppress each other to stabilize the network. If the two B-cells affinity is above certain threshold, they are interconnected. The higher the affinity, the stronger the connection is.

3.3.2.2 Negative Selection mechanism

The purpose of the negative selection mechanism is to protect the self cells. It makes sure that the immune system only reacts to the antigens while not destroying self cells. When T-cells are made through the genetic rearrangement process, they will go through the negative selection process in

thymus. Those T-cells which bind to the self cells will be destroyed during this process. Therefore only those cells which do not bind to self cells can leave the system. Then those T-cells will circulate in the body and perform the immunologic function [103].

3.3.2.3 Clonal Selection Principle

Clonal selection principle is about how the immune system responses to an antigen. The new cells reproduced from their parents by mutation with high rate (hypermutation). When an antibody strongly bind to an antigen, the corresponding B-cell is stimulated to produce more antibodies. On the other hand, the immune system will suppress those antibodies carrying self-reactive receptors. New antibodies will help to explore new search space [104].

3.3.2.4 Vaccination

Vaccination is the injection of an antigen which can stimulate the immune response. Injection of infectious agents can create the immunological memories and let the immune response be more effective against the future encounters with these antigens. The vaccine extraction of a 2 dimensional problem is shown in Fig. 3.7 and the vaccination algorithm consists of following steps [105]

1. Divide the N-dimensional search space into $D_1 \times D_2 \times \ldots \times D_n$ grids. The width of the grid in kth dimension is

$$W_k = \frac{x_k^{max} - x_k^{min}}{D_k}, k = 1, \ldots, n. \tag{3.34}$$

Thus the boundary of the grids can be determined by

$$\begin{aligned} x_k^0 &= x_k^{min}, \\ x_k^i &= x_k^{min} + i \times W_k. \end{aligned} \tag{3.35}$$

2. In each grid, generate a random points $v_k^i \in [x_k^i, x_k^{i+1}]$ for every dimension by

$$v_k^i = x_k^i + Rand(0, W_k), i = 0, \ldots, D_k. \tag{3.36}$$

3. Generate vaccines *vaccine* by combing all the points from step 1 and step 2. For example

$$vaccine_m = v_1^i, \ldots, v_k^j, \ldots, v_n^l, vaccine_m \in vaccine. \tag{3.37}$$

4. Calculate the vaccine affinity $d_m r$ between vaccine $vaccine_m$ and every antibody ab_r in the population using Euclidian distance

$$d_m r = \|vaccine_m - ab_r\|. \tag{3.38}$$

   If the affinity is lower than predefined threshold σ, the vaccine is suppressed. The remaining are stored in the vaccine set.

5. Inject required number of vaccines into population and carry out the artificial immune optimization.

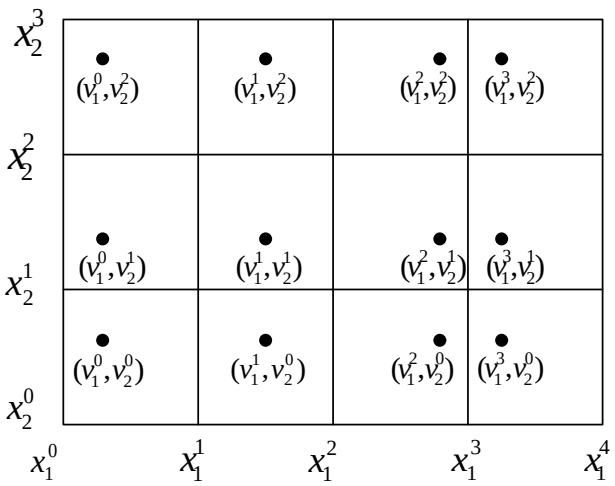


Figure 3.7: Vaccine extraction of a 2 dimensional search space

3.3.2.5 Adaptation

With the increasing number of iterations, the antigens tend to be generated in a relatively small search space. The vaccines are used to balance the global search and local search capabilities. Thus a sparse vaccine generation is needed in the exploration state and a dense vaccine generation is needed in the exploitation state. The adaptive rule of the vaccine size is proposed in (3.39)

$$D_k = floor(\frac{1}{0.04 + 0.215e^{-0.025t}}). \tag{3.39}$$

(3.39) makes sure that the number of grids in each search dimension increases from 4 to 25, which enforces faster exploration state and a more detailed exploitation state.

3.3.2.6 Adaptive vaccine-AIS Algorithm

The detailed algorithm is summarized in the following paragraphs. Fig. 3.8 shows the flowchart of Adaptive Vaccine-AIS algorithm.

1. Extract the vaccines $vaccine_k$ from the n-dimensional search space.

2. Initialize the antibodies randomly $ab_r, r = 1, \ldots, P$.

3. Evaluate the affinity (fitness) of the antibodies $f(ab_r), r = 1, \ldots, P$.

4. Carry out the affinity maturation steps as shown below.

   (a) Asexually clone c copies of each antibody in memory set where

   $$c = round(\beta P). \tag{3.40}$$

   β is a user specified parameter and $round()$ is the function which can round the number to the nearest integer. All the antibodies in set $C_r = \{ab_r^1, ab_r^2, \ldots, ab_r^{nc}\}$ have the same affinity value because they are clones.

   (b) Hypermutate each daughter in every set C_r. The mutation rate is inversely proportional to their affinity value

$$\begin{aligned} f_r(ab_r) &= \frac{f(ab_r)}{max f(ab_r)}, r = 1, \ldots, P \\ \alpha &= exp(-\rho f_r(ab_r)), r = 1, \ldots, P \\ ab_r^{1*} &= ab_r^1 \\ ab_r^{t*} &= ab_r^t + \alpha \cdot rand(-1, 1), t = 2, \ldots, nc; \\ & \quad r = 1, \ldots, P. \end{aligned} \tag{3.41}$$

ρ is a decay constant specified by the user. ab_r^1 does not go through mutation process to keep elitism. The other daughters in the set will mutate according to their affinity value. As a result, the new antibody set $C_r^* = \{ab_r^{1*}, ab_r^{2*}, \ldots, ab_r^{nc*}\}$ is generated.

5. Select the antibodies with the maximum affinity in each set $C_r^* = \{ab_r^{1*}, ab_r^{2*}, \ldots, ab_r^{nc*}\}, r = 1, \ldots, P$ and eliminate the other similar antibodies. The new memory antibodies are generated.

6. Calculate the antibody-to-antibody affinity using Euclidian distance $d_{ri} = \|ab_r - ab_i\|, r = 1, \ldots, P, i = 1, \ldots, P, i \neq r$. If the affinity value is less than a user defined threshold, the antibody with higher value is retained. The other one is deleted from the memory set. After this process, R number of antibodies are deleted and $P - R$ number of antibodies are remained.

7. Adaptively adjust the vaccine grid size according to the adaptation rule and extract vaccines $vaccine_k$.

8. Inject R number of vaccines. The population in memory set becomes P again.

9. If the termination criteria is not met, go to step 3. Otherwise, output the memory set antibodies.

3.3.3 Parametric setting

The parameters of V-AIS are: $\rho = 2, \beta = 0.1, d_{ri} = 0.1, generation = 100$ and $population = 50$. The parameters of PSO are: $\alpha = 1.9, c1 = 1.49, c2 =$

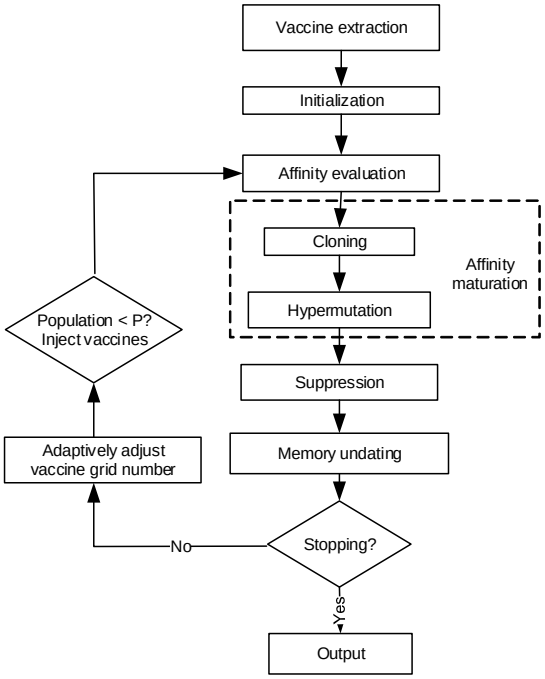


Figure 3.8: Adaptive vaccine-AIS flowchart

$1.49, generation = 100$ and $population = 50$. The parameters of GA are: $\rho_m = 0.005, \rho_c = 0.6, generation = 100$ and $population = 50$. The parameters of AV-AIS are: $\rho = 2, \beta = 0.1, generation = 100$ and $population = 50$. The parameters selected based on breadth search. The parametric space is searched with a small step size. For each method, the parameters which can give the best average results are selected. According to the prediction of Singapore Power and Sustainable Energy Association of Singapore newsletter, the PV electricity and wind electricity tariff is 15 cents per kWh and the utility electricity tariff is 20 cents per kWh in Singapore [106].

3.3.4 Encoding strategy

The encoding strategy of the network reconfiguration must make sure that the network is in radial structure. Fig. 3.9 shows an example of a 3-feeder network. The network has 16 branches (13 switches and 3 tie switches). There are 3 loops as shown. To preserve the radial structure of the network while generating as many feasible solutions as possible, each loop can only have one switch open at a time. The encoding strategy is to use a number to represent the opening switch in each loop. For instance, 3|2|3 represents that the third switch in loop 1, second switch in loop 2, and third switch in loop 3 are open. Each gene in a antibody (number) is between 1 and the number of switches in the loop. The advantage of this encoding strategy is that the percentage of feasible solutions generated are much higher compared to traditional encoding where each switch is represented by 1 binary bit.

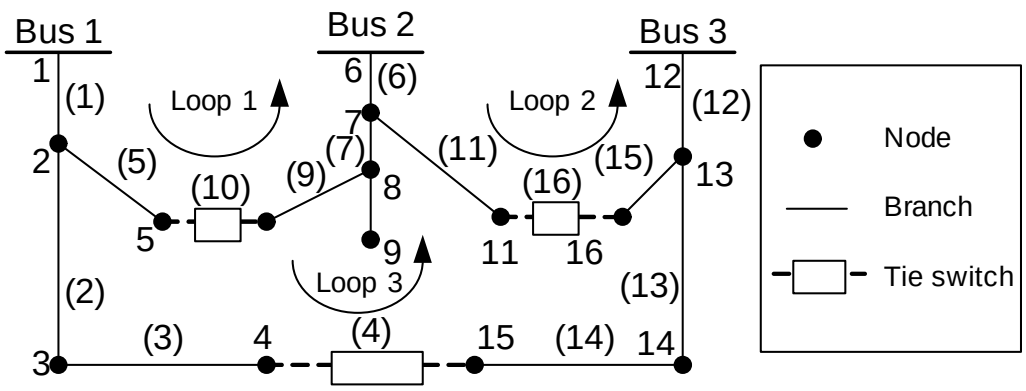


Figure 3.9: An example of 3 feeder network

The encoding strategy for the micro-grid is to use a gene to represent the number of branches to which distributed generators supply power in each direction. Then, the remaining branches in each direction are taken care by

the utility grid. For instance, the chromosome 1|3|2 represents 1 branch in direction A, 3 branches in direction B and 2 branches in direction C are dealt by distributed generators as shown in Fig. 3.10. The remaining branches are taken care by utility grid.

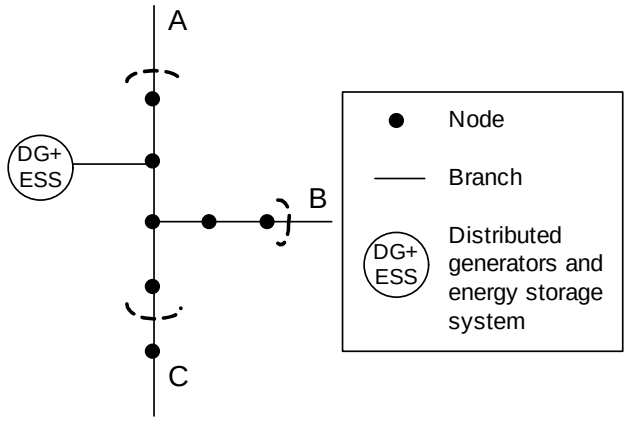


Figure 3.10: An example of nodes supported by DG and ESS

3.3.5 Overall workflow

Fig. 3.11 is the flowchart of the approach. Firstly, the weather and load history data are collected from measuring equipments. After the collections, these data are used to acquire the mean and standard deviation parameters for statistical model building as discussed in Section 3.3.1. Then these data are used to forecast the wind turbine output, PV output, and load demand using stochastic models and support vector regression by using the techniques discussed in Section 3.2.4 and Section 3.2.5. After that, this information is sent to the optimization module which contains optimization techniques as discussed in 3.3.2 together with state-of-charge information of energy storage devices. Then the optimization module takes care of the network encoding

strategy as part of the optimization procedures. After the calculation, it determines which nodes are supported by distributed generators and energy storage devices. It also determines the power output of each distributed generator. Furthermore, it determines which nodes are supported by utility grids to achieve overall operating cost minimization.

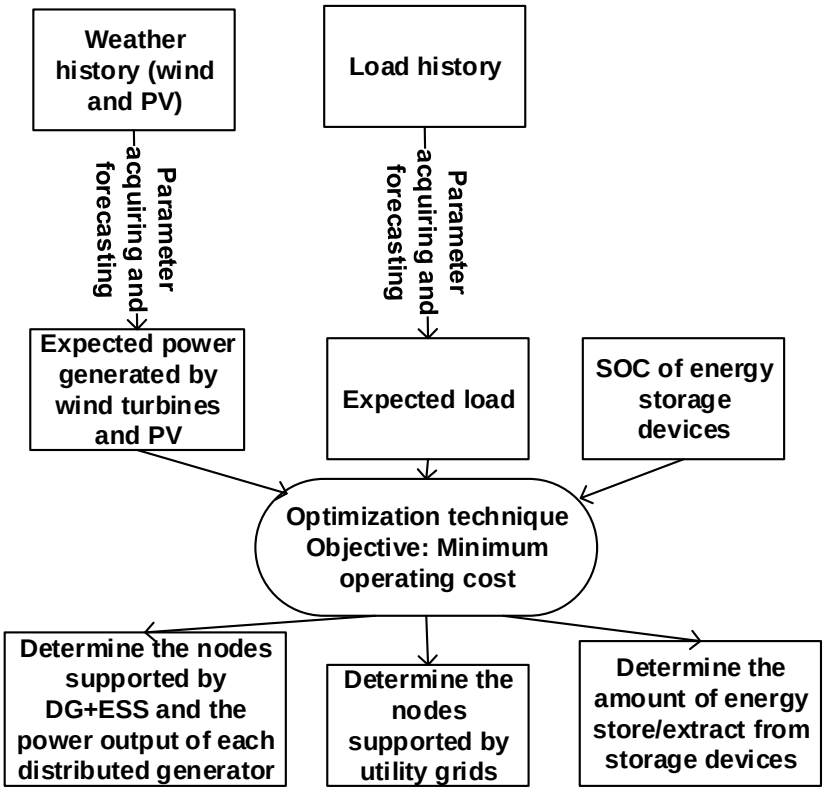


Figure 3.11: Overall system workflow

3.4 Integrated Approach

The novelty of this chapter is by first introducing the integrated formulation of the economic dispatch as well as the network reconfiguration in literature. This integrated approach has also taken stochastic load and generator

forecasting into consideration. The integrated approach allows the network to utilize more renewable energy, while minimizing the operating cost. The reason why it can achieve that is because the reconfigured network flow can significantly reduce the power losses, therefore the overall operating cost can be minimized by reducing the cost of power losses.

3.5 Application Example

For this application example, IEEE 33-node system [107] as shown in Fig. 3.12 is adopted to validate the method presented. This is a hypothetical system. This network consists one main feeder, 33 buses, 3 laterals, and 5 tie lines. The voltage of the transmission is $12.66kV$. Branch 33, 34, 35, 36, and 37 are normally open. The system has power demand $3615kW$ and the initial system power loss is $202.6kW$. Since our calculation time step is half an hour, we can calculate the initial cost for the IEEE 33-node test system as $(3615 + 202.6) * 0.2 * 0.5 = 381.76\ SGD$ per half an hour. The system inside the circle attached to node 15 as shown in Fig. 3.13 denotes distributed generators and energy storage system. The structure of the micro-grid is shown in Fig. 3.13 [108] and includes small generators, storage devices, and local distributed loads, which can operate both, connected to the grid or autonomously in island mode. The power loss on the transmission line is converted into part of the operating cost according to the market price.

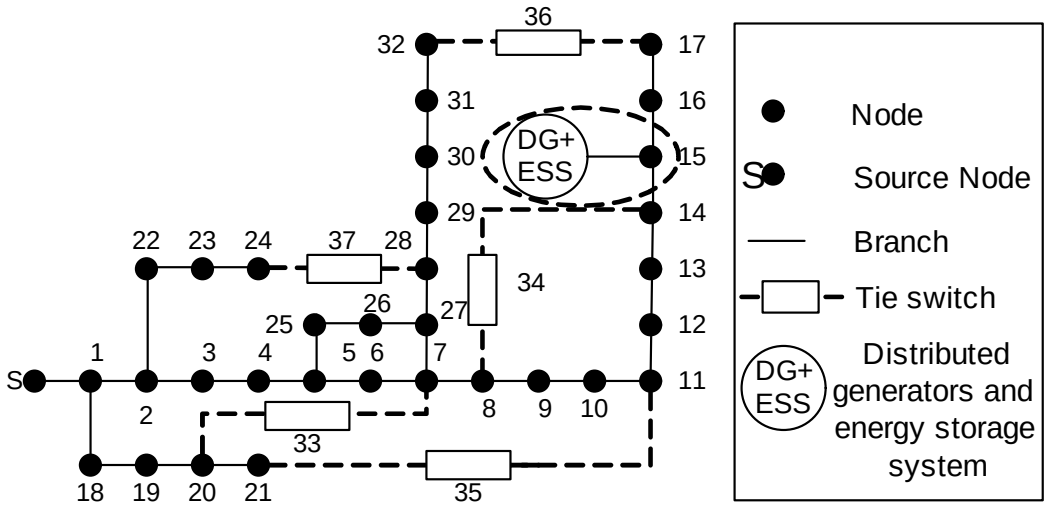


Figure 3.12: IEEE 33-node test system

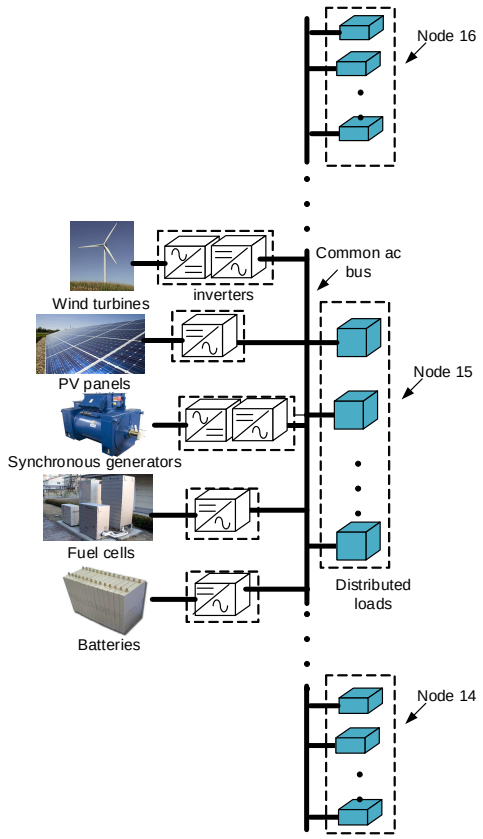


Figure 3.13: Structure inside the dashed ellipse: a flexible micro-grid

These following case studies are carried out on an Intel Core2 Duo E8400 with 4 GB memory machine. To be fair for all algorithms, each test is set to take 400 seconds to finish.

3.5.1 Case study 1: renewable energy at the same place

Firstly, case studies are conducted by using various renewable power values and using different optimization algorithms such as Vaccine-Enhanced Artificial Immune System (V-AIS), Particle Swarm Optimization (PSO), Genetic Algorithm (GA) and Adaptive Vaccine-enhanced Artificial Immune System (AV-AIS). Each type of the optimization techniques is run 10 times to evaluate the average performance. Table 3.1 shows the operating cost, power loss, power delivered by DGs and the opened switches under different renewable power status obtained by AV-AIS. In this case study, we assume that there is certain amount of renewable energy (ranges from 100 kW to 500 kW) available. 50% of the energy is from wind turbine, and another 50% energy is from PV generator. There are 3 diesel generators and 3 fuel cells working as thermal distributed generators attached to node 15. Each of the distributed generators has 30 kW capacity. Table 3.2 shows the operating cost obtained by 4 different optimization algorithms. The operating cost consists of cost of power loss, cost of power from utility grid, and cost of distributed generators as defined in (3.1).

As shown by Table 3.1, the operating cost reduces as the renewable energy power level increases. This is because transmitting electricity from utility grid to destination node incurs energy loss due to resistance of the transmis-

Table 3.1: Case study 1: results obtained by AV-AIS providing different renewable power.

wind/PV (kW)	total cost($)	power loss(kW)	power delivered by DGs (wind/PV/ other DERs)(kW)	switches opened
50/50	365.11	60.14	50/50/20	4-5, 16-17, 13-14, 8-9, 6-7
100/100	362.37	55.31	100/100/70	4-5, 31-32, 13-14, 8-9, 6-7
150/150	359.91	57.64	150/150/30	23-24, 31-32, 13-14, 8-9, 18-19
200/200	357.09	51.94	200/200/80	27-28, 30-31, 13-14, 8-9, 6-7
250/250	354.69	51.94	215/215/50	27-28, 30-31, 13-14, 8-9, 6-7

sion lines. The power transmitted from source node to node 15 incurs high power loss. Using renewable power can be more cost-effective by taking power loss cost into consideration. When there are $50kW$ wind and $50kW$ PV energy available, the system is reconfigured to form a radial structure network by opening switches 4-5, 16-17, 13-14, 8-9, 6-7. The system power loss is reduced to $60.14kW$. The $50kW$ wind energy and $50kW$ PV energy together with $20kW$ diesel/fuel cell energy are delivered to support node 15 ($60kW$) and node 14 ($60kW$). Thus the cost of purchasing energy from utility grid per half an hour is $(3615-120+60.14)*0.2*0.5 = 355.514SGD$. The cost of purchasing energy from micro-grid per half an hour is $365.11-355.514 = 9.596SGD$. Since micro-grid delivers $120kW$ power to utility grid, the average energy price from micro-grid is $9.596/0.5/120 = 0.160SGD/kWh$, which is correct with our price assumptions because we also have higher cost diesel/fuel

cost generators in the micro-grid. The penetration of micro-grid energy is $120/3615 \times 100\% = 3.3\%$, and the improvement of the total operating cost from the initial case is $(381.76 - 365.11)/381.76 \times 100\% = 4.4\%$.

As shown by the last row of Table 3.1, when there are $250kW$ wind and $250kW$ PV energy available, the system is reconfigured to form a radial structure network by opening 27-28, 30-31, 13-14, 8-9, 6-7. The system power loss is further reduced to $51.94kW$. The $215kW$ wind energy and $215kW$ PV energy together with $50kW$ diesel/fuel cell energy are delivered by micro-grid. The cost of purchasing energy from utility grid per half and hour is $(3615-480+51.94)*0.2*0.5 = 318.694SGD$. The cost of purchasing energy from micro-grid per half an hour is $354.69 - 318.694 = 35.996SGD$. Since there is $480kW$ power from micro-grid, the average energy price is $35.996/0.5/480 = 0.150SGD/kWh$, which is correct with our price assumptions. The penetration of micro-grid energy is $480/3615 \times 100\% = 13.3\%$, and the improvement of the total operating cost from the initial case is $(381.76 - 354.69)/381.76 \times 100\% = 7.1\%$.

On the other hand, the operating cost cannot be further reduced after 500 kW renewable power. This is because the system is restricted by voltage constraint and power limits. After 500 kW of renewable power, the transmitting power on branch 15-16 and branch 14-15 reach their limits. The limits are due to their physical material properties. The power flowing out from micro-grid is thus restricted to protect the transmission lines and equipments. Although more renewable energy is available, the system cannot fully utilize the renewable energy due to transmission limits.

In general, islanding operation is not allowed. In the case that the renew-

able power is 100 kW, the micro-grid supplies power to node 15. 100 kW is more than enough to support node 15 as node 15 consumes 60 kW. However, it is not enough to support 2 nodes. Thus, the system only supports node 15, and the excess energy is stored in the battery. In the case that the renewable energy is less than 60 kW, the micro-grid does not support any node. All the nodes are supported by utility grid.

Moreover, different scenarios are taken into consideration by this approach.

Scenario 1: renewable energy is not enough to support any node. In this case, all nodes are supported by utility grid. The available renewable energy is stored into battery.

Scenario 2: renewable energy is able to support a few nodes with some excess energy. Optimization of the whole network is performed to decide which nodes can be supported by distributed generators to minimize the operating cost. Then these decided nodes are supported by distributed generators. The rest are supported by utility grid. The excess energy is stored into battery.

Scenario 3: utility electricity price changes. When market price changes, optimization of the whole network is performed to decide which nodes can be supported by distributed generators to minimize the operating cost. Then these decided nodes are supported by distributed generators. The rest are supported by utility grid. The excess energy is stored into battery.

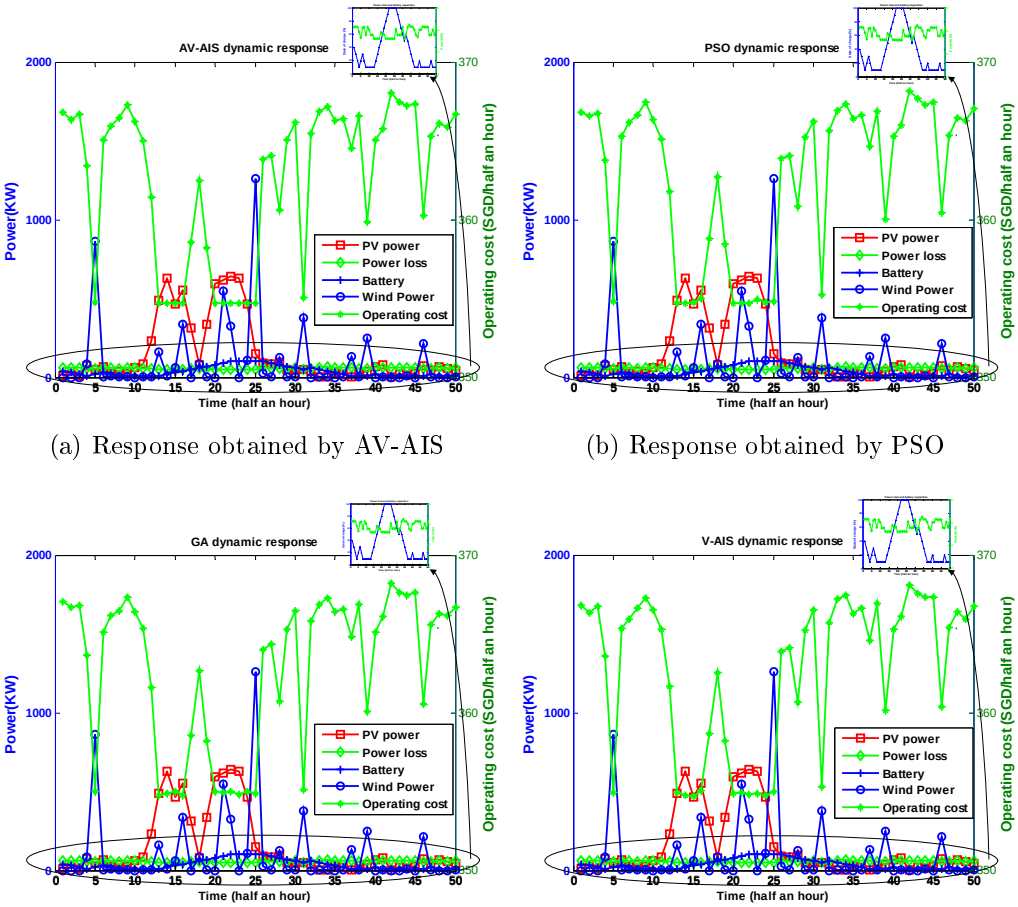
As shown in Table 3.2, four types of optimization algorithms are all capable of completing the task and obtaining convergent results. The results of the four algorithms are close. The reason is because they are all well established stochastic search algorithms and their parameters have been tuned by

Table 3.2: Case study 1: operating cost obtained by different algorithms.

wind/PV (kW)	operating cost by AV-AIS($)	operating cost by PSO($)	operating cost by GA($)	operating cost by V-AIS($)
50/50	365.10	365.23	365.25	365.45
100/100	362.34	362.42	362.65	362.68
150/150	359.97	360.13	359.97	360.24
200/200	357.04	357.15	357.44	357.19
250/250	354.68	354.74	354.93	354.86

breath method before the calculation. It is noticed that the results obtained by AV-AIS are slightly better. This is because AV-AIS is an improved adaptive optimization method based on V-AIS. Thus the performance is more promising.

