Coordination and Optimization of Power Distribution Systems with Stochastic Distributed Energy Resources using Artificial Intelligence

Mohammed Al-Saffar

A thesis submitted in partial fufillment of the requirements for the degree of

Doctor of Philosophy
in
Energy Systems

Department of Electrical and Computer Engineering

University of Alberta

© Mohammed Al-Saffar, 2020

Abstract

High levels of penetration of distributed photovoltaic generators can cause serious overvoltage issues, especially during periods of high power generation and light loads. It is of vital importance to gain more understanding of the system and to prepare mitigation plans before the number of PV installations reaches a critical level. Therefore, properly assessing the PV hosting capacity is necessary. In this thesis, the hosting capacities of several real circuits in Alberta, Canada are evaluated using Monte Carlo simulation-based probabilistic power flow (MCS-based PPF) method. The examined circuits are located in the cities of Fort McMurray, Lloydminster, and Drumheller. These areas represent circuits of different sizes and complexities. The hosting capacities of the three regions were determined to be 10%, 60%, and 70%, respectively. Buses impacted by PV penetration were found in all three distribution networks. Factors influencing the PV hosting capacity are also identified and analyzed.

There have been many solutions proposed to mitigate the voltage problems, some of them using battery energy storage systems (BESS) at the PV generation sites. In addition to their ability to absorb extra power during the light load periods, BESS can also supply additional power under high load conditions. However, their capacity may not be sufficient to allow charging every time when power absorption is desired. Therefore, typical PV/BESS may not fully prevent over-voltage problems in power distribution grids. This thesis develops a co-operative state of charge control scheme to alleviate the BESS capacity problem through Monte-Carlo Tree Search based reinforcement learning (MCTS-RL). The proposed intelligent method coordinates the distributed batteries from other regions to provide voltage regulation in a distribution network. Furthermore, the energy optimization process during the day hours and the simultaneous state of

charge control are achieved using model predictive control (MPC). The proposed approach is demonstrated on two test cases, the IEEE 33 bus system and a practical medium size distribution system in Alberta Canada.

Optimization technology is developing to the point of becoming a cost-effective enabler of increased utilization of power transfer assets. This research presents a smart decomposition technique for the traditional optimal power flow (OPF) algorithm to allow distributed optimal power flow (DOPF) calculations without relying on a centralized controller. Hence, it develops a feasible distributed architectures for the electric power industry. The proposed method is implemented using the same algorithm MCTS-RL. This reduces computational complexity and avoids difficulties associated with stochastic modeling often used to capture the random nature of distributed energy resources (DER) units and loads. The efficiency of the optimization process is improved when the DOPF reflects the fast response capability of the optimal solution. This contribution provides results for a real-time dispatchable resource and demonstrates the flexibility of RL to adapt to changes in system states, ultimately reducing the generation cost while maintaining the system security constraints.

This thesis also develops a decomposition methodology for the traditional optimal power flow. It not only avoids the challenges associated with the stochastic nature of DERs and loads, but it also reduces the computational complexity of the conventional linear programming approach in the optimization problem. It does so using machine learning algorithms employed for two crucial tasks. First, MCTS-RL identifies clusters of network nodes to form a distributed architecture suitable for electric power transactions. Second, the network states updated by RL are used to execute conventional linear programming on a reduced set of lines identified during the previous step. The proposed approach is

demonstrated through a real-time balancing electricity market constructed over the IEEE 69-bus system and enhanced using price signals based on distribution locational marginal prices. This application clearly shows the ability of the new technique to effectively coordinate multiple distribution system entities while maintaining system security constraints.

Preface

This thesis is an original work by Mohammed Al-Saffar performed under the supervision of Dr. Petr Musilek. Parts of the thesis have been published or submitted to journals or conferences, as indicated below.

Chapter 3 of this thesis has been published as M. Al-Saffar, S. Zhang, A. Nassif and P. Musilek, "Assessment of Photovoltaic Hosting Capacity of Existing Distribution Circuits," 2019 IEEE Canadian Conference of *Electrical and Computer Engineering (CCECE)*, Edmonton, AB, Canada, 2019, pp. 1-4. Mohammed Al-Saffar conceived and designed the algorithm models of the stochastic distribution and assessment of photovoltaic hosting capacity. S. Zhang analyzed the data provided data by ATCO. A. Nassif provided the system data and advice on industry needs. P. Musilek took on this study in a supervisory role and contributed to the manuscript composition.

Chapter 4 has been published as M. Al-Saffar and P. Musilek, "Reinforcement Learning based Distributed BESS Management for Mitigating Overvoltage Issues in Systems with High PV Penetration," in *IEEE Transactions on Smart Grid*, 2020.

Chapter 5 has been published as M. Al-Saffar and P. Musilek, "Distributed Optimal Power Flow for Electric Power Systems with High Penetration of Distributed Energy Resources," 2019 IEEE Canadian Conference of *Electrical and Computer Engineering (CCECE)*, Edmonton, AB, Canada, 2019, pp. 1-5.

Chapter 6 of this thesis has been submitted to IEEE Access as M. Al-Saffar and P. Musilek, "Distributed Optimal Power Flow for Distribution Systems with Stochastic DER using Multi-Agent Reinforcement Learning".

Acknowledgements

I would like to thank my supervisor, Dr. Petr Musilek, for tremendous support throughout graduate school and for his dedicated work editing my conference papers and thesis, as well as for always pushing me to find opportunities. I would also like to thank my wife Fadya for her great moral support.

My parents, A.A. and N.K. deserve acknowledgement as they are the main reason behind every step of success in my life.

Thank you to the University of Alberta for the quality working environment, engaged professors and opportunities to work. The nurturing environment and keen colleagues are a derivative of institutional policies.

The support provided by Future Energy Systems under the Canada First Research Excellence Fund (CFREF) and the Natural Science and Engineering Research Council (NSERC) of Canada is gratefully acknowledged.

Contents

List of Figures

List of Tables

Nomenclature

a RL agent action.

a^j The action of the agent under learning period.

a^μ The actions of all agents except the agent under learning period.

B_s BESS capacity (kW).

C_1 Grid energy cost ($/kWh).

C_2 BESS energy cost ($/kWh).

C_p Generation output cost ($).

C_{purch} Electricity prices for the purchased power, ($/kWh).

C_r^b The required BESS capacity for the impacted region (kWh).

C_{sell} Electricity prices for the sold power, ($/kWh).

D_i Full load demand of the impacted buses (kW).

$DLMP$ Distribution locational marginal price (¢).

J_{Batt} BESS time-of-use (TOU) cost objective function ($/kWh).

J_g Grid purchase cost objective function ($/kWh).

k Prediction horizon samples minute.

LF Line Loss Factor.

N Time horizon minute.

N_{reg} The total number of buses in one of the impacted regions.

p Transition probability.

P_{Batt} BESS power generation (kW).

P_{ch} BESS charging power (kW).

P_{dis} BESS discharging power (kW).

P^{DER} DER power injection (kW).

P_g Grid power generation (kW).

P_{ij} Power transfer between the buses i and j (kW).

P_l Active power flow (kW).

P_L Load power (kW).

P_{Losses} Active Power flow losses (kW).

P_{PV} PV power generation (kW).

P_a^{rev} Reverse power of the assisting buses (kW).

P_i^{rev} Reverse power of the impacted buses (kW).

$\mathrm{PTDF}_{ij,k}$ Power Transfer Distribution Factor.

Q The action-value function.

Q_l Reactive power flow (kVAR).

Q_{Losses} Reactive Power flow losses (kVAR).

r Reward value.

s System state.

SOC BESS state of charge.

T_s Sample time (minute).

$\widehat{u}_b$ State space controller for the impacted and assisting regions (kW).

$V(s)$ Value function.

$V^*(s)$ The optimal value function.

y_{Batt} State space system output, BESS energy cost (\$/kWh).

y_g Grid energy cost for the state space system output (\$/kWh).

α Learning rate.

α_a BESS deployment rate of the participation of assisting region at the proposed CSOCC scenario.

α_b BESS deployment rate of without CSOCC scenario.

α_i BESS deployment rate of the impacted region at the proposed CSOCC scenario.

ε Exploration rate.

η_{ch} Charging efficiency of the BESS.

η_{dis} Discharging efficiency of the BESS.

γ Discount factor.

λ_0^P The active power exchange or the reference price (¢/kWh).

Ω_G DER units that belong to the feasible region.

Ω_L set of distribution lines in a feasible region.

φ_n The set of constrain factors under normal conditions.

$\pi(s)$ RL agent policy.

τ Constrain threshold.

θ_a Reverse power threshold of the assisting buses (kW).

θ_i Reverse power threshold of the impacted buses (kW).

Chapter 1

Introduction

Adoption of Distributed Energy Resources (DER)s in the smart grid aims to enhance the reliability, sustainability, and economics of distribution systems. An important example of fast-growing renewable technologies around the world are Photo-Voltaic (PV) generation systems [1]. These roof-mounted PV panels can provide significant cost savings by avoiding purchasing energy from the grid. It is also expected that PV power generation will alleviate the long term load growth in many distribution systems.

Studies have shown that high PV penetration can negatively affect voltage stability in the distribution grid. The main concern is the overvoltage due to the reverse power flow in distribution networks at the light load [2]. As a shortcut solution, utilities typically limit the PV penetration in the distribution feeder. However, this leads to losing a significant amount of PV power generation which might otherwise be beneficial for the system. This thesis uses a stochastic analysis framework to simulate possible PV deployment scenarios, to identify the locations of the critical buses in the grid and to examine the maximum acceptable levels of PV penetrations.

In chapter 3, analysis of PV hosting capacity (PVHC) is conducted using probabilistic power Flow (PPF). This part of the thesis also presents the results of PVHC assessment for three types of distribution networks. The probability distributions of bus voltages are examined with a predetermined threshold to determine the buses affected by the PV penetration levels in the system.

For particular issues, many solutions have been suggested to resolve the voltage rise problem due to high PV penetration. Most researchers rely on reactive power absorption by regulating the local Volt/VAR on the PV sites. Such designs are implemented in many types of equipment such as transformer tap-changers,

smart inverters, capacitor banks, etc. [3–5]. However, although it resolves the over-voltage issue, this solution tends to reduce the feeder power factor and increase of line losses [6,7].

A powerful alternative to maximize the local generation is to couple PV generators with the battery energy storage system (BESS). This technology can increase the local consumption during the light load period and served the increased demand during heavy load periods [8–10]. Moreover, it can contribute to a better balancing of the power grid by smoothing out the demand peaks, and improve the reliability and economy of the distribution system. Unfortunately, BESSs have finite size and may become fully charged or discharged during the control process significantly reducing the efficacy of voltage rise mitigation.

To avoid shortcomings of requiring excessive BESS capacities or setting an upper limit of permitted PV penetration levels, we propose a BESS coordination strategy that does not need any extra equipment to be connected to the power network [11–14]. The DER units are highly stochastic and dynamic in their distribution and behaviour, which may lead to the growth of system uncertainty. This can be addressed using intelligent approach such as Reinforcement Learning (RL) [15–17], to coordinate the DGs so as to mitigate the over voltage problem.

Chapter 4 describes an intelligent technique referred to as Cooperative State of Charge Control (CSOCC). It is well suited for large, complex distribution networks. It suggests that the impacted regions (*i.e.* regions negatively affected by the voltage rise problem caused by high penetration of PV generators) should be supported by an assisting region with normal voltage conditions. This approach combines a centralized controller using Monte-Carlo Tree Search (MCTS) based Reinforcement Learning RL, and a decentralized controller using Model Predictive Control (MPC). The MPC is developed to control the SOC of participating BESS units, and optimize the battery scheduling to minimize energy consumption. To demonstrate the operation of CSOCC and evaluate its efficiency, the proposed approach has been tested on two distribution systems: a test model of IEEE 33 bus system and a medium size distribution system model of the city of Lloydminster, Alberta, Canada.

Available Transfer Capability (ATC) is characterized by the North American Electric Reliability Council (NERC) in conjunction with the Federal Energy Regulatory Commission (FERC). In order to notify all the energy market participants of a power system about the ATC, this information has to be available on an hourly or daily basis [1]. The two major challenges that make the task of ATC

calculation of a nonlinear power system challenging are computing speed and accuracy due to static and dynamic security constraints. Meanwhile, high DER penetration and growing power transfers are necessary for a competitive electrical power market; however, this is leading to the deregulation of power systems. For smooth transactions of power between areas, new technologies and assessment methods are urgently needed. Transfer capability of a power system also points out how much inter-area power transfer can be increased without system security violations. The vital information required for the planning and operation of the power systems can be obtained from these transfer capability calculations. Power transfer capability details provide system bottlenecks to the planners and the limits of the power transfers to the system operators. The repeated estimations of these transfer capabilities allow reduction of the risk of overloads, over-voltages, equipment damage and unexpected blackouts.

Chapter 5, explores the feasibility of fully distributed calculation architectures for the electric power industry. These architectures are represented by the decomposition the large optimal power flow (OPF) problems into distributed optimization and management regions. Moreover, they can handle the physical events and the circumstances of the power grid, and alleviate the massive computations of a centralized optimization problem.

Although the centralized optimization can be decomposed into many smaller distributed optimization problems, the resulting power transactions are very dynamic and stochastic in nature. Thus, forming their stochastic models beforehand is very challenging. As mentioned earlier, RL is a powerful tool for solving complex sequential decision-making problems. It can learn the optimal stochastic policy effectively without relying on prior information. Each distributed optimization problem is led and managed by a local agent that can exchange information with its neighboring agents [18] and limit the power transactions in the network. Chapter 6 proposes a multi-agent RL system which allows the agent controllers to adapt to any changes in the power network without limiting its reliance to the neighbouring buses. In addition, the multi-agent architecture can improve communication among the numerous network segments. This improves the ability of the system to deal with large, complex distribution networks, while maintaining system security.

1.1 Research Objectives

The key objectives of the presented research can be organized in the following three groups:

1. Voltage violation assessment in large distribution systems with high DER penetration:

    - To conduct a comprehensive assessment of the PV hosting capacity impacts on power distribution systems, concentrating on overvoltage problems.
    - To identify main factors affecting hosting capacity, and determine their correlations with the power system operation responses.

2. DER impact mitigation using energy management and coordination methods:

    - To identify regions affected by DER-induced overvoltage, so-called impacted regions, in distribution circuits of substantial size and complexity.
    - To develop an intelligent system for coordinating charging/discharging processes of BESSs to mitigate the overvoltage problem.
    - To minimize energy consumption by developing an optimization strategy that schedules battery operation.

3. Distributed optimal power flow for distribution systems with stochastic DER units:

    - To develop a methodology to divide OPF problems for large circuits into smaller sub-problems through analysis of node cohesiveness and generation of distributed optimization architectures suitable for electric power transactions.
    - To reduce the computational burden of conventional optimization methods such as linear programming.
    - To avoid violations of network security limits, such as congested lines due to the extra power flow and line losses.

1.2 Research Originality

The use of BESS at PV generator sites can significantly improve voltage stability of power systems, because they can absorb extra PV energy generated during periods of light load. In addition, BESS can also assist the generation during peak load periods. However, their extensive use is still limited by their relatively high cost. Consequently, their size limitations may result in a full charge or discharge during the control process and thus inability to further support voltage stability of the system. To address this issue, it has been proposed to use the available storage capacity of neighboring BESS units in distribution networks [11–14, 19]. This way, the charge/discharge power of BESS can be maximized using common battery scheduling. However, the operation strategies of such systems have been programmed in advance, and prior knowledge about the system must be provided to the controllers so they can identify regularities in the environment. Nevertheless, in many cases, the states of the environment are uncertain and/or they may change over time. Consequently, the implemented strategies may be unable to converge. In addition, the above mentioned studies only consider the SOC of the neighbouring BESS units. However, other factors (such as line losses and voltage levels) are not considered, although may also change over time and result in additional uncertainties.

These challenges call for a new system design that can learn strategies on its own, with a minimal prior knowledge about the environment. It is proposed to use an adaptive RL agent model approach that, unlike systems with supervised learning, can learn by trial-and-error without an explicit teacher. In addition, another intelligent technique of MCTS is implemented to build a feasible environment for the agent. This significantly reduces the high computational burden that would occur if the entire network were considered. It also assists and accelerates the progress of storing the state updates of particular buses in a tree-nodes configuration. Each node in the search tree can store multiple system states, such as the information about the available BESS capacities, bus voltage values, and line losses.

Very large, complex distribution networks, may also require detection of impacted buses and identification of assisting buses in the network. The newly proposed approach, *i.e.* CSOCC, is well suited for these tasks as it divides a large grid into multiple smaller networks based on the configuration of so-called impacted regions. Finally, due to the non-trivial nature of system costs, battery generation must be rescheduled using a sound optimization approach and the

overall control strategy must be changed accordingly. Here MPC is used as a decentralized controller with two main tasks: it optimizes battery scheduling to minimize energy consumption, and it estimates and controls the state of charge (SOC) of the BESS units used to mitigate the PV-induced overvoltage problems.

Recently, several distributed optimization architectures have been proposed to solve the problems associated with the use of the centralized optimization in distribution networks [20]. For distribution networks, in contrast to transmission networks, the major challenge is that power transactions are very dynamic. They also occur under varying load and generation conditions, and at different locations. In addition, to avoid limitations of conventional economic power dispatch algorithms, prior statistical characterization of all DER units and loads are needed simultaneously during the optimization process.

As mentioned earlier, RL is powerful tool for solving complex sequential decision-making and control problems. It has been successful in learning stochastic policies and converging in high-dimensional environments, without relying on prior characterization of the system. Liu et al. [18] proposed an approach using distributed RL agents that exchange control information with their neighbours. Each agent controller makes its own action decision based on its states and the states of the neighbouring controllers, in so called distributed cooperative mechanism. However, this study relies on limited information exchanges and power transactions since it utilized just neighbour buses. Moreover, it is based on a base case power flow and does not consider the real-time impact of the line flow variations on the grid.

To address these issues, nodal search associated with the DER optimization in a distribution grid is performed using the proposed intelligent method MCTS-RL. The agent learns the stochastic DER behaviour and network security constraints. As a result, it can optimize DER dispatch based on their price bids. It allows the controller to adapt to any changes in the network to maintain system security. MCTS facilitates navigation through the network (not just the neighbour buses) using tree search. It starts from the buyer bus and allocates the optimal DER units within the feasible region. Compared to centralized OPF, the MCTS-induced search space is significantly reduced. The network states updated by RL with the selected feasible set are then used to accelerate the conventional optimization approach of linear programming.

When dealing with complex, large distribution circuit, it is necessary to simultaneously monitor multiple network constraints and transfer their values

among the network segments during execution. Therefore, it is necessary to establish a suitable communication framework. In this thesis, we form a multi-agent system to facilitate interactions between these segments located in various locations of the circuit. This architecture also includes a central coordinator that can efficiently manages the information exchange among the agents in multi-region systems.

Due to the line flow variations that may occur after each power transaction in the network, the thermal security limit may be jeopardized. To mitigate this risk, line congestion and losses due to the extra flows are calculated in real time through the central coordinator and sent to agents to modify their optimization policy accordingly.

1.3 Thesis Organization

The chapters of the thesis are structured as follows:

1.3.1 Related Work and Theoretical Background (Chapter 2)

This chapter offers a focused literature review and background on the topics relevant to the research presented in this thesis. It also presents details of established algorithms and functionalities which are used in building the research models that are proposed and presented in later sections. The topics covered include decentralized and centralized control strategies, available transfer capability, nodal pricing, Monte Carlo tree search, and reinforcement learning.

1.3.2 Assessment of Photovoltaic Hosting Capacity of Existing Distribution Circuits (Chapter 3)

In recent years, the deployment of distributed energy resources (DERs) has increased significantly, especially the residential photovoltaic (PV) systems. PV power generation has a strictly daytime pattern often coincident with low demand period on most the distribution feeders. As result, it may cause significant impact on feeder voltage. Therefore, it is crucial to assess and quantify the PV system operations to enhance the understanding of these systems at high deployment rates before moving forward to their level. Thus, careful assessment and consideration of the PV hosting capacity limits are necessary. In this chapter, the hosting capacities of three real circuits in Alberta, Canada are evaluated using Monte Carlo simulation-based probabilistic power flow (MCS-based PPF). The

results reveal that the hosting capacities of the three examined circuits (Fort McMurray, Lloydminster, and Drumheller) range significantly between 10%–70%.

1.3.3 Reinforcement Learning based Distributed BESS Management for Mitigating Overvoltage Issues in Systems with High PV Penetration (Chapter 4)

As described in chapter 3, high PV penetration can cause serious overvoltage issues, essentially during periods of high power generation and light loads. Many solutions have been proposed to mitigate this type of voltage problems, for instance, reactive power absorption, control of local Volt/VAR using smart inverters, or on-load tap changer (OLTC) transformers. However, most of these solutions causing reduction of the system power factor. They also lead to an increased flow of the reactive current into the distribution conductors and, consequently, to increased losses. The integration of BESS with PV generator sites may help resolving the overvoltage problem during the light load periods and supply extra generation during the peak periods. However, their widespread deployment has been hindered by their relatively high cost. In addition, their capacity may not be adequate to allow charging every time when power absorption is required.

To compensate for the BESS capacity problem, this chapter develops a new method called cooperative state of charge control. This scheme combines Monte-Carlo Tree Search based Reinforcement Learning (MCTS-RL) and model predictive control (MPC). The proposed intelligent method coordinates distributed battery systems from neighbouring buses to provide power management and voltage regulation in the impacted region of the distribution network. MCTS-RL is used as a centralized controller that optimizes the cooperating buses in each network segment in terms of BESS availability, with minimal line losses. MPC serves as a decentralized controller by determining the amount of energy required to resolve the voltage problem, taking into account the battery state of charge (SOC). Two test circuits are cosndiered in this study: the IEEE 33 bus system, and a real distribution system of the sity of Lloydminster, Alberta, Canada. The results show that in addition to successful overvoltage mitigation, this approach also minimizes the energy consumption.

1.3.4 Distributed Optimal Power Flow for Electric Power Systems with High Penetration of Distributed Energy Resources (Chapter 5)

This chapter presents a smart decomposition technique for the conventional optimal power flow (OPF) to form a distributed optimal power flow (DOPF) architecture. This proposed algorithm demonstrates the usefulness of fully distributed architectures for the electric power industry through the assessment of all the physical entities of the network. It also examines the interaction among these entities to optimize their mutual power transactions.

To alleviate the computational complexity assciated with the stochastic nature of distributed energy resources (DER), this approach uses Monte Carlo Tree Search based reinforcement learning (MCTS-RL). RL also improves the efficiency of the optimization process: it divides the central optimization problem into smaller subproblems and optimizes the economy of regional operations. This presented results demonstrate the ability of the proposed algorithm to dispatch DER units in real time and show its adaptability to changes in the system states. The optimal dispatch minimizes the overall generation cost while maintaining system security constraints.

1.3.5 Distributed Optimal Power Flow for Electric Power Systems with Stochastic Distributed Energy Resources Using Multi-agent Architecture (Chapter 6)

Customer-owned DER entities tend to maximize their profits autonomously and independently. Consequently, the structure of modern grids becomes highly distributed and ensuing electric power markets more complex, and stochastic. Since these entities are physically interconnected in the same network, decision made by any entity may affect the rest of the entities in the network. Collectively, they may affect the power system assets. Hence, the capacity of the system infrastructure should be improved, for example by building new line assets that meet the new peak demand. To address this infrastructural concern, this chapter expands on the results presented in chapter 5 by proposing a new virtual architecture based on a cooperative multi-agent system. The proposed solution promises to mitigate the identified problems without the need to change the system infrastructure.

The proposed approach is based on RL that can efficiently learn the optimal stochastic policy in high-dimensional spaces and reach the goal state in a probabilistic fashion without relying on the prior information. The multi-agent

architecture makes RL even more powerful by facilitating effective communication among agents representing individual regions in the network. Along with multi-agent RL, this approach uses MCTS to increase the RL observability and reduce the system complexity.

