# Joint Optimization of System Hardening and Restoration for Resilience Enhancement against Cascading Failures

by

## Jian Zhou

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Industrial and Systems Engineering

Written under the direction of

Dr. David W. Coit

And approved by

_____

_____

_____

_____

_____

New Brunswick, New Jersey

October, 2020

# ABSTRACT OF THE DISSERTATION

This dissertation presents a framework that determines the optimal integrated system hardening and system restoration strategies. The objective is to improve system resilience while minimizing total cost including investment cost and system damage cost in the event of system failure propagation. The successful functioning of modern society is increasingly dependent upon various crucial infrastructure systems, such as power grids and communication systems, which all comprise a large collection of interconnected sub-systems. The reliability and resilience of these network systems becomes a matter of great concern due to the inevitable occurrence of system failures and their probable disastrous aftereffects. Although it is uncommon for many applications, there are still many examples of massive cascading failures in various real-world network systems. In this research, the mechanism of cascading failures in network systems is investigated, taking into account practical network load dynamics as well as multiple dependencies between and inside systems. A new resilience metric which can be used to evaluate system resiliency loss caused by system disruptions is proposed. Then focusing on electricity system, the influence of cascading failures in power generation and transmission system is extended to local power distribution systems to analyze the resilience of the entire electrical power system. System hardening strategies and system restoration strategies are jointly optimized with the consideration of the existing interaction between each other. The effects of installing distributed energy resources to end users in electricity system, for instance solar array and battery storage, as a type of system hardening measure to improve electric power system resilience are investigated. The effectiveness of restoration strategies with different restoration prioritizations on

reducing the influence of cascading failures on resilience is explored. Finally, an approach to relate the improved system resilience to the reduction in economic losses is developed. Optimization methods are proposed to achieve a balance between the investment of resilience enhancement driven actions, such as system hardening planning and restoration decisions, and system damage cost, for example, unsatisfied customer demand cost. As a result, system resilience targets are better integrated into the investment of resilience enhancement measures. The proposed methodology can be utilized as a decision-making tool for future resilient network systems, for example, electric power system, design and restoration. Together, this research is useful to mitigate and rescue the system from the next cascading failures with the application of effectively integrated system hardening and restoration strategies for resilience enhancement with minimized total cost.

**Acknowledgements**

First I want to thank my family for their endless love to raise me up and their support whenever I need. Without them I could not experience the world and travel such far. My father, Weimin Zhou, is very outgoing, and he always told me that if you don t give up you will make it. Probably that helps me smooth away most of the difficulties that I have ever met. He has a strong sense of justice and he taught me to be a man with courage which impacts me a lot. My mother, Chunfang Wang, has sacrificed all herself for the family and I will never be able to pay her back. She provides me whatever I need as I grow up. She is humble and hard-working, and I am unconsciously influenced. I think they are proud that I am on the way to become a faculty and we will celebrate together when the time comes. I hope I could support them as if they supported me in the last thirty-two years although what they did for me is way more.

I would like to thank my wife, Wenshan Sun. She stands with our long-distance relationship in the last two years and I am grateful that she always stands by me. She is a successful medical doctor, a real doctor who helps others. I am very happy that she also obtained her PhD. degree almost the same time with me. She sets a high standard for me from both personal and academic perspectives, and I will support you whenever you need. Thank you. My love!

I also want to thank my friends that always have my back during these years overseas. Sanling Song, the first friend I made when I came to America. I could not image my beginning days in America that without her helps. Qing Li, as a Chinese student arrived at American yeas before me, both of them help me a lot about understanding the culture difference and getting used to the new environment. I feel warm when I recall the days

that we are sharing the office. Both of them have already obtained their PhD degree, and I wish they all well in their new career. Stam and Nooshin are in the same research group with me, and we have spent four years together. We come from different counties, and each of them has a very long history. We become very good friends and I have a chance to learn and understand their culture. I will always remember the days we spent together and no matter where we will be in the future I wish you all have a wonderful life. I made lots of friends at Rutgers, such as Yaoming Cai, Xianghan Xu and Jia Xue, I will remember the days we spent together during these years. Yi Han, Pei Peng and Xin Dong are my roommates in the last three years, I want to thank for their support whenever I need. With all these friends, I had a wonderful time in America and I will always be proud to be a Scarlet Knight.

I would like to thank my advisor, Professor Coit. His depth of knowledge and academic rigorousness are the keys to help me pursue my PhD degree. His outgoing personality always helps me to get used to the different culture. He truly wants to help me, and supports me whenever I need which is the main reason that I could finish my PhD study. I am sure we will always keep in touch.

I received great supports from many professors during the past five years. Professor Felder could be called my co-advisor. I started to work with him since 2016. His academic rigorousness impacts me a lot. By working on his different projects, my research skills get improved and I learned to think about my research from practical application and economic perspective. I took two reliability courses given by Professor Pham, which build up the foundation of my reliability research. He also gives me a lot of advice about career planning which is very important to me. Professor Xi provides his

# Table of Contents

## List of Figures

## List of Tables

## List of Notations

**Chapter 3**

| | | |
|---|---|---|
| $Net(N, M)$ | = | A network system with $N$ nodes and $M$ edges; |
| $v_i$ | = | System node $i$; |
| $e_i$ | = | System edge $i$; |
| $L_i(t)$ | = | The load on node $i$ at time $t$; |
| $\sigma_{sl}(t)$ | = | The number of shortest paths between nodes $s$ and $l$ at time $t$ in a network system; |
| $\sigma_{sl}(i,t)$ | = | The number of shortest paths between nodes $s$ and $l$ passing through node $i$ at time $t$; |
| $C_i$ | = | The capacity of network node $i$; |
| $\alpha$ | = | The tolerance parameter that decides system node capacity; |
| $\mu$ | = | The nonlinear coefficient regarding system node capacity; |
| $Dep(v_i, l)$ | = | Indicator function denotes whether network node $v_i$ fails given dependence cluster $l$ collapses; |
| $LCC$ | = | The largest connected component of the network; |
| $T$ | = | The duration of cascading failures in the network system until it reaches the predetermined stop criteria; |
| $G$ | = | The parameter indicates network connectivity; |
| $N_I$ | = | The initial size of $LCC$ before cascading failures start; |

$N_F$ = The final size of *LCC* of the network when cascading failures stop;

$P(D)$ = The probability that a node belongs to a dependence cluster of size $D$;

*D-size* = The mean size of dependence clusters in the network system;

$\lambda$ = *D-size* - 1

$K$ = The average degree of network nodes in a network system;

*Nrem* = The number of network nodes initially failed at the beginning of the cascading failures;

*NremC* = The critical number of initially removed nodes which cause network collapse;

$T_c$ = The peak of the changing trend of *T* corresponding to *NremC*;

$b$ = A normalization constant in normal distribution of dependence clusters which makes the summation of all $P(D)$ equals 1;

$\sigma^2/Qvar^2$ = Variance of dependence cluster size which follows normal distribution;

ER network = A random network model proposed by Erdos and Renyi;

$L_{i,c}(t)$ = Local customer load on network node *i* at time *t*;

$L_{i,d}(t)$ = Passing by load on network node *i* at time *t*;

$P_i^c$ = The normalized number of customers connected to node *i*;

$P_i^d$ = The normalized degree of node *i*;

| | | |
|---|---|---|
| $L_{\min}$ | = | The minimum initial passing by load or local customer load; |
| $L_{\max}$ | = | The maximum initial passing by load or local customer load; |
| $\mu_l$ | = | The mean of the truncated normal distribution that is used to describe initial passing by load and local customer load; |
| $\sigma_l^2$ | = | The variance of the truncated normal distribution that is used to describe initial passing by load and local customer load; |
| $\Delta L_{i,j,c}(t)$ | = | The local customer load of failed node $i$ that is equally distributed to its working dependent node $j$ at time $t$; |
| $k_{i,c}(t)$ | = | The set of functional dependent nodes of node $i$ at time $t$; |
| $\Delta L_{i,j,d}(t)$ | = | The passing by load of failed node $i$ that are equally distributed to its working neighbor node $j$ at time $t$; |
| $k_{i,d}(t)$ | = | The set of functional neighbor nodes of node $i$ at time $t$; |
| $NF_c$ | = | The set of nodes that received local customer load from a failed node and can transfer the load fully back to the failed node when it is restored; |
| $ND_c$ | = | The set of nodes that received local customer load from a failed node and transfer a random amount of load back to the failed node when it is restored; |
| $NF_d$ | = | The set of nodes that received passing by load from a failed node and can transfer the load fully back to the failed node when it is restored; |
| $ND_d$ | = | The set of nodes that received passing by load from a failed node and |

transfer a random amount of load back to the failed node when it is restored;

$p_m$ = A variable that decides the amount load transferred back to the recovered node from node $m$;

**Chapter 4**

$ed_{ij}$ = Topological connection between system node $i$ and node $j$;

$E_{ij}$ = Efficiency of the shortest path between node $i$ and node $j$;

$E(G_P)$ = Transmission efficiency of network system $G_p$;

$G$ = The parameter indicates network connectivity;

$T$ = Time to restore a network system from cascading failures to the predetermined level;

$N$ = Number of nodes in a network system;

$CCT$ = Dependence cluster collapsing threshold;

$Q(t)$ = The measurement of a system quality/performance at time $t$;

$t_I$ = A disruptive event occurs in network system at time $t_I$;

$t_e$ = The time when system performance/quality $Q(t)$ is recovered to a predetermined level;

RR = Random repair;

HDFR = High-degree first repair;

STFR = Short time first repair;

| | | |
|---|---|---|
| HLFR | = | High-load first repair; |
| $R_p$ | = | The proportion of failed components, whose repairing has not yet begun, that can start to be repaired; |
| *Nrem* | = | The number of network nodes initially failed at the beginning of the cascading failures; |
| *D-size* | = | The mean size of dependence clusters in the network system; |
| *K* | = | The average degree of network nodes in a network system; |
| BA network | = | A scale-free network model proposed by Barabási and Albert; |
| MVC | = | Minimum vertex cover; |
| RL | = | Reinforcement learning; |
| *u* | = | The percentage of top ranked nodes in MVC which is assigned with a high restoration priority; |

**Chapter 5**

| | | |
|---|---|---|
| PV array | = | Photovoltaic array; |
| DER | = | Distributed energy resources; |
| *STC* | = | System total cost; |
| GHI | = | Global horizontal irradiance; |
| *LOLP* | = | Loss of load probability; |
| *CCP* | = | Chance constraint probability; |

$B_r$        =    Battery capacity (kWh);

$P_b$        =    Battery price ($/kWh);

$VOLL$     =    Value of lost load ($/kWh);

$T_y$        =    The planning time horizon in years;

$r$         =    The discount rate;

$N_k$        =    The random set of outages in year $k$;

$O_{kj}$       =    The random set of time intervals for outage $j$ in year $k$;

$t_{kji}$        =    The $i$th time interval for the $j$th outage in year $k$;

$\Delta t$        =    Time increment, which is 1 hour;

$AEL(t_{kji})$    =    The amount of load demand which is not satisfied by the generation of the PV + battery system at time $t_{kji}$;

$B_{min}$      =    The minimum battery energy value (kWh);

$Q_B(t_{kji})$    =    The energy stored in battery (kWh) at time $t_{kji}$;

$P(t_{kji})$     =    The power generated by PV array (kW) at time $t_{kji}$;

$D(t_{kji})$     =    Load demand (kW) at time $t_{kji}$;

$e$         =    The efficiency of discharging/charging process of battery;

$\eta$         =    The energy conversion efficiency of PV array;

$I(t_{kji})$     =    Solar irradiation (W/m$^2$) at time $t_{kji}$;

$A$         =    PV array area (m$^2$);

| | | |
|---|---|---|
| $DoD$ | = | The maximum depth of discharge of a battery (%); |
| $\beta$ | = | The upper bound for the expected value, E[$LOLP$]; |
| $\tau$ | = | 1 - $\tau$ denotes the minimum desired $CCP$; |
| $\gamma$ | = | The constraint for the $LOLP$ of each grid outage in terms of $CCP$; |
| SAIDI | = | The System Average Interruption Duration Index; |
| CAIDI | = | The Customer Average Interruption Duration Index; |
| $\lambda_s$ | = | Indicator for the achieved $LOLP$ of the $s$th simulated outage; |
| $LOLPs$ | = | Loss of load probability of the $s$th simulated grid outage; |
| $C_s$ | = | Cost of lost load corresponding to the $s$th simulated grid outage; |
| $t_{is}$ | = | The $i$th time interval of the $s$th simulated grid outage; |
| $G_s$ | = | Set of time intervals for the $s$th simulated grid outage; |
| $\delta(t)$ | = | An indicator function for whether the load demand at time $t$ is lost or not, binary; |
| $\zeta$ | = | Constraint limiting $CCP$; |
| $T_{jk}$ | = | Duration of $j$th grid outage of year $k$, hrs |
| $C_p$ | = | Proportion of critical load demand during an outage; |
| $I_c$ | = | Initial charging state of battery when grid outage occurs; |
| $S$ | = | Total number of simulations; |

**Chapter 6**

| | | |
|---|---|---|
| $C_i^{in}$ | = | Hardening investment cost for scenario $i$; |

| | | |
|---|---|---|
| $\varphi$ | = | A specific system hardening strategy; |
| $\omega$ | = | The percentage of top ranked nodes that are considered for hardening; |
| $\boldsymbol{\rho}(\omega)$ | = | The vector that contains the hardening level for each selected node; |
| $\Phi$ | = | The set of available system hardening measures; |
| $u_i$ | = | A specific system disruption severity for a scenario; |
| $\rho_d$ | = | The hardening level of node $d$; |
| $K_c$ | = | Total types of customers are considered in the electricity system; |
| $n_{dk}$ | = | The number of type $k$ customers connected to node $d$; |
| $O_{\varphi\omega i}$ | = | The set of top ranked nodes that are considered for hardening in scenario $i$ basd on the hardening strategy $\varphi$ and the hardening scope $\omega$; |
| $C_k^{ino}$ | = | The cost of accessing the information of type $k$ customer, [\$/cust.]; |
| $C_k^r$ | = | Restoration cost for scenario $i$; |
| $\pi$ | = | A restoration strategy; |
| $c$ | = | The number of hired crews; |
| $C_f^r$ | = | Cost per unit time for hiring the $f$th crew to do repairing; |
| $l_{if}$ | = | Total repairing time performed by the $f$th crew in scenario $i$; |
| $C_b^r$ | = | Cost per unit time for hiring a regular crew to do repairing; |

$C_v^r$ = The increment of hiring cost per unit time as the number of extra hired crews increases;

$c_r$ = The number of regular crews;

$f$ = The sequence number of a crew;

$\Pi$ = The set of available system restoration measures;

$C_i^{los}$ = Economic loss in scenario;

$C_{ijq}^{los}(g_{ijq})$ = Economic loss due to the $q$th failure of system node $j$ with the failure time $g_{ijq}$ in scenario $i$;

$UF_i$ = The set of nodes that failed in scenario $i$;

$Q_{ij}$ = The number of times that node $j$ failed in scenario $i$.

$C_k^f$ = The fixed penalty cost when a type $k$ customer is offline/disconnected from the electric power system, [$/cust.];

$n_{ijk}$ = The number of type $k$ customers connected to node $j$ in scenario $i$.

$C_k^v$ = The variable penalty cost per unit time associated with a type $k$ customer when it is offline, [$/$\Delta t$ cust.];

$\beta$ = The parameter that decides the increase rate of variable penalty cost caused by disconnected customers;

$C_i^{tot}$ = The total cost for scenario $i$ with disturbance severity $u_i$;

$v_i$ = The constraint for *CCP* in scenario $i$;

| | | |
|---|---|---|
| $\theta_i$ | = | Resilience loss constraint for scenario $i$; |
| $\lambda_i$ | = | KKT multiplier for scenario $i$; |
| $FS$ | = | The set of available hardening and restoration measures that make the $CCP$ constraints are satisfied; |
| $I$ | = | The number of different scenarios are considered; |
| $\delta_{im}$ | = | Indicator function tha whether the resiliency loss of the $m$th simulation in scenario $i$ satisfies the constraint; |
| $Q_{im}(t)$ | = | The system performance at time $t$ during the $m$th simulation in scenario $i$; |
| $t_s$ | = | The time when system disruption occurs; |
| $\Re_{im}(T_{im}(1-\tau))$ | = | System resiliency loss up to time $T_{im}$ when system performance is recovered to $1 - \tau$ of its original level during the $m$th simulation for scenario $i$; |
| $T_{im}(1-\tau)$ | = | The time period from the beginning of system interruption till a specific performance is recovered to be a predetermined level, $1 - \tau$ percentage of its initial value; |
| $H$ | = | The predetermined constraint for recovery time $T$ |
| $C_{rw}$ | = | Hard constraint on the maximum number of crews that can be hired for restoration; |
| $g_{imjq}$ | = | The failure time of node $j$ in its $q$th failure during the $m$th simulation of scenario $i$; |

| | | |
|---|---|---|
| $B_d$ | = | The budget constraint for hardening investment; |
| $P_d^a$ | = | Actual probability that node $i$ is selected as initial failure resulting from system disruptions; |
| $P_d^o$ | = | Original probability of node $i$ to be initially failed. |
| $T_i^{ar}$ | = | Actual repair time of node $i$; |
| $T_i^{or}$ | = | Original repair time of node $i$; |
| $e_f$ | = | The decrease rate of crew working efficiency as extra crews are hired; |
| $\Delta t$ | = | The unit simulation time that represents 1 hour; |
| FFFR | = | First fail first repair; |
| NOCR | = | High number of customer-connected nodes first repair; |
| FCR | = | Fastest customer repair; |
| NOCH | = | High number of customer-connected nodes first hardening; |
| HDH | = | High-degree nodes first hardening; |
| HLH | = | High-load nodes first hardening; |
| $MS$ | = | The number of simulations conducted for each scenario with a disruption severity. |

# 1. Introduction

This dissertation presents a framework to jointly optimize system hardening strategies and system restoration strategies to enhance system resilience against cascading failures throughout the system. This is accomplished by taking account of the mixed impacts of network load dynamics and system dependency, with the goal of minimizing total cost which include investment cost and economic loss from system failures.

The functioning of society is more and more dependent on large-scale network systems, such as power grids, water/gas supply networks and transportation networks, to deliver services to consumers. For example, the quality of our daily life largely depends on the continuous availability of the electricity supply. However, because of the increasingly complicated functional and structural dependencies existing within and between network systems, even an initially small part of system failure can cause large scale outages on these real-life network systems via network load redistribution. These failure propagations through diverse types of network systems are named cascading failures, which can be disastrous, affecting hundreds of millions of people with inestimable costs [1-3]. For instance, the great Northeast blackout of 2003 was initiated by several broken power transmission lines in Ohio. It eventually caused a widespread power outage in eight U.S. states and the Canadian province of Ontario, affecting 50 million people [4]. Figure 1.1 indicates the blackout affected area in North America.

Figure 1.1. The 2003 grid outage affected area in the United States and Canada

As a result, increasing system robustness against disruptive events, for instance, natural disasters, and effective response to system failures in order to mitigate the resulting damage of the system are of great significance. In other words, the central question is to find the optimal system hardening strategy and system restoration strategy to enhance system resilience in order to reduce potential system damages caused by cascading failures. Since various investments are involved in system hardening actions and system restoration in practice, the problem translates into a search of an optimal balance between investment cost and economic loss from system damage.

Existing cascading failure research shows that both local load redistribution and global load redistribution after system failure occur have been applied, but both of them can exist in real cases considering the influences of multiple system dependencies. Thus, further work should focus on the mixed network load dynamics and system dependency

in order to better understand the failure propagation mechanism. To be specific, previous research of cascading failures in electric power systems only considers power generation and transmission system while resilience is defined and analyzed for power distribution systems. Therefore, the impacts of cascading failures in the bulk power system should be extended to local power distribution systems and end users in order to enhance resilience of the entire electrical power system. Problems that are related to system hardening actions and system restoration have been widely studied. However, most of the work done neglects the underlying relationship between the effects of system hardening actions and system restoration, and optimizes system hardening strategy and system restoration strategy separately. It is necessary to examine the effects between system hardening and restoration in order to effectively enhance system resilience. Furthermore, there is a practical need to develop an approach to relate improved system resilience to reduction in economic losses in order to optimize the balance between investment of resilience requirement driven actions, such as system hardening planning and restoration decisions, and constrained system damage cost.

Above all, this research mainly focuses on the joint optimization of system hardening and restoration strategies that can effectively improve system resilience against cascading failures by taking account of the underlying relationship between system hardening and restoration actions and practical network load redistribution, as well as multiple system dependencies. Moreover, the optimization framework is able to minimize the total involved cost, which includes investment cost of system hardening and restoration, as well as penalty cost of system performance loss. In this way, the proposed joint optimal system hardening and restoration strategies would contribute to the goals of enhancing

system resilience regarding cascading failures, reducing investment cost and loss of system performance cost, as well as incorporating renewable energy sources into electricity system as a bonus.

## 1.1 Problem statement

The main objective of this research is to jointly optimize system hardening and restoration strategies for system resilience enhancement regarding cascading failures while minimizing system total cost. System total cost which includes investment cost, such as information cost for restoration, recruiting cost of repair crews and distributed energy resources (DER) installation cost for end users, as well as system damage cost, for example, cost of unsatisfied demand because of grid outage. System resilience is measured by a proposed resilience metric for interdependent network systems subjected to mixed cascading failures taking into account the impact of multiple system dependencies and network load dynamics. The influence of cascading failures in power generation and transmission systems is extended to power distribution systems in order to be in accordance with real cases. Except for traditional system hardening measures, DER installation for end users in local power distribution systems is applied to provide power for customers during grid outages in order to improve the entire electricity system resilience. Sustainable and renewable energy are incorporated into the electrical power system in this way.

Figure 1.2 illustrates the framework of this research, which mainly includes five phases: 1) Previous work extension, 2) Cascading failure modeling with system dependency, 3) System resilience-based optimization of restoration, 4) System hardening for resilience enhancement and economic benefits, 5) Validation the integral optimization

of system hardening and restoration on interdependent network systems. The relationship between each part of this research is also illustrated in Figure 1.2.
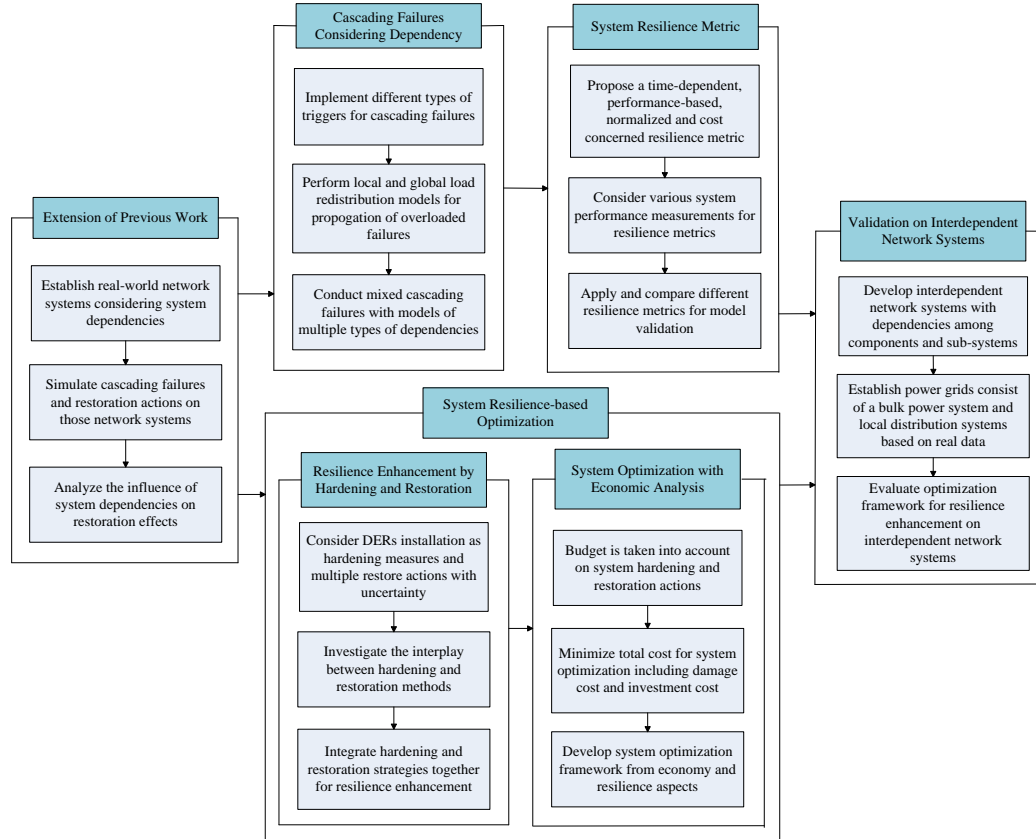


Figure 1.2. The diagram of research components

Extension of my previous work on real-world network systems modeling supports further cascading failure modeling considering multiple system dependencies. Then, the new metric of system resilience regarding cascading failures in network systems can be proposed. Analysis of the influence of system dependency on restoration effects helps us to understand the mechanism of restoration against system failures, which starts the work of resilience enhancement by optimizing system hardening and restoration. Then, the integrated optimization of system hardening and system restoration with minimized total cost can be conducted. Finally, the proposed methodology is validated by testing on

interdependent network systems, for instance, the electric power system. More specifically, the five research phases are as follows:

**Phase I: Extension of previous work**

The first phase of this research is to establish network systems considering system dependency. Mixed cascading failures are simulated in these network systems, and then apply common and typical restoration strategies to examine the difference between restoration effects. Details are provided in Section 3. Eventually, the joint influence of system dependency and restoration strength on the effectiveness of different restoration polices is explored.

**Phase II: Cascading failure modeling with load dynamics and system dependencies**

The second phase in this work is to develop cascading failure models that combine local and global load dynamics for overloaded failure propagations in network systems. Multiple system dependencies and their different impacts are also taken into account in the process of cascading failures. A cascading failure model that integrates the impacts of cascading failures in bulk power system on local customers is developed. In addition, different system disturbances for cascading failures are taken into consideration.

**Phase III: System resilience-based optimization of restoration**

The third phase is to propose a time-dependent, performance-based, normalized and cost concerned resilience metric based on real cases. Different system performance/quality measurements can be adopted for resilience measurement based on the new metric. As a consequence, the resulting resilience-based optimal solutions and findings are more general and more practical. The effectiveness of different restoration

strategies with distinct restoration prioritizations on mitigating resilience loss result from cascading failures are explored.

**Phase IV: Resilience enhancement with backup energy system**

The task is conducted considering the installation of DER to end users in electricity system as a hardening measure. Taking solar array and battery storage as an example of DER, the grid-outage resilience and economic benefits result from DER investment are analyzed. The trade-off between DER investments as a pre-disruption hardening planning and the economic loss of unsatisfied customer load is explored. In addition, the impacts of economic trends of important factors on DER investment optimization are also investigated.

**Phase V: Integral optimization of system hardening and restoration**

New approaches for integrated optimization of hardening and restoration measures to enhance system resilience against cascading failures are proposed. Since a cost budget is applied to system hardening and restoration investment in real cases, the objective is to minimize system total cost, including investment cost and system damage cost. The interplay between the effectiveness of hardening strategies and restoration strategies and the influence of different system disruption is considered. Overall, the task is to develop a system optimization framework from economic and resilience perspectives.

An electric power system model which consists of a bulk power system and several distribution systems is developed. Multiple dependencies exist among system components, and between the bulk system and distribution systems in the electricity system model. The proposed optimization framework regarding system hardening and restoration is applied on the interdependent electric power system for validation.

**1.2 Motivation of research**

Real-world network systems, such as power grids and transportation networks, are critical to people's daily lives and modern economies. However, the increasing system scale and complex dependencies inside of these network systems contribute to the spread of failures within these systems and severe damage of system performance as a result. We have experienced massive cascading failures in different systems all over the world, for instance, high-impact power grid outages, although lots of efforts have been performed to prevent failure propagation throughout the systems.

Thus, investigating the mechanism of cascading failures and understanding the underlying relationship between the effects of system hardening actions and system restoration actions, in response to system catastrophic breakdown, is crucial. In addition, finding approaches to jointly minimize the investment cost of system hardening and restoration measures, which are designed for system resilience enhancement. These measures are implemented to reduce the penalty cost of system performance loss, which is attractive to researchers and practitioners. Overall, optimal integration of system hardening and restoration strategies together, in order to improve system resilience with minimized total cost, are of great significance from both system resilience and economic perspectives. Therefore, this research can be useful for practical applications.

**1.3 Research contributions**

In this research, a general modeling and optimization framework is developed that can be customized or tailored for many different kinds of network system resilience enhancement problems, while the example and case studies mainly focus on electric power system. The scope of this work is to establish an approach to jointly optimize

system hardening and restoration actions to enhance system resilience considering system dependency with minimized total cost. Total cost includes investment and damage cost caused by cascading failures over the system. There are three major research objectives, which are shown as follows:

1) New cascading failure models are developed taking account of the combined impacts of network load dynamics and network dependency on failure propagation.

2) A simulation-based approach is established to integrate the impacts of cascading failures in bulk power system, the effects of system hardening actions and system restoration on the whole electricity system.

3) New methods to simultaneously optimize system hardening actions and restoration in order to minimize investment cost and system damage cost, and to enhance system resilience are developed.

Overall, this research, taking electricity system as an example of network systems, explores the relationship between the cascading failures in bulk power system, the resulting impacts on the entire electrical power system including end users, and the joint effects of system hardening and restoration measures on electricity system resilience enhancement.

## 1.3.1 Theoretical contributions

Three contributions of this work are related to the problem domain, which is flow network system optimization regarding resilience enhancement and economic viability taking into account system dependency. The contributions are shown as follows:

1) Combining local load redistribution and global load redistribution regarding network load dynamics, as well as system dependency in modeling cascading failures in network systems.

2) Linking system failure propagation, system hardening and restoration actions together via hierarchical system modeling.

3) Developing a new iterative method to search the feasible solutions for minimized total cost incorporating resilience enhancement, and to explore the relationship between decision variables regarding system hardening and restoration strategies. As a result, the optimization method can yield robust candidate solutions.

The first theoretical contribution is the development of cascading failure models considering (1) both local load redistribution and global load redistribution after failure occurs, and (2) multiple system dependencies' influences. There are many cascading failure studies that take into consideration network load dynamics, either load redistributed locally or load redistributed globally after failure happens. Nevertheless, these studies seldom consider that network load can be redistributed in the local neighborhood of failure components and redistributed globally together according to the network topology. Moreover, multiple system dependencies within network systems have been proven to be able to significantly influence failure propagation process. Thus, an approach that considers both local and global load redistribution is developed to model network load dynamics regarding cascading failures. A new method is established to describe multiple system dependencies, in which network load is divided into two parts: passing by load and local customer load for certain network systems. Together, new

cascading failures models that describe different failures iteratively occurring in network systems driven by network load dynamics and system dependency is proposed.

The second theoretical contribution arises from the hierarchical structure of many network systems. This work is the first attempt at linking cascading failures in the central system to local subsystems, as well as the joint effects of system hardening and restoration on the entire network system. Thus, the results and findings are more general and realistic than previous research because of the existence of multiple dependencies within and between real-world network systems.

The third theoretical contribution stems from the optimization problem regarding cost-effective system resilience enhancement. In order to find the optimal combination of system hardening strategy and restoration strategy, two methods for sampling system hardening measures and restoration measures in the feasible sets are proposed. Especially, a two-stage optimization method is developed to compensate for the uncertainty of system disruptions. The iterative procedure continues until the predetermined requirements of resilience have been satisfied in order to determine near-optimal (optimal) solutions.

### 1.3.2   Applied contributions

There are two major contributions of this research that are related to electric power system resilience enhancement, and can also be adopted for other real-world applications. These contributions include the following:

1) Extending the impacts of cascading failures in bulk power system to power distribution systems with end users for integrated analysis of electricity system resilience.

2) Collectively optimizing system hardening and system restoration measures for enhancing electricity system resilience, and simultaneously minimizing related investment costs and economic loss from cascading failures.

The first applied contribution is integrating the impacts of cascading failures in bulk power system to end users through local power distribution systems. Previous studies on cascading failures in electricity system have an underlying assumption that cascading failures occur in generation and transmission systems, while the resilience in electricity system is mainly defined in power distribution systems. It is apparent that failures in bulk power system influence the operation of power distribution systems. Therefore, this research establishes and illustrates the link between cascading failures in bulk power system with power distribution system resilience, which offers a way to investigate the resilience of the whole electricity system regarding cascading failures.

The second applied contribution is jointly optimizing electricity system hardening and restoration strategies, and integrating system investment, system damage cost with resilience requirement in the objective function of the problem. The relationship between the effects of system hardening actions and restoration is investigated to effectively enhance system resilience. The minimization of the objective function makes the problem more practical because the trade-off regarding total cost is incorporated with resilience targets, which is important and necessary in real-world cases. It can be used to validate the optimization solutions for enhancing system resilience from an economic perspective.

## 2   Background and Literature Review

This review focuses on the resilience of electric power systems and how to enhance system resilience by developing optimal integrated hardening and restoration policies regarding cascading failures. First, a definition of electric power systems and their constitution is given in Section 1. Existing research on reliability of electric power systems is presented in Section 2. Section 3 reviews the trend of research on cascading failures in complex network systems. Then the state of current research on system resilience is introduced in Section 4. Finally, current research on system resilience enhancement is reviewed in Section 5.

### 2.1 Research on electric power system reliability

The electric power system is one of the most complex manmade technical systems in the world. It consists of the Bulk power system (BPS), which refers to the generation and transmission systems that facilitate the operation of an interconnected electric supply, and several power distribution systems which provide the final links between BPS and the industrial, commercial and residential customers [5, 6]. As of the end of August 2016, there were about 7,658 power plants in the United States that have operational generators with nameplate electricity generation capacities of at least 1 megawatt (MW) [7]. The transmission network in the 48 contiguous states is composed of approximately 697,000 circuit-miles of power lines and 21,500 substations operating at voltages of 100 kilovolts (kV) and above [8].

Electricity is produced at generation facilities and transported to population centers by high-voltage transmission lines. After arriving at population centers, electricity enters local distribution systems where it travels through a series of low-voltage lines in a

process called "stepping down" before reaching homes, offices, and other locations for consumption [9]. Figure 2.1 illustrates the basic structure of the U.S. Electric Grid.
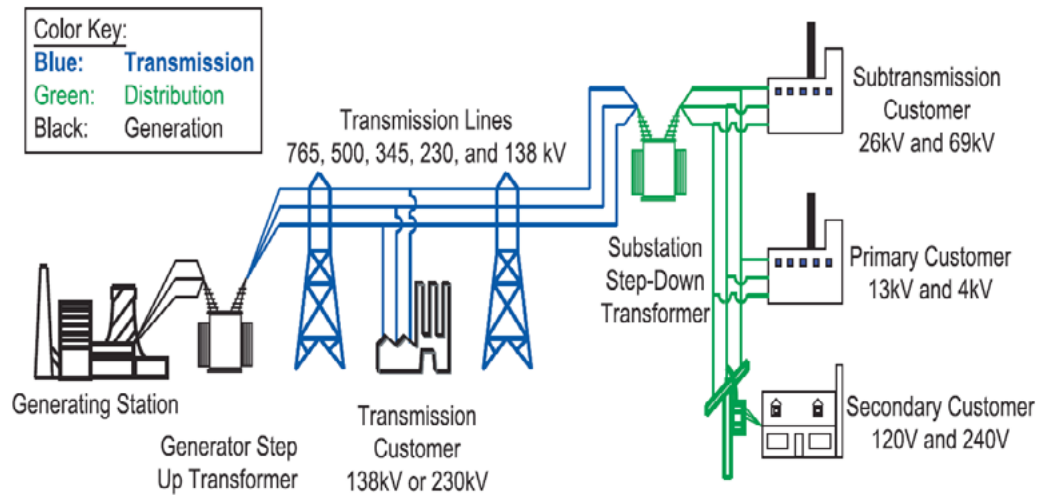


Figure 2.1. Basic structure of U.S. Electric Grid (Canada Power System Outage Task Force [9])

Critical infrastructure sectors have grown increasingly reliant on continued grid operations; for example, the interdependency of electric and natural gas systems is growing [10]. Likewise, in order to function, the electric grid increasingly depends on other infrastructure, such as the communications systems necessary for controlling electric grid systems [11]. Because all economic sectors and critical infrastructures are heavily dependent on electricity to provide essential services to support modern society and economy [10], and power outages would shut down many other major infrastructure components [12], electric power systems serve as a critical foundation for human life.

Electricity is at the center of key infrastructure systems that support these sectors, including transportation, oil and gas production, water, communications and information, and finance. These critical networks are increasingly converging, sharing resources and synergistic interactions via common architectures as shown in Figure 2.2.

**Critical Infrastructure Interdependencies**
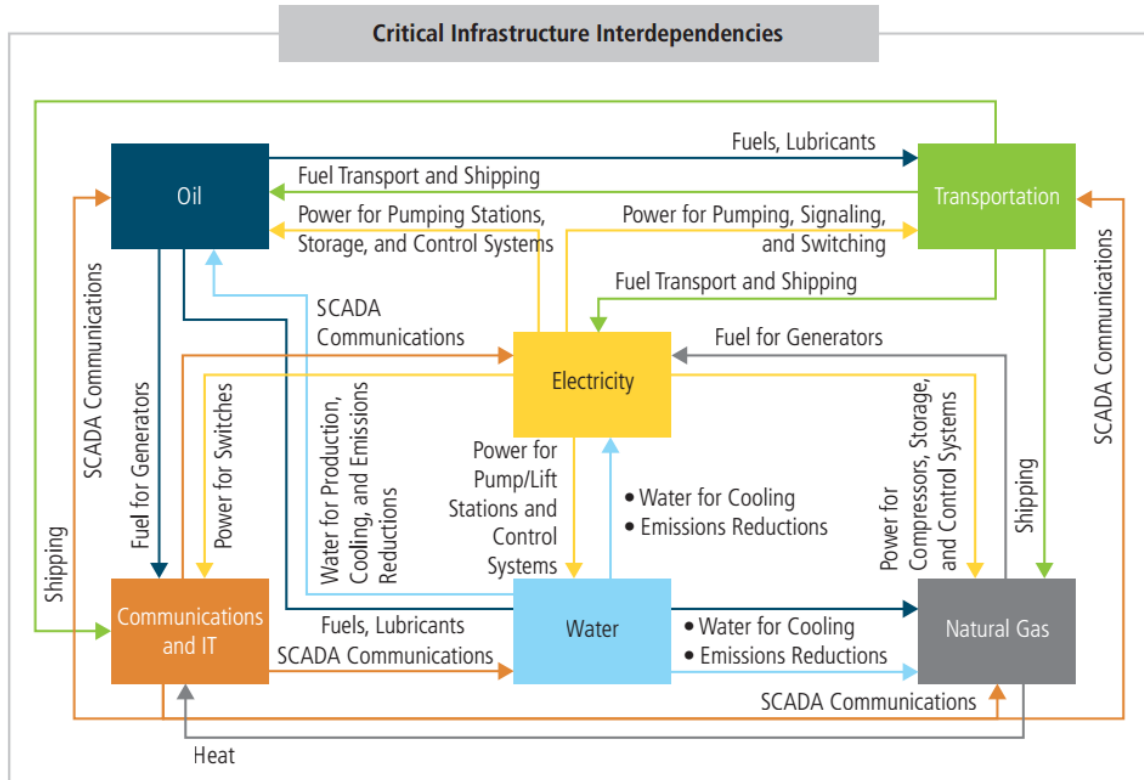
Figure 2.2. U.S. Critical Infrastructure Depend on Electricity (Source: Finster 2016 [10])

### 2.1.1   Introduction to electric power system reliability

Because of the crucial role played by electric power systems in modern society, power system outages can cause devastating impacts on many areas, such as the economy and national security [1, 2, 13]. Data suggest that electricity system outages attributable to weather-related events are increasing, costing the U.S. economy an estimated $20 billion to $55 billion annually [14]. For instance, the 2003 North American Blackout, which originated in Ohio state, finally affected millions of people for up to 4 days [4]. Reliability of electric power systems becomes a fundamental need and requirement across the entire economy [10], and it has been an important focus of researchers and engineers of different fields for many years. For example, by considering hurricane damage, Winkler et al. [15] illustrated that topological properties integrated with power

component fragility models can elucidate the contribution of system topology and spatial features to overall reliability of the power transmission system and distribution system.

The reliability of the BPS is regulated by the U.S. federal government, and the reliability of the power distribution system is regulated at a state level [5]. While there exists a national, North American Electric Reliability Corporation (NERC) - mandated, definition of reliability for the BPS (Reliability Performance Gap, Adequacy Gap, Violation Index and Severity Risk Index), there is no standard definition for reliability of the distribution systems. The reliability of power distribution systems is regulated at the state level and varies greatly. The System Average Interruption Frequency Index (SAIFI), the System Average Interruption Duration Index (SAIDI), the Customer Average Interruption Duration Index (CAIDI) have been the most commonly used outage indices for power distribution systems [16].

NERC's traditional definition of "reliability" was ubiquitous throughout the electric utility industry, and consists of two fundamental concepts − adequacy and operating reliability [17]. Adequacy is the ability of the electric system to supply the aggregate electric power and energy requirements of the electricity consumers at all times, taking into account scheduled and reasonably expected unscheduled outages of system components. Operating reliability is the ability of the electric system to withstand sudden disturbances such as electric short circuits or unanticipated loss of system components.

Nateghi et al. [5] provided recommendations for improving power transmission and distribution system reliability after finding that risk indices for transmission system and reliability indices for distribution system do not fully capture risks and reliability states of

these systems. Both practices fail to adequately account for risks and reliability issues associated with extreme events, and the impact of multiple system dependencies.

### 2.1.2 Reliability analysis of bulk power system

BPS includes generation plants and power transmission network. The transmission system transports electricity from generation facilities to the distributors, which connects power generators across the entire grid. Damage to the transmission system or unscheduled shutdowns of power stations, while infrequent, can result in more widespread major power outages that affect large numbers of customers and large total loads with significant consequences [10, 18].

Reliability of BPS has been studied for decades [19, 20]. Kintner-Meyer et al. [21] divided reliability into two sub-properties, i.e., operational reliability and planning reliability. These two sub-properties of reliability are different in timescales in which they are considered and measured. The metrics used for operational reliability of the BPS are: Area Control Error, Control Performance Standard 1 and Frequency and severity of emergency events. NERC regional entities primarily set standards for long term bulk power system planning reliability, including loss of load probability, loss of load hours, loss of load expectation, expected unserved energy, availability of supply with dual fuel or firm fuel contracts and flexible capacity [21].

Zio et al. [22] proposed an algorithm to evaluate the service reliability performance characteristics of the network and its related vulnerabilities, which consists of three nested cycles of randomization. Lin et al. [23] assumed power transmission network is stochastic and focused on searching for the optimal transmission line assignment such that network reliability is maximized. A genetic algorithm based method was developed

to solve this assignment problem and applied to a real power transmission network to demonstrate the computational efficiency. Rei et al. [24] applied a sequential Monte Carlo to model non-Markovian processes such as control actions, restoration stages and cascading events in BPS, and calculate both the probability for restoration and also reliability indices depending upon consequences of disturbances and the nature of the component failures.

Dong and Cui [25] proposed a normalized CASCADE model, two modified cascading models, and all of which are structure free. The time factor is considered by assuming that there is a period of time for nodes to fail after the load on them exceeds their thresholds. The time period follows an exponential distribution, e.g., the overloaded failure occurs at time $X_i$, $i \geq 1$, then the interval time between two consecutive overloaded failures, $Y_i = X_i - X_{i-1}$, with a probability density function $f(t) = \mu e^{-\mu t}$. Each generation of failures occurs due to the additional load caused by the last round of overloaded failures.

System reliability under these three cascading models is analyzed with the following formulae, e.g., for the system which works if at least one component works, system reliability $R(t)$ is calculated by Eq. (2.1).

$$R(t) = 1 - \sum_{i=0}^{n-1} P(S_i = n, X_i < t) = 1 - \sum_{I=0}^{n-1} P(N(t) > I) \cdot P(N_s = I, S = n) \qquad (2.1)$$

where $P(N(t) > I) = 1 - e^{-\mu t} \sum_{i=0}^{I} \dfrac{(\mu t)^i}{i!}$ denotes cascades have already stopped before time $t$.

It times the later item in above equation means before time $t$, all components have already failed, i.e., the system has failed before time $t$. However, the assumption that additional load caused by disturbances is added to all working components without priority are not

the real cases in most of the time. The current research seldom considers the interplay between multiple dependencies among components and power load dynamics, which impacts BPS reliability a lot and needs more exploration.

### 2.1.3 Reliability analysis of power distribution system

Power distribution system consists of all power system components at voltage levels below the transmission substations [5]. The metrics used for power distribution system reliability are the Average System Availability Index (ASAI), SAIFI, SAIDI and CAIDI as mentioned before [26, 27]. It should be emphasized that understanding and ensuring the reliability of the power distribution systems is very important since roughly 90% of electricity outages occur along distribution systems, although transmission system outages do occur [9]. In fact, failures on the power distribution system are typically responsible for more than 90 percent of electricity interruptions, both in terms of the duration and frequency of outages, which are largely due to weather-related events [10, 28].

Many research works have been focused on power distribution system reliability [29, 30]. By using data mining technologies, Guikema et al. predicted the number of utility electric poles that needed to be replaced based on past damage data prestorm [31]. Li et al. [32] introduced a probabilistic wind storm model and weather-dependent component failure models by using real database. Then, an enhanced sequential Monte Carlo method is applied to carry out risk analysis on distribution systems. Marnay et al. [33] expanded the model to include an assessment of distribution network reliability. However, they did not include siting of resources or contingencies. Hayashi and Matsuki [34] applied an algorithm to decide optimal configuration of a distribution system considering $N - 1$

security. The model determines whether the switches are active or not, which decides whether distribution generators connect to the grid. Bie et al. [35] adopted a non-sequential Monte Carlo simulation method to evaluate the reliability of distribution systems with the consideration of multiple contingencies within networks.

Wang and Rong [36] applied a local load redistribution rule on cascading model of edge overload failures to investigate the effects of three edges attack strategies on robustness of western U.S. power grid [37] which has 4941 nodes and 6594 edges, and identified key edges being prone to trigger universal cascading failures.

Alvehag and Soder [38] proposed a time-sequential Monte Carlo reliability model that treats variations in failure rate and restoration time of distribution system as a function of severe weather intensity and duration. The expected total failure rate of overhead lines under certain wind speed and lightning density is derived as Eq. (2.2), which equals to the summation of probability of weather condition times the expected failure rate under that weather condition.

$$E\left(\lambda(w(t), N_g(t))\right) = \frac{T_{hw}}{T_{tot}} E\left(\lambda_{wind}(w(t))\right) + \frac{T_{hw}}{T_{tot}} E\left(\lambda_{lightning}(N_g(t))\right) + \frac{T_n}{T_{tot}} \lambda_{norm} \qquad (2.2)$$

The restoration time for overhead lines is defined by Eq. (2.3), which is the product of weight factors for severe weather, hourly variations, daily variations, and reference restoration time under normal weather.

$$r(t) = f_w(w(t), N_g(t)) f_d(t) f_h(t) r_{norm} \qquad (2.3)$$

Nonhomogeneous Poisson Process (NHPP) are used to model when severe weather occurs and conclude that weather stochasticity significantly impacts the variance in reliability performance indices for SAIDI and energy not supplied (ENS). Issicaba et al. [39] presented an adequacy and security evaluation of distribution system in order to

assess the impact of device protection and controls on reliability indices and other performance metrics for distribution systems with ongoing integration of DER under various operational states.

According to the literature review, it can be seen that the current research does not consider the relationship between reliability of BPS and reliability of power distribution system even though it exists. In addition, the criteria of electric power system failure and the mechanism of failure propagation on BPS and power distribution system still need further investigation.

## 2.2 Research on cascading failures in complex network systems

As manmade technical systems such as transportation systems, electric power systems and telecommunications are becoming increasingly widely used, economic and social well-being depend on the secure and reliable operation of these complex systems [22]. However, recent worldwide events, such as the 2003 blackout in Italy, the 2003 North American blackout and 2012 Hurricane Sandy blackout have shown the increased system vulnerability [40]. Initial failures of components of these systems, even small ones, can trigger failure propagation by multiple dependencies, which spread adverse effects on a very large scale, severely impacting the whole system performance. That can be called cascading failures.

Cascading failures are network failures which occur in various real-life network systems, such as power grids and transportation systems. Generally, a network system carries a flow of some particular resource (electricity, gas, data packets, information, etc.). Each node individually experiences a load, and in normal circumstances, this load does not exceed the capacity of that node.

Cascade failures are commonly initiated in the following way; when a heavily loaded network node is lost for some reasons, the load on that node (i.e. the flow passing through it) must be redistributed to other nodes in the network system. This load redistribution might cause other nodes to exceed their capacity, triggering their breakdown. Even if an overloaded node does not actually fail, the pre-designed protection mechanisms inside the network may shut it down, in order to prevent damage to this node [41]. Therefore, the number of failed or stressed nodes increases, propagating through the network system. In particularly severe cases, the entire network is influenced.

Although cascading failures occur with a low probability, a sharp degradation of system performance, even the collapse of much of the complex system, will be experienced when they occur [42-45]. Some massive cascading failures have taken place in complex communication network systems, social network systems and economic network systems [46]. Therefore, the study of this phenomenon is particularly significant.

Many researchers have studied cascading failures in complex systems, and many models have been proposed for describing the process of cascades [47, 48]. Current cascading failures models mainly include load-capacity models [49-51], binary influence models [52], sand pile models [53, 54], optimal power algorithm (OPA) models [55, 56] and CASCADE models [57, 58]. OPA models and CASCADE models are focus on cascading failures in electric power systems.

Dependence relationships among network nodes, except for topological connectivity links, also accelerate the propagation of network failures and affect the mechanism of cascading failures [59-61]. Past incidents have demonstrated the cascading impacts posed by interdependencies, by highlighting ways that the electric grid depends on other sectors,

including the communications infrastructure and information systems. For instance, the 2003 Northeastern Blackout, which began initially with power lines in contact with tree branches, was magnified by a series of cascading computer failures that affected airline operations, the financial and banking sector, blood, and potable water supplies, and other critical services [11]. Hence, modeling and simulation of cascading failures considering interdependencies between network systems and dependencies inside the network system becomes a vital field of research [62-66].

Especially for electric power systems, cascading failures are common and have devastating influence [44, 63, 67, 68]. Recent blackouts, for example, the massive blackout in India [69] have illustrated the importance of studying on cascading failures in electric power systems, and other research on power systems with cascading failure analysis has been carried out [70, 71]. Although modeling of the mechanism of cascading failures on BPS has been studied more recently, the impacts of it on power distribution systems, which together form the entire electric power systems, are lack of discussion and need further research.

## 2.3 Research on system resilience

In this section, current definitions of system resilience in different areas and popular resilience models are introduced. Then previous research on resilience enhancement of electric power systems is presented, including current system hardening strategies and restoration strategies.

### 2.3.1   System resilience definitions and models

The general use of the word "resilience" implies the ability of a system to withstand or quickly return to normal condition after the occurrence of an event that disrupts its

state. Resiliency threats tend to be idiosyncratic, of low probability, and of varying degrees of magnitude in terms of scale and duration. The types of uncertainties and the capacities to respond to them differentiate resiliency from reliability [21].

Modeling and evaluating system resilience, especially the complex and large-scale systems, has recently raised significant interest among both practitioners and researchers. It is contributed to several definitions of system resilience and some approaches to measure it, across many disciplines. Hollnagel et al. [72] defined engineering resilience as the intrinsic ability of a system to adjust its functionality in the presence of a disturbance and unpredicted changes. A typical definition of time-dependent resilience metric is illustrated in Figure 2.3 [73],
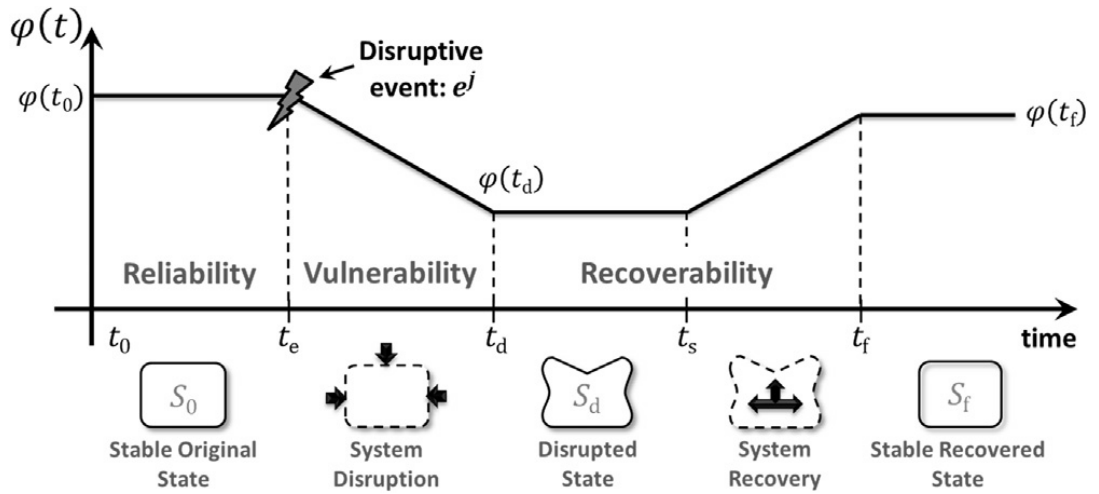


Figure 2.3. System performance for resilience description [73]

$$\Re_{\varphi}(t \,|\, e^j) = \frac{\varphi(t \,|\, e^j) - \varphi(t_d \,|\, e^j)}{\varphi(t_0) - \varphi(t_d \,|\, e^j)} \tag{2.4}$$

Resilience is represented by the ratio of system recovery at time $t$ to the total system loss caused by disruptive event $e^j$, $\varphi$ is a measure of system performance in Eq. (2.4). This system resilience includes reliability, vulnerability, and recoverability. In this work,

a case study performed on a road network system is used to demonstrate the applicability of the proposed resilience metrics and to show how to develop effective resilience design strategies. Time and cost required to restore each component are assumed to be constant in the case study, and restoration strategies decide sequences in which the components are restored.

Some other definitions of system resilience are also proposed as follow.

1) Bruneau et al. [74] introduced resilience Loss (*RL*) which can be illustrated as the shaded area in Figure 2.4 and Eq. (2.5). Larger *RL* value indicates lower system resilience.
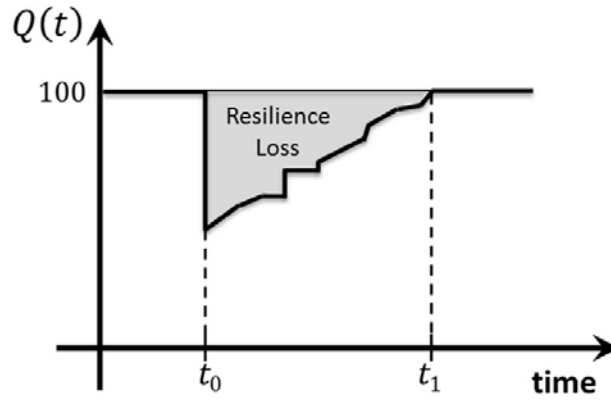


Figure 2.4. Illustration of resilience loss

$$RL = \int_{t_0}^{t_1} [100 - Q(t)]dt \qquad (2.5)$$

where disruption occurs at $t_0$, system returns to pre-disruption state at $t_1$. System quality level at time $t$ is represent by $Q(t)$, which could be several performance parameters. Assume planned quality of system is 100.

2) Rose [75] proposed time-dependent recovery within dynamic resilience. The measure, *DR*, is a function of $SO_{HR}(t_i)$, system output under hastened condition and $SO_{WR}(t_i)$, system output without hastened at the $i$th step of recovery as shown in Figure 2.5 and Eq. (2.6). The value of this dynamic resilience is not bonded between 0 - 1.
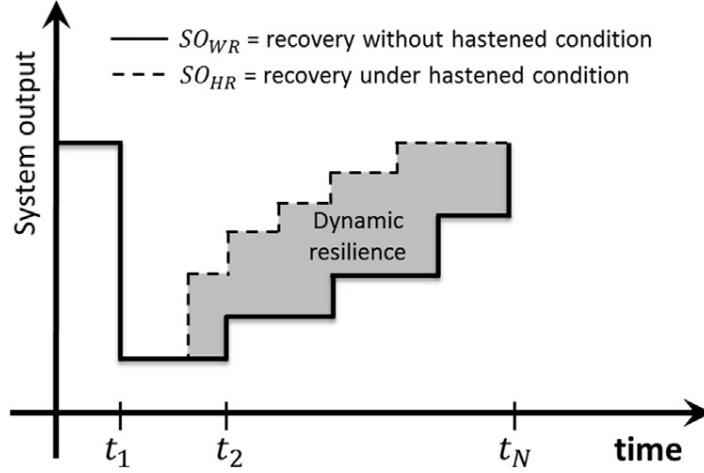
Figure 2.5. Dynamic resilience

$$DR = \sum_{i=1}^{N} SO_{HR}(t_i) - SO_{WR}(t_i) \qquad (2.6)$$

Due to the crucial socioeconomic role of infrastructure systems such as water distribution systems, nuclear plants, and transportation systems, research work has recently focused on infrastructure resilience. Omer M et al. [76] defined the ratio of the closeness centrality of the network before and after disruption is resilience metric for infrastructure network. Closeness centrality is determined based on the accessibility of a node to the rest of the network. Ouyang and Wang [77] applied five types of joint restoration strategies to an interdependent power and gas system under multiple hurricane hazards as an example. Cascading failures within and across this interdependent system are simulated, where the unidirectional links from power system to gas system constitute the only interdependencies.

MacKenzie and Zobel [78] investigated methods of resources allocation prior to disruptions to enhance resilience, either lessen adverse impact from disruption or shorten

recovery time. Resilience is calculated as the product of disruption impact and recovery time. For example, the linear function of allocation effectiveness is provided as Eq. (2.7),

$$\overline{X}\left(z_{\overline{X}}\right) = \widehat{X} - a_{\overline{X}} z_{\overline{X}}, \ \ T\left(z_T\right) = \widehat{T} - a_T z_T, \ \ a_{\overline{X}} > 0, \ \ a_T > 0 \tag{2.7}$$

where $\widehat{X}, \widehat{T}$ are the baseline average loss and recovery time if no resources are allocated. $z_{\overline{X}}$ is the amount of resources allocated to lessen the disruption impact, $z_T$ is the amount allocated to improve the recovery time, $a$ are effectiveness parameters. Three kinds of uncertainties of above four variables are considered to optimize resource allocation for maximum expectation resilience. Specific projects to pursue to restore resilience are also discussed based on cost effectiveness of projects.