Secondly, case studies are conducted by using actual weather data together with IEEE 33-node system to investigate the overall dynamic system response. In this dynamic experiment shown in Fig. 3.14, actual weather data of 14 March 2011 are used to investigate how the network responds to the dynamic weather over one day. In this experiment, there is little renewable energy in the early morning. The energy stored in the battery is used to satisfy load demand and the operating cost of the whole network is high. During time unit 4 to 6 (2:00AM to 3:00AM), there is large amount of wind energy available. This energy is thus used to satisfy more load demand and the operating cost of the network becomes lower. After that, from time unit 7 to 9 (3:30AM-4:30AM), there is little wind and PV energy. The operating cost of the network becomes high again. From time unit 11 to 19 (5:30AM-9:30AM), there is plenty of PV and some wind energy available. The network



(a) Response obtained by AV-AIS

(b) Response obtained by PSO

(c) Response obtained by GA

(d) Response obtained by V-AIS

Figure 3.14: Case study 1: dynamic experiment results

can adjust itself and make use of the renewable energy to reduce the total operating cost. However, the cost reduction is limited by line constraints and it cannot be further reduced below 354. Similarly, from time unit 21 to 24 (10:30AM to 12:00PM), the operating cost is lower due to abundant PV and some wind energy. After time unit 24 (12:00PM), little PV energy is received due to bad weather condition. The operating cost becomes high again. On the other hand, for time unit 31 (3:30PM), 39 (7:30PM) and 47 (11:30PM),

there is small amount of wind energy. The system can adjust itself to lower the operating cost by making use of the wind energy.

3.5.2 Case study 2: renewable energy at different places

In case study 2, wind and PV energy are at two different places as shown in Fig. 3.15. We assume that PV generators with another 6 DERs and energy storage devices are attached to node 15. Wind generators with another 6 DERs and energy storage devices are attached to node 24. In this case study, we assume that there is certain amount of renewable energy (ranges from 100 kW to 900 kW) available. 50% of the energy is from wind turbine, and another 50% energy is from PV generator. Table 3.3 shows the results obtained by AV-AIS with different level of renewable power. Table 3.4 shows the operating cost obtained by 4 different optimization algorithms.

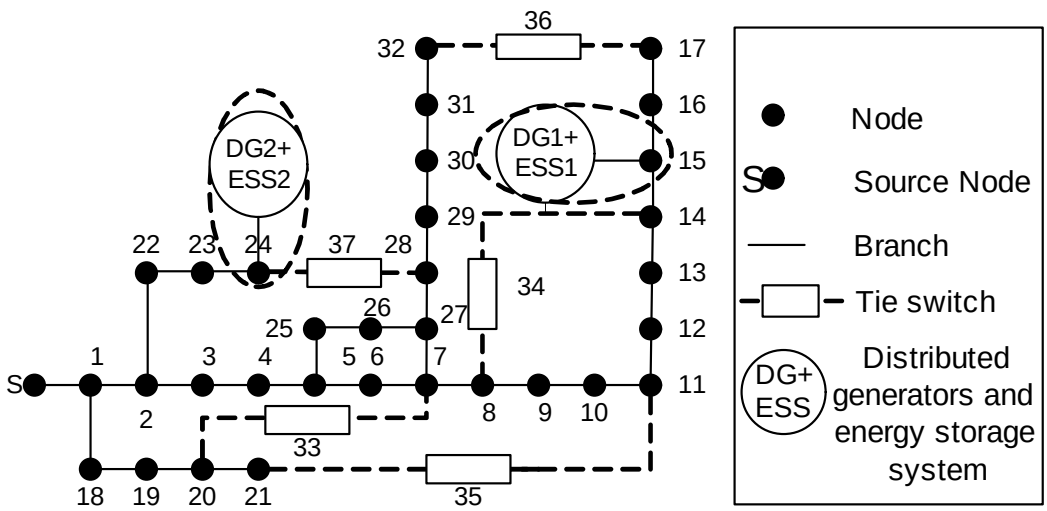


Figure 3.15: IEEE 33-node test system with renewable energy at different places

It is observed that, with the increasing penetration of the renewable ener-

gy, the total operating cost decreases due to reduction of the power loss. As shown in the first row of table 3.3, when there are $50kW$ wind and $50kW$ PV energy available, the system is reconfigured to form a radial structure network by opening switches 37-28, 15-16, 13-14, 8-9, 18-19. The system power loss is reduced to $55.70kW$. The $50kW$ wind energy together with $70kW$ diesel/fuel cell energy are delivered to support node 15 ($60kW$) and node 14 ($60kW$). Thus the cost of purchasing energy from utility grid per half an hour is $(3615 - 120 + 55.70) * 0.2 * 0.5 = 355.07SGD$. The cost of purchasing energy from micro-grid per half an hour is $366.26 - 355.07 = 11.19SGD$. Since micro-grid delivers $120kW$ power to utility grid, the average energy price from micro-grid is $11.19/0.5/120 = 0.187SGD/kWh$, which is correct with our price assumptions because we also have higher cost diesel/fuel cost generators in the micro-grid. The penetration of micro-grid energy is $120/3615 \times 100\% = 3.3\%$, and the improvement of the total operating cost from the initial case is $(381.76 - 366.26)/381.76 \times 100\% = 4.1\%$.

As shown in the last row of Table 3.3, when there are $450kW$ wind and $450kW$ PV energy available, the system is reconfigured to form a radial structure network by opening 37-28, 30-31, 13-14, 8-9, 6-7. The system power loss is further reduced to 42.10. The $450kW$ wind energy together with $30kW$ diesel/fuel cell energy are delivered by micro-grid 1. The $420kW$ PV power is delivered by micro-grid 2. The cost of purchasing energy from utility grid per half and hour is $(3615 - 480 - 420 + 42.10) * 0.2 * 0.5 = 275.71SGD$. The cost of purchasing energy from micro-grid per half an hour is $344.10 - 275.71 = 68.39SGD$. Since there is $900kW$ power from micro-grid, the average energy price is $68.39/0.5/900 = 0.152SGD/kWh$,

Table 3.3: Case study 2: results obtained by adaptive AV-AIS providing different renewable power.

wind/PV (kW)	total cost($)	power loss(kW)	power delivered by DGs (wind/other DERs1/ PV/Other DERs2)(kW)	switches opened
50/50	366.26	56.70	50/70/0/0	37-28, 15-16, 13-14, 8-9, 18-19
100/100	364.76	56.70	100/20/0/0	37-28, 15-16, 13-14, 8-9, 18-19
150/150	363.35	54.61	150/30/0/0	37-28, 16-17, 13-14, 8-9, 18-19
200/200	362.13	52.91	200/70/0/0	37-28, 31-32, 13-14, 8-9, 6-7
250/250	359.33	44.53	250/20/250/170	37-28, 31-32, 13-14, 8-9, 6-7
300/300	352.14	42.92	300/30/300/120	37-28, 31-32, 13-14, 8-9, 6-7
350/350	349.13	42.92	330/0/350/70	37-28, 31-32, 13-14, 8-9, 6-7
400/400	346.20	42.10	400/80/400/20	37-28, 30-31, 13-14, 8-9, 6-7
450/450	344.10	42.10	450/30/420/0	37-28, 30-31, 13-14, 8-9, 6-7

which is correct with our price assumptions because we also have higher cost diesel/fuel cost generators in the micro-grid. The penetration of micro-grid energy is $900/3615 \times 100\% = 24.5\%$, and the improvement of the total operating cost from the initial case is $(381.76-344.10)/381.76\times100\% = 9.9\%$. Compare with Case study 1, by allocating the PV generators to different location, the reduction of total operating cost increases from 7.1% to 9.9%, which is a significant improvement.

It is also noticed that those distributed generators attached to node 24 do not supply power to network from 50 kW wind power to 200 kW wind power. The reason is that the node 24 requires 420 kW. The 200 kW wind power and 180 kW thermal power is not enough to support node 24. Under that circumstance, the energy are stored into energy storage devices. The node 24 is supported by utility grid. From 250 kW wind power to 450 kW wind power, the distributed generators attached to node 24 support node 24. Since node 23 requires 420 kW, those distributed generators can not support two nodes even if wind power goes to $450kW$. After $900kW$ of renewable power, the transmitting power on branch 15-16 and branch 14-15 reach their limits. The limits are due to their physical material properties. Moreover, demand of node 23 and 24 far exceeds the available distributed energy. The power flowing out from micro-grid is thus restricted to protect the transmission lines and equipments. Although more renewable energy is available, the system cannot fully utilize the renewable energy due to transmission limits.

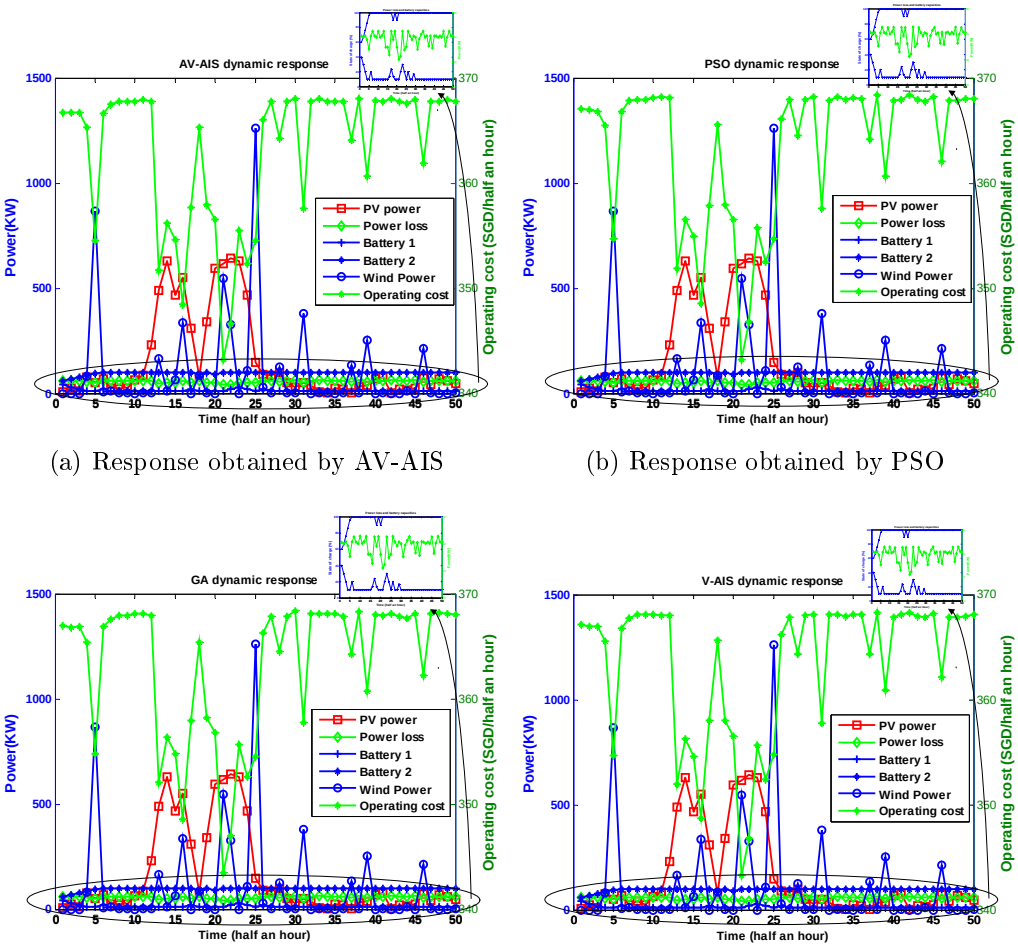
As shown in Table 3.4, four types of optimization algorithms are all capable of completing the task and obtaining convergent results. The results obtained by AV-AIS are slightly better than the other three algorithms. This

Table 3.4: Operating cost obtained by different algorithms.

wind/PV (kW)	operating cost by AV-AIS($)	operating cost by PSO($)	operating cost by GA($)	operating cost by V-AIS($)
50/50	366.23	366.52	366.40	366.43
100/100	364.72	365.02	364.90	364.92
150/150	363.32	363.61	363.49	363.51
200/200	362.10	362.39	362.27	362.30
250/250	359.30	359.59	359.47	359.49
300/300	352.10	352.39	352.27	352.30
350/350	349.10	349.38	349.26	349.29
400/400	346.17	346.45	346.33	346.36
450/450	344.07	344.35	344.23	344.26

is because AV-AIS is an improved algorithm from V-AIS. Thus the performance is more promising.

Furthermore, case studies are conducted by using actual weather data together with IEEE 33-node system to investigate the overall dynamic system response as shown in Fig. 3.16. In this dynamic experiment shown in Fig. 3.16, actual weather data of 14 March 2011 are used to investigate how the network responds to the dynamic weather over one day. In this experiment, there is little renewable energy in the early morning. The energy stored in the battery is used to satisfy load demand and the operating cost of the whole network is high. During time unit 4 to 6 (2:00AM to 3:00AM), there is large amount of wind energy available. This energy is thus used to satisfy more load demand and the operating cost of the network becomes lower. After that, from time unit 7 to 9 (3:30AM-4:30AM), there is little wind and PV energy. The operating cost of the network becomes high again.



(a) Response obtained by AV-AIS

(b) Response obtained by PSO

(c) Response obtained by GA

(d) Response obtained by V-AIS

Figure 3.16: Case study 2: dynamic experiment results

From time unit 11 to 19 (5:30AM-9:30AM), there is plenty of PV and some wind energy available. The network can adjust itself and make use of the renewable energy to reduce the total operating cost. However, the cost reduction is limited by line constraints and it cannot be further reduced below 344. Similarly, from time unit 21 to 24 (10:30AM to 12:00PM), the operating cost is lower due to abundant PV and some wind energy. After time unit 24 (12:00PM), little PV energy is received due to bad weather condition.

The operating cost becomes high again. On the other hand, for time unit 31 (3:30PM), 39 (7:30PM) and 47 (11:30PM), there is small amount of wind energy. The system can adjust itself to lower the operating cost by making use of the wind energy.

3.5.3 Case study 3: fault occurrence

In case study 3, branch 9-10 is assumed to be faulty as shown in Fig. 3.17, which means branch 9-10 is damaged and disconnected due to various reasons. The system is required to reconfigure itself to bypass the faulty branch as well as minimizing operating cost. The results are shown in Table 3.5. As shown in Table 3.5, the operating cost decreases as the renewable energy level increases. The system can adjust itself to lower the operating cost by making use of the renewable energy.

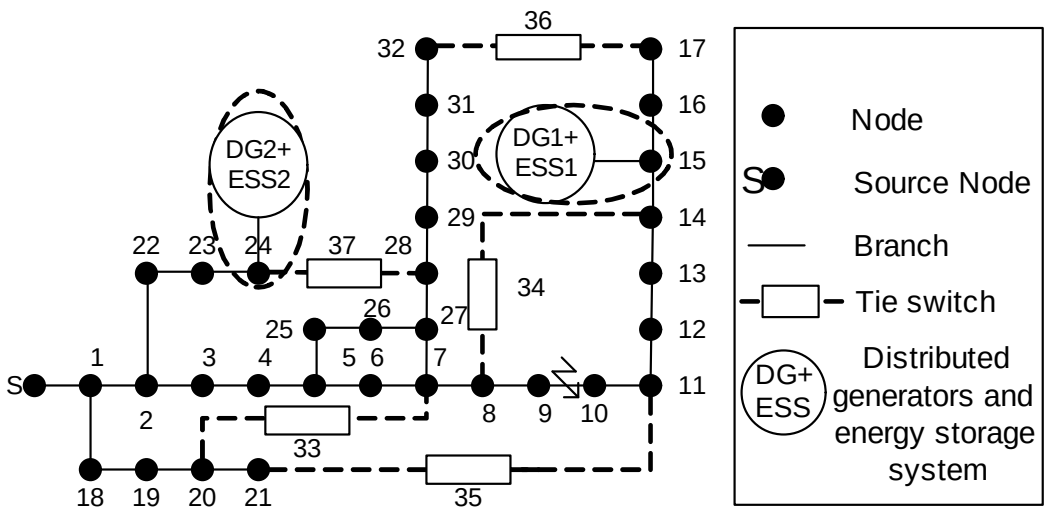


Figure 3.17: IEEE 33-node test system with fault on 9-10

It is observed that, with the increasing penetration of the renewable en-

Table 3.5: Case study 3: results obtained by AV-AIS providing different renewable power.

wind/PV (kW)	total cost($)	power loss(kW)	power delivered by DGs (wind/other DERs1/ PV/other DERs2)(kW)	switches opened
50/50	366.26	56.70	50/70/0/0	37-28, 15-16, 13-14, 9-10, 18-19
100/100	364.76	56.70	100/20/0/0	37-28, 15-16, 13-14, 9-10, 18-19
150/150	363.36	54.61	150/30/0/0	37-28, 16-17, 12-13, 9-10, 18-19
200/200	362.13	52.91	200/70/0/0	37-28, 31-32, 13-14, 9-10, 6-7
250/250	359.33	44.53	250/20/250/170	37-28, 31-32, 13-14, 9-10, 6-7
300/300	352.13	42.92	300/30/300/120	37-28, 31-32, 13-14, 9-10, 6-7
350/350	349.13	42.92	330/0/350/70	37-28, 31-32, 13-14, 9-10, 6-7
400/400	346.19	42.10	400/80/400/20	37-28, 30-31, 13-14, 9-10, 6-7
450/450	344.10	42.10	450/30/420/0	37-28, 30-31, 13-14, 9-10, 6-7

ergy, the total operating cost decreases due to reduction of the power loss. As shown in the first row of Table 3.5, when there are $50kW$ wind and $50kW$ PV energy available, the system is reconfigured to form a radial structure network by opening switches 37-28, 15-16, 13-14, 9-10, 18-19. The system power loss is reduced to $56.70kW$. The $50kW$ wind energy together with $70kW$ diesel/fuel cell energy are delivered to support node 15 ($60kW$) and node 14 ($60kW$). Thus the cost of purchasing energy from utility grid per half

an hour is $(3615-120+56.70)*0.2*0.5 = 355.17SGD$. The cost of purchasing energy from micro-grid per half an hour is $366.26-355.17 = 11.09SGD$. Since micro-grid delivers $120kW$ power to utility grid, the average energy price from micro-grid is $11.09/0.5/120 = 0.185SGD/kWh$, which is correct with our price assumptions because we also have higher cost diesel/fuel cost generators in the micro-grid. The penetration of micro-grid energy is $120/3615 \times 100\% = 3.3\%$, and the improvement of the total operating cost from the initial case is $(381.76-366.26)/381.76 \times 100\% = 4.1\%$.

As shown in the last row of Table 3.5, when there are $450kW$ wind and $450kW$ PV energy available, the system is reconfigured to form a radial structure network by opening 37-28, 30-31, 13-14, 9-10, 6-7. The system power loss is further reduced to 42.10. The $450kW$ wind energy together with $30kW$ diesel/fuel cell energy are delivered by micro-grid 1. The $420kW$ PV power is delivered by micro-grid 2. The cost of purchasing energy from utility grid per half and hour is $(3615-480-420+42.10)*0.2*0.5 = 275.71SGD$. The cost of purchasing energy from micro-grid per half an hour is $344.10-275.71 = 68.39SGD$. Since there is $900kW$ power from micro-grid, the average energy price is $68.39/0.5/900 = 0.152SGD/kWh$, which is correct with our price assumptions because we also have higher cost diesel/fuel cost generators in the micro-grid. The penetration of micro-grid energy is $900/3615 \times 100\% = 24.9\%$, and the improvement of the total operating cost from the initial case is $(381.76-344.10)/381.76 = 9.9\%$. Compare with Case study 2, this case study obtains the same reduction of 9.9%, which shows that, despite the fault occurrence, the proposed technique can still reduce the total operating cost significantly.

The transmitting power on branch 15-16 and branch 14-15 reach their limits after $900kW$ renewable energy. The limits are due to their physical material properties. Moreover, demand of node 23 and 24 far exceeds the available distributed energy. Furthermore, the system structure is more restricted due to the faulty branch 9-10. The power flowing out from micro-grid is thus restricted to protect the transmission lines and equipments.

3.5.4 Case study 4: fixed structure

In case study 4, the network is assumed to have a fixed structure as shown in Fig. 3.18. The purpose of the case study is to show that the proposed technique is able to benefit the network users by optimizing the operating cost even if network reconfiguration is not available due to various reasons.

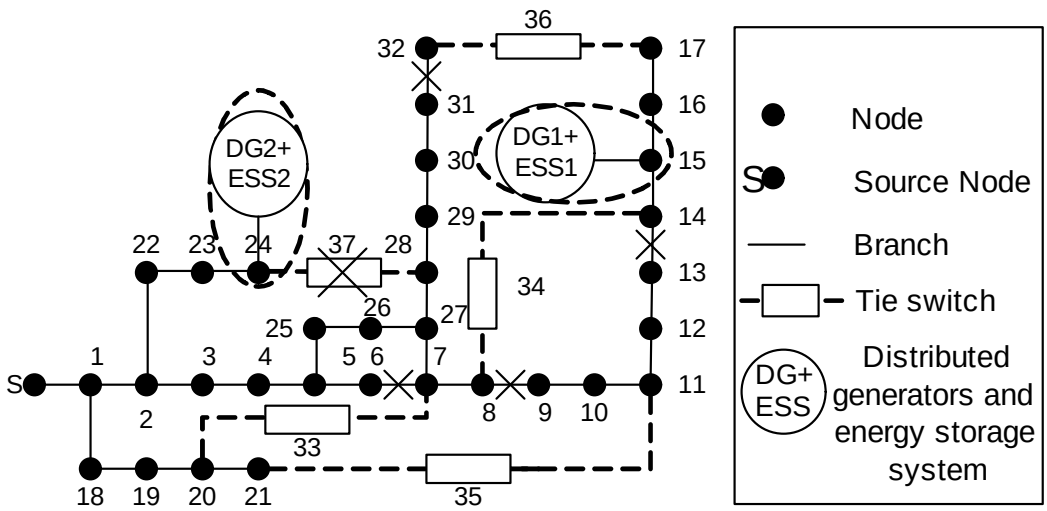


Figure 3.18: IEEE 33-node test system with fixed structure

Table 3.6: Case study 4: results obtained by AV-AIS providing different renewable power.

wind/PV (kW)	total cost($)	power loss(kW)	power delivered by DGs wind/other DERs1/ PV/other DERs2(kW)	switches opened
50/50	375.99	144.90	0/0/0/0	37-28, 31-32, 13-14, 8-9, 6-7
100/100	371.89	71.10	100/170/0/0	37-28, 31-32, 13-14, 8-9, 6-7
150/150	366.06	71.10	150/120/0/0	37-28, 31-32, 13-14, 8-9, 6-7
200/200	363.95	71.10	200/70/0/0	37-28, 31-32, 13-14, 8-9, 6-7
250/250	361.08	62.09	250/20/250/170	37-28, 31-32, 13-14, 8-9, 6-7
300/300	353.88	60.45	300/30/300/120	37-28, 31-32, 13-14, 8-9, 6-7
350/350	350.88	60.45	330/0/350/70	37-28, 31-32, 13-14, 8-9, 6-7
400/400	349.38	60.45	330/0/400/20	37-28, 31-32, 13-14, 8-9, 6-7
450/450	348.79	60.45	330/0/420/0	37-28, 31-32, 13-14, 8-9, 6-7

The results are shown in Table 3.6. Similarly, it is observed that, with the increasing penetration of the renewable energy, the total operating cost decreases due to reduction of the power loss. As shown in the first row of Table 3.6, when there are $50kW$ wind and $50kW$ PV energy available, the system power loss is $144.9kW$. The system power loss is very high due to no node is support by the micro-grids. The penetration of micro-grid energy is $0/3615 \times 100\% = 0\%$, and the improvement of the total operating cost from

the initial case is $(381.76 - 375.99)/381.76 = 1.5\%$.

As shown in the last row of Table 3.6, when there are $450kW$ wind and $450kW$ PV energy available, the system power loss is reduced to 60.45. The $330kW$ wind energy is delivered by micro-grid 1. The $420kW$ PV energy is delivered by micro-grid 2. The cost of purchasing energy from utility grid per half and hour is $(3615-480-420+60.45)*0.2*0.5 = 292.545SGD$. The cost of purchasing energy from micro-grid per half an hour is $348.79 - 292.545 = 56.245SGD$. Since there is $750kW$ power from micro-grid, the average energy price is $56.245/0.5/750 = 0.150SGD/kWh$, which is correct with our price assumptions. The penetration of micro-grid energy is $750/3615 \times 100\% = 20.7\%$, and the improvement of the total operating cost from the initial case is $(381.76 - 348.79)/381.76 \times 100\% = 8.6\%$.

It is also noticed that, compared with Case study 2 and Case study 3, Case study 4's results have less improvements. This is due to the fixed structure of the network. The integrated approach cannot reconfigure the network to further reduce the system power loss. Although there are improvements on the total operating cost, they are not so good as those in previous case studies.

3.6 Conclusion

In this chapter, the network reconfiguration together with micro-grid economic load dispatch are investigated. The stochastic nature of wind, PV and load is taken into consideration by stochastic distribution models. The forecasting of the wind, PV and load data and the energy storage system

is considered. The four bio-inspired optimization techniques are adopted to solve the problem. Optimization of benchmark-link application problems are conducted to investigate the effects of different weather and load conditions on the whole network.

This chapter has integrated micro-grid economic load dispatch and network reconfiguration together for the benefit of the whole network. The results obtained have shown that the four optimization techniques are all capable for this problem. Four techniques are all effective for doing the optimization. It is up to the users preference to choose which technique is more suitable based on the available computation resources. By using the integration approach, micro-grid can be incorporated into the network more effectively even with higher cost renewable energy resources. The network can adjust itself more efficiently to allow the utilization of the renewable energy.