1.3.6 Conclusion and Future Work (Chapter 7)

Chapter 7 summarized the work presented in this thesis and overviews its main contributions. It also identifies the limitations of the current work and shows possible directions for future research on the topics presented in this thesis.

Chapter 2

Related Work and Theoretical Background

This chapter offers a focused literature review on topics relevant to the research presented in this thesis. It also reviews the key concepts employed in the research topics presented in Chapters 3 through 6.

2.1 Related Work

Several studies have assessed PV the impact on the system voltage stability in electric power distribution systems. Tonkoski et al. identified the X/R ratio of the network lines as the main sensitivity factor and analyzed the influence of PV sites on feeder impedance [21]. The authors showed that the bus voltages are inversely proportional to this ratio. A similar study by Coogan et al. examined the locational dependency of PVHC in the network [22]. They determined that there is a correlation between the system voltage and the distance between the feeder and PV locations, and that the overvoltage problem is directly proportional to this distance. Rylander et al. investigated PV interconnections in LV distribution systems in commercial areas through an original streamlined technique [23]. Other studies [22, 24] added that the PVHC depends not only on the PV penetration level, but also on the sizes and locations of PV systems installed in residential areas, as well as on other factors such as circuit load conditions, reactive power of the circuit, and short circuit level of the system.

There have been many solutions proposed to alleviate the voltage rise problems associated with high levels of PV penetration. For example, Hashemi et al. [25] and Stetz et al. [26] proposed to use reactive power absorption and online trans-

former tap changers to deal with the overvoltage issues. However, such solutions may not hold as they could lead to the rise of power system losses and reduce the power factor. Although the use of smart inverters to control local Volt/VAR on the PV sites provides a more beneficial solution [7], high cost prohibits their widespread use.

A number of recent studies use a coordinated control of PV and BESS to improve the voltage stability by locally installed storage systems [9, 10]. However, they do not consider the limitation of full charge and discharge, and ignore the high cost of the BESS-based solutions, especially for battery systems with high capacities.

To minimize the required capacities of BESS, several authors proposed coordination of neighboring BESS units to adjust the charging/discharging rates [11–14, 19]. However, these strategies require prior knowledge about the system to assess its states comprehensively and avoid unstable solutions under the system uncertainties.

At the same time, power distribution systems are becoming more and more distributed. The numbers of DER sites in residential areas are continuously growing. This is associated with massive power transactions that may require additional generation capacity and new line assets to supply the peak demand. Therefore, the tools for calculating distributed OPF became increasingly important to guarantee efficient system operation while maintaining its security constraints.

Recent literature presents several approaches to the distributed OPF problem. Xia et al. [27] and Li et al. [28] proposed fully distributed methods that derive a set of distributed control laws for economic dispatch in a microgrid. In this method, the communication structure is designed in two layers. The first ensures the supply-demand balance in the microgrid and the second coordinates control agents. A decomposition of the centralized optimization problem into many smaller optimization problems has been presented through a divide-and-conquer approach [29, 30] where each agent exchanges information with its network neighbors. However, such approaches lack ways to model the stochastic behavior of the DERs in a distribution system.

Some authors [31–33] used Markov decision processes (MDPs) to resolve the randomness of DER units and loads, while other [34–36] applied distributed model predictive control (MPC) for stochastic dispatch optimization in microgrids. In each entity of a microgrid, there is a local MPC controller that executes a receding-horizon optimization. Although some of these studies consider sys-

tem stochasticity, they all have one shortcoming in common: they require prior knowledge of system parameters to guarantee the optimal solution convergence.

Reinforcement learning (RL) is based on a generalizable and powerful algorithm that can effectively solve complex sequential decision-making problems without relying on prior system information. A distributed cooperative mechanism using RL was proposed by Liu et al. [18]. In this approach, each distributed agent exchanges information with its neighbouring buses in a microgrid. In other words, it makes action decisions based on its state and the neighboring states. However, the system observability is insufficient for a wide range of numbers of buses when considering just neighbours. Therefore, it definitely limits the power transactions and may affect the DER selections in the optimization problem. Moreover, this approach does not consider the thermal security limit.

2.2 Background

2.2.1 Centralized and Decentralized Control Strategies

Coordination methods are usually based on two essential strategies: centralized and decentralized control. In centralized control, a central controller aggregates measurements from the entire network and determines the charging/discharging control tasks for each BESS. This approach is more efficient than the decentralized control as it is based on the current information about the entire network. In the event that a particular BESS cannot meet the requirements for smooth operation, the central controller can successively search for alternative units until the problem is resolved. However, centralized control demands a fast communications infrastructure, increasing the computation requirements and overall costs. Furthermore, in case of communication failure, BESS units may become unresponsive to the central controller [37,38]. In decentralized control, the monitoring and control of charging and discharging actions are performed locally by each BESS. The main advantage of this approach is that it does not require extra communication systems, making it robust and cost-effective. However, when a BESS unit is either fully charged/drained or can only provide a limited power supply/load, it is unable to request support from other BESS units. The unit may even experience a total failure resulting from the lack of communication with other units [39,40].

2.2.2 Available Transfer Capability (ATC)

An ATC assessment is essential for the successful implementation of electric power deregulation where the power producers and customers share a common distribution network for delivering power from the point of generation to the point of consumption. The ATC indicates the amount by which inter-area bulk power transfers can be increased without compromising system security. The value used for ATC affect both system security and the profits made in bulk power transactions. When the upper or lower ATC limits are exceeded, market participants can have conflicting interests. Thus, under deregulation, there is an increasing motivation for defensible calculations of ATC. In this thesis, ATC is calculated in a static power flow as follows [41–43]

$$\text{ATC} = \text{TTC} - \text{TRM} - \text{CBM}$$

where

TTC (Total Transfer Capability) is defined as the quantity of electric power that can be transferred over the interconnected path reliably without violating a predefined set of conditions of the system. The various constraints that limit TTC may be from the physical and/or electrical characteristics of the systems including thermal, voltage, and stability limits

$$\text{TTC} = \text{Minimum of } \{\text{Thermal Limit, Voltage Limit, Stability Limit}\}.$$

TRM (Transmission Reliability Margin) is defined as that amount of transmission, or line transfer capability, and is an essential assessment to ensure that the interconnected network line is secure under a reasonable range of uncertainties in system conditions.

CBM (Capacity Benefit Margin) is defined as the amount of line transfer capability reserved by load-serving entities to ensure access to generation from interconnected systems to meet generation reliability requirements.

The transfer capability is determined for the following reasons:

- The estimation of TTC can be used as a power system reliability indicator.
- It can be used for comparing the special merits of planned transmission or distribution system improvements.

- It can be used to improve reliability and economic efficiency of the electric power markets.
- It can be used for providing a quantitative basis for assessing transmission reservations to drive energy markets.

2.2.3 Power System with Uncertain Variables

With the increased integration of renewable energy resources into power grids such as wind and solar generation, a considerable number of uncertainties are introduced to system power flows. Regarding the uncertainty of power injection caused by intermittent renewable resources and market-driven operation, the active power and reactive power would oscillate within a particular range and can be inspected using the intervals. Uncertainty of power injection is caused intermittency of renewable generation and stochastic behavior of electricity market. As a result, the active and reactive power cannot be determined with certainty, but instead as a range of possible values that can be described using intervals [44, 45]

$$[P_i] = P_i(1 + [\xi_{pi}]); \quad [Q_i] = Q_i(1 + [\xi_{qi}])$$

where $[\xi_{pi}]$, and $[\xi_{qi}]$ are the variation range ratios of the active and reactive power, respectively. They are usually called the perturbed variables of active and reactive power injection, and are represented by a predefined interval.

The Taylor inclusion function is an interval analyzer that is used to solve a set of interval non-linear equations using inclusion functions. To make the interval width narrower, the high-order Taylor series expansion of functions are frequently used. Also, it's used to present the lower and upper bounds of the interval results; where, at the tightest bounds, the probability guarantees to enclose the actual distribution. The interval bounds that an interval variable assumes in the presence of a variation of $\pm\Delta x$ around a nominal value x_c, can be evaluated by the Taylor inclusion function, which is shown in Fig. 2.1.

$$F([x]) = F(x_c) + F^{'}(x_c)[\Delta x] + \cdots + \frac{1}{(n)!}F^n(x_c)[\Delta x]^n + \frac{1}{(n+1)!}F^{n+1}(\lambda)[\Delta x]^{n+1}$$

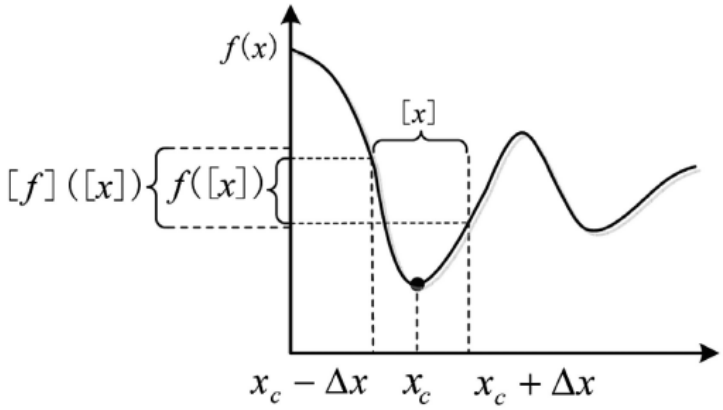


Figure 2.1: Taylor inclusion function.

$$0 = F(x_c, \xi_c) + \sum_{j=1}^{m} \Big(\sum_{i=1}^{N} \frac{\partial F}{\partial x_i} \frac{\partial x_i}{\partial \xi_j} \Big)_{(x_c, \xi_c)} \Delta\xi + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \Big(\sum_{p=1}^{N} \sum_{q=1}^{N} \frac{\partial^2 F}{\partial x_p \partial x_q} \frac{\partial x_p}{\partial \xi_i} \frac{\partial x_q}{\partial \xi_j} + 2 \sum_{p=1}^{N} \frac{\partial^2 F}{\partial x_p \partial \xi_i} \frac{\partial x_p}{\partial \xi_j} + \sum_{p=1}^{N} \frac{\partial F}{\partial x_p} \frac{\partial^2 x_p}{\partial \xi_i \partial \xi_j} + \frac{\partial^2 F}{\partial \xi_i \partial \xi_j} \Big)_{(x_c, \xi_c)}.$$

The first-order derivative of the equation is the conventional power flow and can be represented by the Newton iteration formula

$$-J_k \Delta x_k = \Delta F_k,$$
$$x_{k+1} = xk + \Delta x_k,$$

where J_k represents the Jacobian matrix at iteration k, and ΔF_k is the residual vector at iteration k. The first-order derivative of the equation can be expanded into matrix form as

$$\begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \cdots & \frac{\partial F_1}{\partial x_N} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} & \cdots & \frac{\partial F_2}{\partial x_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_N}{\partial x_1} & \frac{\partial F_N}{\partial x_2} & \cdots & \frac{\partial F_N}{\partial x_N} \end{bmatrix} \begin{bmatrix} \frac{\partial x_1}{\partial \xi_j} \\ \frac{\partial x_2}{\partial \xi_j} \\ \vdots \\ \frac{\partial x_N}{\partial \xi_j} \end{bmatrix} + \begin{bmatrix} \frac{\partial F_1}{\partial \xi_j} \\ \frac{\partial F_2}{\partial \xi_j} \\ \\ \frac{\partial F_N}{\partial \xi_j} \end{bmatrix} = 0$$

for j = 1,2,...,m. Expanding the second-order derivative of the equation also into matrix form gives

$$\begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \cdots & \frac{\partial F_1}{\partial x_N} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} & \cdots & \frac{\partial F_2}{\partial x_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_N}{\partial x_1} & \frac{\partial F_N}{\partial x_2} & \cdots & \frac{\partial F_N}{\partial x_N} \end{bmatrix} \begin{bmatrix} \frac{\partial^2 x_1}{\partial \xi_i \partial \xi_j} \\ \frac{\partial^2 x_2}{\partial \xi_i \partial \xi_j} \\ \vdots \\ \frac{\partial^2 x_N}{\partial \xi_i \partial \xi_j} \end{bmatrix} + \begin{bmatrix} [\frac{\partial x_1}{\partial \xi_i}]^T H(F_1) \frac{\partial F_1}{\partial \xi_j} \\ [\frac{\partial x_1}{\partial \xi_i}]^T H(F_1) \frac{\partial F_2}{\partial \xi_j} \\ \vdots \\ [\frac{\partial x_1}{\partial \xi_i}]^T H(F_1) \frac{\partial F_N}{\partial \xi_j} \end{bmatrix} = 0$$

where $H(F_p)$, for p = 1, 2, ..., N is the Hessian matrix of the function F_p.

According to NERC and FERC, the interactions between two independent bilateral power transactions are required to be checked and evaluated under the power market environment. Hence, it is required to know the maximum power transfer limit.

2.2.4 Nodal Pricing

Nodal pricing or Locational Marginal Pricing (LMP), represents the most efficient expression of locational energy prices. LMP, which is already used by some US utilities, calculates energy costs at specific locations. In current (traditional) networks, pricing schemes can be changed only at the transmission level by the Transmission System Operator (TSO). However, it is expected that a significant growth of DG penetration, combined with both load and generation present in distribution grids, can cause major issues affecting grid operation. Therefore, an appropriate location-based pricing will be a vital clue in our distribution grid. It can be demonstrated that pricing the electricity produced/consumed in each node at the local marginal cost leads to efficient operation decisions. As a result, these variations of the electricity prices from one node to another, would increase the revenues, reduce network losses, and maintain the grid infrastructure [46,47].

2.2.5 Monte Carlo Tree Search (MCTS)

MCTS is a proposed search method that uses random sampling to carry out a search of the paths along a tree. Typically, a game tree with a randomly biased sequence of actions is applied to the given states until a terminal condition is reached. MCTS has demonstrated its success for the game Go which has been considered one of the hardest games for Artificial Intelligence (AI) to solve. MCTS is a method for determining the optimal decisions in a given space by taking random samples in that space and building a search tree according to the sequential decision results. The basic algorithm involves iteratively building a search tree until some predefined computational budget is reached, such as the

depth limit. At this point, the search is stopped, and the best performing action is returned. Each node in the search tree represents a state of the domain and directed links to child nodes represent actions leading to subsequent states. Four steps are required in this search policy [48]

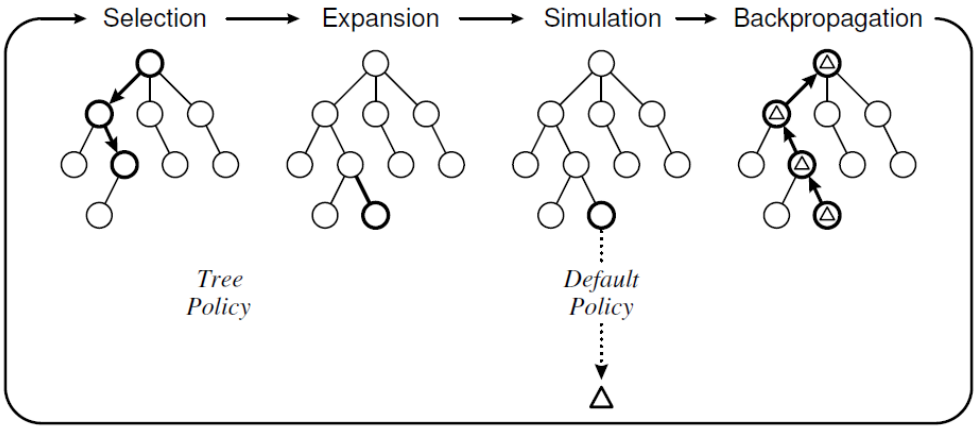


Figure 2.2: MCTS search policy steps.

1. Selection: Starting at the root node, a child selection policy is recursively applied to descend through the tree until the most urgent node is reached. For every stage, the child is randomly selected among the list of available children.

2. Expansion: One or more child nodes are added to expand the tree, according to the available actions.

3. Simulation: A simulation is run from the new node(s) to produce an outcome (determining the goals).

4. Backpropagation: The simulation result is "backed up" through the already selected nodes to update their statistics (number of visits and goals).

2.2.6 Reinforcement Learning (RL)

The interaction between the agent and the environment includes a sequence of actions, state information, and rewards for which the agent is trying to choose optimal actions to maximize the cumulative rewards. A RL agent attempts to learn an optimal strategy from the accumulated trials and the received feedback, and with the optimal strategy, the agent would have successfully adapted to

the environment, meaning the future rewards would be maximized. For RL, the agent would be in one of any number of states ($s \in S$), and will take one action out of a set of possible actions ($a \in A$) to switch from one state to another. The outcome is decided by the transition probabilities between the states ($\mathbf{p}$). Once an action is taken, the reward ($r \in R$) is received from the environment as a feedback. After efficiently learning the optimal solution, the agent's action policy ($\pi(s)$) effectively becomes the agent's optimal solution for which action to take. Each state in the environment is associated with a value function ($V(s)$) which represents the expected future rewards, and quantifies how good the state is. The interaction sequence can be described as [15, 16]

$$S_1, A_1, R_2, S2, A2, .., S_T$$

Transition and Reward. The environment can be described as a model, and this model has two main bases, the transition probability function $\mathbf{p}$ and the reward function R. An example of how this works would be: an agent being in state s and decides to take action a to go to the next state s' while receiving a reward r. This is defined as one episode, and expressed as (s, a, r, s').

The transition probability function $\mathbf{p}$ determines the probability of transitioning from state s to s' and obtaining reward r after taking action a.

$$\mathbf{p}(s', r|s, a) = \mathbb{P}[S_{t+1} = s', R_{t+1} = r|S_T = s, A_t = a],$$

where $\mathbb{P}$ is the probability symbol. Hence, the state-transition function is defined as a function of $\mathbf{p}(s', r|s, a)$,

$$\mathbf{p}^a_{ss'} = \mathbb{P}[S_{t+1} = s'|S_t = s, A_t = a] = \sum_{r \in R}, \mathbf{p}(s', r|s, a).$$

as well as the reward function R predicts the next reward and occurred by one action:

$$R_{(s,a)} = \mathbb{E}[R_{t+1}|S_t = s, At = a] = \sum_{s \in S} \mathbf{p}(s', r|s, a).$$

Policy is the function of the agent's behavior (π), trying to select the optimal action in state s. It is a mapping from state s to action a and can be either of

- deterministic $\pi(s) = a$, or
- stochastic $\pi(a|s) = \mathbb{P}_\pi[A = a|S = s]$.

Value Function of a state s is the expected return of the current state at time t,

$S_t = s$

$$V\pi(s) = \mathbb{E}_\pi[G_t|S_t = s]$$

Similarly, the action-value (quality) can be defined as

$$Q_\pi(s,a) = \mathbb{E}_\pi[G_t|S_t = s, A_t = a]$$

Optimal Value and Policy. The optimal value function produces the maximum return

$$V_*(s) = max_\pi V_\pi(s), \qquad Q_*(s,a) = max_\pi Q_\pi(s,a)$$

Finally, the optimal policy can be achieved by

$$\pi_* = argmax_\pi V_\pi(s), \qquad \pi_* = argmax_\pi Q_\pi(s,a)$$

Temporal-Difference (TD) Learning refers to the method of updating the value function $V(S_t)$ towards an estimated return $R_{t+1} + \gamma V_(S_{t+1})$ (known as the "TD target"). The extent to which the value function is updated is controlled by the learning rate hyper-parameter α. For On-Policy learning, it would be

$$Q(s_t,a_t) \leftarrow Q(s_t,a_t) + \alpha\Big[r_{t+1} + \gamma Q_*(s_{t+}, a_{t+1}) - Q(s_t,a_t)\Big], \tag{2.1}$$

whereas the Off-policy TD is

$$Q(s_t,a_t) \leftarrow Q(s_t,a_t) + \alpha\Big[r_{t+1} + \gamma max_{a\in A} Q_*(s_{t+}, a_{t+1}) - Q(s_t,a_t)\Big], \tag{2.2}$$

where γ is a discount factor that denotes how much impact the current decision can have on the long-term reward. Particularly, a smaller value of γ emphasizes more the immediate rewards and a larger γ gives higher weight to the future rewards. When γ equals 0, the game becomes a one-time-event game.

Exploration-Exploitation. When RL interacts with an unknown environment, and without enough exploration, the agent cannot perform well. However, without enough exploitation, the reward optimization task cannot be completed. Various RL algorithms are trying to compromise between the exploration and exploitation in different ways. In Q-learning, and many other on-policy algorithms, the exploration is commonly associated with the greedy policy.

Chapter 3

Assessment of Photovoltaic Hosting Capacity of Existing Distribution Circuits

The integration of the distributed energy resources (DERs) into the distribution system has been rapidly increasing over the past decade, with photovoltaics (PV) leading the way as the fastest growing renewable energy resource in North America [49]. This development is great for sustainability, but problematic for system operation. Since the current distribution systems are designed for centralized power generation with a unidirectional power flow, a significant level of PV penetration can potentially affect the voltage stability (overvoltage, voltage deviation, and voltage unbalance), and line thermal stress [50]. Moreover, uncertainty in PV power can lead to increases of switching in capacitor banks and regulator tap changes [51] thus leading to more system losses and negatively affecting the operational life expectancy of distribution grid equipment. In order to mitigate these problems early, it is necessary to assess PV hosting capacity (PVHC) of the distribution circuit. PVHC can be defined as the maximum PV generation capacity that can be integrated into an existing distribution network without violating the normal system operating conditions.

There have been several studies conducted to assess different aspects of PVHC in distribution networks. Tonkoski et al. studied the sensitivity of PVHC to feeder impedance [21]. The results showed that the bus voltages increase with reduction of the X/R ratio on network lines. Coogan et al. examined the location dependency of PVHC [22]. They found that the bus voltages are proportional to the distance between the feeder and bus locations. A streamlined technique to

investigate PV interconnections in commercial areas of LV distribution networks was proposed by Rylander et al. [23]. A number of other studies concentrated on hosting capacities in residential areas [52–56].

This contribution presents the results of PVHC determination for three types of distribution networks, using data of three communities that have the potential to adopt PV systems on a large scale. Analysis of PVHC was conducted using probabilistic power Flow (PPF) [57,58]. PPF is an essential tool for power system analysis and planning in situations when some systems parameters are uncertain and can be considered random variables. In this contribution, PPF is implemented through Monte Carlo simulations (MCS) [59] that use repetitive solutions of deterministic power flow with different realizations of random parameters to obtain expected probability distributions of variables of interest. The probability distributions of bus voltages are combined with a predetermined threshold to identify the numbers and locations of buses affected for selected PV penetration levels and load conditions used in PF simulation. Because PVHC can be affected by other factors such as individual installation sizes, locations, system reactive power, and the presence of regulation devices, sensitivity of HC to these factors are also considered. Finally, several recommendations for distribution system operators (DSOs) are made with respect to PV integration.

This research is organized in six sections. Section II defines PV hosting capacity and identifies factors that need to be considered in PVHC studies. Circuits assessed in this study are described in section III. The proposed methodology for calculating probabilistic power flow is outlined in section IV, while section V describes the results of its application to selected circuits. The concluding section VI summarizes the obtained result and lessons learned.