Vugrin et al. [79] introduced resilience cost index, which is composed of loss costs posed by disruptive events and recovery costs. Ash and Newth [80] first proposed an optimized model of links reassignment for resilient complex networks against cascading failures. Then, they analyzed the impact of network topology statistics on resilience. Impacts of interdependency within infrastructure networks on system resilience are translated into loss incurred, extent of and duration of system inoperability by Baroud et al [81]. They presented a stochastic approach to compute metrics of resilience of an interdependent network after a disruption: loss of service cost, total network restoration cost and cost of interdependent impacts. The third one is to measure multi-industry impacts of resilience in infrastructure network by integrating a network resilience model and an economic interdependency model. Barker et al. [82] introduced two resilience-based component importance measures. The first one measures how adverse effect on system resilience could be if one component is affected by a disruption event. The second measures how positive effect on system resilience could be if one component is protected

from being disrupted. Stochastic terms were used to define the recovery time of component and quantity the extent of damage on component. Pant et al. [83] proposed stochastic and time dependent metrics of system resilience, time to total system restoration, time to full system service resilience and time to $a\%$ Resilience. The presented recovery strategies deciding the order of failed components being repaired and recovery time of each failed component. The effects of recovery strategies are evaluated by resilience and the three recovery measures. Zobel [84] extended a multi-dimensional approach to fit resilience function to the preference and priorities of a decision maker. The figure of different combinations of $X$ and $T$ could be obtained with fixed value of resilience function $R$ (proportional to $X \times T$, based on resilience triangle), where $X$ is loss of system performance and $T$ is the recovery time. Then this resilience figure could be divided by different threshold values of $X$ and $T$ based on preference of decision maker.

Many studies have already been conducted on resilience of electric power systems. Resilience reflects the ability of the system to respond to the threat of non-routine disruptions that are difficult to predict or plan for. However, reliability risks are driven by common, internal, but uncertain factors such as generator and transmission line outages, load variability, and intermittent and variable wind and solar generation. The idiosyncratic and low-probability nature of resilience risks makes measuring and valuating resilience challenging [21].

Unlike reliability, there are no commonly used metrics for measuring grid resilience. Several resilience metrics and measures have been proposed; whereas there has been no coordinated industry or government initiative to develop a consensus on or implement standardized resilience metrics [10].

It should be noted that infrastructure resilience which considers both interdependencies and restoration processes is seldom addressed in previous research. More work is needed to incorporate existing interdependent infrastructure models with the study of interdependent resilience [85]. At the same time, planning for system resilience makes resilience quantifying research applicable and usable, which is of great significance and needs exploration.

### 2.3.2 Research on resilience enhancement of electric power systems

Because of the significance of electric power systems and common cascading failures in them, research has been carried out on developing and evaluating methods of maintenance to restore electric power systems from disruptions, and methods of hardening to avoid further system loss [86, 87].

Resilience is typically achieved through hardening or recovery. Hardening refers to physically changing infrastructure to make it better able to withstand the impacts of weather events or attacks. Recovery refers to the ability of systems to recover quickly from damage typically through storage and redundancy. It does not prevent damage but enables the system to continue operating despite damage or rapid return to normal operations [88].

System-wide reliability and resilience can be supported by a diverse portfolio of generation resources that limit over-dependence on any single fuel or technology type, plus demand-side resources that reduce overall demand and better customers in the event of a widespread extreme event [88].

Resilience enhancement initiatives are generally focused on achieving at least one of three primary goals: (1) preventing or minimizing damage to help avoid or reduce

adverse events; (2) expanding alternatives and enabling systems to continue operating despite damage; and/or (3) promoting a rapid return to normal operations when a disruption occurs (i.e., speed the rate of recovery) [10].

The Electric Power Research Institute (EPRI) states that enhanced power system resiliency is based on three elements: 1) damage prevention which includes the application of engineering designs and advanced technologies that harden the power system to limit damage, 2) system recovery which use tools and techniques to restore service as soon as practicable, and 3) survivability, which means use technologies to aid consumers, communities and institutions in continuing some level of normal function without complete access to the normal sources [89]. According to EPRI's research [89], there are fourteen actions can be taken in a number of areas to enhance the transmission system's resiliency, such as targeted transmission undergrounding and vegetation management. Based on the reliability measure of average total duration of the interruptions experienced by a customer, more than 90% of the minutes lost by consumers annually are attributable to distribution events [89]. Several actions can be taken to prevent damage to the distribution system, including vegetation management, targeted undergrounding and overhead distribution reinforcement, etc [89].

Liu et al. [90] studied the effects of restoration strategies executed during cascading failures on power transmission system reliability. The restoration strategies differ in terms of restoration timing in overloaded cascading processes and restoration strength (probability of fully repairing a failed component). Additional disturbances, which could be positive or negative, are also added to functional components due to restoration. When restoration disturbance is positive, restoration will increase load of functional

components while failed components are restored with probability. Total number of failed components *ES* and total system load fluctuation (*SLF*) caused by cascading process are applied to compare effects of different restoration strategies. System load fluctuations *LF* during cascading failures are collected to track system evolution. Load fluctuation of system at time *t* is the absolute difference between total system load at time *t* and initial total system load. Theoretical probability of total number of failed components coincides well with the numerical simulations. The proposed framework of restoration is applied to western U.S. power transmission grid for model validation, except for analysis based on multiple simulated networks.

Tokgoz et al. presented a methodology to assess the resilience of electric power distribution poles against wind-related events. It provided a metric to evaluate the resilience of poles under pre-disaster and post-disaster conditions [91]. Pham et al. [92] presented a new restoration procedure by using dispersed generation (DG) for distribution system based on an adapted branch-and-band algorithm to reduce out-of-service load volume and duration of restoration process. Simulation results illustrated the proposed procedure and benefit of using DG in distribution system restoration.

Mo and Sansavini [93] proposed a defense strategy with regard to optimizing the fraction of resource allocated on enhancing component protection and increasing unit redundancy for minimizing expected damage costs of Cyber Physical System (CPS). The expected damage costs for CPS include unsupplied system demand and inherent value of cyber-physical unit which is composed of a cyber component with several physical components. Uncertainties about cyber-attack time and defender's estimation about it are described by truncated normal distribution. Impacts of these uncertainties on defense

strategies are also analyzed, but they do not consider component dependency, other types of triggering failures and cascading failures. Hou et al. [94] presented generic restoration milestones (GRMs) to provide specific restoration strategies based on actual system conditions. This obtained decision support tool could reduce restoration times to improve system reliability.

Madathil [95] developed a mathematical model to help decision makers optimally design and operate a cost-effective and resilient microgrid with the consideration of $N$ - 1 system security on generator and line contingencies. The solutions from the proposed time-efficient algorithms using real system data recommend that build redundant lines/dispatching power efficiently for off-grid microgrids could help to improve network reliability at reduced cost when compared with installing backup generators. Gutfraind [96] introduced an optimization method for robust constructing networks with high passive cascade resilience. Synthetic networks were able to be constructed to achieve structural cascade resilience and efficiency. It suggests that network optimization can significantly improve the fitness and cascade resilience of networks.

Many efforts have been made to prevent cascading failures on electric power systems, such as capacity redundancy design [97], $N$ - 1 criterion consideration [98] and optimal transmission lines layout for resilient power supply with minimized cost [99]. Recovery actions after cascades occur are also studied, such as microgrids with black start [100, 101], Self-healing system [102], failed components removal [103], load redistribution control or load shedding [104-106], load-capacity optimal distribution with local or global redistribution models [107, 108], as well as other restoration planning and actions [109-111].

Figueroa-Candia et al. [29] proposed a modeling approach to perform optimization on the restoration policies for power distribution systems subjected to extreme weather events in order to meet resilience targets while lowering cost. New resiliency metrics to evaluate the effectiveness of restoration policies are developed, and it is demonstrated that the efficacy of restoration strategies and resiliency metrics are closely linked and there is a trade-off between these two choices. However, the dependencies among system components have not been taken into consideration, so as well cascading failures in electric power systems.

DER are important parts of the electric power system, such as solar array generation and energy storage, which can work in islanded mode during grid outages and supply power to end users instead of power grids [112, 113]. For instance, solar arrays combined with battery systems are a model for strengthening the resilience of distributed power generation to reduce power interruptions of critical facilities [114, 115].

Although there is an increasing growth of DER installation due to DER price decline, they remain relatively expensive, such as solar module and battery. Thus, optimization of DER investment, for example, optimal installing size of battery backup storage, as a hardening strategy to provide power resiliency during grid outage is pressing and significant from both resilience enhancement and economical perspectives, which need further exploration [116].

Additionally, Castillo [117] concluded that there has not been an effective and unanimous approach to relate reliability and resiliency to market efficiency and economic losses. Future work in this area can further address integrating the risk analysis of power outages into investment and restoration planning decisions in order to better incorporate

grid resilience targets, restoration strategies, the adoption of smart grid techniques, and hardening of critical components.

# 3 Cascading failure modeling with system dependency

The preliminary results of this research mainly include three parts. First, the current cascading failure model is improved by taking account of the combined impacts of network load dynamics and system dependency on failure propagation. System dependency is described by dependence clusters and two types of system failures iteratively occur throughout the system. Then, based on the proposed mixed cascading failure model, different restoration strategies are performed to mitigate the impact of system failures and reduce performance loss in network systems. The impacts of dependence characteristics on the effects of restoration strategies are investigated in order to determine the optimal restoration strategies for specific network systems. In the end, economic and resilience benefit analysis of installing battery storage to solar array generation for end users in power distribution systems has been conducted. It is a system hardening measure to provide power resilience during grid outage which can be initiated by cascading failures in power generation and transmission system. The details of each part of research results are introduced as follows.

## 3.1 Modelling mixed cascading failures with local load dynamics and dependency

In this part, a new cascading failure model has been developed to analyze the mixed types of system failures spread over network systems [118]. Previous cascading failure models, which focus on network load dynamics, provide alternative approaches to analyze cascading process of failures in network systems. However, these studies seldom consider the joint impacts of multiple system dependencies among network nodes, which exert a great impact on network system dynamic behaviors. The new model extends

previous research by taking both load dynamics and network dependency which forms

dependence clusters into consideration in a single/isolated nework.

Dependence clusters are proposed to represent the dependence relationships of network nodes apart from the topological connectivity links of the network. These dependence clusters include the network nodes with dependence interactions between each other that are beyond the topological connections. This type of single network containing dependence clusters is presented in Figure 3.1.

The network topological structure has network nodes denoted by points, as depicted in Figure 3.1. The solid lines in Figure 3.1 represent the topological links in the network while the nodes depending on each other by dependence links are represented by dashed lines. The dependence groups are circled, which identify the dependence clusters of nodes in the network.
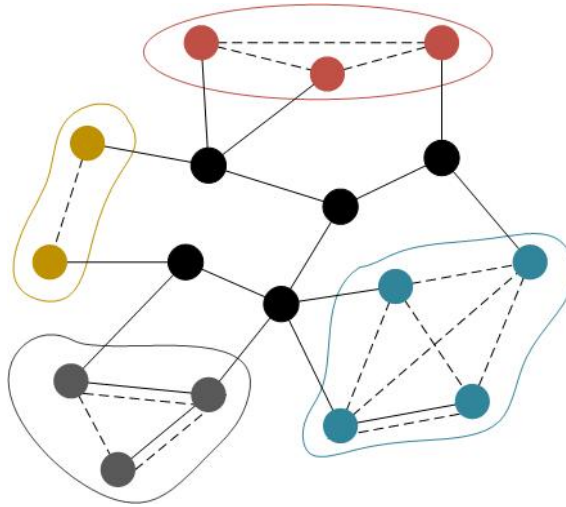


Figure 3.1. Illustration of a single network with dependence clusters of nodes

A single network is defined to be an unweighted, undirected, self-loop free and single edge graph *Net*. It is represented by an adjacent matrix containing *N* nodes $\{v_1, v_2 ..., v_N\}$ and *M* edges $\{e_1, e_2, ..., e_M\}$. For instance, network nodes can represent

power stations (or sub-stations) and edges could represent power transmission elements in power grids. If there is a topological link between two network nodes, the weight of the edge representing this link in an adjacent matrix is 1. If there is no topological link between two nodes, the corresponding edge weight is 0.

It is assumed that network load only transmits along the shortest paths between every pair of network nodes, where paths consist of edges (topological links) between nodes. The length of a path is calculated by summing the weights of edges along the path. If there is no path between a pair of nodes, the distance between these two nodes is assumed to be infinite.

### 3.1.1 Preliminary model

In this section, it is assumed that load on node $i$ at time $t$, $L_i(t)$, is defined as the "betweenness centrality" [119] of node $i$ at time $t$, which can be calculated by Eq. (3.1).

$$L_i(t) = \sum_{s \neq i \neq l} \frac{\sigma_{sl}(i,t)}{\sigma_{sl}(t)}, \ i, s, l \in N \tag{3.1}$$

where $\sigma_{sl}(t)$ is the number of shortest paths between nodes $s$ and $l$ at time $t$. $\sigma_{sl}(i,t)$ denotes the number of shortest paths between nodes $s$ and $l$ passing through node $i$ at time $t$. $s$, $l$ can be any nodes in the network.

Other methods of defining load and modeling flow in network systems have been proposed, but the load adopted here has been widely used to depict the real flow in network systems [103, 105, 120, 121], such as current for power grids and traffic for transportation networks. Node load is calculated based on the algorithm presented in [122]. Node capacity denotes the maximum load that a node can process without congestion. A nonlinear capacity-load model is used to define the capacity of each

network node [108] as Eq. (3.2):

$$C_i = \alpha\left(L_i(0) + L_i(0)^{1-\mu}\right), \quad i = 1, 2, ..., N, \quad \alpha \geq 1, \quad 0 < \mu < 1 \tag{3.2}$$

where $C_i$ denotes the capacity of network node $i$. $L_i(0)$ is the load of node $i$ at initial time $t = 0$. $\alpha$ and $\mu$ denote the parameters to modify network node capacity. The effects of these two parameters were discussed in the papers [50, 108], and both of them are applied to describe the nonlinear characteristic of node capacity and load.

The proposed mixed cascading failures model includes two types of network failures. First, as assumed in traditional cascading failures models, only dynamic network load distribution is treated as the cause of cascading failures. For this kind of failures, failed nodes and their connected links are disconnected from the network. As a result, some shortest paths between pairs of nodes in the network are changed. Network load is then redistributed along the latest shortest paths which can cause more load added to some working nodes. Once the load exceeds node capacity, these nodes are considered to be overloaded breakdown. They are removed from the network and then compositions of some sets of shortest paths in the network are changed again. New rounds of network load redistribution along with the updated shortest paths and overload cascading failures ensued are iteratively triggered in this way. This is the load dynamics-caused failure. Second, taking into account the impacts of dependence relationships among network nodes, immediate failures of entire dependence clusters of network nodes occur if any nodes belonging to these dependence clusters break down. This is dependence-caused failure. Eq. (3.3) defines this type of failure,

$$Dep(v_i, l) = \begin{cases} 0, v_i \notin \{Cluster\, l\} \\ 1, v_i \in \{Cluster\, l\} \end{cases} \tag{3.3}$$

where $Dep(v_i, l)$ denotes whether network node $v_i$ fails given dependence cluster $l$ collapses. 1 means that node $v_i$ immediately fails because it belongs to cluster $l$, while 0 means that node $v_i$ does not belong to dependence cluster $l$ and is not impacted by the collapse of cluster $l$. In actual systems, multiple dependencies are often highly complex, for example, there can be dependencies between different clusters in some systems, but here the impact of dependency inside clusters on cascades is only considered. This is a reasonable assumption that is considered in many common applications [61, 65, 66].

Nodes that break down in any one of these two processes could cause the other failure process. For example, a network node failure caused by overload (i.e., load dynamics-caused failure) leads to direct failures of other nodes which belong to the same dependence cluster with the overloaded nodes, i.e., dependence-caused failures occur. It changes network topology and consequently, a subsequent round of network load redistribution is triggered, which might cause new overloaded node failures. The iterative process of these two types of failures is briefly shown by the four steps in Figure 3.2. It should be noted that in Figure 3.2, dots represent network nodes, and solid lines denote topological links. Circles represent dependence clusters and dashed lines denote dependency links between dependent nodes. Once a node failure occurs, it is disconnected from the network system along with the edges connected to it.

Since dependency among network nodes and topological connectivity links are assumed to be unrelated, node failures caused by node dependency are independent of network structure. According to the proposed model, after a fraction of node breakdown, the failures of nodes caused by dependence clusters and load dynamics recursively occur, leading to a mixed process of cascading failures.
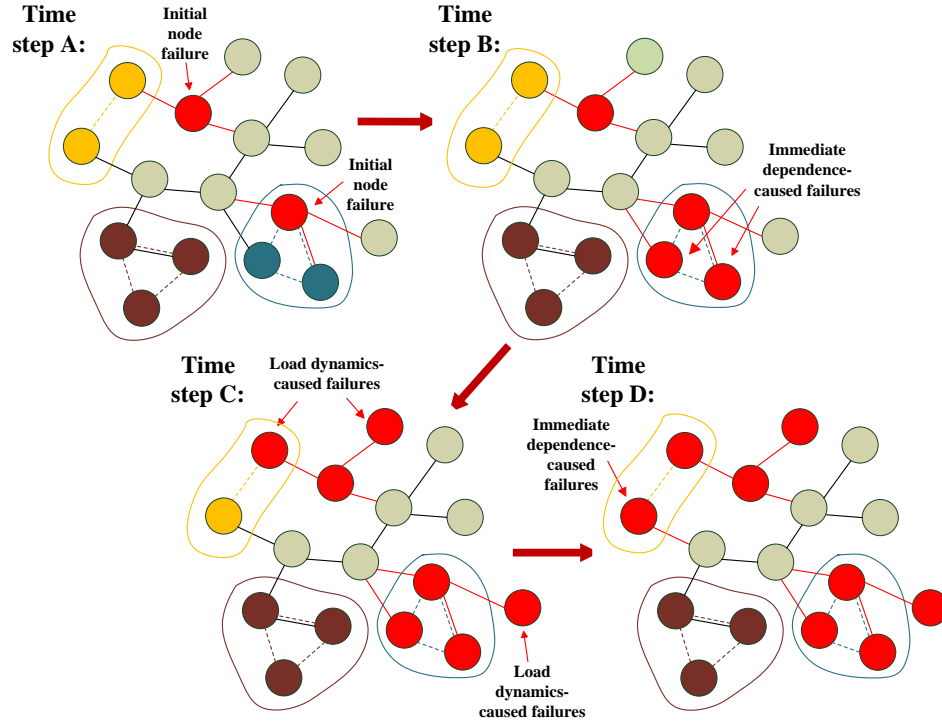
Figure 3.2. Cascading process of two types of failures in a network with dependence

clusters

### 3.1.2 Network cascading failure process

The main steps of the simulation, including two iterative failure processes are depicted as follows,

Step 1)   a. Construct *Net*(*N*, *M*), and calculate the size of largest connected component of the network (*LCC*).

b. Determine dependence clusters in *Net*(*N*, *M*).

Step 2)   Allocate capacity of each node based on Eq. (3.2).

Step 3)   a. Select failures on nodes randomly as initial failures.

b. Remove all failed nodes and their connected edges from the network.

Step 4)   a. Determine network nodes within the dependence clusters which include initially failed nodes.

b. Remove those nodes along with their connected edges from the network.

Step 5)   Update shortest paths in the network based on current network structure.

Step 6)   Calculate load on each node at time $t$, $L_i(t), i \in N$, after new round of load

redistribution.

Step 7)   If $L_i(t) > C_i, i \in N$, remove the overloaded nodes and other nodes of the

same dependence clusters. Then go back to Step 5. If not, go to Step 8.

Step 8)   Calculate the final size of $LCC$ and the number of iterations of cascades, $T$.

Simulation ends.

The parameter used to evaluate the effect of mixed cascading failures on network

structure is $G$, which indicates network connectivity, and is shown as Eq. (3.4),

$$G = \frac{N_F}{N_I} \tag{3.4}$$

where $N_I$ is the initial size of $LCC$ before cascading starts, and $N_F$ is the final size of $LCC$

of the network when cascading stop. As $G$ decreases, a greater proportion of network

collapse is because of cascading failures. $T$ is the duration of cascading failures until it

reaches the stop criteria in the simulation. It is used to evaluate the speed of failure

propagation, which.

### 3.1.3   Numerical examples

Network examples with the sizes of dependence clusters of nodes following Poisson

distribution or approximated by a truncated normal distribution are investigated,

respectively. Simulations are performed on a general network structure, ER random

network [123], which was proposed by Erdos and Renyi. It is a typical model for

constructing random networks, whose load distribution and node degree distribution follow Poisson distributions.

### 3.1.3.1    Poisson distribution of dependence clusters

The case that sizes of dependence clusters follow a shifted or scale adjusted Poisson distribution with $D \geq 1$ is studied, i.e., the probability $P(D)$ that a node belongs to a dependence cluster of size $D$ is given by Eq. (3.5) [66],

$$P(D) = \frac{\lambda^{D-1} e^{-\lambda}}{(D-1)!}, \text{ for } D \geq 1 \tag{3.5}$$

where $\lambda = (D\text{-}size) - 1$. $D\text{-}size$ denotes the mean size of dependence clusters in the network, that is, the average number of nodes included in a dependence cluster.

The impact of the only parameter, mean size of dependence clusters $D\text{-}size$, on network robustness against cascading failures is explored. Figures 3.3 and 3.4 present the changing values of $G$ and $T$ obtained in cascading simulation using the proposed mixed cascading model. The networks in Figures 3.3 and 3.4 have $N = 1000$ and $K = 10$. $K$ is the average degree of network nodes, i.e., average number of topological links that a node has in the ER random network. The parameters for $C_i$ are $\alpha = 1.05$ and $\mu = 0.5$. *Nrem* is the number of network nodes randomly failed at the beginning of the cascading simulation.

Every point in the plotted curves in Figure 3.3 corresponds to an average of 20 random initial failure triggers on 20 different randomly generated ER random networks with the same pair of $(N, K)$. Note that because the curves shown in the figures are average simulation results, those curves may not be as smooth as theoretical results would suggest. The atypical behavior of some curves is attributed to the randomness associated with simulation.

Figure 3.3 shows the results of parameter *G* vs. *Nrem*, while Figure 3.4 shows the results of parameter *T* vs. *Nrem*. Four subgraphs (a), (b), (c), (d) in the figures present the simulation results considering dependence clusters of different mean sizes, i.e., *D-size* = 2 to 5. The plotted curves in the figures mainly present the simulation results (points) which are close to the transition point (critical point).



Figure 3.3. Simulation results of *G* for mixed cascading failures
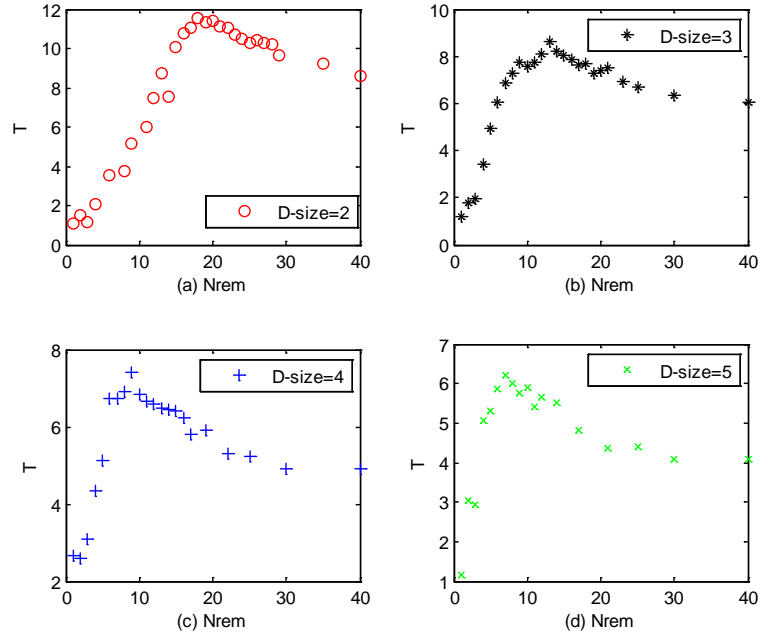
Figure 3.4. Simulation results of *T* for mixed cascading failures

As illustrated in Figure 3.3, phase transition occurs as the changing trend of *G* vs. *Nrem*. Parameter *G* sharply drops as *Nrem* increases to the critical transition threshold value, *NremC*, for *D-size* = 3 to 5, i.e., first-order phase transition occurs, and the network system breaks down directly to the totally collapsed state. *NremC* indicates the critical number of nodes initially removed which cause collapse of the network. The network collapses when the number of initially removed nodes is larger than *NremC*, otherwise, the network does not collapse although cascading failures still can occur.

The network, which undergoes a first-order transition of cascades, indicates that the size of *LCC* abruptly decreases, i.e., indicator *G* goes down discontinuously from large value (for *Nrem* < *NremC*) to almost zero (for *Nrem* > *NremC*). Such a network is obviously more vulnerable than a network that is subjected to a second-order transition of cascading failures, where *G* continuously decreases from a finite value to almost zero at

transition point *NremC*. When the number of network nodes initially failed equals or is larger than *NremC*, the initiated cascading failures collapse the network very quickly, meaning the network is completely fragmented. Thus, *NremC* could be used as an indicator to show the robustness of a network to cascading failures triggered by random node failures.

With *D-size* = 2 or 1 (dependency is not considered in the cascades when *D-size* = 1), *G* changes continually as *Nrem* changes, which indicates the occurrence of a second-order phase transition. In Figure 3.4, the peaks of changing trends of *T* also emphasize the transition points *NremC*, which are consistent with *NremC* as shown in Figure 3.3. Moreover, as illustrated in both figures, as mean size *D-size* increases, the value of critical transition point *NremC* decreases.

The results of transition points *NremC* and evaluation parameter $T_c$ in the cascading simulation in Figures 3.3 and 3.4 are shown in Table 3.1.

Table 3.1. $T_c$ vs. *NremC* with different *D-size* under two cascading failure models

| D-size | Poisson distribution of dependence cluster | | | | Without dependence clusters |
|--------|------|------|------|------|------|
| | 2 | 3 | 4 | 5 | |
| NremC | 18 | 13 | 9 | 7 | 46 |
| $T_c$ | 11.57 | 8.6 | 7.4 | 6.2 | 20.75 |

Based on the simulation results in Table 3.1, conclusions about mixed cascading failures are as follows.

1) When the sizes of dependence clusters of network nodes follow the Poisson distribution, the process of cascading failures changes from a continuous second-order phase transition to a discontinuous first-order phase transition, compared with the cascading process without considering node dependency. As can be seen from

Table 3.1, *NremC* and $T_c$ of the combined cascading networks (first-order phase transition) are much smaller than that of cascades without dependence clusters (second-order phase transition). This indicates that the robustness of the network when exposed to mixed cascading failures is less. Similar conclusions were observed in previous papers which focus on interdependence networks [63]. Change in the type of phase transition explains why mixed cascading failures is much more damaging in the network.

2) *D-size* exerts a notable influence on the effect of dependence clusters on mixed cascading failures. As shown in the results and figures, $T_c$ and *NremC* decrease if *D-size* increases. This means that the effect of dependence clusters on network robustness is higher when *D-size* increases. It indicates that for larger *D-size*, the impact of dependence clusters is more noticeable causing failure propagation. This observation is in agreement with the previous findings about the impact of parameter *D-size* in [51] which assumes all dependence clusters in the network have the same fixed sizes.

3) Compared with the findings in [51] with the same network, the values of $T_c$ obtained in the new mixed cascading simulation are larger given the same values of *D-size*. This is likely because the dependence clusters considered in [26] are mean size, but they have randomly different sizes in the network for the proposed new model.

Then, the impact of the average degree of network node is presented, which is an important network topological property, denoted by *K* on network robustness regarding mixed cascading failures. The obtained simulation results on the ER random network (*N* = 1000) with several values of *K* are illustrated in Figure 3.5.
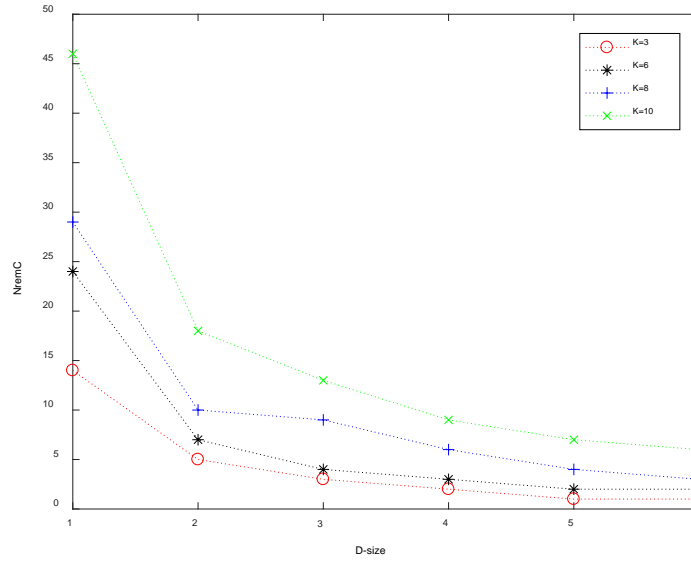
Figure 3.5. Transition points of mixed cascades under different average degree

*NremC* are plotted in Figure 3.5 as a function of *D-size* under different values of *K*. These trends of *NremC* clearly show that *D-size* impacts network robustness to mixed cascading failures in ER random networks with different average degree. Small *NremC* indicates that robustness of the network against mixed cascading failures is poor. In other words, breakdown of just a few nodes could trigger the mixed cascades that collapse the entire network. It should be noted that *D-size* = 1 indicates a network without dependence clusters of nodes. It can be observed that values of *NremC* for *D-size* = 1 differ from that under $D-size \geq 2$. The appearance of change in phase transition of cascading failures is caused by *D-size*. Moreover, it can be seen from Figure 3.5 that, with the same *D-size*, as *K* increases, *NremC* also increases. The ER random networks with larger average degree *K* are more robust to mixed cascading failures triggered by random failures of network nodes.

### 3.1.3.2    Normal distribution of dependence clusters

The next case is that the dependence clusters of nodes in a single network follow a truncated normal distribution. The normal distribution is a continuous random variable distribution, while the number of nodes is discrete. However, the random number of nodes can be approximated using a normal distribution probability density function. Based on the probability density function of the normal distribution, the probability of a random node belong to a dependence cluster of size $D$, $P(D)$, is approximated based on Eq. (3.6),

$$P(D) = \begin{cases} be^{-(D-(D-size))^2/2\sigma^2}, & 1 < D < 2(D\text{–}size) - 1 \\ 0, & \text{otherwise} \end{cases} \tag{3.6}$$

where $b$ is a normalization constant so that the sum of probabilities for all $D$ is one. The purpose of setting $P(D) > 0$ only for the range that $1 < D < 2(D\text{-}size) - 1$ is to have the distribution symmetrical around $D\text{-}size$. The assumption of dependence clusters contains two parameters, mean size of dependence clusters $D\text{-}size$ and variance of cluster size, $\sigma^2$.

First, it is assumed that the variance $\sigma^2$ is fixed, and then the impact of different values of $D\text{-}size$ is investigated on mixed cascading failures, shown as $G$ and $T$ in Figures 3.6 and 3.7. The networks used to conduct the simulations are ER random networks with $N = 1000$, $K = 10$, $\alpha = 1.05$ and $\mu = 0.5$.
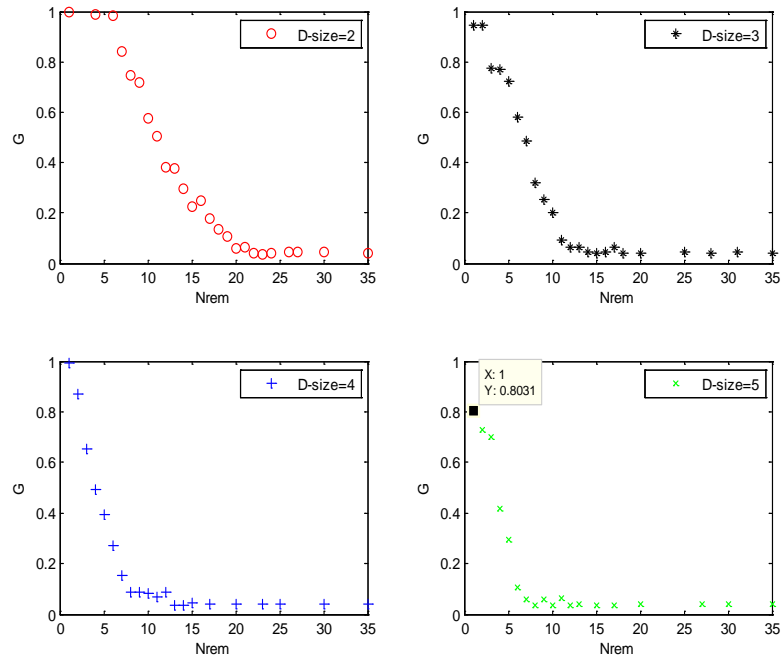
Figure 3.6. Simulation results of *G* for mixed cascading failures
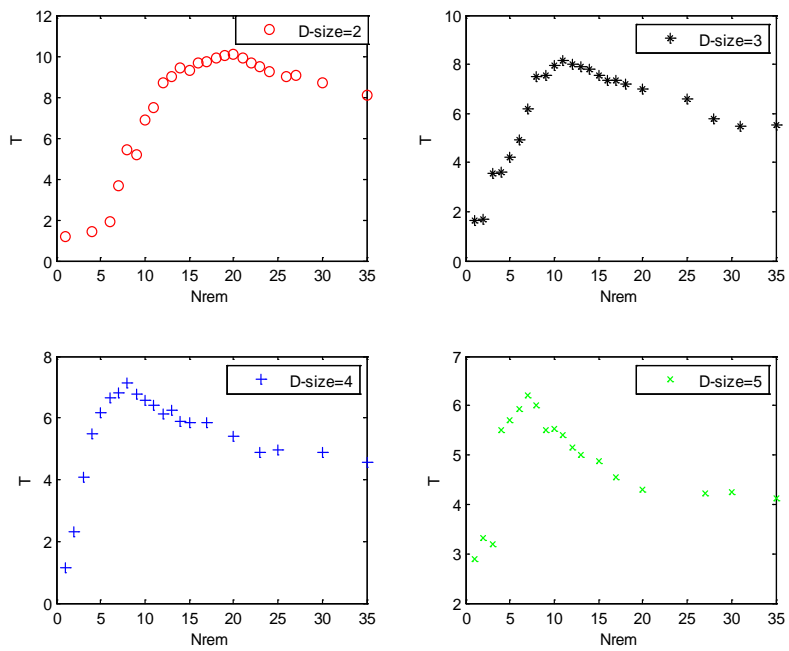


Figure 3.7. Simulation results of *T* for mixed cascading failures

Table 3.2. $T_c$ vs. *NremC* with different *D-size* with two cascading failure models

| | Normal distribution of dependence cluster | | | | Without dependence clusters |
|---|---|---|---|---|---|
| *D-size* | 2 | 3 | 4 | 5 | |
| *NremC* | 20 | 11 | 8 | 7 | 46 |
| $T_c$ | 10.1 | 8.15 | 7.125 | 6.2 | 20.75 |

Similar to the Poisson distribution analysis, the cascading process changes from second-order phase transition to first-order phase transition when considering dependence clusters approximated by a normal distribution (*D-size* > 1) based on the results of transition threshold *NremC*. *NremC* decreases as *D-size* increases, which means the network is less stable. The values of $T_c$ also become smaller as *D-size* increases, which means that the network collapses faster when caused by mixed cascading failures.

The impact of values of parameter $\sigma$ on mixed cascading failures in ER random networks ($N = 1000$, $K = 10$) is discussed, according to the results of evaluation parameters $G$ and $T$, similar to the analysis performed on the impact of *D-size* on mixed cascading networks. Figures 3.8 to 3.9 show the results of $G$ and $T$ vs. *Nrem* with different values of $\sigma$ with fixed *D-size*. NOTE: In the figures, $Qvar = \sigma$.
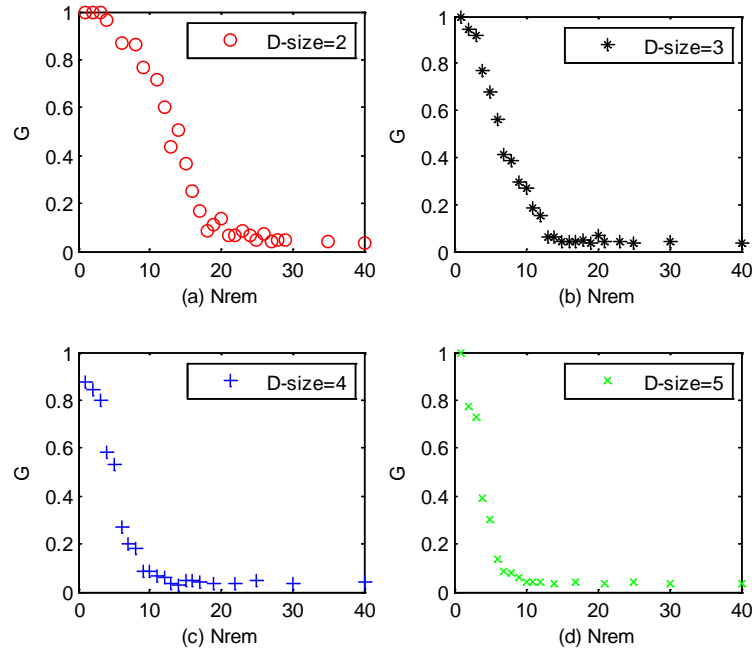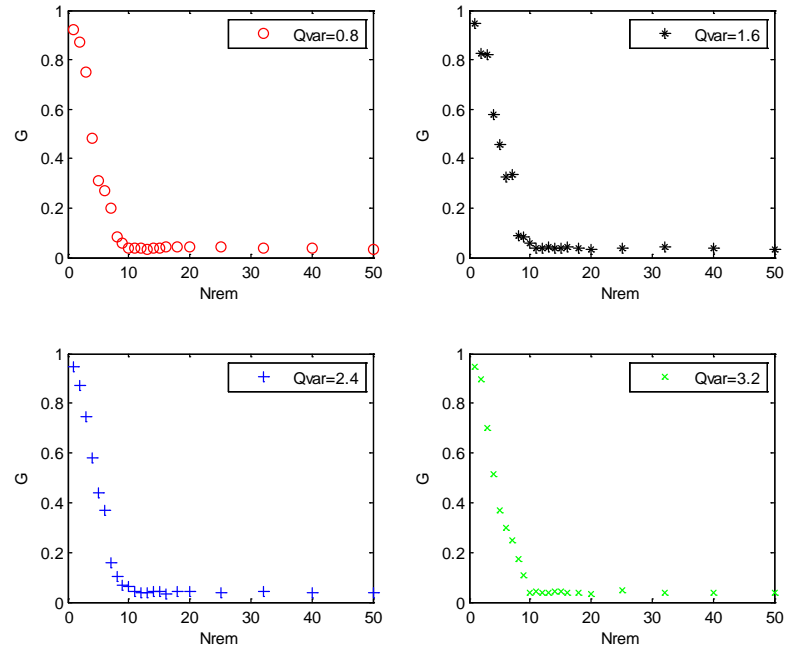
Figure 3.8. Simulation results of *G* for mixed cascading failures (*D-size* = 4)
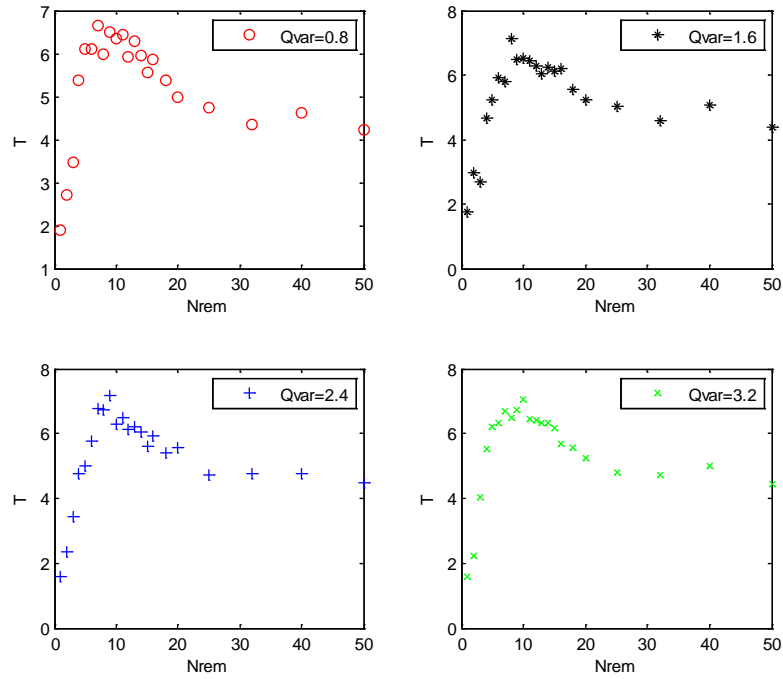


Figure 3.9. Simulation results of *T* for mixed cascading failures (*D-size* = 4)
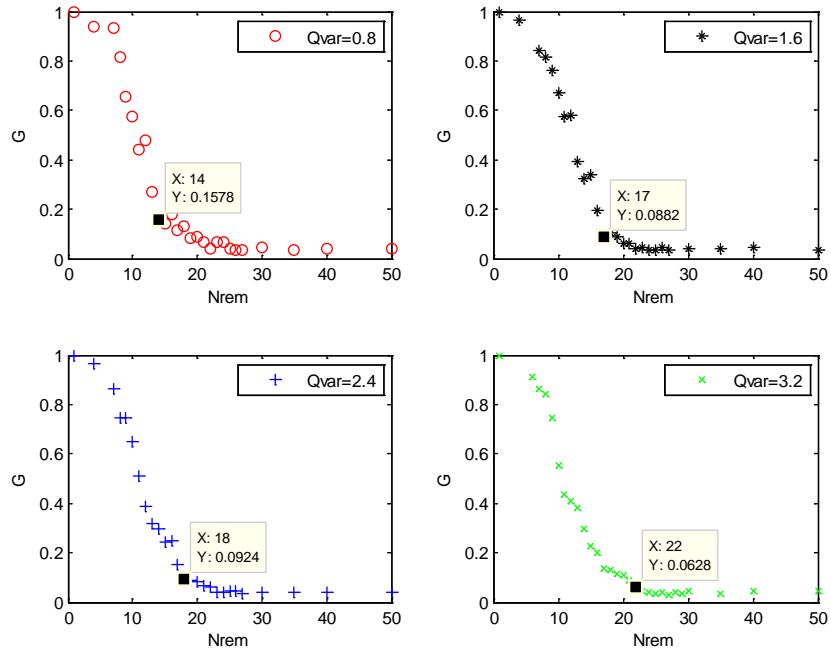
Figure 3.10. Simulation results of *G* for mixed cascades (*D-size* = 2)
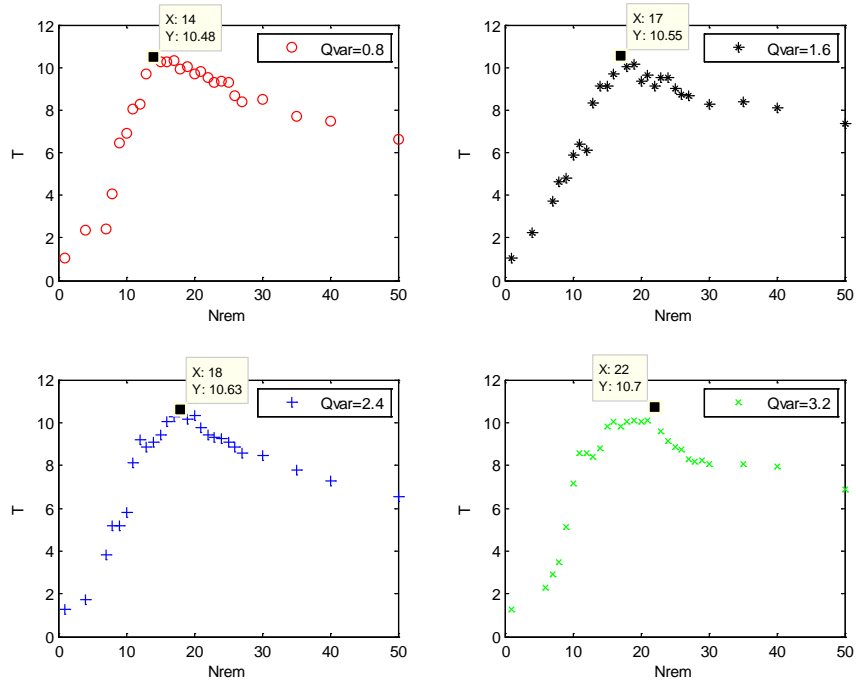


Figure 3.11. Simulation results of *T* for mixed cascades (*D-size* = 2)

Table 3.3. $T_c$ vs. *NremC* with different $\sigma$ and fixed *D-size*

| $\sigma$ | *D-size* = 4 | | | | *D-size* = 2 | | | | Without dependence clusters |
|---|---|---|---|---|---|---|---|---|---|
| | 0.8 | 1.6 | 2.4 | 3.2 | 0.8 | 1.6 | 2.4 | 3.2 | |
| *NremC* | 8 | 8 | 9 | 10 | 14 | 17 | 18 | 22 | 46 |
| $T_c$ | 6.625 | 7.125 | 7.175 | 7.05 | 10.475 | 10.55 | 10.625 | 10.7 | 20.75 |

Figures 3.8 and 3.9 shows the results of *G* and *T* as a function of *Nrem* for four values of $\sigma$, i.e., 0.8, 1.6, 2.4 and 3.2, with *D-size* = 4. It can be observed from the results that the process of cascading failures becomes a first-order phase transition with dependence clusters normally distributed. Figures 3.10 and 3.11 present the changing trends of *G* and *T* as a function of *Nrem* for four different values of $\sigma$ with *D-size* = 2. The process of cascades also becomes a first-order phase transition under this assumption of truncated normal distributions. The size of dependence clusters, which is normally distributed around mean size *D-size*, becomes more highly variable as $\sigma$ becomes larger. Combined with the previous results, it is observed that the first-order transition threshold, *NremC*, increases as $\sigma$ becomes larger. This indicates that the network becomes more stable.

After analysis of the results obtained for the normal distribution, there is an underlying relationship between the impact of variance $\sigma^2$ and mean size *D-size* on network robustness. Figure 3.12 shows the results of *NremC* vs. $\sigma$ under different *D-size* on ER random networks with $N = 1000$, $K = 10$.
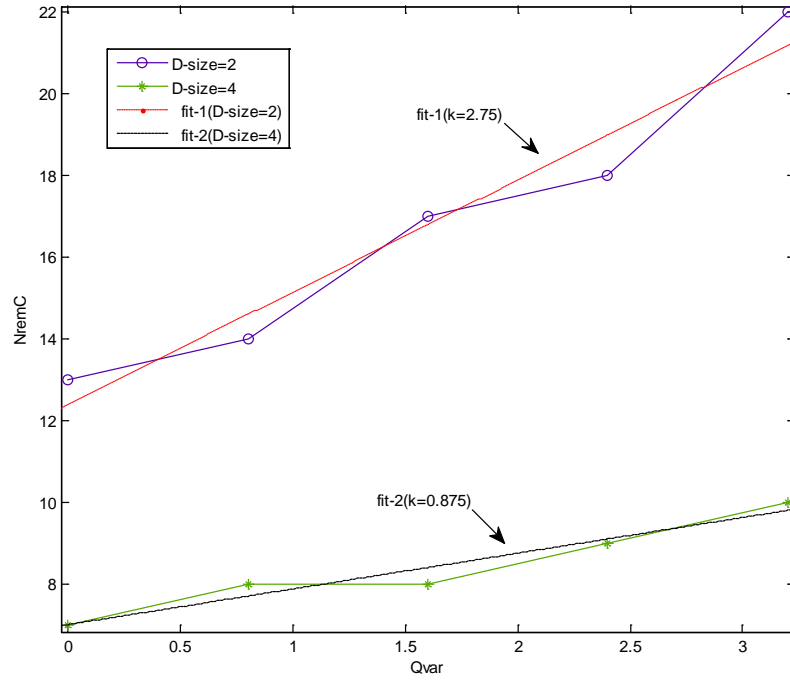
Figure 3.12. *NremC* vs. $\sigma$ with different *D-size* under mixed cascades

As shown in Figure 3.12, aside from the increase of $\sigma$ that contributes to larger *NremC*, different value of *D-size* also impacts the effect of $\sigma$ on network robustness. The two dotted lines, indicated as fit-1 and fit-2 in Figure 3.12, are the lines indicating the changing trends of *NremC* with two values of *D-size*, 2 and 4, respectively. The slope of fit-1 is $k = 2.75$ (for *D-size* = 2), while the slope of fit-2 is $k = 0.875$ (for *D-size* = 4). The effect of $\sigma$ on transition threshold *NremC*, that is the impact of $\sigma$ on network robustness to mixed cascading failures, becomes weaker for larger mean size *D-size*.

The impact of average degree of network, *K*, on network robustness with mixed cascading failures and normally distributed size of dependence clusters is also considered. The simulation results of *NremC* as a function of *D-size* with several values of *K* and fixed $\sigma$, on ER random networks ($N = 1000$) are shown in Figure 3.13.
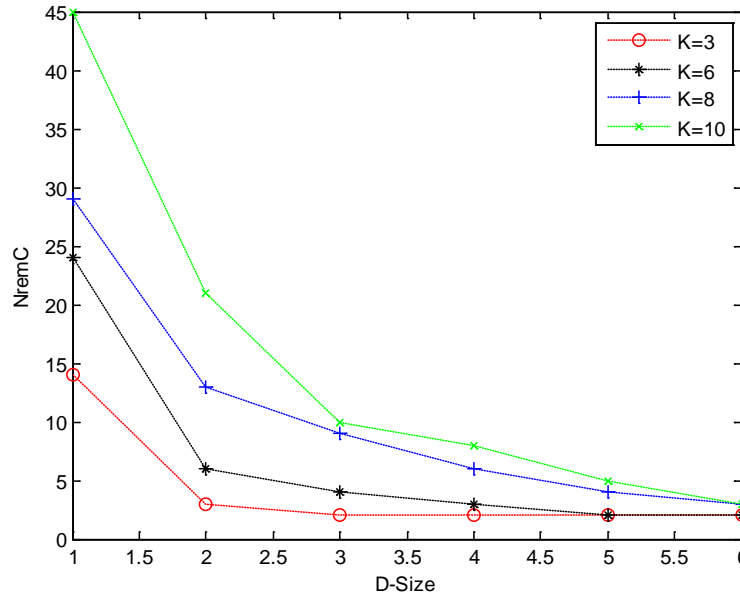
Figure 3.13. Transition points *NremC* vs. *D-size* of mixed cascades under different *K*

It can be seen from Figure 3.13 that the values of *NremC* are larger for ER random networks with larger average degree *K*, which indicates that ER random networks with higher average node degree possess a stronger robustness to mixed cascading failures.

The proposed mixed cascading failure model offers distinct advantages to investigate the combined impacts of network load dynamics and network dependency on failure propagation through network systems, and to explore specific effects of common types of network dependencies on network robustness.

### 3.1.4   Discussion

The combined effects of local load dynamics and node dependency on cascading failures are explored in this work. According to the mixed cascading failure model, cascading simulations are performed on a widely-recognized type of network, i.e., ER random networks. The model can be applied to properly simulate the cascades caused by load dynamics and node dependency, given that many previous models either only

considered load dynamics or the impact of node dependency, but not both. Moreover, statistical distributions are applied to establish dependence clusters of network nodes to depict node dependency, which is consistent with actual functioning networks [66]. Consequently, this model can be used as a basis for analyzing cascading failures and their impacts on network performance, which provide a better understanding of the process of failure propagation in real-life networked systems.

A more realistic simulation of cascading failures is developed so the results and findings are more general than previous research. Research findings show that the robustness of the network, considering dependent nodes with mixed cascading failures, becomes poor. This is a new conclusion for single-structure network with the new model of mixed cascading failures. This conclusion is consistent with the finding of previous papers which test on interdependence network systems. The network deteriorates rapidly in a form of a first-order phase transition when a critical number of nodes initially fail. Both the mean size of dependence clusters and average degree of nodes have a large impact on network robustness to cascading failures. A larger mean size of dependence clusters cause more harm to network robustness, while a larger average degree of node produce stronger network robustness. When the dependence clusters are approximated by a normal distribution, it is observed that mean size of clusters and variance of size of clusters are related to each other. Larger mean size of clusters minimizes the effect of a larger variance of clusters on improving network robustness. Findings from this research can be useful for a better understanding of the impact of important factors on cascading failures. For example, systems with higher average degree are more robust to mixed cascading failures than the networks with lower average degree. Thus average degree in

real network system is an important factor for a network to maintain connectivity when it is subject to cascading failures. This finding could be helpful when evaluating robustness of some real network systems to cascading failures.

In this work, ER random networks subject to mixed cascading failures that are initiated by random failures of nodes are investigated. Other general types of network topological structures could be considered to extend the study of the mixed cascading process. System failures all originated in nodes since the effect of dependence clusters of nodes on network robustness is analyzed. Another case is when failures originate from the links of the network, which exists in some real-world systems, and this could be investigated further. The study of mixed cascading failures triggered by intentional attacks, i.e., cascades triggered by malicious attacks on some important network nodes, is an- other direction for further research.

The assumption that immediate collapse of a dependence cluster occurs if any dependent nodes inside break down is made to depict the impacts of system dependency on failure propagation. It represents very strong dependency strength. They may not be the cases for some real-world applications. Therefore, system dependency strength regarding failure propagation is then described by the dependence cluster collapsing threshold (*CCT*) [124]. It means that a dependence cluster instantly collapses, i.e., all nodes belonging to this cluster break down if the percentage of failed dependent nodes belonging to this dependence cluster exceeds *CCT* (%). A smaller value of *CCT* indicates stronger dependency strength between network nodes, which means that dependent node failures in a dependence cluster have a great impact on other functional nodes in the same dependence cluster. The applications of this improved system dependency modeling on

resilience-based system restoration optimization against cascading failures are presented in Section 4.

## 3.2 Mixed cascading failure model combining local load and global load dynamics

Network system characteristics, system load dynamics and system dependency, which are taken into account in this research, have been proved to have great impacts on cascading failures. In terms of system load dynamics, for instance, electricity transmission in electric power systems influence system overloaded failure propagation process [90]. System load distribution mechanisms vary in different real-world systems, which is dependent on system characteristics, such as system services and system dependencies [125].

In the work that is presented in Section 3.1, system dependencies have been described by dependence clusters which could immediately collapse because of dependent component failures. However, it may not always be the case in reality. For example, system dependency might mainly impact the load sharing mechanism among the dependent nodes inside the dependency clusters. Thus, the propagation of system overload failures is influenced by system dependency indirectly.

Local load redistribution models and global load redistribution models have been used to describe how system load dynamics influence cascading failures [25, 36]. A local load redistribution model means failed system components mainly impact the functional neighboring components. For example, it assumes that failed node load will be locally redistributed so only neighboring components could be impacted and potentially failed due to overload. The local load redistribution model has been used in Section 3.1. The global load redistribution model assumes that system component failures could cause

disturbance to the functional components which are not subject to a localization constraint, i.e., component failures initiate global network load redistribution. These two load redistribution mechanisms can be simultaneously found in some real-world systems based on the system properties [126]. However, current cascading failure models rarely consider nor properly depict this phenomenon. In this section, a mixed cascading failure model that combines local load dynamics and global load dynamics considering the influence of system dependency are presented.

### 3.2.1 Preliminary cascading failure and recovery process

In the new model, specifically for global load redistribution, every time there are new node failures in a dependence cluster, part of the load on the failed nodes will be transferred to the functional nodes in the same dependence cluster. It depicts the impact of global load dynamics on cascading failures resulting from system dependency. Anytime when node load exceeds node capacity, system node fails under overload status. In addition, system nodes are not functional when they are disconnected from *LCC* during the cascading-restoration process. It is based on the percolation-like process that is observed in some real-world applications [127]. Figure 3.14 illustrates when system nodes become non-functional once they are disconnected from *LCC*.
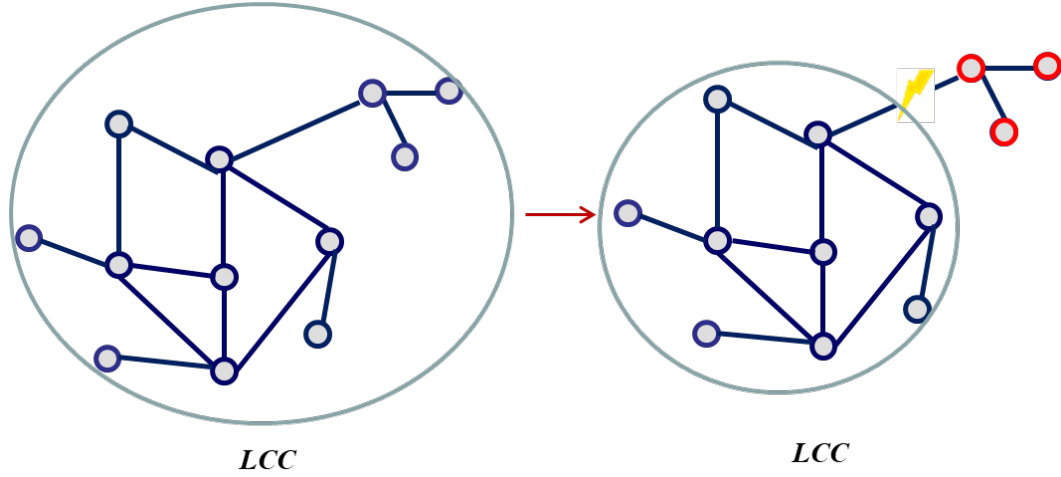
Figure 3.14. Functional failure occurs when system nodes are disconnected from *LCC*

According to the new model, mixed cascading failures are explored taking into account the joint impacts of local load dynamics and global load dynamics result from system dependency, as well as percolation feature.

When a failed system node $i$ is restored, some load is assumed to be transferred back to node $i$ from the functional neighboring nodes and working dependent nodes of node $i$ that have previously received the load from node $i$ when it failed.

This new cascading failure model, including the recovery process regarding load dynamics, captures some essential features of cascading failure process. It is helpful to understand the failure propagation mechanism and recovery process of network systems with load dynamics and multiple dependencies [128].

### 3.2.2   Cascading failure modeling with local load and passing by load

In the cascading failure modeling, different metrics have been applied to measure system load. For example, "betweenness centrality" [119] is used in the mixed cascading failure modeling presented in Section 3.1. CASCADE model is another popular model which is initially proposed to examine overload failure propagation in power

transmission system [129, 130].

In this work, network load on system node *i* at time *t*, $L_i(t)$, is assumed to consist of two parts, local customer load $L_{i,c}(t)$ and passing by load $L_{i,d}(t)$. This assumption is made based on the observation of some real-world network systems. In these network systems, system nodes are not only able to transfer load between each other, they also exchange load with local load points to provide services. A good example is electricity system. The centralized generation units (e.g., power plants) not only transfer electricity between each other, but also provide electricity to customers through local distribution systems. Therefore, it is reasonable to assume that passing by load and local customer load compose the total load on a system node. In some real-world network systems, local customer load of a system node is related to the amount of services provided by the node, while passing by load is related to node degree/connections. A hub node tends to process high passing by load.