The time frame of micro-grid economic load dispatch and network reconfiguration does not match at the moment. As demonstrated by the case studies, the proposed technique can also optimize the power flow and benefit the network users by reducing the operating cost even if the network structure is partially limited of completely fixed.

Chapter 4

Consensus Based Approach for Economic Dispatch Problem in a Micro-grid

4.1 Introduction

Economic dispatch (ED) problem is one of the key problems in power system operation. ED problem is commonly formulated as an centralized optimization problem in the literature. Solving ED problem is to find a power output combination of all generators which gives the lowest operating cost while maintaining system constraints.

However, with the development of DERs, research focus has been shifted to more distributed solutions on ED problem. Related existing literature [75–79] on distributed ED methods have been discussed in Chapter 1 under Literature Review section. In contrast to the existing literature, we still adopt

the equal incremental cost criterion to achieve the optimal dispatch, which means the incremental cost is chosen as the consensus variable. Compared to [76], each generator does not need to know the cost function parameters of other generators in our work. The novelty of the proposed algorithm is that it can estimate the mismatch between demand and total power in a collective sense. Each generator instantiates a local estimation of the mismatch. With a tactical initialization, the local estimated mismatch may not equal to the actual mismatch, but the summation of all the local estimated mismatch is preserved and exactly equal to the actual mismatch. In our case, no leader agent is required to collect all the power generated by each generator. The local estimated mismatch is used to adjust the power generation as if they are the true mismatch. The incremental cost is guaranteed to converge to the optimal value by the algorithm. In addition, the communication graph is assumed to be strongly connected, which is far less restrictive than the bidirectional information exchange in [75][78][109]. Furthermore, our proposed algorithm can be treated as a distributed implementation of the standard Lambda-Iteration method [110].

This chapter is organized as follows. In Section 4.2, graph theory, basic consensus results, and equal incremental cost criterion in traditional ED problem are briefly introduced. Problem description and main results for both constrained and unconstrained cases are presented in Section 4.3. A systematic learning gain design method is developed in Section 4.4. To demonstrate the effectiveness of the proposed algorithms, extensive numerical examples are shown in Section 4.5. Lastly, we conclude the chapter in Section 4.6.

4.2 Preliminary

In this Section, the basic graph terminologies, consensus algorithms, and analytic ED problem solution are introduced.

4.2.1 Graph theory

Let $\mathscr{G} = (\mathscr{V}, \mathscr{E}, \mathscr{A})$ be a weighted directed graph with the set of vertices $\mathscr{V} = \{1, 2, ..., N\}$, and the set of edges $\mathscr{E} \subseteq \mathscr{V} \times \mathscr{V}$. $\mathscr{A}$ is the adjacency matrix. Let $\mathscr{V}$ also be the index set representing the generators in the micro-grid. A directed edge from i to j is denoted by an ordered pair $(i, j) \in \mathscr{E}$, which means that generator j can receive information from generator i. The in-neighbors of the ith generator is denoted by $N_i^+ = \{j \in \mathscr{V} | (j, i) \in \mathscr{E}\}$. Similarly, the out-neighbors of the ith generator is denoted by $N_i^- = \{j \in \mathscr{V} | (i, j) \in \mathscr{E}\}$. Physically, it means a generator can obtain information from its in-neighbors, and send information to its out-neighbors. Since it is reasonable to assume that the generator i can obtain its own state information, we define that each vertex belongs to both its in-neighbor and out-neighbor, i.e., $i \in N_i^+$ as well as $i \in N_i^-$. The in-degree and out-degree of vertex i is defined as $d_i^+ = |N_i^+|$ and $d_i^- = |N_i^-|$ respectively, where $|\cdot|$ denotes the cardinality of a set. A directed graph is said to be strongly connected if there exists a path between any pair of two vertices with respect to the orientation of edges. It is easy to conclude that $d_i^+ \neq 0$ and $d_i^- \neq 0$ in a strongly connected graph.

4.2.2 Consensus algorithm

Let's define two matrices $P,\ Q \in \mathbb{R}^{N\times N}$ associated with a strongly connected graph $\mathscr{G} = (\mathscr{V}, \mathscr{E}, \mathscr{A})$ below,

$$p_{i,j} = \begin{cases} \frac{1}{d_i^+} & \text{if } j \in N_i^+ \\ 0 & \text{otherwise} \end{cases} \quad \forall i, j \in \mathscr{V},$$

similarly,

$$q_{i,j} = \begin{cases} \frac{1}{d_j^-} & \text{if } i \in N_j^- \\ 0 & \text{otherwise} \end{cases} \quad \forall i, j \in \mathscr{V}.$$

From the definition of P and Q, it is not difficult to verify that P is row stochastic, and Q is column stochastic. Note that we actually have much freedom to choose the weights of P and Q, as long as P is row stochastic and Q is column stochastic, and satisfy the following assignments, $p_{i,j} > 0$ if $j \in N_i^+$, $p_{i,j} = 0$ otherwise, $q_{i,j} > 0$ if $i \in N_j^-$, $q_{i,j} = 0$ otherwise. The convergence result of our proposed algorithm is not affected by the weights selections.

Now consider the following two separate discrete-time systems,

$$\xi_i(k+1) = \sum_{j\in N_i^+} p_{i,j}\xi_j(k), \tag{4.1}$$

$$\xi_i'(k+1) = \sum_{j\in N_i^+} q_{i,j}\xi_j'(k), \tag{4.2}$$

where $\xi_i(k)$ and $\xi_i'(k)$ are state variables associated with vertex i in graph $\mathscr{G}$ at time step k. Systems (4.1) and (4.2) have the same structure but using two different sets of weights. They can be written in the following compact

form,

$$\boldsymbol{\xi}(k+1) = P\boldsymbol{\xi}(k), \tag{4.3}$$

$$\boldsymbol{\xi}'(k+1) = Q\boldsymbol{\xi}'(k), \tag{4.4}$$

where $\boldsymbol{\xi}(k)$ and $\boldsymbol{\xi}'(k)$ are the column stack vectors of $\xi_i(k)$ and $\xi_i'(k)$. To investigate the asymptotic behavior of (4.3) and (4.4), the following theorem is needed.

Theorem 1 ([111], pp.516)**.** *If $A \in \mathbb{R}^{N\times N}$ is a nonnegative and primitive matrix, then*

$$\lim_{k\to\infty} (\rho(A)^{-1}A)^k = \boldsymbol{x}\boldsymbol{y}^T > 0$$

where $A\boldsymbol{x} = \rho(A)\boldsymbol{x}$, $\boldsymbol{y}^T A = \rho(A)\boldsymbol{y}^T$, $\boldsymbol{x} > 0$, $\boldsymbol{y} > 0$, $\boldsymbol{x}^T\boldsymbol{y} = 1$, and $\rho(A)$ denotes the spectral radius of A.

The symbol '>' denotes that all the entries in a matrix or vector are greater than zero. Based on the definition, both P and Q are nonnegative and stochastic, so $\rho(P) = \rho(Q) = 1$. Since they are derived from a strongly connected graph, and their diagonal entries are positive as well, then $P^{N-1} > 0$ and $Q^{N-1} > 0$, i.e. P and Q are primitive. From *Theorem 1*, we can derive the following two properties. Let $\mathbf{1}$ denote a vector of length N with all its elements being 1.

Property 1: $\lim_{k\to\infty} P^k = \mathbf{1}\boldsymbol{\omega}^T$ where $\boldsymbol{\omega} > 0$ and $\mathbf{1}^T\boldsymbol{\omega} = 1$.

Property 2: $\lim_{k\to\infty} Q^k = \boldsymbol{\mu}\mathbf{1}^T$ where $\boldsymbol{\mu} > 0$ and $\mathbf{1}^T\boldsymbol{\mu} = 1$.

By using *Property 1* and *2*, we can get $lim_{k\to\infty}\xi_i(k) = \boldsymbol{\omega}^T\boldsymbol{\xi}(0)$ in system (4.1), and $lim_{k\to\infty}\xi'_i(k) = \mu_i \sum_{i=1}^{N} \xi'_i(0)$, where μ_i is the ith element of $\boldsymbol{\mu}$. In system (4.1), all state variables converge to a common value, which depends on the communication topology and initial state values. The algorithm in system (4.1) is the well known consensus algorithm for the first order discrete-time system. In system (4.2), the state variables do not converge to a common value in general, but the summation of all state variable is preserved, i.e., $\sum_{i=1}^{N} \xi'_i(k) = \sum_{i=1}^{N} \xi'_i(0), \forall k$.

These interesting properties will be utilized in the proposed algorithm design in Section 4.3.

4.2.3 Analytic solution to ED problem

A micro-grid usually consists of multiple power generators. Let us assume there are N power generators. The cost function of power generation is given by the following quadratic form

$$C_i(x_i) = \frac{(x_i - \alpha_i)^2}{2\beta_i} + \gamma_i, \tag{4.5}$$

where x_i is the power generated by generator i, $\alpha_i \leq 0$, $\beta_i > 0$, and $\gamma_i \leq 0$.

The traditional ED problem is to minimize the total generation cost

$$\min \sum_{i=1}^{N} C_i(x_i), \tag{4.6}$$

subject to the following two constraints,

Generator constraint:

$$\underline{x}_i \leq x_i \leq \overline{x}_i, \tag{4.7}$$

where $\underline{x}_i$ and $\overline{x}_i$ are the lower and upper bounds of the generator capability.

Demand constraint:

$$\sum_{i=1}^{N} x_i = D, \tag{4.8}$$

where D is the total demand satisfying $\sum_{i=1}^{N} \underline{x}_i < D < \sum_{i=1}^{N} \overline{x}_i$, i.e., the problem is solvable.

The incremental cost for the generator i is $\frac{dC_i(x_i)}{dx_i} = \frac{x_i - \alpha_i}{\beta_i}$. The well known solution to traditional ED problem is the equal incremental cost criterion [110].

$$\begin{cases} \frac{x_i-\alpha_i}{\beta_i} = \lambda^* & \text{for } \underline{x}_i < x_i < \overline{x}_i \\ \frac{x_i-\alpha_i}{\beta_i} < \lambda^* & \text{for } x_i = \overline{x}_i \\ \frac{x_i-\alpha_i}{\beta_i} > \lambda^* & \text{for } x_i = \underline{x}_i \end{cases}, \tag{4.9}$$

where λ^* is the optimal incremental cost.

Note that the parameter γ_i in the cost function does not affect the incremental cost.

Remark 1. *The cost function in (4.5) is slightly different from the one in (4.10), which is commonly used by power engineers.*

$$C_i(x_i) = a_i x_i^2 + b_i x_i + c_i. \tag{4.10}$$

In fact, cost functions (4.5) and (4.10) are equivalent. It is not difficult to convert (4.5) to (4.10). Straightforward manipulation shows by setting $\alpha_i = -\frac{b_i}{2a_i}$, $\beta_i = \frac{1}{2a_i}$, *and* $\gamma_i = c_i - \frac{b_i^2}{4a_i}$, *(4.5) and (4.10) are identical. The*

main motivation of using (4.5) is for notational simplicity in the next section.

4.3 Main Results

Let the communication topology among generators be the strongly connected graph $\mathscr{G}$ described in Section 4.2 Part A. Assume there is a command vertex which distributes the total demand D to a subset of $\mathscr{V}$. Denote the command vertex by vertex 0, and its out-neighbor set N_0^-. Recall that N_0^- is the vertices set which can receive information from vertex 0. For simplicity, let the command vertex distribute the total demand equally among all the generators in N_0^-. By assumption, $1 \leq |N_0^-| \leq N$. In this Section, we firstly ignore the power generation constraints, and develop a linear distributed algorithm to solve the traditional ED problem. Later, we add in the power generation constraints.

4.3.1 Algorithm design without power generation constraints

Assume all generators have no generation constraints. According to the incremental cost criterion (4.9), when all generators operate at the optimal configuration, incremental costs are equal to the optimal value, that is

$$\frac{x_i^* - \alpha_i}{\beta_i} = \lambda^*, \forall i \in \mathscr{V}. \tag{4.11}$$

Hence, the optimal power generation for each individual generator can be calculated if the optimal incremental cost λ^* is known, i.e.,

$$x_i^* = \beta_i \lambda^* + \alpha_i, \forall i \in \mathscr{V}. \tag{4.12}$$

(4.11) and (4.12) motivate us to propose the following algorithm. Denote $\lambda_i(k)$ the estimation of optimal incremental cost by generator i, $x_i(k)$ the corresponding power generation which is an estimation of optimal power generation, $y_i(k)$ the local estimation of the mismatch between demand and total power generation.

Initializations:

$$\begin{cases} \lambda_i(0) = & \text{any fixed admissible value} \\ x_i(0) = & \text{any fixed admissible value} \\ y_i(0) = & \begin{cases} \frac{D}{|N_0^-|} - x_i(0) & \text{if } i \in N_0^- \\ -x_i(0) & \text{otherwise} \end{cases} \end{cases}, \forall i \in \mathscr{V}.$$

We are ready to state the main algorithm.

$$\lambda_i(k+1) = \sum_{j \in N_i^+} p_{i,j} \lambda_j(k) + \epsilon y_i(k) \tag{4.13a}$$

$$x_i(k+1) = \beta_i \lambda_i(k+1) + \alpha_i \tag{4.13b}$$

$$y_i(k+1) = \sum_{j \in N_i^+} q_{i,j} y_j(k) - (x_i(k+1) - x_i(k)) \tag{4.13c}$$

where ϵ is a sufficiently small positive constant.

Remark 2. *The iterative updating algorithm (4.13) only requires the local*

information. Specifically, the updating rule for generator i only requires the information received from its in-neighbor set N_i^+. Hence, this updating rule is a complete distributed algorithm.

Remark 3. *In [78], the estimated incremental cost is updated by*

$$\lambda_q(t+1) = \lambda_q(t) - \beta_t \sum_{r \in \omega_q} (\lambda_q(t) - \lambda_r(t)) - \underbrace{\alpha_t (P_{G_q}(t) - P_{L_q})}_{Innovation\ Term}, \tag{4.14}$$

where t is the time step, α_t is the control gain, $P_{G_q}(t)$ is the power generation at the bus q, and P_{L_q} is the local load at bus q.

The proposed algorithm (4.13) is quite distinct from (4.14). On one hand, the learning gain α_t is vanishing in (4.14), whereas, the learning gain ϵ in (4.13a) is fixed. If α_t is fixed, the method in [78] returns a suboptimal solution. On the other hand, the innovation term in (4.14) is calculated by the solo effort of bus q since both $P_{G_q}(t)$ and P_{L_q} are locally available. However, the feedback term $y_i(k)$ in (4.13a) is obtained by collaborative efforts of all agents in the neighborhood of generator i, see (4.13c).

Furthermore, the method is [78] is fully distributed, but a command vertex 0 is required in the algorithm (4.13). However, this requirement can be relaxed by a different initialization process that achieves fully distributed implementation, see subsection 4.3.3 for detailed discussion.

To analyze the properties and convergence of algorithm (4.13), rewrite it

in the following matrix form

$$\boldsymbol{\lambda}(k+1) = P\boldsymbol{\lambda}(k) + \epsilon\mathbf{y}(k) \tag{4.15a}$$

$$\mathbf{x}(k+1) = B\boldsymbol{\lambda}(k+1) + \boldsymbol{\alpha} \tag{4.15b}$$

$$\mathbf{y}(k+1) = Q\mathbf{y}(k) - (\mathbf{x}(k+1) - \mathbf{x}(k)) \tag{4.15c}$$

where $\mathbf{x}, \mathbf{y}, \boldsymbol{\alpha}, \boldsymbol{\lambda}$ are the column stack vector of $x_i, y_i, \alpha_i, \lambda_i$ respectively, and $B = diag([\beta_1, \beta_2, \ldots, \beta_N])$.

(4.15c) preserves the summation of $x_i(k)+y_i(k)$ over $\mathscr{V}$. It can be verified by premultiplying both sides of (4.15c) by $\mathbf{1}^T$, and noticing that Q is column stochastic, we have

$$\begin{aligned}
\mathbf{1}^T\mathbf{y}(k+1) &= \mathbf{1}^T Q\mathbf{y}(k) - \mathbf{1}^T(\mathbf{x}(k+1) - \mathbf{x}(k)), \\
&= \mathbf{1}^T\mathbf{y}(k) - \mathbf{1}^T(\mathbf{x}(k+1) - \mathbf{x}(k)), \\
\Rightarrow \mathbf{1}^T(\mathbf{y}(k+1) + \mathbf{x}(k+1)) &= \mathbf{1}^T(\mathbf{y}(k) + \mathbf{x}(k)).
\end{aligned}$$

$\mathbf{1}^T(\mathbf{y}(k)+\mathbf{x}(k))$ is a constant for all k. Notice the initialization of $x_i(0)$ and $y_i(0)$, we can obtain $\sum_{i\in\mathscr{V}} x_i(0) + y_i(0) = D$. Hence, $\mathbf{1}^T\mathbf{y}(k) = D - \mathbf{1}^T\mathbf{x}(k)$ is the actual mismatch between demand and total power generation. The mismatch is obtained via a collective effort from all individual generators rather than a centralized method. The first term in the right hand side of (4.15a) is the consensus part, it drives all $\lambda_i(k)$ to a common value. The second term $\epsilon\mathbf{y}(k)$ provides a feedback mechanism to ensure $\lambda_i(k)$ converges to the optimal value λ^*. (4.15b) just updates the estimated power generation $x_i(k)$ to the newest one.

Theorem 2. *In algorithm (4.13), if the positive constant ϵ is sufficiently small, then the algorithm is stable, and all the variables converge to the solution to the traditional ED problem, i.e.,*

$$\lambda_i(k) \to \lambda^*,\ x_i(k) \to x_i^*,\ y_i(k) \to 0, \textit{as } k \to \infty, \forall i \in \mathscr{V}.$$

Proof. We use the eigenvalue perturbation approach [112][113] to analyze the convergence properties. Replace $\mathbf{x}$ in (4.15c) with $\boldsymbol{\lambda}$ by using (4.15a)(4.15b), we have

$$\mathbf{y}(k+1) = (Q - \epsilon B)\mathbf{y}(k) + B(I - P)\boldsymbol{\lambda}(k), \tag{4.16}$$

where I is the identity matrix of appropriate dimension.

Write (4.15a) and (4.16) in matrix form, we get the following composite system,

$$\begin{bmatrix} \boldsymbol{\lambda}(k+1) \\ \mathbf{y}(k+1) \end{bmatrix} = \begin{bmatrix} P & \epsilon I \\ B(I-P) & Q - \epsilon B \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}(k) \\ \mathbf{y}(k) \end{bmatrix} \tag{4.17}$$

Define $M \triangleq \begin{bmatrix} P & 0 \\ B(I-P) & Q \end{bmatrix}$ and $E \triangleq \begin{bmatrix} 0 & I \\ 0 & -B \end{bmatrix}$.

The system matrix of (4.17) can be regarded as M perturbed by ϵE. M is a lower block triangular matrix, the eigenvalues of M is the union of the eigenvalues of P and Q. So M has two eigenvalues $\theta_1 = \theta_2 = 1$, and the rest eigenvalues lie in the open unit disk on the complex plane. Construct vectors

$\mathbf{u}_1$, $\mathbf{u}_2$ and $\mathbf{v}_1^T$, $\mathbf{v}_2^T$ as below.

$$U = [\mathbf{u}_1, \mathbf{u}_2] = \begin{bmatrix} \mathbf{0} & \mathbf{1} \\ \boldsymbol{\mu} & -\eta\boldsymbol{\mu} \end{bmatrix}, \tag{4.18}$$

where $\eta = \sum_{i=1}^{N} \beta_i$, and

$$V^T = \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \end{bmatrix} = \begin{bmatrix} \mathbf{1}^T B & \mathbf{1}^T \\ \boldsymbol{\omega}^T & \mathbf{0}^T \end{bmatrix}, \tag{4.19}$$

which are the two linearly independent right and left eigenvectors of M. Furthermore, $V^T U = I$.

When ϵ is small, the variation of θ_1 and θ_2 perturbed by ϵE can be quantified by the eigenvalues of $V^T E U$, and

$$V^T E U = \begin{bmatrix} 0 & 0 \\ \boldsymbol{\omega}^T\boldsymbol{\mu} & -\eta\boldsymbol{\omega}^T\boldsymbol{\mu} \end{bmatrix}.$$

The eigenvalues of $V^T E U$ are 0 and $-\eta\boldsymbol{\omega}^T\boldsymbol{\mu} < 0$. Thus $\frac{d\theta_1}{d\epsilon} = 0$ and $\frac{d\theta_2}{d\epsilon} = -\eta\boldsymbol{\omega}^T\boldsymbol{\mu} < 0$. That means θ_1 does not change against ϵ, and when $\epsilon > 0$, θ_2 becomes smaller. Let δ_1 be the upper bound of ϵ such that when $\epsilon < \delta_1$, $|\theta_2| < 1$. Since eigenvalues continuously depend on the entries of a matrix, in our particular case, the rest of eigenvalues of $M + \epsilon E$ continuously depend on ϵ. Therefore, there exists an upper bound δ_2 such that when $\epsilon < \delta_2$, $|\theta_j| < 1$, $j = 3, 4, \ldots, 2N$. Hence, if we choose $\epsilon < \min(\delta_1, \delta_2)$, we can guarantee that the eigenvalue $\theta_1 = 1$ is simple, and all the rest eigenvalues lie in the open unit disk.

It can be verified $\begin{bmatrix} \mathbf{1} \\ \mathbf{0} \end{bmatrix}$ is the eigenvector of system matrix in (4.17) associated with $\theta_1 = 1$. Since all the rest eigenvalues are within the open unit disk,

$$\begin{bmatrix} \boldsymbol{\lambda}(k) \\ \mathbf{y}(k) \end{bmatrix} \text{ converges to } span \begin{bmatrix} \mathbf{1} \\ \mathbf{0} \end{bmatrix}$$

as $k \to \infty$. That is $y_i(k) \to 0$. From (4.15c), we can derive $\mathbf{1}^T\mathbf{x}(k) = D$, i.e., the demand constraint is satisfied. From (4.15a), $\lambda_i(k)$ converges to a common value, i.e., the incremental cost criterion is satisfied. Therefore, we can conclude *Theorem 2*. □

4.3.2 Generalization to constrained case

In order to take account of power generation constraints, define the following projection operators.

$$\phi_i(\lambda_i) = \begin{cases} \overline{x}_i & \text{if } \lambda_i > \overline{\lambda}_i \\ \beta_i\lambda_i + \alpha_i & \text{if } \underline{\lambda}_i \leq \lambda_i \leq \overline{\lambda}_i \quad \forall i \in \mathscr{V}, \\ \underline{x}_i & \text{if } \lambda_i < \underline{\lambda}_i \end{cases}$$

where $\underline{\lambda}_i = \frac{\underline{x}_i - \alpha_i}{\beta_i}$ and $\overline{\lambda}_i = \frac{\overline{x}_i - \alpha_i}{\beta_i}$. Now the distributed algorithm becomes

$$\boldsymbol{\lambda}(k+1) = P\boldsymbol{\lambda}(k) + \epsilon\mathbf{y}(k) \tag{4.20a}$$

$$\mathbf{x}(k+1) = \phi(\boldsymbol{\lambda}(k+1)) \tag{4.20b}$$

$$\mathbf{y}(k+1) = Q\mathbf{y}(k) - (\mathbf{x}(k+1) - \mathbf{x}(k)) \tag{4.20c}$$

where $\phi(\boldsymbol{\lambda}(k+1)) = [\phi_1(\lambda_1(k+1)), \phi_2(\lambda_2(k+1)), \ldots, \phi_N(\lambda_N(k+1))]^T$.

The initial value of $\lambda_i(0)$ and $x_i(0)$ can be set to any admissible value. For simplicity, the initial value can be set as follows.

Initializations:

$$\begin{cases} \lambda_i(0) = & \underline{\lambda}_i \\ x_i(0) = & \underline{x}_i \\ y_i(0) = & \begin{cases} \frac{D}{|N_0^-|} - \underline{x}_i & \text{if } i \in N_0^- \\ -\underline{x}_i & \text{otherwise} \end{cases} \end{cases}, \forall i \in \mathscr{V}.$$

Theorem 3. *In algorithm (4.20), if the positive constant ϵ is sufficiently small, then the algorithm is stable, and all the variables converge to the solution to the traditional ED problem.*

Proof. By assumption the total demand $\sum_{i=1}^N \underline{x}_i < D < \sum_{i=1}^N \overline{x}_i$, that means at least one generator is not saturated when the demand constraint is satisfied. If the micro-grid operates in the linear region only, the rest proof follows the proof of *Theorem 2* exactly. So we only consider the saturated case here.

The nonnegative matrix P in (4.20a) tends to map $\boldsymbol{\lambda}(k)$ to the *span*$\{\mathbf{1}\}$. Premultiply $\mathbf{1}^T$ from both sides of (4.20a), we have

$$\sum_i \lambda_i(k+1) = \sum_{i,j} p_{i,j}\lambda_j(k) + \epsilon e(k), \tag{4.21}$$

where $e(k) = D - \sum_i x_i(k)$ is the mismatch between demand and total power generation, and ϵ can be treated as a proportional feedback gain. Without loss of generality, assume $e(k) > 0$, the overall level $\lambda_i(k)$ will increase and it approaches to the same value, and notice that the total power generation

is a monotonically increasing function of incremental cost. Thus, the total power generation will increase. Therefore, the feedback mechanism in (4.21) will reduce the mismatch $e(k)$. In this process, some of the generator may reach its maximum capability. After some sufficiently long time K, if $x_i(K)$ is saturated, then $x_i(k)$ is always saturated for $k > K$. To investigate the transient behavior for $k > K$, algorithm (4.20) can be written in the following composite system.

$$\begin{bmatrix} \boldsymbol{\lambda}(k+1) \\ \mathbf{y}(k+1) \end{bmatrix} = \begin{bmatrix} P & \epsilon I \\ \tilde{B}(I-P) & Q - \epsilon\tilde{B} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}(k) \\ \mathbf{y}(k) \end{bmatrix}, \tag{4.22}$$

where $\tilde{B} = diag([\tilde{\beta}_1, \tilde{\beta}_2, \ldots, \tilde{\beta}_N])$, and

$$\tilde{\beta}_i = \begin{cases} 0 & \text{if } x_i(k) \text{ is saturated,} \\ \beta_i & \text{otherwise.} \end{cases}$$

Based on our assumption, there is at least one $\tilde{\beta}_i$ is nonzero. Follow the similar eigenvalue perturbation analysis, when ϵ is sufficiently small, the above system is stable. In addition, $\boldsymbol{\lambda}(k) \to span\{\mathbf{1}\}$, $\mathbf{y}(k) \to \mathbf{0}$, i.e., solves the traditional ED problem. □

4.3.3 Fully distributed implementation

In the previous two subsections, a command vertex 0 is designed to distribute the total demand into the network, which requires the global information D. Notice that the role of vertex 0 is only activated at the initialization stage, in this subsection a modification of the initialization process is developed to

achieve the fully distributed implementation, and the vertex 0 is no longer required, i.e., the total demand D is not needed.

Let D_i be the local demand associated with generator i at the bus i. Initialize the algorithm with the following method.