3.1 PV Hosting Capacity

PVHC can be defined as the maximum amount of PV capacity that can be integrated into a distribution network without violating any operating thresholds. Meanwhile, PV penetration can be defined as the peak power of all PV generators connected to a power system, represented as a percentage of the system's annual peak load [60]. In this section, the hosting capacity is determined with respect to overvoltage. A simplified feeder model is shown in Figure. 3.1. An equivalent Thevenin impedance of $Z_{th} = R_{th} + jX_{th}$, is used to determine the overvoltage at the point of common coupling (PCC), where the PV unit is connected. The voltage at the PCC can be obtained using methods described

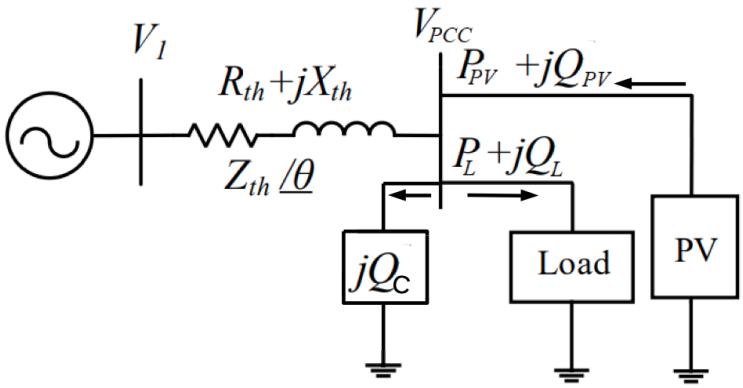


Figure 3.1: Simplified distribution feeder with PV.

in [61, 62]

$$V_{\text{PCC}} \approx V_1 + \frac{P^{'}R_{th} + Q^{'}X_{th}}{V_{\text{PCC}}}, \tag{3.1}$$

where $P^{'} = P_{PV} - P_L$, similarly $Q^{'} = Q_{PV} - Q_L$.

Alternatively, the equations for finding the maximum active power that can be injected by PV unit at the PCC, which results in a maximum voltage at this point, can be derived as follows

$$P^{'} = \frac{V_{\text{PCC}}^{\max}(V_{\text{PCC}}^{\max} - V_1)}{R_{th} + \tan\phi X_{th}}. \tag{3.2}$$

Multiplying and dividing the denominator of equation (3.2) by Z_{th}, the final expression can be written as

$$P_{PV}^{\max} = \frac{V_{\text{PCC}}^{\max}(V_{\text{PCC}}^{\max} - V_1)}{Z_{th}(\cos\theta + \tan\phi\sin\theta)} + P_L, \tag{3.3}$$

where θ is the phase angle of Thevenin impedance, and ϕ is the receiving end power factor.

Several important factors need to be taken into consideration to study the sensitivity of PVHC:

a) Reactive power absorption: Should the reactive power at the PCC be increased by adding Q_C, the term $(-Q^{'}X_{th}/V_{PCC})$ in (3.1) would increase to $(-Q^{''}X_{th}/V_{PCC})$ and cause a reduction in V_{PCC} ($Q^{''} = Q_L + Q_C$ assuming $Q_{PV} = 0$). Therefore, an increase in reactive power could improve PVHC.

However, it would also significantly increase system losses.

b) Network load condition: PVHC linearly increases with network load (3.3). This shows that a higher PVHC can be obtained at the PCC with a higher load demand.

c) Equivalent Thevenin impedance: PVHC is inversely proportional to the equivalent Thevenin impedance at the PCC (3.3). Based on this observation, PVHC can be maximized by reconfiguring the distribution circuit such that the Thevenin impedances are minimized at the corresponding buses. However, this solution invariably involves conductor upgrades, which may be expensive.

d) Voltage regulation can be implemented using voltage regulators or transformer tap changers. These devices can help to maximize PV installations provided that they are placed at optimal locations. However, these regulation devices are not fast enough to reduce transient overvoltages or fast power variations due to rapid insolation changes.

3.2 Circuit description

This study reports on three circuits (Fort McMurray, Drumheller, and Lloydminster). They are the models of three urban networks with different topologies, load conditions, customer types, and populations. The three circuits can be characterizations as follows:

1. Fort McMurray is a 25kV circuit supplied by eight feeders, with a load of 156,726.37 kW and 76,570.24 kVAR. Light load was determined to be 40%.

2. Drumheller is a 25kV distribution circuit supplied by two feeders, with a load of 16,193.7 kW and 6,255.33 kVAR. Light load was determined to be 50%.

3. Lloydminster is a 25kV distribution circuit supplied by four feeders, with a full load of 73,480.28 kW and 27,094.08 kVAR. Light load was determined to be 50%.

3.3 Methodology: MCS-based PPF

The methodology used to determine PVHC can be described using the following four steps:

1. A base model of distribution circuit contains no PVs. Light load is determined by the dominant type of customer (residential or commercial). In most cases, light load is considered to be 50% of the peak load. A size of a single PV panel is also set based on the type of customer.

2. PV panels are randomly placed onto the base model until the desired penetration level is reached (see Figure 3.2). Power flow of the model with PV generators is calculated and bus voltages are recorded. This step is repeated a number of times (depending on the circuit size) to obtain enough data for calculating a voltage probability distribution for each bus. In each repetition, a new random distribution of locations and sizes of PV generators is used.

3. Starting from 0% penetration, step 2 is repeated in 10% increments. Stop after a predefined voltage threshold is attained.

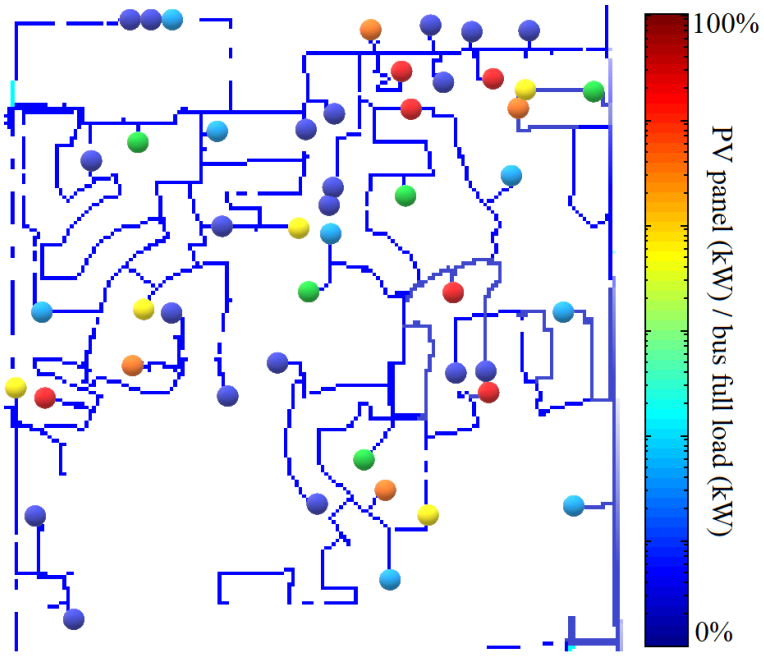


Figure 3.2: An example scenario under the stochastic PV deployment framework.

4. The overall impact of PV penetration on the circuit is assessed and PV hosting capacity evaluated.

3.4 Results and discussion

Power flow was calculated using CYME power engineering software. Unity power factor was assumed. A bus voltage of 1.04 p.u is considered impacted (represented in orange), while voltage of 1.05 p.u is considered critical (represented in red). Figures 3.3, 3.4, and 3.5 show the critical and impacted regions of each circuit, at 70% PV penetration for all circuits. Voltage distribution at the most affected buses is shown using box plots in Figure 3.6 for PV penetration levels as noted, corresponding to the first occurrence of voltage violation for given circuit.

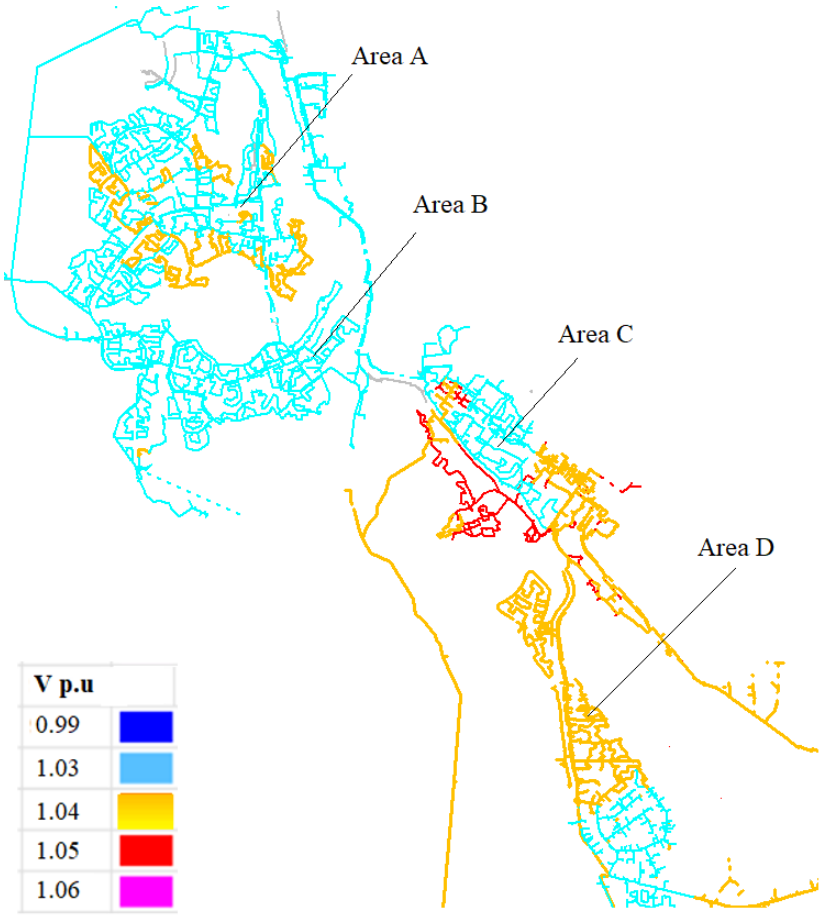


Figure 3.3: Fort McMurray circuit diagrams based on the voltage impact.

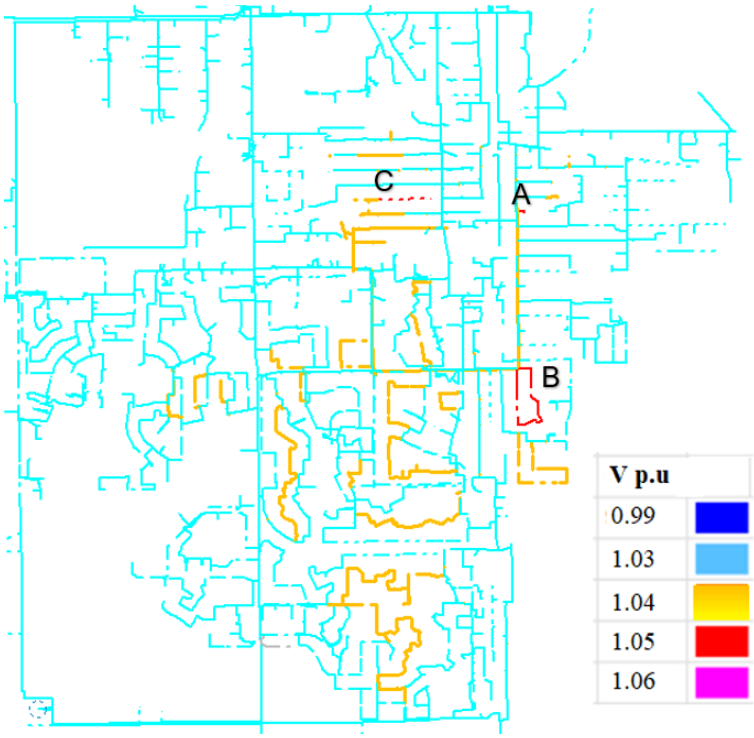
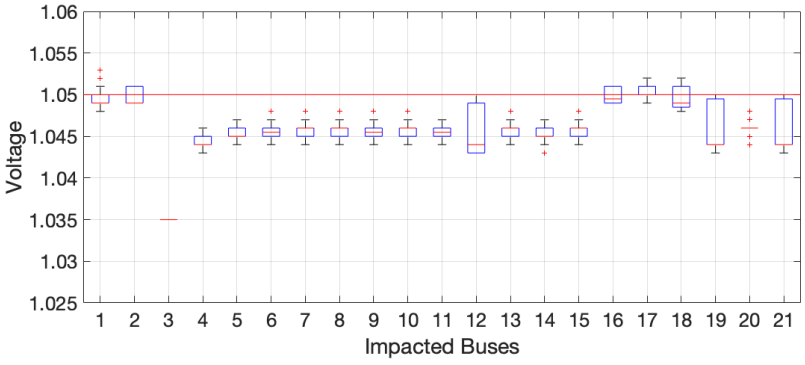


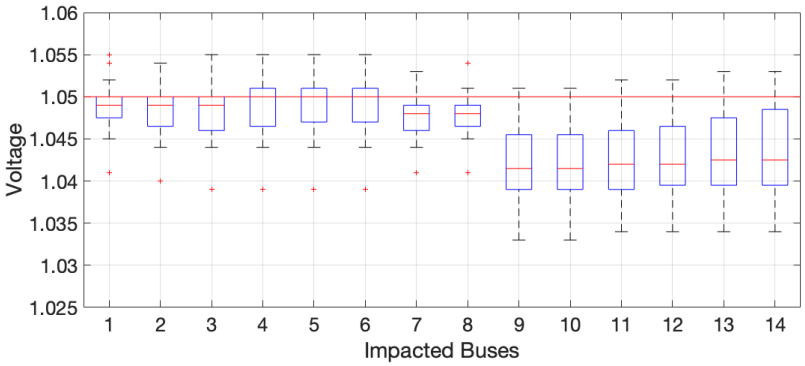
Figure 3.4: Lloydminster circuit diagrams based on the voltage impact.

Figure 3.5: Drumheller circuit diagrams based on the voltage impact.

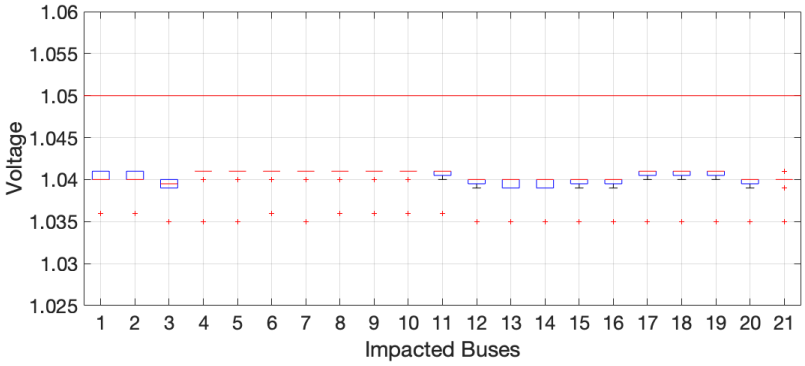
Due to the complexity of *Fort McMurray* circuit, it has been divided into four areas for clarity of analysis and presentation. At 10% penetration, areas 3 and 4 have many critical and impacted regions, while areas 1 and 2 only became problematic at around 40% to 50%. The primary reason that this circuit is affected



(a)



(b)



(c)

Figure 3.6: Voltage violation vs the impacted buses at: a) Fort McMurray, 3rd area (PV 10%). b) Lloydminster (PV 60%). c) Lloydminster (PV 60%).

by low levels of PV penetration is due to the circuit load condition. The highly loaded buses suffer more voltage increase at the PCC since the output current of the PV system is proportional to the loading current of these buses. Although some areas can accommodate higher PV penetration, the HC of the entire circuit

must be taken as 10%.

Lloydminster circuit can be considered moderate in terms of a load demand and complexity. The hosting capacity for this circuit has been determined at 60% PV penetration, when the first voltage violation occurred. The locations of the critical buses are highlighted in Figure. 3.4 by letters A, B, and C.

Drumheller's population and the load demand are significantly lower than those of Fort McMurray and Lloydminster. As a result, the hosting capacity of this region is higher than in the other two cases. PPF calculations were stopped at 70% PV penetration when voltage reached level 1.04 p.u. If 1.05 p.u. was considered, the hosting capacity would have been even greater.

Figures 3.7 and 3.8 show the distribution of voltage levels of the three circuits by using probability density function (PDF) and a cumulative distribution function (CDF). The mean voltages for Fort McMurray and Lloydminster circuits are the same at 1.0461 p.u., while for Drumheller the mean is 1.0401 p.u.

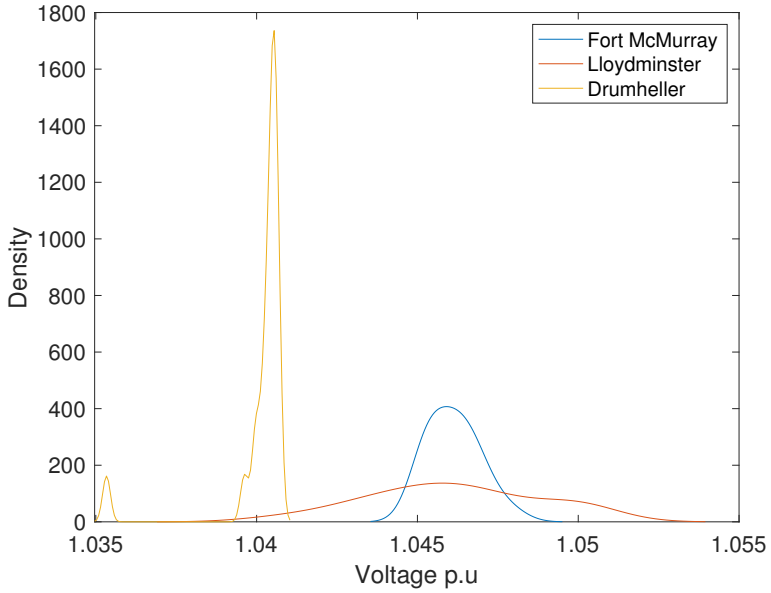


Figure 3.7: PDF of PVHC of the three circuits: Fort McMurray the 3rd area (PV 10%), Lloydminster (PV 60%), and Drumheller (PV 70%).

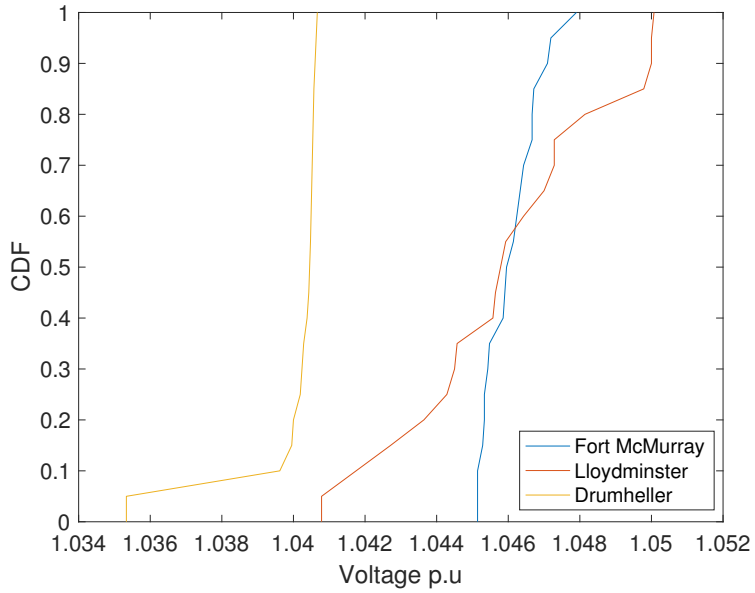


Figure 3.8: CDF (b) of PVHC of the three circuits: Fort McMurray the 3rd area (PV 10%), Lloydminster (PV 60%), and Drumheller (PV 70%).

3.5 Final Remarks

In this contribution, the hosting capacities of three real communities were determined using MCS-based PPF. The presented study examined three circuits of varying size and complexity using the developed MCS-based PPF method. It determined the hosting capacities of Fort McMurray, Lloydminster, and Drumheller circuits as 10%, 60%, and 70%, respectively.

The developed methodology is an effective approach for determining PV hosting capacity that can be used for any distribution circuit model. The use of stochastic approach overcomes shortcomings of deterministic methods and can, in principle, handle circuits of any complexity.

Chapter 4

Reinforcement Learning based Distributed BESS Management for Mitigating Overvoltage Issues in Systems with High PV Penetration

Integration of household-scale photovoltaic (PV) generators and other distributed energy resources (DER) into the existing low voltage (LV) distribution grids around the world has rapidly increased. Many government incentive programs encourage households and businesses to install small-scale, roof-mounted PV panels, driving consumer-led evolution of modern electricity supply systems. Studies have shown that high PV penetration can lead to grid voltage level problems. The main concern is the risk of overvoltage due to reverse power flow in distribution feeders, especially at light loads. To avoid integration-related problems, many utilities limit PV penetration levels. However, this passive approach leads to loss of significant amount of potential PV generation [21,22]. By using a stochastic analysis framework on the simulated PV deployment scenarios, various characteristics can be determined such as the hosting capacity (i.e. the PV penetration limits) or the location of the susceptible buses to overvoltage. The hosting capacity depends not only on the total PV penetration level, but also on the sizes and locations of PV systems installed in residential areas. Additional PV deployment factors that influence the hosting capacity include circuit char-

acteristics and loading conditions [22, 24]. Has been shown that feeder voltage increases with increasing ratio $K = R/X_L$, where R is the line resistance and X_L is its impedance [21].

There have been many solutions proposed to mitigate the voltage rise problems associated with high levels of PV penetration. The reactive power absorption has been proposed in [6]. Although this solution resolves overvoltage issues, it also leads to reduction of the feeder power factor. The use of smart inverters to control local Volt/VAR on the PV connected buses offers a decentralized solution [7]. It is more effective in resolving the overvoltage issue with lower line losses, but at a higher cost. The use of medium voltage to low voltage (MV/LV) transformers to limit the overvoltage has been investigated in [25, 26].

The use of BESS collocated with PV generators can boost local consumption during the periods of light load. In addition, BESS can be used to support generation during the peak load periods. It has been shown that BESS-based approaches can considerably increase the hosting capacity [9, 10]. However, their widespread implementation has been impeded by their relatively high cost. As a result, most proposed approaches combine BESS with PV inverters to absorb part of the active power caused by high PV penetration. The limited size of BESS used in PV applications may result in a full charge or discharge of the batteries during the control process. In addition to the limited BESS capacity, other factors such as initial state of charge, load conditions, and weather-dependent PV energy production may contribute to the shortage of the available energy storage. To address these issues, it has been proposed to use available storage capacity of neighboring BESS units [11–14, 19]. This way, the charge/discharge power in small radial distribution networks can be maximized using common battery scheduling. However, the operation strategies of these systems have been programmed in advance, based on complete prior knowledge of the environment. However, in many cases, the states of the environment are not know and/or may change over time. This affects the applicability and flexibility of the implemented strategies and, in some cases, even their ability to converge. For instance, these methods consider the SOC of the neighbouring bus BESS units, but they do not consider the impact of other factors such as the line losses and voltage levels. These may change over time and lead to growth of system uncertainty and, subsequently, to decrease of reliability. Therefore, it is crucial to design system that can learn the strategies on its own and with limited prior knowledge of the environment. A RL system, unlike systems with supervised learning, operates as an adaptive agent that learns by trial-and-error without an explicit teacher.

A strategy learned by RL is evaluated by the Q-value function that maps each state-action pair to an estimate the reward of the new action using a transition probability. The purpose of MCTS is to build a feasible environment for the RL agent to reduce the high computational burden that would occur if the entire network were considered. It also facilitates storing of the updates of the RL states at particular nodes through the tree search configuration of the MCTS navigation method. Eventually, it considers multiple factors/system states including the sufficiency of the BESS capacities, bus voltage levels and line losses.