Considering the electricity system, local customer load is assumed to depend on the number of customers connected to the node, and passing by load depends on the node degree. The reason lies in that more customer connections on a node tend to result in a larger amount of local customer load on this node, and higher node degree tends to contribute to higher passing by load. In total, the initial load on node *i* at time $t_0$ can be presented as follows,

$$L_i(t_0) = L_{i,c}(t_0)^{1-P_i^c} + L_{i,d}(t_0)^{1-P_i^d} \tag{3.7}$$

where $P_i^c$ denotes the normalized number of customers connected to node *i*. It includes all customers that receive the electricity provided by node *i*. $P_i^d$ denotes the normalized

number of node degree. It includes all neighboring nodes that have direct connections with node *i*. Similar to CASCADE model [129], the initial local customer load $L_{i,c}(t_0)$ and passing by load $L_{i,d}(t_0)$ on node *i* at time $t_0$ are assumed to follow certain distributions [90]. For example, both of them follow a truncated normal distribution with a specific load range [$L_{min}$, $L_{max}$] [131].

$$f(L_i(t_0)) = \begin{cases} \dfrac{e^{-(L_i(t_0)-\mu_l)^2/2\sigma_l^2}\Big/\sigma_l\sqrt{2\pi}}{\Phi\left(\dfrac{L_{max}-\mu_l}{\sigma_l}\right) - \Phi\left(\dfrac{L_{min}-\mu_l}{\sigma_l}\right)}, & L_i(t_0) \in (L_{min}, L_{max}) \\ 0, & \text{otherwise} \end{cases} \tag{3.8}$$

where $f(L_i(t_0)) \sim N(\mu_l, \sigma_l^2)$, and $L_i(t_0)$ represents $L_{i,c}(t_0)$ or $L_{i,d}(t_0)$.

Based on real-world network system features, instead of assuming network components are identical with the same capacity in CASCADE model, a nonlinear capacity-load model is adopted to quantify network node capacity [108].

As of system dependency modeling, network nodes belong to a dependency cluster when they provide services for the same customers. Dependency cluster size follows a certain distribution as mentioned in Section 3.1. The meaning of customers and dependency clusters vary in different real-world applications. In electricity system, customers are different facilities, such as residential and business. Dependency clusters are formed by the power generation units that provide electricity for the same customers through population centers/substations. Figure 3.15 illustrates the basic formation of an electricity system.
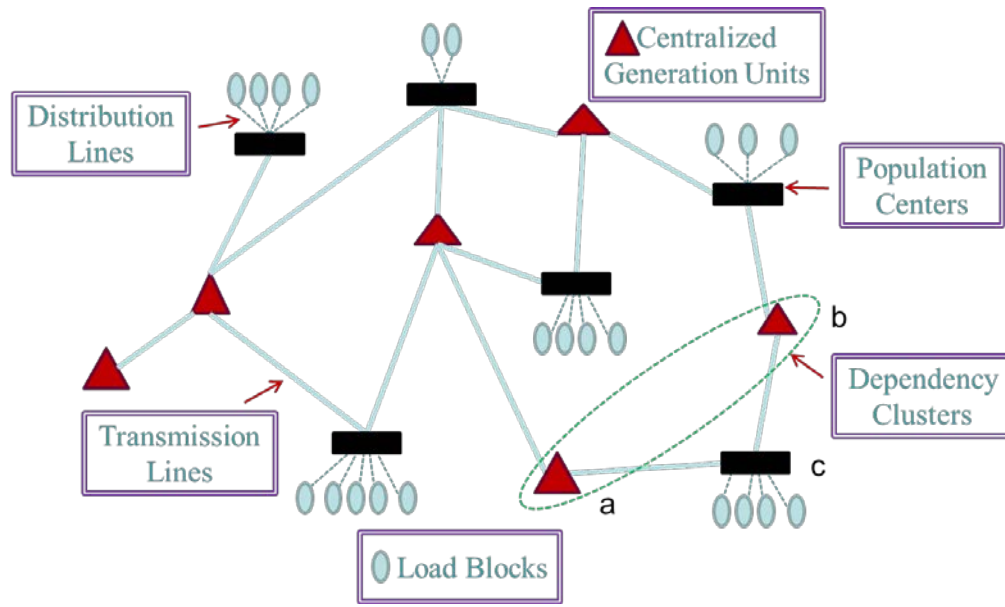
Figure 3.15. Main components of an electricity system

An example of a dependency cluster is indicated in Figure 3.15. There is no direct connection between two centralized generation units *a* and *b*. However, there is dependency relationship between them since they provide electricity for the same customers through population center *c*. Based on the proposed mixed cascading failure model, failure of any one of these two nodes would cause the increase of local customer load on the other node, as dictated by the load sharing mechanism. As can be seen, the dependent nodes in a dependence cluster are not limited by distance but rather depend on the existence of functional dependency relationship.

The model illustrated in Figure 3.15 for the electric power systems is a typical interdependent network system, because system dependencies are considered exist not only among system components, but also between the bulk system and power distribution systems. Electricity transmission is simulated as system load flow.

The cascading-restoration process regarding load dynamics can be described as follows,

$$\Delta L_{i,j,c}(t) = \frac{L_{i,c}(t)}{|k_{i,c}(t)|} \tag{3.9}$$

where $\Delta L_{i,j,c}(t)$ denotes the local customer load of failed node $i$ that is equally distributed to its working dependent node $j$ at time $t$ when node $i$ fails. $k_{i\,c}(t)$ is the set of functional dependent nodes of node $i$ at time $t$. $L_{i,c}(t)$ denotes the local customer load on node $i$ at time $t$ when it fails.

$$\Delta L_{i,j,d}(t) = \frac{L_{i,d}(t)}{|k_{i,d}(t)|} \tag{3.10}$$

Similarly, $\Delta L_{i,j,d}(t)$ denotes the passing by load of failed node $i$ that is equally distributed to its working neighbor node $j$ at time $t$ when node $i$ fails. Nodes $i$ and $j$ are neighbors if there is direct connection between them. $k_{i,d}(t)$ is the set of functional neighbor nodes of node $i$ at time $t$. $L_{i,d}(t)$ denotes the passing by load on node $i$ at time $t$ when node $i$ fails.

The working neighbors and working dependent nodes of node $i$, which received its load when node $i$ broke down, will transfer the load back to node $i$ when it is recovered. For example, node $i$ is restored at time $t'$ after it failed at time $t$. As introduced above, network load on an arbitrary system node $j$ at time $t'$, can be defined as $L_j(t') = L_{j,c}(t') + L_{j,d}(t')$. Local customer load $L_{i,c}(t')$ and passing by load $L_{i,d}(t')$ on a failed node $i$ which is recovered at time $t'$ can be presented as follows,

$$L_{i,c}(t') = \begin{cases} \Delta L_{i,j,c}(t) \, |\, NF_c \,| + \sum\limits_{m \in ND_c} p_m L_{m,c}(t'), & k_{i,c}(t) \neq \varnothing \\ L_{i,c}(t), & k_{i,c}(t) = \varnothing \end{cases} \tag{3.12}$$

where

$$NF_c = \left\{ n \mid L_{n,c}(t^{'}) > \Delta L_{i,j,c}(t) \cap n \in k_{i,c}(t) \right\}, \; ND_c = \left\{ m \mid 0 < L_{m,c}(t^{'}) \leq \Delta L_{i,j,c}(t) \cap m \in k_{i,c}(t) \right\}$$

.

$$L_{i,d}(t^{'}) = \Delta L_{i,j,d}(t) \mid NF_d \mid + \sum_{m \in ND_d} p_m L_{m,d}(t^{'}) \tag{3.13}$$

where $NF_d = \left\{ n \mid L_{n,d}(t^{'}) > \Delta L_{i,j,d}(t) \cap n \in k_{i,d}(t) \right\}$,

$ND_d = \left\{ m \mid 0 < L_{m,d}(t^{'}) \leq \Delta L_{i,j,d}(t) \cap m \in k_{i,d}(t) \right\}$.

$p_m$ is uniformly distributed in (0,1), which decides the amount load transferred back to the recovered node from node $m$. Note that the load on the functional neighbor nodes and functional dependent nodes, which transfer load back to the recovered nodes, is reduced accordingly.

Overall, a new mixed cascading failure model is developed considering the joint impacts of system local load redistribution and global load redistribution that pertain to system dependency. Instead of causing immediate dependence cluster collapse as in section 3.1, system dependency accelerates overload failure propagation indirectly by impacting load sharing in this model. Node load consists of local customer load and passing by load with distinct dynamic distribution mechanisms based on network system characteristics. The new mixed cascading failure model is applied to extend the influence of cascading failures in power transmission system to local customers. It also helps to investigate the interplay between the effects of system hardening strategies and restoration strategies against cascading failures. The applications of this new mixed cascading failure model are presented in section 6.

# 4  Resilience-based restoration selection against cascading failures

In this section, the effects of restoration strategies on recovering network systems from cascading failures are investigated considering the impacts of multiple network system properties. Specifically, the mixed cascading failure model presented in section 3.1 is applied while the impact of system dependence on failure propagation is improved based on real cases. A new system resiliency metric for resiliency loss evaluation, which is time-dependent, performance-based and normalized, is proposed in this work. Two other system performance measurements, system connectivity $G$ and recovery time $T$ are used to evaluate restoration effects. Four restoration strategies in terms of different restoration priority targets are applied to two typical synthetic network models and a real-world network system in the case studies for validation. The influence of system dependency characteristics and the relationship among them and restoration effects on system resiliency against cascading failures are explored. Then, a recovery framework against cascading failures is proposed, which considers the changing importance of system components during the cascading-recovery process for optimizing resource allocation. The recovery framework combines global network connectivity importance with other restoration preference for updating restoration priority targets. Efficient detection of system node importance regarding global network connectivity is provided with graph embedding and deep reinforcement learning. Overall, the results and findings provide references about the selection of resilience-based restoration strategy for real-world network system applications against cascading failures.

**4.1 System modeling and dependence impact**

As introduced in section 3.1, dependence clusters are proposed to represent the dependence relationships of network nodes apart from the topological connections. Since multiple dependencies can exist between network components, they may result in various dependence strengths, which influence cascading failures [59]. In this section, dependency strength is investigated by introducing dependence cluster collapsing threshold, *CCT* [124]. A dependence cluster instantly collapses, i.e., all nodes belonging to this cluster break down, once the proportion of failed nodes belonging to this dependence cluster exceeds *CCT*. Smaller *CCT* means stronger dependency strength among network nodes. It means that failed nodes in a dependence cluster have more significant impact on the functional nodes belonging to the same dependence cluster. The size of system dependence clusters is assumed to follow a shifted/scale adjusted Poisson distribution. The probability of a network node belonging to a dependence cluster of a specific size is determined according to Eq. (3.5).

The mixed cascading failure model proposed in section 3.1, which includes two types of network failures: overloaded failures caused by load dynamics and dependence failures caused by dependence node clusters, is employed in this section. First, some system components break down because of initial interruptions. These failed components can be considered as if they have been removed from the whole network. The network topology then changes, and so does the shortest paths between nodes that network load passes through. Network load carried by the failed components is redirected to nearby functional components. This redistribution may cause these functional nodes to exceed their capacity resulting in overloaded failures. At the same time, immediate collapses of

dependence clusters may result from overloaded failures once *CCT* is reached. In return, dependence failures also accelerate load dynamic redistribution, which triggers more overloaded failures. In this way, two types of failures iteratively propagate through the network system.

In the network system, $ed_{ij}$ depicts the topological connection between system node $i$ and node $j$. $ed_{ij}$ is assumed to be 0 or 1. $ed_{ij} = 1$ if there is a direct topological connection between node $i$ and node $j$. While $ed_{ij} = 0$ indicates that there is no edge between these two nodes. It is assumed that network load transmits on edges (topological connections) that form the network topology, and it only transmits along the shortest paths between every pair of network nodes. A path consists of edges (topological connections) between the two targeted nodes. Thus, the length of a path equals the number of edges along this path. If there is no path between a selected pair of nodes, the distance between the two nodes is infinite. $E_{ij}$, defined in Eq. (4.1), is used to depict the efficiency of the shortest path between node $i$ and node $j$, i.e., the most efficient path regarding load transmission.

$$E_{ij} = \left( \sum \frac{1}{ed_k} \right)^{-1} \tag{4.1}$$

Network efficiency $E(G_P)$ defined in Eq. (4.2) is used as the system performance to evaluate network recovery level in terms of transmission efficiency [132].

$$E(G_p) = \frac{1}{N(N-1)} \sum_{i \neq j \in G_p} E_{ij} \tag{4.2}$$

Higher $E(G_p)$ indicates more efficient transmission of network load. Network load is presented by "betweenness centrality", which is calculated by Eq. (3.1). Network

connectivity is measured by $G$, which is defined in Eq. (3.4). $G$ decreases when failures occur, and the network falls apart. $G$ equals 1 if no failures happen in the network. Here, $T$ denotes the time to restore a network system from cascading failures to the predetermined level. A shorter time $T$ is better.

Node capacity is determined according to Eq. (3.2). Based on the mixed cascading failure model proposed in section 3.1, a node becomes overloaded and breakdown once the load on a node exceeds node capacity. The load on the failed node is preferably redistributed to the functional nodes in its local neighborhood nodes.

The main steps to model the process of mixed cascading failure propagation and restoration implementation are briefly introduced as follows:

Step 1)    Total $N$ network nodes are initially functional with capacity allocated. All dependence clusters in the network are determined.

Step 2)    Randomly select system nodes to break down due to initial interruptions.

Step 3)    Dependence clusters collapse if $CCT$ is exceeded. Update network topology.

Step 4)    Network load are dynamically redistributed over current network structure. Overloaded nodes fail.

Step 5)    Failed nodes are selected to start repair activity with probability p with a required repair time.

Step 6)    Go back to Step 3, until network is recovered to the predetermined level.

Network performance measurements are collected during the cascading-restoration process.

**4.2 Resilience measurement and restoration strategies**

Although some resilience metrics have been presented to describe system resilience in different areas, there are no standardized metrics for system resilience, especially for measuring electric power system resilience. A new resilience metric, $\Re(t)$, which measures system resilience loss due to system failures is proposed. Figure 4.1 shows the changing trend of one type of system quality measures, $Q(t)$, which evolves as time goes on when cascading failures occur in the system and restoration actions are performed as well.
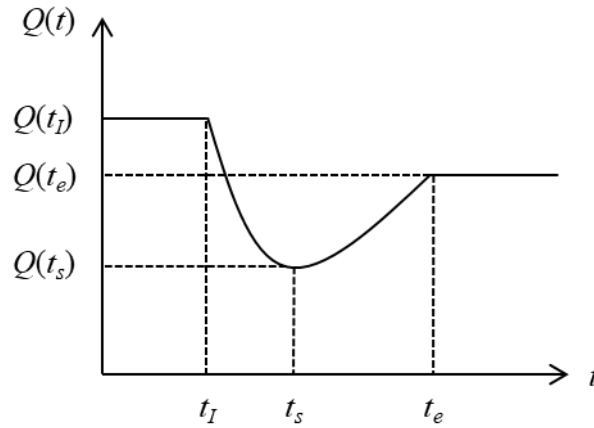


Figure 4.1. Changing trend of quality measure $Q(t)$ after failures occur and restoration

implementation

A disruptive event happens at time $t_I$. It triggers cascading failures in the network system, which result in the degradation of quality measure $Q(t)$. After conducting restoration strategies for a while, $Q(t)$ is recovered to a predetermined level at time $t_e$. In general, $Q(t_I) \geq Q(t_e)$, where $Q(t_I)$ denotes the initial level of $Q(t)$ without any interruptions and $Q(t_e)$ denotes the recovered steady level of system performance. System resilience loss up to time $t$ is measured by a time-dependent metric, $\Re(t)$. $\Re(t)$ is defined

as the proportion of the lost quality measure $Q(t)$ resulted from cascading failures in the system with respect to a comparative $Q(t)$ if there is no failure up to time t. $\Re(t)$ is formulated in Eq. (4.3).

$$\Re(t) = \begin{cases} 0, & t \le t_I \\ \dfrac{\int_{t_I}^{t} \left(Q(t_I) - Q(t)\right)dt}{Q(t_I)(t - t_I)}, & t > t_I \end{cases} \qquad (4.3)$$

where $0 \le \Re(t) \le 1,\ t \in [0, t_e]$. It is clearly that smaller $\Re(t)$ indicates less resilience loss.

First, this new metric considers the time that the system stays in each stage since system resilience changes, such as the stage that cascading failures occur which cause system performance degrading, and the stage that restoration actions are taken which lead to system performance recovering. Second, this system resilience metric is related to the measurement of a specific system performance, $Q(t)$. Because system performance can be measured from different perspectives and resilience is also a type of system property, the considered system performance should be reflected from system resilience metrics. In the end, this resilience metric is quantitative and normalized between [0, 1], which is important to make resilience of different systems comparable and make system resilience to be understandable. Thus, this metric can be applied to different types of systems for resilience analysis given that there is no generally accepted resilience metric currently. In this section, network load is adopted as $Q(t)$ to calculate resiliency loss $\Re(t)$ with respect to system load demand and supply capability.

As mentioned above, four different restoration strategies are adopted to recover the system when cascading failures occur. They are introduced as follows,

(1) Random repair strategy (RR): RR is the default restoration strategy and is applied for

comparison purposes. The failed network nodes, where repair activity has not yet started, are selected to be repaired at each round of inspection based on repair proportion $R_p$. It denotes the proportion of failed components, whose repairing activities have not yet begun, that can start to be repaired;

(2) High degree first repair strategy (HDFR): HDFR is targeted for the repair order assigned according to the degree of failed nodes. Node degree denotes the number of topological connections that a node has with other nodes. Failed components, of which repairing has not yet started, are repaired in descending order of node degree, i.e., network nodes with higher node degrees are repaired with higher priority.

(3) Short time first repair strategy (STFR): STFR involves the repair order assigned based on the required repair time of failed nodes. The restoration prioritizes the failed nodes that require shorter repair time.

(4) High load first repair strategy (HLFR): HLFR strategy prioritizes repair according to the amount of load carried by the failed nodes. Failed nodes of which repair activity has not yet started at inspection are repaired in descending order of the amount of load passed through.

Note that there must be an exact ordered list of repair actions, so ties, which occur when the failed nodes, of which the repairing has not started, at the inspection have the same condition (the same degree, the same required repair time or the same load) are broken according to the first fail first repair policy. The repair activity of failed nodes, once started, will not stop until they are completed in this work. For simplicity, it is assumed that the restoration resources are available to do so. The number of new repair activities of failed nodes, which get started at each round of inspection for different

restoration strategies, is decided by the repair proportion $R_p$ and the total number of failed nodes, where repair activity has not started yet.

## 4.3 Numerical examples on synthetic networks and real-world network system

In this section, the mixed cascading failure model is performed on two synthetic network models and one real-world network topology, respectively. In this way, the impacts of network dependency characteristics and load dynamics on restoration effects regarding system resiliency are investigated.

In this work, the repair proportion $R_p$ is set to be 0.7, i.e., the proportion of failed nodes whose repair activities have not yet started can be selected to start repairing is 0.7 at each round of inspection. For generalization, multiple simulations are conducted under the same condition to calculate the average of resilience loss and other system performance measurements. A simulation process includes failure propagation and restoration implementation. Simulation data is collected since cascading failures happen until network efficiency is recovered to 95% of its initial level, i.e., it denotes the predetermined recovery level. In this work, the required repair time of each failed node is randomly selected between 1 to 3 time units, i.e., time steps in the simulation. In the model, the time to restore a network system to the predetermined level, $T$, is represented by the total number of iterations of the simulated failure propagation process under restoration implementation, until reaching the predetermined recovery level. Each simulation iteration represents a fixed duration of time.

Cascading failures are triggered by randomly selecting failures on network nodes as initial failures. The triggering mechanism depicts random failures of system components to simulate exposure to some extreme event that causes multiple failures. Different

numbers of initially failed nodes (*Nrem*) are considered in this work, which are 9, 15, 21, 27, 33, 39, 45, 51, 57 and 63. To minimize random errors, simulation results for a specific *Nrem* correspond to the average of results over many realizations of randomly selecting nodes to fail.

### 4.3.1 Example of U.S. top 500 airport network

A case study is first conducted on the U.S. top 500 airport network topology [133]. This real-world network system comprises of 500 nodes and 2980 edges. Nodes denote airports and edges denote the connections between airports. 30 simulations are performed to obtain averages of the resiliency metric, resiliency loss, and other system performance measurements. Four restoration strategies are applied to the network system, respectively, to analyze the impacts of dependency characteristics and load dynamics on restoration effects against mixed cascading failures. Figure 4.2 illustrates the changing trend of network efficiency for the U.S. top 500 airport network after simulated cascading failures occur and implementing restoration.
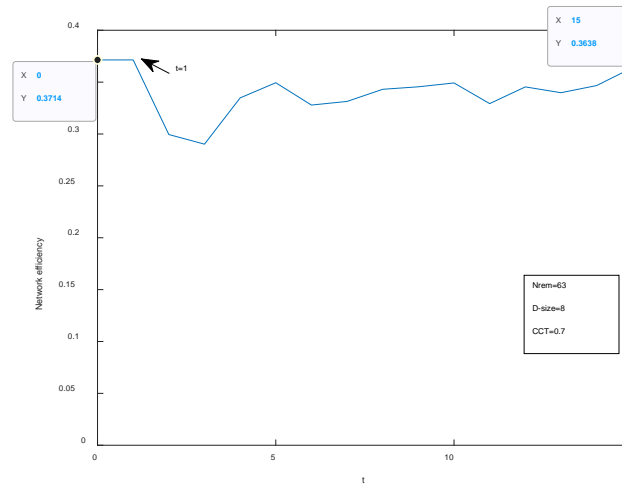


Figure 4.2. Network efficiency vs. time after cascading failures occur and RR restoration implementation. *Nrem* = 63, *D-size* = 8, *CCT* = 0.7.

Based on Figure 4.2, it can be seen that initial failures occur at time $t = 1$, at the same time, restoration strategy is performed on the network system. Initially, network efficiency quickly declines. Then it rises with fluctuations as time goes on. The fluctuation of network efficiency during the process indicates network performance loss due to cascading failures. The results shown in Figure 4.2 are in accordance with that in Figure 4.1.

**1) Test on random repair strategy:**

First, the results of resiliency metric and system performance measurements regarding mixed cascading failures under RR strategy are presented. Figure 4.3 presents average resilience loss and *G* vs. *Nrem* in ER networks with different *CCT* and *D-size*.
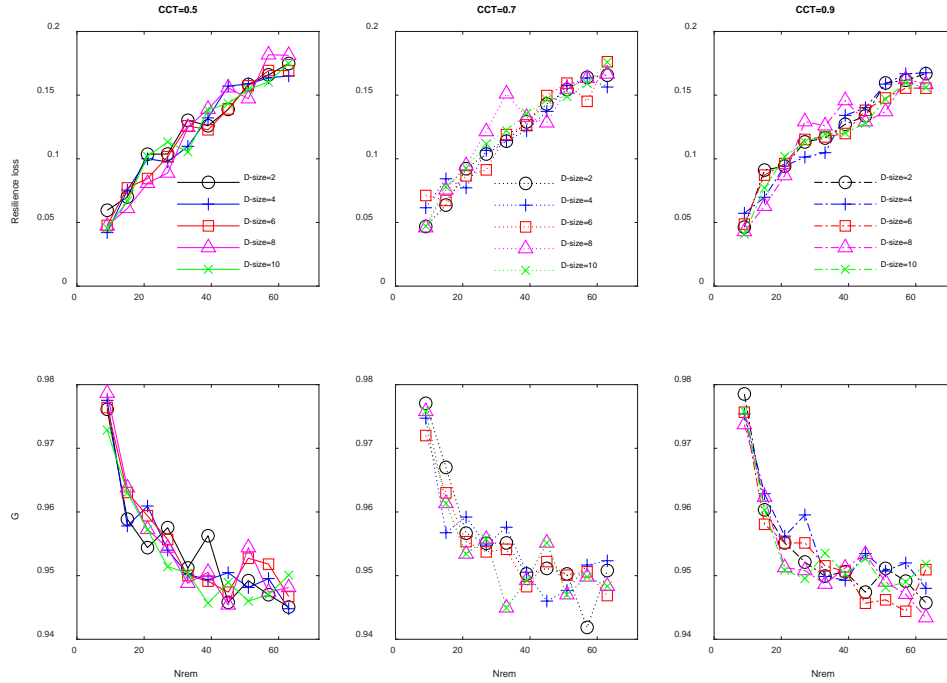


Figure 4.3. Average results of resiliency loss and *G* as a function of *Nrem*. Note: Scenarios are considered with different *CCT* and *D-size*. *CCT* = 0.5, 0.7, 0.9, *D-size* = 2, 4, 6, 8, 10, respectively.

According to Figure 4.3, resiliency loss increases as *Nrem* increase. This trend slows down as *CCT* increases when *Nrem* is large. It can be explained that strong dependency strength, indicated by small *CCT*, can accelerate failure propagation. The three lower subfigures show that *G* decreases with an increasing *Nrem*. It means that larger *Nrem* triggers more devastating cascading failures, which results in more damage to system connectivity. While it can be observed that the rate of *G* declining decreases as *Nrem* increases.

Table 4.1 presents the average results of resiliency loss and *T*. Based on Table 4.1, it can be seen that the smallest resiliency loss is achieved when *CCT* = 0.9, while *D-size* does not show noticeable impacts on resiliency loss or *T*.

Table 4.1. Average resiliency loss and *T* for different *Nrem*

| Resiliency Loss (%) | *D-size* | | | | | Aver-age | *T* | *D-size* | | | | | Aver-age |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 4 | 6 | 8 | 10 | | | 2 | 4 | 6 | 8 | 10 | |
| *CCT* = 0.5 | 12.3 | 12.0 | 12.0 | 12.1 | 12.1 | 12.1 | *CCT* = 0.5 | 3.5 | 3.4 | 3.5 | 3.1 | 3.4 | 3.4 |
| *CCT* = 0.7 | 11.8 | 11.8 | 11.9 | 12.3 | 12.2 | 12.0 | *CCT* = 0.7 | 3.4 | 3.4 | 3.4 | 3.7 | 3.7 | 3.5 |
| *CCT* = 0.9 | 12.1 | 12.0 | 11.8 | 11.8 | 11.6 | 11.9 | *CCT* = 0.9 | 3.4 | 3.6 | 3.4 | 3.2 | 3 | 3.3 |
| Average | 12.1 | 11.9 | 11.9 | 12.1 | 12.0 | 12.0 | Average | 3.4 | 3.5 | 3.4 | 3.3 | 3.4 | 3.4 |

As mentioned above, every point in the plotted curves in the figures in this case study are the simulation results corresponding to an average of 30 random initial failure triggers. It should be noted that since the curves shown in the figures are average simulation results, those curves may not behave smoothly like theoretical results. The atypical behavior of some points results from the randomness associated with simulation.

2)  **Test on high degree first repair strategy:**

The results obtained by applying HDFR strategy are presented in Figure 4.4.

Scenarios with different *CCT* and *D-size* are considered. Figure 4.4 shows the average results of resiliency loss and *G*.
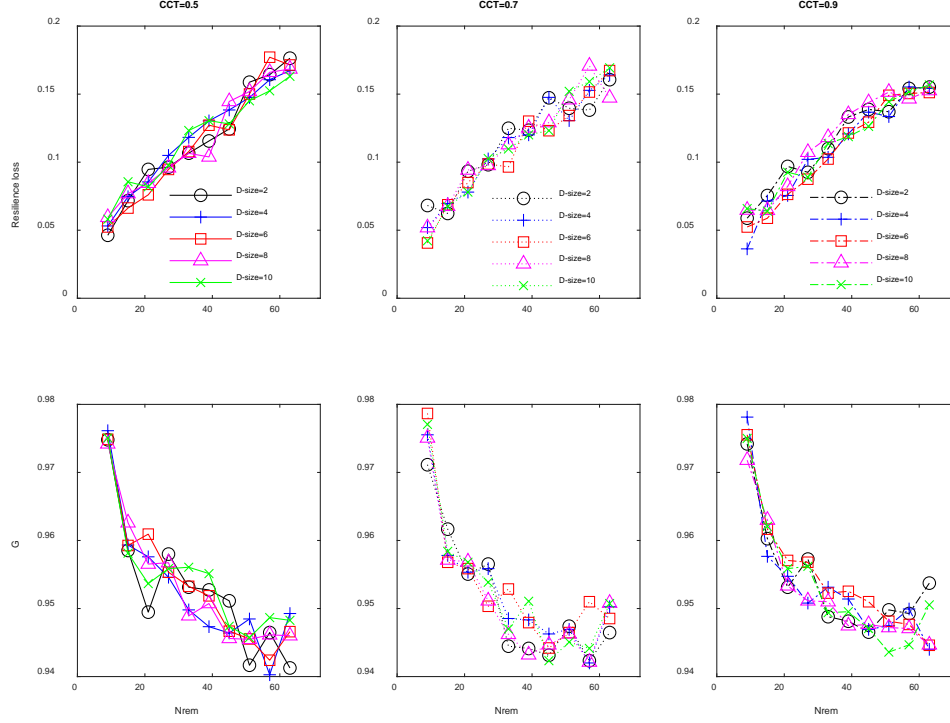


Figure 4.4. Average results of resiliency loss and *G* as a function of *Nrem* for HDFR

Based on three upper subfigures in Figure 4.4, resiliency loss increases when *Nrem* increases, while the trend is abating when *CCT* increases. It can be observed that the smallest resiliency loss is obtained when *CCT* = 0.9 compared to that when *CCT* = 0.5 or *CCT* = 0.7 with large *Nrem*. As to *G*, it shows the same trend as is observed for RR strategy. *G* decreases as *Nrem* increases while the rate of decline for *G* slows down.

Table 4.2 presents the average results of resiliency loss and *T* as follows. The results show that resiliency loss would be aggravated by stronger dependence strength indicated by smaller *CCT*, while *D-size* does not simply have a monotonic influence on system resiliency.

Table 4.2. Average of resiliency loss and *T* for different *Nrem*

| Resiliency Loss (%) | *D-size* | | | | | Average | *T* | *D-size* | | | | | Average |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 4 | 6 | 8 | 10 | | | 2 | 4 | 6 | 8 | 10 | |
| *CCT* = 0.5 | 11.6 | 11.8 | 11.5 | 11.6 | 11.7 | 11.6 | *CCT* = 0.5 | 2.6 | 2.6 | 2.6 | 2.6 | 2.6 | 2.6 |
| *CCT* = 0.7 | 11.6 | 11.4 | 11.0 | 11.4 | 11.2 | 11.3 | *CCT* = 0.7 | 2.5 | 2.5 | 2.4 | 2.6 | 2.5 | 2.5 |
| *CCT* = 0.9 | 11.5 | 10.9 | 10.8 | 11.7 | 11.2 | 11.2 | *CCT* = 0.9 | 2.7 | 2.4 | 2.5 | 2.7 | 2.5 | 2.6 |
| Average | 11.6 | 11.4 | 11.1 | 11.6 | 11.4 | 11.4 | Average | 2.6 | 2.5 | 2.5 | 2.6 | 2.5 | 2.6 |

### 3) Test on short time first repair strategy:

Average resiliency loss and *T* for different *CCT* and *D-size* with STFR strategy are presented in Figure 4.5. Both resiliency loss and *T* increases as *Nrem* increases. It means that larger *Nrem* would aggravate cascading failures. In addition, thresholds can be observed from the changing trends of resiliency loss and *T* as *Nrem* increases. The growth rate of resiliency loss and *T* decreases when *Nrem* exceeds a certain threshold.

Average results over different *Nrem* are presented in Table 4.3. The results in Table 4.3 demonstrate that larger *CCT* contributes to smaller resiliency loss and shorter *T* when STFR strategy is implemented.
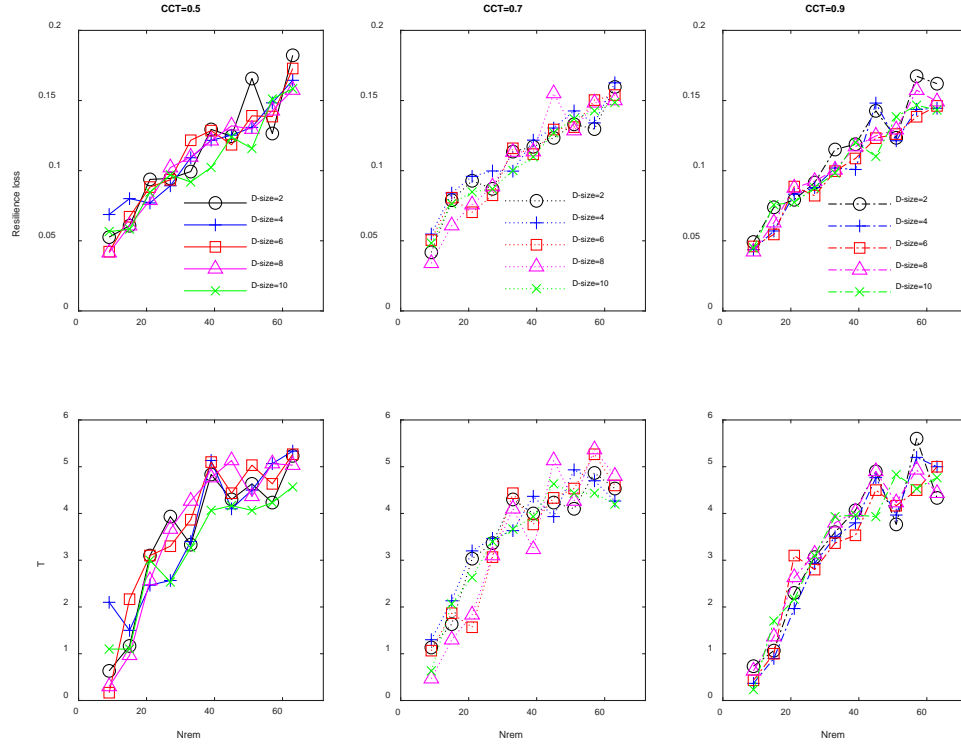
Figure 4.5. Average results of resiliency loss and *T* as a function of *Nrem*. Note: Different

*CCT* and *D-size* are considered, i.e., *CCT* = 0.5, 0.7, 0.9, *D-size* = 2, 4, 6, 8, 10,

respectively.

Table 4.3. Average results of resiliency loss and *T* for different *Nrem*

| Resiliency | *D-size* | | | | | Aver- | *T* | *D-size* | | | | | Aver- |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Loss (%) | 2 | 4 | 6 | 8 | 10 | age | | 2 | 4 | 6 | 8 | 10 | age |
| *CCT* = 0.5 | 11.3 | 11.2 | 11.1 | 10.8 | 10.4 | 11.0 | *CCT* = 0.5 | 3.5 | 3.6 | 3.7 | 3.6 | 3.2 | 3.5 |
| *CCT* = 0.7 | 10.8 | 11.3 | 10.8 | 10.7 | 10.6 | 10.8 | *CCT* = 0.7 | 3.5 | 3.6 | 3.5 | 3.4 | 3.4 | 3.5 |
| *CCT* = 0.9 | 11.2 | 10.3 | 10.1 | 10.6 | 10.4 | 10.5 | *CCT* = 0.9 | 3.3 | 3.2 | 3.2 | 3.4 | 3.3 | 3.3 |
| Average | 11.1 | 10.9 | 10.7 | 10.7 | 10.5 | 10.8 | Average | 3.4 | 3.5 | 3.5 | 3.5 | 3.3 | 3.4 |

**4) High load first repair strategy:**

The trends of average *G* and *T* with increasing *Nrem* under HLFR strategy are presented in Figure 4.6. Scenarios with different assumptions of *CCT* and *D-size* (*CCT* = 0.5, 0.7, 0.9, *D-size* = 2, 4, 6, 8, 10, respectively) are investigated.
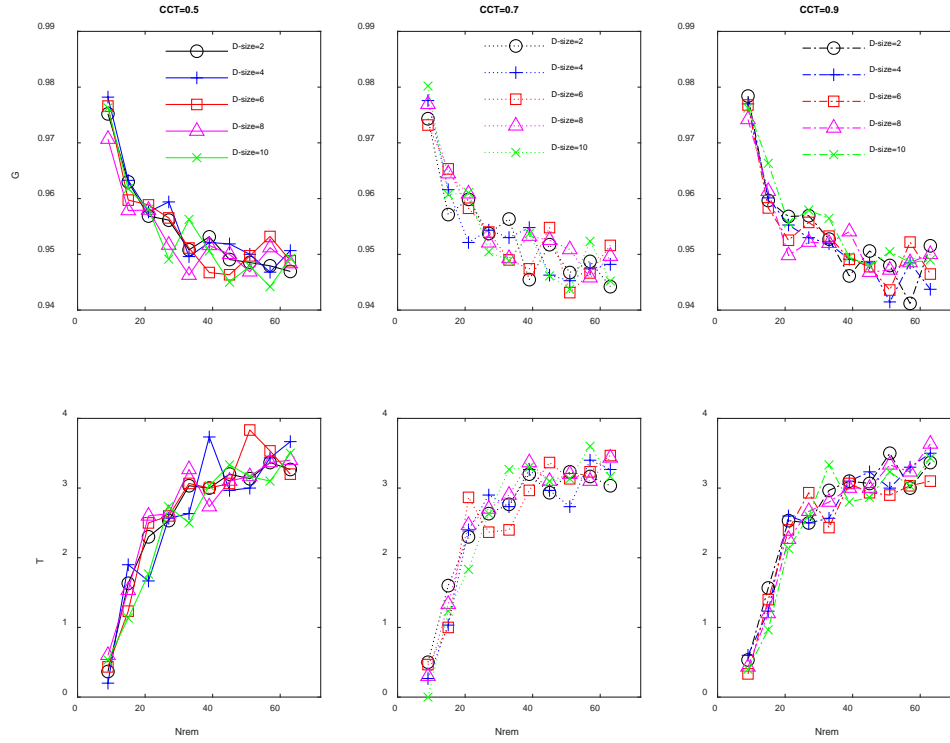


Figure 4.6. Average results of *G* and *T* as a function of number of *Nrem*

According to Figure 4.6, the changing trends of *G* and *T* as *Nrem* increases are in accordance with what was observed previously. There is a noticeable threshold where the growth rate of *T* gets reduced when *Nrem* exceeds. The thresholds in the changing trend of *G* and the changing trend of *T* are almost the same, i.e., *Nrem* is around 20.

The average results for different *Nrem* are presented in Table 4.4. Based on Table 4.4, average resiliency loss reduces as *CCT* increases, whereas the impacts of *D-size* on resiliency loss and *T* do not show a monotonic behavior.

Table 4.4. Average of resiliency loss and *T* for different *Nrem*

| Resilien cy Loss (%) | D-size | | | | | Avera ge | T | D-size | | | | | Avera ge |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 4 | 6 | 8 | 10 | | | 2 | 4 | 6 | 8 | 10 | |
| CCT = 0.5 | 11.3 | 11.2 | 11.5 | 11.4 | 10.7 | 11.2 | CCT = 0.5 | 2.6 | 2.6 | 2.6 | 2.6 | 2.5 | 2.6 |
| CCT = 0.7 | 10.6 | 11.0 | 11.2 | 11.3 | 11.5 | 11.1 | CCT = 0.7 | 2.5 | 2.5 | 2.5 | 2.6 | 2.5 | 2.5 |
| CCT = 0.9 | 11.3 | 11.5 | 10.6 | 11.1 | 10.7 | 11.0 | CCT = 0.9 | 2.6 | 2.6 | 2.5 | 2.6 | 2.5 | 2.6 |
| Average | 11.1 | 11.2 | 11.1 | 11.3 | 11.0 | 11.1 | Average | 2.6 | 2.6 | 2.5 | 2.6 | 2.5 | 2.6 |
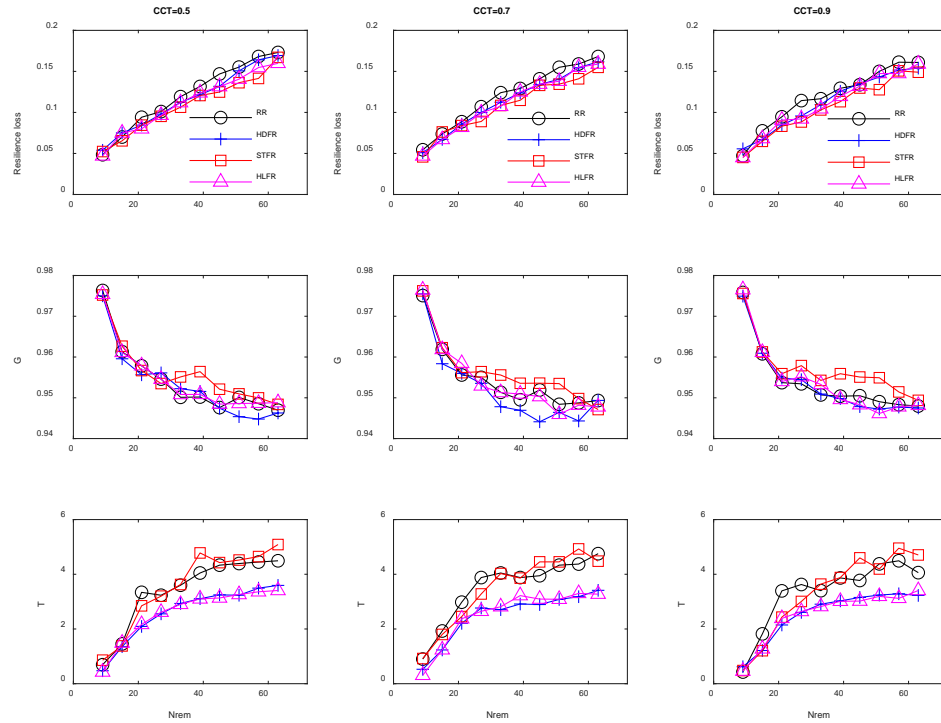


Figure 4.7 Average resiliency loss, *G* and *T* over *D-size* vs. *Nrem* under four restoration

strategies

Figure 4.7 presents the average resiliency loss, *G* and *T* over different *D-size* assumptions (*D-size* = 2, 4, 6, 8, 10) for three *CCT* (*CCT* = 0.5, 0.7, and 0.9) for the four restoration strategies. It can be seen from Figure 4.7 that the largest resiliency loss is

obtained by using RR strategy. The smallest resiliency loss, and almost the best system connectivity (the biggest $G$), are achieved by conducting STFR strategy. Whereas the longest repair time $T$ is incurred by adopting STFR strategy. The shortest repair time $T$ can be achieved by employing HDFR strategy or HLFR strategy.

### 4.3.2 Example of ER random network model

In this section, numeric examples are conducted on ER network models. In order to make a comparison with the results obtained from the U.S. top 500 airport network, the ER networks used in this work contain 500 nodes and average node degree $K$ of 11.92. Note that all the following results from ER networks are averaged over 90 realizations.

#### 1) Test on random repair strategy:

First, RR strategy is performed on ER networks. The numerical results of average resiliency loss and $G$ are illustrated in Figure 4.8. Scenarios for different $CCT$ and $D$-size are considered, i.e., $CCT = 0.5, 0.7, 0.9$, $D$-size $= 2, 4, 6, 8, 10$, respectively.
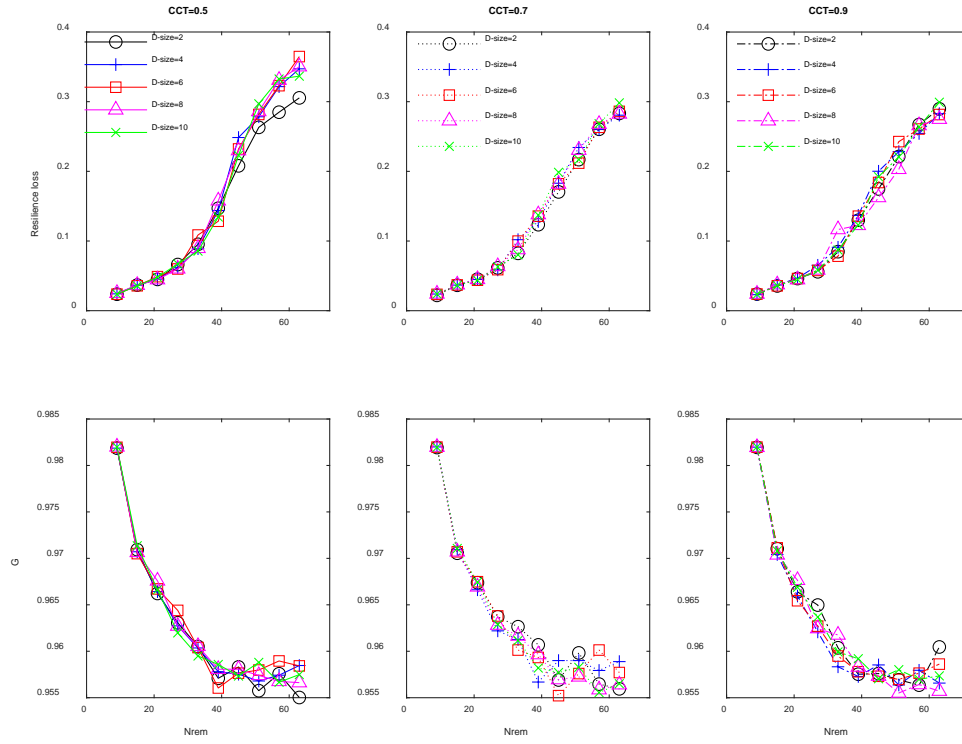
Figure 4.8. Average results of resiliency loss and *G* as a function of *Nrem*

Considering resiliency loss, it increases as *Nrem* increases no matter the value of *CCT* or *D-size*. The incremental speed of resiliency loss increases noticeably when *Nrem* surpasses a threshold. It indicates that the effectiveness of RR strategy on reducing resiliency loss decreases when *Nrem* is large enough to exceed a certain threshold. As for *G*, it declines sharply as *Nrem* increases initially, while the rate of decline decreases when *Nrem* exceeds a threshold. Note that the thresholds of *Nrem* regarding the trends of resiliency loss and *G* are almost the same. It means that system connectivity stays the same after network efficiency is recovered to the predetermined level, although resiliency loss increases as *Nrem* increases.

Table 4.5 presents the average results of resiliency loss and *T* over different *Nrem*. Based on Table 4.5, weak dependence strength corresponding to big *CCT* can reduce

resiliency loss. *T* is decreases as *CCT* increases or *D-size* decreases. It can be an indication that stronger dependence characteristics (either stronger dependence strength or larger scale of dependence cluster) impair restoration effects against cascading failures.

Table 4.5. Average results of resiliency loss and *T* for different *Nrem*

| Resiliency Loss (%) | *D-size* | | | | | Aver-age | *T* | *D-size* | | | | | Aver-age |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 4 | 6 | 8 | 10 | | | 2 | 4 | 6 | 8 | 10 | |
| *CCT* = 0.5 | 14.8 | 16.0 | 16.1 | 16.1 | 15.8 | 15.8 | *CCT* = 0.5 | 10.4 | 17.2 | 18.4 | 19.6 | 20.1 | 17.1 |
| *CCT* = 0.7 | 13.0 | 13.5 | 13.4 | 13.6 | 13.7 | 13.4 | *CCT* = 0.7 | 7.1 | 7.6 | 7.9 | 7.9 | 7.5 | 7.6 |
| *CCT* = 0.9 | 13.3 | 13.7 | 13.5 | 13.1 | 13.5 | 13.4 | *CCT* = 0.9 | 7.6 | 7.5 | 7.6 | 7.1 | 7.8 | 7.5 |
| Average | 13.7 | 14.4 | 14.3 | 14.3 | 14.3 | 14.2 | Average | 8.4 | 10.8 | 11.3 | 11.5 | 11.8 | 10.8 |

**2) Test on high degree first repair strategy:**

Average results of resiliency loss, *G* and *T* vs. *Nrem* for the HDFR strategy are presented in Figure 4.9. According to Figure 4.9, both resiliency loss and *T* increase when *Nrem* increases, while *G* decreases as *Nrem* increases. Table 4.6 shows the average results of resiliency loss and recovery resiliency. It should be noted that different *D-size* does not cause a notable difference regarding three measurements under HDFR strategy, i.e., there is a significant overlap between the plotted curves obtained from different *D-size*.
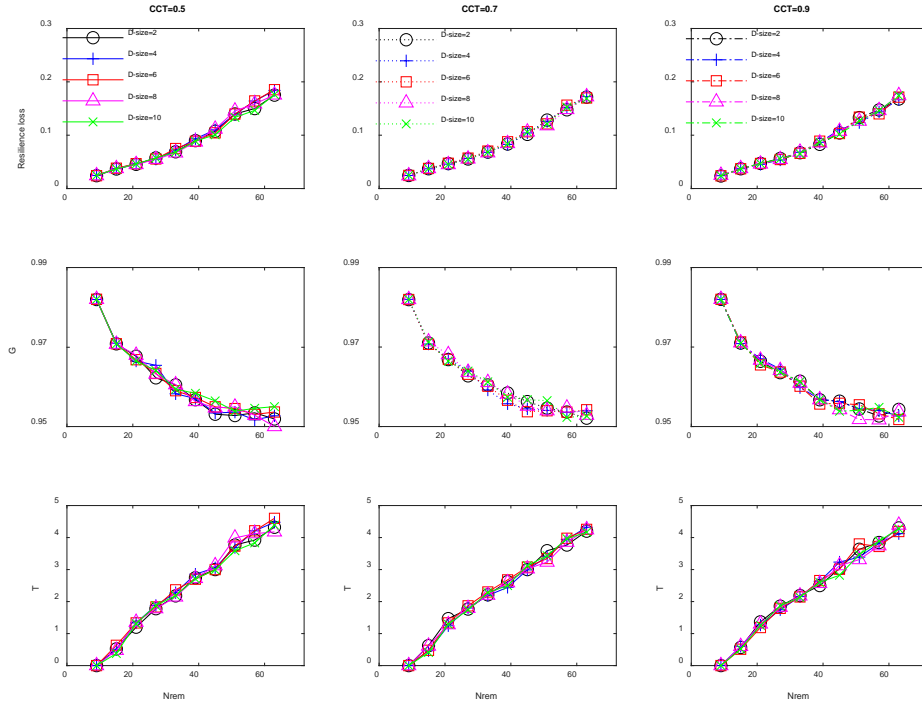
Figure 4.9. Average results of resiliency loss, *G* and *T* as a function of *Nrem*

Table 4.6. Average resiliency loss and *T* for different *Nrem*

| Resilience Loss (%) | *D-size* | | | | | Avera-ge | *T* | *D-size* | | | | | Avera-ge |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 4 | 6 | 8 | 10 | | | 2 | 4 | 6 | 8 | 10 | |
| *CCT* = 0.5 | 9.0 | 9.3 | 9.2 | 9.1 | 8.8 | 9.1 | *CCT* = 0.5 | 2.3 | 2.4 | 2.4 | 2.4 | 2.3 | 2.4 |
| *CCT* = 0.7 | 8.6 | 8.7 | 8.8 | 8.6 | 8.6 | 8.7 | *CCT* = 0.7 | 2.3 | 2.3 | 2.3 | 2.3 | 2.3 | 2.3 |
| *CCT* = 0.9 | 8.7 | 8.6 | 8.7 | 8.7 | 8.7 | 8.7 | *CCT* = 0.9 | 2.3 | 2.3 | 2.3 | 2.3 | 2.3 | 2.3 |
| Average | 8.8 | 8.9 | 8.9 | 8.8 | 8.7 | 8.8 | Average | 2.3 | 2.3 | 2.3 | 2.3 | 2.3 | 2.3 |

Similar to what was observed previously, the biggest resiliency loss and the largest *T* are obtained with the smallest *CCT*. It means that weak dependency strength denoted by large *CCT* contributes to high network resilience against cascading failures.

**3) Test on short time first repair strategy:**

Average resilience loss, *G* and *T* with different number of *Nrem* under STFR strategy are presented in Figure 4.10. Different *D-size* and *CCT* are considered separately.
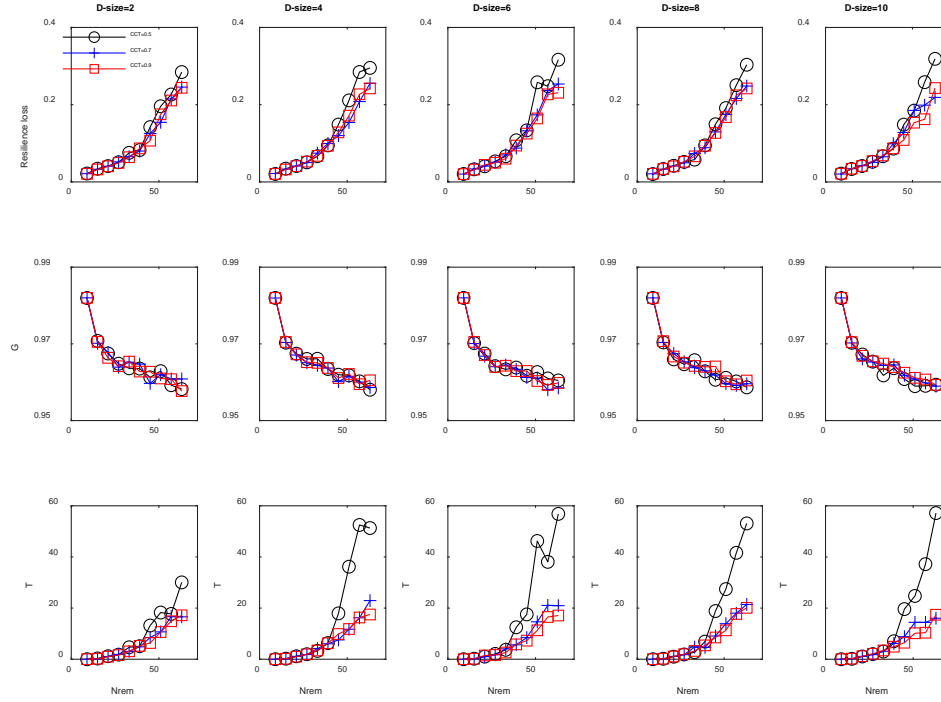


Figure 4.10. Average results of resilience loss, *G* and *T* as a function of *Nrem*. Scenarios with different *CCT* and *D-size* are considered (*CCT* = 0.5, 0.7, 0.9, *D-size* = 2, 4, 6, 8, 10).

Different from the figures obtained from other scenarios, the results presented in Figure 4.10 show the remarkable impacts of *CCT*. Although the changing trends of resilience loss, *G* and *T* as *Nrem* increases are consistent with what was observed before, it can be clearly seen that larger *CCT* contributes to larger resilience loss and longer *T*.

Table 4.7. Average resilience loss and *T* for different *Nrem*

| Resilience Loss (%) | D-size | | | | | Aver-age | *T* | D-size | | | | | Aver-age |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 4 | 6 | 8 | 10 | | | 2 | 4 | 6 | 8 | 10 | |
| *CCT* = 0.5 | 11.5 | 12.5 | 12.8 | 12.0 | 12.1 | 12.2 | *CCT* = 0.5 | 9.2 | 17.1 | 17.8 | 15.4 | 15.2 | 14.9 |
| *CCT* = 0.7 | 10.3 | 10.6 | 11.0 | 10.8 | 10.4 | 10.6 | *CCT* = 0.7 | 6.3 | 7.2 | 7.8 | 7.5 | 6.6 | 7.1 |
| *CCT* = 0.9 | 10.3 | 10.8 | 10.5 | 10.7 | 9.7 | 10.4 | *CCT* = 0.9 | 6 | 6.8 | 6.4 | 7 | 5.6 | 6.4 |
| Average | 10.7 | 11.3 | 11.4 | 11.2 | 10.7 | 11.1 | Average | 7.2 | 10.4 | 10.7 | 10.0 | 9.1 | 9.5 |

The results presented in Table 4.7 illustrate that both resilience loss and *T* can be reduced by increasing *CCT*, i.e., weakening the strength of dependencies among network nodes. Besides, resilience loss, *G* and *T* possess the similar relationship with *Nrem* as observed above. Resilience loss and *T* increase as *Nrem* increases, whereas *G* decreases as *Nrem* increases.

**4) Test on high load first repair strategy:**

Finally, HLFR strategy is applied regarding mixed cascading failures. The numerical results of average *G* and *T* under HLFR strategy are presented in Figure 4.11. The changing trends of *G* and *T* in Figure 4.11 are similar to what was previously presented in Figure 4.9. *G* decreases while *T* increases as *Nrem* increases. We can also observe the thresholds regarding the changing rate of *G* and *T* once *Nrem* exceeds.
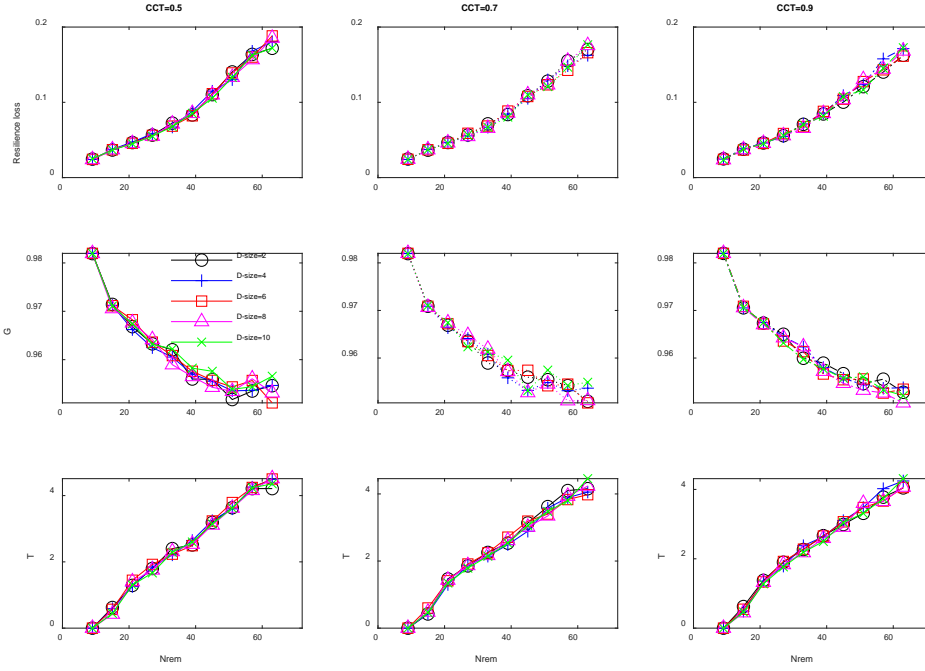
Figure 4.11. Average results of *G* and *T* as a function of *Nrem* under different *CCT* and

*D-size*

Table 4.8. Average of resilience loss and *T* for different *Nrem*

| Resilience Loss (%) | *D-size* | | | | | Aver-age | *T* | *D-size* | | | | | Aver-age |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 4 | 6 | 8 | 10 | | | 2 | 4 | 6 | 8 | 10 | |
| *CCT* = 0.5 | 9.1 | 9.2 | 9.1 | 9.0 | 8.9 | 9.1 | *CCT* = 0.5 | 2.4 | 2.4 | 2.4 | 2.4 | 2.4 | 2.4 |
| *CCT* = 0.7 | 8.8 | 8.6 | 8.6 | 8.8 | 8.7 | 8.7 | *CCT* = 0.7 | 2.4 | 2.3 | 2.3 | 2.3 | 2.3 | 2.3 |
| *CCT* = 0.9 | 8.4 | 8.8 | 8.6 | 8.6 | 8.6 | 8.6 | *CCT* = 0.9 | 2.3 | 2.3 | 2.3 | 2.3 | 2.3 | 2.3 |
| Average | 8.8 | 8.9 | 8.8 | 8.8 | 8.7 | 8.8 | Average | 2.4 | 2.3 | 2.3 | 2.3 | 2.3 | 2.3 |

According to Table 4.8, average resilience loss continues to decrease when *CCT* increases, i.e., weaken dependency strength can help to reduce system load loss caused by cascading failures. It is consistent with the results obtained in above cases.