Initializations:

$$\begin{cases} x_i(0) = \begin{cases} \overline{x}_i & \text{if } \overline{x}_i < D_i \\ D_i & \text{if } \underline{x}_i \leq D_i \leq \overline{x}_i \\ \underline{x}_i & \text{if } D_i < \underline{x}_i \end{cases} & ,\forall i \in \mathscr{V}. \\ \lambda_i(0) = \frac{x_i(0)-\alpha_i}{\beta_i} \\ y_i(0) = D_i - x_i(0) \end{cases}$$

If the bus i contains load only, then $\overline{x}_i = \underline{x}_i = 0$. The above initialization is fully distributed since no global demand is required in the calculation. Summing up $y_i(0)$ over i, we have $\sum_{i\in\mathscr{V}} y_i(0) = \sum_{i\in\mathscr{V}}(D_i - x_i(0)) = D - \sum_{i\in\mathscr{V}} x_i(0)$, which is the actual mismatch between demand and total power generation. In addition, the equality $\sum_{i\in\mathscr{V}} y_i(k) = D - \sum_{i\in\mathscr{V}} x_i(k)$ is preserved over time k by (4.20c). Hence, with the new initialization, the result in *Theorem 3* still holds under the algorithm (4.20).

4.4 Learning gain design

Theorems 2 and *3* provide sufficient conditions for the proposed algorithms to work. However, they do not offer constructive methods to design the learning ϵ. Therefore, we present a systematic design method to select the appropriate learning gain in this section.

Note that the learning gain for all generators are the same in (4.13a). Indeed, the learning gains are not restricted to be identical. All the results in Section 4.3 are still valid when the learning gains are different from each other. Let the generators have distinct learning gains ϵ_i, then (4.13a) becomes

$$\lambda_i(k+1) = \sum_{j \in N_i^+} p_{i,j}\lambda_j(k) + \epsilon_i y_i(k).$$

Follow the similar procedure, we can obtain the composite system below

$$\begin{bmatrix} \boldsymbol{\lambda}(k+1) \\ \mathbf{y}(k+1) \end{bmatrix} = \begin{bmatrix} P & \Omega \\ B(I-P) & Q-\Omega B \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}(k) \\ \mathbf{y}(k) \end{bmatrix}, \tag{4.23}$$

where $\Omega = diag([\epsilon_1, \epsilon_2, \cdots, \epsilon_N])$. Denote the system matrix of (4.23) by H, that is

$$H = \begin{bmatrix} P & \Omega \\ B(I-P) & Q-\Omega B \end{bmatrix}.$$

Based on the our previous analysis, H has an eigenvalue of 1 associated with eigenvector $\boldsymbol{\pi}_1 = \begin{bmatrix} \mathbf{1} \\ \mathbf{0} \end{bmatrix}$. If the moduli of the rest eigenvalues are less than 1, system (4.23) is stable, and solves the ED problem. Thus, our ultimate task now is to find out a suitable Ω such that H is stable.

By using Gram-Schmidt orthonormalization process [111, pp.15], together with $\boldsymbol{\pi}_1$, we can generate a set of $2N-1$ orthonormal bases $\{\boldsymbol{\pi}_2, \boldsymbol{\pi}_3, \cdots, \boldsymbol{\pi}_{2N}\}$, where $\boldsymbol{\pi}_j \in \mathbb{R}^{2N}, j = 2, 3, \cdots, 2N$. Then, we can construct a $2N \times (2N-1)$

projection matrix

$$\Pi = \begin{bmatrix} \Pi_1 \\ \Pi_2 \end{bmatrix} = [\boldsymbol{\pi}_2, \boldsymbol{\pi}_3, \cdots, \boldsymbol{\pi}_{2N}],$$

where $\Pi_1, \Pi_2 \in \mathbb{R}^{N\times(2N-1)}$. Applying the following projection, we have

$$\tilde{H} = \Pi^T H \Pi.$$

Hence, $\tilde{H}$ has the same set of eigenvalues as H except the eigenvalue 1. Therefore, our task is to find an Ω such that $\rho(\tilde{H}) < 1$. This kind of problem is rather difficult to solve by analytic method. From matrix analysis [111], it is well known that

$$\rho(\tilde{H}) \leq \inf_S \|S\tilde{H}S^{-1}\|_2,$$

where S is a non-singular matrix. Hence, if we could find out an Ω so that $\inf_S \|S\tilde{H}S^{-1}\|_2 < 1$, then it is guaranteed that $\rho(\tilde{H}) < 1$. Inspired by D-K iteration method in robust control literature [114], we develop the following algorithm that determines an Ω that minimizes $\inf_S \|S\tilde{H}S^{-1}\|_2$, which is the greatest lower bound of $\rho(\tilde{H})$.

Step 1: *Initialize* Ω*, e.g.,* $\Omega = 0$*;*

Step 2: *Find* S *that minimizes* $\|S\tilde{H}S^{-1}\|_2$ *by given* Ω*;*

Step 3: *Find new* Ω *that minimizes* $\|S\tilde{H}S^{-1}\|_2$*, where* S *is obtained from* Step 2*. Then, goto* Step 2*.*

The above algorithm is essentially an iterative numerical method which

mimics the D-K iteration method. As remarked in [114], up to now, it is not possible to prove the convergence of D-K iteration method. However, in practical applications, it takes only a few iterations to find a solution that is nearly optimal.

The minimization problem in Step 2 is equivalent to the generalized eigenvalue problem below in the field of linear matrix inequality (LMI),

$$\begin{aligned} \underset{S}{\text{minimize}} \quad & t \\ \text{subject to} \quad & 0 \prec S^T S, \\ & \tilde{H}^T S^T S \tilde{H} \prec t S^T S. \end{aligned}$$

The problem in Step 3 can be converted to the following LMI problem by using Schur Complement,

$$\begin{aligned} \underset{\Omega}{\text{minimize}} \quad & r \\ \text{subject to} \quad & \begin{bmatrix} -rS^T S & \star \\ S(H_l + \Pi_1^T \Omega \Pi_2 - \Pi_2^T \Omega B \Pi_2) & -I \end{bmatrix} \prec 0, \end{aligned}$$

where $H_l = \Pi_1^T P \Pi_1 + \Pi_2^T B(I-P)\Pi_1 + \Pi_2^T Q \Pi_2$, and $\star$ represents the corresponding symmetric component in the matrix.

These two LMI problems can be effectively handled by numerical software packages, for instance, MATLAB LMI control toolbox.

Remark 4. *To design an appropriate learning gain Ω that ensures convergence, the proposed design method requires the detailed information of communication P, Q, as well as the generator parameter B. P, Q, B are fixed*

values, and they can be obtained by offline methods. Once the learning gain is calculated, P, Q, B are no longer needed in the implementation. Whereas, a generator only requires the power generation information from its neighbors at the implementation stage. The time-varying global information, such as total power generation, is not required.

4.5 Application Examples

This section first presents two case studies to verify the proposed scheme, namely, the constrained and unconstrained cases. Next, we test the robustness on the variations on command vertex connections, as well as the scalability when new generator unit is added into the micro-grid. Then, the algorithm is slightly modified to accommodate time-varying demand. Next, the relation between learning gain and convergence speed is investigated. Then, we compare the proposed algorithm with Lambda-Iteration method. The fully distributed implementation is applied to the IEEE 14-bus test systems. Lastly, we compare our technique with other distributed optimization technique.

In our case studies the case study 8 and case study 9, we adopt the generator examples from the classic book by Wood and Wollenberg [110, pp.31-32]. Three types of generators are available, namely, Type A (Coal-fired steam unit), Type B (Oil-fired steam unit), and Type C (Oil-fired steam unit). The cost function parameters and generation capabilities are given in Table. 4.1. The cost function parameters are specified in terms of $a - b - c$, which are commonly used by power engineers. Then, they are converted to

$\alpha - \beta - \gamma$ used by our algorithm.

The communication topology is shown in Fig. 4.1. We assume that there are four generators in this micro-grid, labeled as vertices 1, 2, 3, and 4. The four generators are selected from the three types of generators in Table. 4.1. Vertices 1 and 2 are Type A generators, vertex 3 is a Type B generator, and vertex 4 is a Type C generator. The communication among the four generators are denoted by solid lines, which is a strongly connected graph. Vertex 0 is the command vertex. The dashed line represents the communication between command vertex and power generators, that is the command vertex can send information to power generators 1 and 3. Based on Fig. 4.1, matrices P and Q can be defined as

$$P = \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}, \; Q = \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{1}{2} & 0 & \frac{1}{3} \end{bmatrix}$$

respectively.

The initial values in case studies $1 - 4$ are given in Table. 4.2, and total demand $D = 1500MW$. For simplicity, the learning gains for all generators are set to be identical. We apply the design method in Section 4.4, and the iterative method converges in 5 iterations by using MATLAB with initial guess $\epsilon = 0$. The calculated learning gain $\epsilon = 7.026e^{-4}$. Hence, we adopt the obtained learning gain throughout the case studies unless otherwise specified.

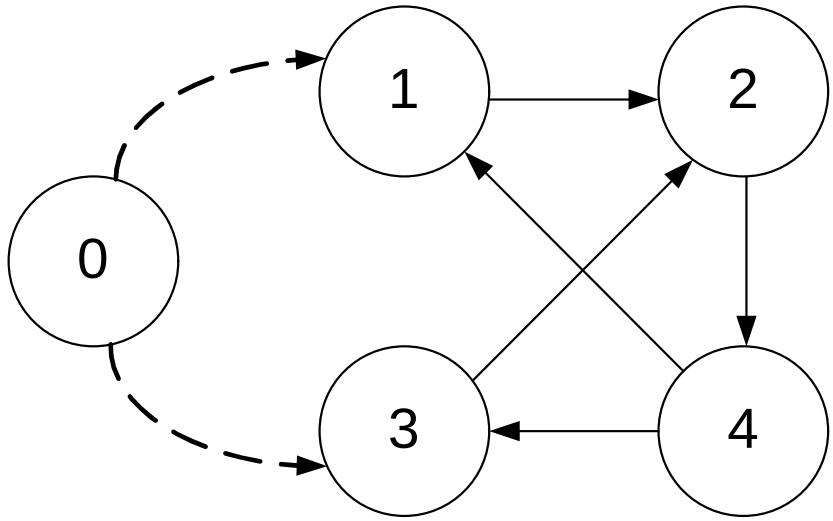


Figure 4.1: Communication topology among generators and command vertex in the network.

Table 4.1: Generator parameters

Generator Type	A (Coal-fired)	B (Oil-fired)	C (Oil-fired)
Range (MW)	[150, 600]	[100, 400]	[50, 200]
a ($\$/MW^2h$)	0.00142	0.00194	0.00482
b ($\$/MWh$)	7.2	7.85	7.97
c ($\$/h$)	510	310	78
α (MW)	-2535.2	-2023.2	-826.8
β ($MW^2h/\$$)	352.1	257.7	103.7
γ ($\$/h$)	-8616.8	-7631.0	-3216.7

4.5.1 Case study 1: without generator constraints

In this case study, the generators' constraints are not imposed. The collective estimated mismatch y_i, generators output x_i, incremental cost λ_i and total power generated are shown in Fig. 4.2. y_i goes to zero after 20 iterations. This means that the mismatch between demand and total power generated goes to zero. This result can be further verified by power balance subplot. More importantly, the incremental costs λ_i of all generators converge to a

Table 4.2: Initializations

Variables	$i = 1$	$i = 2$	$i = 3$	$i = 4$
$x(0)$ (MW)	150	150	100	50
$y(0)$ (MW)	600	-150	650	-50
$\lambda(0)$ $(\$/MWh)$	7.63	7.63	8.24	8.42

common value. Hence, the optimization goal is fulfilled. Based on the results, $\lambda^* = 8.84\$/MWh$, $x_1^* = 577.35MW$, $x_2^* = 577.35MW$, $x_3^* = 255.07MW$, and $x_4^* = 90.22MW$. All the final outputs are within the generators' operational ranges. Careful examination on the plots shows that the fourth generator goes below $50MW$ during the transient response since the generation constraints are not imposed. This is not desirable. We will investigate the results with generator constraints in the next case study.

4.5.2 Case study 2: with generator constraints

In this case study, the generators' constraints are imposed to illustrate a more practical scenario. The results are shown in Fig. 4.3. The collective estimated mismatch y_i goes to zero after 20 iterations. Finally, the estimated incremental costs of all generators converge to the same value while meeting the power balance constraint. From the transient response, we notice that generator 4 gets saturated in the first 4 iterations, and it gradually increases to the final output as the incremental costs are increasing. Based on the results, $\lambda^* = 8.84\$/MWh$, $x_1^* = 577.35MW$, $x_2^* = 577.35MW$, $x_3^* = 255.07MW$, and $x_4^* = 90.22MW$. The results are the same as those obtained from case study 1. All the power generations are within the generation ranges. No power generators exceed the operation ranges even in the transient responses.

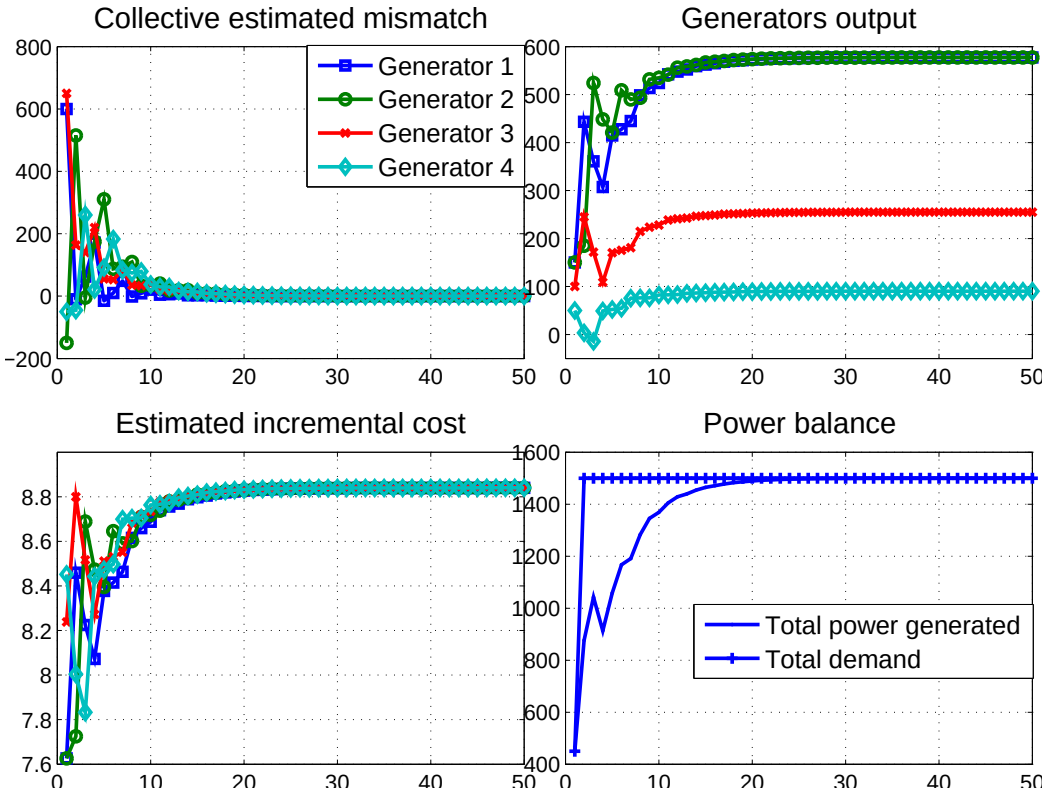


Figure 4.2: Results obtained without generator constraints.

4.5.3 Case study 3: robustness of command node connections

Our proposed algorithm only requires that the communication between generators is strongly connected. If the command vertex has at least one edge to any one of the generators, the algorithm works when the learning gain is appropriately chosen. This is a very flexible communication condition. In the previous two case studies, the command vertex is connected to generators 1 and 3. Now, change the command vertex connection to generators 2 and 4. The purpose of this case study is to demonstrate that the proposed algorithm works regardless of the command node connections. Fig. 4.4 shows the nu-

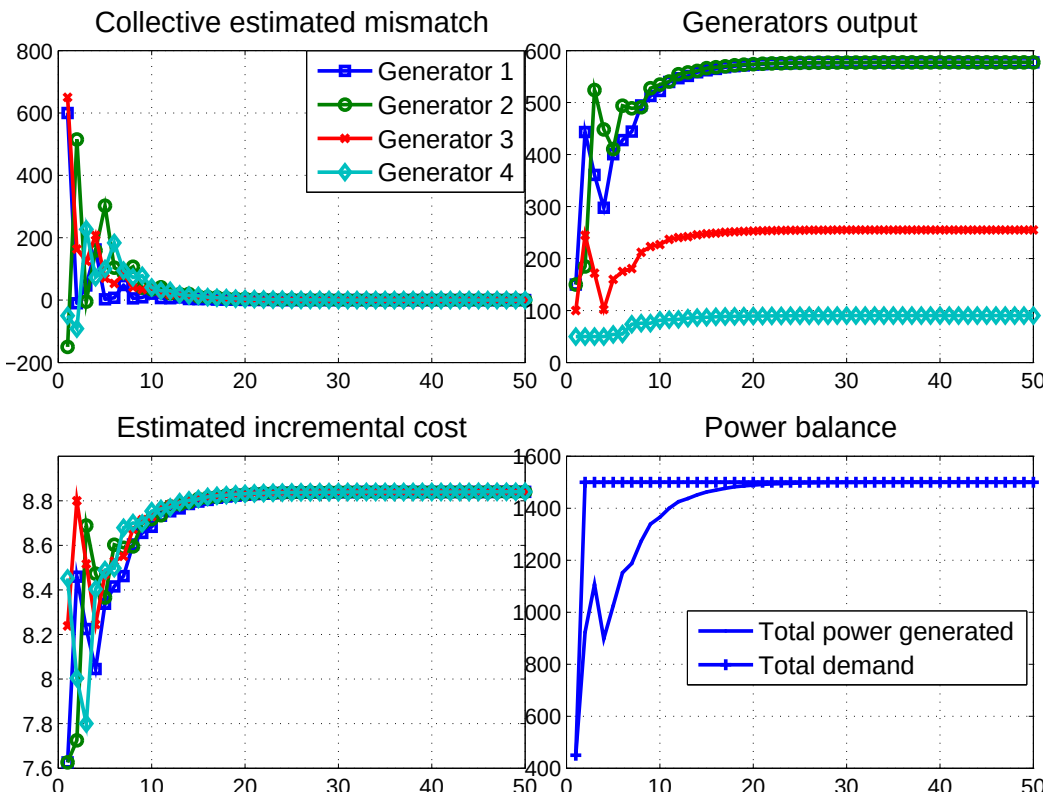


Figure 4.3: Results obtained with generator constraints.

merical results. The mismatch information is collectively minimized by all the generators after 20 iterations. The incremental costs also reach consensus after 20 iterations. As shown in the fourth sub-figure, the total demand and total power are balanced. The initial conditions for $x_i(0)$ are still the same as in Table. 4.2. The initial conditions for $y_i(0)$ are changed to $y_1(0) = -150MW$, $y_2(0) = 600MW$, $y_3(0) = -100MW$, and $y_4(0) = 700MW$, and the final outcomes are $\lambda^* = 8.84\$/MWh$, $x_1^* = 577.35MW$, $x_2^* = 577.35MW$, $x_3^* = 255.07MW$, and $x_4^* = 90.22MW$. The results are identical to Case study 2, which indicates the command vertex connection does not affect the final convergence results.

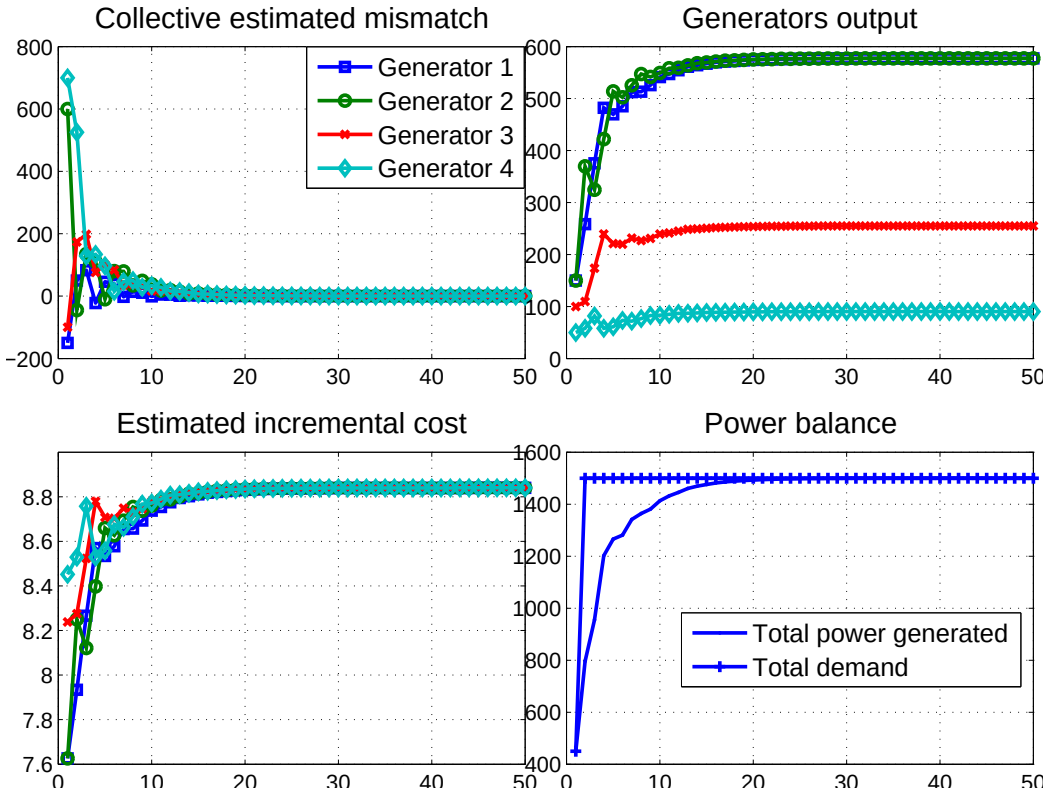


Figure 4.4: Robustness test when the command vertex is connected to generators 2 and 4.

4.5.4 Case study 4: plug and play test

One of the most important features of a micro-grid is its plug and play adaptability. In this case study, the four generators have already reached the optimal states before plugging in the fifth generator. The fifth generator is plugged in at time step $k = 50$. The fifth generator is a Type B generator. The initializations of the generators are the same as in Case 2. The new communication topology is shown in Fig. 4.5. Thus matrices P and Q are

changed to

$$P = \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}, \; Q = \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{3} & 0 \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{2} \\ 0 & \frac{1}{3} & 0 & 0 & \frac{1}{2} \end{bmatrix}$$

respectively. After plugging in the fifth generator at $k = 50$, the output of generator 5 is set to $x_5(50) = 100MW$, $y_5(50) = -100MW$, and $\lambda_5(50) = 8.24\$/MWh$. From the results in Fig. 4.6, we can observe that the local estimated mismatch y_i goes to zero after a short disturbance. The other three generators reduce their outputs in order to accommodate the fifth generator output. Therefore, the incremental cost drops due to lower average output. Finally, the estimated incremental costs of all generators converge to the same value while meeting the power balance constraint. Based on the results, $\lambda^* = 8.647\$/MWh$, $x_1^* = 509.49MW$, $x_2^* = 509.49MW$, $x_3^* = 205.40MW$, and $x_4^* = 70.22MW$, and $x_5^* = 205.40MW$. All the power outputs are within the generation range. The optimization goal is fulfilled and the fifth generator is well adapted into the system.

4.5.5 Case study 5: time-varying demand

In this case study, the generators' setup is the same as in Case 2. In a practical situation, it is very likely that the demand is not a constant over time. Slight modification of our proposed algorithm can handle the demand

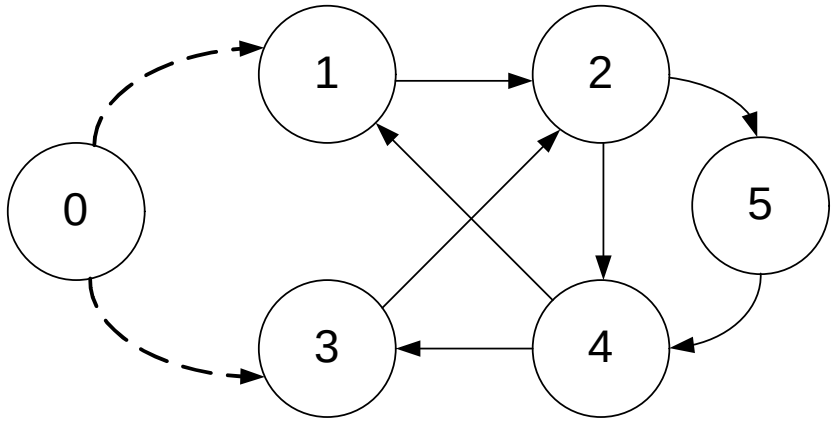


Figure 4.5: Communication topology with the fifth generator.

change effectively. Let the initial demand be $D = 1500MW$ as usual, we purposely change it to $D' = 1000MW$ at time step $k = 50$. The demand change is only known to generators 1 and 3 since they are connected to the command vertex. Thus, the algorithm needs to modify the local estimated mismatch at $k = 50$ before continuing updating the variables. Keep y_2, y_4 unchanged at $k = 50$, and update y_1, y_3 as follows at $k = 50$ before proceeding to the next updating iteration.

$$\begin{aligned} y_1(50) &= y_1(50) + \frac{D' - D}{m}, \\ y_3(50) &= y_3(50) + \frac{D' - D}{m}. \end{aligned}$$

Fig. 4.7 shows the results. After the demand changes from $1500MW$ to $1000MW$ at $k = 50$, the algorithm responds to the change of demand quickly. The mismatch information is obtained at $k = 50$, and the generators minimize the mismatch information collectively. The new incremental cost consensus is reached at $k = 65$. The algorithm asymptomatically converges

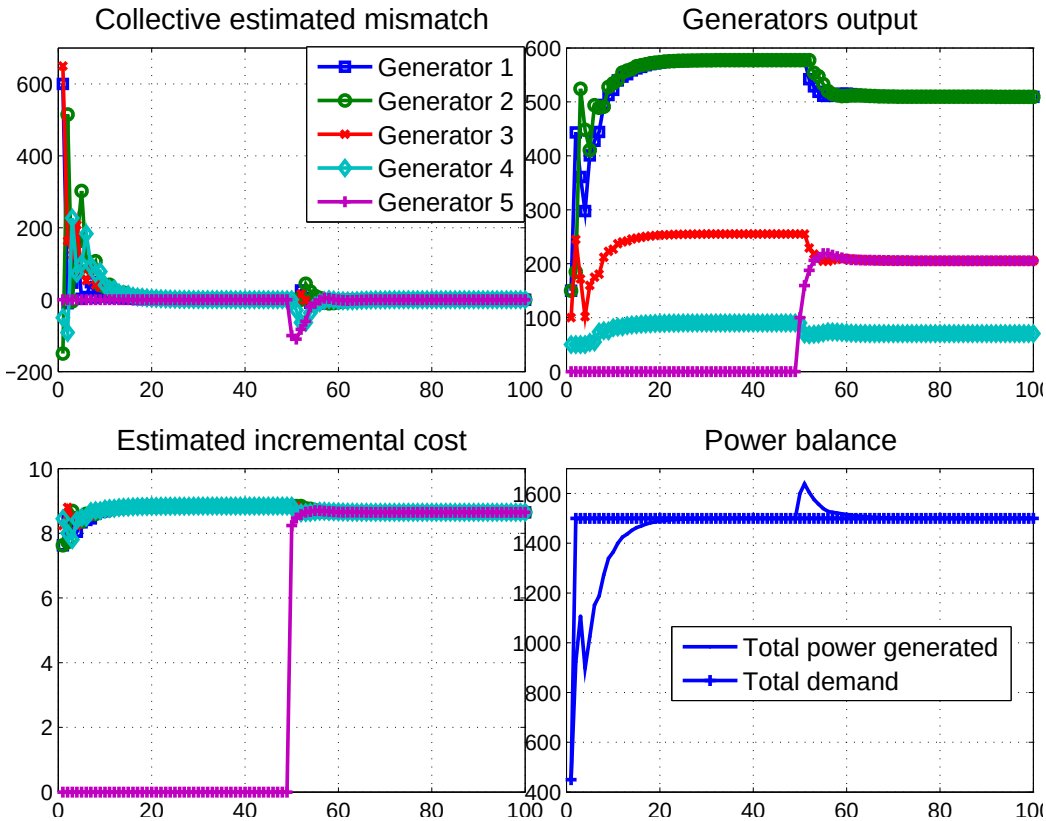


Figure 4.6: Results obtained with the fifth generator.

to the new optimal solution at iteration $k = 65$, i.e., $\lambda^* = 8.3617\$/MWh$, $x_1^* = 409.06MW$, $x_2^* = 409.06MW$, $x_3^* = 131.89MW$, and $x_4^* = 50.00MW$. Compared to Case 2 results, all the outputs decrease since the demand is reduced. The new power balance is also reached at $k = 65$.