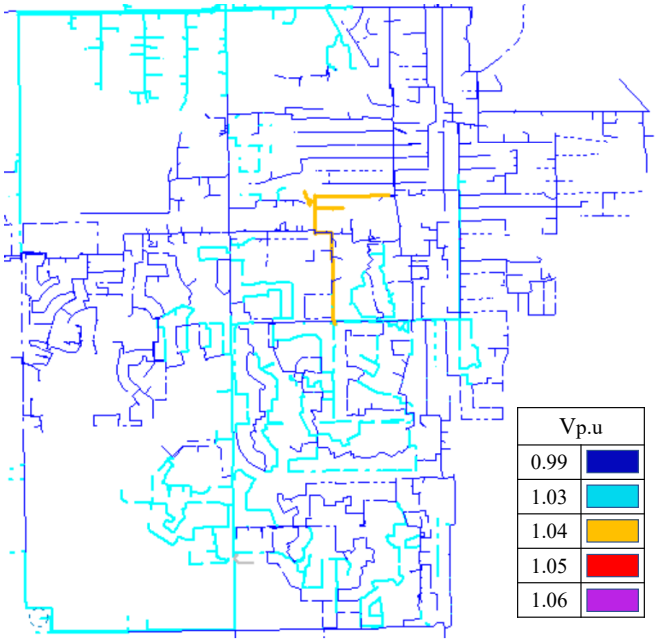
Finally, because of the non-trivial nature of the system cost function, the battery operation must be scheduled using a sound optimization approach and the overall system control strategy must be modified accordingly.

The new approach proposed in this research, termed *cooperative state of charge control* (CSOCC), is well suited for large and complex distribution networks. It divides a network into multiple smaller segments based on the configuration of so called *impacted regions* (i.e. regions negatively affected by voltage raise problems due to the high penetration of PV generators). This approach relies on the intelligent technique of Monte-Carlo tree search-based reinforcement learning (MCTS-RL) and on model predictive control (MPC), to resolve the PV-induced overvoltage problem. MCTS-RL is used as a centralized controller that optimizes the cooperating buses in each of the partitioned network structures in terms of BESS availability, and line losses that are subjected to the voltage level. MPC is used as a decentralized controller to mitigate the overvoltage by estimating and controlling the battery state of charge (SOC). Battery scheduling is optimized not only to mitigate the PV-induced overvoltage problems, but also to minimize energy consumption. To demonstrate the operation of CSOCC and evaluate its efficiency, the proposed approach has been implemented on a model of the distribution system of Lloydminster, Alberta Canada.

This research is organized as follows. Section II presents the system model description of the Lloydminster distribution circuit and the formulation of the economic dispatch problem. The proposed cooperative state of charge control algorithm, RL-MCTS algorithm, and the system modules are presented in Section III. Section IV provides a summary of the proposed algorithm. Simulation results are presented and discussed in Section V, followed by conclusions in Section VI.

4.1 Case description

In order to understand the performance of the proposed method the simulations are conducted on two distribution feeders. The first is the 33 bus distribution feeder [63]. The second, which is the practical case; the distribution feeder of the Lloydiminster city. Lloydiminster, is served by a 25-kV distribution circuit supplied by four substations. The total generation at full load is 75 MW and 23 MVAR with the total load equal to 73 MW and 27 MVAR. The distribution circuit with no PV generator installed, used as a base case, is shown in Figure 4.1. This colored schematic diagram shows relative voltage levels corresponding to the full load for the circuit. Load flow analysis of the circuit under different PV penetration levels has been simulated using power engineering software CYME [64] and will be shown in section. V.



(a)

Figure 4.1: A circuit diagram of the Lloydminster distribution system showing voltage levels at the full load.

PV penetration is defined as the total peak power of all installed PV generators expressed as a percentage of the annual peak load of the same system [60]. PV hosting capacity is then the maximum level of PV penetration that can be integrated into the system without violating a predefined performance index (e.g., the voltage level) [65]. The impact of PV systems on a distribution circuit can

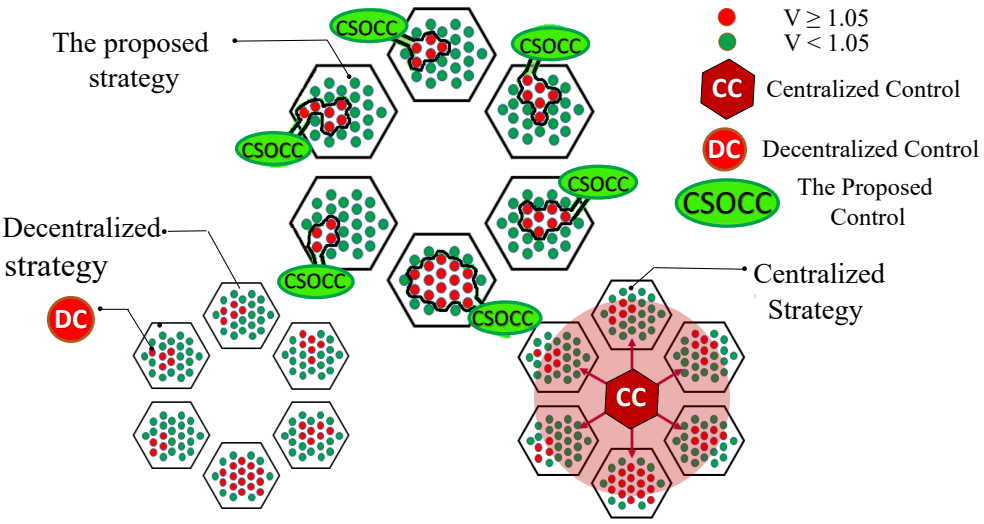


Figure 4.2: Conceptual comparison of the proposed distributed control strategy against the conventional centralized and decentralized strategies.

be determined using a probabilistic deployment framework that simulates various PV deployment scenarios and examines the impact of their variation on the variable(s) of interest. The outcome of this stochastic modelling is then used to determine the PV penetration level defined as the ratio of the PV panel rated power (kW) to the value of the full-load (kW) at the connected bus [24].

4.2 cooperative state of charge control

The proposed approach to increase PV hosting capacity is based on a combination of centralized and decentralized control strategies. In decentralized control, the monitoring and control of charging and discharging actions are performed locally by each BESS. The main advantage of this approach is that it does not require extra communication systems. This makes decentralized approach robust and cost-effective. However, when a BESS unit either is fully charged or can only provide a limited power, it cannot communicate with other BESS units to request support. The unit may even experience a total failure resulting from the lack of communication with other units [39,40]. In centralized control, a central controller aggregates measurements from the entire network and determines the charging/discharging control tasks for each BESS. This approach is more efficient than the decentralized control, as it is based on the current information about the entire network. In the event of unavailability of particular BESS unit,

the central controller can search for an alternative unit and resolve the problem. However, centralized control demands fast communications infrastructure which results in higher computational burden and overall costs. Furthermore, in case of communication failure, a BESS units may not respond to the central controller [37,38], or the entire system may collapse.

The proposed CSOCC strategy forms multiple distributed sections within an entire network [66, 67]. Integrating the advantages of the centralized and decentralized strategies, this approach is highly robust and tolerant to disconnections of network components. The robustness is achieved through the node feedback responses of the multiple distributed network sections, while the connection/disconnection tolerance comes directly from the presence of multiple smaller sections rather than the entire network. Similar to the decentralized control method, the proposed strategy takes into account the local measurements of the BESS capacity and voltage sensitivity (VS) for each node to control each BESS unit. At the same time, communication with the central controller offers an opportunity to optimize BESS charging/discharging to not only mitigate the overvoltage problems, but also to minimize the overall costs of system operation.

BESS can be used to mitigate voltage rise under high penetration of rooftop PV systems in LV distribution networks. However, technical limitations of BESS systems, such as their finite capacity, may prevent them from maintaining sufficient amounts of energy throughout the day. Thus, participation of BESS units from different regions in the network is necessary to correct the voltage profile, to maximize the utilization of available BESS capacity in the network, and to prevent its premature depletion or saturation. The proposed CSOCC approach addresses coordination of distributed BESS units for an overall voltage regulation in distribution networks [68,69]. The proposed approach identifies an assisting region (with normal voltages) surrounding an impacted region (with voltage limit violation) using MCTS-RL. It acts as a centralized consumption planner and controller for choosing the optimal BESS buses in the assisting region. In the decentralized controller, the MPC controls the SOC of each BESS bus to accommodate the daily system operation in both assisting and impacted regions. Figure 4.3a shows the discretized SOC curve of the impacted region without the use of the CSOCC. In this figure, $K - \Delta t$, K, and $K + \Delta t$ are, respectively, the time instants when past, present and predicted (future) states are sampled. The future SOC state in the impacted region is estimated using the MPC controller. Consequently, the amount of energy required from BESS by the impacted region can also be estimated. This energy can be obtained from the assisting regions

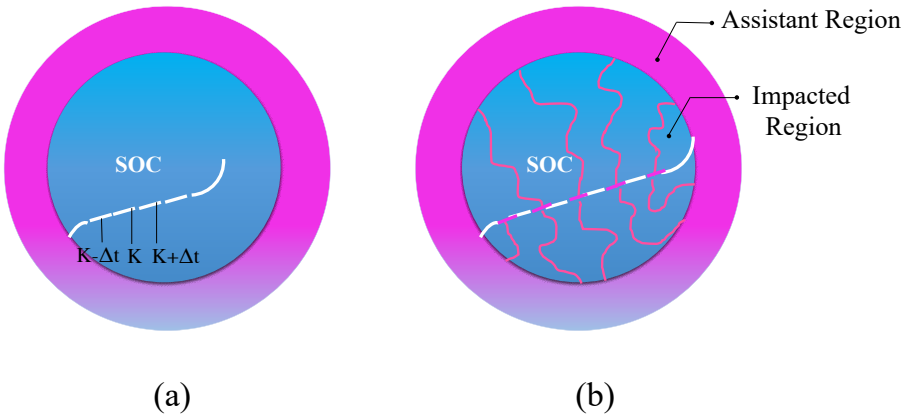


Figure 4.3: Principle of the proposed CSOCC approach: BESS in the assisting regions effectively stretch the available storage capacity to avoid its premature saturation or depletion.

using the centralized controller. The energy transferred from other regions is represented by the wavy lines in Figure 4.3b. Eventually, the MPC-driven energy transfers will extend the SOC curve of the impacted region to fit the required time period and prevents premature saturation of the BESS. The main task of CSOCC is to select the optimal BESS units using MCTS-RL. The selected units work as a coherent group and are able to effectively share their capacity. This strategy functions properly even in situations when the impacted buses have BESS units with insufficient capacities.

4.2.1 MCTS-RL

Methods of machine learning and artificial intelligence (AI) offer many valuable tools to address a variety of issues in dynamic and complex networks of contemporary power systems. Using machine learning techniques, a control system can learn without being explicitly programmed. Among machine learning mechanisms, Monte Carlo tree search (MCTS) based reinforcement learning (RL) [48] [15, 16] is particularly suitable for solving sequential decision-making problems. RL provides an agent with the ability to learn the state variations and to find potential solutions. By interacting with the environment, RL agents gain powerful experience for sequential decision making under uncertainty. The MCTS-RL algorithm aims to find desirable resource diffusion strategies. For

example, in application described in this research, this algorithm can identify the best available BESS buses in the assisting region and optimal power transfer paths. It navigates through the network and gradually builds up its experience (i.e., learns from the results) to further optimize its own decisions in an unsupervised fashion. This approach uses the temporal difference RL methods such as SARSA [17]. The accumulated experience is the result of biased random sampling in the decision space during the policy optimization and exploration process. Another advantage of the MCTS, compared to conventional exhaustive search methods, is its narrow area of exploration obtained using the subset search method. In RL, an agent is in a state s from a set of possible states S, and takes an action a out of a set of possible actions A. It moves between states according to transition probabilities p. Once an action is taken, the agent receives a reward $r \in R$ from the environment. In MCTS, each node (representing a network bus) contains state and action edges (s, a) of a tree. In addition, each edge stores a set of statistical parameters $\{N(s,a), r(s,a), Q(s,a)\}$, where $N(s,a)$ is the visit counter (initialized at zero), $r(s,a)$ is the instant reward, and $Q(s,a)$ is the action-value pair obtained from the value network. The value of Q is updated by Q-learning. Q-learning is an algorithm to learn a policy for selecting actions an agent can take within its environment. It is a model-free learning approach that can handle problems under stochastic conditions. The learned action-value function, Q, directly approximates the optimal action-value function, Q_*, of the policy being followed [17]. It can be described using the following equation:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \Big[r_{t+1} + \gamma Q_*(s_{t+}, a_{t+1}) - Q(s_t, a_t) \Big], \tag{4.1}$$

where α is a learning rate hyper-parameter that controls the extent to which the value function is updated, and γ is a discount factor that denotes the impact of the current decision on the long-term reward. A lower value of γ results in more immediate rewards, while a higher value of γ gives a higher weight to the future rewards. Discount factor $\gamma = 0$ corresponds to a one-time-event.

The tree search state is randomly initialized. When the network state changes to s', agent n selects an action a. The Q-value of the agent is then updated according to the preliminary Q-function, and the number of visits is incremented by 1 for the visited nodes. The learning process ends either when the algorithm converges, or when the depth limit is reached. The proposed method is illustrated in Figure 4.4. The MCTS-RL model receives the state from the environment, transforms it into the transition probability, computes $Q(s,a)$ values, and selects the

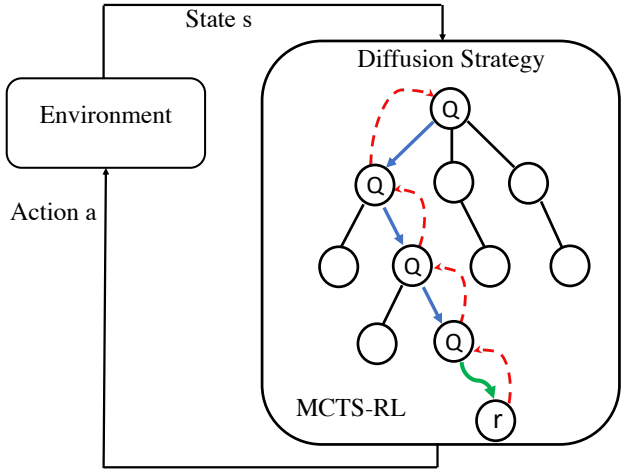


Figure 4.4: The proposed algorithm based on MCTS-RL method.

diffusion strategy according to a greedy policy. In this policy, each strategy corresponds to a different set of actions A upon which RL will be executed in the search tree. In a nutshell, MCTS constructs its search tree, generates many DG allocation strategies, and through the back propagation process, it returns the Q-values for each state after the RL has executed the actions and generated the transition probabilities.

4.2.2 Battery modeling

The dynamic behavior of the BESS unit can be modeled using the following discrete-time equation

$$SOC(k+1) = SOC(k) + \frac{T_s}{B_s}\left(\eta_{ch,} P_{ch}(k) - \eta_{dis} P_{dis}(k)\right), \tag{4.2}$$

where SOC is the state variable (state of charge), η_{ch}, η_{dis} are the charging and discharging efficiencies (respectively), T_s is the sampling time, B_s is the battery capacity, $P_{ch}(k)$ and $P_{dis}(k)$ are the charging (withdrawing) and discharging (injecting) power at each bus and time k. The variance of SOC is proportional to the charging/discharging current [70].

The SOC of the battery bank is subject to the upper and lower boundaries [71]

$$SOC_{min} \leqslant SOC(k) \leqslant SOC_{max}. \tag{4.3}$$

The boundary constraints for the charging and discharging power can be rep-

resented as

$$0 \leqslant P_{ch}(k) \leqslant P_{max}, \tag{4.4}$$

$$0 \leqslant P_{dis}(k) \leqslant P_{max}, \tag{4.5}$$

where P_{max} is the rated charging or discharging power of the battery.

Finally, the limitation on the simultaneous charging and discharging of the BESS is shown by

$$P_{ch}(k) \cdot P_{dis}(k) = 0. \tag{4.6}$$

4.2.3 MPC Design

This section develops MPC controller design procedure [72] that implements the proposed method. The dynamic behavior of SOC (4.2) can be simplified in terms of the state space model

$$SOC(k+1) = SOC(k) + B_b(k) \cdot u_b(k) \tag{4.7}$$

The estimated parameters of the next state of charge $SOC(k+1)$ are $B_b(k) = [0, \eta_{ch}, -\eta_{dis}]$, and $u_b(k) = [0, P_{ch}(k), P_{dis}(k)]^T$. The control variable u_b can be derived as follows

$$\widehat{u_b}(k) = [0, \big(P_{ch,i}(k) + P_{ch,a}(k)\big), \big(P_{dis,i}(k) + P_{dis,a}(k)\big)]^T \tag{4.8}$$

where the subscripts i and a designate battery systems in the impacted and assisting region, respectively. The final expression can be written as

$$SOC(k+1) = SOC(k) + B_b(k) \cdot \widehat{u}_b(k). \tag{4.9}$$

The error between two samples is defined as the cost function to be minimized

$$J_e = SOC(k+1) - \big(SOC(k) + B_b(k) \cdot \widehat{u}_b(k)\big). \tag{4.10}$$

4.2.3.1 MIMO System Modeling

To facilitate MPC design, the MIMO system (i.e., Multi-Input-Multi-Output) has to be transformed into a state-space model [73]. The energy balance within the

system is maintained thorugh the following equality constraint

$$P_g(k) = P_L(k) - \big(P_{PV}(k) + P_{Batt}(k)\big), \tag{4.11}$$

where $P_{Batt}(k) = -P_{ch}(k) + P_{dis}(k)$, and P_g and P_L are the grid power and the load demand power respectively.

Considering grid energy cost C_1 and battery energy cost C_2 [74], the goal of the MPC assignment can be expressed as minimization of the following two variables [75]

$$y_g(k) = C_1(k)P_g(k), \tag{4.12}$$

$$y_{Batt}(k) = C_2(k)\big(-P_{ch}(k) + P_{dis}(k)\big). \tag{4.13}$$

The augmented system state can be expressed as

$$x(k+1) = [SOC(k+1), y_g(k), y_{Batt}(k)]^T, \tag{4.14}$$

and transformed into state-space model [76]

$$x(k+1) = Ax(k) + B \cdot \widehat{u_b}(k) \tag{4.15}$$

to calculate the output

$$y(k) = Cx(k) \tag{4.16}$$

where the model matrices are

$$A = \begin{bmatrix} 1 & 0_{1\times 2} \\ 0_{2\times 1} & 0_{2\times 2} \end{bmatrix}, B = \begin{bmatrix} 0 & \eta_{ch} & -\eta_{dis} \\ C_1 & 0 & 0 \\ 0 & -C_2 & C_2 \end{bmatrix},$$

$$C = \begin{bmatrix} 0_{2\times 1} & I_{2\times 2} \end{bmatrix}.$$

4.2.3.2 MPC Objective Functions

In addition to determining CSOC control signal, MPC is also capable of optimizing the overall cost through minimizing the grid cost. This is achieved by defining the MPC objective functions used for minimizing the purchase period and battery time-of-use (TOU) cost as shown below

$$J_g(k) = \min \sum_{k}^{k+N_p} C_1(k) P_g(k), \tag{4.17}$$

and

$$J_{Batt}(k) = \min \sum_{k}^{k+N_p} C_2(k) P_{Batt}(k), \tag{4.18}$$

where the battery TOU cost is

$$C_2(k) P_{Batt}(k) = C_{purch}(k) * P_{ch}(k) * T_s - C_{sell}(k) * P_{dis}(k) * T_s, \tag{4.19}$$

where C_{purch} and C_{sell} are the grid purchase and selling price, respectively.

4.2.3.3 MPC Constraints

In this system, MPC is modelled with two major types of constraints applied to the output and control signal, respectively. From the dynamic equation of BESS (3), predicted values of $SOC(k+1)$ can be calculated as follows [75]:

$$\begin{bmatrix} SOC(k+1) \\ SOC(k+2) \\ \vdots \\ SOC(k+Nc) \end{bmatrix} = \begin{bmatrix} SOC(k) \\ SOC(k+1) \\ \vdots \\ SOC(k+Nc-1) \end{bmatrix} + \begin{bmatrix} B_b \end{bmatrix} \cdot \begin{bmatrix} \widehat{u}_b(k) \\ \widehat{u}_b(k+1) \\ \vdots \\ \widehat{u}_b(k+Nc-1) \end{bmatrix} \tag{4.20}$$

where B_b dimension is $(Nc-1) \times (Nc-1)$.

Maximum and minimum constraints are applied to all future values of control signals, also called manipulated variables. They can be expressed in matrix form as follows.

$$\begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \\ \cdots & \cdots & \cdots \\ 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix} \widehat{u}_b(k) \leq \begin{bmatrix} P_{max}(k) \\ \vdots \\ P_{max}(k+Nc-1) \\ \cdots \\ P_{min}(k) \\ \vdots \\ P_{min}(k+Nc-1) \end{bmatrix} \tag{4.21}$$

where, P_{max} and P_{min} are column vectors with $N_c - 1$ elements of the maximum and minimum limits of the rated charging or discharging power of the battery, respectively. In this system, $P_{min} = 0$.

Similarly, the output constraints applied to the controlled variables can be represented as follows:

$$\begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \\ \cdots & \cdots & \cdots \\ 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix} SOC(k+1) \leq \begin{bmatrix} SOC_{max}(k+1) \\ \vdots \\ SOC_{max}(k+Nc) \\ \cdots \\ SOC_{min}(k+1) \\ \vdots \\ SOC_{min}(k+Nc) \end{bmatrix} \tag{4.22}$$

where, SOC_{max} and SOC_{min} are column vectors with N_c elements of the maximum and minimum limits of the observed $SOC(k+1)$. In this system, $SOC_{max} = 90\%$ and $SOC_{min} = 20\%$.

4.3 Summary of the Proposed Algorithm

The proposed hybrid strategy is implemented by a Distribution System Operator (DSO). It combines centralized and decentralized components: MCTS-RL is used as a centralized controller that optimizes cooperating buses in each network partition in terms of BESS availability and line losses, and subject to a voltage level constraint; MPC is used as a decentralized controller to mitigate overvoltage by estimating and controlling the battery state of charge as well as to optimize battery operation from the perspective of energy consumption. The profiles of load and PV generation, required as input by the MPC algorithm, are assumed to be available. While development of the forecaster is outside the scope of this work, there are a number of existing forecasting algorithms that can

be used for this purpose, such as [77]. The proposed CSOCC algorithm is shown in the flowchart of Figure 4.5. In this algorithm, the bus terminal voltages are compared to the threshold voltage ($1.05 p.u.$) to identify the impacted buses. All voltages and their buses coordinates are sent to the centralized controller to start specifying the assisting region using MCTS. In the proposed approach, MPC operates as a decentralized controller to achieve two objectives: a) to control the SOC, b) to optimize the daily schedule of BESS, based on the information from the centralized controller implemented using MCTS-RL.