Figure 4.12 Average resilience loss, *G* and *T* over *D-size* vs. *Nrem* for three *CCT* under

four restoration strategies

Figure 4.12 presents average resilience loss, *G* and *T* for different *D-size* (*D-size* = 2, 4, 6, 8, 10) under four restoration strategies. Based on Figure 4.12, the biggest resilience loss is incurred by using RR strategy, while the smallest resilience loss and the shortest *T* are achieved by conducting HDFR strategy or HLFR strategy. The largest *G* is achieved by adopting STFR strategy, whereas the corresponding *T* and resilience loss is less desirable compared with that under HDFR strategy or HLFR strategy. In addition, resilience loss and *T* under RR strategy or STFR strategy are the most undesirable when *CCT* = 0.5. It indicates that the restoration effects of these two strategies are weakened by strong dependence strength.

### 4.3.3 Example of BA scale free network model

Another numerical example is performed on Barabási-Albert (BA) scale-free network [134]. BA network model is a widely used model to depict scale-free networks. These types of networks have a power-law distribution of node degree. The distribution probability that network nodes have degree $k$ is $P(k): k^{-r}, r \approx 3$. This property has been observed in many real-life networks such as power grids, communication networks, and the internet [135]. In this work, the adopted BA networks that have the same number of nodes and edges with the networks used above (500 nodes and 2980 edges). Note that all the following results are average results over 90 simulation realizations.

### 1) Test on random repair strategy:
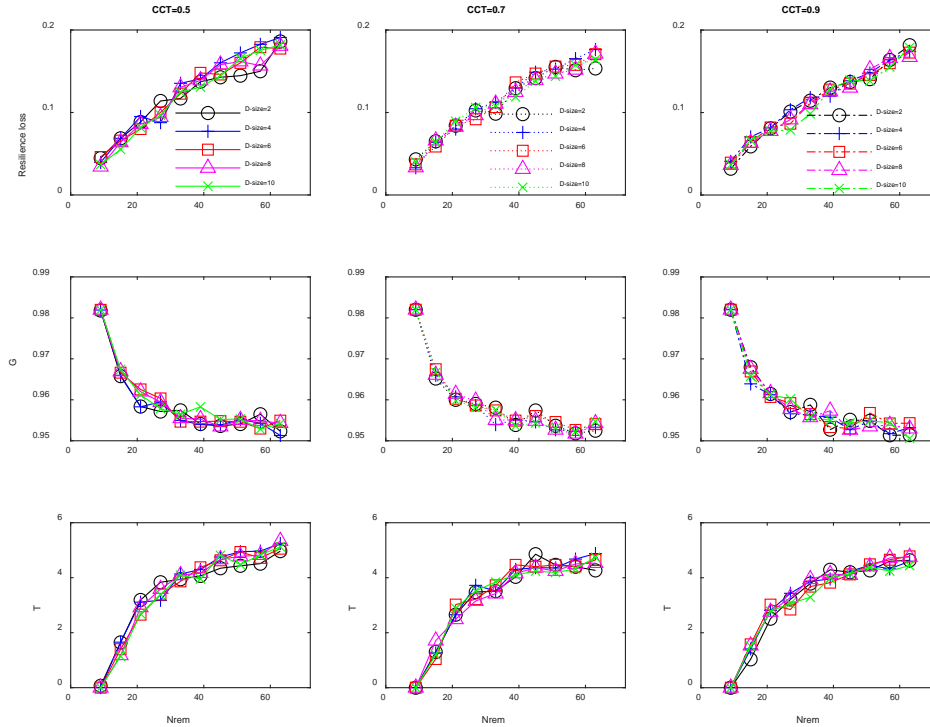


Figure 4.13. Average results of resilience loss, *G* and *T* as a function of *Nrem*

Firstly, RR strategy is applied to the BA network. The results of average resilience loss, $G$ and $T$ are presented in Figure 4.13. Average resilience loss and $T$ increase while $G$ decreases as *Nrem* increases. The trends are in accordance with the results obtained previously. It means that a larger number of initially failed nodes contribute to higher resilience loss and longer repair time to recover the system to the predetermined level. It can be observed that the growth rate of $T$ and the rate of decline of $G$ dramatically change when *Nrem* exceeds the same threshold.

Average of resilience loss and $T$ for different *Nrem* are presented in Table 4.9. According to Table 4.9, resilience loss decreases when *CCT* increases, and the longest average $T$ results from the network with the smallest *CCT*. It can be explained that network systems with weak dependency strength, denoted by big *CCT*, have better resilience against cascading failures.

Table 4.9. Average of resilience loss and $T$ for different *Nrem*

| Resilience Loss (%) | D-size | | | | | Aver-age | $T$ | D-size | | | | | Aver-age |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 4 | 6 | 8 | 10 | | | 2 | 4 | 6 | 8 | 10 | |
| *CCT* = 0.5 | 11.9 | 12.7 | 12.3 | 12.1 | 11.9 | 12.2 | *CCT* = 0.5 | 3.5 | 3.6 | 3.5 | 3.6 | 3.5 | 3.5 |
| *CCT* = 0.7 | 11.2 | 11.7 | 11.5 | 11.2 | 11.3 | 11.4 | *CCT* = 0.7 | 3.3 | 3.4 | 3.3 | 3.3 | 3.3 | 3.3 |
| *CCT* = 0.9 | 11.4 | 11.6 | 11.3 | 11.2 | 11.0 | 11.3 | *CCT* = 0.9 | 3.3 | 3.3 | 3.3 | 3.4 | 3.2 | 3.3 |
| Average | 11.5 | 12.0 | 11.7 | 11.5 | 11.4 | 11.6 | Average | 3.4 | 3.4 | 3.4 | 3.4 | 3.3 | 3.4 |

## 2) **Test on high degree first repair strategy:**
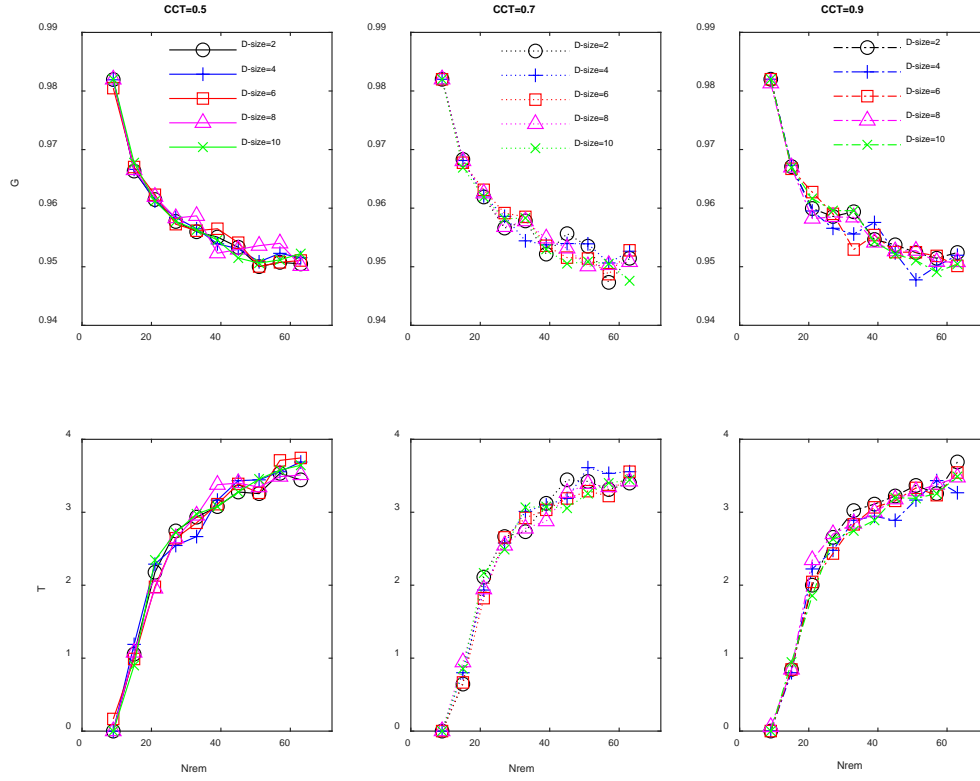


Figure 4.14. Average results of *G* and *T* as a function of number of *Nrem*

HDFR strategy is employed on BA networks regarding cascading failures. The average results of *G* and *T* are shown in Figure 4.14. Similar to what was observed before about the changing trends of *G* and *T* vs. *Nrem*, there is a noticeable threshold that the decreasing of *G* and the increasing of *T* slows down when *Nrem* surpasses that threshold.

The average numerical results for different *Nrem* are presented in Table 4.10. Based on Table 4.10, larger *CCT* leads to smaller resilience loss and shorter *T*. It is consistent with the results obtained from other network cases.

Table 4.10. Average results of resilience loss and *T* for different *Nrem*

| Resilienc e Loss (%) | D-size | | | | | Aver-age | T | D-size | | | | | Aver-age |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 4 | 6 | 8 | 10 | | | 2 | 4 | 6 | 8 | 10 | |
| CCT = 0.5 | 10.9 | 11.1 | 11.3 | 11.0 | 11.0 | 11.1 | CCT = 0.5 | 2.6 | 2.6 | 2.6 | 2.6 | 2.6 | 2.6 |
| CCT = 0.7 | 10.6 | 10.9 | 10.1 | 10.4 | 10.5 | 10.5 | CCT = 0.7 | 2.5 | 2.5 | 2.4 | 2.5 | 2.5 | 2.5 |
| CCT = 0.9 | 10.8 | 10.3 | 10.6 | 10.6 | 10.2 | 10.5 | CCT = 0.9 | 2.5 | 2.4 | 2.4 | 2.5 | 2.4 | 2.4 |
| Average | 10.8 | 10.8 | 10.7 | 10.7 | 10.6 | 10.7 | Average | 2.5 | 2.5 | 2.5 | 2.5 | 2.5 | 2.5 |

**3) Short time first repair strategy:**
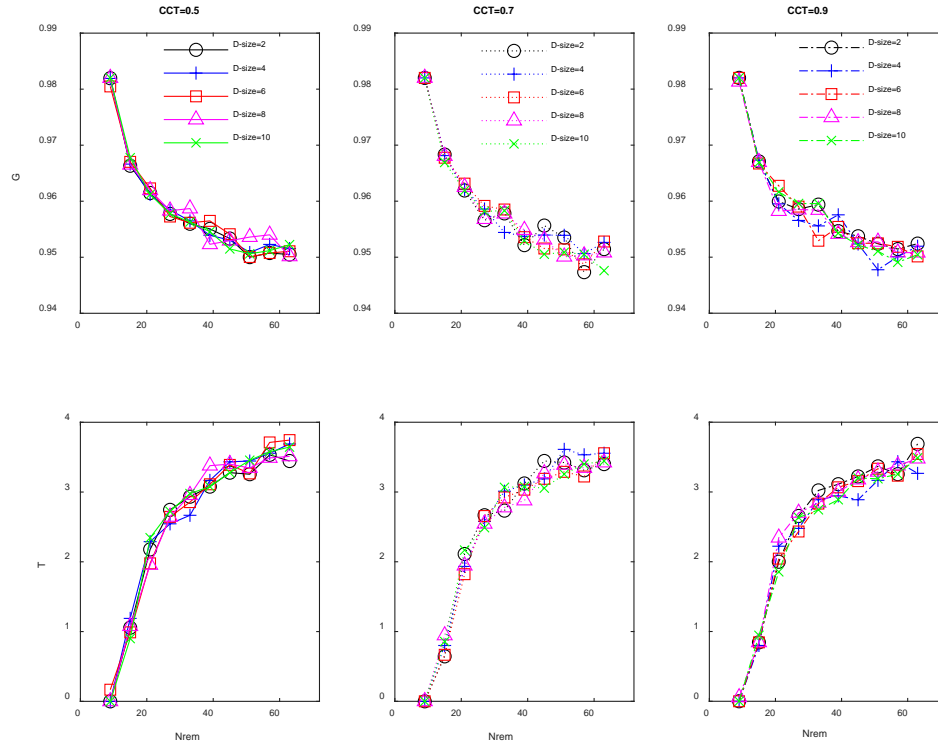


Figure 4.15. Average results of *G* and *T* as a function of number of *Nrem*

The results of average *G* and *T* obtained from using STFR strategy with different *CCT* and *D-size* are presented in Figure 4.15. *G* decreases and *T* increases as *Nrem* increases based on Figure 4.15, whereas the decreasing rate of *G* and increasing rate of *T*

wane when *Nrem* exceeds a certain threshold. It can be an indication that the effectiveness of STFR strategy is gradually diminished regarding the mixed cascading failures caused by a larger number of initially failed nodes.

Table 4.11. Average of resilience loss and *T* for different *Nrem*

| Resilience Loss (%) | D-size | | | | | Avera-ge | *T* | D-size | | | | | Avera-ge |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 4 | 6 | 8 | 10 | | | 2 | 4 | 6 | 8 | 10 | |
| *CCT* = 0.5 | 10.4 | 10.9 | 10.5 | 11.1 | 10.9 | 10.8 | *CCT* = 0.5 | 3.5 | 3.7 | 3.6 | 3.6 | 3.7 | 3.6 |
| *CCT* = 0.7 | 10.0 | 10.5 | 10.3 | 10.3 | 10.2 | 10.3 | *CCT* = 0.7 | 3.4 | 3.5 | 3.4 | 3.5 | 3.4 | 3.4 |
| *CCT* = 0.9 | 10.2 | 9.9 | 10.3 | 10.5 | 10.2 | 10.2 | *CCT* = 0.9 | 3.3 | 3.4 | 3.4 | 3.5 | 3.4 | 3.4 |
| Average | 10.2 | 10.4 | 10.4 | 10.6 | 10.4 | 10.4 | Average | 3.4 | 3.5 | 3.5 | 3.5 | 3.5 | 3.5 |

Corresponding average results of *G* and *T* for different values of *Nrem* are presented in Table 4.11. The results in Table 4.11 show that resilience loss decreases as *CCT* increases, and the longest *T* is achieved with the smallest *CCT*. It can be explained that weak dependency strength, reducing the impacts of failed components on other functional ones, leading to less resilience loss due to cascading failures and shorter system repair time as well.

**4) Test on high load first repair strategy:**

In the end, HLFR strategy is performed on BA networks with different dependence characteristics, i.e., *CCT* = 0.5, 0.7, 0.9, *D-size* = 2, 4, 6, 8, 10, respectively. The average results of *G* and *T* are presented in Figure 4.16. According to Figure 4.16, *G* decreases while *T* increases when *Nrem* increases. The changing rates of *G* and *T* decrease when *Nrem* exceeds a certain threshold.

Average resilience loss and *T* are presented in Table 4.12. The results shown in Table 4.12 illustrate that resilience loss is reduced if *CCT* is increased. It is in accordance with the results obtained from the above cases.
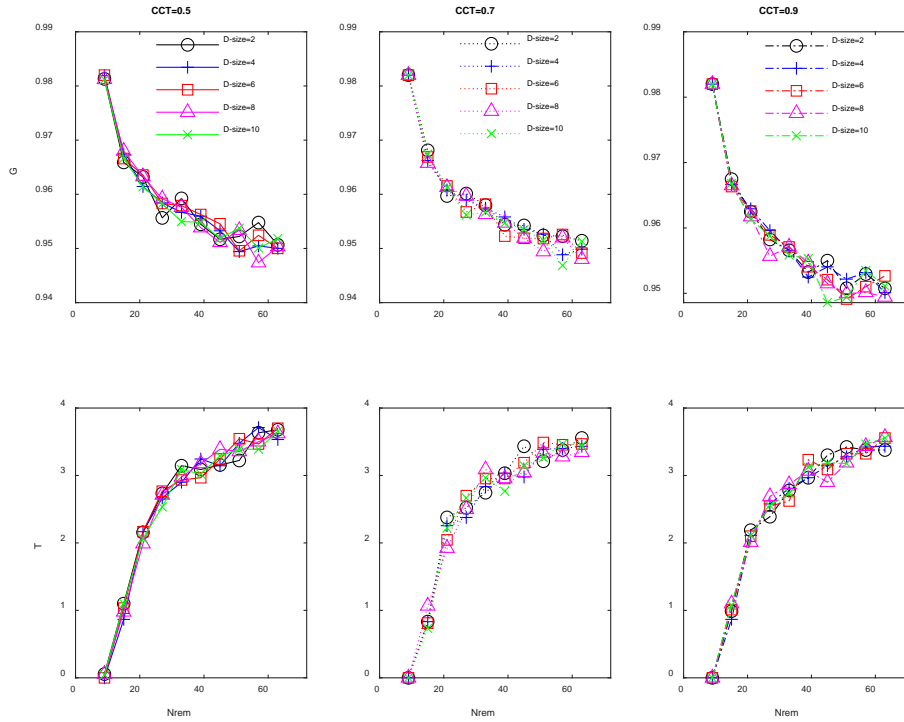


Figure 4.16 Average results of *G* and *T* as a function of number of *Nrem*

Table 4.12. Average of resilience loss and *T* for all considered *Nrem*

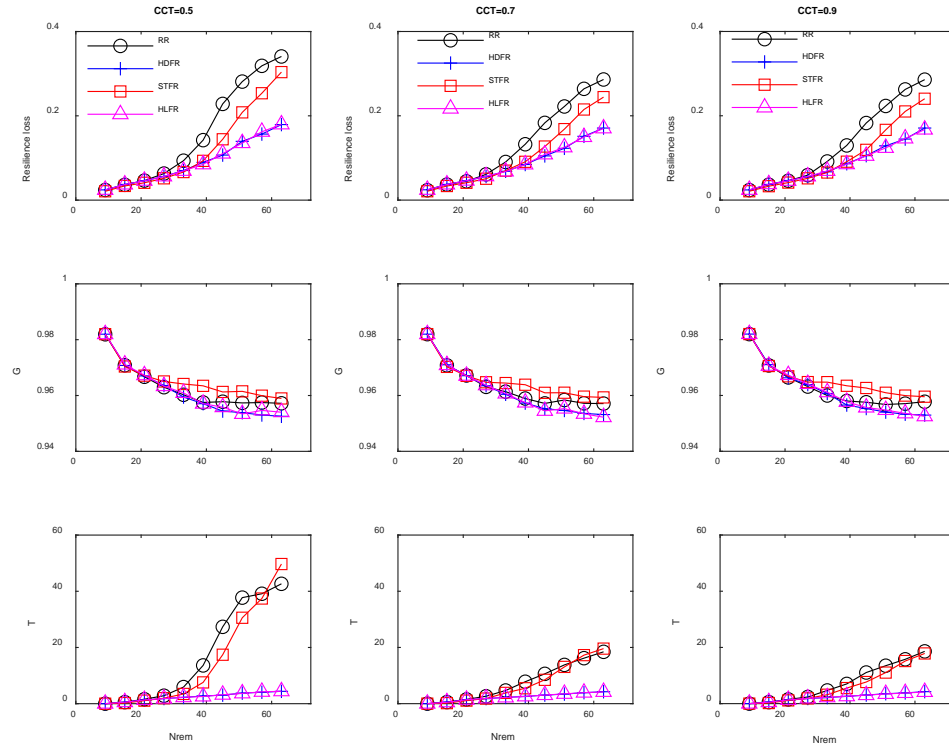| Resilience Loss (%) | D-size | | | | | Aver-age | T | D-size | | | | | Aver-age |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 4 | 6 | 8 | 10 | | | 2 | 4 | 6 | 8 | 10 | |
| CCT = 0.5 | 11.0 | 11.2 | 11.0 | 11.2 | 10.9 | 11.1 | CCT = 0.5 | 2.6 | 2.6 | 2.6 | 2.6 | 2.6 | 2.6 |
| CCT = 0.7 | 10.7 | 10.5 | 10.7 | 10.5 | 10.6 | 10.6 | CCT = 0.7 | 2.5 | 2.5 | 2.5 | 2.5 | 2.5 | 2.5 |
| CCT = 0.9 | 10.5 | 10.3 | 10.6 | 10.4 | 10.7 | 10.5 | CCT = 0.9 | 2.5 | 2.5 | 2.5 | 2.5 | 2.5 | 2.5 |
| Average | 10.7 | 10.7 | 10.8 | 10.7 | 10.7 | 10.7 | Average | 2.5 | 2.5 | 2.5 | 2.5 | 2.5 | 2.5 |
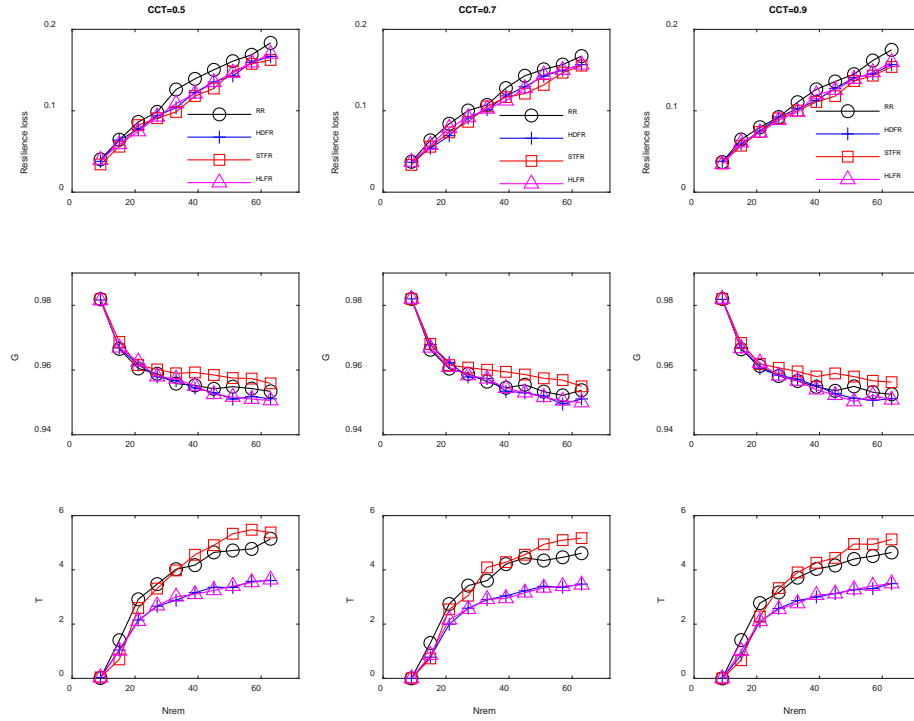
Figure 4.17 Average resilience loss, *G* and *T* over *D-size* vs. *Nrem* under different

restoration strategies

Figure 4.17 presents average resilience loss, *G* and *T* over different *D-size* (*D-size =* 2, 4, 6, 8, 10) for the four restoration strategies for three different *CCT*. It can be seen from Figure 4.17 that RR strategy leads to the biggest resilience loss, whereas the resilience loss resulted from other restoration strategies does not have a remarkable difference. It is clear that the largest *G* is obtained by employing STFR strategy, especially, when *Nrem* is large. For *T*, the shortest *T* is achieved by applying HDFR strategy or HLFR strategy, whereas it takes longer *T* for network efficiency to be recovered to the predetermined level when using STFR strategy or RR strategy.

Overall, the average resilience loss, *G* and *T* over different dependence characteristics, *CCT* and *D-size*, for each network system under four different restoration

strategies are presented in Table 4.13. It should be mentioned again that the restoration process is stopped when network efficiency is recovered to 95% of its initial value in this work.

Table 4.13. Average measurements under four restoration strategies for three network systems

| Average results | U.S. top500 airport network | | | ER network | | | BA network | | |
|---|---|---|---|---|---|---|---|---|---|
| | Resilience loss | $G$ | $T$ | Resilience loss | $G$ | $T$ | Resilience loss | $G$ | $T$ |
| RR | 12.0% | 0.954 | 3.4 | 14.2% | 0.964 | 10.8 | 11.6% | 0.960 | 3.4 |
| HDFR | 11.4% | 0.954 | 2.6 | 8.8% | 0.962 | 2.3 | 10.7% | 0.958 | 2.5 |
| STFR | 10.8% | 0.956 | 3.4 | 11.1% | 0.966 | 9.5 | 10.4% | 0.962 | 3.5 |
| HLFR | 11.1% | 0.954 | 2.6 | 8.8% | 0.962 | 2.3 | 10.7% | 0.958 | 2.5 |
| Average | 11.3% | 0.955 | 3 | 10.7% | 0.964 | 6.2 | 10.9% | 0.960 | 3 |

Based on Table 4.13, the following findings can be obtained,

(1) As the baseline restoration strategy, the biggest resilience loss and the longest total repair time $T$ (except for BA network) are obtained by applying RR strategy.

(2) The shortest total repair time $T$ can be achieved by employing HDFR or HLFR strategies. Meanwhile, less resilience loss is obtained when compared with that of the RR strategy. However, the smallest $G$ results from applying HDFR strategy or HLFR strategy.

(3) As to STFR strategy, it contributes to the largest $LCC$ and the least resilience loss except for ER network, while the total repair time $T$ is longer than that under HDFR or HLFR strategy.

(1) In comparison with the U.S. top 500 airport topology and BA networks, ER networks

possess better resilience against mixed cascading failures regarding resilience loss and network connectivity after applying restoration strategies. However, it takes a longer time for the network system to be recovered to the predetermined level.

### 4.3.4    Summary

In this work, optimal selection of resilience-based restoration against cascading failures considering the impacts of network dependence characteristics is performed. Dependence clusters are modeled and taken into account for accelerating failure propagation, which contributes to mixed cascading failures. Numerical examples are performed on the U.S. top 500 airport network structure and synthetic network models with employing four restoration prioritization strategies. The impacts of system dependence characteristics and their interactions with the effectiveness of restoration strategies regarding cascading failures are explored. In order to perform a comprehensive evaluation, network resilience loss, caused by cascading failures, is evaluated in terms of network load. Two measurements with respect to network connectivity and restoration time-efficiency are adopted, and network transmission efficiency is applied to define recovery level.

In summary, strong dependence strength, which may enhance system operation efficiency, impairs system resilience against cascading failures because the influence of failed components is aggravated. According to the results, larger $CCT$ leads to smaller resilience loss, shorter system repair time $T$ no matter which restoration strategy is applied. In other words, weak dependence strength also helps to improve restoration effects to reduce resilience loss and shorten repair time. However, $D\text{-}size$, which indicates

the dependence scale, does not show a monotonic impact on system resilience or restoration effects.

Based on the results, restoration prioritization strategy selection should be carefully made according to a specific preference. The best recovery of network connectivity can be achieved by applying STFR strategy, while the expected system repair time is longer. HDFR strategy and HLFR strategy are the choices if shorter system repair time is preferred with an acceptable resilience loss. The results can be useful to provide a basis for selecting restoration prioritization strategy regarding resilience enhancement for different network systems with specific properties. It also provides insights for integrating restoration effects with system dependency impacts to mitigate the intensity and extent of cascading failures effectively.

Since available budgets for restoration investment are limited in reality, budget constraints can be incorporated regarding restoration optimization from an economic aspect for future research. More restoration strategies can be applied to other real-world network systems with different modeling of system dependency and network load dynamics.

## 4.4  A recovery framework with deep graph learning

Although efforts have been made to prevent cascading failures, high-impact cascading failures still occur for various reasons. One of the main reasons is that a large network system comprises many different components. System component importance for restoration changes as system structure changes by the ongoing failures. Therefore, identifying important system components from the real-time system structure during failure propagation and system recovery is critical to allocate limited recovery resources

efficiently. However, most current restoration strategies determine restoration priority based on a steady or fixed network property.

Among network topology features, minimum vertex cover (MVC) - a key system component/node subset with a minimized number of nodes so that all system edges are covered - can be used to identify critical system components in terms of global network connectivity importance. MVC has already been applied to different research fields such as network security [136], computational biology [137], and text summarization [138].

In this section, a recovery framework for system resilience enhancement against cascading failures by considering real-time MVC importance is proposed. The framework is developed based on quick system MVC identification using structural graph embedding technique and deep reinforcement learning (RL)-based graph learning. By incorporating MVC into existing restoration strategies, the framework can optimize resource allocation during the recovery process by real-time updating restoration priority targets.

### 4.4.1 Main assumptions

The mixed cascading failure model proposed in section 3.1 is also applied in this work. The sizes of dependence clusters in the network system follows a shifted/scale adjusted Poisson distribution, which is described based on Eq. (3.5), and the impact of dependence clusters on failure propagation is implemented by *CCT*. The resilience loss metric that is proposed in section 4.2 is adopted to evaluate the restoration effects of the recovery work. Network load is also used as $Q(t)$ for calculating resilience loss in terms of network load demand and supply capability. The predetermined system recovery level is measured by network efficiency, as was described previously. The time to restore the

network system to the predetermined level during the cascading-restoration process, $T$, reflects the restoration effects in terms of downgrading operation time. It is obviously that shorter $T$ is more desirable.

For comparison, three different existing restoration strategies with specific preferences are employed to restore the network system against cascading failures. (1) The high-degree first repair (HDFR) strategy is designed for the repair order assigned based on the degree (the number of topological connections) of failed nodes (i.e., this strategy has the preference that network nodes with higher node degrees are repaired with higher priority). (2) The shortest-time first repair (STFR) strategy assigns the repair order according to the required repair time of failed nodes. The restoration prioritizes the failed nodes that require shorter repair time. (3) The high load first repair (HLFR) strategy prioritizes repair based on the amount of load transmitted through the failed nodes (i.e., this strategy has the preference that network nodes that carry higher loads are repaired with higher priority).

Ties, which occur when failed nodes have the same conditions (i.e., degree, required repair time, and load), are broken according to the first fail first repair policy. Once started, the repair activities of failed nodes do not stop until they are completed. For simplicity, it is assumed that the restoration resources are available to complete repairs. The number of new repairing activities of failed nodes, which begin at each round of inspection for different restoration strategies, is decided by the repair proportion $R_p$ and the total number of failed nodes in which repair activities have not yet started at the time of inspection.

As was mentioned previously, existing restoration strategies with specific preferences rarely, if ever, consider component importance changing during the cascading-restoration process. Therefore, identifying important components based on MVC from the real-time network structure is of practical significance for updating restoration priority targets during the recovery process for effective restoration. A recovery framework is proposed that can combine real-time MVC importance with other preferences according to customers' desirability (e.g., network transmission importance) to restore network system from the cascading failures effectively. Case studies are implemented on different network systems to demonstrate the effectiveness of the proposed framework. The details of the recovery framework are presented in the following sections.

### 4.4.2 Deep graph learning for MVC detection

MVC is an NP (non-deterministic polynomial-time) - hard computational problems, which means it requires exponential time algorithms to search for the optimal solution. Traditional methods to solve MVC can be divided into three categories: exact algorithms, approximation algorithms, and heuristic algorithms. Exact algorithms are based on enumeration or linear programming, which have exponential time complexity and are not practical for large-scale graphs. Approximation algorithms (if they are available) are faster with polynomial time complexity but do not guarantee solution quality and are not practically applicable. Heuristic algorithms are fast but do not guarantee solution quality, and they require expert knowledge and repeated designs for different problems. RL has shown its potential in optimization and management [139], and can be combined with graph embeddings to derive an MVC solution efficiently. A state-of-the-art, high-

performance deep graph analysis environment (OpenGraphGym) [140] is adopted to support MVC estimation in a very short period of time to provide a real-time reference for restoration priority updates in practice. The OpenGraphGym helps to use several graph embeddings to represent graph attributes and features; it adopts several RL algorithms for an optimal solution for graph-based programs. Structure2vec graph embedding is used [141], and a deep Q learning algorithm is adopted [142] to find an optimal solution of MVC (Figure 4.18).

As shown in Figure 4.18, two major parts of the framework are graph embedding and the RL Q function. Two steps are illustrated in the figure. At each step, graph embedding takes the graph as the input and produces the embeddings for all nodes in the graph. Then, the embeddings for all nodes are sent to the RL Q function. The RL Q function computes the scores for all nodes in the graph (shown as $q$ in Figure 4.18). The node with the highest score ($q_2$) in the first round is marked and added to the partial solution. In Figure 4.18, two blue nodes are included in the partial solution.
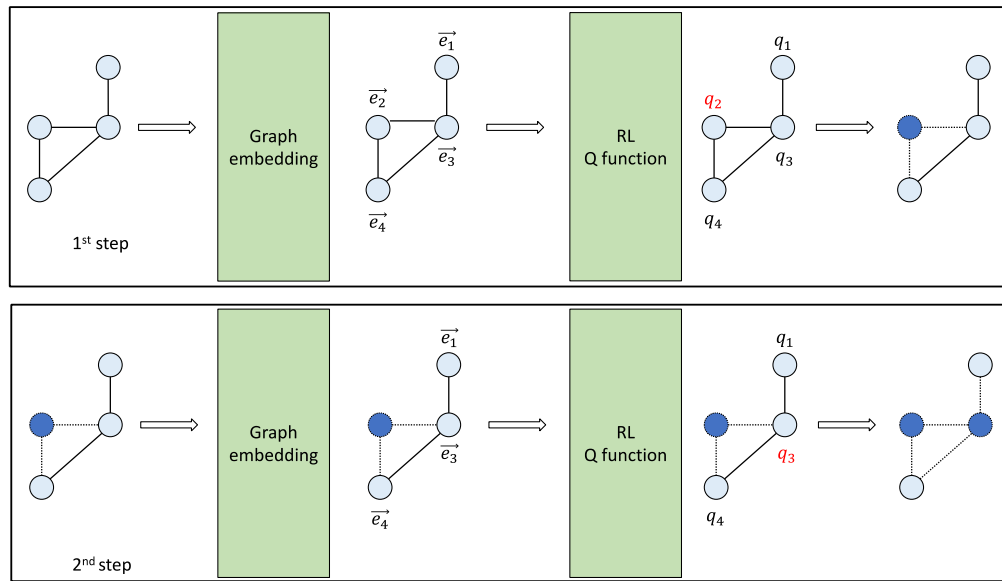


Figure 4.18. Diagram of the advanced RL-based graph learning for MVC detection.

By adding a node to the partial solution at each step, this approach provides a ranked MVC solution, which means that early selected nodes in an MVC solution generally contribute more to the final optimal MVC solution than late selected nodes. To further improve the computational efficiency to deliver a reasonable MVC solution, small graphs are used to train an MVC graph agent within the OpenGraphGym that can provide a reasonable MVC solution for relatively large graphs, which enables real-time MVC detection for real-world large scale networks. This acceleration is valid when the training graphs and testing graphs are of the same graph type and degree distribution.

Because MVC identifies important network nodes from the perspective of global network connectivity, the proposed recovery framework combines it with the existing restoration strategies with other prioritization preferences for optimal restoration resource allocation. In this work, real-time MVC is incorporated with HLFR and STFR, respectively. The first/top $u$ percentage of the nodes in MVC (the most vital nodes in an MVC optimal solution) is assigned with a high restoration priority. Then, the repairing of other failed nodes, which do not belong to these top $u$ percentage of MVC nodes, started according to HLFR or STFR. The order to start the repairing activities of the top u percentage MVC nodes, if their repairing had not yet started, is assigned based on their ranking inside of MVC. The restoration prioritization of MVC in the recovery framework is indicated as $u$.

The main steps of implementing the recovery framework with real-time MVC detection for system restoration prioritization against cascading failures are briefly presented as follows:

Step 1)    All network nodes are initially functional with limited capacity. Dependence clusters are created. Detect MVC node set of the network system.

Step 2)    Randomly select nodes to break down because of initial disturbances.

Step 3)    Dependence clusters collapse if *CCT* is exceeded. Update the network structure.

Step 4)    Network loads are dynamically redistributed over the current network structure. Overloaded nodes break down.

Step 5)    Failed nodes are arranged to start the repairing activities based on the restoration strategy considering MVC importance.

Step 6)    Quickly detect MVC from the current network structure.

Step 7)    Return to Step 3 until system is recovered to the predetermined level.

The performance of the network system is recorded during the process of cascading failures with restoration implementation.

### 4.4.3   Case studies

Two case studies are conducted on a synthetic network structure and on a real-world network system structure. The results are presented in the following sections. To prove the effectiveness of the proposed recovery framework, three existing restoration strategies introduced previously are implemented for comparison. The results under the existing restoration strategies are used to compare with results from the MVC-based recovery framework.

### 4.4.3.1 Experiments with synthetic networks

In this case study, the mixed cascading failures are performed on BA scale-free networks with maintenance implementation. The main assumptions made in the

numerical simulation are as follows, the required repair time for failed nodes are independent random variables that are uniformly distributed in [1, 4] in terms of simulation steps. Each simulation step represents a fixed duration of time. Once started, the repairing of a node does not stop until it is completed. *CCT* is 0.7, i.e., dependence clusters instantly collapse once the failed nodes inside exceeds 70%. The maintenance process stops when network efficiency is recovered to 95% of its initial value. The proportion of total failed network nodes, where repair activities have not yet started, are selected to start repairing at each round of inspection, $R_p = 0.6$.

The adopted BA network examples have 250 nodes with an approximate edge probability of 0.1. To minimize random errors, the cascading failures triggered by 10 different sets of initial random failures of nodes, for each of which they randomly generated a different number of BA network realizations following the specific network scale is implemented. The 10 sets of different numbers of randomly initial failed nodes (*Nrem*) are 9, 15, 21, 27, 33, 39, 45, 51, 57 and 63, accounting for about 4 to 25% of the total network nodes.

### 1) MVC graph agent development and implementation

To improve computational efficiency, 400 BA graphs of 20 nodes with an approximate edge probability of 0.1 is created to train the MVC graph agent. These training BA graphs are generated using the function "barabasi_albert_graph" in the graph manipulation library networkX [143]. This function can be used to create scale-free networks using BA network model [134].

The trained MVC graph agent is tested on random graphs. For the targeted 250-node BA networks with edge probability around 0.1, the agent provided the MVC solution

with the average size of 202, whereas the average size of MVC solutions given by a 2-opt

MVC solver in networkX is around 225. The solution size obtained from the networkX

solver is theoretically guaranteed to be less than two times the optimal solution size [144].

Based on the trained MVC graph agent, MVC is estimated from the real-time BA

network structure changed by the ongoing cascading failures and maintenance

implementation. Then, the restoration priority is updated according to the current MVC

and the other maintenance preference.

## 2) Results and discussion

MVC is combined with HLFR and STFR, respectively, to perform the proposed

recovery framework to dynamically update the maintenance priority. The parameter $u$,

which determines the top percentage of MVC nodes that have high maintenance priorities,

are top 10%, top 30%, top 50% of the nodes in MVC, respectively. Figure 4.19 presents

the average resilience loss and $T$ vs. *Nrem* in BA networks under the proposed recovery

framework and three existing restoration strategies. The presented results under the

existing restoration strategies for each *Nrem* correspond to an average of over 90 network

realizations. The results under MVC-based recovery strategies are averaged over 20

realizations. Scenarios with different *D-size* are also considered, i.e., *D-size* = 4, 8, 12, 16,

respectively. The curves shown in Figure 4.19 are averaged over the results of the

scenarios with the four values of *D-size*.
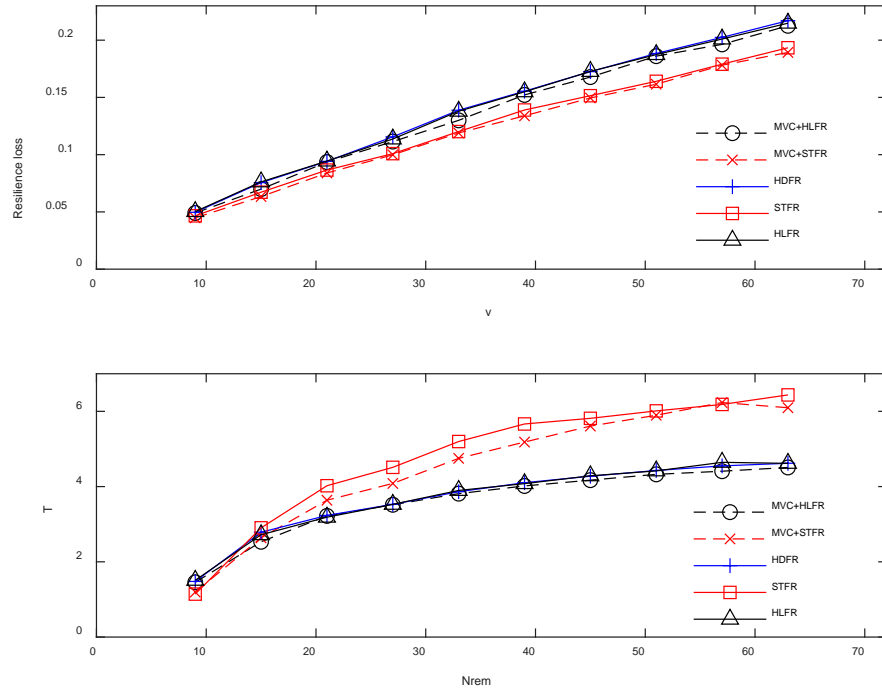
Figure 4.19 Average results of resilience loss and *T* as a function of *Nrem* for the

proposed recovery framework vs. existing restoration strategies

As shown in Figure 4.19, resilience loss increases as *Nrem* increases because a larger number of nodes randomly selected as initial failures would trigger more dramatic cascading failures, which leads to more damage to the system performance. STFR shows better maintenance effects on reducing resilience loss than those of HDFR and HLFR. MVC+STFR presents the best maintenance effects on mitigating resilience loss, and MVC+HLFR performs better than HLFR. *T* also increases as *Nrem* increases, which is consistent with the changing trend of resilience loss, $\Re(T)$. Restoring network system takes more time from the cascading failures caused by a larger *v* to the predetermined level. The longest *T* is incurred using STFR, and the shortest *T* is achieved with HDFR or HLFR. The MVC-based recovery framework clearly shortens *T* when compared with the

corresponding existing restoration strategies. Additionally, the threshold can be observed from the changing trend of *T* as *Nrem* increases. The growth rate of *T* decreases when *v* exceeds a certain threshold. Table 4.14 presents the average results of *T* and resilience loss for different scenarios under the existing restoration strategies and the proposed recovery framework incorporating MVC.

Table 4.14: Average results under different restoration strategies

| Restoration strategy | Average *T* | Average resilience loss (%) |
|---|---|---|
| HDFR | 3.7 | 14.10 |
| HLFR | 3.7 | 14.03 |
| MVC+HLFR | 3.6 | 13.69 |
| STFR | 4.8 | 12.49 |
| MVC+STFR | 4.5 | 12.23 |

According to Table 4.14, MVC+STFR, which combining global network connectivity importance with repair time preference, performs better than STFR. MVC+HLFR, which considers the importance of global network connectivity and transmission capability, has better restoration effects than HLFR. These results indicate that the proposed recovery framework updating the restoration priority by incorporating real-time MVC importance with other restoration prioritizations has better maintenance effects on reducing resilience loss and shortening *T* than the existing strategies.

### 4.4.3.2 Experiments with the US Top 500 airport network

Disruptions of air transportation systems, caused by events, such as extreme weather and attacks, can lead to huge economic losses [145]. Studies have been conducted on the robustness of air transportation networks subject to interruptions [146, 147]. A case study is conducted on the US top 500 airport network with the largest amount of traffic from publicly available data [133]. This real-world network system consists of 500 nodes and 2,980 edges. Network nodes denote airports, and edges represent air routes between

airports. Mixed cascading failures and restoration are implemented into this system to investigate the maintenance effects of different strategies. Figure 4.20 depicts the US top 500 airport network using the visualization tool Gephi [140]. The darker colors represent the nodes with more significant degrees. Some hub nodes have larger degrees than other nodes.



Figure 4.20. Visualization of the US top 500 airport infrastructure network

### 1) MVC graph agent development and implementation

1,000 training networks by randomly removing a part of the nodes from the original airport network is generated. The number of nodes in the training networks ranged from 300 to 480 based on the observation of the number of functional nodes in the system during the recovery process. The number for different sizes of training networks is presented in Table 4.15.

Table 4.15 The percentage of training graphs with different sizes

| Number of nodes | Percentage of the networks accounted for total training networks (%) |
| --- | --- |
| 480 | 40 |
| 460 | 10 |
| 440 | 10 |
| 420 | 10 |
| 400 | 7.5 |
| 380 | 7.5 |
| 360 | 5 |
| 340 | 5 |
| 320 | 2.5 |
| 300 | 2.5 |

The average size of the MVC solution found by the trained agent (around 234) is better than that from the 2-opt solver in the networkX library, which is 294.

The maintenance actions are implemented once the cascading failures are triggered. MVC is calculated from the real-time BA network structure during the recovery process based on the trained MVC agent. Then, restoration priority targets are updated according to the current MVC and the other maintenance preference that is considered.

**2) Results and discussion**

Similar to the previous case study, the values of parameter $u$ are set to be top 10%, 30%, 50%, respectively, for simulation in this case study. Figure 4.21 shows the average resilience loss $\Re(T)$ and $T$ vs. *Nrem* for the proposed recovery framework and three existing restoration strategies. The results shown in Figure 4.21 under each *Nrem* are averaged over the results of 90 network realizations for each existing restoration strategy, whereas the results under the recovery framework considering MVC are averaged over 10 realizations because it is a fixed network topology. To reduce the bias result from network dependency, scenarios with different *D-size* (*D-size* = 4, 8, 12, 16, respectively) are considered, and the results are averaged for these different scenarios regarding *D-size*.
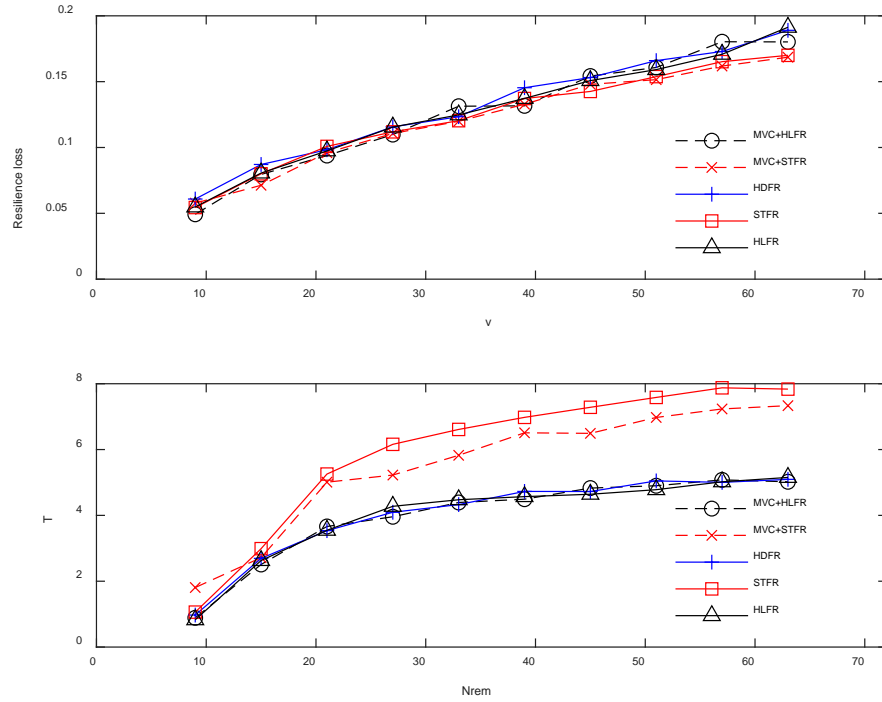
Figure 4.21. Average results of resilience loss and *T* as a function of *Nrem* for different

restoration strategies

Similar to what is observed from the results on BA networks, resilience loss $\Re(T)$

and *T* increase as *Nrem* increases in the airport network. Figure 4.21 shows that the

difference between the maintenance effects of different restoration strategies regarding

resilience loss are not as large as those in BA networks. However, the restoration

strategies that incorporate MVC importance (i.e., MVC+STFR and MVC+HLFR) still

contribute to less resilience loss in most of the *Nrem*. A shorter *T* is achieved by applying

HDFR, HLFR, and MVC+HLFR. A longer *T* is needed for network efficiency to be

recovered to the predetermined level when applying STFR, whereas *T* is shortened with

MVC+STFR.

It also can be seen from Figure 4.21 that the growth rate of $T$ is reduced when $v$ exceeds a noticeable threshold. The threshold of $v$ in the changing trend of $T$ is around 20 (i.e., about 4% of the total network nodes). This finding is in accordance with the results in [148] under different repair strength and dependency strength. It indicates the robustness of the US top 500 airport network to cascading failures regarding recovery time. Table 4.16 presents the average results regarding *D-size* for different scenarios under the existing restoration strategies and the proposed recovery framework incorporating MVC.

Table 4.16: Average results under different restoration strategies

| Restoration strategy | Average $T$ | Average resilience loss (%) |
|---|---|---|
| HDFR | 4 | 13.11 |
| HLFR | 4 | 12.83 |
| MVC+HLFR | 4 | 12.71 |
| STFR | 6 | 12.36 |
| MVC+STFR | 5.5 | 12.19 |

Based on Table 4.16, HLFR and STFR show better maintenance effects when considering global connectivity importance based on MVC detection, which is consistent with the results presented in Figure 4.16.

### 4.4.4 Summary

In this work, a network recovery framework that updates restoration priority targets based on the changing importance of system components during the process of cascading failures and restoration is presented. The recovery framework incorporates global network connectivity importance represented by MVC to existing restoration prioritization strategies to optimize the restoration priority at each round of inspection. Efficient MVC calculation from the updating network structure is the key step, and a

desirable solution through OpenGraphGym with graph embedding and deep reinforcement learning is provided.

Case studies are conducted on BA networks and the US top 500 airport network by employing the proposed recovery framework against cascading failures. MVC estimation is incorporated into two existing restoration strategies with different restoration prioritization weights. Restoration effects in terms of resilience loss regarding network load and recovery time are remarkably improved by considering real-time MVC in the restoration prioritization strategy, which demonstrates the effectiveness of the proposed recovery framework.

This work supports the necessity and significance of updating the restoration priority based on the importance of system components during system recovery. The recovery framework can be applied to other real-life network systems in the next work and will include monetary costs and benefits to make a real-world optimization problem. To make the framework more practical, quick detection of MVC in a large network system is one direction of the future work. How to determine the importance of MVC during joint restoration prioritization remains an unexplored question, which is complicated by extending to a general network topology. This question will be another direction of future research.

## 5   Resilience enhancement with optimal backup energy system

Power systems with solar photovoltaic (PV) arrays combined with battery storage are becoming increasingly used because of their capability of working in power island mode, especially during power grid outages, to provide backup energy for customers. Thus, PV + battery system, as an example of DER, can be installed as a hardening measure to

enhance the resilience of electricity system subject to cascading failures. Although a battery has many advantages, its current price is still relatively high when compared with that of other types of energy storages. In this work, the benefit of installing PV + battery system as a hardening measure to help customer access electricity when grid outage occurs is explored. Both power resilience and economic factors are taken into account in order to obtain optimal solution. The impacts of the changing trends of some important factors, for example, battery price, on system optimization to achieve resilience and economic benefits are also investigated.

**5.1 Economic and resilience benefit of incorporating battery to photovoltaic array**

First, the research focuses on optimizing battery sizes for PV + battery systems with given solar arrays, i.e., solar array is already installed. The optimal investment on battery storage is to make the backup system provide power resilience economically.

**5.1.1    Problem formulation and main assumptions**

An optimization method based on PV + battery system operation simulation during grid interruption is developed to investigate the effects of battery size on system output reliability level of meeting load demand with minimum cost. The main variables and parameters considered in the optimization problem are as follows:

1) System total cost (*STC*), which is the objective of the optimization problem that needs to be minimized. It has two parts, investment cost of adding a battery, and loss of load cost which is the incurred cost of unsatisfied load demand of customers.

2) Solar irradiation variation, which represents the solar radiation input of solar PV array. In this part, real solar insolation data of hourly GHI [149] (global

horizontal irradiance, combining both direct and scattered light hitting a level local reference plane) is taken into account.

3) Load demand variation, which is the electricity usage demand of customers. Actual data of load demand of specific model facilities from [150] are applied in this part.

4) Loss of load probability (*LOLP*), which is calculated as the proportion of time when load demand of facilities cannot be met by PV + battery system during grid outage. It is adopted as one of the reliability metrics for the electricity supply of PV + battery system, which has already been used by some researchers to develop sizing algorithms to minimize system cost [151, 152].

5) Chance constraint probability (*CCP*), which is estimated as the proportion of grid outages in which the *LOLP* constraint is satisfied by the islanding operation of the PV + battery system with a specific battery size, i.e., an estimate of the probability that the *LOLP* constraint can be met by the PV + battery system energy supply during system outage.

The optimization problem is to minimize system total cost incurred when a PV + battery system works in an island mode during grid outage. The corresponding optimization formulation is present as follows [153].

$$\min \quad B_r b + \sum_{k=1}^{T_y} \left( VOLL \sum_{j \in N_k} \sum_{i \in O_{kj}} AEL(t_{kji}) \right) (1+r)^{-k}$$

$$(5.1)$$

$$\text{s.t.} \quad AEL(t_{kji}) = \begin{cases} D(t_{kji}), & \text{if } Q_B(t_{kji}) + \int_{t_{kji}}^{t_{kji}+\Delta t} \left( P(u) - D(u) \right) e\, du < B_{\min} \\ 0, & \text{otherwise} \end{cases}$$

$$Q_B(t_{kji} + \Delta t) = \begin{cases} \min \left\{ Q_B(t_{kji}) + P(t_{kji})e, \ B_r \right\}, & \text{if } Q_B(t_{kji}) + \int_{t_{kji}}^{t_{kji}+\Delta t} \left( P(u) - D(u) \right) e\, du < B_{\min} \quad (5.2) \\ Q_B(t_{kji}) + \int_{t_{kji}}^{t_{kji}+\Delta t} \left( P(u) - D(u) \right) e\, du, & \text{if } B_{\min} \le Q_B(t_{kji}) + \int_{t_{kji}}^{t_{kji}+\Delta t} \left( P(u) - D(u) \right) e\, du \le B_r \\ B_r, & \text{if } Q_B(t_{kji}) + \int_{t_{kji}}^{t_{kji}+\Delta t} \left( P(u) - D(u) \right) e\, du > B_r \end{cases}$$

$$P(t_{kji}) = \eta I(t_{kji}) A, \quad \forall t_{kji}, \quad k \in T_y, \quad j \in N_k, \quad i \in O_{kj} \tag{5.3}$$

$$B_{\min} = B_r \left( 1 - DoD \right) \tag{5.4}$$

$$E[LOLP] = \frac{\sum_{k=1}^{T_y} \sum_{j \in N_k} LOLP_{kj}}{\sum_{k=1}^{T_y} |N_k|} \le \beta, \quad \text{where} \quad LOLP_{kj} = \frac{\sum_{i \in O_{kj}} AEL(t_{kji})}{|O_{kj}|} \tag{5.5}$$

$$CCP = \Pr \left\{ LOLP_{kj} \le \gamma \right\} \ge 1 - \tau, \quad k \in T_y, \quad j \in N_k \tag{5.6}$$

$$0 \le \tau, \ \beta, \ \gamma, \ DoD \le 1, \ B_{\min} \ge 0, \ B_r \ge 0$$

The objective function consists of two items, $B_r b$ which represents battery

investment cost and $\sum_{k=1}^{T_y} \left( VOLL \sum_{j \in N_k} \sum_{i \in O_{kj}} AEL(t_{kji}) \right) (1+r)^{-k}$ which denotes the net present

value of total cost of unsatisfied load demand during grid failure. $B_r$ is battery capacity

(kWh), and $P_b$ denotes battery price (\$/kWh). $VOLL$ is the penalty cost of lost load

(\$/kWh), which is a typical monetary expression for the costs caused by electricity

interruptions [154, 155]. $T_y$ is the planning time horizon in years, and $r$ is the discount rate. $N_k$ is the random set of outages in year $k$, while $O_{kj}$ is the random set of time intervals for outage $j$ in year $k$. $t_{kji}$ denotes the $i$th time interval for the $j$th outage in year $k$. Time is a discrete parameter in the formulation with increments of $\Delta t$, which means one hour in this formulation.

Eq. (5.1) presents the amount of load demand, $AEL(t_{kji})$, which is not satisfied by the PV + battery system at time $t_{kji}$. $B_{min}$ is minimum battery energy value (kWh). $Q_B(t_{kji})$ is the energy stored in battery (kWh) at time $t_{kji}$. $P(t_{kji})$ is the power generated by PV array (kW) at time $t_{kji}$. $D(t_{kji})$ is load demand (kW) at time $t_{kji}$. $e$ represents the efficiency of discharging/charging process of battery. If the load demand for an outage hour is fully satisfied by the PV + battery system, there is no loss of load cost for this hour. However, if the hourly load demand is not fully satisfied, then it is assumed that the whole demand for this hour is lost. Eq. (5.2) presents the change of energy stored in battery over a short time period $\Delta t$ due to power generation of PV array and load demand of facilities. It describes the energy balance of the PV + battery system, and many previous methods are proposed based on this topic [156]. Eq. (5.3) denotes the energy produced by the PV array at time $t_{kji}$. $\eta$ is the energy conversion efficiency of PV array. $I(t_{kji})$ is solar irradiation (W/m$^2$) at time $t_{kji}$. $A$ is PV array area (m$^2$). Eq. (5.4) denotes how $B_{min}$ is calculated with $DoD$, which is the abbreviation of the maximum depth of discharge of a battery (%). Eq. (5.5) defines the calculation of $LOLP$ of each grid outage and the expected value of $LOLP$, which indicate the islanded generation reliability of the PV + battery system. $\beta$ is the upper bound for the expected value of $LOLP$. $LOLP_{kj}$ is defined as the proportion of time when load demand of customers cannot be met by PV + battery

system during the *j*th grid outage of year *k*. Eq. (5.6) presents the definition of *CCP*, which incorporates the *LOLP* constraint for each grid outage. The *LOLP* constraint for each grid outage is limited to be not more than $\gamma$. $1 - \tau$ denotes the minimum desired *CCP*. It is apparent that the smaller $\tau$, $\beta$ and $\gamma$ are, the stricter power supply reliability constraints are.