4.5.6 Case study 6: relation between convergence speed and learning gain

The learning gain is the only design parameter that we can manipulate, and it plays a significant role in convergence speed and also the performance. If the learning gain is not properly selected, the results may oscillate, or even

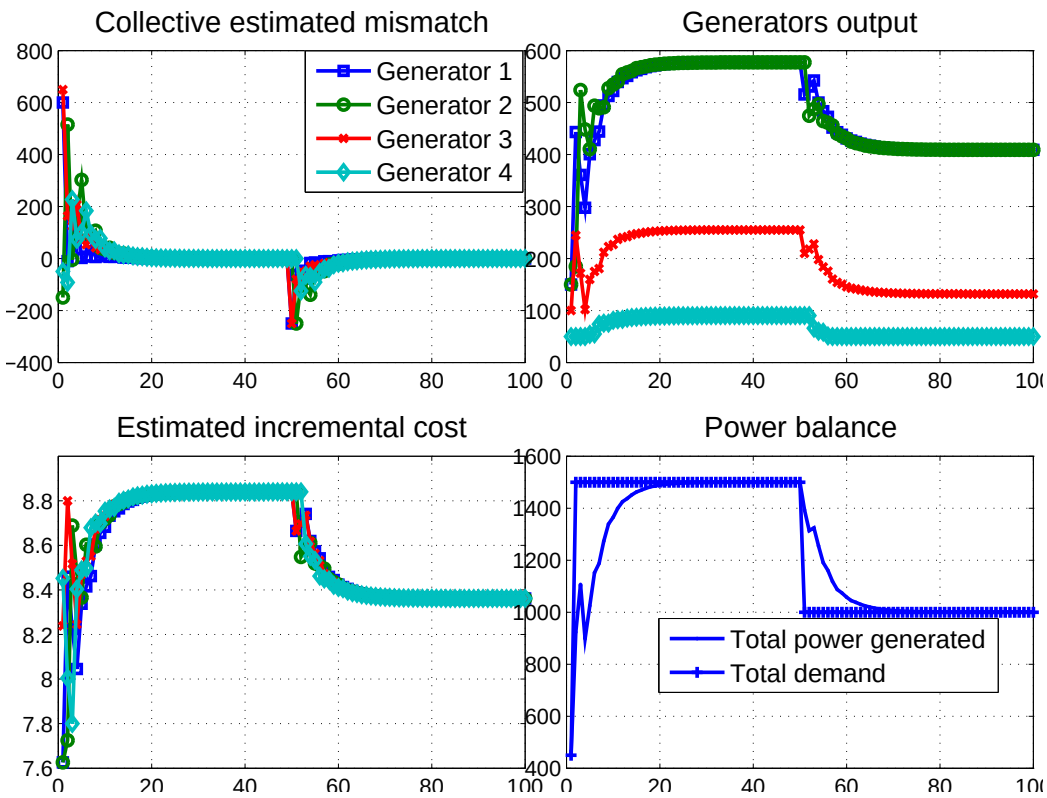


Figure 4.7: Results obtained with time-varying demand.

diverge. Thus, it is important to investigate the relation between convergence properties and learning gain. We select four sets of learning gains, i.e., $\epsilon = 0.5e^{-3}, 1.0e^{-3}, 1.5e^{-3}$, and $2.0e^{-3}$, and study the performances of the learning gains. Let the initial conditions be the ones specified in Table. 4.2. The mismatch between demand and total power generation is depicted in Fig. 4.8. In general, when the learning gain is small, the convergence speed is relatively slow, and the transient response is smooth. A large gain learning results in a faster convergence speed. However, the performance may become oscillatory. In addition, if the learning goes beyond certain limit, the algorithm diverges as in the experiment with $\epsilon = 2.0e^{-3}$.

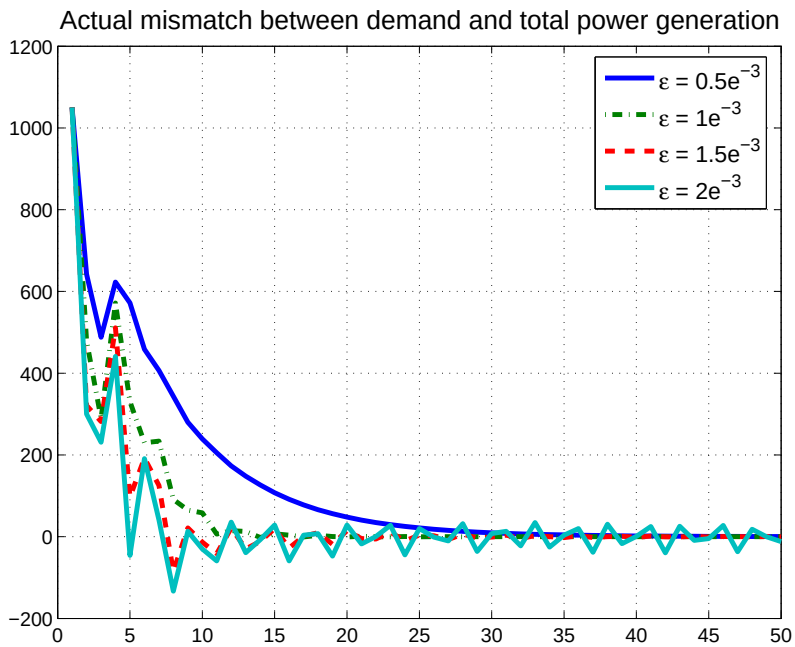


Figure 4.8: Relation between convergence speed and learning gain.

4.5.7 Case study 7: comparison with Lambda-Iteration method

In the Lambda-Iteration method [110], there is an independent system operator (ISO) that broadcasts the current estimate of the optimal increment cost λ to all the generators, and then collects the power generations from all power generators at the current cost λ. Next, the ISO calculates the mismatch between demand and total power generation. Based on the mismatch, the ISO broadcast a new estimate of the optimal increment cost to all generators. Such a mechanism repeats until the optimal incremental cost is obtained.

There are three differences between Lambda-Iteration method and our

proposed method. (1) The role of ISO and command vertex are distinct. In Lambda-Iteration method, the ISO is responsible for all calculations. Whereas, the command vertex is only responsible for setting the demand references for a subset of generators, and all computations are done by individual generators at local level. (2) Information requirements are different. The total power generated by all generators and the estimate of λ are both global information in the Lambda-Iteration method. The ISO needs to collect the power generation from each generator, and also broadcast λ to each generator. However, in our method, generator parameters and communication graph are only required in the design phase. Once the learning gain is obtained, they are no longer needed. No global information is needed in the implementation phase, specifically, both the optimal incremental cost, and mismatch between demand and total power generation are estimated by each generator through local interactions. (3) Communication topologies are different. In Lambda-Iteration method, the graph should be a star, i.e. bidirectional communication between ISO and all individual generators. No communication is involved among generators. This would increase the communication burden for ISO. In contrast, any strongly connected graph is sufficient in our method. The command vertex only connects to a subset of the generators. The communication burden is spread among all generators. This is one of the advantages of the distributed algorithms.

To see how Lambda-Iteration method performs, let the ISO update λ by

the following feedback rule,

$$\lambda(k+1) = \lambda(k) + \theta(D - \sum_{i=1}^{N} x_i(k)),$$

where θ is a positive learning gain, and $x_i(k)$ is the power generation by the ith generator. Set the problem formulation the same to Case 2, with the initial guess $\lambda(0) = 5$, learning gain $\theta = 7e^{-4}$, total demand $D = 1500MW$. Fig. 4.9 shows the performance of Lambda-Iteration method. The algorithm converges very fast since global information is available to the ISO. In contrast, our results in Case 2 are identical to the results in Fig. 4.9. It shows that our distributed method can obtain exactly the same results even without global information.

Remark 5. *In the proposed algorithms, the learning gain should be sufficiently small in order to stabilize the whole system. The upper bound of ϵ depends on both the communication topology and the cost function parameters. In our proposed design method, the learning gain is calculated by solving a LMI problem, which requires complete information of communication and generator parameters. To achieve better robustness and scalability of the distributed solution to the traditional ED problem, the choice of learning gain should be independent of the communication topology and generator parameters. This is our ongoing research work.*

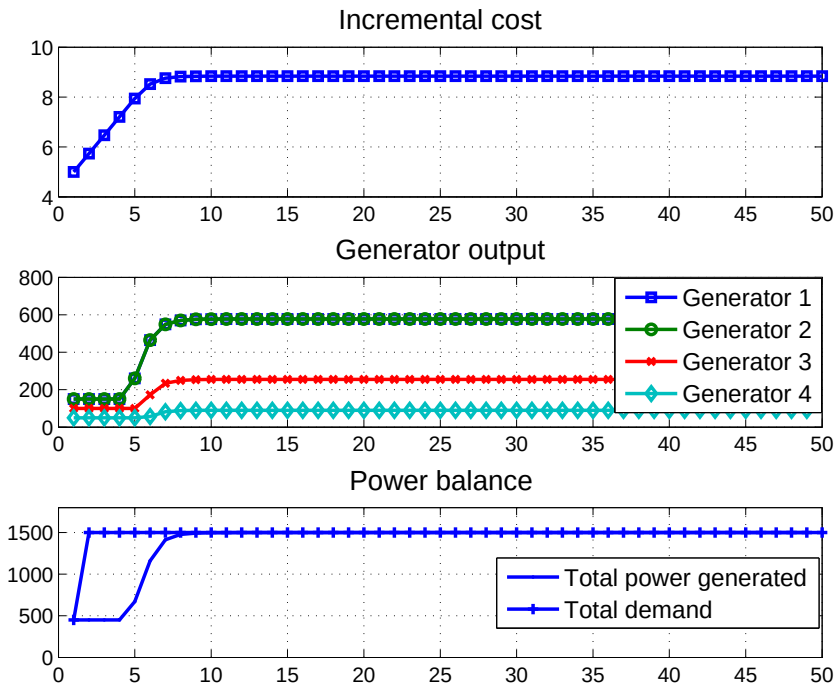


Figure 4.9: Results obtained by Lambda-Iteration method.

4.5.8 Case study 8: fully Distributed Implementation with IEEE 14-Bus Test Systems

This case study demonstrates the fully distributed implementation of the proposed methods to IEEE 14-bus test systems. This test system represents a portion of the American Electric Power System which is located in the Midwestern US as of February in 1962. This system has 14 buses, 5 generators and 11 loads [115]. The generator parameters are adopted from the examples in [78], which are restated in Table. 4.3.

Notice that Buses 1, 2, 3, 6, and 8 contain generators. When a bus contains load only, the power generation at that bus is set to zero. The

Table 4.3: IEEE 14-bus test systems generator parameters

Bus	a ($\$/MW^2h$)	b ($\$/MWh$)	c ($\$/h$)	Range (MW)
1	0.04	2.0	0	[0, 80]
2	0.03	3.0	0	[0, 90]
3	0.035	4.0	0	[0, 70]
6	0.03	4.0	0	[0, 70]
8	0.04	2.5	0	[0, 80]

initial local demands at each buses are given as $D_1 = 0MW$, $D_2 = 21.7MW$, $D_3 = 66.2MW$, $D_4 = 47.8MW$, $D_5 = 7.6MW$, $D_6 = 11.2MW$, $D_7 = 0MW$, $D_8 = 0MW$, $D_9 = 29.5MW$, $D_{10} = 9.0MW$, $D_{11} = 3.5MW$, $D_{12} = 6.1MW$, $D_{13} = 13.5MW$, and $D_{14} = 14.9MW$. It is easy to calculate that the total demand is $D = 231MW$, and it is not required by the algorithm.

The communication among buses can be independent from the actual bus connections. Consider each bus as a vertex, and a vertex only sends information to the next two vertices, i.e., the edge set is $\mathscr{E} = \{(i, i+1), (i, i+2) | 1 \leq i \leq 12\} \bigcup \{(13, 14), (13, 1), (14, 1), (14, 2)\}$.

Now increase the demand by 10%, and the new total demand is $D' = D * (1 + 10\%) = 254.1MW$. Initialize $\mathbf{x}(0)$ and $\mathbf{z}(0)$ with the previous optimal dispatch (historical data). The estimated mismatch is initialized as below

$$y_i(0) = D_i * 10\%, \forall i \in \mathscr{V},$$

which is the fully distributed implementation.

By using the proposed algorithm (4.20) with learning gain $\epsilon = 5e^{-3}$, the numerical results are presented in Fig. 4.10. The generators outputs are confined within the operation ranges. The collective estimated mismatches converge to zero. Total power generated by the micro-grid converges to the

target demand, and all the incremental costs converge to the same value, i.e., the ED problem is solved. In this particular case, the optimal incremental cost is $\lambda^* = 6.6665\$/MWh$, generator outputs are $x_1^* = 58.33MW$, $x_2^* = 61.11MW$, $x_3^* = 38.09MW$, $x_6^* = 44.44MW$, and $x_8^* = 52.08MW$.

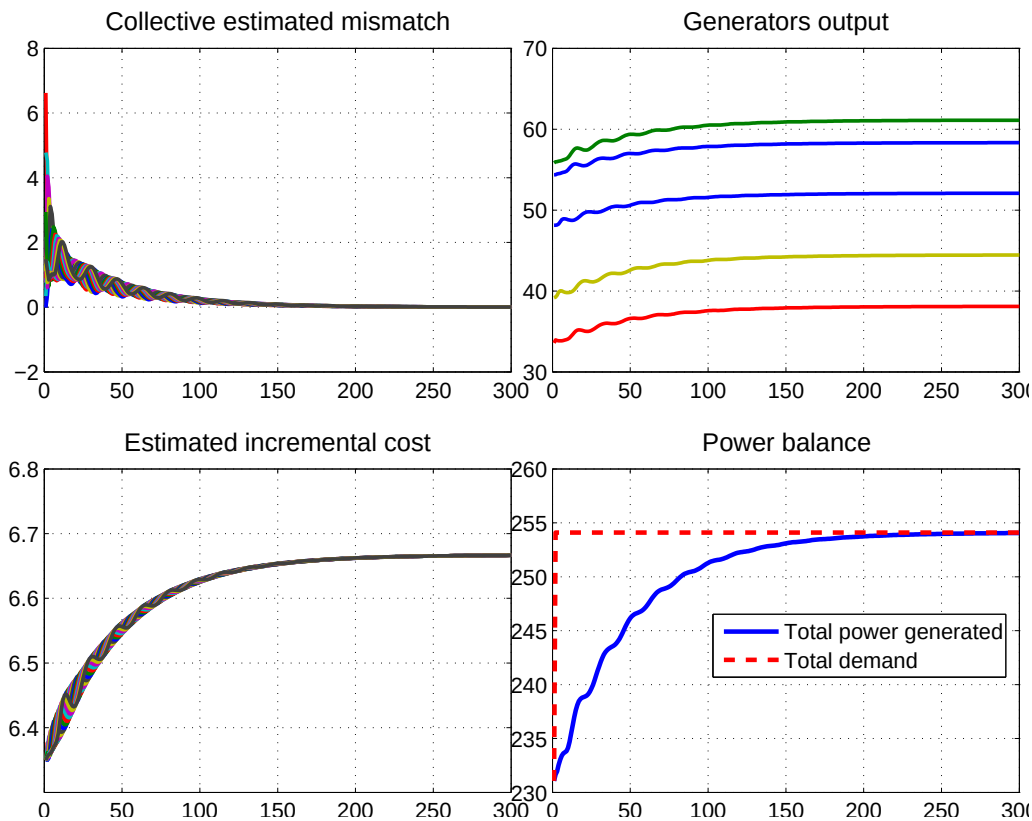


Figure 4.10: Results obtained with IEEE 14-bus test systems.

4.5.9 Case study 9: comparison with Kar's Work

Based on the (25) in Kar's work [78], the incremental cost is updated as below

$$\lambda_q(t+1) = \underbrace{\lambda_q(t) - \beta_t \sum_{r \in \omega_q} (\lambda_q(t) - \lambda_r(t))}_{\text{Consensus Part}} - \underbrace{\alpha_t (P_{G_q}(t) - P_{L_q})}_{\text{Innovation Term}}, \tag{4.24}$$

where α_t is the **vanishing gain**, $P_{G_q}(t)$ is the power generation at the bus q, and P_{L_q} is the local load at bus q. The innovation term is calculated by the **solo effort** of bus q since both $P_{G_q}(t)$ and P_{L_q} are locally available.

In our work, the incremental cost is updated by the following equation,

$$\lambda_i(k+1) = \underbrace{\sum_{j \in N_i^+} p_{i,j} \lambda_j(k)}_{\text{Consensus Part}} + \underbrace{\epsilon y_i(k)}_{\text{Feedback Term}}, \tag{4.25}$$

where ϵ is a **fixed learning gain**, and $y_i(k)$ is the estimate of mismatch between total demand and total power generation in the whole network by generator i. Notice that $y_i(k)$ cannot be calculated by the solo effort of generator i, instead it is calculated from

$$y_i(k+1) = \sum_{j \in N_i^+} q_{i,j} y_j(k) - (x_i(k+1) - x_i(k)), \tag{4.26}$$

which is the collaborative effort of all agents in the neighborhood of generator i.

Furthermore, the innovation term in (4.24) is obtained by algebraic ma-

nipulation, whereas, the feedback term in (4.25) is obtained by the dynamic (4.26).

A fully distributed implementation of our method is detailed in subsection 4.5.8, and the total demand is not required at all. While control performance is still preserved.

The parameters setup in 4.5.8 is very close to Kar's paper. The maximum degree of a vertex is 3 in Kar's work, i.e., a vertex sends information to at most 3 vertices. To be comparable to Kar's work, we design the communication graph in a way such that a vertex sends information to 2 vertices. Therefore, the actual number of edges is $14 * 2 = 28$, and maximum possible number of edges is $14 * 14 = 196$, thus, the density of the graph is $28/196 = 0.143$, which is quite sparse.

Based on the convergence criteria, $\|\lambda(k) - \lambda^* \mathbf{1}\| \leq 0.03$, the convergence is achieved in 300 steps in Kar's work as shown in Fig 4.11. Whereas, it is achieved in 142 steps by our approach, more than 50% improvement.

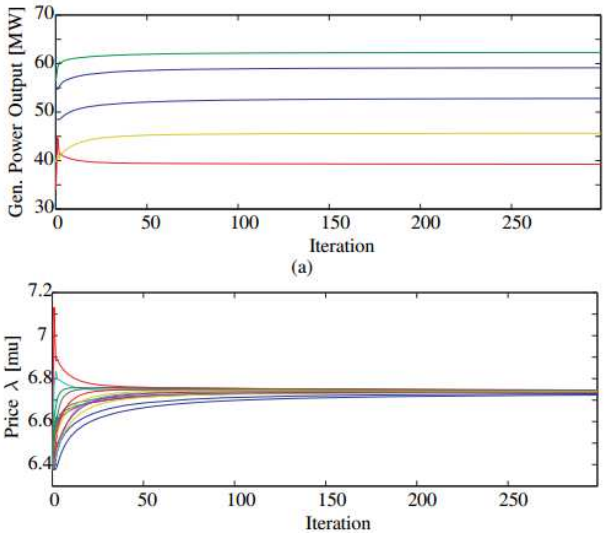


Figure 4.11: (a) Power output from generators and (b) λ over iterations with no generator reaching its limit and strong network connections

4.5.10 Case study 10: Application in large micro-grid

This case study demonstrates the scalability of the proposed algorithm to large micro-grids. Consider that there are 100 generators in the micro-grid. The communication is defined as below. Let the out-neighbor of generator i be $N_i^- = \{ \mod (i+k, 100) \mid k = 0, 1, \ldots, 20\}$, i.e., the communication is a circular graph, which is strongly connected. In the case study, the learning gain $\epsilon = 1.0e-4$. The initialization of $x_i(0), \lambda_i(0), y_i(0)$ are determined by the method in *Theorem 3*. The target demand $D = 25,000kW$. Fig. 4.12 shows the numerical results. The generators outputs are confined within the operation ranges. The collective estimated mismatches converge to zero. Total power generated by the micro-grid converges to the target demand, and all the incremental costs converge to the same value. In this particular case, the optimal incremental cost is $\lambda^* = 8.5074\$/kWh$, Type 1 generator

outputs are $x_1 = 460.36kW$, Type 2 generator outputs are $x_2 = 169.44kW$, and Type 3 generator outputs are $x_3 = 55.75kW$.

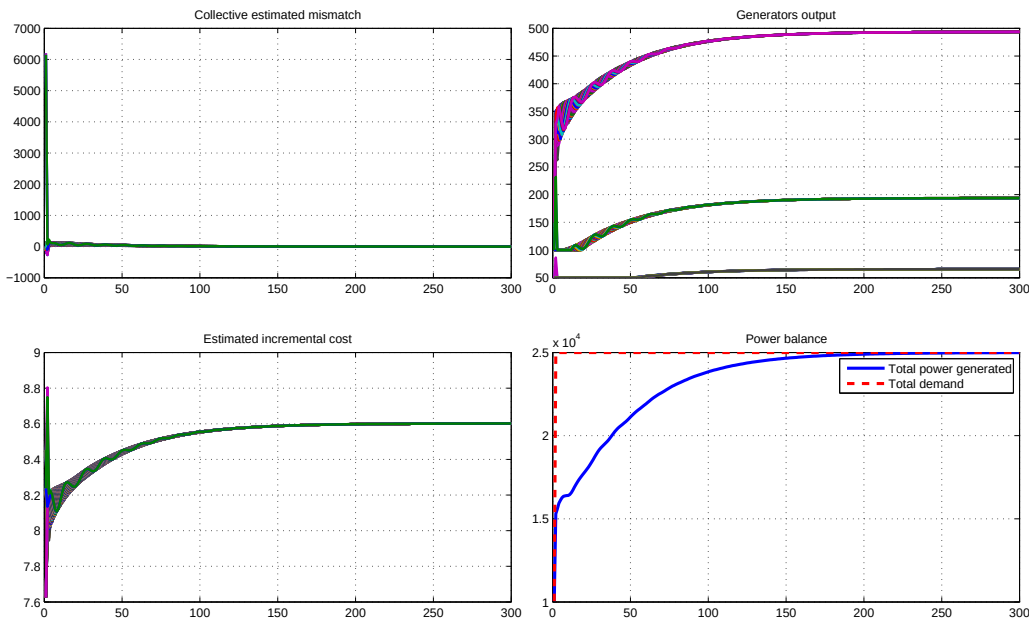


Figure 4.12: Results obtained with 100 generators.

4.6 Conclusion

In this chapter, a novel consensus based algorithm is proposed to solve ED problem in a distributed manner. The quadratic convex cost function models are used in the problem formulation, and strongly connected communication is sufficient for the information exchange. By the proposed scheme, the distributed algorithm enables generators to collectively learn the mismatch between demand and total power generation. The estimated mismatch is used to adjust current power generation by each generator. All generators are able to collectively minimize the total cost while satisfying power balance constraints. In addition, a design method is developed to calculate

the appropriate learning gains. Furthermore, the proposed algorithm can be implemented in a fully distributed fashion with a modified initialization. Numerous examples are presented to illustrate the effectiveness of the proposed algorithm. Comparisons with centralized Lambda-Iteration method as well as distributed ED method "consensus + innovation" method are also carried out in the application examples. These comparisons also demonstrate that our technique requires less restrictive communication networks and provides better convergence results.

Chapter 5

Hierarchical Consensus Based Approach with Loss Consideration for Economic Dispatch Problem under Micro-grid Context

5.1 Introduction

Economic dispatch (ED) problem is one of the important problems in power system operation. Climate change and rising fuel costs have changed the mindset of power system stakeholders. It is hoped that the distributed and renewable energy resources are going to play a more important role in reducing the emissions. With the increasing demand for renewable and distributed

energy resources, the grid also needs improved controlling and monitoring of existing networks. To better integrate the distributed energy resources, the system operators need to operate the system actively using local resources. This indicates a shift from traditional control strategy to distributed control paradigm [116].

In all the existing related works, the power line loss is not considered. However, the power line loss affects the actual optimal power dispatch. To design a power network that operates in optimal condition, the power line loss has to be carefully taken care of.

In this work, we no longer adopt the incremental cost criteria as consensus variable because we take power line loss calculation into account. The consensus variable is chosen as the product of the penalty factor and the incremental cost. Our approach is designed to have two levels of algorithms. The upper level constructs a set of distributed linear loss functions based on the previous "optimal power combination", which are supplied to the lower level algorithm. The lower level makes sure the consensus variable is driven to reach agreement among all agents while satisfying system constraints with the linear loss function, and generate a new "optimal power combination". Then, the new "optimal power combination" is fed back to the upper level algorithm. The whole process repeats until the "optimal power combination" converges. It turns out that the limit is the actual optimal power combination with loss consideration. Furthermore, each generator does not need to know the cost function parameters of other generators in our work. The novelty of the proposed algorithm is that it can estimate the mismatch between demand, loss and total generated power in a collective sense. The

local estimated mismatch may not equal to the actual mismatch in the beginning, but the summation of all the local estimated mismatch is preserved and exactly equal to the actual mismatch. The local estimated mismatch is used to adjust the power generation as if they are the true mismatch. The consensus variable is guaranteed to converge to the optimal value by the algorithm. In addition, in order to demonstrate that our proposed technique can work under weak communication environment, the communication graph is assumed to be strongly connected, which is less restrictive than the bidirectional information exchange in [75][109].

This chapter is organized as follows. In Section 5.2, graph theory, basic consensus results, and equal incremental cost criterion in ED problem are briefly introduced. Problem description and main results for both constrained and unconstrained cases are presented in Section 5.3. To demonstrate the effectiveness of the proposed algorithms, numerical example is shown in Section 5.4. Lastly, we conclude the chapter in Section 5.5.

5.2 Preliminary

The basic graph terminologies and consensus algorithms are introduced in Chapter 4. Thus, in this section, only analytic ED problem solution is introduced.

5.2.1 Analytic solution to ED problem with loss calculation

A micro-grid usually consists of multiple power generators. Let's assume there are N power generators. The operating cost function of power generation is given by the following quadratic form

$$C_i(x_i) = \frac{(x_i - \alpha_i)^2}{2\beta_i} + \gamma_i, \tag{5.1}$$

where x_i is the power generated by generator i, $\alpha_i \leq 0$, $\beta_i > 0$, and $\gamma_i \leq 0$.