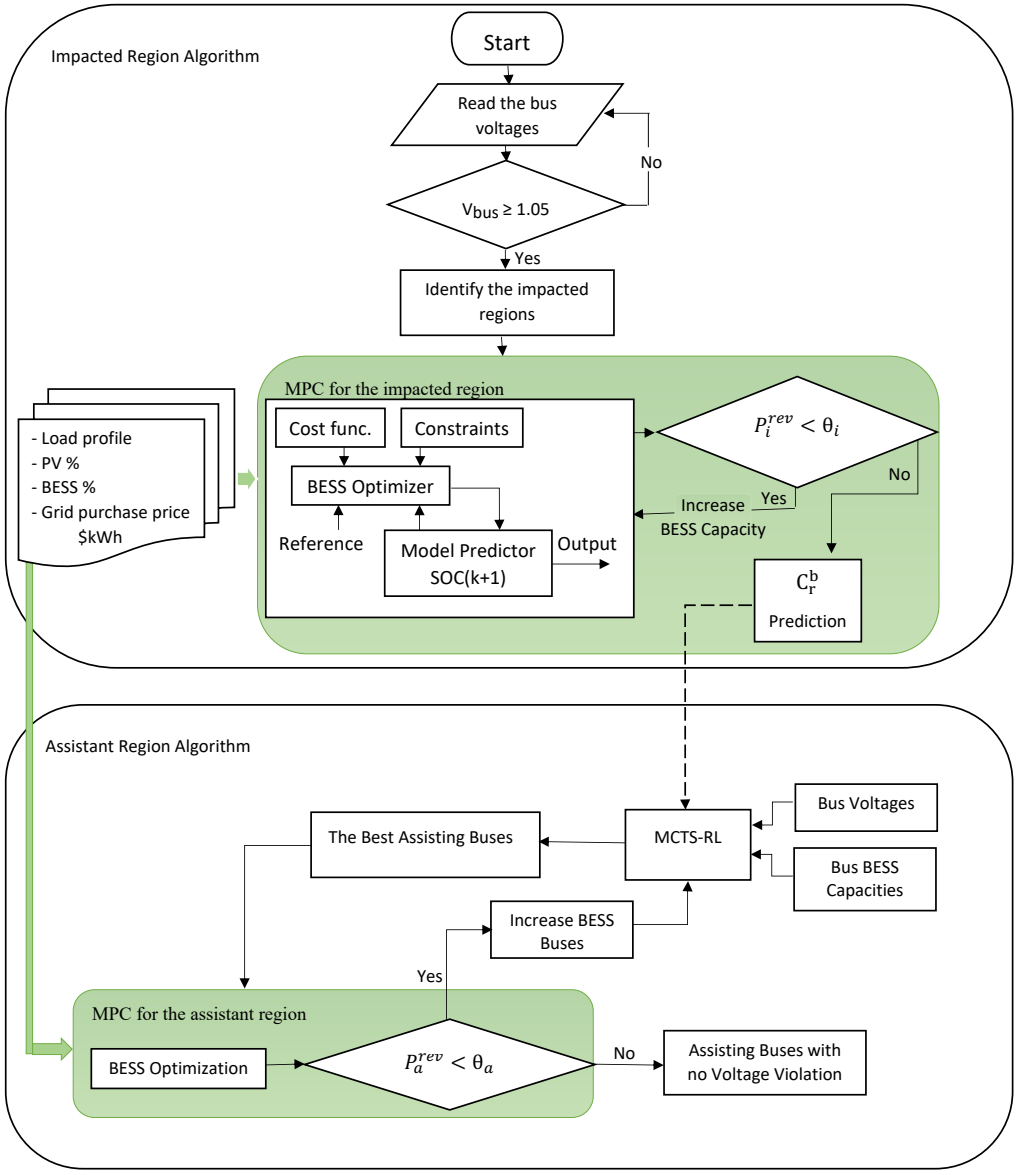


Figure 4.5: A flowchart of the proposed algorithm.

Each decentralized controller monitors the reverse power P_i^{rev} of the respective bus; the reverse direction compared to the original flow in the circuit is represented by a negative sign. This reverse power is often the main reason for overvoltage. This can be accomplished through a comparison of the reverse power for given impacted buses, P_i^{rev}, against a predetermined threshold, θ_i. Accordingly, there will be a required BESS capacity C_r^b,, computed by the MPC, to maintain the reverse power limit at the impacted region. MPC will provide

this information to the centralized controller to manage the energy from the assisting region(s). In other words, there will be an information packet sent from the impacted region to the centralized controller regarding its energy reduction requirement. θ_i is determined as follows [78]:

$$P_{PV}^{\max} = \frac{V_{\text{PCC}}^{\max}(V_{\text{PCC}}^{\max} - V_1)}{R_{th}(1 + \tan(\cos^{-1}(PF))\frac{X_{th}}{R_{th}})} + P_L, \tag{4.23}$$

$$\theta_i = P_L - P_{PV}^{\max} \tag{4.24}$$

where R_{th} and X_{th} are, respectively, the Thevenin resistance and reactance between the PCC and the feeder. P_{PV}^{max} and P_L are, respectively, the powers at the PCC of the maximum PV penetration level and the bus load. Accordingly, θ_i usually has very small value. Therefore, for simplicity, its value is approximated by zero in this study.

The role of the centralized controller is to determine the availability of neighboring BESSs to store energy using MCTS-RL algorithm. This algorithm uses three values of each bus: BESS capacity, voltage level and line losses determined when MCTS navigates through the paths to the buses with DG. The tree states of MCTS are updated through RL as follows

$$Q(s,a) \leftarrow Q(s,a) + \alpha\Big[r_{t+1}(1-\varphi_n) + \gamma Q_*(s^{'},a) - Q(s_t,a_t)\Big]. \tag{4.25}$$

This way, the algorithm selects the best buses in the assisting region with the lowest path losses. In equation (4.25), φ_n represents the percentage level of the voltage violation – the stability factor under normal power system conditions that maintains the security limit. Symbol s denotes the state of the line segment losses that leads to the next bus through the selected path. Finally, reward r represents the BESS capacity of the selected bus in the algorithm.

The state-value $V(s)$, under the optimal policy π for given state s and the state-action-value $Q(s,a)$ for taking action a, is calculated as follows [15]

$$V^{\pi}(s) = \max_{a \in A} Q^{\pi}(s,a), \tag{4.26}$$

where π is the optimal policy followed by the RL agent to optimize the action-value function Q. To mitigate the potential excessive voltage rise in assisting regions, engagement is monitored for each assisting BESS to check if its participation does not cause voltage constraint (1.05 p.u) violation within its own

region. Participation of each assisting bus is determined using BESS deployment rate α_a

$$\alpha_a = 1/D_a * \Big(\max(P_{ch,a,without\ CSOCC}) - \max(P_{ch,a,with\ CSOCC}) \Big) \tag{4.27}$$

where D_a is the full load demand of the assisting bus (subscript a denotes variables related to the assisting region). α_a may be reduced if the reverse power of an assisting bus P_a^{rev} exceeds a predetermined threshold θ_a which is determined by the same manner of θ_i of the impacted buses as mentioned above so θ_a equals to zero also. The resulting shortage of power injection capacity of the assisting buses is compensated by adding more BESS units.

4.4 Results and Discussion

4.4.1 Stochastic framework for PV deployment modelling

The impact of PV systems on a distribution circuit can be determined using a probabilistic deployment framework which is implemented by a probabilistic power Flow-based Monte Carlo simulations (PPF-MCS) [24, 59]. This method is used when some systems parameters are uncertain and can be considered random variables. PPF-MCS uses repetitive solutions of deterministic power flow with different realizations of random parameters to obtain expected probability distributions of variables of interest.The outcome of this stochastic modelling is then used to determine the PV penetration level defined as the ratio of the PV panel rated power (kW) to the value of the full-load (kW) at the connected bus [24].The probability distributions of bus voltages are compared with a predetermined threshold to identify the numbers and locations of buses affected for selected PV penetration levels.The stochastic framework is implemented in the following four steps:

1) A base case model of the selected distribution feeder is developed, assuming that there is no PV generation installed in the system that operates under a light load. The load level is derived from a known load profile. Depending on the feeder type (e.g. residential, commercial or industrial), the light load is defined as a percentage of the peak load.

2) Using the base case model, multiple PV deployment scenarios are considered for different PV penetration levels. For each level, a number of cases

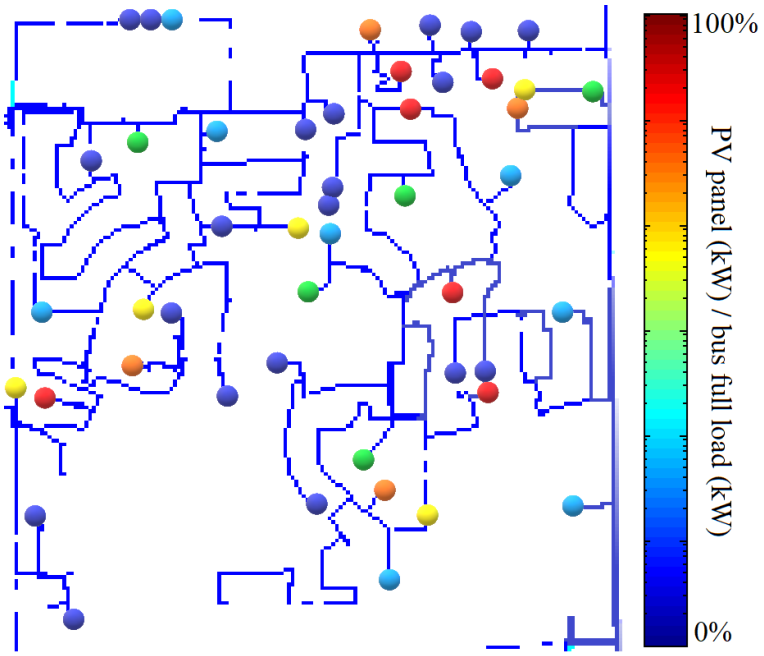


Figure 4.6: An illustration of the stochastic PV deployment scenario.

(e.g., 100) are considered, each with a random distribution of locations and sizes of PV generators which at the individual points of supply. An illustration of PV deployment scenario is shown for a small of the circuit in Figure 4.6.

3) Step 2) is repeated for each penetration level in selected increments (e.g. 10%) until the predefined performance index of the feeder is violated. Due to the stochastic nature of the modelling framework, this violation must also be defined in probabilistic terms taking into account the prescribed reliability indices.

4) When modelling the PV deployment scenarios is completed, the overall impact of PV penetration is assessed using a probabilistic evaluation of multi-phase load flow analysis to determine the PV hosting capacity.

It has been assumed that the PV bus is of PQ type [79] and the PV inverters cannot control voltage as there is no reactive injection or absorption at a majority of locations where PV generators are installed. In addition, the effects of capacitors on complex power are negligible and it can be safely assumed that very few nodes are near synchronous generators or motors. Consequently, the reactive power injection (negative VARs) is not considered in the analysis (i.e. only zero

or positive VAR values are considered) [14, 80]. In simulations, the load power factor (PF) of 1.0 is used, that corresponds to the worst-case scenario.

4.4.2 Test Case 1: IEEE 33 Bus System

4.4.2.1 PV hosting capacity

Since IEEE 33 bus system is a small circuit, it can accommodate high PV penetration without violating the commonly used voltage rise threshold of 1.05 p.u. Therefore, a lower threshold of 1.045 p.u. is considered in this study. The voltage violation starts at 90% PV penetration at 3 critical buses 16, 17, and 18 where voltages exceed 1.045 p.u.

4.4.2.2 Battery implementation without CSOCC

As discussed earlier, the overvoltage issue in distribution systems can be addressed using BESS with optimal charge/discharge energy scheduling. In addition to absorbing high instantaneous power due to PV generation at light load, this approach can also support heavy system demand during peak load periods. This BESS charge/discharge optimization is carried out by a decentralized controller using MPC without a central coordinator. To make the this test case consistent with the real case (circuit in Lloydminster, Alberta, Canada) described next, the hourly averaged electricity pool prices [81, 82] provided by the Alberta Electric System Operator (AESO) are considered. The energy price is high during the peak hours and low off-peak. Similarly, the 24-hour residential load profile for the city of Lloydminster is considered in both cases, scaled to the peak load of each circuit. The load demand and PV power generation profiles on the 3 critical buses are shown in Figure 4.7.

In this scenario, the BESS deployment rate (α_b) of each bus in the circuit (including the critical buses) is determined using BESS data collected over a 24h period as follows

$$\alpha_b = \max(P_{ch,i})/D_i, \tag{4.28}$$

where D_i is the full load demand of given critical bus. For instance, consider the bus causing the highest voltage violation in this region, bus 18. The MPC optimization results, shown in Figure 4.8, has $\max(P_{ch,i}) = 18$ kW. After dividing this value by the full load $D_i = 90$ kW (see Figure 4.7), the value of deployment rate is obtained as $\alpha_b = 20\%$. Figure 4.8 also shows that the BESS units discharge when the energy price rate is high. The units charge during two distinct periods:

1) when the energy price rate is low (during the night), and 2) when PV energy is generated (during the day).

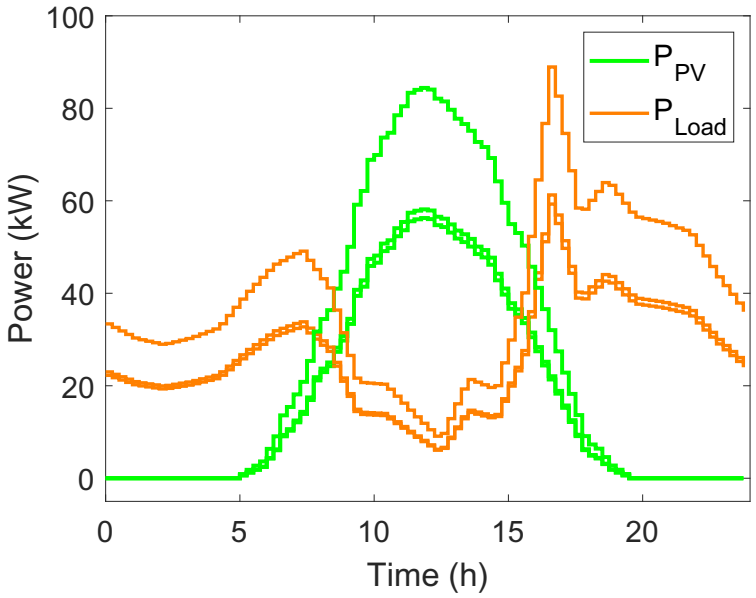


Figure 4.7: The load demand and PV power generation of the 3 critical buses 16, 17, and 18.

By examining the power flows in Figure 8, it is obvious that the capacities of individual BESS are not sufficient to allow charging throughout the entire period of high solar PV generation. This leads to a premature full use of battery charging capacity during the overvoltage mitigation processes. As a result, there is still large amount of reverse power fed to the bus terminal during the time period between 11:00 and 13:00. The highest impact occurs at bus 18 at 12:15, when reverse power of -54 kW leads to overvoltage of 1.047 p.u. (corresponding to the pronounced peak in Figure 4.17).

4.4.2.3 Battery implementation with CSOCC

To demonstrate the operation of the proposed approach, consider MCTS process applied to the IEEE 33 bus system. Each bus in the circuit can be considered a state, and each move from one bus to another an action of a virtual agent. The goal of the agent is to reach the state with the highest reward. It acts according to a policy learned through experience. At the beginning, the agent traverses from one bus to another with no knowledge about the environment and without knowing the correct sequence of buses. What the agent needs to learn is repre-

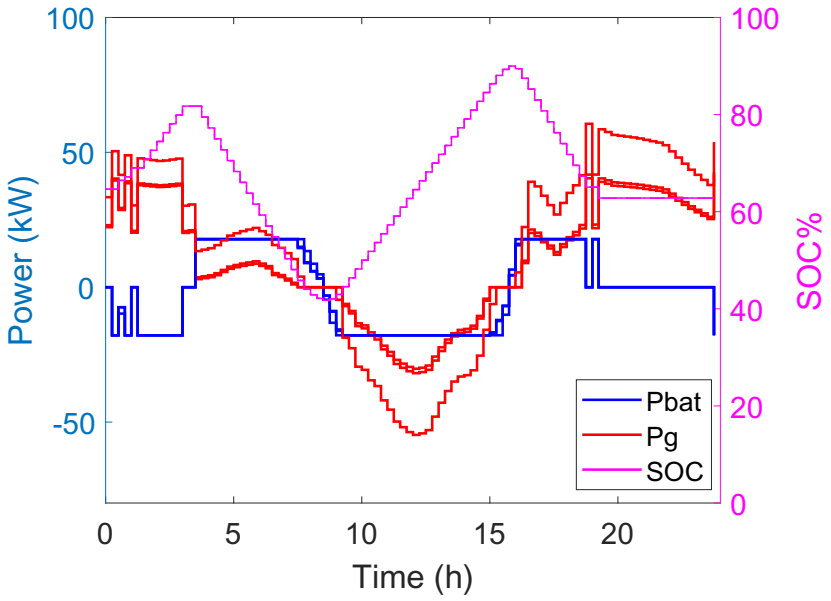


Figure 4.8: BESS powers controlled by MPC at the critical buses and their corresponding grid terminal power flows and SOC plots (without CSOCC).

sented by so called Q matrix [83] illustrated in Table 4.1. Each row of this matrix represents a current state and its columns represent possible actions that lead to the next state (i.e. the links between the buses). In the IEEE 33-bus circuit used in this illustration, each node has at most two branches that lead to other possible buses (called *children* in MCTS terminology). An action is represented by a random selection of one of the children. As shown in Table 4.1, the algorithm starts from a group of the impacted buses and navigates in a descending order of bus numbers. Initially, the action set has only one child to chose from, until it reaches state 6 that has two possible next states (buses 5 and 26). The agent chooses bus 26, and the algorithm continues.

This matrix is initialized with zero values corresponding to *tabula rasa* – an agent with no knowledge. A new value is assigned to each element as the agent explores the environment from state to state until the goal is reached, according to equation (4.25). This way, the virtual agent is learning through experience without a teacher. Each exploration process is represented by an episode and it is equivalent to one training session. More training results in more enhancement of the Q matrix that facilitates finding the fastest route to the optimal bus. The Q-values over the training episodes for three selected buses are shown in Figure 4.9. It can be seen that different Q-values are assigned to different buses. Bus 8 has

Table 4.1: State-Action

	Action	
State	Child 1	Child 2
16	15	-
15	14	-
14	13	-
13	12	-
12	11	-
11	10	-
10	9	-
9	8	-
8	7	-
7	6	-
6	-	26
5, 26	-	27
27	-	28
28	-	29
⋮	⋮	⋮

the highest value compared to buses 24 and 25. In addition, buses 24 and 25 have very similar line losses and identical BESS capacities that result in very close Q-values at the end of the training process. All three nodes converged after about 600 episodes.

Since the 33 bus circuit is a small system, all its buses are selected in the MCTS search space to determine the bus candidates that construct the assisting region.

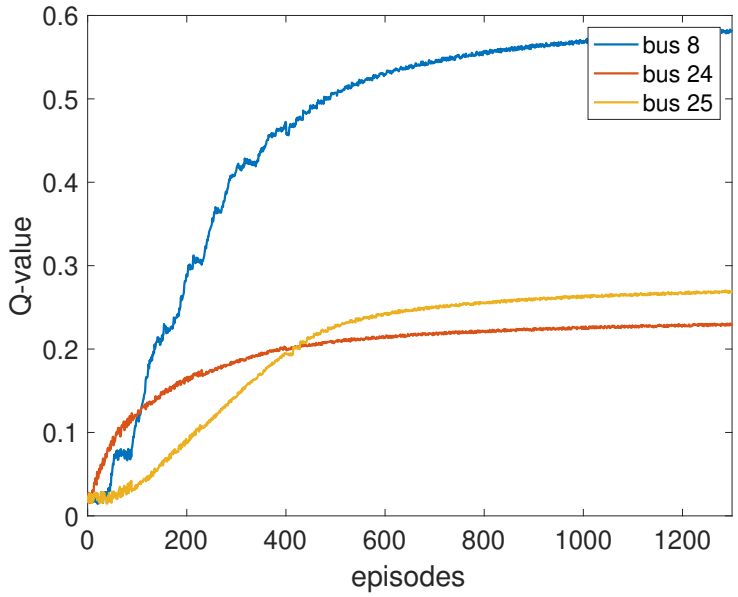


Figure 4.9: Plots of Q-values of buses 8, 24, and 25

For simplicity, and to provide a clear illustration of the MCTS navigation process, it is considered that a BESS is deployed on every bus of the assisting and the impacted region in both cases (without and with CSOCC). The size of BESS in the first case (without CSOCC) is uniform across the system, determined as α_b=20% of the full load using equation (4.28). With CSOCC, the BESS deployment rate in the impacted region is increased to a new value α_i calculated from the required energy obtained by MPC (462 kWh). On the other hand, the deployment rate of the assisting region is reduced so that each assisting bus provides 10% of its original BESS capacity to the impacted region.

Column 1 of Table 4.2 shows bus labels (2-33) in the search space. MCTS uses two objective functions: the line losses (column 3), and BESS power flow capacities (column 4) determined as 20% of the full load. Column 2 shows the bus voltages (with PV implemented) that are subject to the voltage constraint limit of 1.045 p.u. Reverse powers in assisting buses, if they occur, are subject to reverse power threshold $\theta_a = 0$. Line losses [79] are calculated through the MCTS navigation process over all line segments between the starting bus (the impacted region) and the bus selected as the assisting bus (a part of the assisting region). Finally, the bus prioritization is shown in column 5 (MCTS results). In this column, the candidate buses (encircled by a red ellipse) are selected based on the capacity required to store the excess energy in the impacted region.

Table 4.2: MCTS-RL Result

Bus No	Voltage (p.u)	Losses (kW)	BESS (kW)	MCTS
2	1.002	60.873	20	8
3	1.01	38.641	18	24
4	1.014	30.951	24	25
5	1.018	23.797	12	7
6	1.027	9.298	12	14
7	1.028	8.392	40	32
8	1.031	6.12	40	30
9	1.034	4.172	12	15
10	1.038	2.536	12	13
11	1.038	2.285	9	12
12	1.039	1.882	12	31
13	1.043	0.654	12	10
14	1.044	0.314	24	9
15	1.044	0.136	12	11
16	1.045	-	-	29
17	1.046	-	-	6
18	1.047	-	-	26
19	1.002	60.968	18	27
20	1.004	61.458	18	28
21	1.005	61.517	18	33
22	1.005	61.543	18	4
23	1.012	40.343	18	5
24	1.016	43.104	84	3
25	1.018	43.791	84	23
26	1.028	10.018	12	2
27	1.029	10.898	12	19
28	1.033	13.731	12	20
29	1.036	15.57	24	21
30	1.038	16.384	40	22
31	1.04	17.1	30	-
32	1.04	17.194	42	-
33	1.04	17.199	12	-

The required BESS capacities resulting from the MPC optimization process correspond to the energy differences between the two scenarios shown in Figures 4.8 and 4.10, over the time period when SOC changes from its minimum to maximum value. These capacities are determined based on the optimal BESS cost (4.13) while using the MPC controller (4.8). Deployment rate of the critical buses in the impacted region, α_i, is calculated through (4.28) for the amount of energy required by the impacted region. For example, in the scenario with CSOCC and considering the power flows shown in Figures 4.7 and 4.10, the deployment rate is α_i = 72kW/90kW = 80%. To determine the total required capacity of BESS for critical buses, C_r^b, one has to calculate the incremental amount of energy with respect to the base case without CSOCC. This amount depends on the difference between the two deployment rates and the load demand of the impacted region

$$C_r^b \approx \sum_{n_{\text{bus}}=1}^{N_{\text{reg}}} \int_{t_{\text{soc,min}}}^{t_{\text{soc,max}}} D_i(t)(\alpha_i(t) - \alpha_b(t))dt), \tag{4.29}$$

where N_{reg} is the total number of critical buses within a particular region, and $t_{\text{soc,min}}/t_{\text{soc,max}}$ are the times when SOC reaches its minimum/maximum value, respectively (see Figure 4.10). The total required BESS capacity is $C_r^b = 462.8$kWh, and the individual required capacities for buses 16, 17, and 18 are 116.2 kWh, 118.4 kWh, and 228.2 kWh, respectively. Figure 4.10 shows that the negative power flows in all three regions have been resolved. Since the reverse power is always checked, this method guarantees voltage rise mitigation. The BESS deployment rate of the assisting buses (encircled by the red ellipse in Table 4.2 column 5) is α_a=10% of the existing BESS units in the assisting region. The reduction of the deployment rate in this scenario, compared to the original value $\alpha_b = 20\%$ (without CSOCC) is $\alpha_b' = \alpha_b - \alpha_b \cdot \alpha_a = 20\% - 20\% \cdot 10\% = 18\%$ (10% is level of participation of BESS installed in the assisting region towards the impacted region).