Like other industrial systems, the life cycle of a PV + battery system should be taken into consideration when analyzing cost-benefit problem. Whereby, the investment of battery is designed for 20-year operation of a PV + battery system in the simulation model [157], i.e., decisions for the PV + battery system are made considering 20 years of operation, and the corresponding loss of load cost is accumulated for these 20 years. It should be noted that loss of load cost in each year is calculated as the present value with discount rate of 4% in order to be considered as one part of objective function, because the investment of battery is at the beginning of the first year.

Three metrics are used to evaluate the islanding capability of the PV + battery system from both economical and reliability perspectives, system total cost, *CCP* and system achieved *LOLP*. System achieved *LOLP* denotes the actual expected *LOLP* achieved by islanding mode of PV + battery system during 20 years of operation. *CCP* and system achieved *LOLP* together represent the reliability level of the PV + battery system power supply during grid outages. Note that each curve shown in the following figures are the average results over 100,000 simulations.

It is assumed that grid outage duration follows a shifted or scale adjusted Poisson distribution in which the duration is larger than 0, and the expected outage duration is approximately equal to CAIDI. The number of outages per year equals to SAIFI. Note

that CAIDI is applied with the unit of hours in this section. The starting time of a grid outage is assumed to be uniformly distributed throughout a year. It is also assumed that *VOLL* is uniformly distributed in a specific range, which is determined based on the criticality of facilities. The value of *VOLL* is randomly selected within a range because the economic impact of each outage varies depending on the specific functions or unique characteristics of different facilities.

## 5.1.2    Simulation process

The main procedures of time series simulation of PV + battery system working in island mode during grid outage are shown in Figure 5.1. The main steps of the proposed simulation method are described as follows,

Step 1)    Select the values of random parameters, i.e., the starting time of each grid outage, outage duration and number of grid outages each year.

Step 2)    Incrementally enumerate and evaluate for all considered array sizes.

Step 3)    Incrementally enumerate and evaluate for all simulations of grid outages.

Step 4)    The hourly electricity generation of PV array during the outage is calculated based on solar irradiation data.

Step 5)    The *VOLL* is chose in a specified range for each grid outage.

Step 6)    Simulate the charging and discharging process of the battery over outage duration. The amount of lost load and the associated cost are then computed.

Step 7)    Return to Step 3.

Step 8)    Compute *STC* and reliability metrics (i.e., system achieved *LOLP* and *CCP*) for the specific array size.

Step 9)    Return to Step 2.
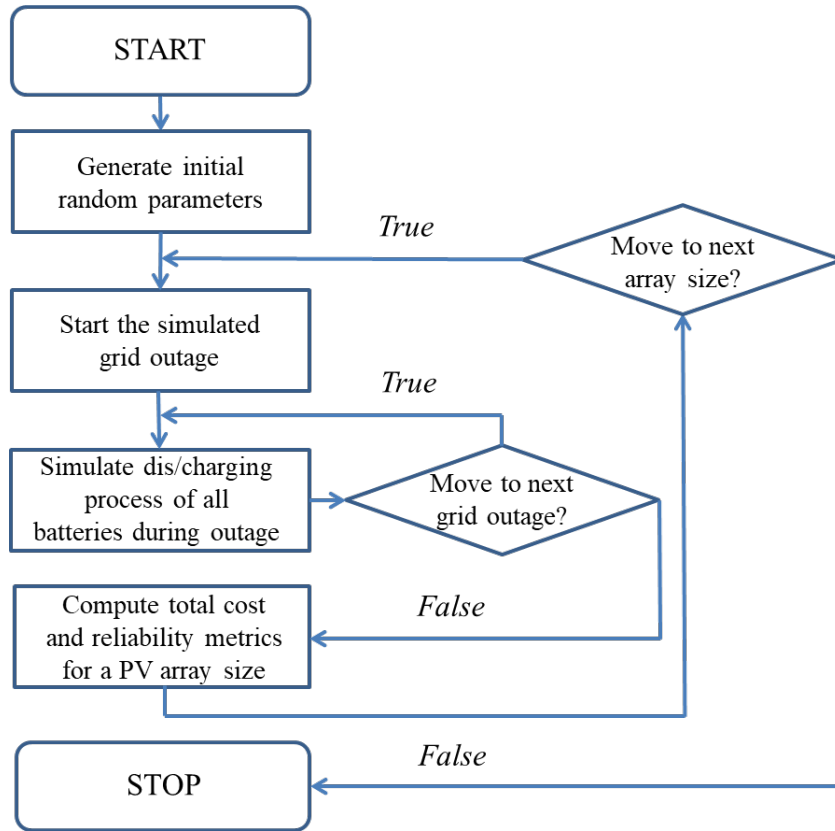
Figure 5.1. Diagram of the main simulation steps

### 5.1.3 Numerical examples and analysis

In this section, numerical examples are performed based on the simulation-based optimization method. A hospital facility, subject to grid interruption and served only by the PV + battery system during outage, is selected for demonstration purposes. The hospital is located in Islip, Long Island, NY, U.S.

The existences of natural and unavoidable variabilities of some realistic factors that are associated with PV + battery system operation have been analyzed by some researchers [158, 159]. Here the impact of an important factor, battery price, on optimal battery sizing of PV + battery systems is explored. Furthermore, sensitivity analysis of the relationship between *STC* and power supply reliability of the PV + battery system is

carried out.

**1) Case study**

For the following numerical examples, the current battery price equals 162 $/kWh [160]. *VOLL* is within the selected 90% confidence-level (*CL*) range of (5 $/kWh-25 $/kWh), which is based on the expectation of *VOLL* level in 2030 [161]. The solar radiation profile and load profile of the facilities that used in this part are for Islip, Long Island, NY, U.S. [149, 150]. It is commonly assumed that the generator efficiency remains constant [162], here the PV array conversion efficiency is set to be 16% [163]. Efficiency of the battery charging/discharging process is set to be 85%, and *DoD* is 70%. Values of CAIDI and SAIFI are 8.23 and 1.17 for the Long Island, NY area [164].

The analysis is performed first on the hospital application. The selected, model hospital is a large (average load demand is 1,138 kW) and critical facility, so *VOLL* for this facility is selected between 22 $/kWh and 25 $/kWh. The constraint for system achieved *LOLP* is set to be 10%. The optimization model is to determine optimal battery size for a PV array that has already been selected or installed. Here three specific solar array sizes, 200 $m^2$, 5200 $m^2$ and 10,200 $m^2$ are considered with battery capacities ranging from 200 kWh to 20,000 kWh. It should be noted that each grid outage is considered to be independent.
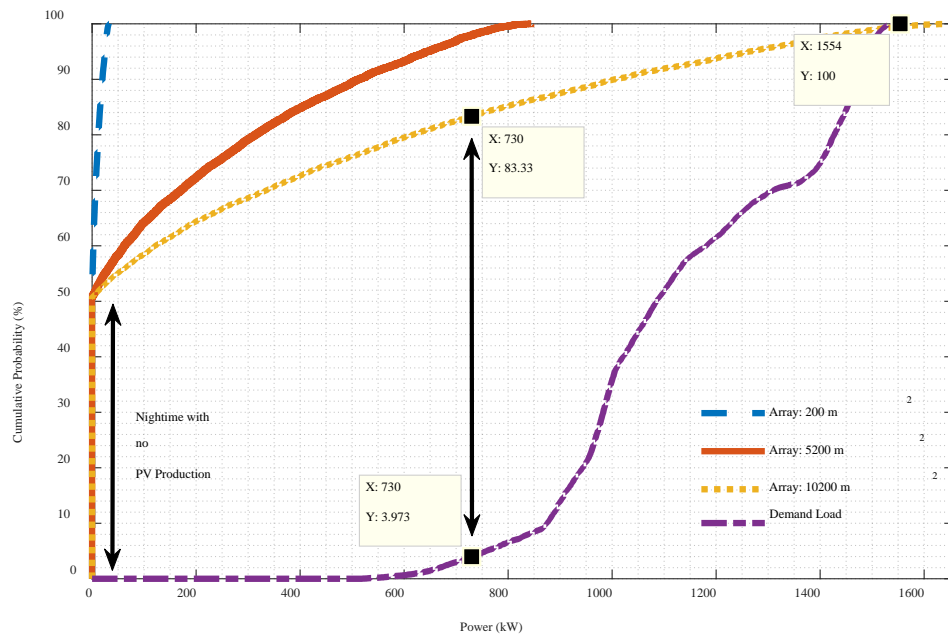
Figure 5.2. Probability distributions of hourly PV generation and demand power through

an entire year

Figure 5.2 shows the probability distributions for hourly energy generated by the PV array and hourly load demand vs. the fraction of hours in a year that experience less PV energy generation or less load demand. As the labeled dots show, the probability that a randomly chosen hourly PV generation from the largest PV array size of 10,200 m$^2$ is less than 730 kW is about 83%, while the probability that a randomly chosen hourly demand power is less than 730 kW is around 4%. Thus, the hourly energy generated by the three considered solar array sizes can rarely satisfy the hourly load demand. Note that since there is no sunlight during night, there is an approximately 50% probability that a randomly chosen PV hourly energy generation is 0 kWh.
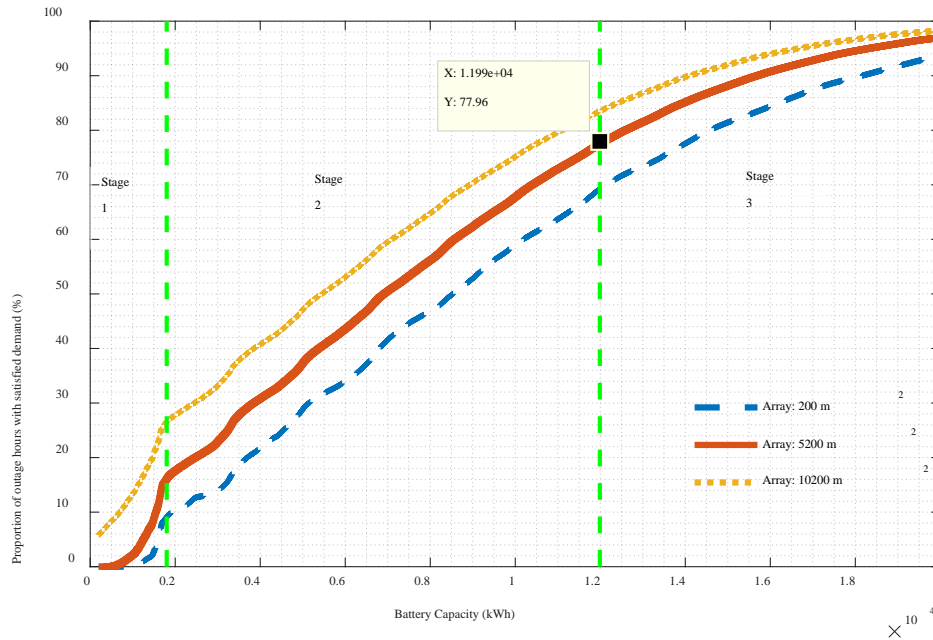
Figure 5.3. Proportion of outage hours when load demand is satisfied by different PV +

battery systems through all simulated grid outages

Figure 5.3 presents the simulation results that the proportion of outage hours when

load demand is satisfied by the PV + battery system as a function of battery capacity

under three considered PV array sizes. As it can be observed, the increasing trend of each

curve can be divided into three different stages. Stage 1 shows the initial sharp increasing

trend of the curves. This sharp increase can be explained that the proportion of outage

hours when load demand is satisfied increases remarkably if a battery is added to a PV

system. The larger the battery capacity, the higher the proportion of outage time when

load demand is satisfied. According to Figure 5.2, the maximum hourly load demand for

this hospital is 1,554 kW. Thus, if the battery capacity exceeds 1,554 kWh, the fully

charged battery alone can provide enough energy for a single hour outage, whereas

smaller batteries might fail to do this unless the solar array output is large enough to fill

the gap for the first hour (and perhaps successive hours). Then stage 2 comes as battery

capacity increases, which shows a reasonably stable linear relationship between battery capacity and the proportion of outage hours when load demand is satisfied. When battery capacity goes beyond the battery capacity corresponding to the lowest *STC*, the growth trend of the curves becomes flatter. The proportion of outage time when load demand is satisfied finally goes up to near 100 % as battery capacity continues to increase, which denotes stage 3.
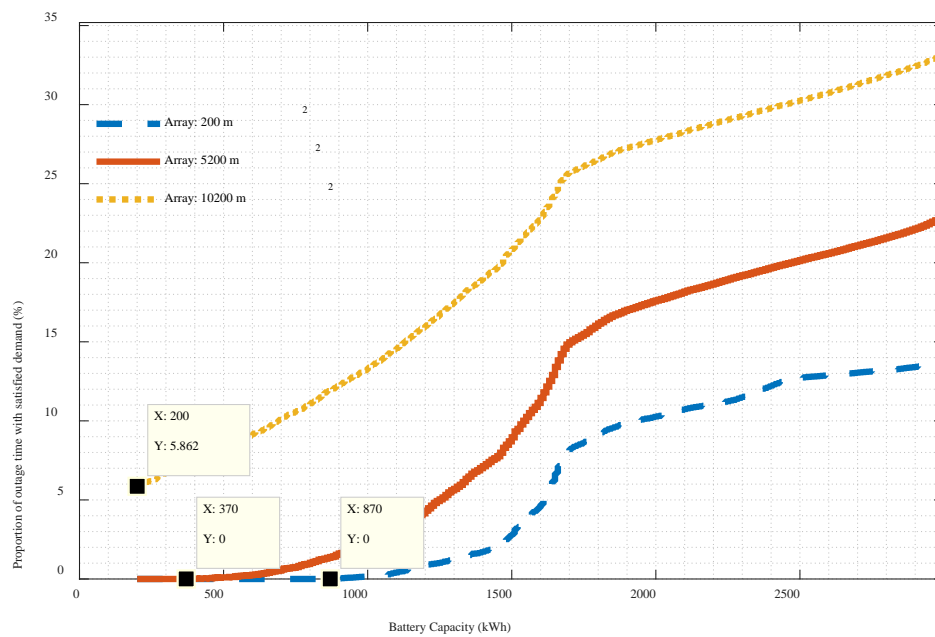


Figure 5.4. The enlarged view of initial part of curves in Figure 5.5

The initial part of Figure 5.3 is illustrated in an enlarged view in Figure 5.4, which highlights the change of slope after battery capacity becomes large enough to meet average hourly load demand. Besides, generation of the increasing sized solar array compensates for batteries that are not big enough.
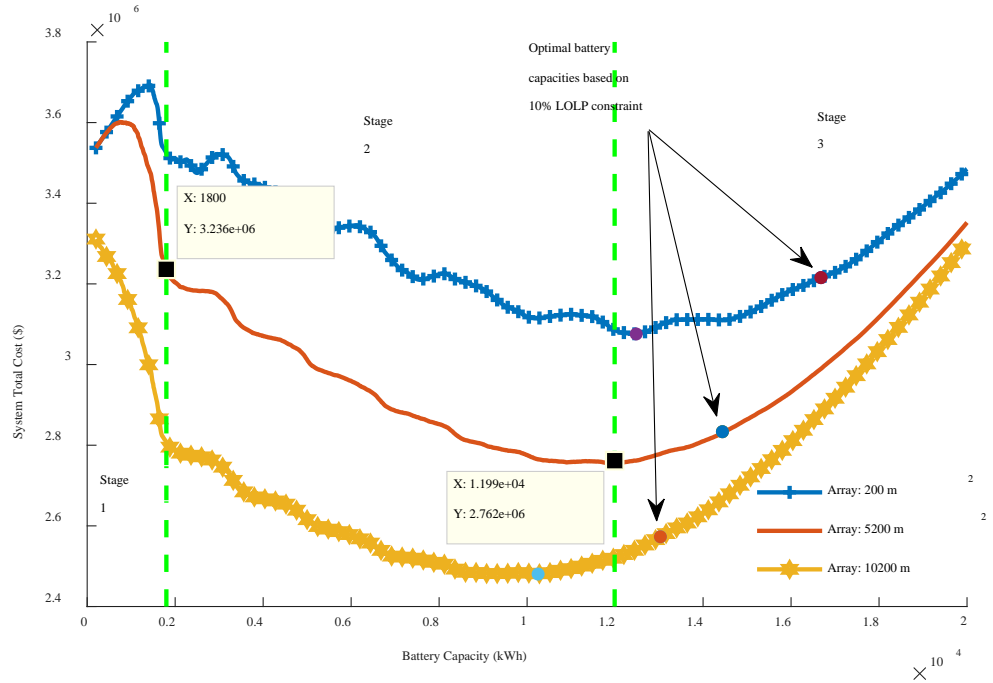
Figure 5.5. Total cost for islanding operation of PV + battery systems with different

battery capacities and array sizes

*STC* is graphed in Figure 5.5 which shows the three distinct stages as well. In stage 1, it can be seen that *STC* for the cases with 200 m$^2$ and 5,200 m$^2$ array sizes actually increases when a small battery is added to the PV system. It is attributed to the large difference between load demand and PV energy generation as shown in Figure 5.2. As it can be seen from Figure 5.4, until battery capacity goes up to around 900 kWh for the array size of 200 m$^2$ and 400 kWh for the array size of 5,200 m$^2$, the proportion of outage time when load demand is satisfied stays at 0%. Thus, the addition of a small battery does not help to noticeably reduce the unsatisfied load demand, and the increasing investment of battery contributes to a higher system total cost. After battery capacity exceeds these threshold values (as it is shown in Figure 5.4), the proportion of outage time with satisfied demand increases as battery capacity increases, and the obtained significant reduction of unsatisfied load demand drives system total cost down sharply. Considering

the case with 10,200 m$^2$ array size, the initial system total cost without investment of battery (i.e., only loss of load cost incurred) is much smaller than that of other two cases, since the PV system with an array size of 10,200 m$^2$ alone can approximately satisfy the load demand of 4% of the outage time. According to Figure 5.4 and Figure 5.5, the increasing battery capacity continually decreases system total cost in stage 1 for this large-array case.

Considering an array size of 5,200 m$^2$ as an example, after the battery capacity exceeds the threshold value of 1,800 kWh, it enters into stage 2, where the decreasing rate of system total cost slows down and there is a linear trend as battery capacity increases. It corresponds to the stage 2 indicated in Figure 5.3, where the proportion of outage time when load demand is satisfied increases linearly with the increase of battery capacity. After the battery capacity exceeds about 12,000 kWh, as it is labeled in Figure 5.5, it comes to stage 3, where the increasing investment cost of adding more battery capacity leads to a higher system total cost instead. This is because increasing battery capacity in stage 3 does not increase the proportion of outage hours when load demand is satisfied as efficiently as it does in stage 2, as it is shown in stage 3 in Figure 5.3.

Table 5.1. Total cost of PV + battery system operation in island mode with combination of battery and array sizing

| System total cost ($\times 10^6$\$) | | Array size (m$^2$) | | |
|---|---|---|---|---|
| | | 200 | 5,200 | 10,200 |
| Battery Capacity (kWh) | 200 | 3.54 | 3.54 | 3.31 |
| | 10,250 | 3.11 | 2.77 | **2.48** |
| | 11,990 | 3.09 | **2.75** | 2.52 |
| | 12,480 | **3.08** | 2.76 | 2.54 |

Note: Values in bold are the lowest system total cost corresponding to each PV array size.

Table 5.1 presents the system total cost for the three considered array sizes and the

battery capacities which correspond to the lowest system total cost for each specific array. The smallest battery tested (200 kWh) is also included as this nearly represents the baseline lost-load system cost as a comparison. Based on the results in Table 5.1, it can be observed that the lowest system total cost for a given PV array size decreases as the PV array size increases, and the corresponding battery capacity decreases as well. This is because a larger PV array generates more electricity (subject to the sunlight variability), which contributes to less loss of load cost leading to a lower system total cost and a smaller required battery capacity to address power needs during an outage. All of these minimum cost battery/array configurations yield substantial financial savings compared to the baseline. The optimal battery capacities satisfying 10% *LOLP* constraints are also labeled in Figure 5.5, which, for the present scenario, are larger than the battery capacities corresponding to the lowest system total cost.
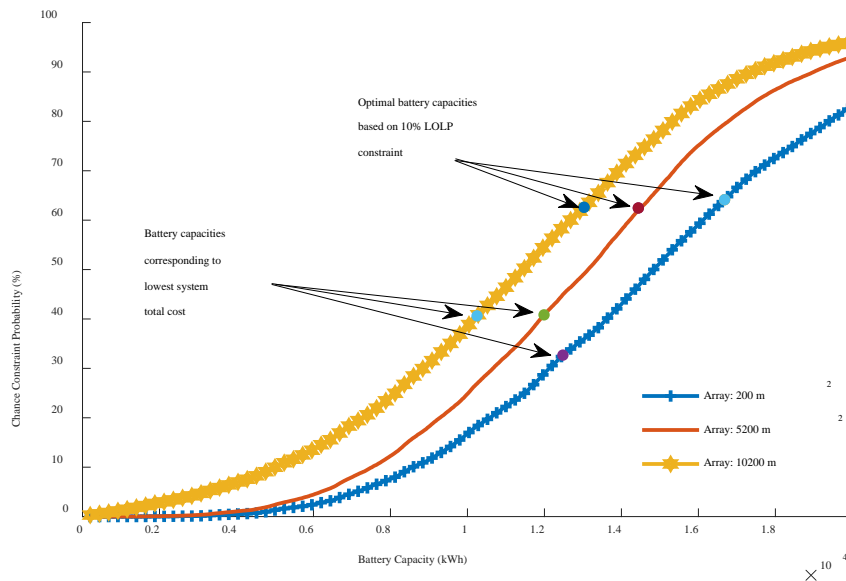


Figure 5.6. Chance constraint probability for islanding operation of PV + battery system with different array sizes

Table 5.2. Chance constraint probability of PV + battery system islanding operation with different combinations of battery and array sizing

| **Chance constraint probability** (%) | | Array size (m$^2$) | | |
|---|---|---|---|---|
| | | 200 | 5,200 | 10,200 |
| Battery Capacity (kWh) | 10,250 | 18.32 | 26.90 | **40.63** |
| | 11,990 | 28.92 | **40.84** | 54.65 |
| | 12,480 | **32.69** | 44.95 | 58.57 |

Note: Values in bold are chance constraint probabilities corresponding to each scenario with the lowest system total cost.

It can be seen from Figure 5.6 that the optimal battery capacities which meet the 10% *LOLP* constraint are larger than the battery capacities which achieve the lowest system total cost. Higher chance constraint probabilities, which indicate more reliable islanding energy supply of PV + battery systems, could also be achieved with these optimal battery capacities.
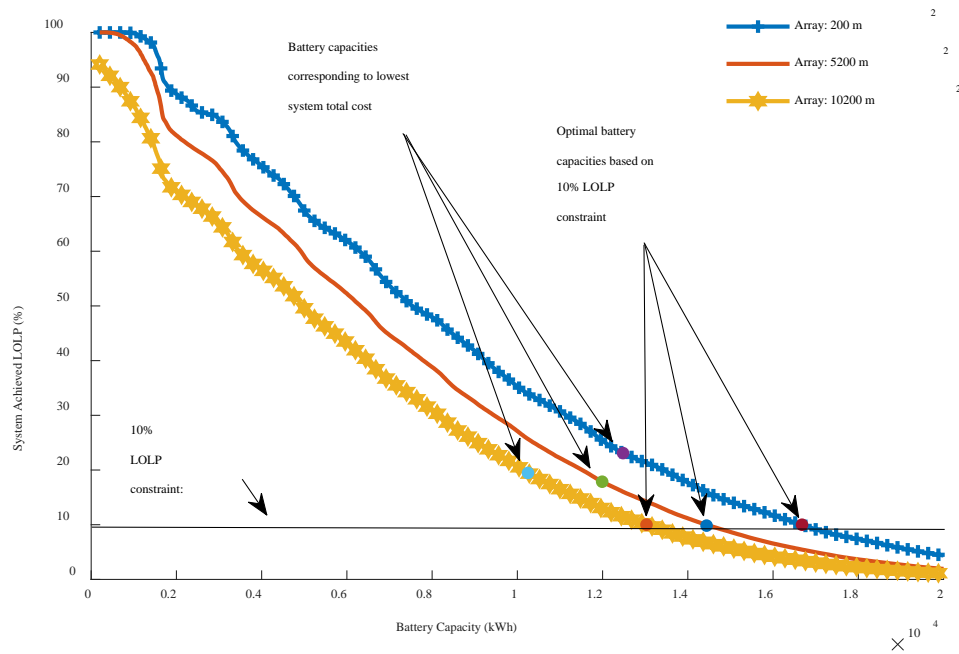


Figure 5.7. Achieved *LOLP* for islanding operation of PV + battery system with different array sizes

Table 5.3. Achieved *LOLP* of PV + battery system islanding operation with combination

of battery and array sizing

| System achieved *LOLP* (%) | | Array size (m$^2$) | | |
|---|---|---|---|---|
| | | 200 | 5,200 | 10,200 |
| Battery Capacity (kWh) | 10,250 | 33.95 | 25.67 | **19.46** |
| | 11,990 | 25.39 | **17.86** | 12.95 |
| | 12,480 | **23.09** | 16.02 | 11.44 |

Note: Values in bold are system achieved *LOLP* corresponding to each scenario with optimal system total

cost.

The upper labeled dots in Figure 5.7 are system achieved *LOLP* corresponding to the

lowest system total cost. System achieved *LOLP* denotes the real *LOLP* achieved by

islanding mode of PV + battery system during grid outage. The optimal battery capacities

which meet the 10% *LOLP* constraint are also highlighted in the figure, which (in this

case) are larger than the battery capacities achieving the lowest system total cost.

### 2) Impact of battery price

This analysis is performed by exploring the scenario of a decrease in battery price,

while other parameters stay the same as they are used before. Results are shown in Figure

5.8 to 5.10, which are compared with the figures obtained in the above case study section.

The impact of battery price is investigated in terms of system total cost, chance

constraint probability and system achieved *LOLP*. Battery price is changed from 162

$/kWh (the current price) to 74 $/kWh (a forecasted price, which comes from Bloomberg

New Energy Finance for 2030 [160]). For the following Figure 5.8 to 5.10, Figures 5.8

(a), 5.9 (a) and 5.10 (a) are obtained by using the current battery price (162 $/kWh) while

Figures 5.8 (b), 5.9 (b) and 5.10 (b) are obtained by adopting the forecasted battery price

(74 $/kWh).

The simulation results presented in Figures 5.8 to 5.10 show the difference in system total cost, chance constraint probability and system achieved *LOLP* caused by a likely future reduction in battery prices.
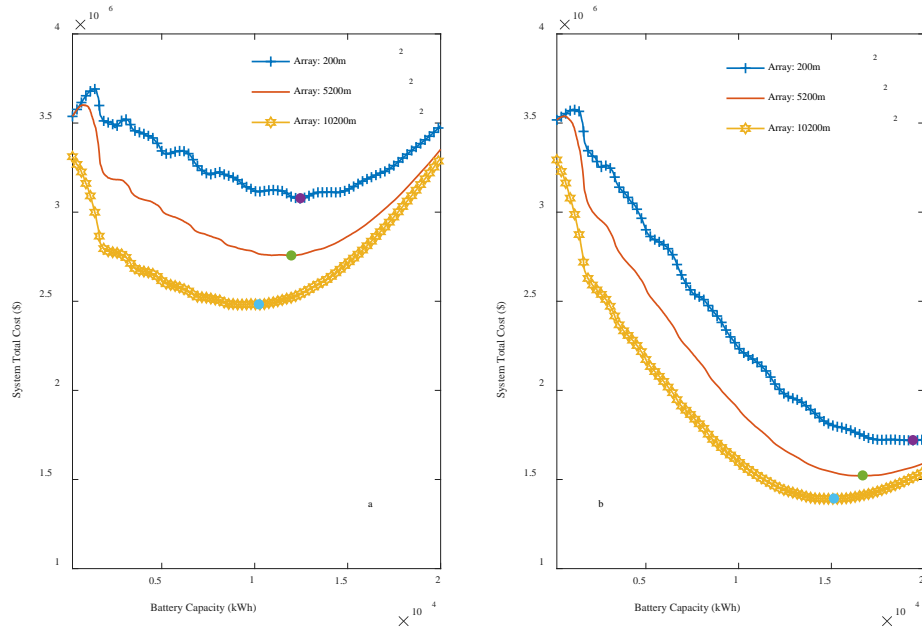


Figure 5.8. Total cost for islanding operation of PV + battery system with different array sizes vs. battery capacity for the case of current battery price 162 $/kWh (a) and the case of forecasted battery price 74 $/kWh (b)
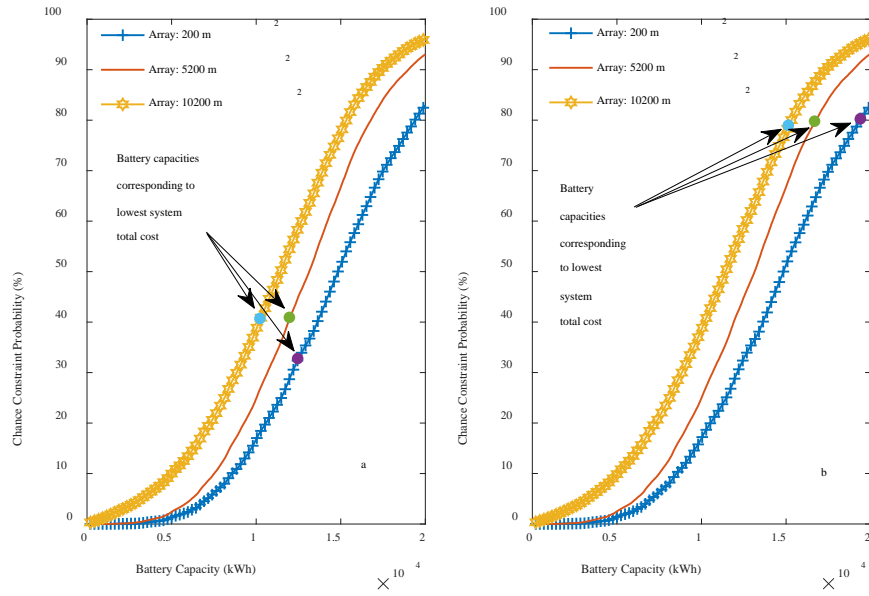
Figure 5.9. Chance constraint probability for islanding operation of PV + battery system with different array sizes vs. battery capacity for the case of current battery price 162 $/kWh (a) and the case of future battery price 74 $/kWh (b)
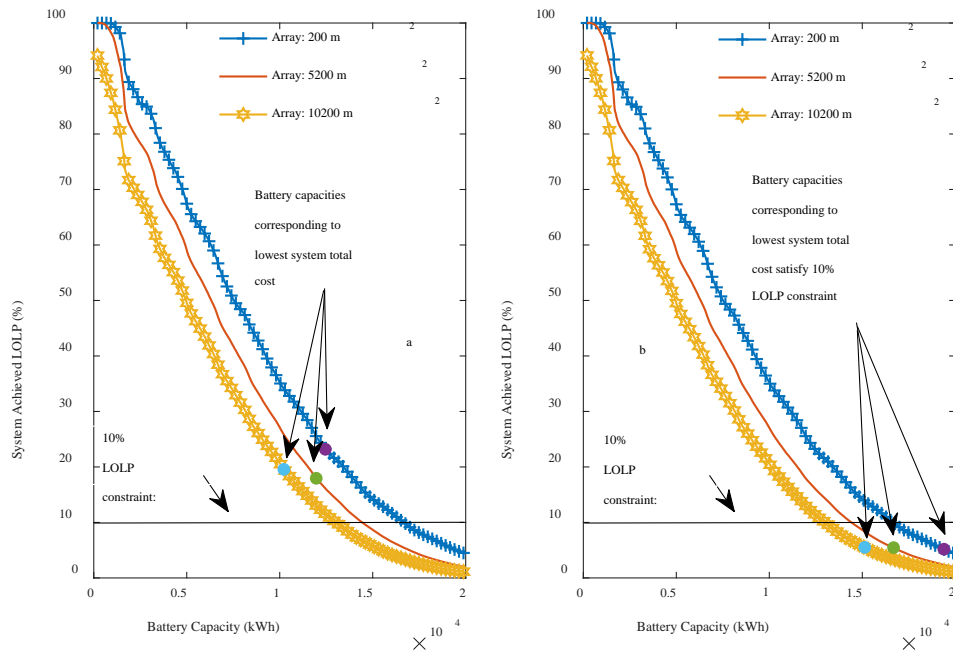


Figure 5.10. Achieved *LOLP* for islanding operation of PV + battery system with

different array sizes vs. battery capacity for the case of current battery price 162 $/kWh (a)

and the case of future battery price 74 $/kWh (b)

The simulation results showed in Figures 5.8 to 5.10 illustrate the effect that battery price has on system total cost and system power supply reliability. The change in terms of system total cost is clear, as it can be seen from Figure 5.8. The system total cost figure, shown in Figure 5.8 (b), becomes less symmetrical, and the lowest system total cost is achieved for larger battery capacities compared to the previous case shown in Figure 5.8 (a). Chance constraint probability corresponding to the lowest system total cost with PV array size 10,200 m$^2$ changes from 40.63% to 78.88%. System achieved *LOLP* with the lowest system total cost changes from 19.46% to 5.41%. The most important thing to notice in Figure 5.10 (b) is that the battery capacities corresponding to the lowest system total cost are actually the optimal battery capacities based on the 10% *LOLP* constraint.

The reason for these changes can be found in the definition of system total cost in the formulation. System total cost consists of two parts, investment cost of battery capacity and loss of load cost. Investment cost of battery capacity decreases as the battery price decreases, which means system total cost depends more on loss of load cost than before, so the need to meet the demand becomes more urgent.

### 3) Sensitivity analysis

A sensitivity analysis has also been conducted in terms of system total cost and the two reliability metrics. The relationship between every 1% decrease of system achieved *LOLP* (or 1% increase of chance constraint probability) and the corresponding change of system total cost has been investigated, i.e., how system total cost changes (higher or lower) is examined while the reliability performance of the system is improving. One specific array size is taken as an example. Figure 5.11 and Figure 5.12 are provided to demonstrate sensitivity analysis results.

Figure 5.11 shows the results of sensitivity analysis of system total cost vs. system achieved *LOLP*. An example of how Figure 5.11 should be interpreted as the following. The circle point (13, 12,490) indicates that a positive cost increase of $12,490 will be incurred if system achieved *LOLP* decreasing from 13% to 12% is desired because more battery capacity would need to be purchased.
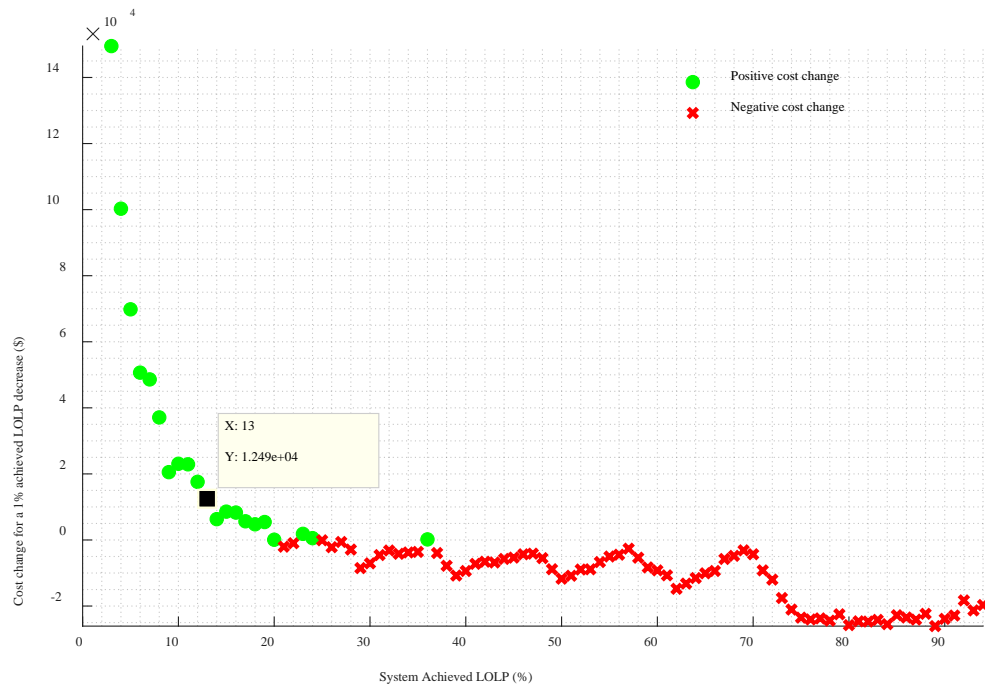


Figure 5.11. Sensitivity analysis of cost change and achieved *LOLP*

Figure 5.12 (a) shows the results of sensitivity analysis of system total cost and chance constraint probability. An important characteristic that can be observed from both Figure 5.11 and Figure 5.12 (a) is the presence of cross points, i.e., points where the system reliability improvement is accompanied with a decrease in system total cost. That is justified by simultaneously looking at Figure 5.12 (a) and Figure 5.12 (b).
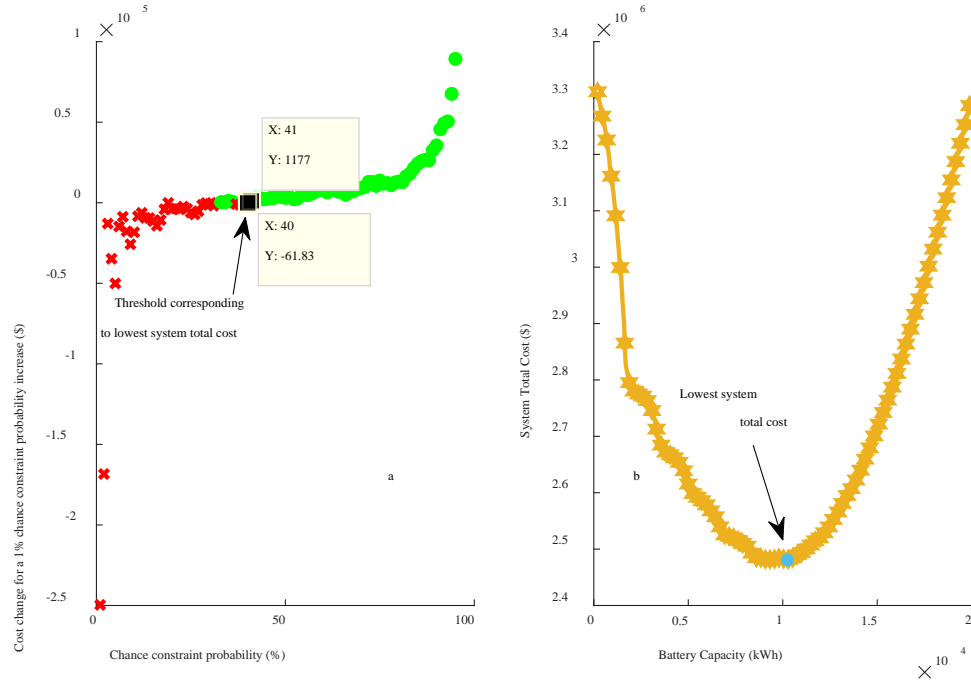
Figure 5.12. (a) Sensitivity analysis of cost change and chance constraint probability. (b) System total cost vs. battery capacity

It can be observed that the lowest system total cost in Figure 5.12 (b), which is labeled with an arrow, corresponds to the threshold where cross points end and circle points begin in Figure 5.12 (a). It demonstrates the argument that the optimal battery size with respect to system power output performance should be always greater than or equal to this threshold value, given the fact that chance constraint probability is a non-decreasing function of battery capacity. As a result, a battery which is smaller than this threshold value not only causes worse system islanding operation capability, but also incurs higher system total cost.

**4) Scaling of battery sizing**

Another possible extension of the proposed method is scaling the simulation results, i.e., divide both battery capacity (kWh) and system total cost ($) by the annual average hourly load demand (power) of the considered facility (kW). By using this conversion,

the simulation results obtained from different facilities might be compared. A necessary assumption here is that the facilities under comparison should have a reasonably similar load profile throughout the whole year.

The comparison is extended to add three other facilities to test the extension of the proposed method. An illustration of the above-mentioned procedure is given in Figure 5.13, by presenting the simulation results from four different types of facilities. The considered facilities are a hospital, a large hotel, a primary school, and a small office all located in Islip, Long Island, NY, U.S., though they each have different load profiles and *VOLL* significance.
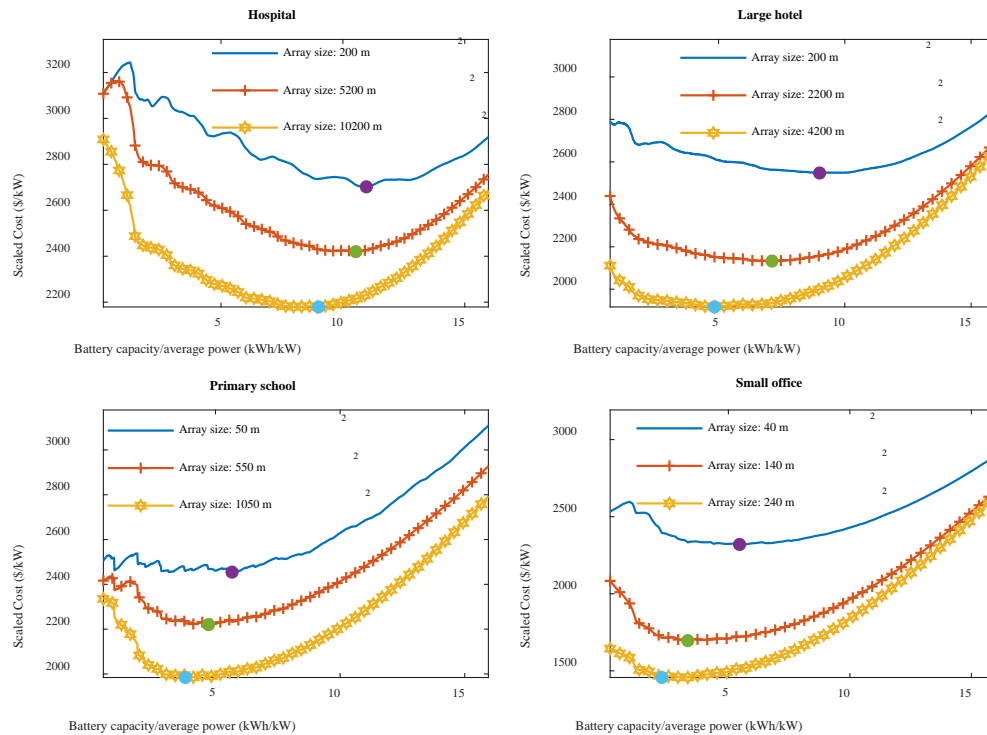


Figure 5.13. Scaled system total cost vs. battery capacity/average hourly load demand (kWh/kW) for all four facilities

Based on Figure 5.13, it can be noticed that although the range of the lowest system

total cost for each facility is distinct, the scaled optimal battery capacity for each facility is quite similar. Therefore, by using the procedure described in this section, decision-makers could quickly make comparisons between a reference facility and a facility of their interest with a given PV + battery system from output reliability and economic perspectives.

Overall, the finding shows that the decline of battery price not only reduces the lowest system total cost but also improves islanded system generation reliability. The proposed methodology for optimizing battery capacity added to PV array systems can make them grid-outage resilient and economically viable, which can be utilized as a decision-making tool for future PV + battery system expansion.

Based on the preliminary results, following are my current research contributions:

1) A mixed cascading failure model is proposed, which can be applied to properly simulate the cascades in network systems, like power grids, caused by load dynamics and node dependency. Moreover, statistical distributions are used to establish dependence clusters of network nodes for the modeling of node dependency, which is consistent with real-world functioning networks. Consequently, the proposed model can be used as a basis for analyzing cascading failures and their impacts on network system performance, which provide a better understanding of the process of failure propagation in real-life networked systems.

2) The effects of different restoration strategies regarding restoration strength are evaluated based on system resilience loss during the process of mixed cascading failures. How system dependency, which is described by multiple dependence characteristics, influences the effects of restoration actions against cascading

failures is also investigated. Accordingly, optimal system restoration strategies for system resilience enhancement by effectively reducing the intensity and extent of cascading failures can be designed for specific real-world network systems.

3) A mathematical model has been formulated to optimize battery capacity for a PV + battery system starting with a given PV array size in the face of grid outage. Moreover, a simulation-based optimization method has been developed according to the mathematical model. Based on the proposed approach, the optimum system is able to meet load demand of facilities during grid outage for a certain reliability level with minimal cost. As a result, the presented methodology is helpful for future design of grid outage-resilient PV + battery systems from both economical and reliability perspectives.

## 5.2 Interplay between storage investment and resilience loss considering economic trends

In this section, the influence of battery price and *VOLL*, affecting PV + battery system optimal sizing, are explored from power supply resilience and economic perspectives. A case study which considers a PV + battery system working in island mode following a grid outage is modeled, according to a simulation-based optimization method [153]. The findings can be used for the future optimal design of PV + battery systems taking into account the trade-off between the impact of battery price and *VOLL*.

### 5.2.1   Methodology

First, the modeling of power grid outages is introduced. The rigorous mathematical formulation of PV + battery system optimization regarding energy supply resilience

during grid outages is presented. Then the simulation-based method to implement system optimization and main assumptions are presented.

### 5.2.1.1 Grid outage modeling

It is necessary to introduce the approach used for modeling grid outages. Firstly, $N_k$ is defined as a random set of all grid outages that occurred in year $k$, while $O_{jk}$ is defined as a random set of all discrete time intervals of grid outage $j$ in year $k$. Time, with increments of $\Delta t$, is a discrete variable in the formulation. Therefore, $t_{ijk}$ denotes the $i^{th}$ time interval for the $j^{th}$ outage in year $k$. Finally, the indicator function $\delta(t_{ijk})$ shows whether the demand at the $i^{th}$ time interval for the $j^{th}$ outage in year $k$ is satisfied or not. It takes the value of 1 when the demand is lost for the corresponding time interval, while it is 0 in the opposite case.

Furthermore, it is assumed that the duration of the $j^{th}$ grid outage in year $k$, $T_{jk}$, is a random variable, following a scale adjusted Poisson distribution in which $T_{jk} > 0$. Therefore, the total grid outage time in year $k$ is $\sum_{j \in N_k} T_{jk}$. $K$ denotes the planning time horizon in years, so the cumulative outage time during the planning horizon is $\sum_{k=1}^{K} \sum_{j \in N_k} T_{jk}$.

Based on the definition of $\delta(t_{ijk})$, the outage time with unmet load demand in year $k$ can now be defined as $\sum_{j \in N_k} \sum_{i \in O_{jk}} \delta(t_{ijk}) \, \Delta t$, while the cumulative outage time with unmet load demand can be defined as $\sum_{k=1}^{K} \sum_{j \in N_k} \sum_{i \in O_{jk}} \delta(t_{ijk}) \, \Delta t$. Similarly, the unmet load demand in year $k$ is $\sum_{j \in N_k} \sum_{i \in O_{jk}} \delta(t_{ijk}) \, D(t_{ijk}) \, \Delta t$ where $D(t_{ijk})$ is the demand, and the cumulative unmet load

can be calculated as $\sum_{k=1}^{K}\sum_{j\in N_k}\sum_{i\in O_{jk}}\delta(t_{ijk})\,D(t_{ijk})\,\Delta t$ . Meanwhile, the starting time of each

simulated grid outage is assumed to be uniformly distributed throughout the year.

In this work, CAIDI and SAIFI data are also applied to model power grid outages. In order to be consistent with real-world cases, grid outages are modeled based on CAIDI and SAIFI as follows:

$$E\left[T_{jk}\right] = \text{CAIDI}, \quad \forall\, j \in N_k,\, k \in \{1,2,3,...,K\}$$

$$E\left[\left|N_k\right|\right] = \text{SAIFI}, \quad \forall\, k \in \{1,2,3,...,K\}$$

$$\left|N_k\right| = \text{number of grid outages in year } k$$

The expected time duration of a grid outage equals CAIDI, while the expected number of grid outages per year equals SAIFI.

### 5.2.1.2 System optimization formulation

The problem formulation is presented in this section. The objective is to minimize *TSC*, including battery investment combined with the cost of unsatisfied customer demand when the PV + battery system works in the island mode during grid outages (for as long as the available battery charge allows). In the proposed model, there are constraints for energy balance, energy generation and the reliability metrics, i.e., *LOLP* and *CCP*.

The formulation of the PV + battery optimization problem, which is an extension of the model presented by Zhou et al., [153] is illustrated as follows,

$$\min \quad B_r b + \sum_{k=1}^{K} VOLL \,(1+r)^{-k} \sum_{j\in N_k}\sum_{i\in O_{jk}}\delta(t_{ijk})\,D(t_{ijk})\,\Delta t$$

$$\text{s.t. } P(t_{ijk}) = \eta\, I(t_{ijk})\, A, \quad \forall i \in O_{jk},\ j \in N_k,\ k \in \{1,2,3,...,K\} \tag{5.7}$$

$$B_{\min} = B_r \left(1 - DoD\right) \tag{5.8}$$

$$Q_B(t_{1jk}) = I_c \, B_r, \quad \forall \, j \in N_k, \ k \in \{1, 2, 3, ..., K\} \tag{5.9}$$

$$Q_B(t_{(l+1)jk}) = \begin{cases} \min\left\{Q_B(t_{ljk}) + P(t_{ljk}) \, e, \ B_r\right\}, \ \text{if} \ Q_B(t_{ljk}) + \int_{t_{ljk}}^{t_{ljk}+\Delta t} \left(P(u) - C_p D(u)\right) e \, du < B_{\min} \\[2ex] Q_B(t_{ljk}) + \int_{t_{ljk}}^{t_{ljk}+\Delta t} \left(P(u) - C_p D(u)\right) e \, du, \\[2ex] \qquad\qquad\qquad\qquad \text{if} \ B_{\min} \le Q_B(t_{ljk}) + \int_{t_{ljk}}^{t_{ljk}+\Delta t} \left(P(u) - C_p D(u)\right) e \, du \le B_r \\[2ex] B_r, \qquad\qquad\qquad \text{if} \ Q_B(t_{ljk}) + \int_{t_{ljk}}^{t_{ljk}+\Delta t} \left(P(u) - C_p D(u)\right) e \, du > B_r \end{cases}$$

$$\forall \, l \in O'_{jk}, \ j \in N_k, k \in \{1, 2, 3, ..., K\} \ \text{where} \ O'_{jk} = O_{jk} - \left\{\max O_{jk}\right\} \tag{5.10}$$

$$\delta(t_{ijk}) = \begin{cases} 1, \ \text{if} \ Q_B(t_{ljk}) + \int_{t_{ljk}}^{t_{ljk}+\Delta t} \left(P(u) - C_p D(u)\right) e \, du < B_{\min} \\[2ex] 0, \ \text{otherwise} \end{cases} \tag{5.11}$$

$$\forall \, i \in O_{jk}, \ j \in N_k, \ k \in \{1, 2, 3, ..., K\}$$

$$LOLP_{jk} = \frac{\sum\limits_{i \in O_{jk}} \delta(t_{ijk}) \, \Delta t}{T_{jk}}$$

$$T_{jk} = \left|O_{jk}\right| \Delta t \qquad\qquad \forall \, j \in N_k, \ k \in \{1, 2, 3, ..., K\} \tag{5.12}$$

$$\frac{\sum\limits_{k=1}^{K} \sum\limits_{j \in N_k} LOLP_{jk}}{\sum\limits_{k=1}^{K} \left|N_k\right|} \le \alpha$$

$$\Pr\left\{LOLP_{jk} \le \beta\right\} = \frac{\sum\limits_{k=1}^{K} \sum\limits_{j \in N_k} \lambda_{jk}}{\sum\limits_{k=1}^{K} \left|N_k\right|} \ge \gamma, \quad \text{where} \quad \lambda_{jk} = \begin{cases} 1, \ \text{if} \ LOLP_{jk} \le \beta \\ 0, \ \text{otherwise} \end{cases} \tag{5.13}$$

$$\forall \, j \in N_k, \ k \in \{1, 2, 3, ..., K\}, \ B_{\min} \ge 0, \quad B_r \ge 0, \quad 0 \le \alpha, \ \beta, \ \gamma \le 1$$

The objective function consists of two different terms. The first term, $B_r b$, refers to

the battery investment cost. The second term, $\sum_{k=1}^{K} VOLL\,(1+r)^{-k} \sum_{j \in N_k} \sum_{i \in O_{jk}} \delta(t_{ijk})\,D(t_{ijk})\,\Delta t$,

refers to the present value of the cost of lost load demand of customers during grid

outages. These two terms of the objective function compose *TSC*, which needs to be

minimized. The objective function is stochastic due to the random sets $N_k$ and $O_{jk}$

involved in the second term. As mentioned earlier, $N_k$ is the random set of grid outages

for year $k$ and $O_{jk}$ is the random set of time intervals of outage $j$ for year $k$. It should be

noted that this stochastic problem is transformed to be solvable by minimizing the

expected value of the stochastic cost objective. More specifically, the transformation was

done by calculating the estimated expected values instead while using the proposed

simulation-based optimization method and numerous iterations. Concerning the grid

outage modeling adopted in this work, CAIDI and SAIFI are used as the expected values

of the duration of a grid outage and the number of outages per year, respectively. Thus,

based on the grid outage modeling method described in section 5.2.1.1, it is possible to

calculate the outage time with unmet load, the amount unmet load itself and the

associated penalty costs of unmet load.

The first constraint determines the energy generated by the PV array at time $t_{ijk}$. The

second constraint denotes the calculation of $B_{\min}$ based on *DoD*. The third constraint

defines the amount of energy stored in the battery at the beginning of each outage, where

$I_c$ is used to decide the initial charging state of the battery. It depends on several realistic

factors, for example, self-discharge of a battery (%/month), which is an important feature

of a battery. For the case of a lead-acid battery, it is equal to 2-5%/month, while for

lithium-ion battery, self-discharge is 1%/month [165]. The fourth constraint focuses on

the energy balance of PV + battery systems during grid outages. It describes the change of energy stored in the battery over a short time period $\Delta t$ due to battery charging from PV array generation or battery discharging to satisfy the load demand of customers. When a PV + battery system islands during an outage, non-critical load demand of facilities may be shed, so the PV + battery system only needs to meet the proportion of critical load demand, which is denoted by $C_p$ [166]. The fifth constraint presents the condition under which the indicator function $\delta(t_{ijk})$ equals 1 or 0, which indicates whether the load demand at time $t_{ijk}$ is lost or not. The sixth constraint indicates the reliability of the islanding generation of the PV + battery system, with $\alpha$ as the upper bound for the expected value of *LOLP*. *LOLP*$_{jk}$ is defined as the proportion of time when load demand of customers cannot be met by PV + battery system during the $j^{th}$ grid outage for year $k$. The definition of *CCP* is given in the seventh constraint, which incorporates the *LOLP* constraint.

The simulation-based optimization approach that proposed in section 5.1.2 is used here to solve the problem. Based on the proposed approach, optimal battery capacity combined with a predetermined PV array can be obtained. The simulations are conducted under different combinations of battery price and *VOLL* in order to investigate the interplay between them. Different battery capacities are evaluated from both economic and reliability aspects according to the simulated grid outages.

To obtain the expected values of stochastic functions used in the analyses, it is necessary to introduce some new notation. $S$ denotes the total number of simulated grid outages. $G_s$ is the random simulated set of time intervals for the $s$th simulated outage and $t_{is}$ is the $i$th time interval of the $s$th simulated outage. $C_s$ and *LOLP*$_s$ are the cost of lost

load and loss of load probability corresponding to the *s*th simulated outage. Note that in the simulation model $\Delta t = 1$ hour, but it could be any other incremental value. The estimators for the expected values of the cost of lost load, *LOLP* and *CCP*, which is defined as the probability that the *LOLP* constraint is satisfied, regarding the simulated grid outages are presented as follows:

$$\hat{E}[C] = VOLL \frac{\sum_{s=1}^{S} C_s}{S}, \quad \text{where } C_s = \sum_{i \in G_s} \delta(t_{is}) D(t_{is}) \tag{5.14}$$

$$\hat{E}[LOLP] = \frac{\sum_{s=1}^{S} LOLP_s}{S}, \quad \text{where } LOLP_s = \frac{\sum_{i \in G_s} \delta(t_{is})}{|G_s|} \tag{5.15}$$

$$\hat{E}[CCP] = \hat{\Pr}\{LOLP_{jk} \le \beta\} = \frac{\sum_{s=1}^{S} \lambda_s}{S}, \quad \text{where } \lambda_s = \begin{cases} 1, & \text{if } LOLP_s \le \beta \\ 0, & \text{otherwise} \end{cases} \tag{5.16}$$

Other important assumptions are made as follows, the efficiency of PV array conversion in the simulation, $\eta$, is assumed to be 16% [167]. The efficiency of battery charging/discharging process, $e$, is assumed to be 85% and *DoD* is 70%. As to the reliability constraints of this problem, $\alpha$ is set to be 0.13, $\beta$ is 0.2 and $\gamma$ is assumed to be 0.7. Because of the intended use of battery storage as an energy backup system for facility electrification, it should be classified as an energy-oriented and long-duration battery storage system [161]. For all the above reasons, discharging the battery is not allowed during normal grid operation, which apparently leads to $I_c = 1$, i.e., the battery is considered fully charged at the starting time of each outage. Because of the criticality of a hospital as a test case, $C_p$ is set to be 80%, which determines the proportion of critical load demand of hospital that needs to be met during grid outages.

**5.2.2   Case study**

A hospital is an excellent and insightful test case to serve as the facility to implement the proposed mathematical model. During the design of energy backup systems, the criticality of buildings is one of the most significant independent variables because it dictates the strictness of the system modeler. Therefore, the high-reliability standards required for the operation of a hospital and the high penalty costs for unmet load demand make this specific facility a good test case. In this case study, the adopted CAIDI and SAIFI are 8.23 hours/interruption and 1.17 interruptions/year for the Long Island, NY area [164], where the actual historical solar irradiation data and the hospital load profile are selected [149, 150]. NREL provides data for a typical meteorological year (*TMY*), which are exactly the data used in this case study for the entire 20-year planning horizon.

**5.2.2.1 Main assumptions**

Battery price and *VOLL* are the most significant parameters of the proposed model. Moreover, the anticipated decline in battery price could even be accelerated in the near future, for example, the second life of electrical vehicle batteries [168]. To consider price fluctuations and variations, it is assumed that battery price is varying within the range of 20 \$/kWh to 220 \$/kWh in the simulation. In addition, *VOLL* is varying within the selected 90% confidence level (*CL*) range of 5 \$/kWh to 25 \$/kWh, which is based on the expectation of *VOLL* level in 2030 [161].