The ED problem is to minimize the total generation cost

$$\min \sum_{i=1}^{N} C_i(x_i), \tag{5.2}$$

subject to the following two constraints,

Generator constraint:

$$\underline{x}_i \leq x_i \leq \overline{x}_i, \tag{5.3}$$

where $\underline{x}_i$ and $\overline{x}_i$ are the lower and upper bounds of the generator capability.

Demand constraint:

$$\sum_{i=1}^{N} x_i = D + L, \tag{5.4}$$

where D is the total demand and L is the total loss on the transmission lines satisfying $\sum_{i=1}^{N} \underline{x}_i < D + L < \sum_{i=1}^{N} \overline{x}_i$, i.e., the problem is solvable.

The incremental cost for the generator i is $\frac{dC_i(x_i)}{dx_i} = \frac{x_i - \alpha_i}{\beta_i}$. The solution

to ED problem with loss calculation is [110].

$$\begin{cases} PF_i \frac{x_i - \alpha_i}{\beta_i} = \lambda^* & \text{for } \underline{x}_i < x_i < \overline{x}_i \\ PF_i \frac{x_i - \alpha_i}{\beta_i} < \lambda^* & \text{for } x_i = \overline{x}_i \\ PF_i \frac{x_i - \alpha_i}{\beta_i} > \lambda^* & \text{for } x_i = \underline{x}_i \end{cases}, \tag{5.5}$$

where λ^* is the consensus variable and PF_i is the penalty factor of unit i given by

$$PF_i = \frac{1}{1 - \frac{\partial L}{\partial x_i}} \tag{5.6}$$

and $\frac{\partial L}{\partial x_i}$ is unit i incremental loss. The penalty factors are computed from losses represented using B coefficients [86]:

$$L = X^T[B]X + B_0^T X + B_{00} \tag{5.7}$$

where $X = [x_1, \ldots, x_N]^T$ is the vector of all generators' outputs, $[B]$ is the square matrix, B_0^T is the vector of the same length as X and B_{00} is a constant.

Note that the parameter γ_i in the cost function does not affect the incremental cost.

Remark 6. *The cost function in (5.1) is slightly different from the one in (5.8), which is commonly used by power engineers.*

$$C_i(x_i) = a_i x_i^2 + b_i x_i + c_i. \tag{5.8}$$

In fact, cost functions (5.1) and (5.8) are equivalent. It is not difficult to convert (5.8) to (5.1). Straightforward manipulation shows by setting $\alpha_i =$

$-\frac{b_i}{2a_i}$, $\beta_i = \frac{1}{2a_i}$, *and* $\gamma_i = c_i - \frac{b_i^2}{4a_i}$, *(5.1) and (5.8) are identical. The main motivation of using (5.1) is for notational simplicity in the next section.*

5.3 Main Results

Assume there is a command vertex which distributes the total demand D to a subset of $\mathscr{V}$. Denote the command vertex by vertex 0, and its out-neighbor set N_0^-. Recall that N_0^- is the vertices set which can receive information from vertex 0. For simplicity, let the command vertex distribute the total demand equally among all the generators in N_0^-. By assumption, $1 \leq |N_0^-| \leq N$. In this Section, we propose a two-level consensus algorithm that solves the economic dispatch problem distributively with loss consideration. The upper level algorithm estimates the power loss on the transmission lines and send the loss information to the lower level algorithm. The lower level algorithm then learns and eliminates the mismatch information between power output, demand and power loss in a distributed manner. After that, the results of the lower level computation are fed back to the upper level. This process repeats until there is no discrepancy between two latest iterations.

5.3.1 Upper level: estimating the power loss

The loss on the transmission lines is computed using B coefficients [86]:

$$L(X) = X^T[B]X + B_0^T X + B_{00}. \tag{5.9}$$

Note that the loss is in quadratic form of X. It is difficult to directly

incorporate the nonlinear loss function to design a distributed algorithm that solves the ED problem. The task of upper level algorithm is to generate a set of distributed linear loss functions. To achieve the target, we approximate the loss function with first order Taylor approximation.

The gradient of $L(X)$ is

$$\begin{aligned} \mathbf{g} = [g_1, g_2, \ldots, g_N] &= \left.\frac{\partial L}{\partial X}\right|_{X=X_{new}} \\ &= (2BX_{new} + B_0)^T, \end{aligned}$$

where X_{new} is the current "optimal power combination" estimation obtained from lower level algorithm.

The distributed linear loss function for generator i is

$$L_i(x_i) = g_i x_i + d,$$

where g_i is the ith component of $\mathbf{g}$, and

$$d = \frac{L(X_{new}) - \mathbf{g}X_{new}}{N}.$$

The loss function $L_i(x_i)$ depends only on x_i, and

$$\sum_{i=1}^{N} L_i(X_{new}^i) = L(X_{new}),$$

where X_{new}^i is the ith component of X_{new}. It implies that when each generator works at the operation point X_{new}^i, the linear loss function is a perfect estimator of the quadratic loss function.

Then, the linear loss function is passed to the lower level algorithm, and the lower level algorithm would return a power combination X_{lower}. We update the new "optimal power combination" by the following equation,

$$X_{new} = wX_{old} + (1 - w)X_{lower}, \tag{5.10}$$

where w is the weighted value that determines how fast X_{new} is moving away from the X_{old}. If X_{new} converges, that is $X_{new} = X_{old} = X_{lower}$, then X_{new} is the actual optimal power combination with loss consideration, provided that all the constraints are satisfied.

With the linear loss function, the penalty factor in (5.6) becomes,

$$PF_i = \frac{1}{1 - g_i}.$$

5.3.2 Lower level: solving economic dispatch distributively

According to the incremental cost criterion (5.5), when all generators operate at the optimal configuration, equal incremental costs equal to the optimal value, that is

$$PF_i \frac{x_i^* - \alpha_i}{\beta_i} = \lambda^*, \forall i \in \mathscr{V}. \tag{5.11}$$

Hence, the optimal power generation for each individual generator can be calculated if the optimal incremental cost λ^* is known, i.e.,

$$x_i^* = \frac{\beta_i \lambda^*}{PF_i} + \alpha_i, \forall i \in \mathscr{V}. \tag{5.12}$$

(5.11) and (5.12) motivate us the following algorithm. Let $\lambda_i(k)$ be the estimation of optimal incremental cost by generator i, $x_i(k)$ be the corresponding power generation which is an estimation of optimal power generation, $l_i(k)$ be the corresponding loss information, and $y_i(k)$ be the collective estimation of the mismatch between demand and total power generation.

Initializations:

$$\begin{cases} \lambda_i(0) = & \text{any fixed admissible value} \\ x_i(0) = & \text{any fixed admissible value} \\ l_i(0) = & 0 \\ y_i(0) = & \begin{cases} \frac{D}{|N_0^-|} - (x_i(0) - l_i(0)) & \text{if } i \in N_0^- \\ -(x_i(0) - l_i(0)) & \text{otherwise} \end{cases} \end{cases} , \forall i \in \mathscr{V}.$$

We are ready to state the main algorithm.

$$\lambda_i(k+1) = \sum_{j \in N_i^+} p_{i,j}\lambda_j(k) + \epsilon y_i(k) \tag{5.13a}$$

$$x_i(k+1) = \frac{\beta_i \lambda_i(k+1)}{PF_i} + \alpha_i \tag{5.13b}$$

$$l_i(k+1) = g_i x_i(k+1) + d \tag{5.13c}$$

$$\begin{aligned} y_i(k+1) = & \sum_{j \in N_i^+} q_{i,j} y_j(k) + (l_i(k+1) - l_i(k)) \\ & - (x_i(k+1) - x_i(k)) \end{aligned} \tag{5.13d}$$

where ϵ is a sufficiently small positive constant.

To analyze the properties and convergence of algorithm (5.13), rewrite it

in the following matrix form

$$\boldsymbol{\lambda}(k+1) = P\boldsymbol{\lambda}(k) + \epsilon\mathbf{y}(k) \tag{5.14a}$$

$$\mathbf{x}(k+1) = \mathbf{PF}\beta\boldsymbol{\lambda}(k+1) + \boldsymbol{\alpha} \tag{5.14b}$$

$$\mathbf{l}(k+1) = \mathbf{G}\mathbf{x}(k+1) + d \tag{5.14c}$$

$$\begin{aligned}\mathbf{y}(k+1) &= Q\mathbf{y}(k) + (\mathbf{l}(k+1) - \mathbf{l}(k)) \\ &\quad - (\mathbf{x}(k+1) - \mathbf{x}(k))\end{aligned} \tag{5.14d}$$

where $\mathbf{x}, \mathbf{y}, \mathbf{l}, \boldsymbol{\alpha}, \boldsymbol{\lambda}$ are the column stack vector of $x_i, y_i, l_i, \alpha_i, \lambda_i$ respectively, $\mathbf{PF} = diag([1/PF_1, 1/PF_2, \ldots, 1/PF_N])$, $\mathbf{G} = diag([g_1, g_2, \ldots, g_N])$, and $\beta = diag([\beta_1, \beta_2, \ldots, \beta_N])$.

(5.14d) preserves the summation of $x_i(k) + y_i(k) + l_i(k)$ over $\mathscr{V}$. It can be verified by premultiply both sides of (5.14d) by $\mathbf{1}^T$, and noticing that Q is column stochastic, we have

$$\begin{aligned}&\mathbf{1}^T\mathbf{y}(k+1) = \mathbf{1}^T Q\mathbf{y}(k) + \mathbf{1}^T(\mathbf{l}(k+1) - \mathbf{l}(k)) \\ &-\mathbf{1}^T(\mathbf{x}(k+1) - \mathbf{x}(k)), \\ &\mathbf{1}^T\mathbf{y}(k+1) = \mathbf{1}^T\mathbf{y}(k) + \mathbf{1}^T(\mathbf{l}(k+1) - \mathbf{l}(k)) \\ &-\mathbf{1}^T(\mathbf{x}(k+1) - \mathbf{x}(k)), \\ &\mathbf{1}^T(\mathbf{y}(k+1) + \mathbf{x}(k+1)) = \mathbf{1}^T(\mathbf{l}(k+1) - \mathbf{l}(k)) \\ &+\mathbf{1}^T(\mathbf{y}(k) + \mathbf{x}(k)).\end{aligned}$$

Notice the initialization of $x_i(0)$, $y_i(0)$ and $l_i(0)$, we can obtain $\sum_{i\in\mathscr{V}} x_i(0) + y_i(0) = D + L(0)$. The terms $\sum_{i\in\mathscr{V}}(l_i(k+1) - l_i(k))$ estimate the loss $L(k)$

on the transmission lines. Hence, $\mathbf{1}^T\mathbf{y}(k) = D + L(k) - \mathbf{1}^T\mathbf{x}(k)$ is the actual mismatch between sum of demand and loss and total power generation. The mismatch is obtained via a collective effort from all individual generators rather than a centralized method. That is the reason why $y_i(k)$ is called collective estimation of the mismatch. The first term in the right hand side of (5.14a) is the consensus part, it drives all $\lambda_i(k)$ to a common value. The second term $\epsilon\mathbf{y}(k)$ provides a feedback mechanism to ensure $\lambda_i(k)$ converges to the optimal value λ^*. (5.14b) just updates the estimated power generation $x_i(k)$ to the newest one.

Theorem 4. *In algorithm (5.13), if the positive constant ϵ is sufficiently small, then the algorithm is stable, and all the variables converge to the solution to the ED problem, i.e.,*

$$\lambda_i(k) \to \lambda^*,\ x_i(k) \to x_i^*,\ y_i(k) \to 0, \textit{as } k \to \infty, \forall i \in \mathscr{V}.$$

Proof. We use the eigenvalue perturbation [111] approach to analyze the convergence properties. Replace $\mathbf{x}$ in (5.14c) with $\boldsymbol{\lambda}$ by using (5.14a)(5.14b), we have

$$\begin{aligned} \mathbf{y}(k+1) &= (Q - \epsilon\mathbf{PF}\beta + \epsilon\mathbf{GPF}\beta)\mathbf{y}(k) + \\ &\quad (I - \mathbf{G})\mathbf{PF}\beta(I - P)\boldsymbol{\lambda}(k), \end{aligned} \tag{5.15}$$

where I is the identity matrix of appropriate dimension.

Write (5.14a) and (5.15) in matrix form, we get the following composite

system,

$$\begin{bmatrix} \boldsymbol{\lambda}(k+1) \\ \mathbf{y}(k+1) \end{bmatrix} = \begin{bmatrix} P & \epsilon I \\ (I-\mathbf{G})\mathbf{PF}\beta(I-P) & Q-\epsilon\mathbf{PF}\beta+\epsilon\mathbf{GPF}\beta \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}(k) \\ \mathbf{y}(k) \end{bmatrix} \tag{5.16}$$

Define $M \triangleq \begin{bmatrix} P & 0 \\ (I-\mathbf{G})\mathbf{PF}\beta(I-P) & Q \end{bmatrix}$ and $E \triangleq \begin{bmatrix} 0 & I \\ 0 & -\mathbf{PF}\beta+\mathbf{GPF}\beta \end{bmatrix}$.

The system matrix of (5.16) can be regarded as M perturbed by ϵE. M is a lower block triangular matrix, the eigenvalues of M is the union of the eigenvalues of P and Q. So M has two eigenvalues equal to 1, and the rest eigenvalues lie in the open unit disk on the complex plane. Denote the two values by $\theta_1 = \theta_2 = 1$. It is easy to verify that $\mathbf{u}_1$, $\mathbf{u}_2$ and $\mathbf{v}_1^T$, $\mathbf{v}_2^T$ are the two linearly independent right and left eigenvectors of M.

$$U = [\mathbf{u}_1, \mathbf{u}_2] = \begin{bmatrix} \mathbf{0} & \mathbf{1} \\ \boldsymbol{\mu} & -\eta\boldsymbol{\mu} \end{bmatrix}, \tag{5.17}$$

where $\eta = [\sum_{i=1}^{N}(1-a_i)][\sum_{i=1}^{N}\beta_i]$.

$$V^T = \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \end{bmatrix} = \begin{bmatrix} \mathbf{1}^T(I-\mathbf{G})\mathbf{PF}\beta & \mathbf{1}^T \\ \boldsymbol{\omega}^T & \mathbf{0}^T \end{bmatrix}. \tag{5.18}$$

Furthermore, $V^T U = I$.

When ϵ is small, the variation of θ_1 and θ_2 perturbed by ϵE can be

quantified by the eigenvalues of $V^T EU$.

$$V^T EU = \begin{bmatrix} 0 & 0 \\ \boldsymbol{\omega}^T \boldsymbol{\mu} & -\eta \boldsymbol{\omega}^T \boldsymbol{\mu} \end{bmatrix}.$$

The eigenvalues of $V^T EU$ are 0 and $-\eta \boldsymbol{\omega}^T \boldsymbol{\mu} < 0$. So $\frac{d\theta_1}{d\epsilon} = 0$ and $\frac{d\theta_2}{d\epsilon} = -\eta \boldsymbol{\omega}^T \boldsymbol{\mu} < 0$. That means θ_1 does not change against ϵ, and when $\epsilon > 0$, θ_2 becomes smaller. Let δ_1 be the upper bound of ϵ such that when $\epsilon < \delta_1$, $|\theta_2| < 1$. Since eigenvalues continuously depend on the entries of a matrix, in our particular case, the rest of eigenvalues of $M + \epsilon E$ continuously depend on ϵ. Therefore, there exists an upper bound δ_2 such that when $\epsilon < \delta_2$, $|\theta_j| < 1$, $j = 3, 4, \ldots, 2N$. Hence, if we choose $\epsilon < \min(\delta_1, \delta_2)$, we can guarantee that the eigenvalue $\theta_1 = 1$ is simple, and all the rest eigenvalues lie in the open unit disk.

It can be verified $\begin{bmatrix} \mathbf{1} \\ \mathbf{0} \end{bmatrix}$ is the eigenvector of system matrix in (5.16) association with $\theta_1 = 1$. Since all the rest eigenvalues are within the open unit disk,

$$\begin{bmatrix} \boldsymbol{\lambda}(k) \\ \mathbf{y}(k) \end{bmatrix} \text{ converges to } span \begin{bmatrix} \mathbf{1} \\ \mathbf{0} \end{bmatrix}$$

as $k \to \infty$. That is $y_i(k) \to 0$. From (5.14c), we can derive $\mathbf{1}^T \mathbf{x}(k) = D$, i.e., the demand constraint is satisfied. From (5.14a), $\lambda_i(k)$ converges to a common value, i.e., the incremental cost criterion is satisfied. Therefore, we can conclude *Theorem 2*. □

5.4 Application Examples

The communication topology is shown in Fig. 5.1. We assume that there are four generators in this micro-grid, labeled as vertex 1, 2, 3, and 4. The four generators are selected from the three types of generators in Table. 5.1. Vertices 1 and 2 are Type A generators, vertex 3 is a Type B generator, and vertex 4 is a Type C generator [110]. The communication among the four generators are denoted by solid lines, which is a strongly connected graph. Vertex 0 is the command vertex. The dashed line represents the communication between command vertex and power generators, that is the command vertex can send information to power generator 1 and 3. Based on Fig. 5.1, matrices P and Q can be defined as

$$P = \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}, \; Q = \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{1}{2} & 0 & \frac{1}{3} \end{bmatrix}$$

respectively. The power lines loss matrix is defined as [110]

$$B = \begin{bmatrix} 0.00676 & 0.00953 & -0.00507 & 0.00211 \\ 0.00953 & 0.0521 & 0.00901 & 0.00394 \\ -0.00507 & 0.00901 & 0.0294 & 0.00156 \\ 0.00211 & 0.00394 & 0.00156 & 0.00545 \end{bmatrix},$$

$$B_0 = \begin{bmatrix} -0.0766 \\ -0.00342 \\ 0.0189 \\ 0.0173 \end{bmatrix}, \; B_{00} = 0.0595.$$

The initial values in case studies $1-3$ are given in Table. 5.2 [110], w is chosen to be 0.8, and total demand $D = 1300MW$.

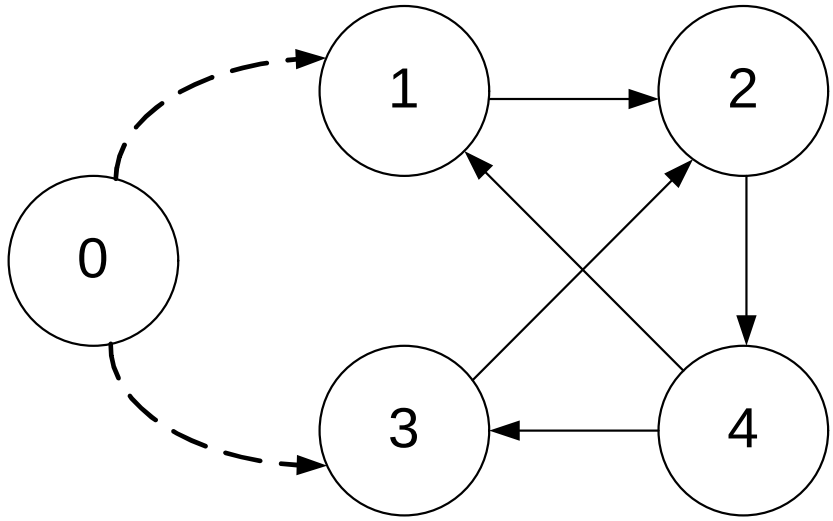


Figure 5.1: Communication topology among generators and command vertex in the network.

Table 5.1: Generator parameters

Generator Type	A (Coal-fired)	B (Oil-fired)	C (Oil-fired)
Range (MW)	[150, 600]	[100, 400]	[50, 200]
a ($\$/MW^2h$)	0.00142	0.00194	0.00482
b ($\$/MWh$)	7.2	7.85	7.97
c ($\$/h$)	510	310	78
α (MW)	-2535.2	-2023.2	-826.8
β ($MW^2h/\$$)	352.1	257.7	103.7
γ ($\$/h$)	-8616.8	-7631.0	-3216.7

Table 5.2: Initializations

Variables	$i = 1$	$i = 2$	$i = 3$	$i = 4$
$x(0)$ (MW)	150	150	100	50
$y(0)$ (MW)	600	-150	650	-50
$l(0)$(MW)	0	0	0	0
$\lambda(0)$	736.29	957.38	906.46	871.24

5.4.1 Case study 1: convergence test

In this case study, we want to verify the convergence property. The results are shown in Fig. 5.2. The sub-figure on the top shows the estimated power loss and actual power loss calculated by the upper level. The estimated power loss reaches a steady state and the value is equal to the actual calculated loss after 14 iterations. This demonstrates the convergence of the upper level algorithm. The other four sub-figures at the bottom show the dynamics of the lower level algorithm. The collective estimated mismatch y_i goes to zero after 30 iterations. The estimated λ of all generators converge to the same value while meeting the power balance constraint. From the transient response, we notice that generator 1, 3, and 4 are saturated after the first 25 iterations, and they gradually increase to the final output as the λ is increasing. Based on the results, $L = 76.02MW$, $\lambda^* = 1246.3$, $x_1^* = 600MW$, $x_2^* = 176.01MW$, $x_3^* = 400MW$, and $x_4^* = 200MW$. All the power generations are within the generation ranges. No power generators exceed the operation ranges even in the transient responses.

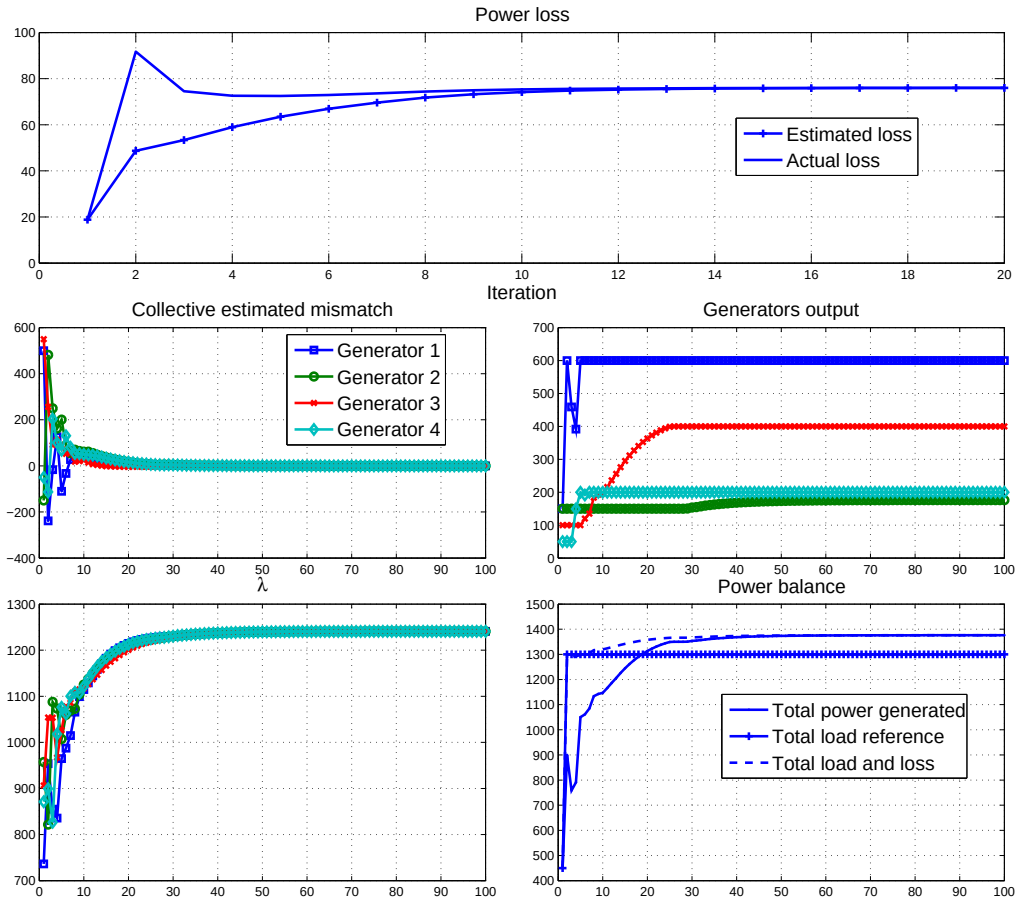


Figure 5.2: Results obtained with generator constraints.

5.4.2 Case study 2: robustness of command node connections

Our proposed algorithm only requires that the communication between generators is strongly connected. If the command vertex has at least one edge to any one of the generators, the algorithm works when the learning gain is appropriately chosen. This is a very flexible communication condition. In the previous case study, the command vertex is connected to generators 1 and 3. Now, change the command vertex connection to generators 2 and 4.

The purpose of this case study is to demonstrate that the proposed algorithm works regardless of the command node connections. Fig. 5.3 shows the numerical results. The top sub-figure shows the upper level of the algorithm. The estimated loss matches the actual loss information after 14 iterations. The four sub-figures below show the lower level of the algorithm. The mismatch information is collectively minimized by all the generators after 30 iterations. The incremental costs also reach consensus after 30 iterations. As shown in the fourth sub-figure, the total demand and loss are balanced by total power. The initial conditions for $x_i(0)$ are still the same as in Table. 5.2. The initial conditions for $y_i(0)$ are changed to $y_1(0) = -150MW$, $y_2(0) = 500MW$, $y_3(0) = -100MW$, and $y_4(0) = 600MW$, and the final outcomes are $L = 76.02MW$, $\lambda^* = 1246.3$, $x_1^* = 600MW$, $x_2^* = 176.01MW$, $x_3^* = 400MW$, and $x_4^* = 200MW$. The results are identical to Case study 1, which indicates the command vertex connection does not affect the final convergence results.