Figure 4.11 shows the 33-bus voltages for three cases: i) 90% PV penetration with no storage; ii) 90% PV penetration with 20% BESS penetration but without CSOCC; and iii) 90% PV penetration with 20% BESS penetration and with CSOCC. It can be clearly seen that the proposed method significantly improves the voltage levels of the circuit, even compared to the case with BESS but without CSOCC. In this specific case, the use of storage alone (without CSOCC) resolves the overvoltage issues caused by high PV penetration. This can be attributed to the small size and low line impedances of the 33-bus circuit where the branches

between the feeder and the PV buses are short.

This is not the case for a real system of substantial extent and complexity, the aim of using this circuit is to explain the procedure steps and the feasibility of the proposed method. However, a practical test case of the Lloydminster circuit is described in the next section and the voltage threshold 1.05 p.u is considered.

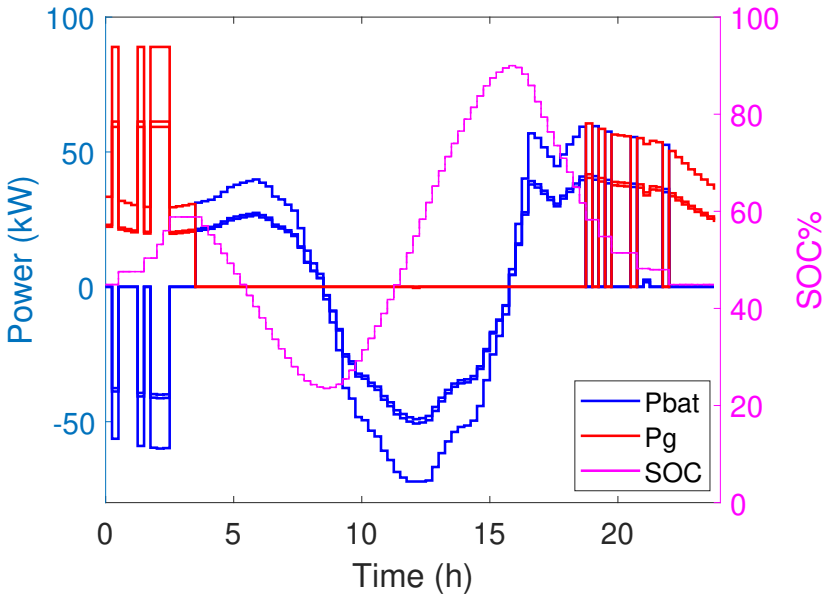


Figure 4.10: BESS powers controlled by MPC at the critical buses and their corresponding grid terminal power flows and SOC plots (with CSOCC).

4.4.3 Test Case 2: Lloydminster Circuit

4.4.3.1 PV hosting capacity

Similarly, as in the previous test case, the load flow analysis of the distribution circuit was simulated by PPF-MCS. However, in this test case with a real and complicated network CYME power engineering software is used to perform the power flow analysis. Also, the same assumptions of the previous case in terms of the reactive power and the power factor is considered here. Certain penetration levels cause the occurrence of impacted regions (regions with marginal, but still acceptable, voltage increase) shown in orange in Figure 4.12. The voltage violation starts at 60% PV penetration at 14 buses called critical buses that located at three regions A, B, and C called the impacted regions, shown in red where voltages exceed 1.05 p.u.

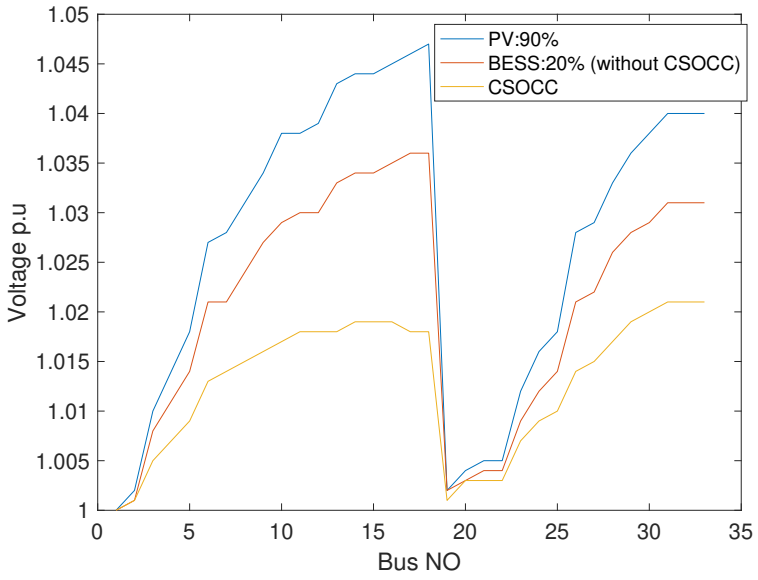


Figure 4.11: 33-bus system voltages in the following cases: i) 90% PV penetration with no storage; ii) 90% PV penetration with 20% BESS penetration but without CSOCC; and iii) 90% PV penetration with 20% BESS penetration and with CSOCC.

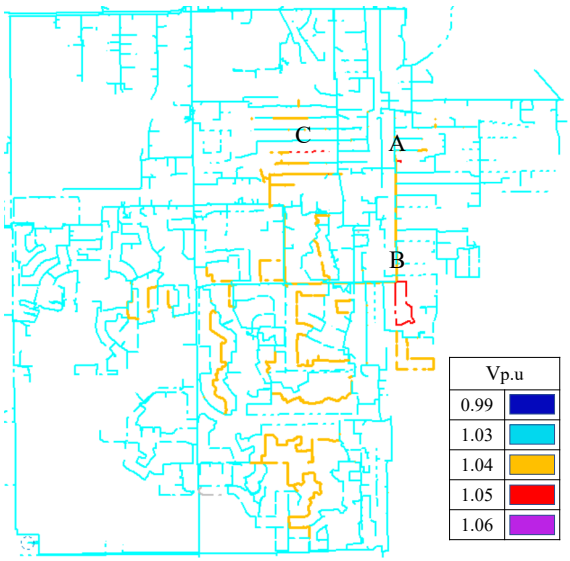


Figure 4.12: The impacted regions (orange) and critical buses (red); the impacted regions contain the following buses: a) bus 1 at (A), b) buses 2-8 at (B), c) buses 9-14 at (C).

4.4.3.2 Battery implementation without CSOCC

The load demand and PV power generation profiles of the 14 critical buses are shown in Figure 4.13, grouped into the three regions as in Figure 4.12. The load profiles are obtained as averages of the residential and industrial load curves from the circuit data. In this study case, only one region is enough to explain the simulation. For demonstration, area C has been selected as in this area the assisting buses fittingly surround the impacted buses where will be needed in the next scenario (with CSOCC). The worst bus in this area is bus 4, which is causing the highest voltage violation in this area. Based on the worst bus, the MPC optimization results, shown in Figure 4.14, has $\max(P_{ch,i}) = 5.34$ kW. and $D_i = 62$ kW (see Figure 4.13), the value of deployment rate is obtained as $\alpha_i = 9\%$.

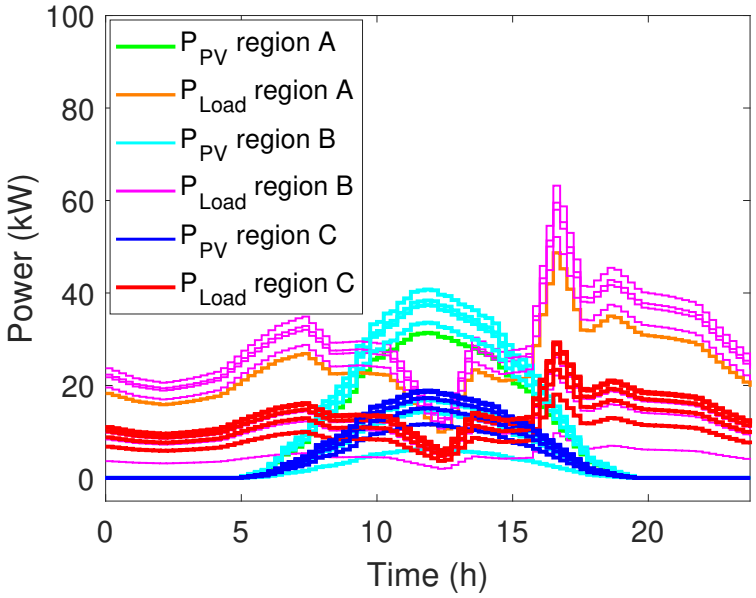


Figure 4.13: The load demand and PV power generation of the 14 critical buses in the three regions A, B, and C.

The same figure also shows that the reverse power is fed to the bus terminal between 11:00 and 13:00. The highest impact occurs at bus 4 (in region C) at 12:15, when reverse power of -6.1 kW leads to overvoltage of 1.052 p.u. (the pronounced peak in Figure 4.17).

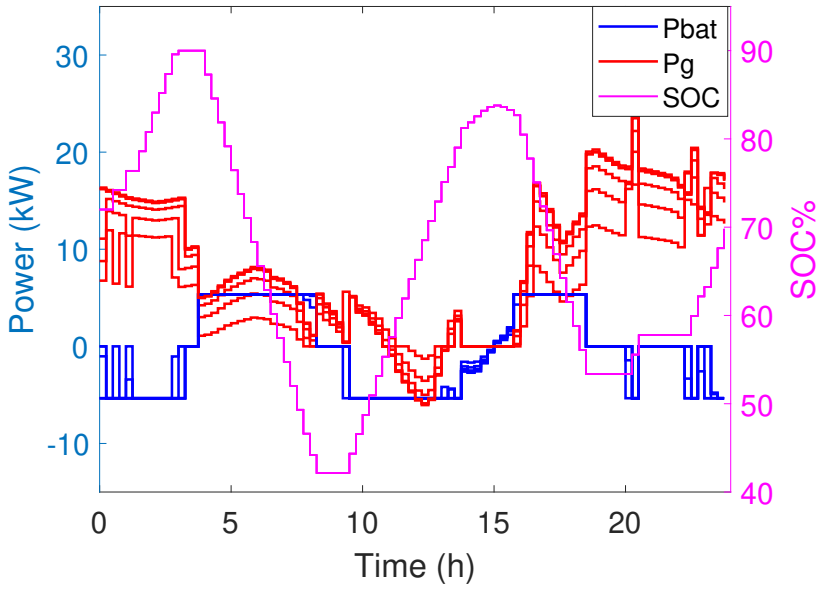


Figure 4.14: Terminal power flows at the critical buses and corresponding BESS flows (without CSOCC).

4.4.3.3 Battery implementation with CSOCC

As mentioned in the previous scenario, area C is considered in this test case. Figure 4.15d shows the buses in the search space that are selected as the best candidates by following the two objective functions the line losses and the BESS capacities, where their values are shown through the Figures 4.15a and b respectively, and the bus voltages that are checked with the voltage constraint are also shown in Figure 4.15c, and the revers powers if happen in the assisting buses will be subjected to θ_a which is zero. Similar to the previous test case, the excess energy in the impacted region may not need to use all assisting buses available in the search space. The participating buses in the assisting region are shown in different colors according to the priority of their selection determined by the optimization process (MCTS-RL algorithm). The required BESS capacities resulting from the MPC optimization process correspond to the energy differences between the two scenarios shown in Figures 4.14 and 4.16, over the time period of SOC maximum to minimum value.

In this test case, the deployment rate for the impacted buses is α_i = 11.5kW/62kW = 18% (obtained using information from Figures 4.13 and 4.16, as in the previous case).

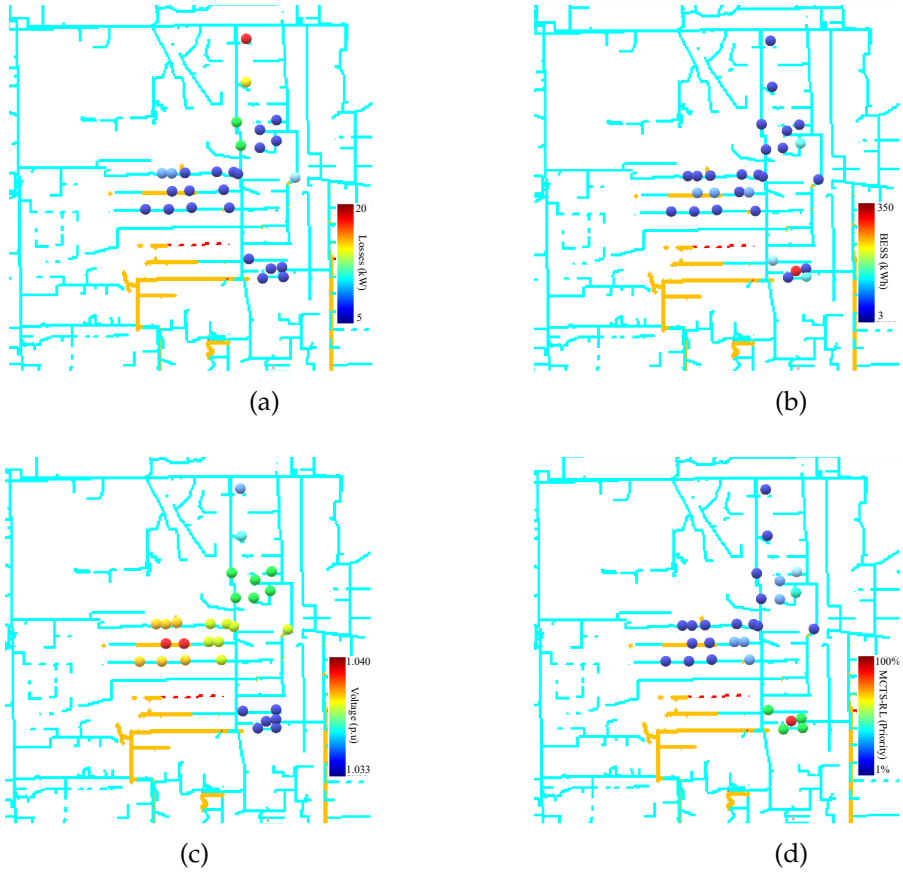


Figure 4.15: States at the assisting regions: a) line losses, b) bus BESS capacities, c) bus Voltages, d) the optimized assisting buses obtained using MCTS-RL d) the final results of the assisting buses corresponding to the critical region required energy

C_r^b values for impacted regions A, B, and C are determined as 42.2 kWh, 197.75 kWh, and 124.3 kWh, respectively, using equation (4.29). Moreover, in this practical case, the overvoltages are resolved and PV penetration of 60% can be achieved with a low BESS deployment rate of the assisting buses (20% of the existing BESS units in the assisting region). This low required penetration level can be considered an additional advantage of a robust power management in the centralized controller that reduces the overall fixed cost of this BESS-based solution. Voltages at the critical buses for the two scenarios (without and with CSOCC), obtained using CYME software with a 15 minutes resolution, are shown in Figure 4.17. The semitransparent blue plane represents the voltage limit of 1.05 p.u.

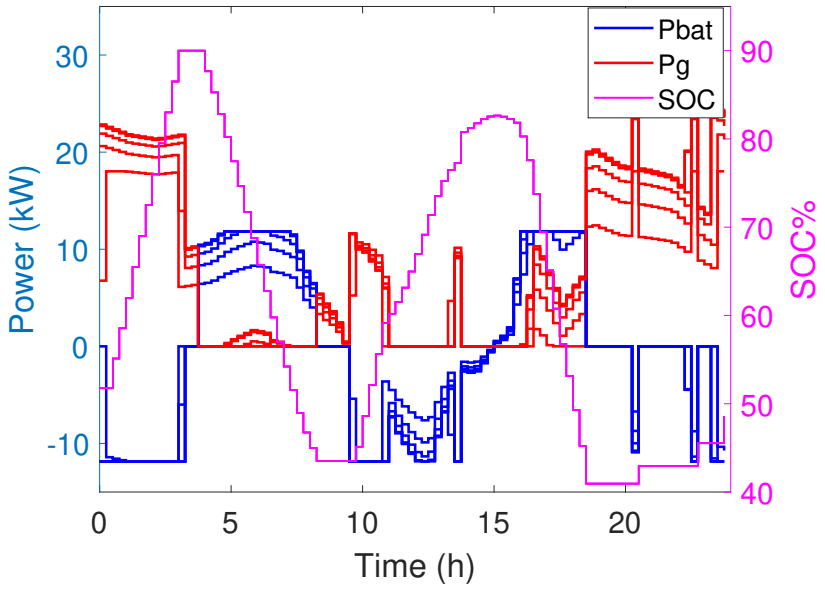


Figure 4.16: Terminal power flows at the critical buses and corresponding BESS flows (with CSOCC) .

r

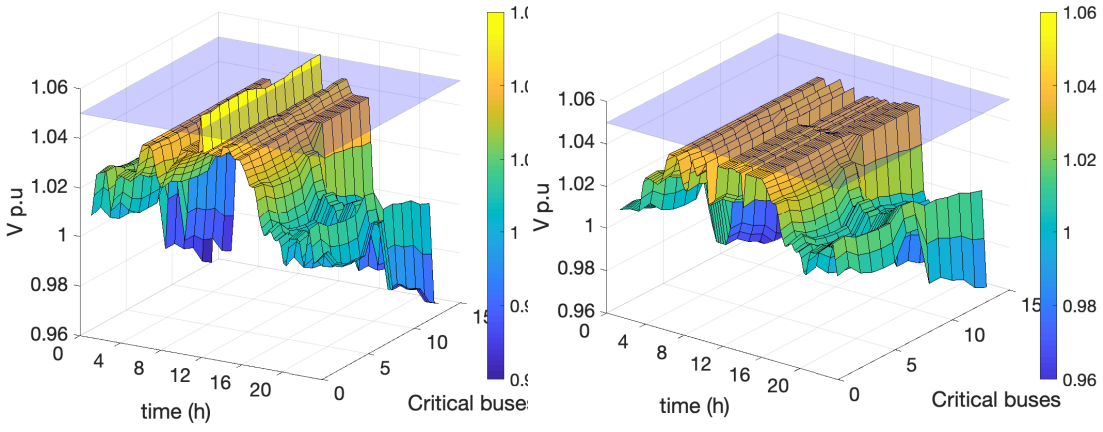
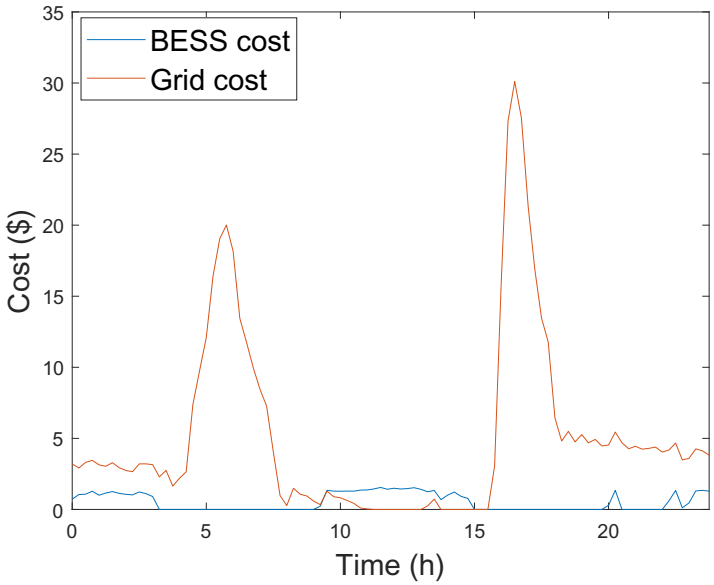
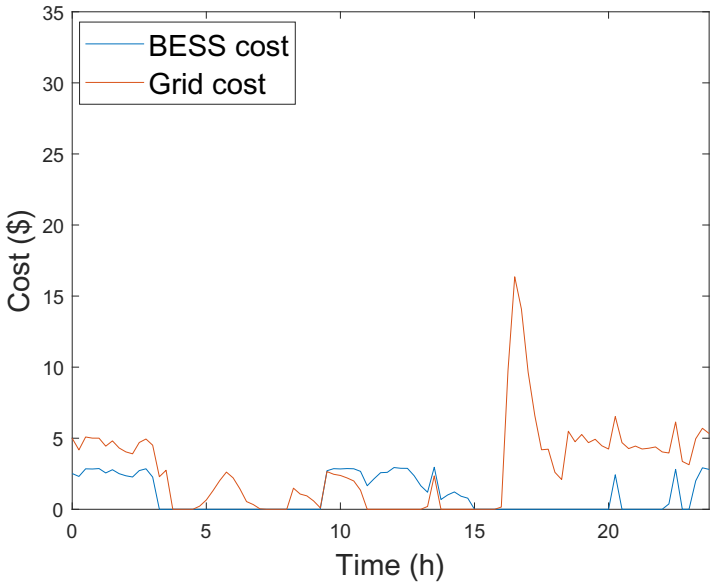


Figure 4.17: Voltages levels at the critical buses during a 24 hour operation without CSOCC (a), and with CSOCC (b).



(a)



(b)

Figure 4.18: The critical buses costs: a) without CSOCC, b) with CSOCC.

Finally, Table 4.3 shows the system costs for the following three scenarios: conventional method (without MPC), with MPC but without CSOCC and with MPC and CSOCC are calcualted for the region C. The conventional method is clarified in [10], [12], and [13], where the BESS charges at the PV time period and

Table 4.3: System secanrio costs

Scenario A	Conventional method	
	Impacted region	Assisting region
Grid Cost ($)	654.54	2919.00
BESS Cost ($)	16.61	47.73
Total Cost ($)	3637.90	
Scenario B	Without CSOCC (with MPC)	
	Impacted region	Assisting region
Grid Cost ($)	497.31	1332.80
BESS Cost ($)	50.49	340.51
Total Cost ($)	2221.10	
Scenario C	With CSOCC	
	Impacted region	Assisting region
Grid Cost ($)	269.19	1590.50
BESS Cost ($)	94.29	278.66
Total Cost ($)	2232.60	

discharges at the peak load at night time. Since the costs incurred by the assisting region must also be considered (in addition to the costs in the impacted region), the total cost of the area that includes both impacted and assisting region costs are shown in the last row of each scenario. It is obvious that the total costs in the two scenarios without and with CSOCC are significantly reduced compared to scenario of the conventional method. The use of CSOCC reduces the impacted region cost even further, as also shown in plots a) and b) of Figure 4.18. The total costs of both MPC-based scenarios are very close to each other. This is due to the fact that the BESS operation in both impacted and the assisting regions are optimized by the MPC that tracks the electricity pool prices and optimizes the BESS accordingly. The additional savings gained using MCTS-RL are relatively small because this algorithm only coordinates the BESS units in the two regions without changing energy adding new energy to the system.

There are other published works that also use coordination strategies to share energy among neighbouring BESS units [11–14, 19]. However, the do not consider the stochasticity of DER. Therefore, to facilitate coordination, these strategies methods require a priori knowledge that may not be available in practical

systems. On the other hand, the proposed model-free algorithm demonstrates robust performance under stochastic conditions of real system. This can be attributed to the efficiency of the multistage stochastic optimization based on RL and the diffusion strategy (i.e. the mechanism of allocating nodes to assist the impacted region). This strategy is based on randomly selecting routes between the impacted region and the assisting buses in a way similar to how Monte Carlo methods build probabilistic models for each state. Eventually, the diffusion strategy finds the optimal BESS capacities along routes that are optimal in terms of line losses and also subject to considering voltage constraints. Because the RL-MCTS controller has a global view of the system, the energy management over feasible region at each node is effectively obtained even though the BESS units in the system are only described by explicit stochastic models. The proposed method also facilitates BESS optimization through economic dispatch so that the cost of the system operation is minimized using MPC, while respecting system operation constraints. Finally, the use of diffusion strategy reduces the time required to find solution, because it restricts search to a small region surrounding the impacted region that examining the
entire search space.