For the purposes of research, a 2-way factorial experimental design is used and analyzed, with battery price and *VOLL* being the two factors. In total, 11 distinct *VOLL* values and 11 distinct battery prices are considered, which means that an 11x11 factorial design with a total of 121 treatments is applied. The measurements are the battery

capacities required for the optimal *TSC*, the actual optimal *TSC*, the expected *LOLP,* and the *CCP*, which are obtained from 1,000 independent simulation runs for each treatment. Convergence analysis has been done to determine that 1,000 is the proper simulation number for each combination of *VOLL* and battery price to obtain the representative aggregated values. For instance, the standard deviation for optimal *TSC* is $853.29 and for the expected *LOLP* is 0.0063 in the constrained scenario when battery price is 60 $/kWh and *VOLL* is 17 $/kWh. The photovoltaic array size considered in the analysis is 6,000 m$^2$.

The problem is analyzed under two different scenarios, which are defined as follows:

1) The *unconstrained* case, in which the optimal *TSC* corresponds to the minimum *TSC* of the PV + battery system, without imposing any reliability criteria or constraints.

2) The *constrained* case, in which the optimal *TSC* is the minimum *TSC* of the PV + battery system which also satisfies all the reliability criteria of the problem.

The distinction between the two scenarios is noticeable. Because the battery sizes contributing to the unconstrained optimal *TSC* are different from that contributing to the constrained optimal *TSC* unless the former ones are also able to satisfy the reliability constraints.

Figures 5.14 to 5.21 are 3D figures (presented in subfigures (a)) and contour figures (presented in subfigures (b)) of the experiment results regarding the two factors, i.e., battery price and *VOLL*, being indicated in the horizontal axis and vertical axis, respectively. In the third axis, the presented variables are the battery size, *TSC*, the expected *LOLP,* and the *CCP*, respectively.

### 5.2.2.2 Results and discussion

First, the unconstrained case with optimal *TSC* is conducted. Then the optimal *TSC* in the constrained case is presented, as they are defined before.

**1) Unconstrained scenario**

The results shown in Figures 5.14 to 5.17 correspond to the unconstrained case where the optimal *TSC* is achieved. Battery capacities that contribute to the unconstrained optimal *TSC* under different combinations of battery price and *VOLL* are shown in Figure 5.14 a) (3D figure), and Figure 5.14 b) (2D figure with contours). In Figures 5.15 to 5.17, the horizontal and the vertical axes are the same with Figure 5.14, but the third axes are the *TSC*, the expected *LOLP* and the *CCP*, respectively. Figures 5.14 to 5.17 are presented in pairwise comparisons, which is easy to be interpreted individually.



Figure 5.14 a). Battery capacity for the unconstrained scenario as a function of battery price and *VOLL*

Figure 5.14 b). Contours corresponding to Figure 5.14 a)



Figure 5.15 a). Optimal *TSC* for the unconstrained scenario as a function of battery price

and *VOLL*

Figure 5.15 b). Contours corresponding to Figure 5.15 a)
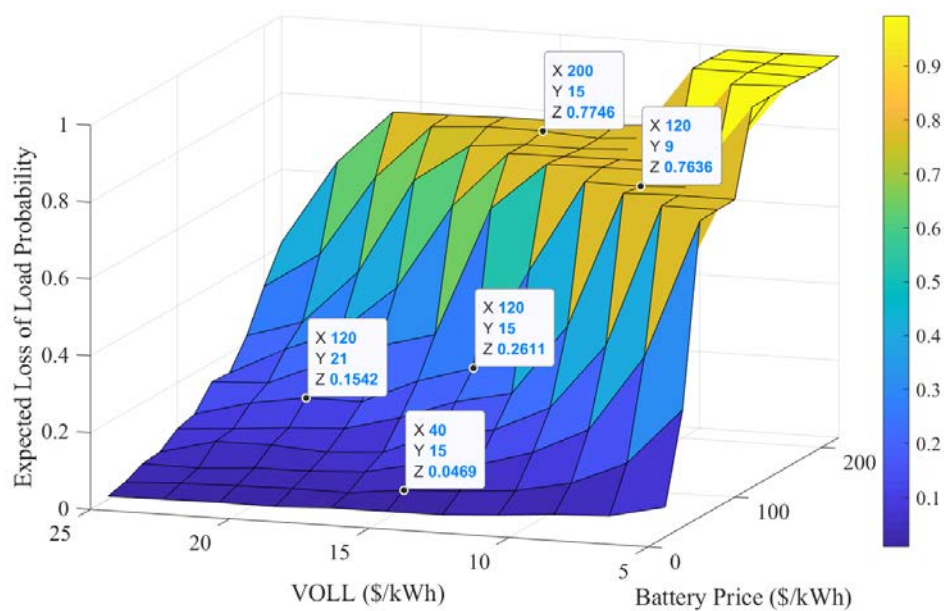


Figure 5.16 a). Expected *LOLP* for the unconstrained scenario as a function of battery
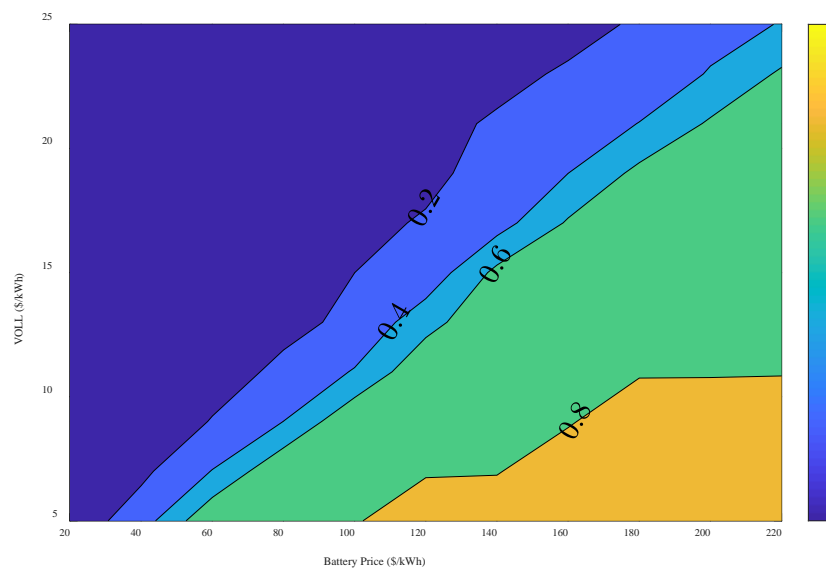
price and *VOLL*

Figure 5.16 b). Contours corresponding to Figure 5.16 a)
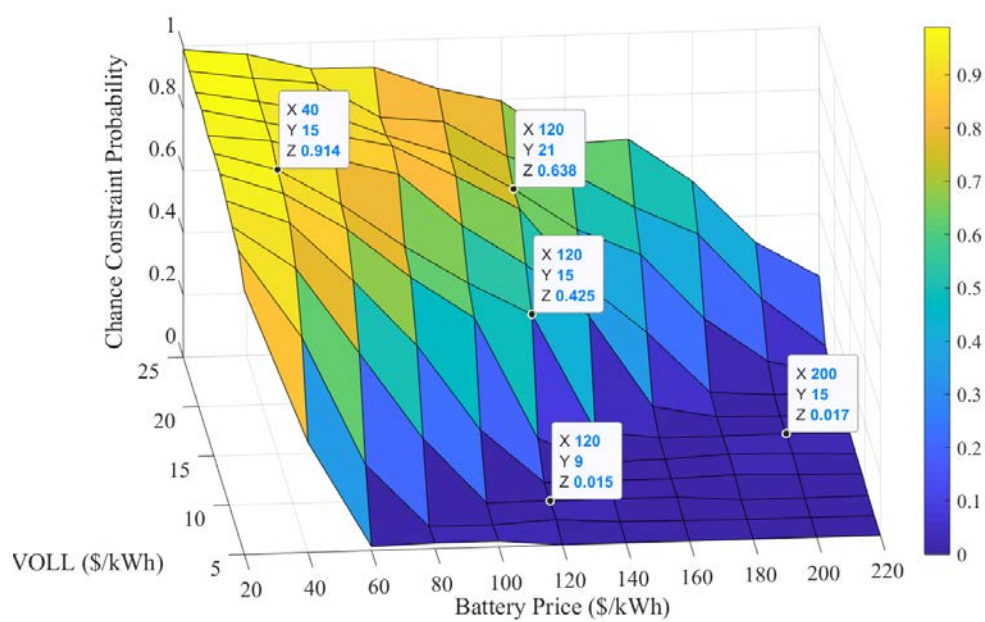


Figure 5.17 a). *CCP* for the unconstrained scenario as a function of battery price and
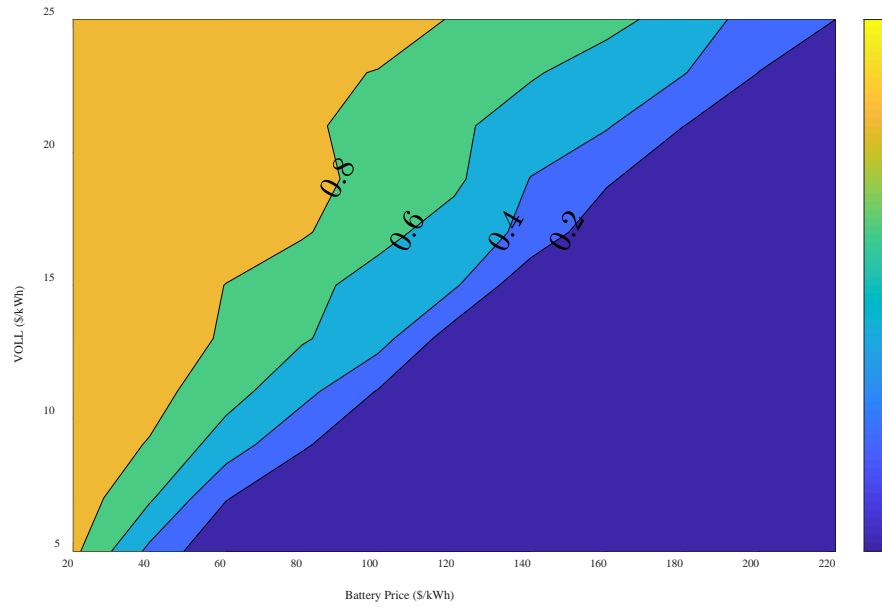
*VOLL*

Figure 5.17 b). Contours corresponding to Figure 5.17 a)

An important observation from Figure 5.14 a) is that the battery capacities, which are required for the optimal *TSC*, increase when battery price decreases or when *VOLL* increases. As can be observed, the required optimal *TSC* increases when battery price decreases from 200 $/kWh to 40 $/kWh with the same *VOLL* of 15 $/kWh or when *VOLL* increases from 9 $/kWh to 21 $/kWh with the same battery price of 120 $/kWh. These results are in agreement with previous findings [153] that lower battery price leads to a larger required battery to achieve the optimal *TSC*. When Figure 1a) is analyzed together with Figure 5.14 b), a big plateau is noticeable in the region of low values of *VOLL* and high values of battery price. This plateau clearly indicates that if it is more affordable to allow demand lost during grid outages than to invest batteries, only small batteries are required for optimal *TSC* in the unconstrained scenario.

In Figure 5.15 a) it can also be observed that the optimal *TSC* increases when the battery price and/or *VOLL* increases, which is actually in accordance with what would be

expected. However, an important feature of Figure 5.15 b) is that in the upper-left part, the change rate of the optimal *TSC* with respect to battery price is much greater than that of the optimal *TSC* with respect to *VOLL*. Noticing the lower-right of Figure 5.15 b), the change rate of the optimal *TSC* with respect to battery price becomes smaller than that of the optimal *TSC* with respect to *VOLL*. This phenomenon can be explained by observing Figures 5.14 a) and 5.14 b). When it is economical to invest in batteries, the optimal battery capacities are close to the upper limit required by the facility load profile and outage statistics and are independent of the *VOLL*. It leads to the similar investment cost and outage costs, which is not the case when battery investment costs become higher.

According to Figure 5.16 a), the expected *LOLP* corresponding to the optimal *TSC* decreases only when battery price decreases or *VOLL* increases. On the contrary, it can be observed from Figure 5.17 a) that the *CCP* shows the opposite behavior. Nevertheless, the most interesting feature that should be highlighted is the two plateaus that exist in Figures 5.16 a) and 5.16 b). The first one, which occupies the high *VOLL* and low battery price area, indicates that it is cost-effective to invest in bigger batteries to achieve a very low expected *LOLP*. The other plateau, found in the low *VOLL* and high battery price area, shows the opposite results. If the penalty cost of lost demand is low while the batteries are expensive, it is preferable to endure lost demand during outages instead of investing more in battery storage. These two plateaus can also be observed in Figures 5.17 a) and 5.17 b) but in different and opposite regions. The *CCP* is very high when it is economical to invest in battery storage and simultaneously expensive to afford the lost demand during outages, while the *CCP* becomes very low when the opposite conditions are considered.

## 2) Constrained scenario

The results shown in Figures 5.18 to 5.21 correspond to the case where the optimal *TSC* is achieved while adhering to the predetermined reliability constraints. Battery capacities that contribute to the constrained optimal *TSC* with different combinations of battery price and *VOLL* are shown in Figure 5.18 a) (3D figure) and Figure 5.18 b) (2D figure with contours). In Figures 5.19 to 5.21, the horizontal and the vertical axes are the same as Figure 5.18, but the third axis of these figures is the constrained optimal *TSC*, the expected *LOLP,* and the *CCP*, respectively.
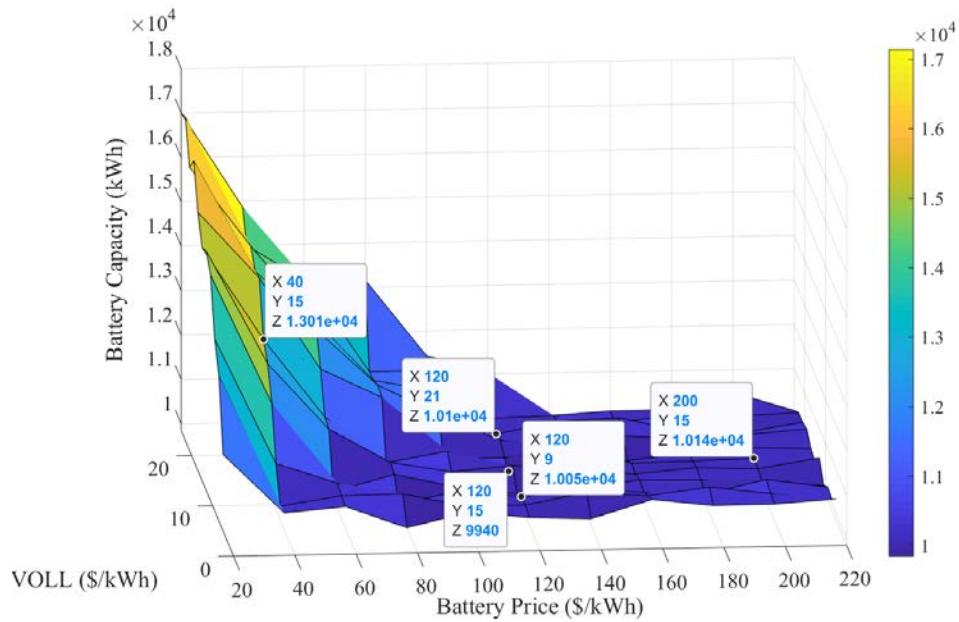


Figure 5.18 a). Battery capacity for the constrained scenario as a function of battery price
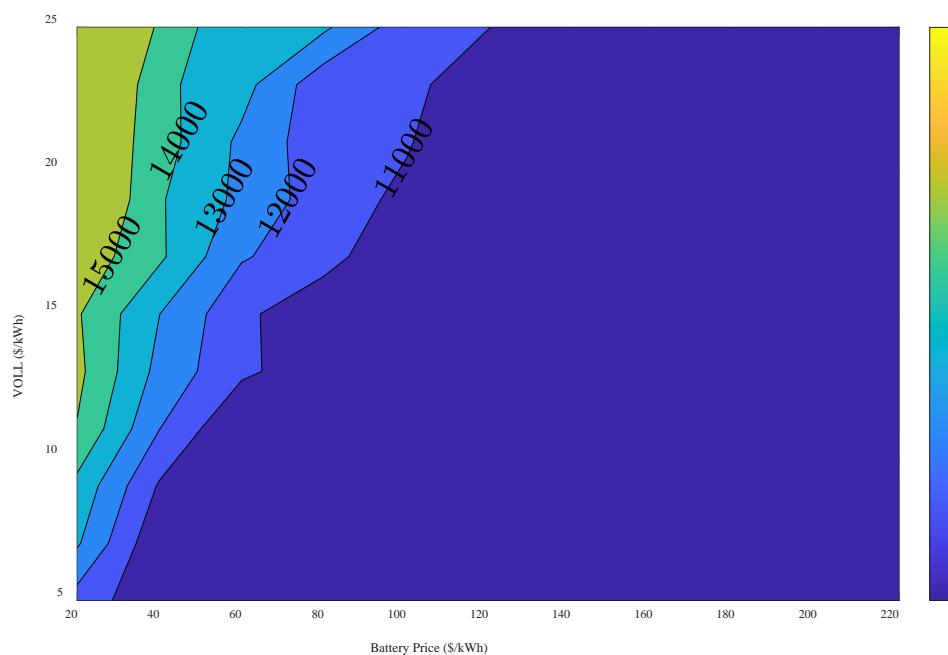
and *VOLL*

Figure 5.18 b). Contours corresponding to Figure 5.18 a)
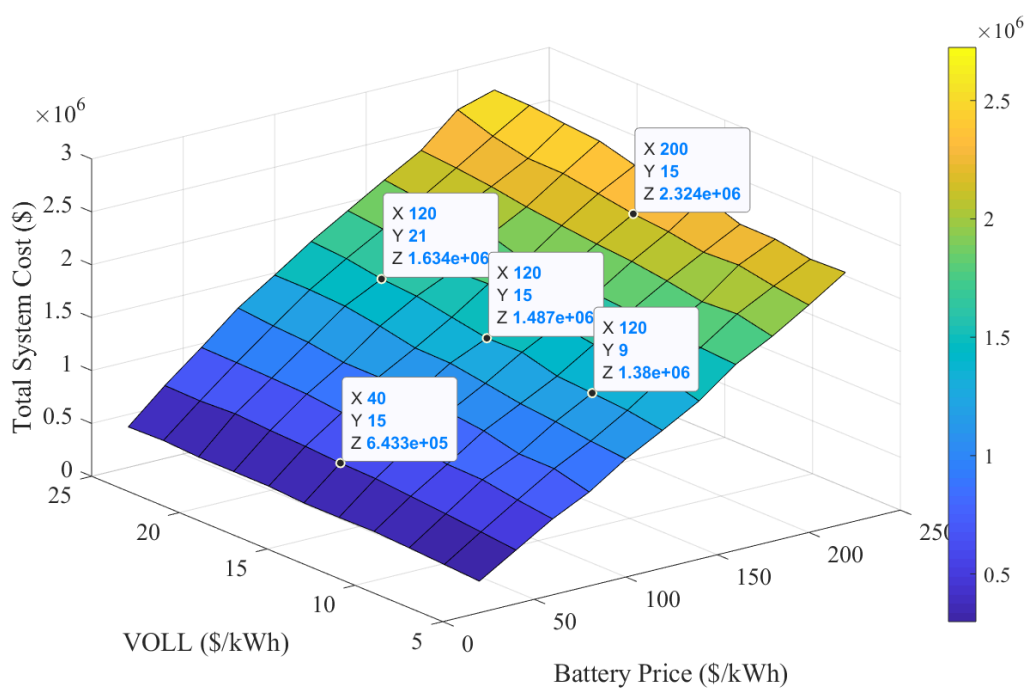


Figure 5.19 a). Optimal *TSC* for the constrained scenario as a function of battery price
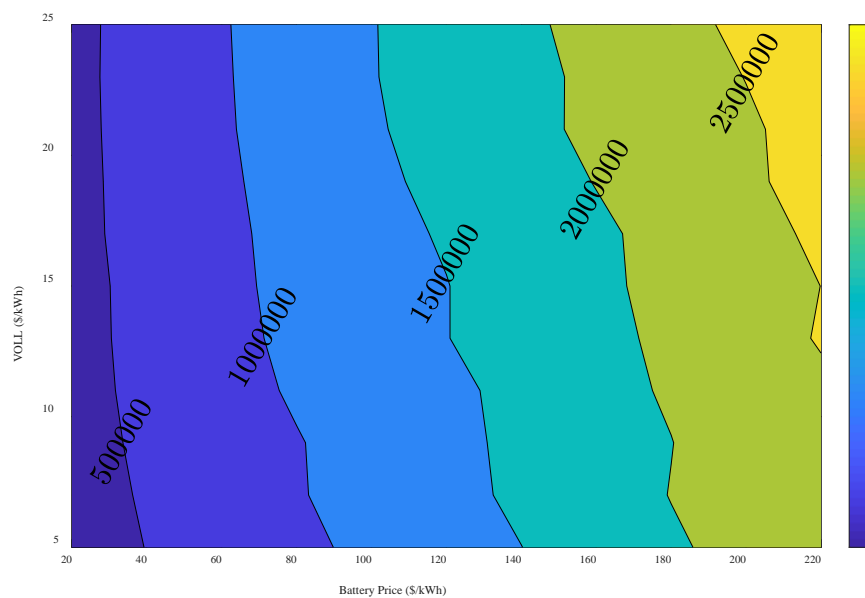
and *VOLL*

Figure 5.19 b). Contours corresponding to Figure 5.19 a)
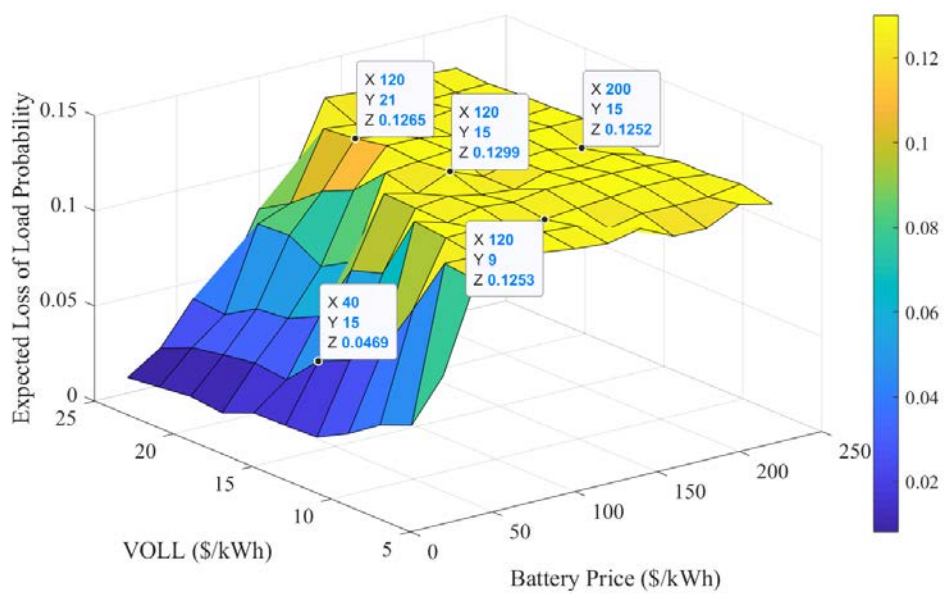


Figure 5.20 a). Expected *LOLP* for the constrained scenario as a function of battery price
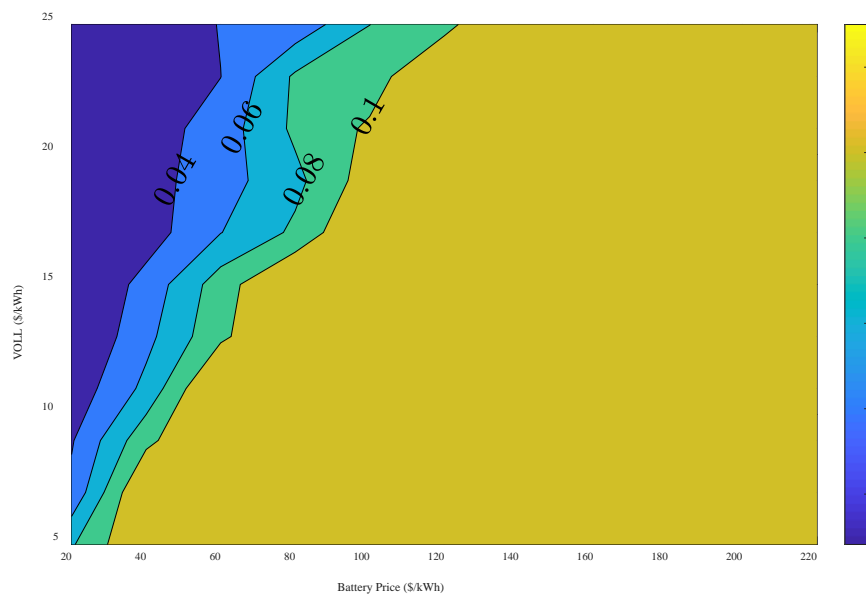
and *VOLL*

Figure 5.20 b). Contours corresponding to Figure 5.20 a)
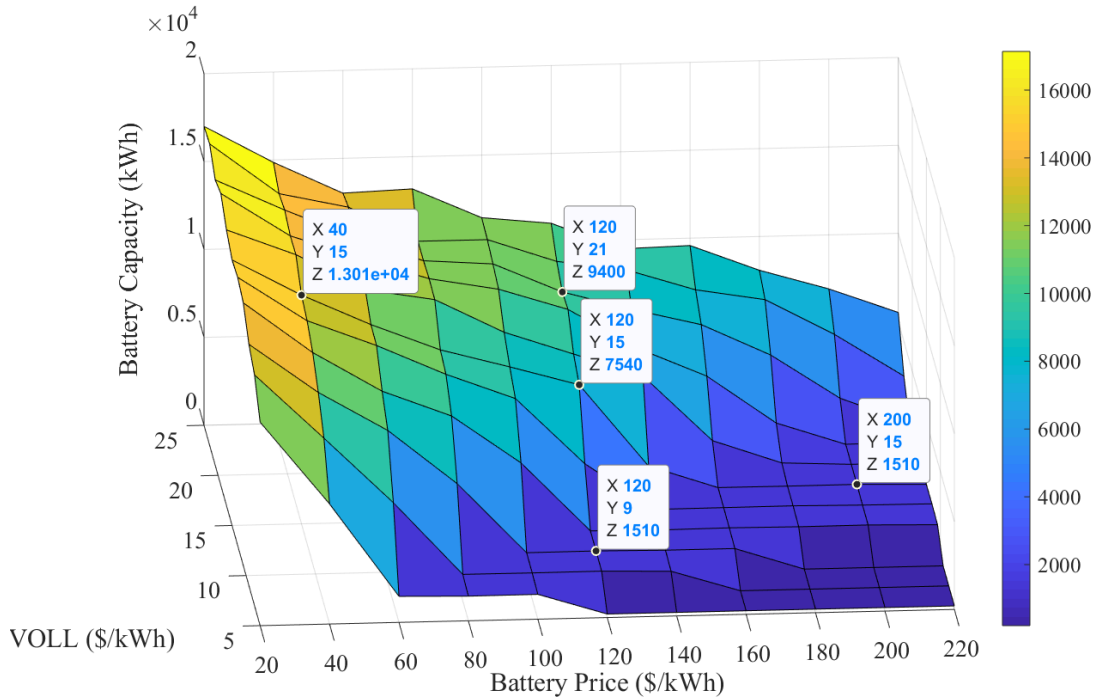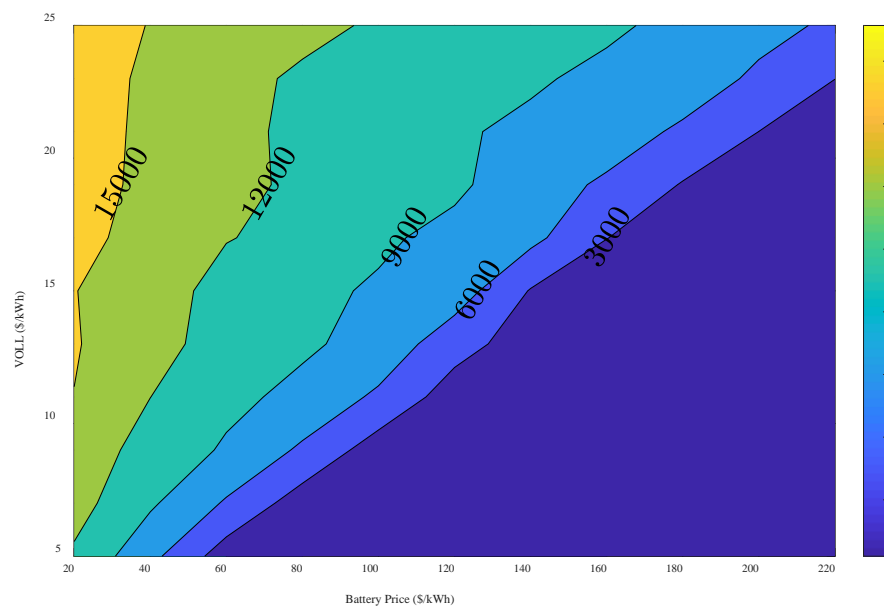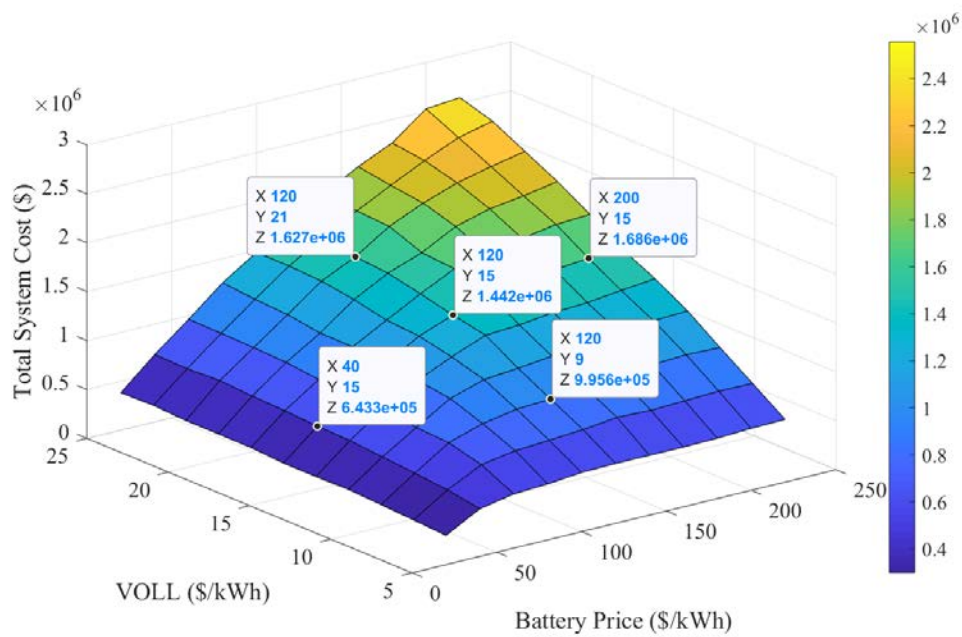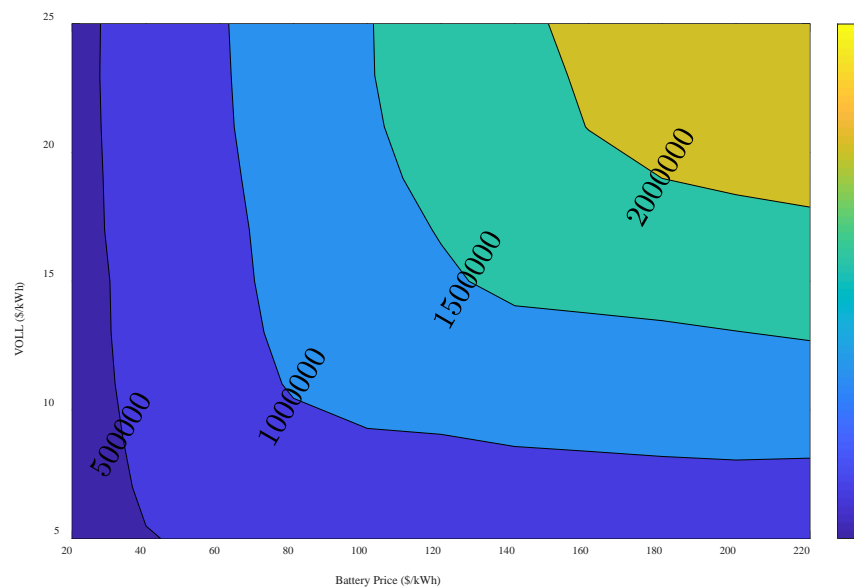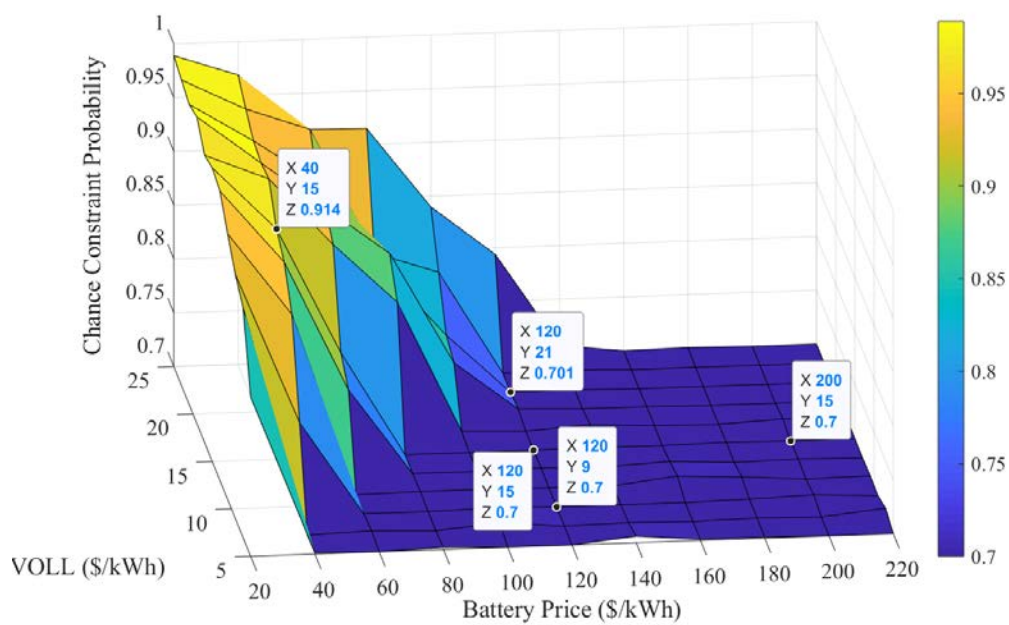


Figure 5.21 a). *CCP* for the constrained scenario as a function of battery price and *VOLL*
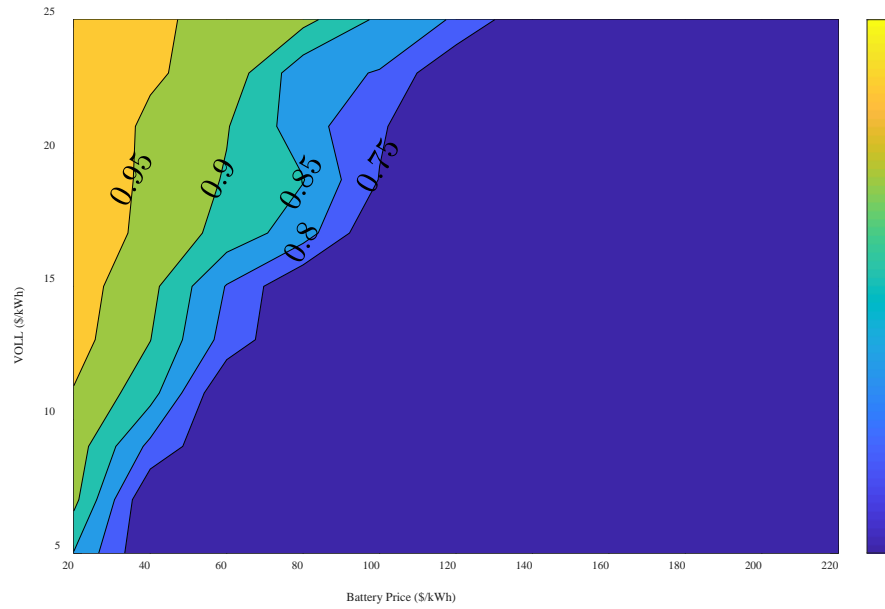
Figure 5.21 b). Contours corresponding to Figure 5.21 a)

Figure 5.18 a) indicates that for battery prices higher than 100 $/kWh, the required battery capacities for the optimal *TSC* have very small variations among them. This fact is independent of the *VOLL*. Furthermore, the battery capacity that is required to achieve the optimal *TSC* increases negligibly as *VOLL* increases, for example, when *VOLL* increases from 9 $/kWh to 21 $/kWh with the battery price of 120 $/kWh.

Figures 5.19 a) and 5.19 b) present the most consistent and smoothest trend among all figures. In Figure 5.19 b), the change rate of the optimal *TSC* with respect to battery price is greater than the change rate of the optimal *TSC* with respect to *VOLL*. However, the latter rate of change becomes greater in the high battery price area.

Finally, the big plateau appearing in Figures 5.20 a) and 5.20 b) should be compared with the plateau in Figures 5.18 a) and 5.18 b). When it is expensive to invest in a battery and the economic losses of load demand are relatively low, the required battery capacities barely satisfy the reliability constraints. When compared with Figure 5.18 a), it can be

observed that the PV + battery systems with the lowest *TSC* can also satisfy the predetermined *LOLP* constraint when battery price is low. It denotes that the decline of battery price not only decreases the lowest *TSC* but also enhances system output capability in island mode because the system planner can afford to buy a larger battery for the system. Conclusions extracted from the existing plateau in Figures 5.21 a) and 5.21 b) are consistent with the previous results observed in Figures 5.18 and 5.20. In the region where battery price is high, the *CCP* constraint is almost tight for the system with optimal battery size. Similarly, it can be observed in the plateau of Figure 5.18 a) that when it is not cost-efficient to invest in battery storage, the systems with optimal battery capacities can barely satisfy the *CCP* constraint. Lastly, for the same region, the expected *LOLP* is very close to its upper limit, as it can be seen in Figure 5.20 a).

### 3) Discussion

Besides the general conclusions described above, it would be valuable if the obtained results are interpreted from a more specific scope of the hospital. It means that the attention should be given to the high *VOLL* and the low battery price based on the forecasted decline in battery prices [169]. Furthermore, the discussion only focuses on Figures 5.18 to 5.21, which correspond to the constrained cases where a relatively strict *LOLP* constraint is satisfied. This is supported by the high priority that resilience possesses in critical facilities, like a hospital. The results obtained from the two scenarios are provided in Tables 5.4 and 5.5:

Table 5.4 Results of optimal *TSC* and corresponding batteries under the two scenarios

| ($/kWh) | Optimal *TSC* ($\times 10^5$ \$) | | | | | | Corresponding batteries ($\times 10^3$ kWh) | | | | | |
| | Unconstrained scenario | | | Constrained scenario | | | Unconstrained scenario | | | Constrained scenario | | |
| *VOLL* \ *b* | 9 | 15 | 21 | 9 | 15 | 21 | 9 | 15 | 21 | 9 | 15 | 21 |
| 120 | 9.96 | 14.42 | 16.27 | 13.8 | 14.87 | 16.34 | 1.51 | 7.54 | 9.4 | 10.05 | 9.94 | 10.1 |
| 200 | 11.12 | 16.86 | 22.11 | 22.53 | 23.24 | 24.37 | 0.20 | 1.51 | 2.99 | 10.34 | 10.14 | 10.03 |

Table 5.5 Results of system achieved expected *LOLP* and *CCP* corresponding to

optimal *TSC* in the two scenarios

| ($/kWh) | Expected *LOLP* | | | | | | System achieved *CCP* | | | | | |
| | Unconstrained scenario | | | Constrained scenario | | | Unconstrained scenario | | | Constrained scenario | | |
| *VOLL* \ *b* | 9 | 15 | 21 | 9 | 15 | 21 | 9 | 15 | 21 | 9 | 15 | 21 |
| 120 | 0.76 | 0.26 | 0.15 | 0.13 | 0.13 | 0.13 | 0.02 | 0.43 | 0.64 | 0.70 | 0.70 | 0.70 |
| 200 | 0.99 | 0.77 | 0.63 | 0.13 | 0.13 | 0.13 | <0.01 | 0.02 | 0.06 | 0.70 | 0.70 | 0.70 |

It can be seen from Tables 5.4 and 5.5 that optimal *TSC* obtained in the constrained scenario is higher than that in the unconstrained scenario. The corresponding batteries in the two scenarios show a similar difference. It indicates that in order to meet the predetermined reliability constraints, i.e., $LOLP \leq 0.13$ and $CCP > 0.7$, in the constrained scenario, larger batteries should be installed which contribute to higher *TSC*. Note that the expected *LOLP* and system achieved *CCP* in Tables 5.4 and 5.5 are the same under the considered battery price and *VOLL* in the constrained scenario. It is because when it is more affordable to endure the economic losses of lost load demand than to invest in a large battery, the battery capacities required for the optimal *TSC* are small, a fact that forces the system to barely satisfy the given reliability constraints.

There is a pressing need to install relatively large batteries for such critical applications, as can be observed in Figure 5.18 a). This trend is expected to remain the same or even to be intensified in the future when battery prices decrease, which makes

energy storage systems more cost-efficient. However, there is also great potential ahead because when battery investment costs decrease, the corresponding *TSC* can be anticipated to be reduced accordingly, as Figures 5.19 a) and 5.19 b) suggest. To further strengthen this argument, it can be concluded that the forecasted changing trend of battery price would also bring positive influence on system resilience. As Figure 5.20 a) illustrates, the system designer would be able to decrease the budget and also enhance the energy system resilience.

To summarize, there are some interesting relationships among the contour trends shown in Figure 5.18 and Figures 5.20 to 5.21, which are collectively presented in Figure 5.22.
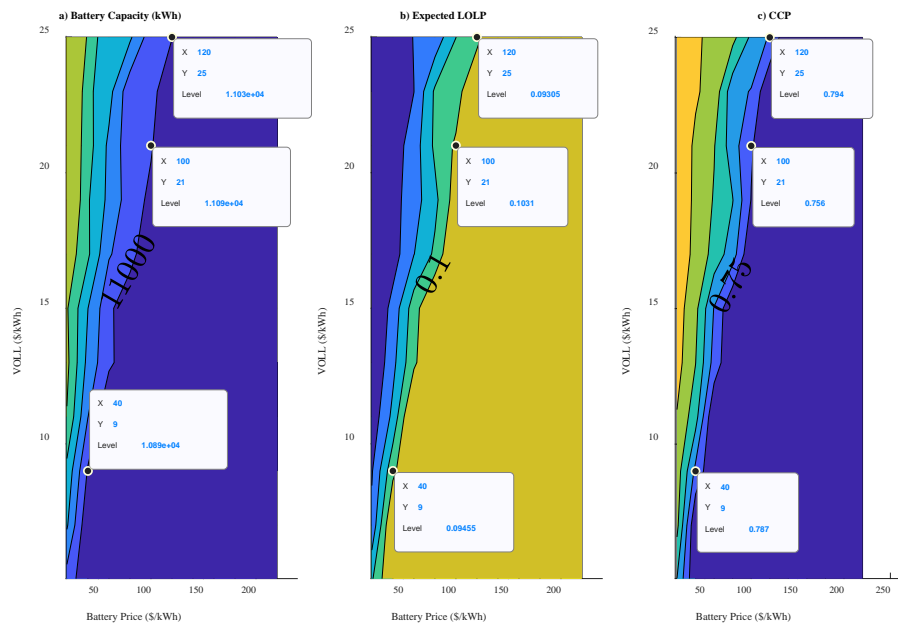


Figure 5.22. Comparison among battery capacity, expected *LOLP* and *CCP* for the

constrained case

There is a plateau that can be clearly seen in the lower-right part of Figures 5.22 a), b) and c). Based on Table 5.6 and the data labels shown in Figure 5.22, there are three

critical contours in Figure 5.22 a): the required battery is close to 11,000 kWh; b): the expected *LOLP* is close to 0.1; and c): the *CCP* is close to 0.75.

Table 5.6 Selected points in three critical contours with the same battery price and *VOLL* in Figure 5.22

| Evaluation Metrics | | Figure 5.22 a) (kWh) | Figure 5.22 b) | Figure 5.22c) |
|---|---|---|---|---|
| | [40,9] | 10890 | 0.095 | 0.787 |
| [Battery price, *VOLL*] | [100,21] | 11090 | 0.103 | 0.756 |
| | [120,25] | 11030 | 0.093 | 0.794 |

The impacts of *VOLL* and battery price on the required batteries, the expected *LOLP,* and the *CCP* become insignificant after exceeding the critical contours. The explanation for this phenomenon is that there are specific combinations of *VOLL* and battery price, upon which the battery capacities that contribute to the optimal *TSC* make both the expected *LOLP* and the *CCP* constraints not binding. It indicates that even lower *TSC* can be achieved with the desired system islanding output reliability.

**4) Analysis of variance (ANOVA)**

In this section, a linear regression model is fit using *VOLL* and battery price as the independent variables, and the *TSC* for the unconstrained case as the response variable. The purpose is to prove the significance of these two parameters and their interaction in determining the unconstrained optimal *TSC*. In Table 5.7, a summary of an ANOVA Table for estimated coefficients of the linear model is presented.

Table 5.7 Estimated coefficients of the linear regression model

| | *Estimate* | *Standard Error* | *tStatistic* | *p-Value* |
|---|---|---|---|---|
| *Intercept* | 5.76e+5 | 68,072 | 8.45 | <0.01 |
| *VOLL* | -9,207.7 | 4,181.6 | -2.2 | 0.02963 |
| *Battery price* | -1,239.5 | 501.83 | -2.47 | 0.01495 |
| *VOLL* ✕ *Battery price* | 509.6 | 30.827 | 16.531 | <0.01 |

The ANOVA Table shows some interesting and conclusive results. The $p$-value for both linear terms and the interaction term, i.e., *VOLL*, battery price, and *VOLL* $\times$ Battery price is lower than a significance level of 0.05. The findings regarding the main effects of *VOLL* and battery price are expected by considering the previous results because both factors are apparently of high significance. The analysis of variance provides strong evidence that both terms play a crucial role in the final output. Moreover, the *R*-squared is 0.946, which clearly indicates that the linear model is indeed a good model choice.

The interaction between these two factors requires a more thorough investigation. There is interplay between the impact of *VOLL* and battery price on system optimal configuration, and the ANOVA table demonstrates this in a most convincing way. Figures 5.23 and 5.24, which refer to the unconstrained scenario, are shown as follows to illustrate the inter-relationship. It can be observed that the optimal *TSC* is higher with a higher *VOLL*. Besides, the slope of a curve that is obtained with a specific *VOLL* increases as *VOLL* increases, i.e., the optimal *TSC* increases faster as battery prices increase with a larger *VOLL*. It indicates that the effect that the optimal *TSC* decreases as battery price decreases becomes stronger for a higher *VOLL*. We can also see that the optimal *TSC* stabilizes after battery price exceeds a threshold, as shown in Figure 5.23.
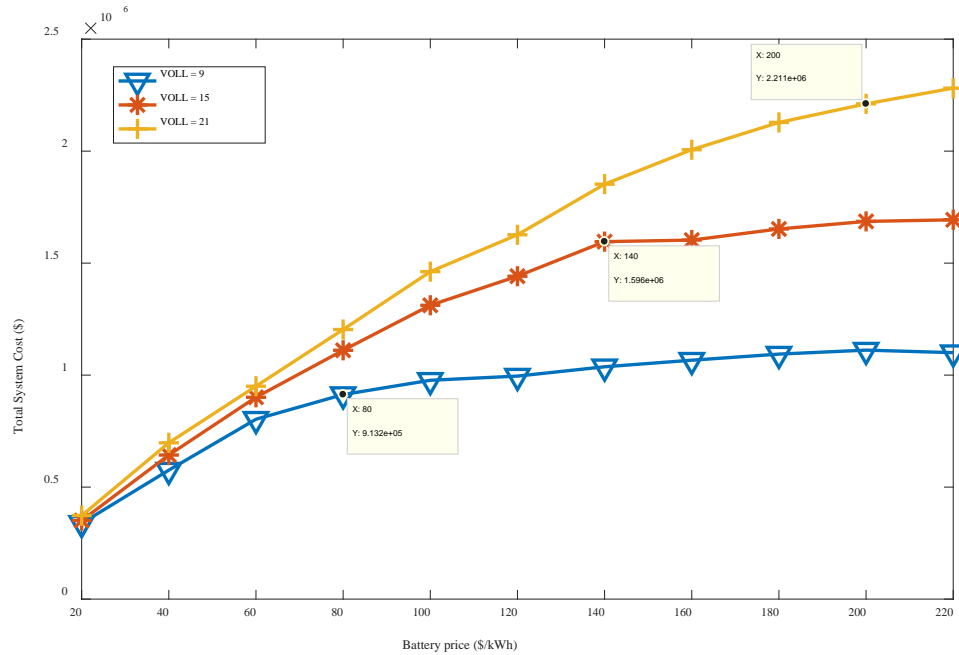
Figure 5.23. Optimal *TSC* for the unconstrained case as a function of battery price
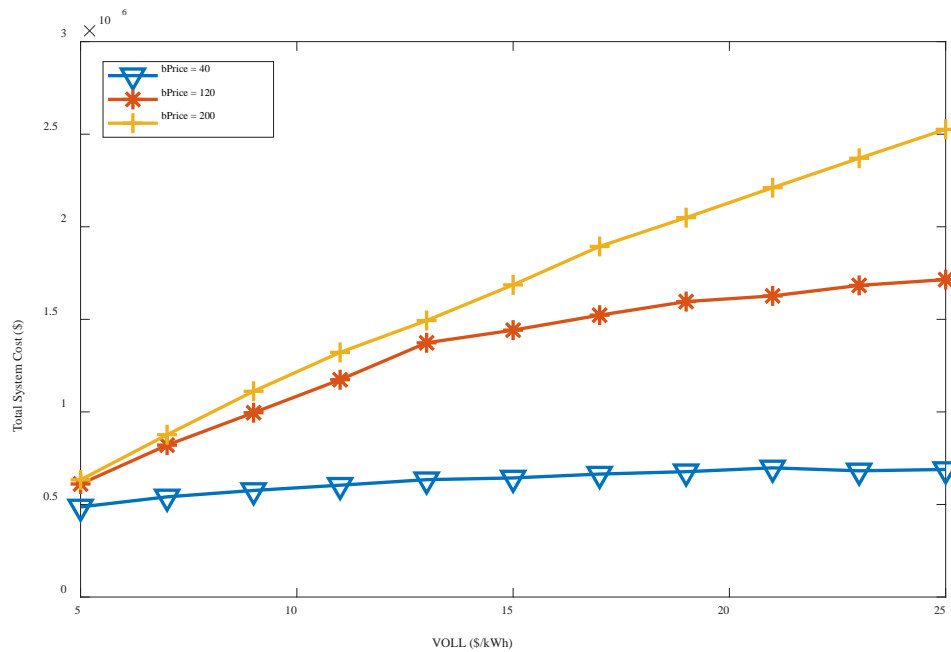


Figure 5.24. Optimal *TSC* for the unconstrained case as a function of *VOLL*

It is apparent that the optimal *TSC* is higher with a higher battery price. Similarly to

Figure 5.23, the slope of a curve that is obtained with a specific battery price in Figure 11

increases as battery price increases, i.e., the optimal *TSC* increases faster as *VOLL* increases with a larger battery price. It denotes that the effect that the optimal *TSC* increases as *VOLL* increases becomes stronger for a higher battery price. After *VOLL* exceeds a threshold, the increasing trend of *VOLL* does not notably impact the *TSC*.

The noticeable difference of changing trends of *TSC* as battery capacity changes caused by the change of *VOLL* can be observed from Figures 5.25 a) and c). It shows that investing in battery storage can actually reduce *TSC* when *VOLL* is high. While it becomes less economically viable to install batteries as battery price increases.
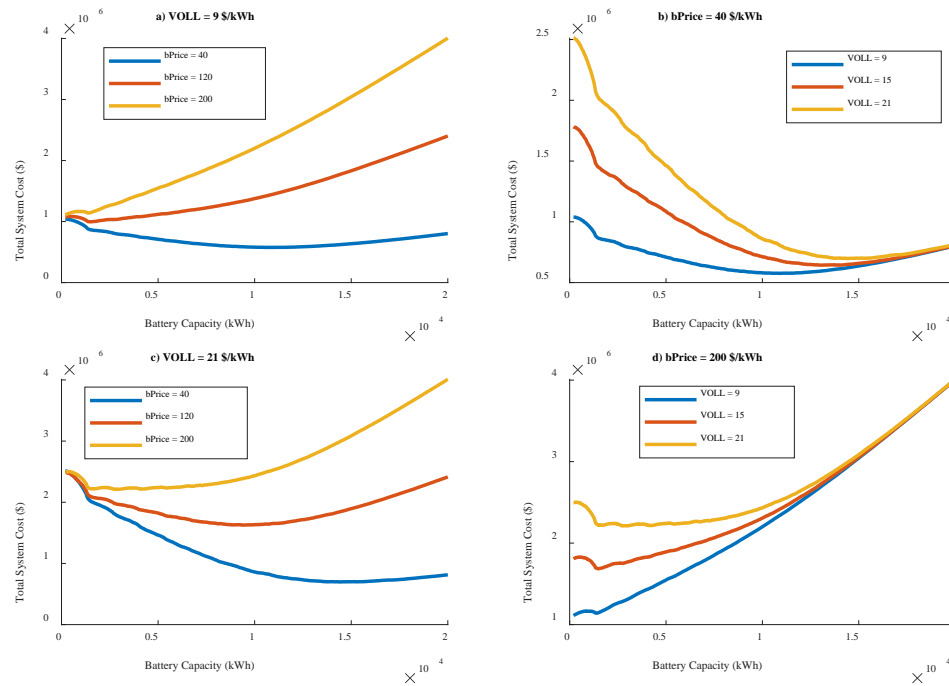


Figure 5.25. *TSC* as a function of battery capacity with different *VOLL* and battery price

It can be seen from Figure 5.25 that *TSC* for the cases with the battery capacity of 1,000 – 2,000 kWh actually increases when *VOLL* is small and battery price is high. It is attributed to the large difference between load demand and PV energy generation. The addition of a small battery does not help to noticeably reduce the unsatisfied load demand

as well as the penalty cost of unsatisfied load, while the investment of battery contributes to a higher total system cost. After battery capacity exceeds the threshold values around 1,000 to 2,000 kWh, the proportion of outage time with satisfied demand increases as battery capacity increases, and the obtained significant reduction of unsatisfied load demand drives total system cost down sharply.

Similarly, there is a distinct difference between the changing trends of *TSC* observed in Figs. 5.25 b) and d), which is attributed to the change in battery price. *VOLL* influences if it is cost-effective to invest in battery storage, especially when battery price is high. It can be concluded that high *VOLL* drives *TSC* up no matter how large battery is invested.

It is clearly shown in the analyses that the future decline in battery price will crucially affect applications where the associated value of lost load is high. Consequently, it is necessary to combine the latter result with the interaction that exists between value of lost load and battery price for future optimization of PV + battery system. Although in general it is observed that higher value of lost load results in higher total system cost, it is also showed that this relationship becomes weaker as the battery price decreases. This finding brings great potential for all kinds of applications especially for applications whose energy backup systems need to be highly reliable and efficient.

Section 4 presents the research work about optimal resilience-based restoration selection against cascading failures. Section 5 investigates the impacts of system hardening by investment of backup energy system against grid outages. These two parts of work are separately considered when performing optimization. Therefore, it leads to a further research direction: integrated optimization of system hardening and restoration for resilience enhancement against cascading failures. This research focus on the decision

making considering the total cost about system investment, restoration cost and economic loss caused by system disturbances simultaneously. This work is presented in section 6.

# 6 Joint optimization of system hardening and restoration

In this section, an optimization framework is proposed against cascading failures in network systems by jointly optimizing system hardening and system restoration strategies regarding resilience enhancement and economic benefits. Traditionally, system hardening or design and restoration actions are employed sequentially. In this work, an integrated optimization approach is developed to solve the problem with the objective function that minimizes total cost, including hardening cost, restoration cost and economic loss caused by cascading failures.

In this work, system hardening planning refers to the actions that are conducted before the occurrence of system disruptions, such as natural disasters or man-made hazards [170]. For example, the deployment of hardening or prevention resources, such as fire and chemical protection facilities, can provide hazard control in the early stages of fire and leakage of dangerous goods at transportation stations. Access of system component characteristics, such as operation and service information, is helpful for hardening resources allocation, which can mitigate system vulnerability and reduce the potential loss by quickly returning system to normal operations after a failure. System hardening planning is related to diverse factors, such as component hardening priority, component recovery time and system information availability. The foundation of reasonable allocation of hardening resources is to identify important system components.

System restoration actions, involving many factors, such as repairing prioritization and repair crew planning, which impact restoration efficacies, are usually performed after

the occurrence of system disruptions. Restoration actions, as post disruption responses, aim to reduce the impact of system disruptions and restore the system back to the desired performance.

Most of current research focuses on either hardening planning optimization or restoration measure optimization [108, 153, 171-173]. For example, system hardening decisions in redundancy allocation problem are made independently from restoration measures. However, there are interplays between the effects of system hardening measures and that of restoration measures, especially considering the uncertainty associated with system disruptions. Hardening measures are designed to prevent system failure from happening or mitigate the losses caused by potential risks. Proper deployment of hardening resources, such as security resources and hot standby equipment, can reduce the failure of alternative equipment or shorten its recovery time, which provides better external conditions for restoration implementation. While the effect of restoration actions on reducing the influence of system disruptions is impacted by distinct pre-disruption hardening planning. Therefore, the interaction between system hardening and restoration should be considered in order to integrate them for effectively reducing system damage caused by system disruptions with uncertainty.

In this section, the proposed optimization framework is applied to the modelled electricity system which is subjected to the mixed cascading failures proposed in section 3.2. The influence of cascading failures in power generation and transmission system on local customers through population centers is explored. The impacts of multiple system characteristics, such as system dependency and system load level, on the effectiveness of system hardening and restoration actions are also investigated. Except for resilience loss

metric $\Re(t)$, other measurements regarding system reliability and system performance, such as *CCP* and system connectivity *G*, are applied to evaluate the efficacy of integrated hardening and restoration. Different system hardening strategies and restoration strategies are taken into consideration against failure propagation. The application of the integrated optimization approach is tested and validated on the network systems subjected to different disruptions in numerical examples.

## 6.1 Combined optimization of hardening and restoration incorporating resilience benefits

In this section, system hardening actions combined with restoration measures are optimized to mitigate the impacts of cascading failures triggered by different system disruptions. In real cases, the budget for system hardening investment and restoration measures is limited. For hardening planning, investment cost includes cost of adding redundant components, cost for obtaining information of system characteristics and failure events, and cost of DER installation for energy system, etc. As to restoration actions, the related cost includes cost of maintaining repair crews. System damage and other adverse impacts of cascading failures usually lead to economic loss. *CCP* constraint with respect to resilience loss $\Re(t)$ is incorporated into the objective function. In this way, the combined optimization of system hardening and restoration can meet the desired resilience requirement with economic viability.