5.4.3 Case study 3: plug and play test

One of the most important features of a micro-grid is its plug and play adaptability. In this case study, the four generators have already reached the optimal states before plugging in the fifth generator. The four generators steady states are the same as shown in case study 1. The fifth generator is plugged in at time step $k = 20$. The fifth generator is a Type B generator. The initializations of the generators are the same as in case study 1. The new communication topology is shown in Fig. 5.4. Thus matrices P and Q

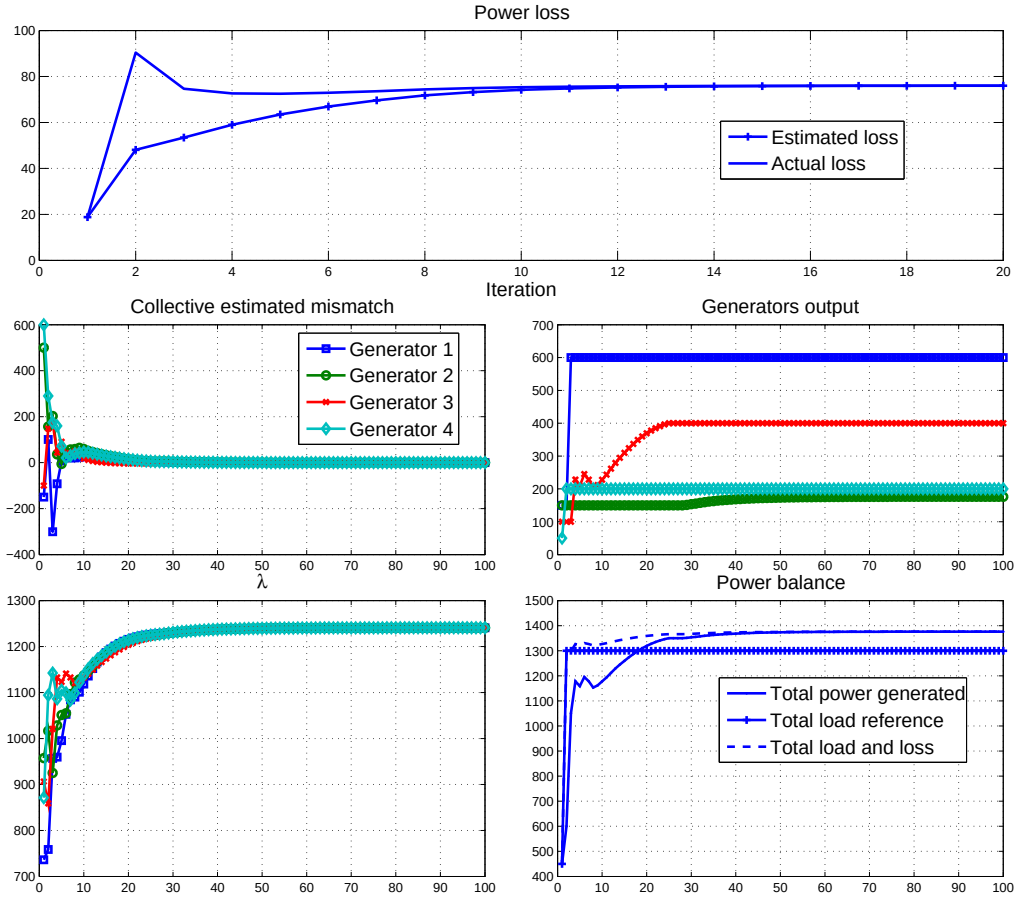


Figure 5.3: Robustness test when the command vertex is connected to generators 2 and 4.

are changed to

$$P = \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}, \; Q = \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{3} & 0 \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{2} \\ 0 & \frac{1}{3} & 0 & 0 & \frac{1}{2} \end{bmatrix}$$

respectively. After plugging in the fifth generator at $k = 20$, the output of generator 5 is set to $x_5(20) = 100MW$ and $y_5(20) = -100MW$. From the results in Fig. 5.5, we can observe that the local estimated mismatch y_i goes to zero after a short disturbance. The other three generators reduce their outputs in order to accommodate the fifth generator output. Therefore, the incremental cost drops due to lower average output and lower loss on the transmission lines. Finally, the estimated incremental costs of all generators converge to the same value while meeting the power balance constraint. Based on the results, $L = 37.27MW$, $\lambda^* = 950.68$, $x_1^* = 600MW$, $x_2^* = 150MW$, $x_3^* = 178.94MW$, and $x_4^* = 86.80MW$, and $x_5^* = 321.53MW$. All the power generations are within the generation range. The optimization goal is fulfilled and the fifth generator is well adapted into the system.

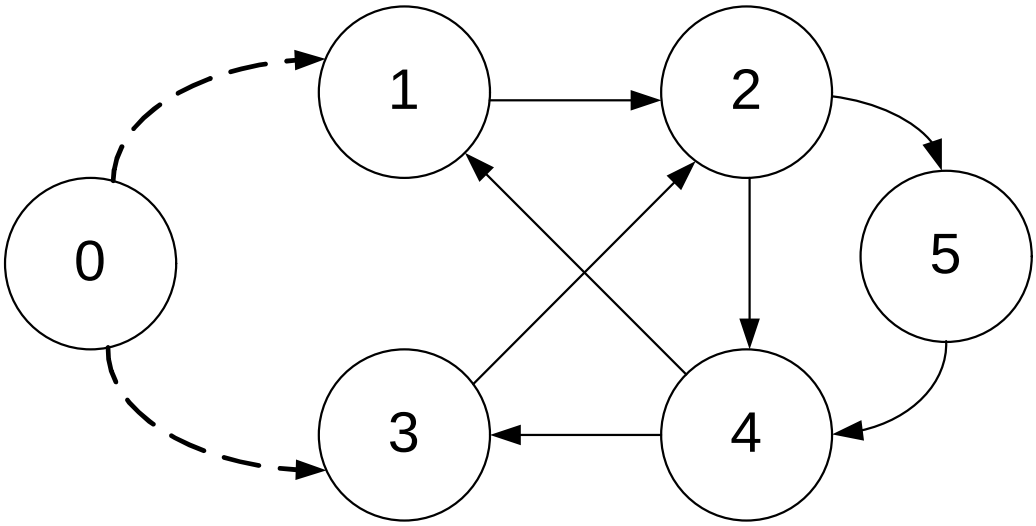


Figure 5.4: Communication topology with the fifth generator.

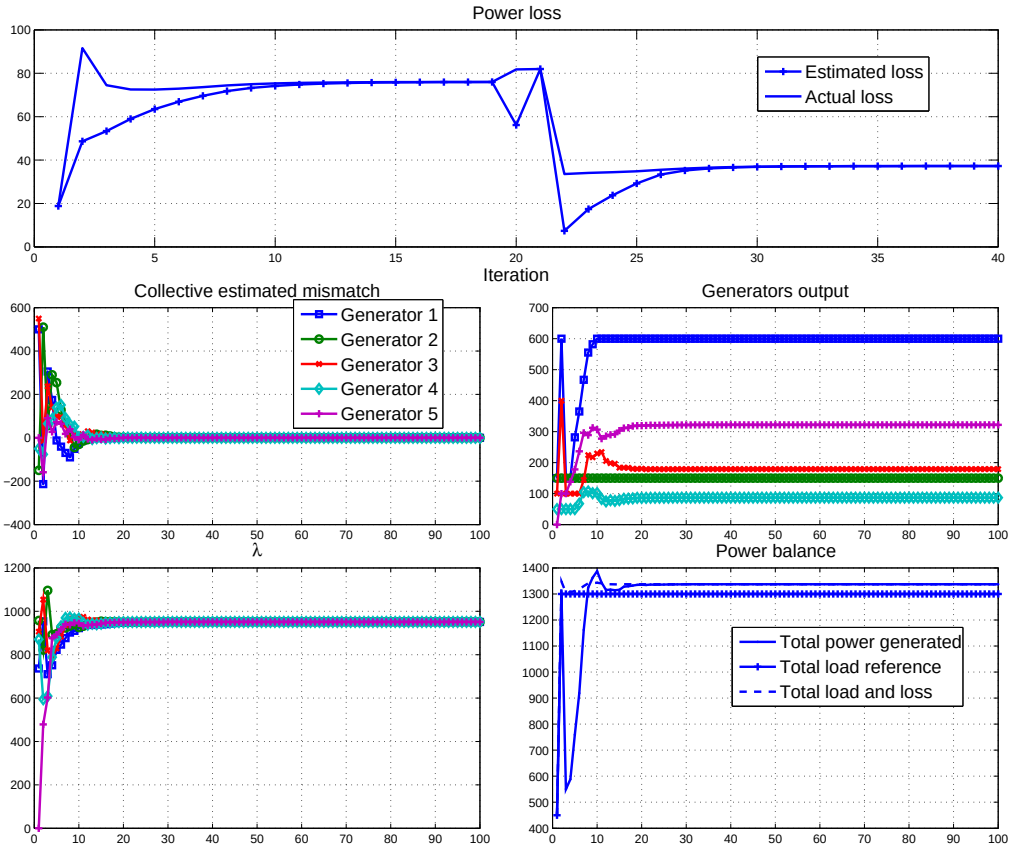


Figure 5.5: Results obtained with the fifth generator.

5.4.4 Case study 4: time-varying demand

In this case study, the generators' setup is the same as in case study 1. In a practical situation, it is very likely that the demand is not a constant over time. Slight modification of our proposed algorithm can handle the demand change effectively. Let the initial demand $D = 1300MW$ as usual, we purposely change it to $D' = 800MW$ at upper loop time step $k = 20$. The demand change is only known to generators 1 and 3 since they are connected to the command vertex. Thus, the algorithm needs to modify the local estimated mismatch at $k = 20$ before continuing updating the variables.

Keep y_2, y_4 unchanged at $k = 20$, and update y_1, y_3 as follows at $k = 20$ before proceeding to the next updating iteration.

$$\begin{aligned} y_1(20) &= y_1(20) + \frac{D' - D}{m}, \\ y_3(20) &= y_3(20) + \frac{D' - D}{m}. \end{aligned}$$

Fig. 5.6 shows the results. After the demand changes from $1300MW$ to $800MW$ at $k = 20$, the algorithm quickly responds to the change of demand. The top sub-figure shows the upper level algorithm. The estimated loss match actual loss at upper level iteration $k = 30$. The four sub-figures below show the lower level algorithm. The mismatch information is obtained, and the generators minimize the mismatch information collectively. The new incremental cost consensus is reached at lower level iteration $k = 20$. The algorithm asymptomatically converges to the new optimal solution at iteration lower level iteration $k = 20$, i.e., $\lambda^* = 875.89$, $x_1^* = 510.32MW$, $x_2^* = 150MW$, $x_3^* = 104.66MW$, $x_4^* = 50.00MW$, and $L = 14.98MW$. Compared to case study 1 results, all the outputs decrease since the demand is reduced.

5.4.5 Case study 5: relation between convergence speed and learning gain

The learning gain is the only design parameter that we can manipulate, and it plays a significant role in convergence speed and also the performance. If the learning gain is not properly selected, the results may oscillate, or even

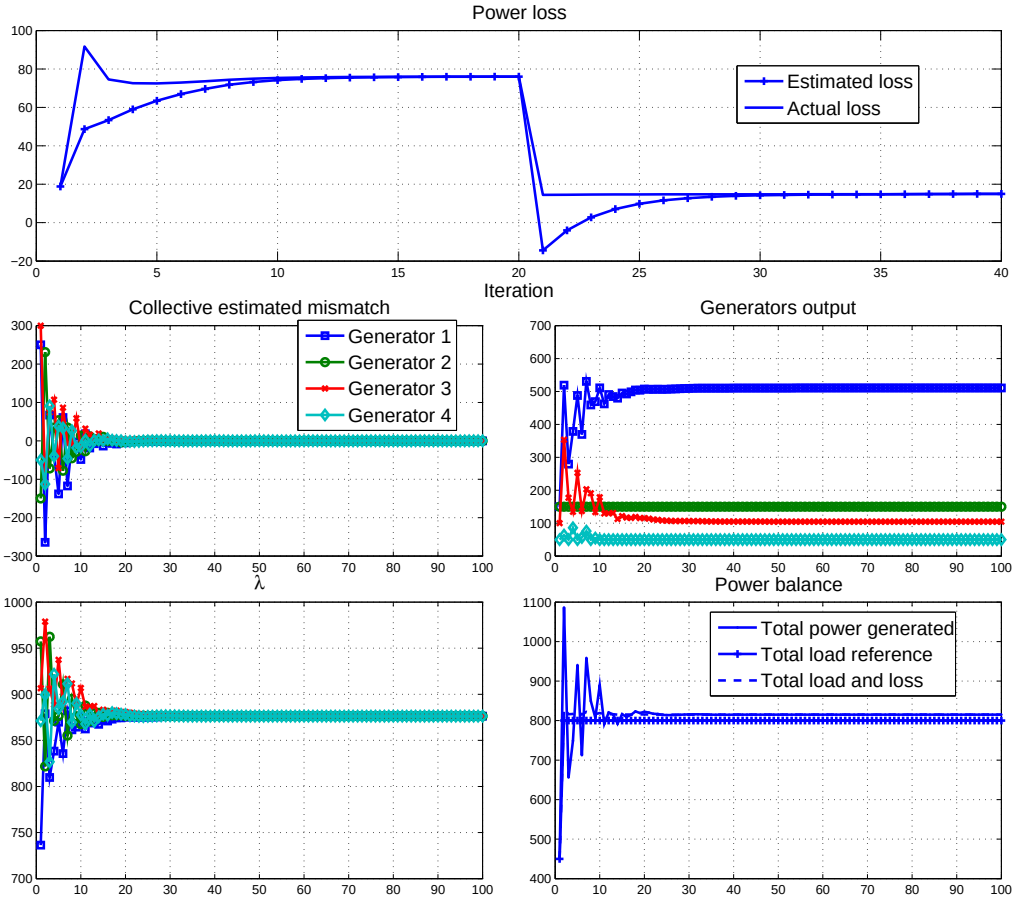


Figure 5.6: Results obtained with time-varying demand.

diverge. Thus, it is important to investigate the relation between convergence properties and learning gain. We select four sets of learning gains, i.e., $\epsilon = 20, 50, 80,$ and 110, and study the performances of the learning gains. Let the initial conditions be the ones specified in Table. 5.2. The mismatch between demand and total power generation is depicted in Fig. 5.7. In general, when the learning gain is small, the convergence speed is relatively slow, and the transient response is smooth. A large gain learning results in a faster convergence speed. However, the performance may become oscillatory. In

addition, if the learning goes beyond certain limit, the algorithm diverges as in the experiment with $\epsilon = 110$.

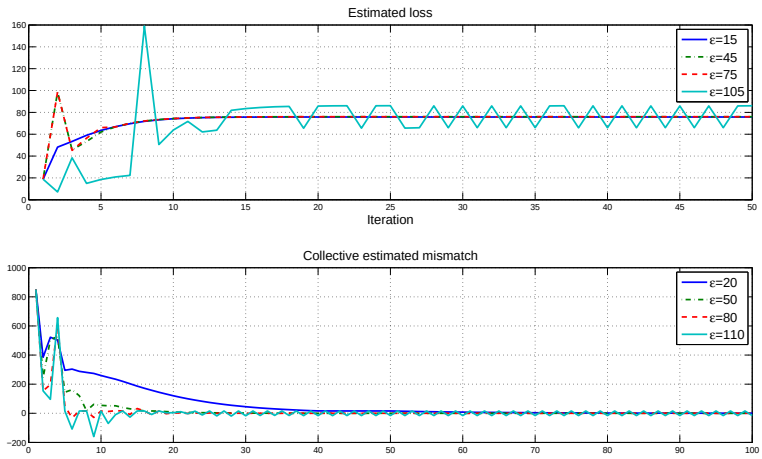


Figure 5.7: Relation between convergence speed and learning gain.

5.4.6 Case study 6: IEEE 14-bus system test

This case study demonstrates the fully distributed implementation of the proposed methods to IEEE 14-bus test systems. This test system represents a portion of the American Electric Power System which is located in the Midwestern US as of February in 1962. This system has 14 buses, 5 generators and 11 loads [115]. The parameters of the generators are listed in Table. 5.3 [78].

Notice that Buses 1, 2, 3, 6, and 8 contain generators. When a bus contains load only, the power generation at that bus is set to zero. The initial local demands at each buses are given as $D_1 = 0MW$, $D_2 = 21.7MW$, $D_3 = 66.2MW$, $D_4 = 47.8MW$, $D_5 = 7.6MW$, $D_6 = 11.2MW$, $D_7 = 0MW$, $D_8 = 0MW$, $D_9 = 29.5MW$, $D_{10} = 9.0MW$, $D_{11} = 3.5MW$, $D_{12} =$

Table 5.3: IEEE 14-bus test system generator parameters

Bus	a ($\$/MW^2h$)	b ($\$/MWh$)	c ($\$/h$)	Range (MW)
1	0.04	2.0	0	[0, 80]
2	0.03	3.0	0	[0, 90]
3	0.035	4.0	0	[0, 70]
6	0.03	4.0	0	[0, 70]
8	0.04	2.5	0	[0, 80]

$6.1MW$, $D_{13} = 13.5MW$, and $D_{14} = 14.9MW$. Thus the total demand is $D = 231MW$..

The communication among buses can be independent from the actual bus connections. Consider each bus as a vertex, and a vertex only sends information to the next two vertices, i.e., the edge set is $\mathscr{E} = \{(i, i+1), (i, i+2) | 1 \leq i \leq 12\} \bigcup \{(13, 14), (13, 1), (14, 1), (14, 2)\}$. The graph is strongly connected, and the total number of edges is $|\mathscr{E}| = 28$. The maximum possible number of edges is $|\mathscr{V}| \cdot |\mathscr{V}| = 196$. So the density of the graph is $42/196 = 0.214$, which is quite sparse.

By using the proposed algorithm with learning gain $\epsilon = 5e^{-3}$, the numerical results are presented in Fig. 5.8. The generators outputs are confined within the operation ranges. The collective estimated mismatches converge to zero. Total power generated by the micro-grid converges to the target demand, and all the λ converge to the same value, i.e., the ED problem is solved. In this particular case, the consensus variable is $\lambda^* = 8.15$, generator outputs are $x_1^* = 68.43MW$, $x_2^* = 70.32MW$, $x_3^* = 31.88MW$, $x_6^* = 38.42MW$, and $x_8^* = 43.20MW$.

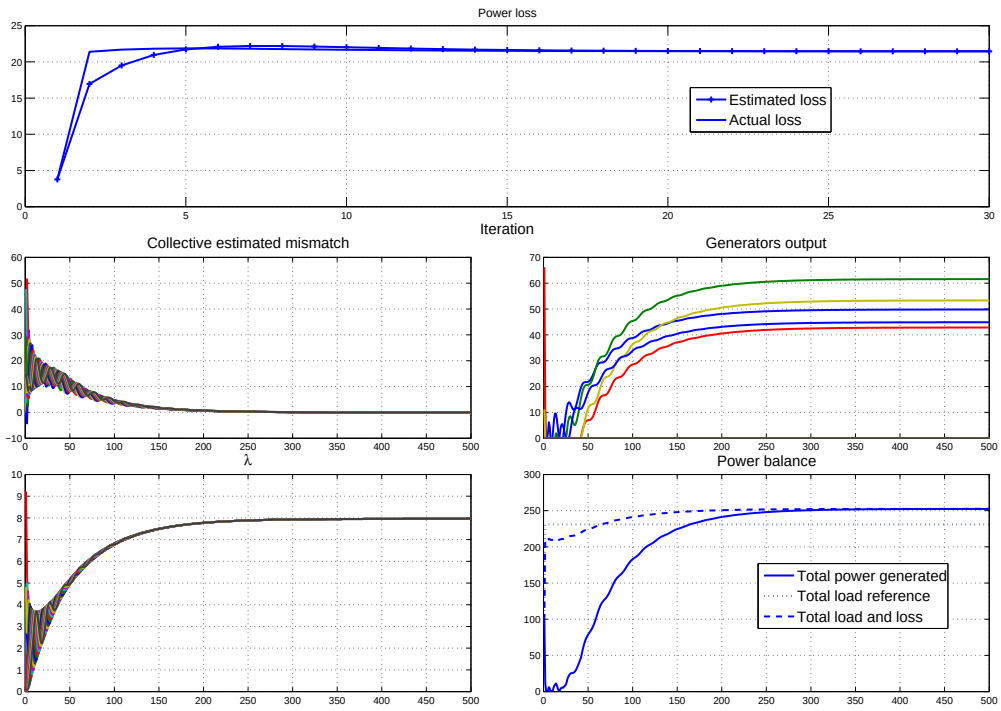


Figure 5.8: Results obtained with IEEE 14-bus test system.

5.4.7 Case study 7: IEEE 57-bus system test

The IEEE 57-bus test system represents a portion of the American Electric Power System in the early 1960s. This system has 57 buses, 7 generators and 42 loads [117].

This case study demonstrates the scalability of proposed algorithm on large systems with low level of connectivity. The communication among buses can be independent from the actual bus connections. Consider each bus as a vertex, and a vertex only sends information to the next two vertices, i.e., the edge set is $\mathscr{E} = \{(i, i{+}1), (i, i{+}2) | 1 \leq i \leq 55\} \bigcup \{(56, 57),\ (56, 1),\ (57, 1),\ (57, 2)\}$. The graph is strongly connected, and the total number of edges is $|\mathscr{E}| = 114$.

The maximum possible number of edges is $|\mathscr{V}| \cdot |\mathscr{V}| = 3249$. So the density of the graph is $171/3249 = 0.0526$, which is quite sparse.

The results of the proposed algorithm are presented in Fig. 5.9. The generators outputs are confined within the operation ranges. The collective estimated mismatches converge to zero. Total power generated by the micro-grid converges to the target demand, and all the λ converge to the same value, i.e., the ED problem is solved. In this particular case, the consensus variable is $\lambda^* = 25.1840$, generator outputs are $x_1^* = 137.13MW$, $x_2^* = 100MW$, $x_3^* = 44.11MW$, $x_6^* = 100MW$, $x_8^* = 485.27MW$, $x_9^* = 100MW$, and $x_12^* = 328.92MW$.

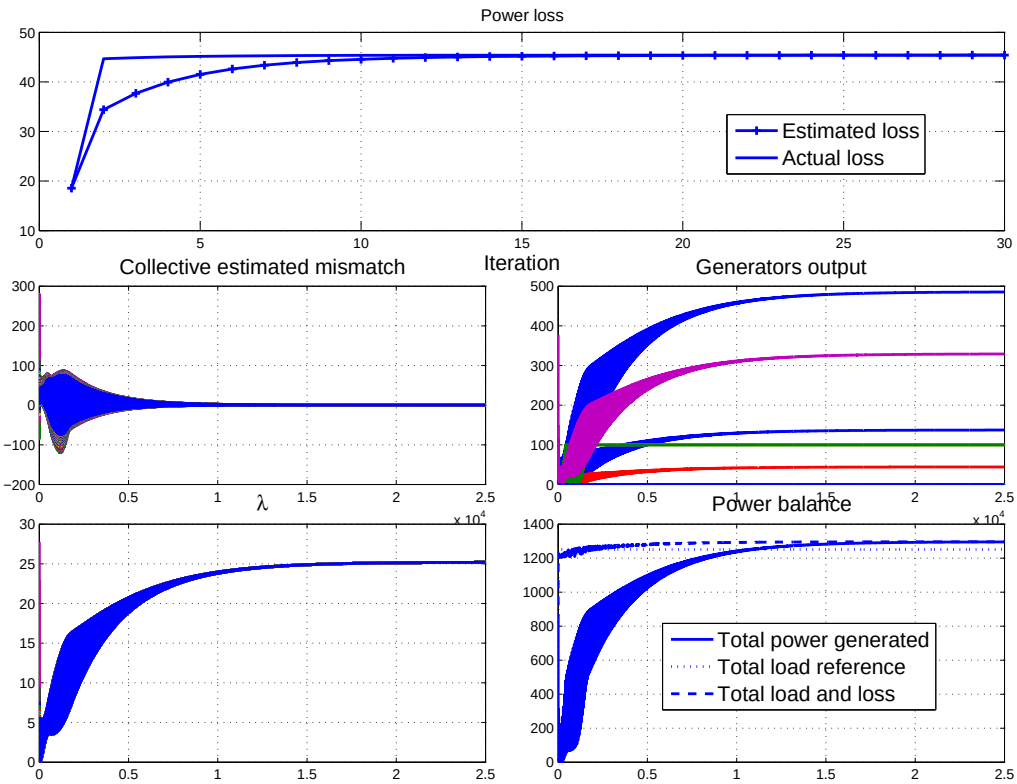


Figure 5.9: Results obtained with IEEE 57-bus test system.

5.4.8 Case study 8: Application in large micro-grid

This case study demonstrates the scalability of the proposed algorithm to very large networks. Consider that there are 100 generators in the micro-grid. The communication is defined as below. Let the out-neighbor of generator i be $N_i^- = \{ \mod (i+k, 100) \mid k = 0, 1, \ldots, 20\}$, i.e., the communication is strongly connected. Let the learning gain $\epsilon = 2$. The initialization of $x_i(0), \lambda_i(0), y_i(0)$ are determined by the method in *Theorem 3*. The target demand $D = 25000kW$. Fig. 5.10 shows the numerical results. The generators outputs are confined within the operation ranges. The collective estimated mismatches converge to zero. Total power generated by the micro-grid converges to the target demand, and all the incremental costs converge to the same value, the EDP is solved. In this particular case, the optimal incremental cost is $\lambda^* = 878.89$, Type 1 generator outputs are near $x_1 = 480kW$, Type 2 generator outputs are $x_2 = 180kW$, and Type 3 generator outputs are $x_3 = 62kW$.

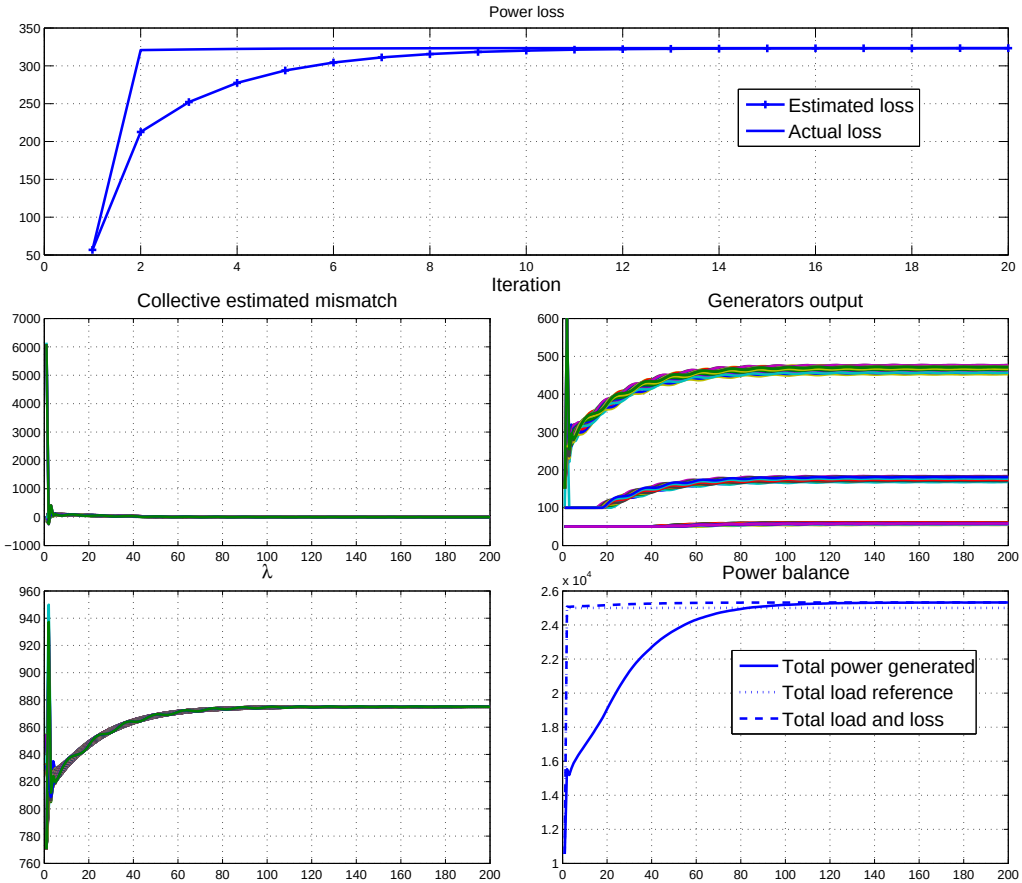


Figure 5.10: Results obtained with 100 generators.

5.5 Conclusion

In this work, a hierarchical consensus based algorithm is proposed to solve economic dispatch problem with loss calculation. The operating cost functions are modeled as quadratic functions. The loss information on the transmission lines is modeled using B coefficients. The upper level algorithm estimates the loss information and distribute to the lower level. The lower level algorithm drives the product of penalty factor and incremental cost to reach an

agreement among all agents. Sparse communication networks are adopted to demonstrate the convergence speed of the proposed algorithm. The case studies have shown that the total operating cost can be minimized while maintaining the power balance constraint. The proposed technique can also take care of demand variations and generator variations. In addition, the scalability of our technique is demonstrated by solving the large network problem.

Chapter 6

Conclusion and Future Work

ED problem has been discussed in literature more than 90 years. Various aspects of the problem such as security, reliability, and economy were investigated by previous researchers. Recently, with the increasing penetration of distributed generation, renewable energy plays a more and more important role in power network. To integrate the DERs into traditional grid, the economic issues such as uncertain power generation due to stochastic weather conditions and higher price renewable energy due to smaller generator scale must be carefully solved. This thesis addresses the economic dispatch issues of integrating renewable energy sources from two different aspects. The first part of the work presents two centralized approaches to solve the problem. Furthermore, the second part of the work proposes two distributed techniques to tackle the problem.