4.5 Final Remarks

This research introduces a novel method for distributed control to mitigate voltage rise in power distribution networks caused by high penetration of residential photovoltaic generators. The proposed method, dubbed coordinated state of charge control (CSOCC), combines an intelligent method MCTS-RL and cooperative SOC control using MPC. In addition to preventing voltage violations through optimal use of network-wide installed battery capacity, it also minimizes losses due to power transfer. The described case study considers two scenarios of the battery implementation, without and with CSOCC, and corresponding energy management strategies. To prevent a premature saturation or depletion of BESS, the first scenario (without CSOCC) requires a significant increase of BESS penetration level from initial 9% to 18% at each impacted bus. The second scenario, using the proposed CSOCC approach, successfully mitigates voltage rise issues using the the assisting buses with only a low BESS participation of 20%. This is because the higher number of assisting buses with lower BESS participation compensates for high BESS penetration required for a lower number of impacted buses in the non-coordinated case. In the proposed approach, distributed

storage units are coordinated to form an assisting region that is optimized using MCTS-RL method considering the capacity of each available BESS, the voltage levels at the individual buses, and the line losses along the paths of required energy transfers. As a result, it resolves the voltage issues with a low BESS penetration level while inflicting minimal system losses. The effectiveness of the proposed method is examined on a circuit model of the distribution system of city of Lloydminster in Alberta, Canada. The results obtained using power engineering software package CYME show that the proposed distributed control approach is effective in mitigating over-voltages with a guaranteed performance since the values of reverse power are always checked. The proposed method can be implemented in real, complicated networks with multiple laterals. In the future, the algorithm can also be extended to incorporate other practical factors such as thermal constraints or ability to handle unbalanced three-phase systems.

Chapter 5

Distributed Optimal Power Flow for Electric Power Systems with High Penetration of Distributed Energy Resources

In the liberalized electricity market of North America, the NERC in conjunction with FERC introduced the concept of Available Transfer Capability (ATC). It is provided to all energy market participants of a power system through Open Access Same-time Information System (OASIS) [1]. This information must be available on hourly or daily basis for system optimization processes. Therefore, a strong economic incentive is necessary to improve the effectiveness of power system optimization and thus the transfer capability for the Distribution System Operators (DSO)s and power markets. Meanwhile, high Distributed Energy Resources (DER) penetration and their growth in power transactions will definitely affect the power system operation and the competitive power markets. Furthermore, the power transactions by these DERs are very dynamic, occur under varying load and generation conditions, and at different locations [84]. Therefore, the interactions between two independent bilateral power transactions under the power market environment need to be optimized and checked using Optimal Power Flow (OPF) method. However, the centralized OPF method would be very complicated under these conditions. The centralized optimization architecture already exhibits the following limitations [85]:

a) The massive number of DER units rapidly increases system uncertainty;

this results in control problems that require immense power computations and thus became fundamentally very demanding and hard to solve.

b) Control of the massive numbers of DER units requires massive implementation of communication systems and extensive data processing.

c) Data privacy and cyber-security risks become substantially more serious in a centralized optimization control architecture.

d) The market complexity that accompanies new technologies and new business models leads to suboptimal outcomes or even to intractable results.

Distributed optimization and management architectures represent a transitioning support towards a highly-efficient industry. These architectures can capture all the physical realities of the network and relieve a centralized agent from massive computing. This research explores the feasibility of fully distributed architectures for the electric power industry. This architecture is based on various electric power ecosystems that can interact with each other to achieve their power transaction objectives without the need for real-time centralized coordination or optimization. Moreover, decomposition technique is used to divide the large OPF problems into smaller subproblems [86, 87]. Spatial coordination method based on reinforcement learning (RL) can separate the challenge of optimizing interconnection operation into smaller units and optimize the internal operation economically. Ideally, the same system overall operating point is achieved while satisfying the constraints of power network boundaries, mainly the voltage stability limit and the line thermal limit or contingency.

5.1 DCOPF constraints Modeling in a Dynamic Power Market

In a distribution network with high DER penetration, power flow patterns may change after each power transaction due to the dynamic behaviour of DER. If the network is very large and complex, the AC OPF is a nonlinear, non-convex problem. Finding feasible solutions for such problems is a very difficult task [88]. A widespread approach in many industrial applications is so-called DC approximation for the OPF. This simplification leads to a convex optimization problem that can be solved quickly not only for a single transaction, but even for multiple simultaneous transactions. The DC OPF approximation assumes that the angle

differences θ_{ij} are small so that $\sin(\theta_{ij}) \approx \theta_{ij}$. It also assumes that the voltage is equal one at each bus in the network buses set N, and $G_l \ll B_l$ for all l lines in the set L where G_l and B_l are the conductance and susceptance of line l [89]. The key to applying the line losses constraint and the marginal loss price is the loss factor (LF) based on the line flow shift factor (also called power transfer distribution factor, PTDF). Thus, in order to formulate the DOPF using the linearized DCOPF, the PTDF and LF need to be established first.

5.1.1 Line Flow Shift Factor

When a DER injects power in the distribution network with N buses and L lines, a power transfer will change the network active power flow. Assuming P^{DER} is a random value of the DER power injected by a seller, and the power flow P between buses i and j is continuous and differentiable in P^{DER}, the distribution factor can be represented as [90]

$$\text{PTDF}_{ij,k} = \frac{\partial P_{ij}}{\partial P^{\text{DER}}}, \tag{5.1}$$

where, $\text{PTDF}_{ij,k}$ is the portion of the flow on line ij due to the injection at node k. In the linear approximation, the active power flow from i to j can be expressed as [89]

$$P_{ij} = B_{ij}[A\theta_{ij}], \tag{5.2}$$

where B is a susceptance matrix with a diagonal dimension of $L \times L$, A is an $L \times N$ incidence matrix that encodes how the buses are connected by the lines, and θ_{ij} is the voltage angle difference between the buses i and j. Solving for equations (5.1) and (5.2), the PTDF can be obtained as [91]

$$\text{PTDF} = B_d A B^{-1}. \tag{5.3}$$

Starting from the pre-transfer or base case $P^{\text{DER}} = 0$ and the pre-transfer flow of line ij being $P_{ij}(P^{\text{DER0}})$, the post-transfer flow of the line is determined as

$$P_{ij} = P_{ij}(P^{\text{DER0}}) + \int_{P^{\text{DER0}}}^{P^{\text{DER}}} \rho_{ij,T}(P^{\text{DER}})dp. \tag{5.4}$$

For a particular power transfer, the PTDF measures the sensitivity of a line's kW

flow to a kW transfer, so (5.4) can be rewritten as

$$P_{ij} = P_{ij}(P^{\text{DER0}}) + \text{PTDF}_{ij,T} \times P^{\text{DER}}. \tag{5.5}$$

5.1.2 Line Loss Factor

Another important factor is the line losses factor (LF). Losses can be quite significant for a large scale power system, and their impact on the DC OPF and locational marginal price (LMP) cannot be ignored. The LF, a function of a PTDF, is calculated using the iterative DC technique. It can be determined from the line flow losses as follows [92, 93]

$$P_{\text{Losses},l} = \frac{P_l^2 + Q_l^2}{V_l^2} \cdot r_l, \tag{5.6}$$

$$Q_{\text{Losses},l} = \frac{P_l^2 + Q_l^2}{V_l^2} \cdot x_l, \tag{5.7}$$

where P_l and Q_l are the real and reactive power flow, respectively. LF is defined as a linear sensitivity of the total system losses to the real power injections at each bus. Therefore sensitivity w.r.t. (5.6) can be determined as

$$\frac{\partial P_{\text{Losses},l}}{\partial P^{DER}} = \left(2 \cdot P_l \cdot \frac{\partial P_l}{\partial P^{\text{DER}}} + 2 \cdot Q_l \cdot \frac{\partial Q_l}{\partial P^{\text{DER}}}\right) \cdot \frac{r_l}{V_l^2}. \tag{5.8}$$

It will be assumed that all voltages are equal to 1 p.u. and the reactive power is constant during the DER power transactions, the reactive power derivative is zero and can be excluded from (5.8). This allows us to reduce (5.8) to

$$\frac{\partial P_{\text{Losses},l}}{\partial P^{\text{DER}}} = 2 \cdot P_l \cdot \frac{\partial P_l}{\partial P^{\text{DER}}} \cdot r_l. \tag{5.9}$$

The final expression of the n^{th} element of the LF can be represented as a function of PTDF

$$\text{LF}_n = 2 \sum_{l=1}^{L} r_l \cdot P_l \cdot \text{PTDF}_{l,n}. \tag{5.10}$$

5.2 Proposed MCTS-RL Algorithm

Machine learning has started to play an important role in maintaining the stability of autonomous micro-grids (MG)s and complex power networks. The primary reason is that it provides these systems with the ability to learn without

being explicitly programmed. Typically, a multi-agent system (MAS)-based distributed control scheme can interact with the environment to learn the state variations and find potential solutions. The interaction between the environment (the power grid) and the reinforcement learning RL (representing the system operator) can be formulated as a stochastic game. Let S be a set of system states in this environment. Each state $s \in S$ is a vector that indicates the current status of all the DER units in the search space. Next, A be a the set of all possible actions, $a \in A$, that the agent may take to achieve the optimal decision. In response to action a taken and state s traversed, the agent will receive an immediate reward $R(a, s)$ [94].

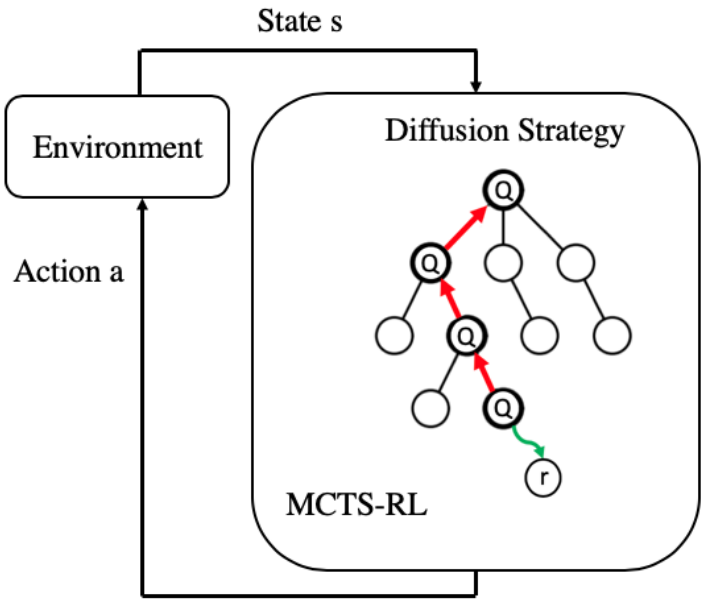


Figure 5.1: Proposed MCTS-RL algorithm.

The MCTS-RL algorithm aims to find the desired resource diffusion strategies, such as the best available DER units and the optimal power transfer paths. The use of MCTS provides the proposed algorithm with ability to navigate through the network and gradually build experience by performing tree search [95]. The addition of RL allows learning from the results to further optimize decisions taken by the system in an unsupervised fashion similar to temporal difference RL methods such as SARSA [96]. Figure 3 illustrates the proposed method.

Each bus in a power network can be represented as a node in MCTS, and each node contains state and action edges (s, a) of the tree. Each edge stores a set of statistical parameters: visit count $N(s, a)$, action-value $Q(s, a)$ pair updated by

Q-learning (5.11), and instant reward $r(s,a)$. These parameters are defined as follows

$$Q(s_t,a_t) \leftarrow Q(s_t,a_t) + \alpha\Big[r_{t+1} + \gamma Q(s_{t+1},a_{t+1}) - Q(s_t,a_t)\Big], \tag{5.11}$$

where α is a learning rate which controls the extent to which the value function is updated, and γ is a discount factor which indicates the effect of the current decision on the long-term reward.

The agent selects an action a, and the network state is changed from s to s'. Accordingly, the Q-value of the agent is updated, and the number of visits on the visited nodes is increased by one. The tree state is randomly built up and the accumulated experience in each state is updated by the random sampling during the process of the optimization and exploration policy and store the information in the node states by the back propagation process.

5.3 The proposed DOPF based on MCTS-RL Model

The distribution network is modeled as a directed tree graph $T(\Omega_k, L)$. It consists of a set of buses Ω_k (part of the set of all network buses N), also referred to as nodes, and a set of distribution lines L that link these nodes. The subset Ω_k contains K linked nodes that are indexed as $k = 0, 1, ..., K$. Node 0 denotes the root node, the starting point of the tree search. In general, this node can represent either a buyer or seller. However, under the scenarios considered in this study, the root node is restricted to being a buyer looking for the best seller(s). The remaining nodes in Ω_k (except the root node j) are branch nodes. All pairs of branch nodes are linked together by a branch line l. All nodes (except the leaf node) in this tree are considered parent nodes since they have a set of child nodes C_i connected to the parent by the branch lines. The child nodes may have connected generators. For each line in L that links a parent node to a child node, there is an impedance $z_i = r_i + x_i$. Power injection from i to j is computed using (5.5).

The objective of the DOPF is to determine the generation dispatch that results in the lowest cost to supply the load while all system constraints are satisfied. Based on the DC OPF approximation, the DOPF problem formulation will be determined as follows

$$\min \sum_{x \in \Omega_k} f_i(x_i), \tag{5.12}$$

$$f(x) = f^1(C_P) + f^2(\mathrm{LF}_{ij}),$$

$$h(x) = \begin{cases} P_i - d_j = \sum_{i \in \Omega_k} \text{PTDF}_{ij,x} + \text{LF}_{ij} \, \forall j \in N, & (5.13) \\ \theta_{ref} = 0, & (5.14) \end{cases}$$

$$g(x) = \begin{cases} \underline{P}_x \leqslant P_x \leqslant \overline{P}_x \quad \forall x \in \Omega_G, & (5.15) \\ \underline{\text{PTDF}}_{ij,x} \leqslant \text{PTDF}_{ij,x} \leqslant \overline{\text{PTDF}}_{ij,x} & (5.16) \\ \text{LF}_{ij} \leqslant \phi \qquad \forall l_{ij} \in N. & (5.17) \end{cases}$$

Here x is the DER state variable which is updated by RL. It belongs to the feasible set Ω_k of the MCTS region which is part of the original distribution network. This variable includes the generation output cost C_p, and the loss factor LF_{ij} between the load bus and the DER bus. Functions $h(x)$ and $g(x)$ are the sets of equality and inequality constraints, respectively. Ω_G denotes all DER units that belong to the feasible set of the MCTS region Ω_k, Ω_i refers to the set of all buses connected to the load bus j, and Ω_L refers the set of all branches of these buses. The balance nodal power flow is enforced by constraint (5.13). Constraint (5.14) keeps the reference bus voltage angle at zero. The inequalities (5.15) and (5.16) ensure that power flow in the branches and the power output of the DER unit are within their upper and lower bounds. Constraint ϕ provides coupling between region Ω_k and its neighboring regions. According to the DC OPF and the concepts of electric distance [97], ϕ represents the threshold value of the LF where the latter is a function to the voltage angle difference between the load bus j and the cross-bordering buses separating region Ω_k from its neighbors. Thus it identifies the feasible space of the OPF problem in multiple regions.

RL changes its policy through on-line learning to allow adaptation to any changes in power system states and maintaining system security. This cean be expressed through a modified state-action-value update equation

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \Big[r_{t+1}(m - \varphi_n^{'}) + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t) \Big], \quad (5.18)$$

with additional parameters

$$\varphi_n^{'} = \begin{cases} 1 & \text{if } m \geqslant \tau, \\ 0 & \text{otherwise} \end{cases}$$

$$m = 1 - (\varphi_n + \text{LF}),$$

where φ_n is the set of all constrain factors under normal power system condi-

tions, and τ is the constrain threshold that maintains the security limit. The state-value $V(s)$, under the optimal strategies for a given state s, is represented as

$$V_*(s) = \max \sum_{a \in A, s'} \pi_*(s), Q_*(s', a), \tag{5.19}$$

where $\pi_*(s)$ is the optimal state policy.

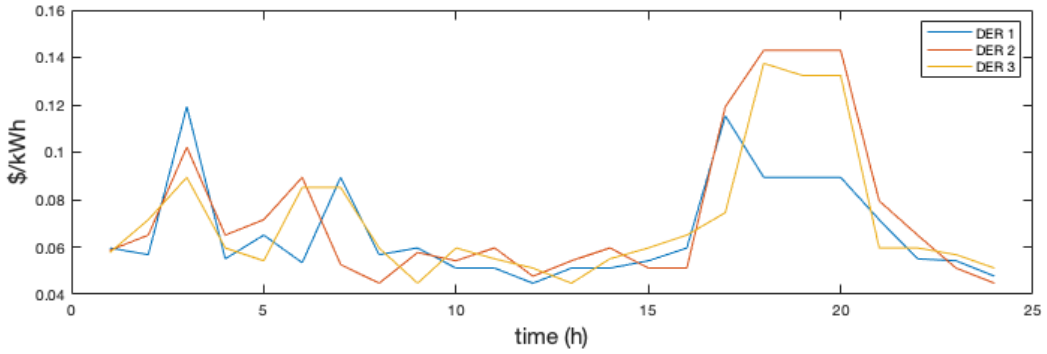
5.4 Results and Discussion

To demonstrate the proposed MCTS-RL algorithm, a real-world network, IEEE 33-bus test system is considered as a case study [98]. The testing DER dynamic price bid profiles and corresponding generation profiles are shown in Figures 5.2a and 5.2b, respectively. Load profile at node 2 is shown in Figure 5.2c. DER units 1,2 and 3 are connected to buses 25, 6, and 22, respectively. The algorithm progressively builds a partial game tree [95], with a greedy and randomly biased sequence of actions applied to a given series of states, is simulated for the MCTS and trained on using RL.

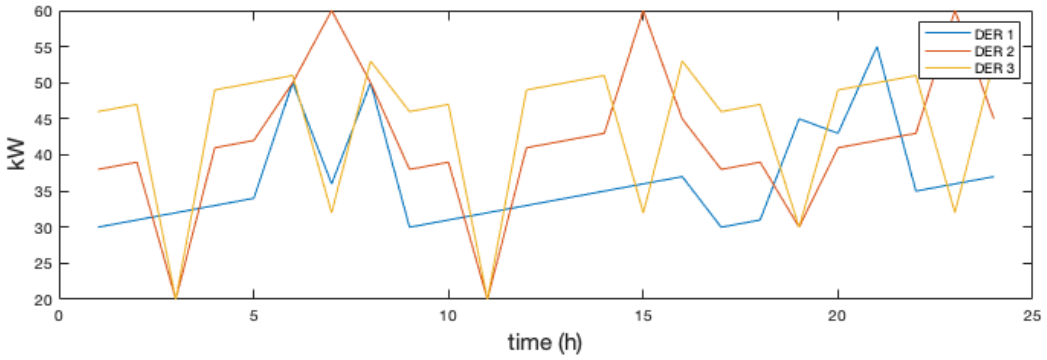
MCTS-RL attempts to obtain the optimal selection of DER price bids through building a search tree until the predefined constrain ϕ is reached. Figure 5.3a shows that different DERs are selected at each time slot depending on the best price bids following the load demand shown in Figure 5.2c. When a single DER is not capable to cover the demand at a particular time slot, more than one DER is selected to participate in the transaction. The best DER prices that have met the demand during the sample day are plotted in Figure 5.3b.

The scenario described above does not consider contingency analysis and may lead to line congestion as shown in Figure 5.5a. To resolve this problem under non-linear load and DER behavior, the contingency constraint is added to the RL update algorithm (5.18), and implemented during the optimization process. In this new scenario, generation of DER units that compromise network security is reduced even if they offer energy at the best prices. Instead, other DER units that do not influence system security while providing energy at satisfactory price offers are selected to satisfy the load as shown in Figure 5.4a. This demonstrates that the proposed DOPF model can effectively consider system contingencies and ensure efficient power market solutions that do not jeopardize system security. It should be noted that this added security comes at the price of a slight increase of the system overall generation cost, as shown in Figure 5.4b. For a more detailed comparison of the contingency analysis in the two considered sce-

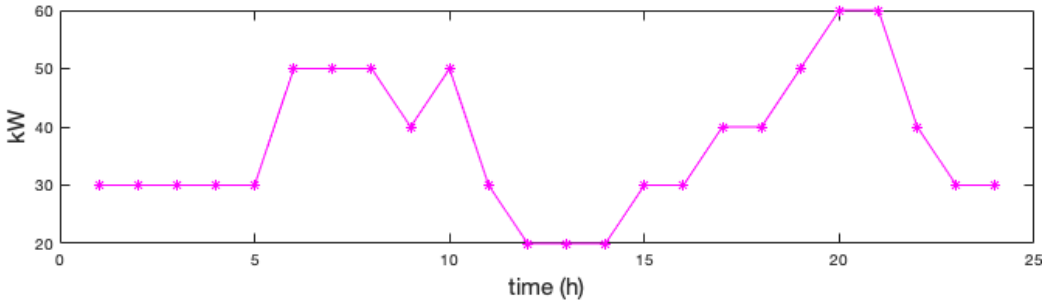
narios, refer to Figure 5.5.



(a)

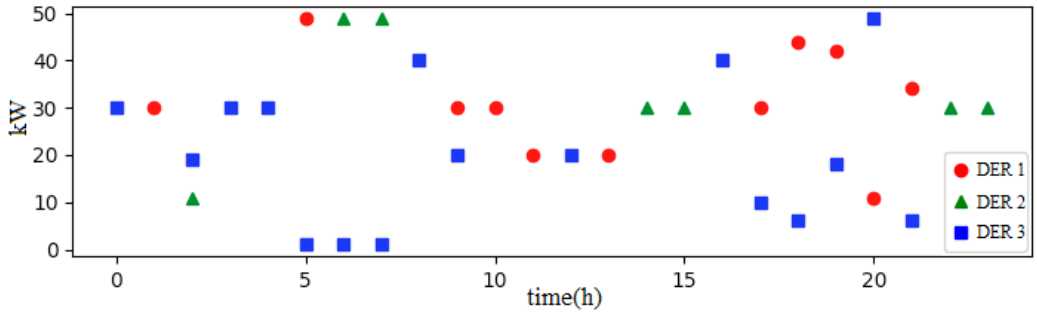


(b)

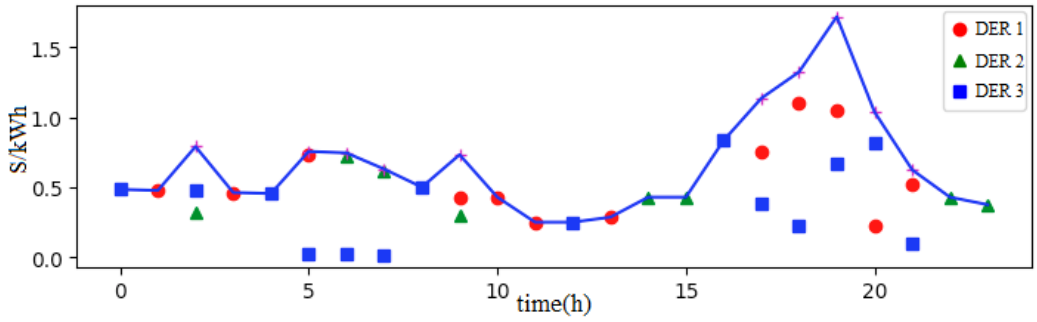


(c)

Figure 5.2: DER units and load used for testing: (a) DER dynamic prices profile, (b) DER dynamic power output, (c) Load profile at node 2.