The electricity system model introduced in section 3.2 is applied as the targeted network system in this work. The resilience-based problem is to determine the optimal integrated system hardening measures and restoration measures with minimum total cost, which includes cost of restoration exertion, cost of information acquisition for hardening

and cost of damage caused by cascading failures. System disturbances with uncertainty due to unplanned variation or changing operating conditions emerge during real-world system operation. The problem is subject to *CCP* requirements regarding resilience loss constraints for scenarios with different disturbances. The disturbances can be represented by a random vector $\mathbf{U}$, where $\mathbf{U} = (U_1, U_2, ..., U_k)$ include $k$ different disturbance factors, such as disturbance type and severity, with uncertainties. Different scenarios in terms of disturbance would impact system hardening and restoration optimization results.

In the following case studies, initial system disruption severity is the only disturbance factor used to separate scenarios. $\mathbf{u} = (u_1, u_2, ..., u_l)$, where $u_l$ is a determined level of initial system disturbance severity in scenario *l*. To be more specific, initial system disruption severity is evaluated by the number of initially failed system components. It should be noted that the initially failed system components may result from random internal failures or external disturbances in real cases. More initial system component failures means a more severe disturbance which tends to trigger more dramatic system failure propagation.

### 6.1.1 Problem formulation

The problem is to jointly optimize system hardening and restoration measures with economic viability, i.e., minimizing investment cost and economic loss. Resilience constraints are predetermined for different scenarios. Firstly, the components of the objective function in the optimization problem are introduced as follows,

### 1) Hardening cost

Access of system component information, for instance system component work-load, is helpful to deploy prevention resources which can reduce the potential of component failure and shorten the recovery time of the component when it fails. As to the electricity system model, there is investment cost of collecting the information of customers connected to the system component in the power transmission system. It determines the level of customer information available to the system component. As mentioned above, with the availability of customer information, the corresponding system component can be hardened up to a certain level. In this work, system component hardening is able to reduce the component initial failure probability and shorten its recovery time. Therefore, more customer information accessed, which requires more information investment cost, leads to better system component hardening. The hardening cost for scenario $i$ is defined as follows,

$$C_i^{in}\left(\Phi\left(\varphi,\mathbf{\rho}(\omega)\right),u_i\right) = \sum_{d \in O_{\varphi\omega i}} \rho_d \sum_{k=1}^{K_c} n_{dk} C_k^{ino} \tag{6.1}$$

$\Phi\left(\varphi,\mathbf{\rho}(\omega)\right)$ is the set of available hardening measures for decision makers. Decision variable $\varphi$ denotes a particular hardening strategy, which ranks system components based on their importance with a prioritization preference. $\mathbf{\rho}(\omega)$ denotes the vector that contains the specific hardening level for each considered component, i.e., the level of customer information available to the corresponding component. In this work, the hardening level accociated with hardening information cost decides the reduction of initial component failure probability and the reducion of component repair time. $\omega$ denotes the percentage of highly ranked system components by hardening strategy that are considered for hardening. $\rho_d$ denotes the hardening level of node $d$, $0 \le \rho_d \le 1$. $O_{\varphi\omega i}$ is the set of highly

ranked nodes that are considered for hardening in scenario $i$ based on the hardening strategy $\varphi$ and the hardening scope $\omega$. $C_k^{ino}$ is the cost of accessing the information of type $k$ customer, which is a fixed value. $n_{dk}$ is the number of type $k$ customers connected to node $d$. Total $K_c$ types of customers exist in the network system. The hardening cost for node $d$ is proportional to the number of connected customers and factored by the hardening level $\rho_d$. $\Phi(\varphi, \mathbf{\rho}(\omega))$ denotes the decision variables for sytem hardening.

### 2) Restoration cost

In this work, the cost of hiring restoration crews to repair failed nodes is named restoration cost. It depends on many factors, such as restoration strategy which ranks the failed nodes for repairing prioritization based on their importance and the number of available crews that are hired. Total restoration cost for scenario $i$ is defined in Eq. (6.2),

$$C_i^r\left(\Phi(\varphi, \mathbf{\rho}(\omega)), \Pi(\pi, c), u_i\right) = \sum_{f=1}^{c} C_f^r l_{if} \tag{6.2}$$

$$C_f^r = \begin{cases} C_b^r, & f = [1, c_r] \\ C_b^r + C_v^r(f - c_r), & c \ge f \ge c_r \end{cases} \tag{6.3}$$

$\Pi(\pi, c)$ is the set of available restoration measures for decision makers. $\pi$ represents a particular restoration strategy which ranks the failed nodes for repairing based on node importance. $c$ denotes the number of hired crews. It is assumed that $c_r$ crews are regularly hired under normal circumstances. The cost for each regular crew to do repairing is $C_b^r$ per unit time. More restoration crews can be hired with higher cost for repairing per unit time and less working efficiency, which is defined by Eq. (6.3). $C_v^r$ denotes the increment of hiring cost per unit time as the number of hired crews increases. $l_{if}$ denotes the total repairing time performed by the $f$th crew in scenario $i$. It should be

noted that $l_{if}$ is influenced by system disturbance, hardening measures and restoration measures, which also impact restoration cost.

### 3) Economic loss

As to the electricity system, when cascading failures occur in power transmission system, customers might be disconnected from the power grids because of tripped system components, such as transformers and transmission lines. It leads to economic loss since the power load of customers is no longer satisfied. The economic loss result from disconnected customers due to failed system nodes is presented as follows,

$$C_i^{los}\left(\Phi\left(\varphi,\mathbf{\rho}(\omega)\right),\Pi(\pi,c),u_i\right)=\sum_{j\in UF_i}\sum_{q=1}^{Q_{ij}}C_{ijq}^{los}(g_{ijq}) \tag{6.4}$$

$$C_{ijq}^{los}(g_{ijq})=\sum_{k=1}^{K_c}n_{ijk}C_k^f+\sum_{k=1}^{K_c}n_{ijk}C_k^v\left(2g_{ijq}-g_{ijq}^{1-\beta}\right) \tag{6.5}$$

$C_{ijq}^{los}(g_{ijq})$ denotes the economic loss due to the $q$th failure of system node $j$ with the failure time $g_{ijq}$ in scenario $i$. $UF_i$ is the set of nodes that failed in scenario $i$. $Q_{ij}$ denotes the number of times that node $j$ failed in scenario $i$. There are two types of cost associated with customers when they are disconnected from the power grids. $C_k^f$ is the fixed penalty cost when a type $k$ customer is offline/disconnected from the electric power system. $n_{ijk}$ denotes the number of type $k$ customers connected to node $j$ in scenario $i$. $C_k^v$ is the baseline variable cost per unit time associated with a type $k$ customer when it is offline. $C_k^v(2g_{ijq}-g_{ijq}^{1-\beta})$ denotes the variable cost of a type $k$ disconnected customer, which nonlinearly increases as the failure time of node $j$, $g_{ijq}$, increases. $\beta$ is the parameter that decides the increase rate of variable cost associated with disconnected customers, $0\le\beta\le1$.

### 4) Objective function including resilience constraint

The total cost for scenario $i$ with disturbance severity $u_i$ that needs to be minimized in the optimization problem is presented as follows,

$$
\begin{aligned}
C_i^{tot}\left(\Phi\left(\varphi,\boldsymbol{\rho}(\omega)\right),\Pi(\pi,c),u_i\right) &= C_i^{in}\left(\Phi\left(\varphi,\boldsymbol{\rho}(\omega)\right),u_i\right) + C_i^{r}\left(\Phi\left(\varphi,\boldsymbol{\rho}(\omega)\right),\Pi(\pi,c),u_i\right) \\
&+ C_i^{los}\left(\Phi\left(\varphi,\boldsymbol{\rho}(\omega)\right),\Pi(\pi,c),u_i\right)
\end{aligned}
\tag{6.6}
$$

Total cost includes information investment of hardening, restoration cost for hiring crews and economic loss when customers are disconnected from the system. The *CCP* constraint for scenario $i$ with disruption severity $u_i$ in terms of resilience loss is provided by Eq. (6.7),

$$
\gamma_i\left(\Phi\left(\varphi,\boldsymbol{\rho}(\omega)\right),\Pi(\pi,c),u_i\right) = v_i - \Pr\left\{\mathfrak{R}_i\left(\Phi\left(\varphi,\boldsymbol{\rho}(\omega)\right),\Pi(\pi,c),u_i\right) \le \theta_i\right\} \le 0
\tag{6.7}
$$

where $v_i$ denotes the constraint for *CCP* in scenario $i$, $0 \le v_i \le 1$. $\theta_i$ is the resilience loss constraint for scenario $i$, $0 \le \theta_i \le 1$. Resilience loss $\mathfrak{R}_i(\Phi\left(\varphi,\boldsymbol{\rho}(\omega)\right),\Pi(\pi,c),u_i)$ in scenario $i$ is obtained with a targeted system performance according to Eq. (4.3). System resilience loss during the cascading-restoration process depends on the selected hardening measures, restoration measures and system disruption. In the following case studies, the number of customers functionally connected to the electricity system is adopted as the system performance to evaluate system resilience loss. It reflects system service level with respect to customer accessibility.

Different from the work presented in sections 4 and 5, where resilience loss $\mathfrak{R}(t)$ and *CCP* constraint are applied separately regarding resilience enhancement, *CCP* constraint in terms of resilience loss $\mathfrak{R}(t)$ is incorporated directly into the objective function of the optimization problem in this work.

In order to incorporate resilience constraint into the objective function, Karush-Kuhn-Tucker (KKT) approach is applied, which converts the problem to be an unconstrained one. The new variables $\lambda$ named KKT multipliers are introduced in the problem which forms a new objective function, the generalized Lagrangian. The generalized Lagrangian over all considered scenarios are defined as follows,

$$L\left(\varphi,\boldsymbol{\rho}(\omega),\pi,c,\mathbf{u},\boldsymbol{\lambda}\right) = \sum_{i=1}^{I}\left(-C_i^{tot}\left(\Phi(\varphi,\boldsymbol{\rho}(\omega)),\Pi(\pi,c),u_i\right) - \lambda_i\gamma_i\left(\Phi(\varphi,\boldsymbol{\rho}(\omega)),\Pi(\pi,c),u_i\right)\right)$$

(6.8)

where $\lambda_i$ is given for scenario $i$. $I$ is the number of considered scenarios with different disturbance severities. The feasible solution of the new problem $\max\limits_{\varphi,\boldsymbol{\rho}(\omega),\pi,c}\min\limits_{\lambda,\lambda\geq\mathbf{0}}L\left(\varphi,\boldsymbol{\rho}(\omega),\pi,c,\mathbf{u},\boldsymbol{\lambda}\right)$ is also the optimal solution of the original optimization problem $\max\limits_{\varphi,\boldsymbol{\rho}(\omega),\pi,c}\sum\limits_{i=1}^{I}-C_i^{tot}\left(\Phi(\varphi,\boldsymbol{\rho}(\omega)),\Pi(\pi,c),u_i\right)$ which is subject to *CCP* constraints defined in Eq. (6.7). It can be proved with proper KKT multipliers $\lambda$ as follows,

$$\min\limits_{\lambda,\lambda\geq\mathbf{0}}L\left(\varphi,\boldsymbol{\rho}(\omega),\pi,c,\mathbf{u},\boldsymbol{\lambda}\right) = \begin{cases} -\infty, & \text{if } \varepsilon \notin FS \\ -\sum\limits_{i=1}^{I}C_i^{tot}\left(\Phi\left(\varphi,\boldsymbol{\rho}(\omega)\right),\Pi(\pi,c),u_i\right), & \text{if } \varepsilon \in FS \end{cases}$$

(6.9)

where $FS = \left\{\varepsilon|\gamma_i\left(\Phi\left(\varphi,\boldsymbol{\rho}(\omega)\right),\Pi(\pi,c),u_i\right)\leq 0, i\in\{1,2,...,I\}\right\}$. Resilience loss becomes cost-concerned when it is incorporated into the objective function since it impacts the optimized total cost. Note that KKT multipliers $\lambda_i$ denote the penalty weight when *CCP* constraint regarding resilience loss for scenario $i$ is violated. In some real-world cases, $\lambda_i$ represents the penalty with monetary cost when the resilience requirement is not satisfied.

### 6.1.2   Simulation-based optimization

A simulation-based optimization method is used to solve the problem. The major simulation procedures and main assumptions are presented in this section. The simulation describes the cascading-restoration process with pre-disruption system hardening. The optimal combination of system hardening and restoration considering the impacts of different system disruptions can be determined from both economic and resilience aspects. Note that different system disruptions are simulated by conducting scenarios with distinct disruption severities, which is used to mimic real-world system disturbances with uncertainty.

To obtain the expected values of stochastic functions used in the optimization problem, multiple simulation results are obtained to approximate the expected values. $I$ is the number of scenarios with different disruption severities $\mathbf{u}$ are considered in the simulations. $MS$ denotes the number of simulations conducted for each scenario with a specific disruption severity. The estimators for the expected values of resilience loss and $CCP$ regarding the simulated cascading failures are presented as follows. Based on Eq. (6.7), $CCP$ is an estimated probability, that the resilience loss for the simulated cascading failures in scenario $i$ is not more than $\theta_i$, should be larger than $v_i$.

$$E[CCP_i] = \Pr\left\{ \Re_i\left(\Phi\left(\varphi,\boldsymbol{\rho}(\omega)\right), \Pi(\pi,c), u_i\right) \le \theta_i \right\} = \frac{\sum_{m=1}^{MS} \delta_{im}}{MS}, \quad \delta_{im} = \begin{cases} 1, & \text{if } \Re_{im}\left(T_{im}(1-\tau)\right) \le \theta_i \\ 0, & \text{otherwise} \end{cases}$$

(6.10)

$$\Re_{im}(T_{im}(1-\tau)) = \begin{cases} 0, & T_{im}(1-\tau) \le t_s \\ \dfrac{\int_{t_s}^{T_{im}} \left(Q_{im}(t_s) - Q_{im}(t)\right) dt}{Q_{im}(t_s)(T_{im} - t_s)}, & T_{im}(1-\tau) > t_s \end{cases}$$

(6.11)

As mentioned above, $\theta_i$ is the resilience loss constraint for scenario $i$ with disruption severity $u_i$, $0 \leq \theta_i \leq 1$. $\delta_{im}$ is the indicator function and equals 1 if the resilience loss of the $m$th simulation in scenario $i$ satisfies the constraint, otherwise it is 0. $Q_{im}(t)$ is the system performance, which is adopted to evaluate resilience loss, at time $t$. $t_s$ is the time when system disruption occurs. $T_{im}(1 - \tau)$ is the time period from the beginning of system interruption till a specific performance, is recovered to be a predetermined level, $1 - \tau$ percentage of its initial value. $0 \leq \tau \leq 1$. It denotes the stop criteria. There is a hard constraint, $H$, on the recovery time, i.e., $T_{im}(1 - \tau) \leq H$.

Based on Eqs. (6.1) - (6.5), the estimators for the expected values of total cost, including hardening cost, restoration cost and economic loss of customers, resulting from the simulated cascading failures are shown as follows,

$$E[C_i^{tot}\left(\Phi\left(\varphi,\boldsymbol{\rho}(\omega)\right),\Pi(\pi,c),u_i\right)] = \frac{\sum_{m=1}^{MS} C_{im}^{tot}\left(\Phi\left(\varphi,\boldsymbol{\rho}(\omega)\right),\Pi(\pi,c),u_i\right)}{MS} \tag{6.12}$$

$$C_{im}^{tot}\left(\Phi\left(\varphi,\boldsymbol{\rho}(\omega)\right),\Pi(\pi,c),u_i\right) = C_{im}^{in}\left(\Phi\left(\varphi,\boldsymbol{\rho}(\omega)\right),u_i\right) + C_{im}^{r}\left(\Phi\left(\varphi,\boldsymbol{\rho}(\omega)\right),\Pi(\pi,c),u_i\right) \\ + C_{im}^{los}\left(\Phi\left(\varphi,\boldsymbol{\rho}(\omega)\right),\Pi(\pi,c),u_i\right) \tag{6.13}$$

$$C_{im}^{in}\left(\Phi\left(\varphi,\boldsymbol{\rho}(\omega)\right),u_i\right) = \sum_{d \in O_{\varphi\omega im}} \rho_d \sum_{k=1}^{K} n_{imdk} C_k^{ino} \leq B_d \tag{6.14}$$

$$C_{im}^{r}\left(\Phi\left(\varphi,\boldsymbol{\rho}(\omega)\right),\Pi(\pi,c),u_i\right) = \sum_{f=1}^{c} C_f^r l_{imf}(u_i, 1-\tau), \quad C_f^r = \begin{cases} C_b^r, & f = [1, c_r] \\ C_b^r + C_v^r(f - c_r), & c \geq f \geq c_r \end{cases} \tag{6.15}$$

$$C_{im}^{los}\left(\Phi\left(\varphi,\boldsymbol{\rho}(\omega)\right),\Pi(\pi,c),u_i\right) = \sum_{j \in UF_{im}} \sum_{q=1}^{Q_{imj}} C_{imjq}^{los}(g_{imjq}) \tag{6.16}$$

$$C_{imjq}^{los}\left(g_{imjq}\right) = \sum_{k=1}^{K} n_{imjk} C_k^f + \sum_{k=1}^{K} n_{imjk} C_k^v \left(2g_{imjq} - g_{imjq}^{1-\beta}\right) \tag{6.17}$$

$$1 \leq c, c_r \leq C_{rw}, 0 < \beta < 1, 0 \leq \rho < 1, 0 \leq \omega < 1, 0 \leq \tau < 1, 1 \leq T_{im} \leq H, I \geq 1, MS \geq 1, K_c \geq 1 \tag{6.18}$$

As introduced previously, $\Phi\left(\varphi, \boldsymbol{\rho}(\omega)\right)$ represents the set of available hardening measures, including hardening ranking strategy and hardening level. $\Pi(\pi, c)$ denotes the set of restoration actions available to be employed, including restoration ranking strategy and the number of hiring crews. These are decision variables of the problem. There is a hard constraint on the maximum number of crews that can be hired for restoration, $C_{rw}$. It indicates the capability of assigned or available restoration resources. Total repairing time performed by the $f$th crew in the $m$th simulation of scenario $i$, $l_{imf}(u_i, 1 - \tau)$, is mainly impacted by the predetermined recovery level $1 - \tau$ and disruption severity $u_i$, but it is also impacted by $\varphi$, $\boldsymbol{\rho}(\omega)$, $\pi$ and $c$. $g_{imjq}$, denoting the failure time of node $j$ in its $q$th failure during the $m$th simulation of scenario $i$, is impacted by external factors, $u_i$, $1 - \tau$, and hardening and restoration decisions, $\varphi$, $\boldsymbol{\rho}(\omega)$, $\pi$ and $c$. As presented in Eq. (6.14), the hardening investment cost is subject to a budget $B_d$.

As mentioned above, pre-disruption hardening can reduce the potential of system node failure and shorten the recovery time of failed nodes. In this work, the original probability of a system node that breaks down as initial failure to trigger cascading failures follows a certain distribution. While the actual probability of node $d$ selected as initial failure to simulate system disruption depends on its hardening level $\rho_d$ in the $m$th simulation under scenario $i$,

$$P_d^a = P_d^o \left(1 - \rho_d\right), d \in O_{\varphi\omega im} \tag{6.19}$$

where $P_d^a$ is the actual probability that node $i$ is selected as initial failure resulting from system disruptions. $P_d^o$ denotes the original probability of node $i$ to be initially failed. Higher hardening level contributes to lower potential of node failure.

Similarly, the original repair time of failed system nodes is assumed to follow a certain distribution. While the actual repair time of a failed node $i$ depends on the hardening level applied to node $i$, $\rho_i$, and the working efficiency of the $f$th crew that repairs node $i$. Actual repair time of node $i$, $T_i^{ar}$, based on original repair time, $T_i^{or}$, is obtained as follows,

$$T_i^{ar} = T_i^{or}(1-\rho_i)(1+e_f(f-c_r))$$ (6.20)

where $e_f$ denotes the decrease rate of crew working efficiency as extra crews are hired, i.e., when $c > c_r$. Less crew working efficiency indicates longer time for the crew to fully restore a failed system node. It is obvious that actual repair time of failed system nodes is related to the failure time of system nodes and the working time of crews, which influences total cost. As a result, there is a trade-off between hardening investment, restoration investment and system damage cost.

The main steps of the simulation, including pre-disruption hardening, cascading failures and post-disruption restoration response are depicted as follows,

Step 1)    a. Establish *Net*(*N*, *M*). Calculate the initial size of *LCC*.

b. Determine dependence clusters in *Net*(*N*, *M*).

Step 2)    a. Assign capacity to system node according to Eq. (3.2).

b. Apply hardening measures.

Step 3)    a. Initial system node failures occur associated with probabilities result from system disruption.

b. Failed nodes and their connected edges are not functional.

Step 4)    a. Update network structure. System components which do not belong to

current   *LCC* are not functional.

b. Failed node loads are redistributed locally and globally.

Step 5)    a. Failed nodes are repaired by crews based on restoration strategy.

b. Overload system nodes break down and their connected edges are not

functional. Go back to Step 4 until a specific system performance is

recovered to a predetermined level. Then simulation ends.

### 6.1.3   Case study

In this section, the optimization problem is solved numerically to show some insights

on the integration of system hardening and restoration to enhance system resilience

against cascading failures in different scenarios.

The electricity system model proposed in section 3.2 is used to build power grids

including centralized generation units, power transmission system and customers

connected into the power grids through population centers and distribution systems (See

Figure 3.15). Different kinds of customers associated with multiple types of monetary

cost are taken into account in the problem. As introduced in section 6.1.1, these types of

monetary cost include fixed cost once a customer is disconnected from the power grids

[$/cust.], variable cost per unit time when a customer becomes disconnected from the

power grids [$/$\Delta$t cust.], and investment cost [$/cust.] of collecting customer information

to provide reference for hardening the system nodes that provide them with electricity.

**6.1.3.1 Main assumptions and setting**

BA network model is adopted to build up power generation and transmission system. In the simulation, BA network contains 250 nodes and 498 edges. Nodes denote centralized generation units and edges denote transmission lines. Four types of customers, residential, small business, commercial and Gov./Public, are taken into consideration of which the electricity demand should be met by the power grids [29]. There are in total 1,250 customers in the network system and these four types of customers account for 65%, 20%, 10%, 5% of the entire customers, respectively. The unit simulation time $\Delta t$ is assumed to represent 1 hour. Three types of cost associated with different customers are given in Table 6.1.

Table 6.1 Three types of costs associated with four kinds of customers

|  | Residential | Small business | Commercial | Gov./Public |
|---|---|---|---|---|
| Variable cost [$/hr cust.] | 100 | 200 | 300 | 1000 |
| Fixed cost [$/cust.] | 200 | 20000 | 5000 | 50000 |
| Information cost [$/cust.] | 100 | 400 | 800 | 3000 |

The assumptions about decision variables regarding hardening and restoration are presented as follows,

A hardening strategy $\varphi$ ranks system nodes based on a specific preference, and higher ranking nodes have higher priority to be hardened. Only the top $\omega$ percentage of ranked nodes are considered for hardening. $\omega$ is a continuous variable which is between [0.1, 0.7]. The hardening level for any selected node is also a continuous variable, which is between [0.05, 0.5]. The candidate hardening strategies are:

1) High number of customer-connected nodes first hardening (NOCH): System nodes are ranked based on the number of customers that are connected to them. More connected customers indicate broader services provided by the system node.

2) High-degree nodes first hardening (HDH): System nodes are ranked according to their node degree. Higher degree indicates more importance on transferring system load. Node with higher degree has priority to be hardened.

3) High-load nodes first hardening (HLH): System nodes are ranked according to their load. Higher node load indicates more importance of the node in terms of transferring system load and satisfy customer load requirement. Node with higher load has priority to be hardened.

A restoration strategy $\pi$ also ranks system nodes based on a specific preference. Higher ranked nodes are prioritized to start repairing when they failed. The candidate restoration strategies are:

1) High degree first repair strategy (HDFR): Details of this restoration strategy are presented in section 4.2.

2) Short time first repair strategy (STFR): Details of this restoration strategy are presented in section 4.2.

3) High load first repair strategy (HLFR): This restoration strategy is presented in section 4.2.

4) First fail first repair (FFFR): This strategy means that the repairing of failed nodes is started follows the order that they failed. This is a default restoration strategy and is applied for comparison purposes.

5) High number of customer-connected nodes first repair (NOCR): Failed nodes, of which repairing have not yet started, are repaired in a descending order of the number of connected customers, i.e., network nodes with more connected customers are repaired with higher priority.

6) Fastest customer repair (FCR): The repairing of failed nodes is ordered by most customers restored per time unit, i.e., the failed node that has the largest customer-reconnected to the power grids to repair time ratio is repaired first.

It should be noted that there must be an exact ordered list of repair actions. Ties, which occur when the failed nodes of which the repairing has not started at the inspection have the same condition, e.g., the same degree, are broken according to FFFR. The repair activity of failed nodes, once started, will not stop until they are completed.

In practice, there are more candidate hardening and restoration strategies. However, the adopted ones help to illustrate the effects of joint hardening and restoration that can also be generalized for other strategies.

The availability of restoration resources depends on the number of hired crews. In this case study, the maximum number of crews that can be hired $c$ is between [10, 36]. The number of regular crews $c_r = 13$. The cost for hiring a regular crew to do repairing $c_b^r$ = 200 $/hr. The extra cost for hiring each extra crew to perform repairing $c_v^r = 100$ $/hr, and the decrease rate of repairing efficiency $e_f = 5\%$ for each extra crew. In this case study, the original node repair time is distributed uniformly in [1, 8], which is based on CAIDI data, 8.23 hr/interruption [164]. The original probability that a node is selected as initial failure is uniformly distributed in [0, 1]. Actual node repair time and initial failure probability are determined based on Eq. (6.19) and Eq. (6.20). Hardening cost constraint $B_d = \$ 90,000$.

System node local customer load and passing by load depend on its customer connections and node degree. According to Eqs. (3.7) and (3.8), system node load can be obtained. The base load follows a truncated normal distribution, where $\mu = 0.5$, $\sigma = 0.45$.

$L_{min} = 0$, $L_{max} = 1$. When a node fails, its passing by load is equally distributed to its functional neighboring nodes, i.e., the nodes that have direct connections via transmission lines. Its local customer load is equally distributed to its dependent nodes, i.e., the centralized generation units that provide electricity for the same customers (See Figure 3.15). The cascading and recovery process regarding load dynamics is presented in section 3.2.2.

The number of online customers, i.e., the customers that are connected to the electricity system and can access the electricity, is adopted as the system performance to evaluate resilience loss. All customers are connected to the electricity system before system disruption. The number of functional system nodes is used as the system performance to evaluate the recovery level. $\tau = 2\%$, denotes the stop criteria of the recovery process. The hard constraint on the recovery time $H = 100$.

The number of initially failed nodes is used to describe system disruption severity, which is the considered disruption factor in this work. More initially failed nodes indicate a more severe system disruption, which would trigger larger scale failure propagation. Since system disruptions are associated with uncertainty, the combined optimization of system hardening and restoration is performed for cascading failures caused by disruptions with different severities.

This mixed-integer optimization problem is solved using genetic algorithm in C++. The decision variables are $\varphi$, $\omega$, $\boldsymbol{\rho}(\omega)$, $\pi$, $c$. Since the number of variables in $\boldsymbol{\rho}(\omega)$ depends on $\omega$, the total number of decision variables is also a decision variable. These decision variables are a string of genes of each member of the population. The fitness value of each member is the objective function, which is calculated based on Eq. (6.8).

Tournament selection is applied to make sure that the better members in the generation likely survive.

### 6.1.3.2 Numerical examples

In the first example, the values of parameters used in GA implementation are introduced. *Pos* denotes the population size, *Pos* = 15. *Gen* is the maximum number of generations, *Gen* = 80. *Pxo* is the probability of crossover, *Pxo* = 0.8. *Pmu* = 0.25 denotes the probability of mutation. *Tor* denotes the number members randomly selected at each round of tournament selection, *Tor* = 3. GA algorithm is performed multiple times in order to check the convergence of the optimal solution.

In this example, the combined optimization of hardening and restoration is targeted for two scenarios with different system disruption severities, $u_1 = 9$, $u_2 = 51$. The first scenario has 9 system nodes breakdown as initial disruptions, and the second scenario has 51 system nodes breakdown as initial disruptions. Based on Eq. (6.19), the set of initially failed nodes are not only determined by their initial failure probabilities, but also impacted by the corresponding hardening levels. It leads to uncertainty regarding system disruptions, so simulations are conducted multiple times for each scenario. In this example, $MS = 100$, i.e., 100 simulations are performed for scenario with specific number of initially failed nodes. Resilience loss constraints $\theta_1 = 0.005$, $\theta_2 = 0.3$. *CCP* constraints $v_1 = 0.7$, $v_2 = 0.8$. KKT multiplier $\lambda_i = 10^8$. $i = 1, 2$, denotes scenario 1 and 2. Mean dependence cluster size, *D-size* = 8. The load-capacity model parameters $\alpha = 0.75$, $\mu = 0.2$. Figure 6.1 shows the fitness value of the best member in each generation in 6 GA implementations.
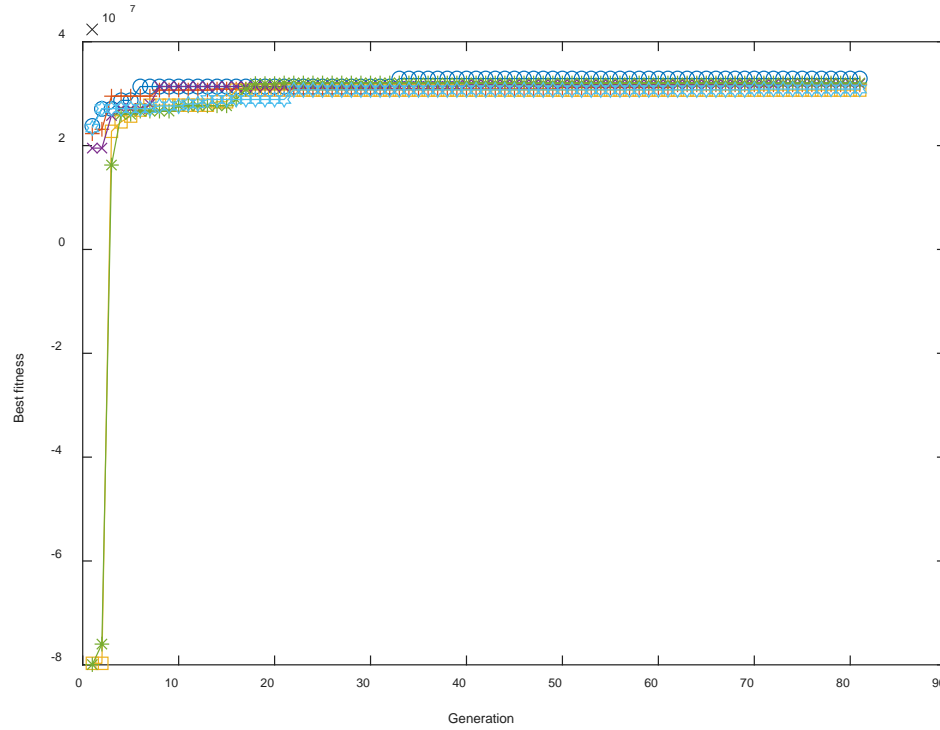
Figure 6.1 Fitness of the best member vs. generation in 6 GA implementations

It can be observed from Figure 6.1 that the fitness of the best member dramatically

increases as generation increases in the early stage. Then the best fitness turns to be stable

as generation continues to increase. It can be seen that the best fitness of different GA

implementation converges after about 30 generations. The best solutions of 6 GA

implementations are presented in Table 6.2.

Table 6.2. Best solutions of 6 GA implementations

| GA | $\varphi$ | $\pi$ | $c$ | $\omega$ | Best fitness | Total cost ($) |
|----|-----------|-------|-----|----------|--------------|----------------|
| 1 | 2 | 2 | 15 | 0.53 (132) | 32906212 | 16093787 |
| 2 | 2 | 2 | 14 | 0.647 (161) | 31712357 | 16287642 |
| 3 | 2 | 2 | 30 | 0.155 (38) | 30695065 | 16304934 |
| 4 | 3 | 2 | 20 | 0.638 (159) | 31902220 | 16097779 |
| 5 | 2 | 2 | 22 | 0.325 (81) | 32056749 | 15943250 |
| 6 | 2 | 2 | 29 | 0.406 (101) | 30888800 | 16111199 |

According to Table 6.2, HDFR restoration strategy is selected for all GA

implementations, which indicates that HDFR strategy performs better in recovery system

from cascading failures. The desired number of hired crews is bigger than the number of regular crews ($c_r = 13$). It means extra crews are worth to be fired to recover system from cascading failures. The best fitness is the objective function of the problem, and total cost includes hardening investment, restoration cost and economic loss. As shown in Table 6.2, best fitness and total cost converge from different GA implementations. The difference in the selection of hardening strategy, the number of hired crews, and the percentage of nodes to be hardened results from the uncertainty associated with the set of initially failed nodes, repair time of failed nodes, etc.

In the second example, *Pos* = 15. *Gen* = 40. *Pxo* = 0.8. *Pmu* = 0.25. *Tor* = 2. GA algorithm is performed 5 times to check the convergence of the optimal solution. The targeted two scenarios with disruption severities, $u_1 = 9$, $u_2 = 51$. 80 simulations are also performed for each scenario. For the two scenarios, resilience loss constraints $\theta_1 = 0.005$, $\theta_2 = 0.3$. *CCP* constraints $v_1 = 0.7$, $v_2 = 0.8$. KKT multiplier $\lambda_i = 10^8$. $i = 1, 2$. Mean dependence cluster size, *D-size* = 8. The load-capacity model parameter $\alpha = 0.75$. In this case, the average node capacity is reduced which indicates the increased node load level. Figure 6.2 shows the fitness value of the best member in each generation in 5 GA implementations.
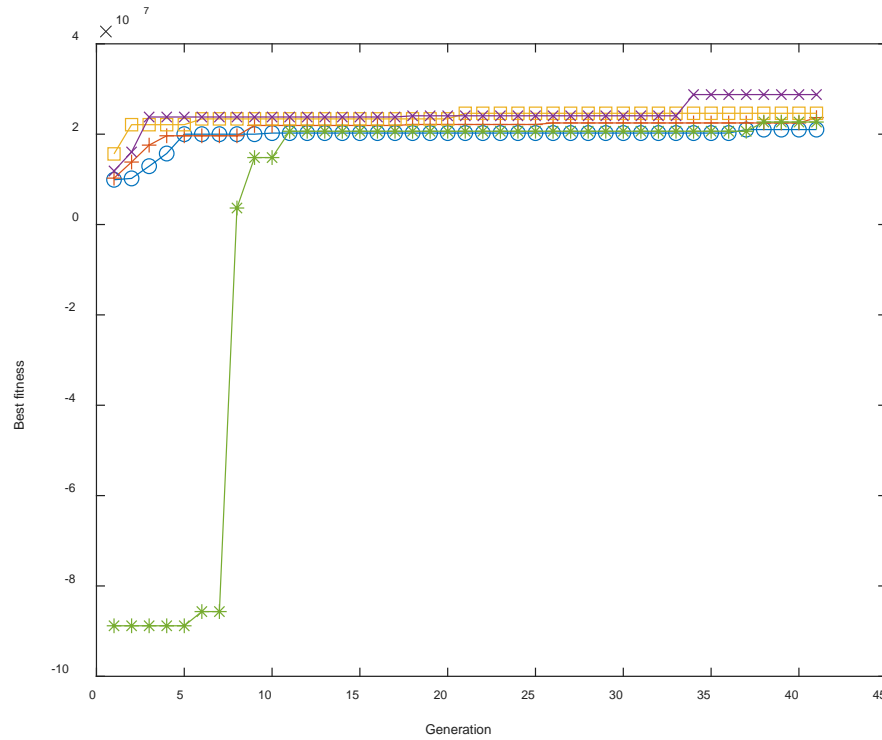
Figure 6.2 Fitness of the best member vs. generation in 5 GA implementations

Based on Figure 6.2, that fitness of the best member also dramatically increases as generation increases at the beginning. The best fitness of different GA implementations converges as generation increases. The best solutions of 5 GA implementations are presented in Table 6.3.

Table 6.3. Best solutions of 5 GA implementations

| GA | $\varphi$ | $\pi$ | $c$ | $\omega$ | Best fitness | Total cost ($) |
|----|-----------|-------|-----|----------|--------------|----------------|
| 1 | 2 | 2 | 20 | 0.227 (56) | 21,016,212 | 16,483,787 |
| 2 | 3 | 2 | 17 | 0.569 (142) | 23,653,151 | 16,346,848 |
| 3 | 3 | 2 | 22 | 0.327 (81) | 24,617,962 | 16,632,037 |
| 4 | 3 | 2 | 24 | 0.608 (152) | 28,759,033 | 16,240,966 |
| 5 | 2 | 2 | 30 | 0.241 (60) | 22,677,975 | 17,322,024 |

It can be observed from Table 6.3 that HDFR restoration strategy is selected as the best solution for all GA implementations, which is consistent with the findings from the last example. The desired hardening strategy is either HDH strategy or HLH strategy.

The uncertainty associated with the set of initially failed nodes in each scenario contributes to the difference of the desired hardening measure selection. However, the best fitness and total cost obtained from different GA implementations are very close. The hardening levels for the top $\omega$ percentage of system nodes that are ranked by hardening strategy are also obtained. Table 6.4 presents the desired hardening levels from the 1$^{st}$ node to the 56$^{th}$ nodes ($\omega = 0.227$, total 56 nodes are hardened), from no. 1 GA implementation.

Table 6.4. The best hardening levels for the top ranked nodes from no. 1 GA

implementation

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 0.184 | 0.304 | 0.264 | 0.473 | 0.078 | 0.35 | 0.321 | 0.408 |
| 0.061 | 0.491 | 0.466 | 0.078 | 0.076 | 0.147 | 0.141 | 0.125 |
| 0.119 | 0.255 | 0.278 | 0.338 | 0.211 | 0.439 | 0.175 | 0.05 |
| 0.445 | 0.125 | 0.177 | 0.051 | 0.343 | 0.128 | 0.169 | 0.301 |
| 0.311 | 0.113 | 0.479 | 0.104 | 0.336 | 0.387 | 0.276 | 0.494 |
| 0.485 | 0.368 | 0.257 | 0.259 | 0.381 | 0.296 | 0.204 | 0.095 |
| 0.125 | 0.086 | 0.361 | 0.285 | 0.398 | 0.058 | 0.476 | 0.468 |

According to Eq. (6.19) and Eq. (6.20), the actual probability of a system node breaking down as an initial failure and the actual required repair time of a failed node depends on its hardening level.

### 6.1.3.3 Impact of system dependence

According to the model proposed in section 3.2, system dependence is described by dependence clusters. In this section, the impact of mean dependence cluster size, *D-size*, on the combined optimization solution of hardening and restoration is explored.

In the first case, *D-size* = 8 and 16 are taken into account, separately. *Pos* = 15. *Gen* = 45. *Pxo* = 0.8. *Pmu* = 0.25. *Tor* = 3. GA is performed 6 times for each *D-size* to check the convergence of the optimal solution. The considered two scenarios have disruption

severities, $u_1 = 39$, $u_2 = 51$, respectively. 100 simulations are performed for each scenario. For the two scenarios, resilience loss constraints $\theta_1 = 0.25$, $\theta_2 = 0.3$. *CCP* constraints $v_i = 0.8$, KKT multiplier $\lambda_i = 10^8$. $i = 1, 2$. The load-capacity model parameter $\alpha = 0.75$. Figure 6.3 shows the best member of each generation from different GA implementations under two *D-size*. Table 6.5 presents the results of GA optimization under two *D-size* assumptions.
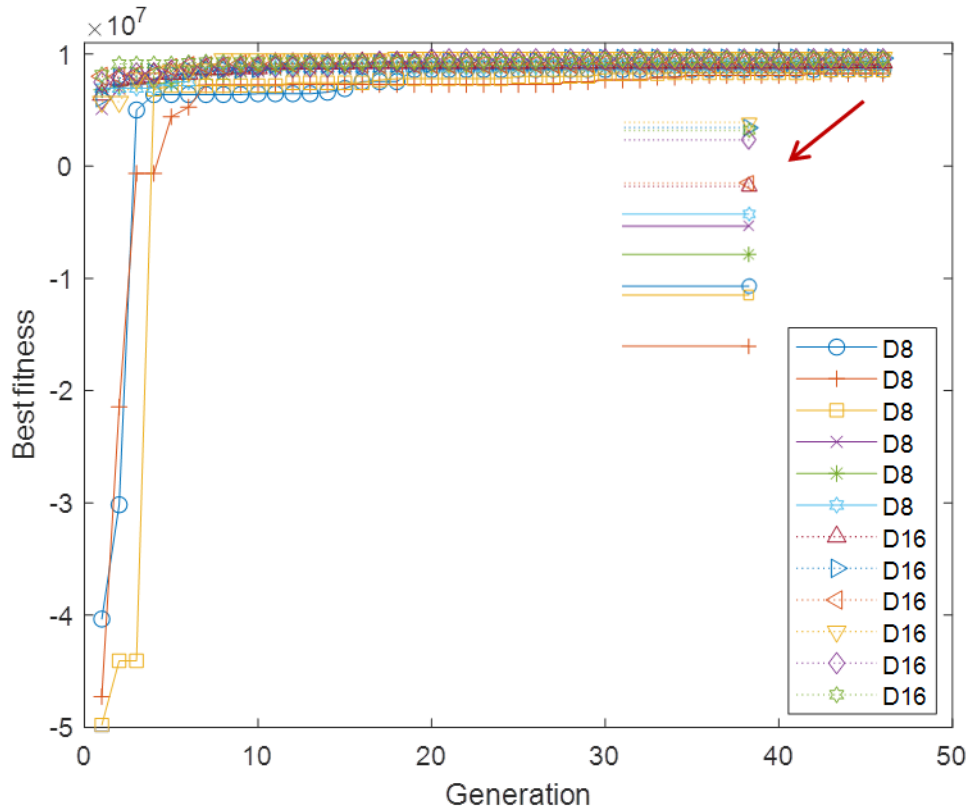


Figure 6.3 Fitness of the best member vs. generation from different GA implementations under different *D-size*

Table 6.5. Best solutions of 6 GA implementations under different *D-size*

| GA | $\varphi$ | $\pi$ | $c$ | $\omega$ | Best fitness | Total cost ($) | $\varphi$ | $\pi$ | $c$ | $\omega$ | Best fitness | Total cost ($) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | *D-size* = 8 | | | | | | | *D-size* = 16 | |
| 1 | 2 | 2 | 22 | 0.644 | 8625996 | 31374002 | 2 | 2 | 25 | 0.457 | 9253339 | 30746659 |
| 2 | 2 | 2 | 23 | 0.628 | 8248724 | 30751274 | 2 | 2 | 18 | 0.561 | 9624012 | 30375986 |
| 3 | 2 | 2 | 18 | 0.541 | 8568942 | 31431056 | 2 | 2 | 16 | 0.575 | 9272982 | 30727016 |
| 4 | 2 | 2 | 19 | 0.525 | 9003355 | 30996643 | 2 | 2 | 21 | 0.662 | 9655526 | 30344472 |
| 5 | 2 | 2 | 20 | 0.452 | 8824330 | 31175668 | 2 | 2 | 18 | 0.648 | 9544954 | 30455044 |
| 6 | 2 | 2 | 22 | 0.479 | 9079421 | 30920577 | 2 | 2 | 22 | 0.478 | 9605371 | 30394627 |

According to Figure 6.3, the best fitness values of different GA implementations converge as generation increases no matter what *D-size* is. Based on the results presented in Table 6.5, the best fitness under *D-size* = 16 is bigger than that obtained from *D-size* = 8 on average, while the trend of total cost under these two *D-size* examples is opposite. A bigger *D-size* contributes to smaller total cost on average and larger fitness, which are desired.

The other case is also conducted to explore the impact of mean dependence cluster size, *D-size* = 8 and 16, respectively. In this case, *Pos* = 15. *Gen* = 60. *Pxo* = 0.8. *Pmu* = 0.25. *Tor* = 2. GA is performed 5 times for each *D-size* to check the convergence of the optimal solutions. The considered two scenarios have different disruption severities, $u_1 = 9$, $u_2 = 51$, respectively. 80 simulations are performed for each scenario. For the two scenarios, resilience loss constraints $\theta_1 = 0.005$, $\theta_2 = 0.3$. *CCP* constraints $v_1 = 0.7$, $v_2 = 0.8$. KKT multiplier $\lambda_i = 10^8$, $i = 1, 2$. The load - capacity model parameter $\alpha = 0.75$. Figure 6.4 illustrates the best fitness of each generation from different GA

implementations under two *D-size*. Table 6.6 presents GA optimization results under two *D-size* examples.
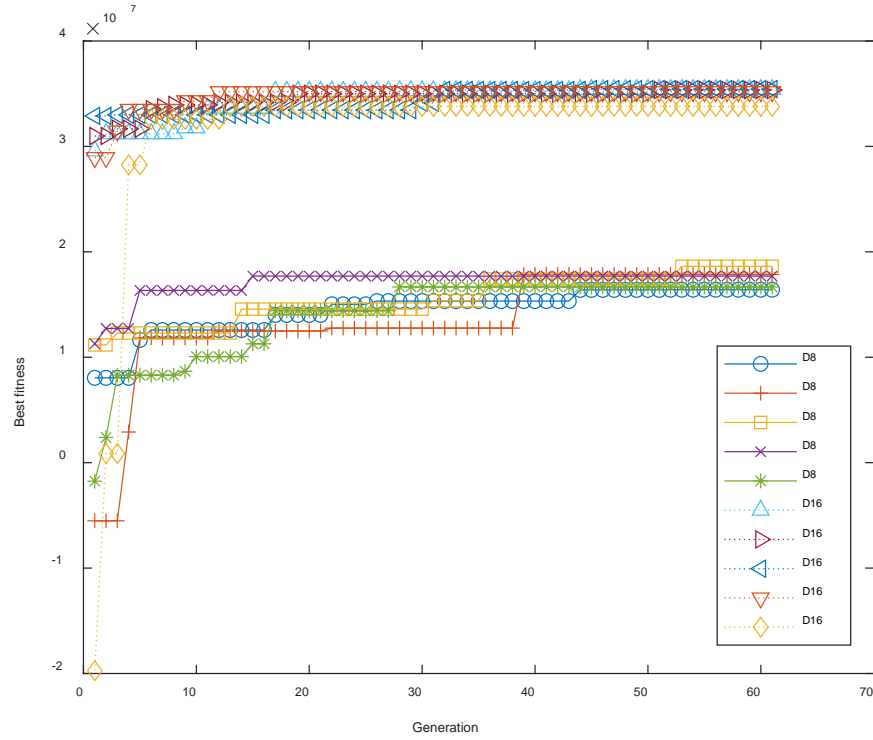


Figure 6.4 Fitness of the best member vs. generation from different GA implementations

with *D-size* = 8 and *D-size* = 16, respectively

Table 6.6 Best solutions of 5 GA implementations under different *D-size*

| GA | *D-size* = 8 | | | | | | *D-size* = 16 | | | | | |
|----|-----|-----|-----|-----|------|------|-----|-----|-----|------|------|------|
| | $\varphi$ | $\pi$ | $c$ | $\omega$ | Best fitness | Total cost ($) | $\varphi$ | $\pi$ | $c$ | $\omega$ | Best fitness | Total cost ($) |
| 1 | 2 | 2 | 15 | 0.566 | 16406445 | 14843554 | 3 | 2 | 23 | 0.46 | 35393634 | 14606365 |
| 2 | 3 | 2 | 21 | 0.388 | 17851637 | 14648362 | 3 | 2 | 18 | 0.529 | 35343310 | 14656689 |
| 3 | 2 | 2 | 33 | 0.6 | 18621216 | 15128783 | 3 | 2 | 22 | 0.528 | 35428779 | 14571220 |
| 4 | 2 | 2 | 30 | 0.55 | 17694914 | 14805085 | 3 | 2 | 23 | 0.655 | 35186550 | 14813449 |
| 5 | 2 | 2 | 29 | 0.554 | 16666721 | 14583278 | 1 | 2 | 35 | 0.517 | 33781703 | 16218296 |

It can be seen from Figure 6.4 that the best fitness values of GA implementations are becoming stable and close as generation increases, which indicates convergence of the optimal solution. Based on the results in Table 6.6, the best fitness values under *D-size* = 16 are bigger than that under *D-size* = 8 on average. While the average total cost under *D-size* =8 is bigger than that obtained when *D-size* = 16. According to Eq. (6.8), these results indicate that larger *D-size* leads to better system resilience recovery with smaller total cost. The findings in the above two cases are consistent.

When *D-size* is decreased, it has two implications. First, smaller mean size of dependence clusters means that customers receive electricity from fewer system nodes (centralized generation units). It may increase the chance of customer disconnections when system node failure propagates. Second, smaller mean dependence cluster size decreases the local customer load sharing scale, which helps to reduce the occurrence of overload node failures. Therefore, a trade-off results from decreasing mean dependence cluster size. In the above two cases, the impact of smaller mean dependence cluster size on customer disconnections is overwhelming. It increases recovery time and resilience loss, so achieved *CCP* is decreased while total cost is increased. As a result, the fitness of the desired hardening and restoration is decreased with a smaller *D-size*.

As mentioned before, distinct system disturbances might occur in real cases. The proposed combined optimization method simultaneously selects a set of hardening and restoration measures for all system disruptions. This method is helpful to minimize total cost while satisfying resilience requirements. However, post-disruption response needs not to be the same for different system disruptions. It is reasonable to implement restoration measures according to each specific system disruption while considering the

impact of pre-disruption hardening. Further research about two-stage optimization of system hardening and restoration is presented in the following section.

## 6.2 Two-stage optimization of system hardening and restoration

The integrated optimization of system hardening and restoration includes two stages, where the first stage focus on pre-disruption hardening measures. After system disruption occurs, the second stage decisions about restoration as post-disruption response are determined. The second stage decision is able to compensate for the impacts that result from the first stage decisions according to the system disruptions. Therefore, the first stage decision variables include hardening strategy $\varphi$, the percentage of top ranked nodes considered for hardening, $\omega$ (%) and the hardening levels of selected nodes, $\mathbf{\rho}(\omega)$. These decisions are made before system disruptions occur. The second stage decision variables include restoration strategy $\pi$, and the number of hired crews $c$. They are determined after the disruptions.

The objective function of the two-stage optimization problem is formulated as follows,

$$\max \ -C^{in}\left(\Phi\left(\varphi,\mathbf{\rho}(\omega)\right)\right) - \sum_{i=1}^{I} p_i\left(C_i^{ad}\left(\Phi(\varphi,\mathbf{\rho}(\omega)),\ \Pi(\pi,c),\mathbf{u}_i\right) - \lambda_i\gamma_i\left(\Phi(\varphi,\mathbf{\rho}(\omega)),\ \Pi(\pi,c),\mathbf{u}_i\right)\right)$$

where

$$C_i^{ad}\left(\Phi(\varphi,\mathbf{\rho}(\omega)),\ \Pi(\pi,c),\mathbf{u}_i\right) = C_i^{los}\left(\Phi\left(\varphi,\mathbf{\rho}(\omega)\right),\Pi(\pi,c),\mathbf{u}_i\right) + C_i^{r}\left(\Phi\left(\varphi,\mathbf{\rho}(\omega)\right),\Pi(\pi,c),\mathbf{u}_i\right).$$

It denotes the post-disruption cost which includes economic loss and restoration cost. The difference between this objective function and the previous one is that the restoration measures are determined based on different system disruptions given that hardening measures have been decided. $p_i$ denotes the probability that scenario $i$ occurs. System disruption severity represented by the number of initially failed nodes is also considered

as the system disruption factor. Other assumptions are the same with the problem formulated in section 6.1. GA is used to solve the mixed-integer optimization problem. Case studies are conducted to check the convergence of the optimal solutions and illustrate the impacts of system dependence. The two-stage optimization method is also compared with the combined optimization method proposed in section 6.1 to evaluate the effectiveness.

## 6.2.1 Convergence of the optimal solution

In the first case, *D-size* = 16. *Pos* = 15. *Gen* = 50. *Pxo* = 0.8. *Pmu* = 0.2. *Tor* = 3. GA is performed 8 times to check the convergence of the optimal solution. Two scenarios with disruption severities, $u_1 = 9$, $u_2 = 51$, are considered. $p_1 = 50\%$, $p_2 = 50\%$. 80 simulations are performed for each scenario. For the two scenarios, resilience loss constraints $\theta_1 = 0.005$, $\theta_2 = 0.3$. *CCP* constraints $v_1 = 0.7$, $v_2 = 0.8$, KKT multiplier $\lambda_i = 10^8$. $i = 1, 2$. The load-capacity model parameter $\alpha = 0.8$. Figure 6.5 shows the best member of each generation from different GA implementations. Table 6.7 presents the results of GA optimization.

Figure 6.5 The best fitness vs. generation from 8 GA implementations

It can be seen from Figure 6.5 that the fitness value of the best member dramatically increases as generation increases. Then the best fitness becomes stable as generation continues to increase. The fitness values of some optimal solutions from different GA implementations are very close while the difference caused by the uncertainties associated with the factors, such as specific initially failed nodes and required node repair time. The best solutions of 8 GA implementations are presented in Table 6.7.

Table 6.7. Best solutions of 8 GA implementations

| GA | $\varphi$ | $\omega$ | $\pi_1$ | $c_1$ | $\pi_2$ | $c_2$ | Best fitness |
|----|-----------|----------|---------|-------|---------|-------|--------------|
| 1 | 3 | 0.212 | 3 | 29 | 2 | 19 | 15,703,357 |
| 2 | 2 | 0.415 | 5 | 26 | 2 | 26 | 15,566,674 |
| 3 | 3 | 0.678 | 5 | 18 | 2 | 23 | 15,022,402 |
| 4 | 2 | 0.481 | 2 | 31 | 2 | 24 | 15,289,276 |
| 5 | 3 | 0.239 | 7 | 13 | 2 | 31 | 15,513,261 |
| 6 | 2 | 0.526 | 3 | 18 | 2 | 16 | 16,003,121 |
| 7 | 1 | 0.122 | 7 | 11 | 2 | 18 | 16,286,233 |
| 8 | 2 | 0.257 | 5 | 21 | 2 | 27 | 15,726,109 |

As shown in Table 6.7, different restoration strategies are selected for recovering the system subjected to different disruptions with the same hardening measures. It indicates that different restoration measures are implemented to compensate for the influence of distinct disruptions in order to minimize the total cost.

In the second case, $D\text{-}size = 16$. $Pos = 13$. $Gen = 100$. $Pxo = 0.8$. $Pmu = 0.2$. $Tor = 3$. GA is performed 10 times to check the convergence of the optimal solution. Two scenarios with disruption severities, $u_1 = 39$, $u_2 = 51$, are considered. $p_1 = 50\%$, $p_2 = 50\%$. 90 simulations are performed for each scenario. For the two scenarios, resilience loss constraints $\theta_1 = 0.25$, $\theta_2 = 0.3$. $CCP$ constraints $v_i = 0.8$, KKT multiplier $\lambda_i = 10^8$. $i = 1, 2$. The load-capacity model parameter $\alpha = 0.85$. Figure 6.6 shows the best member of each generation from 10 GA implementations. Table 6.8 presents the corresponding GA optimization results.
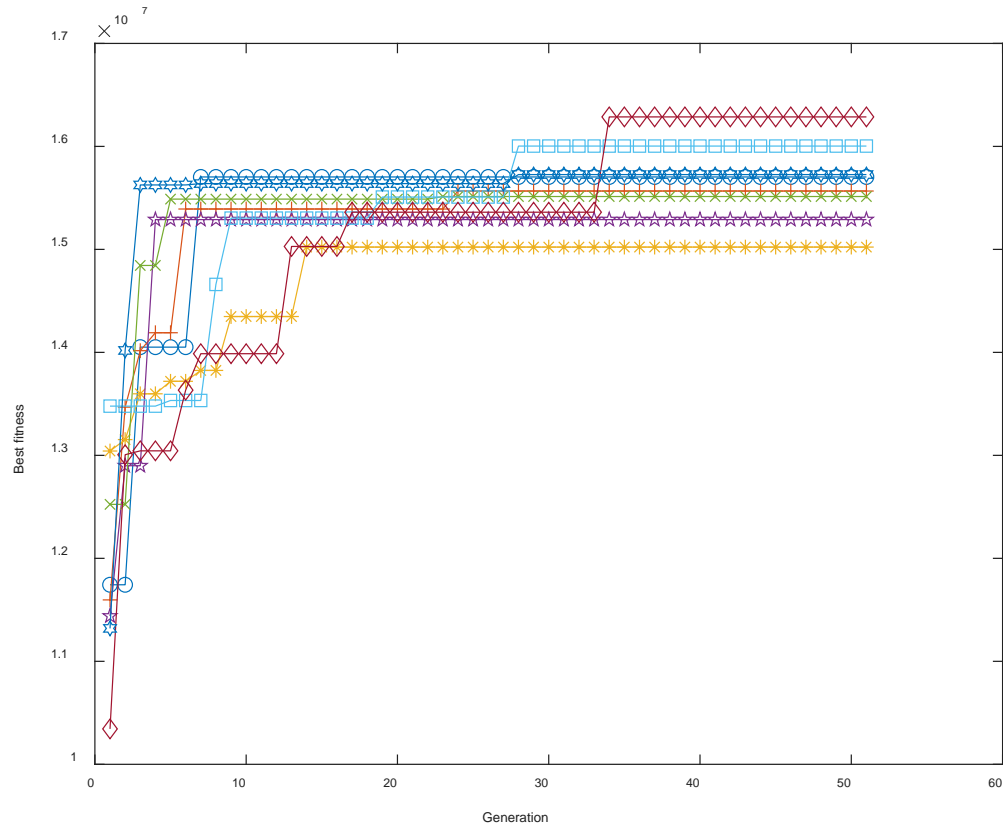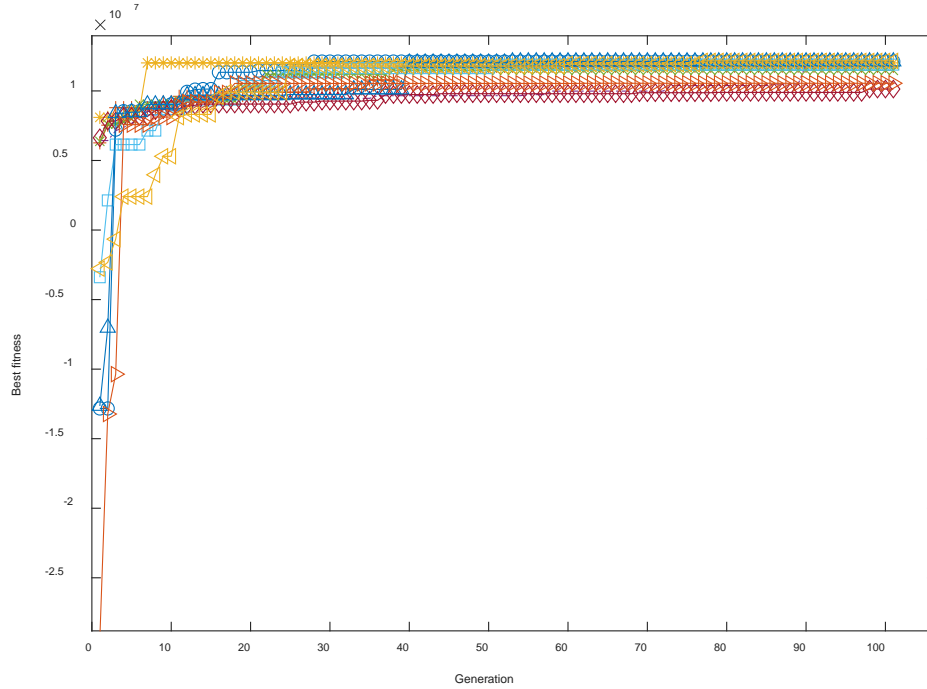
Figure 6.6 The best fitness vs. generation from 10 GA implementations

Based on Figure 6.6, the fitness value of the best member in each generation increases as generation increases at the early stage. Then the best fitness converges, and the best fitness values of the optimal solutions from different GA implementations are very close. The best solutions of 10 GA implementations are shown in Table 6.8.