In Chapter 2, an economic load dispatch method for micro-grid using multi-objective optimization algorithms is presented. The emission and valve-point loading fuel cost of thermal generators are considered. The penalty and

reserve cost of the wind turbines are used to encourage the usage of renewable energy. The fuel cost and hydrogen management cost of PEM fuel cell generators are adopted. The objectives are minimizing operating cost as well as minimizing emission. Two state-of-the-art multi-objective optimization techniques are applied on this problem. Pareto fronts are obtained from three different systems to provide the operator a variety of choices. The results show that SPEA2 has a faster convergence when generation number is small and NSGA-II can perform slightly better for large number of generations. NSGA-II provides more diverse solutions than SPEA2.

In Chapter 3, an integrated approach which considers both reconfiguration of network and economic load dispatch is presented. The stochastic nature of wind, PV and load is taken into consideration by stochastic generation models. The forecasting of the wind, PV and load data and the energy storage system is considered to facilitate the dispatch process. The four bio-inspired optimization techniques are adopted to solve the problem. Optimization of benchmark-link application problems are conducted to investigate the effects of different weather and load conditions on the whole network. By using the integration approach, micro-grid can be incorporated into the network more effectively. Higher price renewable energy can be introduced into the grid to minimize the operating cost by reducing the power losses significantly. The network can adjust itself more efficiently to allow the utilization of the renewable energy.

In Chapter 4, a novel consensus based algorithm is proposed to solve ED problem in a distributed manner. Quadratic cost functions and strongly connected communication graph are used in the problem formulation. By

the proposed scheme, the distributed algorithm enables generators to collectively learn the mismatch between demand and total power generation. The estimated mismatch is used to adjust current power generation by each generator. All generators are able to collectively minimize the total cost while satisfying power balance constraints through the sparse communication channels. Comparisons with centralized Lambda-Iteration method as well as distributed ED method "consensus + innovation" method are also carried out in the application examples. These comparisons also demonstrate that our technique requires less restrictive communication networks and provides better convergence results.

In Chapter 5, the problem of considering power loss on transmission lines is first introduced into distributed optimization. In this Chapter, a hierarchical consensus based algorithm is proposed to solve economic dispatch problem with loss calculation. The quadratic cost functions are used to model the generators' operating costs. The B matrix is used to model the transmission losses. The upper level algorithm estimates the loss information and distribute to the lower level. The lower level algorithm drives the product of penalty factor and incremental cost to reach an agreement among all agents. Sparse communication networks are adopted to demonstrate the convergence speed of the proposed algorithm. The case studies have shown that the total operating cost can be minimized while maintaining the power balance constraint. The proposed technique can also take care of demand variations and generator variations. In addition, the scalability of our technique is demonstrated by solving the large network problem.

For future work, we notice that the distributed optimization of the ED

problem is a emerging area in power system research. The research work on this field is still at its starting phase. The B matrix may not be an accurate representation of the power flow analysis. Thus, the future work will focus on bringing in more comprehensive power flow analysis and more realistic power system constraints. Furthermore, from our literature studies, we find that load shedding and load recovering can be addressed in distributed manner. We will also consider these problems in our future work.

Bibliography

[1] G. R. Davison, "Dividing load between units," *Electrical World*, 1922.

[2] A. Wilston, "Dividing load economically among power plant," *A.I.E.E. Journal*, 1928.

[3] E. C. M. Stahl, ""load division in interconnections"," *Electrical World*, 1930.

[4] ——, "Economic loading of generating stations," *Electrical Engineering*, 1931.

[5] G. R. Hahn, "Load division by the increment method," *Power*, 1931.

[6] H. Estrada, "Economical load allocation," *Electrical World*, 1930.

[7] M. J. Steinberg and T. H. Smith, "Incremental loading of generating stations," *Electrical Engineering*, 1933.

[8] ——, "The theory of incremental rates," *Electrical Engineering*, 1934.

[9] ——, "Economy loading of power plants and electric systems," *John Wiley & Sons*, 1943.

[10] M. Huneault and F. Galiana, "A survey of the optimal power flow literature," *IEEE Transactions on Power Systems*, vol. 6, no. 2, pp. 762–770, May 1991.

[11] E. E. George, "Intrasystem transmission losses," *Electrical Engineering*, vol. 62, no. 3, pp. 153–158, March 1943.

[12] E. George, H. Page, and J. Ward, "Co-ordination of fuel cost and transmission loss by use of the network analyzer to determine plant loading schedules," *Transactions of the American Institute of Electrical Engineers*, vol. 68, no. 2, pp. 1152–1163, July 1949.

[13] G. Kron, "Tensorial analysis of integrated transmission systems part i. the six basic reference frames," *Transactions of the American Institute of Electrical Engineers*, vol. 70, no. 2, pp. 1239–1248, 1951.

[14] ——, "Tensorial analysis of integrated transmission systems; part ii. off-nominal turn ratios [includes discussion]," *Transactions of the American Institute of Electrical Engineers Power Apparatus and Systems, Part III*, vol. 71, no. 1, 1952.

[15] ——, "Tensorial analysis of integrated transmission systems; part iii. the "primitive" division [includes discussion]," *Transactions of the American Institute of Electrical Engineers Power Apparatus and Systems, Part III*, vol. 71, no. 1, pp. –, Jan 1952.

[16] ——, "Tensorial analysis of integrated transmission systems; part iv. the interconnection of transmission systems [includes discussion]," *Transactions of the American Institute of Electrical Engineers Power Apparatus and Systems, Part III*, vol. 72, no. 2, pp. –, Jan 1953.

[17] L. Kirchmayer and G. Stagg, "Evaluation of methods of co-ordinating incremental fuel costs and incremental transmission losses [includes discussion]," *Transactions of the American Institute of Electrical Engineers Power Apparatus and Systems, Part III*, vol. 71, no. 1, Jan 1952.

[18] A. Glimn, L. Kirchmayer, R. Habermann, and R. Thomas, "Automatic digital computer applied to generation scheduling [includes discussion]," *Transactions of the American Institute of Electrical Engineers Power Apparatus and Systems, Part III*, vol. 73, no. 2, Jan 1954.

[19] J. B. Ward, "Economy loading simplified," *Electrical Engineering*, vol. 73, no. 1, pp. 38–38, Jan 1954.

[20] J. Tudor and W. Lewis, "Transmission losses and economy loading by the use of admittance constants," *IEEE Transactions on Power Apparatus and Systems*, vol. 82, no. 68, pp. 676–683, Oct 1963.

[21] J. E. Van Ness, "A note on incremental loss computation," *Transactions of the American Institute of Electrical Engineers Power Apparatus and Systems, Part III*, vol. 81, no. 3, pp. 735–737, April 1962.

[22] H. Happ, "Optimal power dispatch," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-93, no. 3, pp. 820–830, May 1974.

[23] R. Shoults, "A simplified economic dispatch flow algorithm using decoupled network," *IEEE PES Summer Meeting*, 1977.

[24] B. F. Wollenber and W. Stadlin, "A real time optimum for security dispatch," *IEEE Transactions on PAS*, vol. 93, pp. 1640–1649, May 1974.

[25] D. Deo, "Economic load dispatch of interconnected power system by linear programming," *J. Inst. Engr. (India)*, vol. 53, pp. 117–122, 1973.

[26] R. Ringlee and D. Williams, "Economic system operation considering valve throttling losses ii-distribution of system loads by the method of dynamic programming," *Transactions of the American Institute of Electrical Engineers Power Apparatus and Systems, Part III*, vol. 81, no. 3, pp. 615–620, April 1962.

[27] J. Carpentier, "Contribution a l¡¯etude du dispatching economique," *Francaise Electricians*, vol. 8, pp. 431–447, Aug 1962.

[28] J. Carpentier, C. Cassapoglou, and C. Hensgen, "Differential injections, a method for general resolution of problems of economic dispatching without complete variables using the generalizes reduced gradient method," *Proc. of Conf. by Hellenic Operational Res. Soc.*, pp. 4–8, Nov 1968.

[29] J. Carpentier, "Results and extensions of the methods of differential injections," *Proc. of 4th PSCC Conf.*, vol. 2, pp. 1–8, 1972.

[30] ——, "System security in the differential injections method for optimal load flows," *Proc. 03 PSCC Conf.*, 1975.

[31] R. Andreani, J. Martinez, and M. Schuverdt, "On the relation between constant positive linear dependence condition and quasinormality constraint qualification," *Journal of Optimization Theory and Applications*, vol. 125, no. 2, pp. 473–483, 2005.

[32] J. Peschon, H. Dommel, W. Powell, and D. Bree, "Optimum power flow for systems with area interchange controls," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-91, no. 3, pp. 898–905, May 1972.

[33] H. Dommel and W. Tinney, "Optimal power flow solutions," *Power Apparatus and Systems, IEEE Transactions on*, vol. PAS-87, no. 10, pp. 1866–1876, Oct 1968.

[34] F. F.Wu, G. Gross, J. Luini, and P. M. Look, "A two-stage approach to solving large-scale optimal power flows," in *IEEE Conference Proceedings Power Industry Computer Applications Conference, 1979. PICA-79*, May 1979, pp. 126–136.

[35] R. Burchett, H. Happ, D. Vierath, and K. Wirgau, "Developments in optimal power flow," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-101, no. 2, pp. 406–414, Feb 1982.

[36] M. Gent and J. Lamont, "Minimum-emission dispatch," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-90, no. 6, pp. 2650–2660, Nov 1971.

[37] J. K. Delson, "Controlled emission dispatch," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-93, no. 5, pp. 1359–1366, Sept 1974.

[38] C. Shen and M. Laughton, "Determination of optimum power-system operating conditions under constraints," *Proceedings of the Institution of Electrical Engineers*, vol. 116, no. 2, pp. 225–239, February 1969.

[39] A. H. El-Abiad and F. Jaimes, "A method for optimum scheduling of power and voltage magnitude," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-88, no. 4, pp. 413–422, April 1969.

[40] N. Nabona and L. Freris, "Optimisation of economic dispatch through quadratic and linear programming," *Proceedings of the Institution of Electrical Engineers*, vol. 120, no. 5, pp. 574–580, May 1973.

[41] H. Nicholson and M. Sterling, "Optimum dispatch of active and reactive generation by quadratic programming," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-92, no. 2, pp. 644–654, March 1973.

[42] T. Dillon, K. Morsztyn, and T. Tun, "Sensitivity analysis of the problem of economic dispatch of thermal power systems," *International Journal of Control*, vol. 22, no. 2, pp. 229–248, 1975.

[43] T. Dillon, "Rescheduling, constrained participation factors and parameter sensitivity in the optimal power flow problem," *IEEE Transactions on Power Apparatus and Systems*, no. 5, pp. 2628–2634, 1981.

[44] M. Biggs and M. Laughton, "Optimal electric power scheduling: A large nonlinear programming test problem solved by recursive quadratic programming," *Mathematical Programming*, vol. 13, no. 1, pp. 167–182, 1977.

[45] J. Lipowski and C. Charalambous, "Solution of optimal load flow problem by modified recursive quadratic-programming method," in *IEE Proceedings C (Generation, Transmission and Distribution)*, vol. 128, no. 5. IET, 1981, pp. 288–294.

[46] B. Cova, G. Granelli, M. Montagna, A. Silvestri, M. Innorta, and P. Marannino, "Large-scale application of the han-powell algorithm to compact models of static and dynamic dispatch of real power," *International Journal of Electrical Power & Energy Systems*, vol. 9, no. 3, pp. 130 – 141, 1987. [Online]. Available: http://www.sciencedirect.com/science/article/pii/014206158790010X

[47] T. Giras and S. Talukdar, "Quasi-newton method for optimal power flows," *International Journal of Electrical Power & Energy Systems*, vol. 3, no. 2, pp. 59 – 64, 1981.

[48] S. N. Talukdar and F. Wu, "Computer-aided dispatch for electric power systems," *Proceedings of the IEEE*, vol. 69, no. 10, pp. 1212–1231, Oct 1981.

[49] R. Burchett, H. Happ, and K. Wirgau, "Large scale optimal power flow," *IEEE Transactions on Power Apparatus and Systems*, no. 10, pp. 3722–3732, 1982.

[50] D. I. Sun, B. Ashley, B. Brewer, A. Hughes, and W. F. Tinney, "Optimal power flow by newton approach," *IEEE transactions on power apparatus and systems*, no. 10, pp. 2864–2880, 1984.

[51] A. Monticelli, M. Pereira, and S. Granville, "Security-constrained optimal power flow with post-contingency corrective rescheduling," *IEEE Transactions on Power Systems*, vol. 2, no. 1, pp. 175–180, 1987.

[52] G. Schnyder and H. Glavitsch, "Integrated security control using an optimal power flow and switching concepts," *IEEE Transactions on Power Systems*, vol. 3, no. 2, pp. 782–790, 1988.

[53] ——, "Security enhancement using an optimal switching power flow," *IEEE Transactions on Power Systems*, vol. 5, no. 2, pp. 674–681, May 1990.

[54] H.-T. Yang, P.-C. Yang, and C.-L. Huang, "Evolutionary programming based economic dispatch for units with non-smooth fuel cost functions," *IEEE Transactions on Power Systems*, vol. 11, no. 1, pp. 112–118, 1996.

[55] K. Wong and C. Fung, "Simulated annealing based economic dispatch algorithm," in *IEE Proceedings C (Generation, Transmission and Distribution)*, vol. 140, no. 6. IET, 1993, pp. 509–515.

[56] D. C. Walters and G. B. Sheble, "Genetic algorithm solution of economic dispatch with valve point loading," *IEEE Transactions on Power Systems*, vol. 8, no. 3, pp. 1325–1332, 1993.

[57] 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.

[58] K. P. Wong, B. Fan, C. Chang, and A. C. Liew, "Multi-objective generation dispatch using bi-criterion global optimisation," *IEEE Transactions on Power Systems*, vol. 10, no. 4, pp. 1813–1819, 1995.

[59] D. B. Das and C. Patvardhan, "New multi-objective stochastic search technique for economic load dispatch," in *IEE Proceedings-Generation, Transmission and Distribution*, vol. 145, no. 6. IET, 1998, pp. 747–752.

[60] K. Chandrasekaran and S. P. Simon, "Optimal deviation based firefly algorithm tuned fuzzy design for multi-objective ucp," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 460–471, 2013.

[61] A. M. Jubril, O. A. Komolafe, and K. O. Alawode, "Solving multi-objective economic dispatch problem via semidefinite programming," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2056–2064, 2013.

[62] J. A. Carr, J. C. Balda, and H. A. Mantooth, "A survey of systems to integrate distributed energy resources and energy storage on the utility grid," in *IEEE Energy 2030 Conference, 2008.* IEEE, 2008, pp. 1–7.

[63] A. Abo-Khalil and D.-C. Lee, "Dynamic modeling and control of wind turbines for grid-connected wind generation system," in *37th IEEE Power Electronics Specialists Conference*, June 2006, pp. 1–6.

[64] C. Hernandez-Aramburo and T. Green, "Fuel consumption minimisation of a microgrid," in *IEEE Industry Applications Conference, 2004. 39th IAS Annual Meeting*, vol. 3, Oct 2004, pp. 2063–2068 vol.3.

[65] C. Chen, S. Duan, T. Cai, B. Liu, and G. Hu, "Smart energy management system for optimal microgrid economic operation," *IET Renewable Power Generation*, vol. 5, no. 3, pp. 258–267, May 2011.

[66] S. Civanlar, J. J. Grainger, H. Yin, and S. S. H. Lee, "Distribution feeder reconfiguration for loss reduction," *IEEE Transactions on Power Delivery*, vol. 3, no. 3, pp. 1217–1223, Jul 1988.

[67] R. Srinivasa Rao, S. V. L. Narasimham, M. Ramalinga Raju, and A. Srinivasa Rao, "Optimal network reconfiguration of large-scale distribution system using harmony search algorithm," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1080–1088, Aug 2011.

[68] Y.-C. Huang, "Enhanced genetic algorithm-based fuzzy multi-objective approach to distribution network reconfiguration," *IEE Proceedings-Generation, Transmission and Distribution*, vol. 149, no. 5, pp. 615–620, Sep 2002.

[69] W.-C. Wu and M.-S. Tsai, "Application of enhanced integer coded particle swarm optimization for distribution system feeder reconfiguration," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1591–1599, Aug 2011.

[70] P. Gupta, "A stochastic approach to peak power-demand forecasting in electric utility systems," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-90, no. 2, pp. 824–832, March 1971.

[71] J. Taylor and R. Buizza, "Neural network load forecasting with weather ensemble predictions," *IEEE Transactions on Power Systems*, vol. 17, no. 3, pp. 626–632, Aug 2002.

[72] E. Elattar, J. Goulermas, and Q. Wu, "Electric load forecasting based on locally weighted support vector regression," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 40, no. 4, pp. 438–447, July 2010.

[73] D. Srinivasan, C. Chang, and A. Liew, "Demand forecasting using fuzzy neural computation, with special emphasis on weekend and public holiday forecasting," *IEEE Transactions on Power Systems*, vol. 10, no. 4, pp. 1897–1903, 1995.

[74] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, 2004.

[75] 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 Powers Systems*, vol. 27, no. 4, pp. 1761–1768, 2012.

[76] A. D. Dominguez-Garcia, S. T. Cady, and C. N. Hadjicostis, "Decentralized optimal dispatch of distributed energy resources," in *Proceedings of IEEE Conference on Decision and Control*, Maui, Hawaii, USA, 10-13 December 2012, pp. 3688–3693.

[77] M. Madrigal and V. H. Quintana, "An analytical solution to the economic dispatch problem," *IEEE Power Engineering Review*, vol. 20, no. 9, pp. 52–55, 2000.

[78] S. Kar and G. Hug, "Distributed robust economic dispatch in power systems: A consensus + innovations approach," in *IEEE Power and Energy Society General Meeting*, 22-26 July 2012, pp. 1–8.

[79] L. Du, S. Grijalva, and R. Harley, "Game-theoretic formulation of power dispatch with guaranteed convergence and prioritized best response," *IEEE Transactions on Sustainable Energy*, vol. PP, no. 99, pp. 1–9, 2014.

[80] A. Al-Awami, E. Sortomme, and M. El-Sharkawi, "Optimizing economic/environmental dispatch with wind and thermal units," in *IEEE Power Energy Society General Meeting*, July 2009, pp. 1–6.

[81] J. Hetzer, D. Yu, and K. Bhattarai, "An economic dispatch model incorporating wind power," *IEEE Transactions on Energy Conversion*, vol. 23, no. 2, pp. 603–611, June 2008.

[82] M. El-Sharkh, M. Tanrioven, A. Rahman, and M. Alam, "Cost related sensitivity analysis for optimal operation of a grid-parallel {PEM} fuel cell power plant," *Journal of Power Sources*, vol. 161, no. 2, pp. 1198 – 1207, 2006.

[83] L. dos Santos Coelho and V. Mariani, "Combining of chaotic differential evolution and quadratic programming for economic dispatch optimization with valve-point effect," *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 989–996, May 2006.

[84] R. Yokoyama, S. H. Bae, T. Morita, and H. Sasaki, "Multiobjective optimal generation dispatch based on probability security criteria," *IEEE Transactions on Power Systems*, vol. 3, no. 1, pp. 317–324, Feb 1988.

[85] S. Hemamalini and S. Simon, "Maclaurin series-based lagrangian method for economic dispatch with valve-point effect," *IET Generation, Transmission Distribution*, vol. 3, no. 9, pp. 859–871, September 2009.

[86] 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.

[87] K. L. Cheong, P. Li, and J. Xia, "Control oriented modeling and system identification of a diesel generator set (genset)," in *American Control Conference (ACC)*, June 2010, pp. 950–955.

[88] K. E. Yeager and J. R. Willis, "Modeling of emergency diesel generators in an 800 megawatt nuclear power plant," *IEEE Transactions on Energy Conversion*, vol. 8, no. 3, pp. 433–441, Sep 1993.

[89] D. H. Wang, C. V. Nayar, and C. Wang, "Modeling of stand-alone variable speed diesel generator using doubly-fed induction generator," in *2nd IEEE International Symposium on Power Electronics for Distributed Generation Systems (PEDG)*, June 2010, pp. 1–6.

[90] D. McGowan, D. Morrow, and M. McArdle, "A digital pid speed controller for a diesel generating set," in *IEEE Power Engineering Society General Meeting*, vol. 3, July 2003, pp. –1477 Vol. 3.

[91] M. Al-Hajri and M. Abido, "Multiobjective optimal power flow using improved strength pareto evolutionary algorithm (spea2)," in *11th International Conference on Intelligent Systems Design and Applications (ISDA), 2011*, Nov 2011, pp. 1097–1103.

[92] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: Nsga-ii," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, Apr 2002.

[93] S. Larsson and E. Ek, "The black-out in southern sweden and eastern denmark, september 23, 2003," in *IEEE Power Engineering Society General Meeting, 2004*, June 2004, pp. 1668–1672 Vol.2.

[94] M. Baran and F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Transactions on Power Delivery*, vol. 4, no. 2, pp. 1401–1407, Apr 1989.

[95] B. Chen, M. Chang, and C. Lin, "Load forecasting using support vector machines: a study on eunite competition 2001," *IEEE Transactions on Power Systems*, vol. 19, no. 4, pp. 1821–1830, Nov 2004.

[96] R. Billinton and R. Allan, *Reliability Evaluation of Power Systems*. Kluwer Academic Pub, 1984. [Online]. Available: http://books.google.com.sg/books?id=3qHCQgAACAAJ

[97] Z. Salameh, B. Borowy, and A. Amin, "Photovoltaic module-site matching based on the capacity factors," *IEEE Transactions on Energy Conversion*, vol. 10, no. 2, pp. 326–332, Jun 1995.

[98] J. Dhillon, J. Dhillon, and D. Kothari, "Economic-emission load dispatch using binary successive approximation-based evolutionary search," *IET Generation, Transmission Distribution*, vol. 3, no. 1, pp. 1–16, January 2009.

[99] S. A. Kazarlis, A. Bakirtzis, and V. Petridis, "A genetic algorithm solution to the unit commitment problem," *IEEE Transactions on Power Systems*, vol. 11, no. 1, pp. 83–92, 1996.

[100] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *IEEE International Conference on Neural Networks, 1995*, vol. 4, Nov 1995, pp. 1942–1948 vol.4.

[101] K. Woldemariam and G. Yen, "Vaccine-enhanced artificial immune system for multimodal function optimization," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 40, no. 1, pp. 218–228, Feb 2010.

[102] N. K. Jerne, "Toward a network theory of the immune system," *Ann. Immunol.*, vol. 125, pp. 373–389, 1974.

[103] S. Forrest, A. Perelson, L. Allen, and R. Cherukuri, "Self-nonself discrimination in a computer," in *1994 IEEE Computer Society Symposium on Research in Security and Privacy*, May 1994, pp. 202–212.

[104] L. N. De Castro and F. J. Von Zuben, "Learning and optimization using the clonal selection principle," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 3, pp. 239–251, 2002.

[105] A. S. Perelson, "Immune network theory," *Immunological reviews*, vol. 110, no. 1, pp. 5–36, 1989.

[106] "Optimization of power system operation," 2014. [Online]. Available: http://www.responsiblebusiness.com/news/singapore-sustainable-energy-trends-2015/

[107] X. Yu, H. Jia, C. Wang, W. Wei, Y. Zeng, and J. Zhao, "Network reconfiguration for distribution system with micro-grids," in *2009 International Conference on Sustainable Power Generation and Supply*, April 2009, pp. 1–4.

[108] J. Guerrero, J. Vasquez, J. Matas, M. Castilla, and L. de Vicuna, "Control strategy for flexible microgrid based on parallel line-interactive ups systems," *IEEE Transactions on Industrial Electronics,*, vol. 56, no. 3, pp. 726–736, March 2009.

[109] Z. Zhang and M.-Y. Chow, "Incremental cost consensus algorithm in a smart grid environment," in *IEEE Power and Energy Society General metting*, 24-27 July 2011, pp. 1–6.

[110] A. J. Wood and B. F. Wollenberg, *Power generation, operation & control*, 2nd ed. New York: Wiley, 1996.

[111] R. A. Horn and C. R. Johnson, *Matrix Analysis.* Cambridge University Press, 1985.

[112] K. Cai and H. Ishii, "Average consensus on general digraph," in *Proceedings of IEEE Conference on Decision and Control*, Orlando, FL, USA, 12-15 December 2010, pp. 1956–1961.

[113] A. P. Seyranian and A. A. Mailybaev, *Multiparameter Stability Theory with Mechanical Applications*, ser. Series on Stability, Vibration and Control of Systems. World Scientific, 2004.

[114] U. Mackenroth, *Robust Control Systems: Theory and Case Studies.* Springer, 2004.

[115] N. Hashim, N. Hamzah, M. Latip, and A. Sallehhudin, "Transient stability analysis of the ieee 14-bus test system using dynamic computation for power systems (dcps)," in *Third International Conference on Intelligent Systems, Modelling and Simulation (ISMS), 2012*, Feb 2012, pp. 481–486.

[116] M. Hashmi, S. Hanninen, and K. Maki, "Survey of smart grid concepts, architectures, and technological demonstrations worldwide," in *2011 IEEE PES Conference on Innovative Smart Grid Technologies (ISGT Latin America)*, 2011, pp. 1–7.

[117] E. Caro, R. Singh, B. Pal, A. Conejo, and R. Jabr, "An incremental p-mu placement method for power system state estimation," *IET Generation Transmission and Distribution*, vol. 6, pp. 922–929, 2012.

Appendix A

Publication

A.1 Journal Publication

1. Sicong Tan, Jian-Xin Xu and Panda, S.K., "Optimization of Distribution Network Incorporating Distributed Generators: An Integrated Approach," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2421, 2432, Aug. 2013

2. Shiping Yang, Sicong Tan and Jian-Xin Xu, "Consensus Based Approach for Economic Dispatch Problem in a Smart Grid," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4416, 4426, Nov. 2013

A.2 Conference Publication

1. Jianxin Xu, Sicong Tan and Panda, S.K., "Optimization of economic load dispatch for a microgrid using evolutionary computation," *37th Annual Conference on IEEE Industrial Electronics Society*, pp. 3192, 3197, 7-10 Nov. 2011

2. Geramifard, O., Jian-Xin Xu, Tan Sicong, Jun-Hong Zhou and Xiang Li, "A multi-modal hidden Markov model based approach for continuous health assessment in machinery systems," *37th Annual Conference on IEEE Industrial Electronics Society*, pp. 2294, 2299, 7-10 Nov. 2011

3. Sicong Tan, Jianxin Xu and Panda, S.K., "Optimization of distribution network incorporating microgrid using vaccine-AIS," *38th Annual Conference on IEEE Industrial Electronics Society*, pp.1381,1386, 25-28 Oct. 2012

4. Sicong Tan, Shiping Yang and Jian-Xin Xu, "Consensus based approach for economic dispatch problem in a smart grid," *39th Annual Conference of the IEEE Industrial Electronics Society*, pp. 2011, 2015, 10-13 Nov. 2013

A.3 Under Revision

1. S.C. Tan, S.P. Yang, J.X. Xu and S.K. Panda, "Hierarchical Consensus Based Approach with Loss Consideration for Economic Dispatch Problem under Smart Grid Context," *IEEE Transaction on Neural Network.*