Table 6.8. Best solutions of 10 GA implementations

| GA | $\varphi$ | $\omega$ | $\pi_1$ | $c_1$ | $\pi_2$ | $c_2$ | Best fitness |
|----|-----------|----------|---------|-------|---------|-------|--------------|
| 1  | 3 | 0.527 | 7 | 33 | 2 | 16 | 12,110,872 |
| 2  | 3 | 0.594 | 3 | 29 | 2 | 18 | 11,772,188 |
| 3  | 3 | 0.464 | 3 | 26 | 2 | 16 | 12,014,471 |
| 4  | 3 | 0.626 | 3 | 32 | 2 | 18 | 11,907,575 |
| 5  | 3 | 0.568 | 3 | 26 | 2 | 33 | 11,479,518 |
| 6  | 3 | 0.519 | 3 | 33 | 2 | 24 | 11,892,123 |
| 7  | 3 | 0.616 | 2 | 27 | 2 | 20 | 10,134,641 |
| 8  | 3 | 0.562 | 3 | 31 | 2 | 20 | 12,115,718 |
| 9  | 3 | 0.51  | 2 | 29 | 2 | 18 | 10,550,590 |
| 10 | 3 | 0.539 | 3 | 22 | 2 | 17 | 12,156,483 |

Similar to the results of the first case, different restoration measures are conducted for two scenarios with different disruption severities given the same hardening measures

are employed before system disruptions occur. Based on Table 6.8, HLH strategy is selected as the best hardening strategy, and HDFR strategy is selected as the restoration strategy for the second scenario with the highest disruption severity for all GA implementations. It indicates that the HDFR strategy recovers the system from severe disruptions effectively. Although there are uncertainties brought by initially failed node sets, required node repair time and load dynamics, the best fitness of different GA implementations are very close.

## 6.2.2   Impact of dependence

In this section, the impact of mean dependence cluster size, *D-size*, on the two-stage optimization regarding system hardening and restoration is analyzed.

In the numeric example, *D-size* = 8 and 16 are taken into consideration. *Pos* = 13. *Gen* = 80. *Pxo* = 0.8. *Pmu* = 0.2. *Tor* = 3. GA is implemented 10 times for each *D-size* to check the convergence of the optimal solution. Three scenarios have disruption severities, $u_1 = 39$, $u_2 = 51$, $u_3 = 63$. 80 simulations are performed for each scenario. For three scenarios, resilience loss constraints $\theta_1 = 0.25$, $\theta_2 = 0.3$, $\theta_3 = 0.35$. *CCP* constraints $v_i = 0.8$, KKT multiplier $\lambda_i = 10^8$. $i = 1, 2, 3$. The load-capacity model parameter $\alpha = 0.85$. Figure 6.7 illustrates the best member of each generation from 10 GA implementations under *D-size* = 8. Figure 6.8 presents the best fitness of each generation from 10 GA implementations under *D-size* = 16. Table 6.9 and Table 6.10 present GA optimization results under two *D-size* assumptions, respectively.
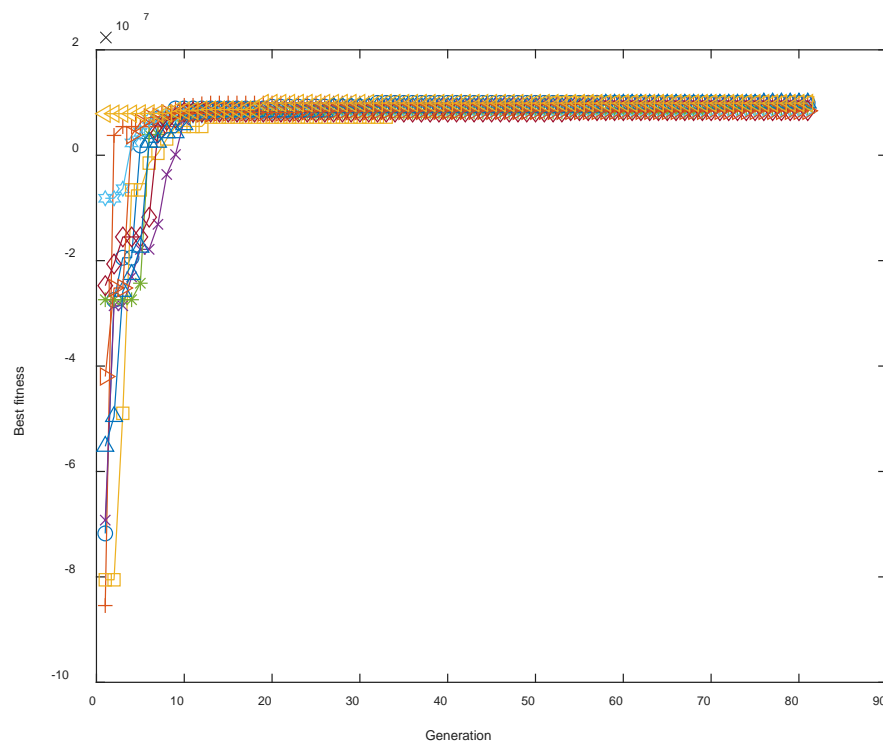
Figure 6.7 The best fitness vs. generation from 10 GA implementations (*D-size* = 8)
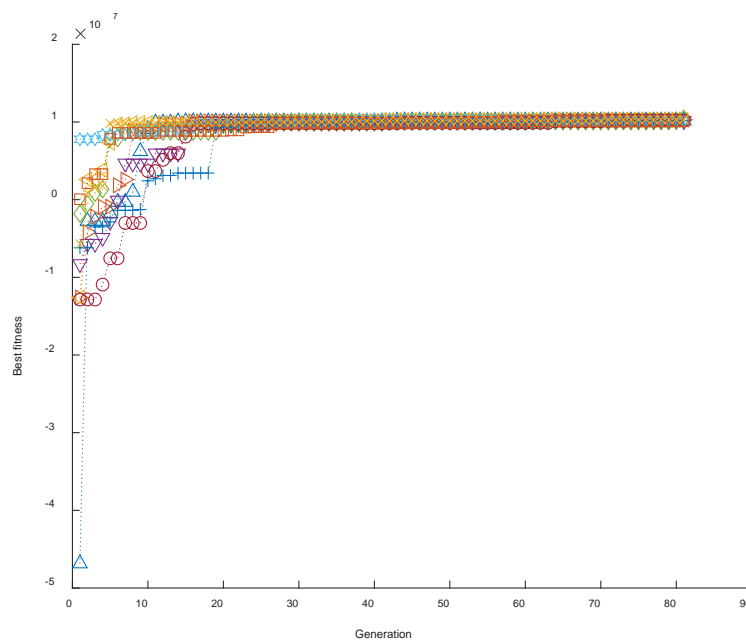


Figure 6.8 The best fitness vs. generation from 10 GA implementations (*D-size* = 16)

Both Figure 6.7 and 6.8 show that the best fitness converges as generation continuously increases after the dramatic changes with small generations. Table 6.9 presents GA optimization results under *D-size* = 8. Table 6.10 presents GA optimization results under *D-size* = 16.

Table 6.9. Best solutions of 10 GA implementations with *D-size* = 8

| GA | $\varphi$ | $\omega$ | $\pi_1$ | $c_1$ | $\pi_2$ | $c_2$ | $\pi_3$ | $c_3$ | Best fitness |
|----|-----|-------|-----|-----|-----|-----|-----|-----|--------------|
| 1 | 3 | 0.466 | 3 | 22 | 2 | 19 | 2 | 22 | 9,921,656 |
| 2 | 3 | 0.637 | 3 | 31 | 2 | 21 | 2 | 26 | 10,062,186 |
| 3 | 3 | 0.605 | 3 | 32 | 2 | 23 | 2 | 26 | 9,949,473 |
| 4 | 3 | 0.666 | 7 | 29 | 2 | 27 | 2 | 19 | 10,030,016 |
| 5 | 3 | 0.601 | 7 | 23 | 2 | 29 | 2 | 26 | 8,971,830 |
| 6 | 2 | 0.572 | 2 | 22 | 2 | 18 | 2 | 26 | 8,647,374 |
| 7 | 2 | 0.525 | 2 | 29 | 2 | 28 | 2 | 34 | 8,429,707 |
| 8 | 3 | 0.498 | 7 | 35 | 2 | 19 | 2 | 27 | 9,870,839 |
| 9 | 2 | 0.456 | 2 | 26 | 2 | 19 | 2 | 32 | 8,395,838 |
| 10 | 3 | 0.615 | 7 | 23 | 2 | 27 | 2 | 26 | 9,861,324 |

Table 6.10. Best solutions of 10 GA implementations with *D-size* = 16

| GA | $\varphi$ | $\omega$ | $\pi_1$ | $c_1$ | $\pi_2$ | $c_2$ | $\pi_3$ | $c_3$ | Best fitness |
|----|-----|-------|-----|-----|-----|-----|-----|-----|--------------|
| 1 | 2 | 0.367 | 3 | 20 | 2 | 19 | 2 | 17 | 10,163,213 |
| 2 | 3 | 0.545 | 7 | 25 | 2 | 19 | 2 | 17 | 10,285,389 |
| 3 | 3 | 0.534 | 7 | 25 | 2 | 20 | 2 | 17 | 10,272,538 |
| 4 | 1 | 0.609 | 3 | 23 | 2 | 22 | 2 | 27 | 10,145,439 |
| 5 | 3 | 0.678 | 7 | 27 | 2 | 25 | 2 | 16 | 10,431,056 |
| 6 | 2 | 0.584 | 3 | 30 | 2 | 21 | 2 | 13 | 10,230,641 |
| 7 | 3 | 0.625 | 3 | 28 | 2 | 19 | 2 | 19 | 10,200,502 |
| 8 | 3 | 0.695 | 3 | 29 | 2 | 18 | 2 | 17 | 10,183,446 |
| 9 | 3 | 0.43 | 3 | 27 | 2 | 17 | 2 | 16 | 10,116,163 |
| 10 | 3 | 0.604 | 7 | 26 | 2 | 33 | 2 | 18 | 10,305,952 |

According to Tables 6.9 and 6.10, higher best fitness value is obtained with bigger *D-size,* which is in accordance with the findings obtained in section 6.1.3.3. It can be observed that different restoration measures are adopted for different scenarios with distinct severities with the same pre-disruption hardening measures. HDFR strategy is selected for the scenarios with more initially failed nodes, i.e., more severe system disruptions. It means that HDFR strategy is more effective to rank failed nodes for repairing prioritization compared with other restoration strategies.

### 6.2.3 Comparison of two integrated optimization methods

In this section, numeric examples are conducted to compare the effectiveness of the combined optimization method proposed in section 6.1 and the two-stage optimization method. Note that the adopted network system structure, dependence clusters, customer connections are the same for the two integrated optimization methods. The uncertainties are mainly associated with the set of initially failed nodes, actual required repair time of system nodes and local customer load recovery level.

In the numeric example, *D-size* = 8. *Pos* = 15. *Gen* = 50. *Pxo* = 0.8. *Pmu* = 0.25. *Tor* = 3. Two scenarios have disruption severities, $u_1 = 9$, $u_2 = 51$ are considered. 80 simulations are performed for each scenario. For two scenarios, resilience loss constraints $\theta_1 = 0.005$, $\theta_2 = 0.3$. *CCP* constraints $v_1 = 0.7$, $v_2 = 0.8$, KKT multiplier $\lambda_i = 10^8$. $i = 1, 2$. The load-capacity model parameter $\alpha = 0.8$. GA is implemented 6 times for the two-stage optimization for each *D-size* while 4 GA implementations are conducted for the combined optimization proposed in section 6.1. Figure 6.9 illustrates the best fitness of each generation from different GA implementations.

Figure 6.9 The best fitness vs. generation from GA implementations under two

optimization methods

According to Figure 6.9, the best fitness values obtained from the two optimization methods are different. The best fitness from the two-stage optimization method is larger than that obtained from the combined optimization method. It demonstrates the higher effectiveness of two-stage optimization method compared with that of the combined optimization method. Post-disruption response made according to the specific system disruption contributes to better restoration effects. The best fitness firstly dramatically increases and then converges as generation increases no matter which optimization method is applied. Table 6.11 presents GA optimization results under the two-stage optimization method while Table 6.12 presents the results under the combined optimization method.

Table 6.11. Best solutions of 6 GA implementations under the two-stage optimization

method

| GA | $\varphi$ | $\omega$ | $\pi_1$ | $c_1$ | $\pi_2$ | $c_2$ | Best fitness | Total cost ($) |
|----|-----------|----------|---------|-------|---------|-------|--------------|----------------|
| | | | | | | Two-stage optimization method | | |
| 1 | 3 | 0.354 | 7 | 16 | 2 | 35 | 29,813,408 | 15,186,591 |
| 2 | 3 | 0.583 | 7 | 10 | 2 | 17 | 30,206,355 | 14,793,644 |
| 3 | 2 | 0.606 | 4 | 25 | 2 | 29 | 27,896,282 | 14,603,717 |
| 4 | 2 | 0.603 | 4 | 27 | 2 | 23 | 29,483,082 | 14,266,917 |
| 5 | 3 | 0.684 | 6 | 31 | 2 | 19 | 31,666,285 | 14,583,714 |
| 6 | 1 | 0.511 | 3 | 12 | 2 | 25 | 27,711,897 | 14,788,102 |

Table 6.12. Best solutions of 4 GA implementations under the combined optimization

method

| GA | $\varphi$ | $\pi$ | $c$ | $\omega$ | Best fitness | Total cost ($) |
|----|-----------|-------|-----|----------|--------------|----------------|
| | | | Combined optimization method | | | |
| 1 | 1 | 2 | 29 | 0.601 | 31,398,243 | 14,851,756 |
| 2 | 3 | 2 | 18 | 0.6 | 28,641,764 | 15,108,235 |
| 3 | 3 | 2 | 26 | 0.475 | 27,544,700 | 14,955,299 |
| 4 | 2 | 2 | 17 | 0.633 | 29,356,332 | 14,393,667 |

Based on Tables 6.11 and 6.12, the best fitness obtained from the two-stage optimization method is bigger than that obtained from the combined optimization method on average, which is more desirable. It proves that the two-stage optimization method is more efficient to optimize system hardening and restoration when compared with the combined optimization method.

In the second example, $D\text{-}size = 8$. $Pos = 15$. $Gen = 50$. $Pxo = 0.8$. $Pmu = 0.25$. $Tor = 3$. Two scenarios have disruption severities, $u_1 = 9$, $u_2 = 51$ are considered. 80 simulations

are performed for each scenario. For two scenarios, resilience loss constraints $\theta_1 = 0.005$, $\theta_2 = 0.3$. *CCP* constraints $v_1 = 0.7$, $v_2 = 0.8$, KKT multiplier $\lambda_i = 10^8$. $i = 1, 2$. The load-capacity model parameter $\alpha = 0.75$. GA is implemented 6 times for both two-stage optimization method and combined optimization method for each *D-size*. Figure 6.10 illustrates the best fitness of each generation from different GA implementations.
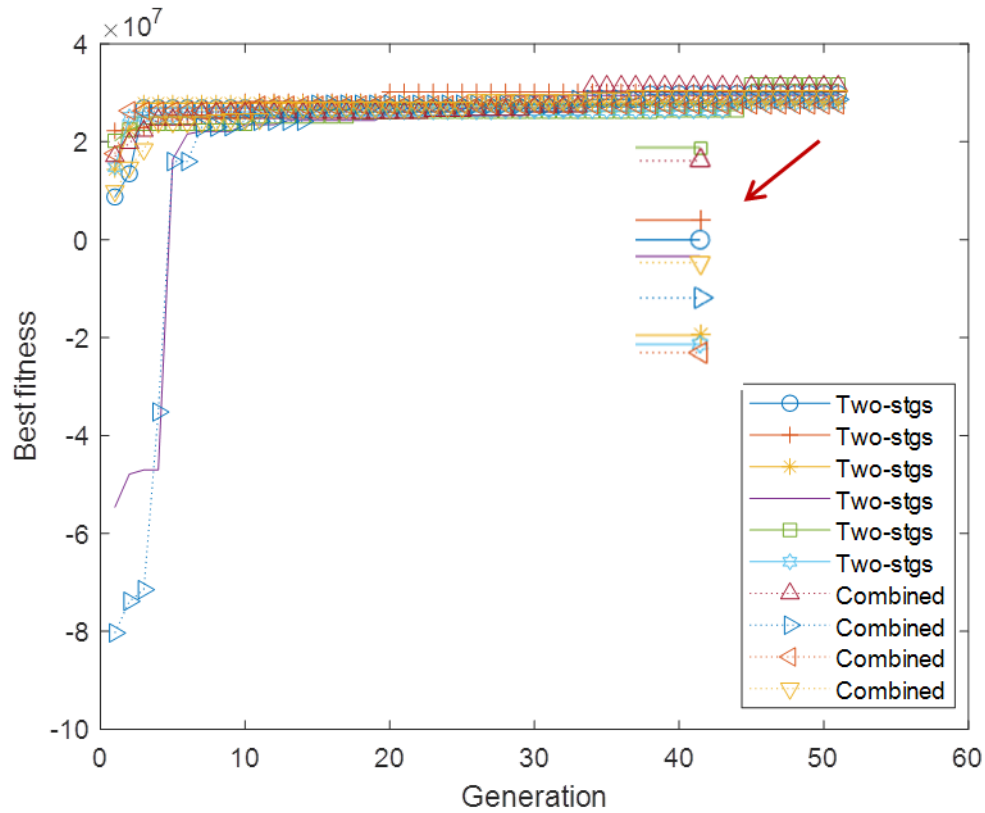


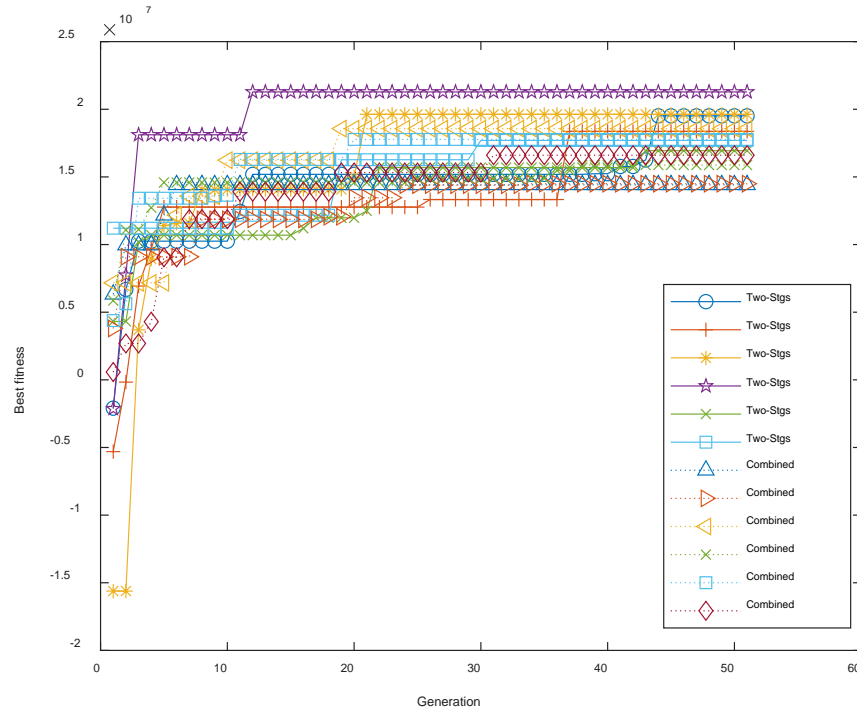Figure 6.10 The best fitness vs. generation from GA implementations under two optimization methods

As shown in Figure 6.10, the best fitness of each generation under each GA implementation becomes stable gradually as the number of generation increases. The best fitness values of some GA implementations are very close. Table 6.13 and Table 6.14 present the optimization results under two different optimization methods.

Table 6.13. Best solutions of 6 GA implementations under the two-stage optimization

method

| GA | Two-stage optimization method | | | | | | | |
| | $\varphi$ | $\omega$ | $\pi_1$ | $c_1$ | $\pi_2$ | $c_2$ | Best fitness | Total cost ($) |
|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 0.64 | 6 | 29 | 2 | 17 | 19,513,472 | 16,736,527 |
| 2 | 3 | 0.489 | 3 | 20 | 2 | 26 | 18,365,547 | 16,634,453 |
| 3 | 2 | 0.156 | 3 | 22 | 2 | 16 | 19,627,534 | 17,872,465 |
| 4 | 2 | 0.566 | 4 | 23 | 2 | 23 | 21,283,420 | 16216579 |
| 5 | 2 | 0.442 | 7 | 35 | 2 | 17 | 16,941,406 | 16808593 |
| 6 | 2 | 0.584 | 3 | 35 | 2 | 28 | 17,700,804 | 16049195 |

Table 6.14. Best solutions of 6 GA implementations under the combined optimization

method

| GA | Combined optimization method | | | | | |
| | $\varphi$ | $\pi$ | $c$ | $\omega$ | Best fitness | Total cost ($) |
|---|---|---|---|---|---|---|
| 1 | 3 | 2 | 22 | 0.522 | 14,416,787 | 16,833,213 |
| 2 | 2 | 2 | 17 | 0.54 | 14,499,439 | 16,750,560 |
| 3 | 2 | 2 | 21 | 0.469 | 18,572,326 | 16,427,673 |
| 4 | 2 | 2 | 25 | 0.562 | 15,8843,55 | 16,615,644 |
| 5 | 2 | 2 | 24 | 0.626 | 17,775,396 | 15,974,603 |
| 6 | 2 | 2 | 16 | 0.342 | 16,609,312 | 17,140,687 |

Based on Table 6.13, different restoration strategies and the number of crews are

hired to recover the system subjected to two disruptions with different severities.

Although the decisions about hardening and restoration are not exactly the same, which is

attributed to the uncertainties of the initially failed nodes, actual node repair time and

customer load recovery level, the best fitness values of different GA implementations are

still close. According to the results in the two tables, the best fitness obtained from the two-stage optimization method is larger than that obtained from the combined optimization method on average, which is consistent with the results from the first case.

Overall, the two-stage optimization method is demonstrated to be effective on integrally optimizing system hardening and restoration. Post-disruption restoration responses are employed considering the impacts of system disruptions given a pre-disruption hardening, which has better performance on minimizing total cost when compared with the combined optimization method. The results demonstrate that it is profitable to consider the restoration planning at the system hardening/design stage.

# 7  Conclusions

This dissertation presents a framework to jointly optimize pre-disruption system hardening and post-disruption restoration strategies to enhance system resilience against cascading failures with minimized total cost. Costs include investment associated with system hardening and restoration actions, and penalty cost of system performance loss resulting from system failures are taken into account. Network systems are the research subjects since large-scale network systems at any stage of the life-cycle are of great interest nowadays. The existence of network load dynamics and system dependency can endanger system stability by increasing the risk for failure spread. Therefore, it is necessary and important to analyze failure propagation mechanism in network systems and to enhance system resilience considering economic viability.

Since there are multiple dependencies among components and complex load dynamics in real-world network systems, new mixed cascading failure models, taking into account the joint impacts of system dependency, local and global load redistributions,

are proposed to describe failure propagation in systems. Based on the properties of a certain type of network systems, for example, electricity system, system load is divided into passing by load and local customer load, which is helpful to depict the impact of system dependence on overload failure propagation. A new resilience metric is proposed to evaluate resilience loss due to system disruptions. This metric provides opportunities to combine system performance with resilience evaluation and makes the resilience of different systems comparable and understandable. According to the proposed cascading failure models, the optimal selection of restoration strategies with distinct restoration preferences are performed for resilience enhancement. The impacts of system dependency on the effects of system restoration actions are also investigated. Then, the grid-outage resilience and economic benefits from installing solar array and battery size to customers, as hardening measures from customer perspective, are analyzed via a simulation-based optimization approach. This research is the first comprehensive attempt to investigate the manner in which cascading failures in bulk power system influence the entire electrical power system, including local customers connected by power distribution systems. New modeling approaches are presented to optimally and simultaneously perform system hardening and restoration under system disruptions with uncertainty. The integrated optimization of system hardening and restoration is to achieve an optimal balance between investment cost and economic loss while satisfying the resilience requirements.

In summary, the main contributions of this research include: 1) Combining the joint impacts of load dynamics and system dependency in modeling cascading failures in flow network systems; 2) Linking failure propagation with system hardening planning and

restoration implementation via hierarchical modeling of network systems; 3) Developing new optimization methods to search the optimal solutions for total cost minimization incorporating resilience requirements considering the uncertainty of system disruptions. The relationship between decision variables regarding system hardening and restoration measures is taken into account to make solutions robust; 4) Extending the impacts of cascading failures in bulk power system to local customers through power distribution systems to conduct integrated analysis of electricity system resilience; 5) Collectively and simultaneously optimizing system hardening and system restoration for the purpose of resilience enhancement, and minimizing investment costs and system damage cost caused by cascading failures together.

Ultimately, the proposed joint optimization framework for system hardening and restoration strategies in terms of resilience enhancement can be utilized to provide suggestions for future effective and economical system hardening and restoration regarding different kinds of real-life network systems against cascading failures.

There are great potentials of research extensions based on the current work. The possible extensions are introduced as follows:

Realistic characteristics of different parts of electric power systems can be taken into account in the future works. For example, system dependence clusters can be established according to real datasets regarding system failure propagation, which also can be used to modify the cascading failure model. Different popular DERs can be attached to local distribution systems as hardening investment to provide grid-outage resilience. In addition to the trade-off between investment and outage loss, the environmental benefits from utilizing renewable energy for power resilience and the incurred production

variability can be incorporated into the optimization problem. Furthermore, the influence of different types of system disturbances, such as extreme events and malicious attacks with uncertainty, on network systems can be investigated using the proposed framework based on real system datasets. How to use the data obtained from the real-world system cascading-restoration process, which reflects the interaction between failure propagation and action response, to improve future system hardening planning and restoration planning can be another further step. Using deep learning-based techniques for forecasting the effectiveness of system hardening and restoration regarding specific system disruptions based on previous failure datasets is another attractive direction.

# 8   References

[1] NERC, State of Reliability 2013, 2013.

[2] L.B. Lave, J. Apt, A. Farrell, M.G. Morgan, Increasing the Security and Reliability of the US Electricity System, The Economic Impacts of Terrorist Attacks, Cheltenham, UK, and Northampton, Mass  (2005) 57-69.

[3] K. Zhao, A. Kumar, T.P. Harrison, J. Yen, Analyzing the resilience of complex supply network topologies against random and targeted disruptions, IEEE Systems Journal 5(1) (2011) 28-39.

[4] B. Liscouski, W. Elliot, Final report on the august 14, 2003 blackout in the united states and canada: Causes and recommendations, A report to US Department of Energy 40(4) (2004).

[5] R. Nateghi, S.D. Guikema, Y.G. Wu, C.B. Bruss, Critical assessment of the foundations of power transmission and distribution reliability metrics and standards, Risk analysis 36(1) (2016) 4-15.

[6] N.L.a.S. Department, NERC Memorandum, 2012,  (2012).

[7] F.A.Q. Enery Information Administration, "How many power plants are there in the United States?",  (2016).

[8] E.F. Giles, K.L. Brown, UDI Directory of Electric Power Producers and Distributors: 123rd Edition of the Electrical World Directory,  (2014).

[9] U.S.D.o.E.O.o.E.D.a.E.R. President's Council of Economic Advisers, Economic Benefits of Increasing Electric Grid Resilience to Weather Outages, The Council, 2013.

[10] t.Q.E.R.Q.T. Force, Quadrennial Energy Review Second Installment: Transforming The Nation's Electricity System, Quadrennial Energy Review (QER), 2017.

[11] G.o.t.U.S.a. Canada, Joint United States-Canada Electric Grid Security and Resilience Strategy,  (2016).

[12] O.o.E.D.a.E.R. Department of Energy (DOE), Comparing the Impacts of Northeast Hurricanes on Energy Infrastructure,  (2013).

[13] B.T. Eaton, United States Annual Report 2012, Tech. Rep., 2013 [Online]. Available: http://powerquality. eaton. com/blackouttracker, 2014.

[14] R. Campbell, Weather-Related Power Outages and Electric System Resiliency, Congressional Research Service, 2012.

[15] J. Winkler, L. Duenas-Osorio, R. Stein, D. Subramanian, Performance assessment of topologically diverse power systems subjected to hurricane events, Reliab. Eng. Syst. Saf. 95(4) (2010) 323-336.

[16] I.S. ASSOCIATION, IEEE Guide for Electric Power Distribution Reliability Indices, IEEE, 2012.

[17] NERC, Definition of "Adequate Level of Reliability",  (2007).

[18] EPSA, Analysis of Electric Emergency Incident and Disturbance Report (Form OE-417), Washington DC, 2012.

[19] D. Henry, J.E. Ramirez-Marquez, On the Impacts of Power Outages during Hurricane Sandy—A Resilience-Based Analysis, Systems Engineering 19(1) (2016) 59-75.

[20] Y. Fang, N. Pedroni, E. Zio, Optimization of Cascade-Resilient Electrical Infrastructures and its Validation by Power Flow Modeling, Risk Analysis 35(4) (2015) 594-607.

[21] Kintner-Meyer, M.C. W., Homer, J. S, Balducci, P. J, Weimar, M. R, Valuation of Electric Power System Services and Technologies, Pacific Northwest National Lab. (PNNL), Richland, WA (United States), 2017.

[22] E. Zio, R. Piccinelli, Randomized flow model and centrality measure for electrical power transmission network analysis, Reliab. Eng. Syst. Saf. 95(4) (2010) 379-385.

[23] Y.-K. Lin, C.-T. Yeh, Maximal network reliability for a stochastic power transmission network, Reliab. Eng. Syst. Saf. 96(10) (2011) 1332-1339.

[24] A.M. Rei, A.L. Da Silva, J.L. Jardim, J. Mello, Static and dynamic aspects in bulk power system reliability evaluations, IEEE Transactions on Power Systems 15(1) (2000) 189-195.

[25] H. Dong, L.R. Cui, System Reliability Under Cascading Failure Models, IEEE Transactions on Reliability 65(2) (2016) 929-940.

[26] A.H. Michael Hyland, Tyler Doyle, Ji Yoon Lee, APPA Distribution system reliability & operations survey report, 2014.

[27] W. Warwick, T. Hardy, M. Hoffman, J. Homer, Electricity Distribution System Baseline Report, (2016).

[28] J. Eto, "How Reliable Is Transmission Compared to Distribution and What Do Power Interruptions Really Cost Customers", 2016.

[29] M. Figueroa-Candia, F.A. Felder, D.W. Coit, Resiliency-based optimization of restoration policies for electric power distribution systems, Electric Power Systems Research 161 (2018) 188-198.

[30] Y. Jiang, C.-C. Liu, Y. Xu, Smart Distribution Systems, Energies 9(4) (2016) 297.

[31] S.D. Guikema, S.M. Quiring, S.R. Han, Prestorm estimation of hurricane damage to electric power distribution systems, Risk analysis 30(12) (2010) 1744-1752.

[32] G. Li, P. Zhang, P.B. Luh, W. Li, Z. Bie, C. Serna, Z. Zhao, Risk analysis for distribution systems in the northeast US under wind storms, IEEE Transactions on Power Systems 29(2) (2014) 889-898.

[33] C. Marnay, J.S. Chard, K.S. Hamachi, T. Lipman, M.M. Moezzi, B. Ouaglal, A.S. Siddiqui, Modeling of customer adoption of distributed energy resources, Lawrence Berkeley National Laboratory (2001).

[34] Y. Hayashi, J. Matsuki, Loss minimum configuration of distribution system considering N-1 security of dispersed generators, IEEE Transactions on power systems 19(1) (2004) 636-642.

[35] Z. Bie, P. Zhang, G. Li, B. Hua, M. Meehan, X. Wang, Reliability evaluation of active distribution systems including microgrids, IEEE Transactions on power systems 27(4) (2012) 2342-2350.

[36] J.-W. Wang, L.-L. Rong, Robustness of the western United States power grid under edge attack strategies due to cascading failures, Safety science 49(6) (2011) 807-812.

[37] D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world' networks, Nature 393(6684) (1998) 440.

[38] K. Alvehag, L. Soder, A reliability model for distribution systems incorporating seasonal variations in severe weather, IEEE Transactions on Power Delivery 26(2) (2011) 910-919.

[39] D. Issicaba, J.A.P. Lopes, M.A. da Rosa, Adequacy and security evaluation of distribution systems with distributed generation, IEEE Transactions on Power Systems 27(3) (2012) 1681-1689.

[40] P. Hines, J. Apt, S. Talukdar, Trends in the history of large blackouts in the United States, Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century, IEEE, 2008, pp. 1-8.

[41] V. Rosato, L. Issacharoff, F. Tiriticco, S. Meloni, S. Porcellinis, R. Setola, Modelling interdependent infrastructures using interacting dynamical models, International Journal of Critical Infrastructures 4(1-2) (2008) 63-79.

[42] J.J. Romero, Blackouts illuminate India's power problems, IEEE Spectrum 49(10) (2012) 11-12.

[43] M.L. Sachtjen, B.A. Carreras, V.E. Lynch, Disturbances in a power transmission system, Physical Review E 61(5) (2000) 4877-4882.

[44] I. Dobson, B.A. Carreras, V.E. Lynch, D.E. Newman, Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization, Chaos 17(2) (2007) 13.

[45] G. Andersson, P. Donalek, R. Farmer, N. Hatziargyriou, I. Kamwa, P. Kundur, N. Martins, J. Paserba, P. Pourbeik, J. Sanchez-Gasca, R. Schulz, A. Stankovic, C. Taylor, V. Vittal, Causes of the 2003 major grid blackouts in North America and Europe, and recommended means to improve system dynamic performance, IEEE Transactions on Power Systems 20(4) (2005) 1922-1928.

[46] S.N. Dorogovtsev, J.F. Mendes, Evolution of networks, Advances in physics 51(4) (2002) 1079-1187.

[47] A.E. Motter, Y.-C. Lai, Cascade-based attacks on complex networks, Physical Review E 66(6) (2002) 065102.

[48] P. Crucitti, V. Latora, M. Marchiori, Model for cascading failures in complex networks, Physical Review E 69(4) (2004) 045104.

[49] X.-y. LI, H.-l. WANG, Urban public transport network study of cascading failure, Machinery & Electronics (2010) S1.

[50] J. Zhou, N. Huang, X. Wang, F. Zhao, An improved model for cascading failures in complex networks, 2012 IEEE 2nd International Conference on Cloud Computing and Intelligence Systems, 2012, pp. 721-725.

[51] J. Zhou, N. Huang, X. Sun, K. Wang, H. Yang, A new model of network cascading failures with dependent nodes, Reliability and Maintainability Symposium (RAMS), 2015 Annual, IEEE, 2015, pp. 1-6.

[52] D.J. Watts, A simple model of global cascades on random networks, Proceedings of the National Academy of Sciences 99(9) (2002) 5766-5771.

[53] S. Hong, C. Lv, T. Zhao, B. Wang, J. Wang, J. Zhu, Cascading failure analysis and restoration strategy in an interdependent network, Journal of Physics A: Mathematical and Theoretical 49(19) (2016) 195101.

[54] K.M. Lee, K.I. Goh, I.M. Kim, Sandpiles on Multiplex Networks, J. Korean Phys. Soc. 60(4) (2012) 641-647.

[55] I. Dobson, J. Chen, J. Thorp, B.A. Carreras, D.E. Newman, Examining criticality of blackouts in power system models with cascading events, System Sciences, 2002. HICSS. Proceedings of the 35th Annual Hawaii International Conference on, IEEE, 2002, p. 10 pp.

[56] I. Dobson, B.A. Carreras, V.E. Lynch, B. Nkei, D.E. Newman, Estimating failure propagation in models of cascading blackouts, Probability in the Engineering and Informational Sciences 19(4) (2005) 475-488.

[57] J. Kim, I. Dobson, Approximating a loading-dependent cascading failure model with a branching process, IEEE Transactions on Reliability 59(4) (2010) 691-699.

[58] I. Dobson, B.A. Carreras, D.E. Newman, A loading-dependent model of probabilistic cascading failure, Probability in the Engineering and Informational Sciences 19(1) (2005) 15-32.

[59] R. Parshani, S.V. Buldyrev, S. Havlin, Interdependent Networks: Reducing the Coupling Strength Leads to a Change from a First to Second Order Percolation Transition, Phys. Rev. Lett. 105(4) (2010) 048701.

[60] A. Vespignani, Complex networks: The fragility of interdependency, Nature 464(7291) (2010) 984-985.

[61] Y.N. Bai, N. Huang, L. Wang, Z.X. Wu, Robustness and Vulnerability of Networks with Dynamical Dependency Groups, Sci Rep 6 (2016) 9.

[62] F. Radicchi, Percolation in real interdependent networks, Nat Phys 11(7) (2015) 597-602.

[63] S.V. Buldyrev, R. Parshani, G. Paul, H.E. Stanley, S. Havlin, Catastrophic cascade of failures in interdependent networks, Nature 464(7291) (2010) 1025-1028.

[64] G.J. Baxter, S.N. Dorogovtsev, A.V. Goltsev, J.F.F. Mendes, Avalanche Collapse of Interdependent Networks, Phys. Rev. Lett. 109(24) (2012) 248701.

[65] R. Parshani, S.V. Buldyrev, S. Havlin, Critical effect of dependency groups on the function of networks, Proc. Natl. Acad. Sci. U. S. A. 108(3) (2011) 1007-1010.

[66] A. Bashan, R. Parshani, S. Havlin, Percolation in networks composed of connectivity and dependency links, Physical Review E 83(5) (2011) 8.

[67] A. Atputharajah, T.K. Saha, Power system blackouts-literature review, Industrial and Information Systems (ICIIS), International Conference on, IEEE, 2009, pp. 460-465.

[68] X. Weng, Y. Hong, A. Xue, S. Mei, Failure analysis on China power grid based on power law, Journal of Control Theory and Applications 4(3) (2006) 235-238.

[69] A. Bakshi, A. Velayutham, S. Srivastava, K. Agrawal, R. Nayak, S. Soonee, B. Singh, Report of the enquiry committee on grid disturbance in northern region on 30th July 2012 and in northern, eastern & north-eastern region on 31st July 2012, New Delhi, India (2012).

[70] H. Qi, L. Shi, Y. Ni, L. Yao, B. Masoud, Study on power system vulnerability assessment based on cascading failure model, PES General Meeting| Conference & Exposition, IEEE, 2014, pp. 1-7.

[71] J. Yan, Y. Tang, H. He, Y. Sun, Cascading failure analysis with DC power flow model and transient stability analysis, IEEE Transactions on Power Systems 30(1) (2015) 285-297.

[72] E. Hollnagel, D.D. Woods, N. Leveson, Resilience engineering: Concepts and precepts, Ashgate Publishing, Ltd.2007.

[73] D. Henry, J.E. Ramirez-Marquez, Generic metrics and quantitative approaches for system resilience as a function of time, Reliab. Eng. Syst. Saf. 99 (2012) 114-122.

[74] M. Bruneau, S.E. Chang, R.T. Eguchi, G.C. Lee, T.D. O'Rourke, A.M. Reinhorn, M. Shinozuka, K. Tierney, W.A. Wallace, D. Von Winterfeldt, A framework to quantitatively assess and enhance the seismic resilience of communities, Earthquake spectra 19(4) (2003) 733-752.

[75] A. Rose, Economic resilience to natural and man-made disasters: Multidisciplinary origins and contextual dimensions, Environmental Hazards 7(4) (2007) 383-398.

[76] M. Omer, A. Mostashari, U. Lindemann, Resilience analysis of soft infrastructure systems, Procedia Computer Science 28 (2014) 565-574.

[77] M. Ouyang, Z. Wang, Resilience assessment of interdependent infrastructure systems: With a focus on joint restoration modeling and analysis, Reliab. Eng. Syst. Saf. 141 (2015) 74-82.

[78] C.A. MacKenzie, C.W. Zobel, Allocating Resources to Enhance Resilience, with Application to Superstorm Sandy and an Electric Utility, Risk Analysis 36(4) (2016) 847-862.

[79] E.D. Vugrin, D.E. Warren, M.A. Ehlen, R.C. Camphouse, A framework for assessing the resilience of infrastructure and economic systems, Sustainable and resilient critical infrastructure systems, Springer2010, pp. 77-116.

[80] J. Ash, D. Newth, Optimizing complex networks for resilience against cascading failure, Physica A: Statistical Mechanics and its Applications 380(Supplement C) (2007) 673-683.

[81] H. Baroud, K. Barker, J.E. Ramirez-Marquez, C.M. Rocco, Inherent costs and interdependent impacts of infrastructure network resilience, Risk Analysis 35(4) (2015) 642-662.

[82] K. Barker, J.E. Ramirez-Marquez, C.M. Rocco, Resilience-based network component importance measures, Reliab. Eng. Syst. Saf. 117 (2013) 89-97.

[83] R. Pant, K. Barker, J.E. Ramirez-Marquez, C.M. Rocco, Stochastic measures of resilience and their application to container terminals, Computers & Industrial Engineering 70 (2014) 183-194.

[84] C.W. Zobel, Representing perceived tradeoffs in defining disaster resilience, Decision Support Systems 50(2) (2011) 394-403.

[85] S. Hosseini, K. Barker, J.E. Ramirez-Marquez, A review of definitions and measures of system resilience, Reliab. Eng. Syst. Saf. 145 (2016) 47-61.

[86] C. Balderer, M. Guarisco, M. Laumanns, R. Zenklusen, Repair strategies for minimising the risk of cascading failures in electricity networks, International Journal of Critical Infrastructures 5(1-2) (2009) 51-71.

[87] R. Francis, B. Bekera, A metric and frameworks for resilience analysis of engineered and infrastructure systems, Reliab. Eng. Syst. Saf. 121 (2014) 90-103.

[88] U.S.D.o. Energy, Staff Report to the Secretary on Electricity Markets and Reliability, (2017).

[89] E.P.R. Institute, Electric Power System Resiliency: Challenges and Opportunities, (2016) 56.

[90] C. Liu, D. Li, E. Zio, R. Kang, A modeling framework for system restoration from cascading failures, PloS one 9(12) (2014) e112363.

[91] B.E. Tokgoz, M. Safa, S. Hwang, Resilience Assessment for Power Distribution Systems, World Academy of Science, Engineering and Technology, International Journal of Civil, Environmental, Structural, Construction and Architectural Engineering 11(6) (2017) 806-811.

[92] T.T.H. Pham, Y. Bésanger, N. Hadjsaid, New challenges in power system restoration with large scale of dispersed generation insertion, IEEE Transactions on Power Systems 24(1) (2009) 398-406.

[93] H. Mo, G. Sansavini, Dynamic Defense Resource Allocation for Minimizing Unsupplied Demand in Cyber-Physical Systems Against Uncertain Attacks, IEEE Transactions on Reliability (2017).

[94] Y. Hou, C.-C. Liu, K. Sun, P. Zhang, S. Liu, D. Mizumura, Computation of milestones for decision support during system restoration, Power and Energy Society General Meeting,, IEEE, 2011, pp. 1-10.

[95] S.C. Madathil, Modeling and Analysis of Remote, Off-grid Microgrids, Industrial Engineering, Clemson University, 2017.

[96] A. Gutfraind, Optimizing topological cascade resilience based on the structure of terrorist networks, PloS one 5(11) (2010) e13448.

[97] F. Cadini, E. Zio, C.-A. Petrescu, Optimal expansion of an existing electrical power transmission network by multi-objective genetic algorithms, Reliab. Eng. Syst. Saf. 95(3) (2010) 173-181.

[98] H. Ren, I. Dobson, B.A. Carreras, Long-term effect of the n-1 criterion on cascading line outages in an evolving power transmission grid, IEEE transactions on power systems 23(3) (2008) 1217-1225.

[99] Y. Zinchenko, H. Song, W. Rosehart, Optimal Transmission Network Topology for Resilient Power Supply, ILS Conference, 2016, p. 9.

[100] J.P. Lopes, C. Moreira, F. Resende, Microgrids black start and islanded operation, 15th Power systems computation conference (PSCC), Liege, 2005.

[101] C. Moreira, F. Resende, J.P. Lopes, Using low voltage microgrids for service restoration, IEEE Transactions on Power Systems 22(1) (2007) 395-403.

[102] W. Quattrociocchi, G. Caldarelli, A. Scala, Self-healing networks: redundancy and structure, PloS one 9(2) (2014) e87986.

[103] H.P. Ren, J.H. Song, R. Yang, M.S. Baptista, C. Grebogi, Cascade failure analysis of power grid using new load distribution law and node removal rule, Physica a-Statistical Mechanics and Its Applications 442 (2016) 239-251.

[104] M. Parandehgheibi, E. Modiano, D. Hay, Mitigating Cascading Failures in Interdependent Power Grids and Communication Networks, International Conference on Smart Grid Communications, IEEE, New York, 2014, pp. 242-247.

[105] S. Hong, X.J. Zhang, J.X. Zhu, T.D. Zhao, B.Q. Wang, Suppressing failure cascades in interconnected networks: Considering capacity allocation pattern and load redistribution, Mod. Phys. Lett. B 30(5) (2016) 16.

[106] M.Q. Ahsan, A.H. Chowdhury, S.S. Ahmed, I.H. Bhuyan, M.A. Haque, H. Rahman, Technique to develop auto load shedding and islanding scheme to prevent power system blackout, IEEE transactions on Power Systems 27(1) (2012) 198-205.

[107] Y.R. Zhang, O. Yagan, Optimizing the robustness of electrical power systems against cascading failures, Sci Rep 6 (2016) 15.

[108] J. Zhou, N. Huang, X.L. Sun, L.D. Xing, S. Zhang, Network Resource Reallocation Strategy Based on An Improved Capacity-Load Model, Eksploat. Niezawodn. 17(4) (2015) 487-495.

[109] M. Adibi, L. Fink, Overcoming restoration challenges associated with major power system disturbances-Restoration from cascading failures, IEEE Power and Energy Magazine 4(5) (2006) 68-77.

[110] E. Zio, L.R. Golea, G. Sansavini, Optimizing protections against cascades in network systems: A modified binary differential evolution algorithm, Reliab. Eng. Syst. Saf. 103 (2012) 72-83.

[111] M. Adibi, N. Martins, Power system restoration dynamics issues, Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century, IEEE, 2008, pp. 1-8.

[112] B. Zhang, P. Dehghanian, M. Kezunovic, Optimal allocation of pv generation and battery storage for enhanced resilience, IEEE Transactions on Smart Grid (2017).

[113] N.D. Nordin, H.A. Rahman, A novel optimization method for designing stand alone photovoltaic system, Renew. Energy 89 (2016) 706-715.

[114] D.P. Birnie, Optimal battery sizing for storm-resilient photovoltaic power island systems, Sol. Energy 109 (2014) 165-173.

[115] E.W. Prehoda, C. Schelly, J.M. Pearce, US strategic solar photovoltaic-powered microgrid deployment for enhanced national security, Renewable and Sustainable Energy Reviews 78 (2017) 167-175.

[116] T. Khatib, I.A. Ibrahim, A. Mohamed, A review on sizing methodologies of photovoltaic array and storage battery in a standalone photovoltaic system, Energy Conv. Manag. 120 (2016) 430-448.

[117] A. Castillo, Risk analysis and management in power outage and restoration: A literature survey, Electric Power Systems Research 107 (2014) 9-15.

[118] J. Zhou, N. Huang, D.W. Coit, F.A. Felder, Combined effects of load dynamics and dependence clusters on cascading failures in network systems, Reliab. Eng. Syst. Saf. 170 (2018) 116-126.

[119] K.I. Goh, B. Kahng, D. Kim, Universal behavior of load distribution in scale-free networks, Phys. Rev. Lett. 87(27) (2001) 4.

[120] S.D. Li, L.X. Li, Y.X. Yang, Q. Luo, Revealing the process of edge-based-attack cascading failures, Nonlinear Dyn. 69(3) (2012) 837-845.

[121] R. Kinney, P. Crucitti, R. Albert, V. Latora, Modeling cascading failures in the North American power grid, Eur. Phys. J. B 46(1) (2005) 101-107.

[122] T. Zhou, J.G. Liu, B.H. Wang, Notes on the algorithm for calculating betweenness, Chin. Phys. Lett. 23(8) (2006) 2327-2329.

[123] P. Erdos, A. Rényi, On the evolution of random graphs, Bull. Inst. Internat. Statist 38(4) (1961) 343-347.

[124] Y.-N. Bai, N. Huang, L. Wang, Z.-X. Wu, Robustness and vulnerability of networks with dynamical dependency groups, Sci Rep 6 (2016) 37749.

[125] J.-W. Wang, L.-L. Rong, Edge-based-attack induced cascading failures on scale-free networks, Physica A: Statistical Mechanics and its Applications 388(8) (2009) 1731-1737.

[126] T.C. Gulcu, V. Chatziafratis, Y. Zhang, O. Yağan, Attack Vulnerability of Power Systems Under an Equal Load Redistribution Model, IEEE/ACM Transactions on Networking 26(3) (2018) 1306-1319.

[127] D.Q. Li, Q. Zhang, E. Zio, S. Havlin, R. Kang, Network reliability analysis based on percolation theory, Reliab. Eng. Syst. Saf. 142 (2015) 556-562.

[128] E. Zio, G. Sansavini, Modeling failure cascades in networks systems due to distributed random disturbances and targeted intentional attacks, Proceeding of the European Safety and Reliability Conference (ESREL 2008), 2008.

[129] I. Dobson, J. Chen, J. Thorp, B.A. Carreras, D.E. Newman, Examining criticality of blackouts in power system models with cascading events, Proceedings of the 35th annual Hawaii international conference on system sciences, IEEE, 2002, p. 10 pp.

[130] I. Dobson, B.A. Carreras, D.E. Newman, A branching process approximation to cascading load-dependent system failure, 37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the, IEEE, 2004, p. 10 pp.

[131] J. Zhong, H. Sanhedrai, F. Zhang, Y. Yang, S. Guo, S. Yang, D. Li, Network endurance against cascading overload failure, Reliab. Eng. Syst. Saf. (2020) 106916.

[132] V. Latora, M. Marchiori, Efficient behavior of small-world networks, Phys. Rev. Lett. 87(19) (2001) 198701.

[133] V. Colizza, R. Pastor-Satorras, A. Vespignani, Reaction–diffusion processes and metapopulation models in heterogeneous networks, Nature Physics 3(4) (2007) 276-282.

[134] A.-L. Barabási, R. Albert, Emergence of scaling in random networks, science 286(5439) (1999) 509-512.

[135] D.T. Nguyen, Y. Shen, M.T. Thai, Detecting critical nodes in interdependent power networks for vulnerability assessment, IEEE Transactions on Smart Grid 4(1) (2013) 151-159.

[136] E. Filiol, E. Franc, A. Gubbioli, B. Moquet, G. Roblot, Combinatorial optimisation of worm propagation on an unknown network, International Journal of Computer Science 2(2) (2007) 124-130.

[137] S. Pirzada, Applications of graph theory, PAMM: Proceedings in Applied Mathematics and Mechanics, Wiley Online Library, 2007, pp. 2070013-2070013.

[138] A. Gupta, M. Kaur, S. Mirkin, A. Singh, A. Goyal, Text summarization through entailment-based minimum vertex cover, Proceedings of the Third Joint Conference on Lexical and Computational Semantics (* SEM 2014), 2014, pp. 75-80.

[139] M. Liu, F.R. Yu, Y. Teng, V.C. Leung, M. Song, Performance optimization for blockchain-enabled industrial internet of things (iiot) systems: A deep reinforcement learning approach, IEEE Trans. Ind. Inform 15(6) (2019) 3559-3570.

[140] M. Bastian, S. Heymann, M. Jacomy, Gephi: an open source software for exploring and manipulating networks, Third international AAAI conference on weblogs and social media, 2009.

[141] E. Khalil, H. Dai, Y. Zhang, B. Dilkina, L. Song, Learning combinatorial optimization algorithms over graphs, Advances in Neural Information Processing Systems, 2017, pp. 6348-6358.

[142] S. Gu, T. Lillicrap, I. Sutskever, S. Levine, Continuous deep q-learning with model-based acceleration, International Conference on Machine Learning, 2016, pp. 2829-2838.

[143] A. Hagberg, P. Swart, D. S Chult, Exploring network structure, dynamics, and function using NetworkX, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2008.

[144] R. Bar-Yehuda, S. Even, A Local-Ratio Theorm for Approximating the Weighted Vertex Cover Problem, Computer Science Department, Technion, 1983.

[145] X. Sun, S. Wandelt, Complementary strengths of airlines under network disruptions, Safety science 103 (2018) 76-87.

[146] X. Fang, Q. Yang, W. Yan, Modeling and analysis of cascading failure in directed complex networks, Safety Science 65 (2014) 1-9.

[147] J. Skorupski, The simulation-fuzzy method of assessing the risk of air traffic accidents using the fuzzy risk matrix, Safety science 88 (2016) 76-87.

[148] J. Zhou, D.W. Coit, F.A. Felder, D. Wang, Resiliency-based restoration optimization for dependent network systems against cascading failures, Reliab. Eng. Syst. Saf. (2020 (under review)).

[149] N.R.E. Laboratory, PSM Global Horizontal Irradiance, in: N.D. Viewer (Ed.) 2016.

[150] N.R.E. Laboratory, Commercial and Residential Hourly Load Profiles for all TMY3 Locations in the United States, in: OpenEI (Ed.) 2013.

[151] T. Khatib, A. Mohamed, K. Sopian, M. Mahmoud, A new approach for optimal sizing of standalone photovoltaic systems, International Journal of Photoenergy 2012 (2012).

[152] H.A. Kazem, T. Khatib, A novel numerical algorithm for optimal sizing of a photovoltaic/wind/diesel generator/battery microgrid using loss of load probability index, International Journal of Photoenergy 2013 (2013).

[153] J. Zhou, S. Tsianikas, D.P. Birnie, D.W. Coit, Economic and resilience benefit analysis of incorporating battery storage to photovoltaic array generation, Renew. Energy 135 (2019) 652-662.

[154] S. Mandelli, C. Brivio, E. Colombo, M. Merlo, A sizing methodology based on Levelized Cost of Supplied and Lost Energy for off-grid rural electrification systems, Renew. Energy 89 (2016) 475-488.

[155] E. Leahy, R.S. Tol, An estimate of the value of lost load for Ireland, Energy Policy 39(3) (2011) 1514-1520.

[156] R. Khalilpour, A. Vassallo, Planning and operation scheduling of PV-battery systems: A novel methodology, Renew. Sust. Energ. Rev. 53 (2016) 194-208.

[157] E. Koutroulis, D. Kolokotsa, A. Potirakis, K. Kalaitzakis, Methodology for optimal sizing of stand-alone photovoltaic/wind-generator systems using genetic algorithms, Sol. Energy 80(9) (2006) 1072-1088.

[158] G. Merei, J. Moshövel, D. Magnor, D.U. Sauer, Optimization of self-consumption and techno-economic analysis of PV-battery systems in commercial applications, Applied Energy 168 (2016) 171-178.

[159] H. Louie, P. Dauenhauer, Effects of load estimation error on small-scale off-grid photovoltaic system design, cost and reliability, Energy Sustain Dev. 34 (2016) 30-43.

[160] C. Curry, Lithium-ion Battery Costs and Market, 2017.

[161] A. Van Der Welle, B. Van Der Zwaan, An overview of selected studies on the value of lost load (VOLL), Energy Research Centre of the Netherlands (ECN) (2007).

[162] G. Ambrosone, S. Catalanotti, U. Coscia, G. Troise, G. Cocurullo, Comparison between power and energy methods of analyses of photovoltaic plants, Solar Energy 34(1) (1985) 1-8.

[163] V. Aggarwal, What are the most efficient solar panels on the market? <https://news.energysage.com/what-are-the-most-efficient-solar-panels-on-the-market/>, (accessed January 1, 2018.).

[164] G. Electric Distribution Systems. Office of Electric, and Water, 2016 Electric Reliability Performance Report, Department of Public Service, 2017.

[165] K.C. Divya, J. Østergaard, Battery energy storage technology for power systems— An overview, Electric Power Systems Research 79(4) (2009) 511-520.

[166] I. Alsaidan, A. Khodaei, W. Gao, A comprehensive battery energy storage optimal sizing model for microgrid applications, IEEE Transactions on Power Systems 33(4) (2018) 3968-3980.

[167] V. Aggarwal, What are the most efficient solar panels on the market? <https://news.energysage.com/what-are-the-most-efficient-solar-panels-on-the-market/>, 2018 (accessed June 2, 2018.).

[168] B. Nykvist, M. Nilsson, Rapidly falling costs of battery packs for electric vehicles, Nature climate change 5(4) (2015) 329.

[169] P. Ralon, M. Taylor, A. Ilas, H. Diaz-Bone, K. Kairies, Electricity storage and renewables: costs and markets to 2030, International Renewable Energy Agency: Abu Dhabi, UAE (2017).

[170] B. Schäfer, D. Witthaut, M. Timme, V. Latora, Dynamically induced cascading failures in power grids, Nature communications 9(1) (2018) 1-13.

[171] J.G. Jin, K.M. Teo, A.R. Odoni, Optimizing bus bridging services in response to disruptions of urban transit rail networks, Transportation Science 50(3) (2016) 790-804.

[172] N. Yousefi, D.W. Coit, X. Zhu, Dynamic maintenance policy for systems with repairable components subject to mutually dependent competing failure processes, Computers & Industrial Engineering (2020) 106398.

[173] N. Chatwattanasiri, D.W. Coit, N. Wattanapongsakorn, System redundancy optimization with uncertain stress-based component reliability: Minimization of regret, Reliab. Eng. Syst. Saf. 154 (2016) 73-83.