(a)



(b)

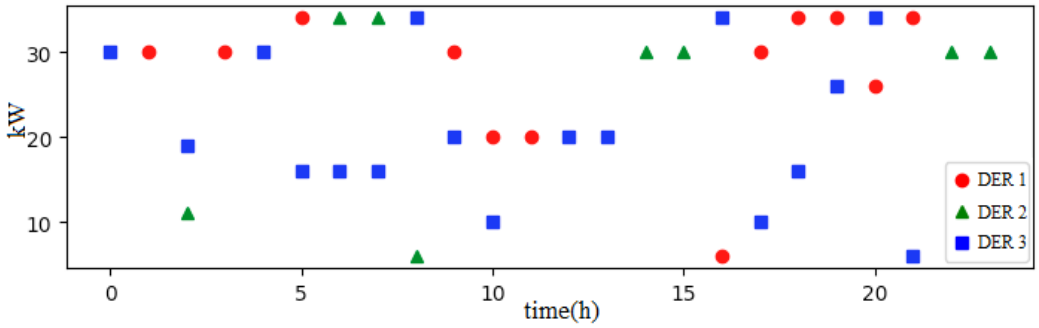
Figure 5.3: Results of MCTS-RL optimization when line congestion is not considered: (a) Optimal DER units, (b) DER generation cost.

5.5 Final Remarks

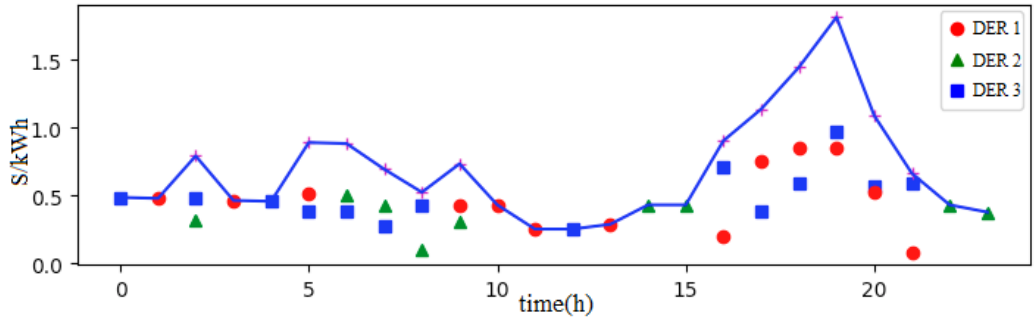
This contribution introduced an efficient method for distributed optimal power flow (DOPF) calculation that overcomes shortcomings of the state-of-the-art OPF methods. As such, this new methodology will find applications in the emerging distribution grids with high penetration of distributed energy resources. The developed DOPF has a scalable architecture for energy optimization and management at all levels.

The proposed methods is based on a powerful combination of Monte Carlo tree search (MTCS) and reinforcement learing (RL). MCTS-RL leverages a diffusion strategy to generate a set of optimal routes through navigating actions/lines that link to multiple candidate/child nodes in the microgrid. The speed at which MCTS-RL can obtain optimal solution is one of its main advantages that makes it suitable for real-time system with dispatchable resources. This algorithm can avoid the dimensionality problem and guarantee convergence.

Simulation results also demonstrate that the proposed DOPF can achieve cost



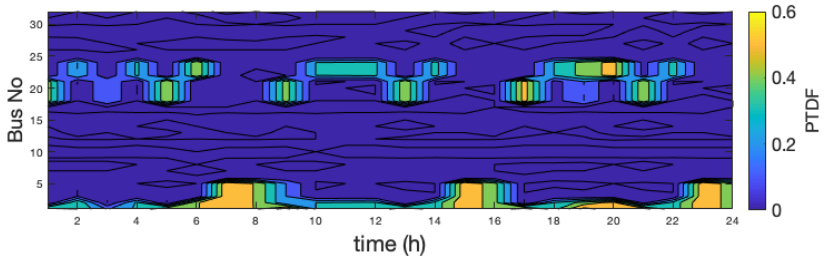
(a)



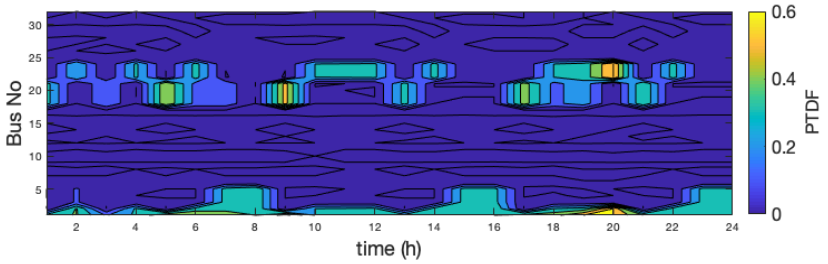
(b)

Figure 5.4: Results of MCTS-RL optimization when line congestion is considered: (a) Optimal DER units, (b) DER generation cost.

minimization for distributed economic dispatch in microgrids while maintaining the system security limits. The presented analysis can be of use to many practitioners who are interested in designing and developing power distribution markets. The key advantage of the proposed method is that it can be efficiently integrated into power market software.



(a)



(b)

Figure 5.5: Contingency analysis (a) Congested lines are unconsidered in MCTS-RL. (b) Congested lines are considered in MCTS-RL.

Chapter 6

Distributed Optimization for Distribution Systems with Stochastic DER using Multi-Agent Deep Reinforcement Learning

Optimal power flow (OPF) is an essential tool for managing energy in electric power systems. It seeks the least-cost operation of a power system by dispatching generation for given power demand while satisfying the system constraints. The changing nature of modern power grids brings new entities into electric power markets. They include owners of distributed energy resources (DERs), and even so called prosumers - individual customers equipped with self-owned DER units. The new market participants are interested in autonomous maximization of their profits. Therefore, they can be considered independent entities of the system [99]. However, a decision made by a single entity may affect the decisions of the remaining entities that are physically interconnected in the same system.

As power distribution systems are becoming more and more dispersed, they may require additional generation capacity and new line assets to supply the peak demand. The network participants may need to cooperate with each other to achieve reliable and effective operation of the network without changing the system infrastructure. The incremental dispersion of new network entities will

also affect the electric power markets. In this new scenario, the interactions between two independent bilateral power transactions in the network need to be checked and optimized using OPF. However, the conventional centralized OPF method poses a number of problems [85]. To avoid these issues and provide the power industry with tools to support highly efficient system operation, distributed optimization architectures are required. Such architectures can capture all physical realities of a dispersed network and alleviate a centralized optimization agent from tremendous amount of computing.

Recent literature presents several approaches to distributed economic dispatch [31, 100]. They resolve the randomness of DER units and loads in microgrids through the use of Markov decision processes (MDPs). Fully distributed methods for economic dispatch in microgrids have been proposed by Xia et al. [27] and Li et al. [28]. They derive a set of distributed control laws for microgrid agents. The underlying communication network has two layers. One ensures the supply-demand balance in the microgrid, while the other one coordinates control agents. Distributed model predictive control (MPC) for stochastic dispatch optimization in microgrids have been proposed by several authors [34–36]. They use a local MPC for each entity to implement receding-horizon optimization.

Other authors use a divide-and-conquer approach [29, 30]. They decompose the centralized optimization problem into many smaller optimization problems executed by local agents. Each agent can exchange information with its network neighbors. After the information is processed, agents adjust production of their DER units in a distributed manner with limited communication among the entities.

A common shortcoming of all these approaches to distributed economic dispatch is that they require prior statistical information on all DER units and loads. In addition, they cannot effectively cope with the dynamic nature of power transactions that occur under varying load and generation conditions, and at different locations.

Reinforcement learning (RL) is a powerful tool to solve complex sequential decision-making and control problems. RL can effectively learn optimal stochastic policies, even in high-dimensional or dynamic action spaces. It can reach the goal state in a few steps, with high probability, and without relying on prior information or complex stochastic modeling. These properties make RL a suitable tool to address the multistate stochastic optimization problems in modern distribution grids.

A simple solution is to use tabular Q-learning [100]. In this approach, RL only

finds a feasible region that contains DERs that are implicitly considered optimal. However, it does not find DERs that can contribute to the acceleration of the optimization process. In addition, tabular Q-learning does not work well with continuous observations in complex systems with many DERs. A cooperative RL approach has been proposed by Liu et al. [18]. The authors suggest that each distributed controller exchanges information with its neighbours, makes action decision based on its own state and the neighbourhood states, and performs so called distributed cooperative mechanism. However, the system observability is limited to the neighbouring buses, leading to limited power transactions. In addition, this approach does nor consider the real-time impact of the line flow variations due to the power transactions.

In practice, operation of power systems relying on machine learning may be affected by approximation errors. This may increase the cumulative operation costs of the system or even cause damage to the equipment connected to the circuit. On the other hand, the use of exact optimization methods in complex distribution systems with stochastic DER units is often impractical due to the increase of computational burden associated with such methods. For example, the use of linear programming (LP) may not be possible due to a massive number of control variables and associated conditional statements. Hence, the proposed model presents a hybrid approach that avoids the drawbacks of both constituents: machine learning errors and lack of scalability of conventional optimizers. The proposed system, called Multi Leader-Follower Actors under Centralized Critic (MLFACC), can fully capture the environment states and learn from the behavior of network participants to determine the optimal DERs before they are sent to LP for power generation optimization. To expand the capability of distributed OPF algorithms from microgrids to large, complex distribution circuits, it is necessary to monitor network states and communicate them among the microgrids. This can be accomplished through the proposed multi-agent system (MAS) architecture.

This research proposes a multi-agent RL system that allows agent controllers to adapt to changes in the power distribution network as a means to maintain system security. The feasible region in a large system is obtained using Monte Carlo Tree Search (MCTS) as a first step. This is followed by deep RL-based optimization to navigate from a buyer bus through the entire network (i.e. beyond the local neighbourhood). This procedure finds the best DER units to buy power from, while reducing the search space compared to the centralized OPF.

This research is organized as follows. Section 6.1 describes a power flow

linearization method used to calculate power injection in the proposed OPF approach introduced in section 6.2. Illustration of the new approach is based on a model of real-time balancing electricity market constructed in section 6.3. Simulation results based on IEEE 69-bust system are presented and analyzed in section 6.4.

6.1 Power Flow Linearization

In large distribution networks with high penetration of intermittent DER units (such as photovoltaic and wind generators), the power flow computational burden becomes substantial. In addition to the impact of scale, OPF needs to be checked more frequently due to the dynamic behaviour of DER units. However, the OPF in AC systems is a nonlinear, non-convex problem [101]. Therefore, finding feasible solutions for such problems is a very difficult task. A common approach is DC OPF approximation that leads to a convex optimization problem which can be solved quickly. However, its use for practical large distribution networks with a high system R/X ratio negatively affects the accuracy of OPF computations.

Yang et al. [102] illustrate the impact of several approximations used in the linearization process on branch power flows. They start from the well known polar AC power flow model

$$P_i = \sum_{j=1}^{N} G_{ij} V_i V_j \cos\theta_{ij} + \sum_{j=1}^{N} B_{ij} V_i V_j \sin\theta_{ij} \tag{6.1}$$

$$Q_i = -\sum_{j=1}^{N} B_{ij} V_i V_j \cos\theta_{ij} + \sum_{j=1}^{N} G_{ij} V_i V_j \sin\theta_{ij} \tag{6.2}$$

where N is the bus number, and G_{ij} and B_{ij} are the conductance and susceptance of the line. There are three main approximations [102] of the expression for branch power flow $G_{ij}V_i(V_i - V_j cos\theta_{ij}) \approx$

$$\text{a) } 0, \quad \text{b) } G_{ij}(V_i^2 - V_j^2), \quad \text{c) } G_{ij}(V_i - V_j).$$

Based on voltage computation results [102], the third approximation (c) provides the best accuracy. Using this simplification, the linearized models of the

active and reactive power injections at bus i are [103]

$$P_i = \sum_{j=1, j\neq i}^{N} \frac{k_{ij_2}}{x_{ij}}(\delta_i - \delta_j) + \frac{k_{ij_1}}{x_{ij}}(V_i - V_j), \tag{6.3}$$

$$Q_i = \sum_{j=1, j\neq i}^{N} -\frac{k_{ij_1}}{x_{ij}}(\delta_i - \delta_j) + \frac{k_{ij_2}}{x_{ij}}(V_i - V_j), \tag{6.4}$$

where

$$k_{ij_1} = \frac{r_{ij}x_{ij}}{r_{ij}^2 + x_{ij}^2}, k_{ij_2} = \frac{x_{ij}^2}{r_{ij}^2 + x_{ij}^2}. \tag{6.5}$$

To solve equations (6.3) and (6.4), the node voltages have to be obtained first

$$\begin{bmatrix} P' \\ Q' \end{bmatrix} - \begin{bmatrix} B_2^c \\ -B_1^c \end{bmatrix} \delta_1 - \begin{bmatrix} B_1^c \\ -B_1^c \end{bmatrix} V_1 = \begin{bmatrix} B_2' & B_1' \\ B_1' & B_2' \end{bmatrix} \begin{bmatrix} \delta' \\ V' \end{bmatrix}, \tag{6.6}$$

where P', Q', δ', and V' are vectors of real power injection, reactive power injection, voltage angle, and voltage magnitude, respectively. Matrices B_1^c, B_2^c, B_1' and B_2' can be found in [103].

In a large scale power system, losses can be quite significant and their impact on the OPF and locational marginal price (LMP) cannot be ignored [103]. The flow losses for line l can be determined as follows [79, 100]

$$P_{\text{loss},l} = \frac{P_{\text{flow}}^2 + Q_{\text{flow}}^2}{V_l^2} r_l, \; Q_{\text{loss},l} = \frac{P_{\text{flow}}^2 + Q_{\text{flow}}^2}{V_l^2} x_l, \tag{6.7}$$

where P_{flow} and Q_{flow} are the real and reactive power flow, respectively. The loss factor is defined as a linear sensitivity of the total system losses to the real power injections at each bus with connected DER, i.e. $\text{LF} = \partial P_{\text{loss},l} / \partial P^{\text{DER}}$. Substituting (6.7) for $P_{\text{loss},l}$, one gets

$$\text{LF} = \left(2P_{\text{flow}} \frac{\partial P_{\text{flow}}}{\partial P^{\text{DER}}} + 2Q_{\text{flow}} \frac{\partial Q_{\text{flow}}}{\partial P^{\text{DER}}} \right) \frac{r_l}{V_l^2}. \tag{6.8}$$

Assuming that the reactive power is constant during DER power transactions, its derivative is zero and thus the second term can be excluded from formula (6.8), reducing it to

$$\text{LF} = 2P_{\text{flow}} \frac{\partial P_{\text{flow}}}{\partial P^{\text{DER}}} \cdot r_l. \tag{6.9}$$

6.2 Description of the Algorithm

6.2.1 Monte Carlo Tree Search based Reinforcement Learning

RL and game theory can be used to develop optimization strategies for stochastic games. If considered a stochastic game, the problem of integrating intermittent, weather-dependent DERs on the grid can benefit from the developments in these areas. In conventional strategies, description of the system must be programmed in advance with sufficient prior knowledge. However, in a dynamic environment with stochastic behavior, the system itself changes over time making the optimization problem very hard to solve. In such situations, the optimization strategy can be developed by an agent through a learning process, without being explicitly programmed.

In a power system with high penetration of DERs, let S be a finite or infinite set of environment states. Each state $s \in S$ is a vector that refers to the current status of a DER unit in the search space. An agent may take an action $a \in A$ from a set of all possible actions A. The transition probability p determines the likelihood of the agent traversing from state s to s' under the joint action of all agents. In response to action a taken and state s traversed, the agent will receive an immediate local reward $r(s, a, s')$ [104]. Eventually, the learning objective of the agent is to maximize the discounted cumulative reward at each time step as follows

$$R(t) = r(t+1) + \gamma r(t+2) + \gamma^2 r(t+3) + \ldots, \tag{6.10}$$

where $\gamma \in [0, 1]$ is a discount factor expressing the effect of the current decision on the long-term reward. A small value of γ means that rewards in the near future are more important.

Applied to power systems, feasible regions with suitable energy resources can be identified using Monte Carlo tree search based Reinforcement Learning (MCTS-RL) [105]. This search algorithm provides the proposed approach with ability to navigate through the power network and gradually build experience.

The region feasible for power transactions with optimal power transfer trajectories to the DERs are determined using diffusion strategies illustrated in Figure 6.1. Each bus in a power network is modeled as a node in the MCTS graph [48]. Each edge stores a set of parameters: the state-action pair (s, a) and the visit count $N(s, a)$. A learned strategy is represented by a Q-value function that maps each state-action pair to a value estimating goodness of the action in the next

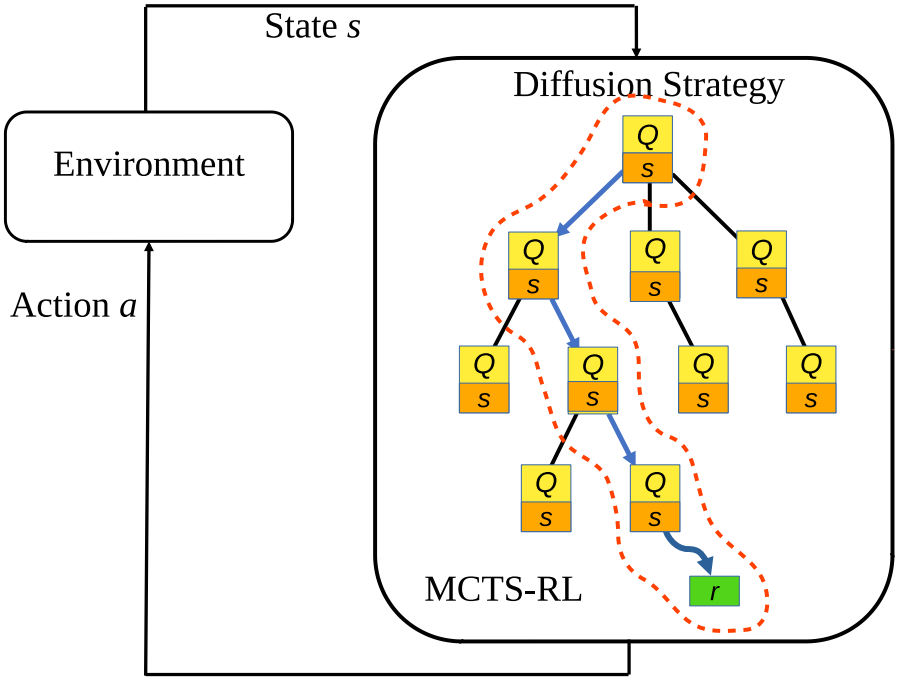


Figure 6.1: Proposed MCTS-RL algorithm.

state s'. The Q-value function is obtained as follows

$$Q(s_t, a_t) \quad \leftarrow \quad Q(s_t, a_t) \; + \; \alpha\Big[r_{t+1} \; + \; \gamma Q(s_{t+1}, a_{t+1}) \; - \; Q(s_t, a_t)\Big], \quad (6.11)$$

where α is the learning rate which controls the extent of the value function update.

The next joint action is selected by the ε-greedy policy

$$a = \begin{cases} \max Q(s, a) & \text{with probability } 1 - \varepsilon \\ \text{random } a \in A & \text{with probability } \varepsilon \end{cases} \quad (6.12)$$

where $\varepsilon \in [0, 1]$ is the exploration rate used to balance the exploration and exploitation policies during the process of learning the Q value function. This way, the state tree is randomly built up and the experience accumulated in each state is updated by random sampling and stored in the node states by a back propagation process.

To enhance the learning capability in terms of DER optimization in complex power systems, the described agent-based algorithm can be expanded to multi-agent case through the proposed MLFACC method. A theoretical framework of this method is introduced in the next section.

6.2.2 Multi Leader-Follower Actors under Centralized Critic

The proposed MLFACC algorithm relies on deep reinforcement learning using an actor-critic [104, 106]. Three actor networks train decentralized policies in a multi-agent framework, and share information using a centralized critic network. The main idea of using a critic network is to learn a centralized policy with an attention mechanism. In complex multi-agent environments, the attention mechanism has shown effective and scalable learning [107]. The intuition behind this idea is that the centralized critic can dynamically evaluate each agent's action; eventually, it sends attention to the agents to adjust their actions according to the environment need.

Another crucial approach to obtain the optimal variables of interest to accelerate the LP method is the leader-follower policy. The idea of the leader-follower game policy is inspired by Stackelberg game model [108]. In order to take an optimal action, it is necessary that a leader fully understand the environment and not only learns from its own actions but also the follower's actions. Typically, the leader acts first, then announces its action. At this point, the game rule allows the followers to make their decisions. In the proposed method, the roles of the players in the game change: if the number of agents is more than two, every follower agent can be a leader to the next agent. However, the first agent is always a leader, and the last agent is always a follower. Also, it is worth noting that the follower's action is estimated as a function of the leader's actions since the goal's reservation of the previous leader is already made. Thus, in this game, the leader uses a competitive policy, while the follower is expected to use a cooperative policy.

The main question that arises in this algorithm is how the agents learn from each other the optimal policies and get higher rewards. If an agent performs an action, it is based on the states mapped through the critic network. These states are always changing by the actions of the agent itself as well as the other agents. In other words, all agents should take optimal policies by their action probabilities as in $(\nabla \log \pi_\theta(a_t, s_t))$ of (6.13) in order to increase the return in the critic network in $(r_{t+1} + \gamma v(s_{t+1}, w_{t+1}))$ of the same equation. Since agents seek their own, unique goals, each agent has its own loss gradient $(\nabla_\theta J^{\mathcal{E}}_{\pi\theta})$ that is sent to the critic network to allow estimation of the policy probabilities advantage $(\hat{A}_t)$ as in (6.14), where index $\mathcal{E}$ refers to a particular agent under policy estimation. This way, the central critic can teach the agents based on the experience of the other agents and the state updates of the system.

$$\nabla_\theta J^{\mathcal{E}}_{\pi\theta} \quad \leftarrow \quad \nabla \log \pi_\theta(a^{\mathcal{E}}_t, s^{\mathcal{E}}_t)(r_{t+1} \; + \; \gamma \hat{v}(s_{t+1}, w_{t+1}) \; - \; \hat{v}(s_t, w_t)), \quad (6.13)$$

$$\nabla_\theta J^{\mathcal{E}}_{\pi\theta} \leftarrow \nabla_\theta \log \pi_\theta(a^{\mathcal{E}}_t, s^{\mathcal{E}}_t)\hat{A}_t. \quad (6.14)$$

To prevent the follower agents from seeking the same leader's policy trajectories, a tracing constraint $(\nabla_\theta \log \pi_\theta(a^{\mathcal{L}}_t, s^{\mathcal{L}}_t))$ is added to (6.14), as

$$\nabla_\theta J^{\mathcal{E}}_{\pi\theta} \quad \leftarrow \quad \nabla_\theta \log \pi_\theta(a^{\mathcal{E}}_t, s^{\mathcal{E}}_t)\hat{A}_t \; - \; \mu\big[\max\big(\nabla_\theta \log \pi_\theta(a^{\mathcal{L}}_t, s^{\mathcal{L}}_t), \Psi\big)\big], \quad (6.15)$$

where $\mathcal{L}$ is the index of all agent policies in a leader position, and μ is a lagrangian multiplayer. This constraint forces the follower agent that intends to choose the best goal to be right inferior to the leader's goal. However, this constraint may result in inefficient policies by the follower agents and slow learning. Since the leader plans its strategy to propel the followers to take actions in its favour, it may pick a trivial trajectory; consequently, the followers are constrained to choose other trajectories with even less importance. Hence, to relax this constraint at the beginning of the learning process, a relaxation factor (Ψ) is used. This factor is also an action probability that allows the followers to break the tracing constrained to a particular limit; once the leader finds a proper trajectory that leads to obtain a better DER price, the relaxation factor vanishes. At this point, just the first term of the maximization operator is valid.

This mechanism enables the proposed algorithm to identify the best group of DERs in a descending order, and without overlap. The pseudocode is presented as follows.

Algorithm 1 MLFACC algorithm

Initialize actors' weights; $(\theta_1, ..., \theta_N), \theta \in \mathbb{R}^n$
Initialize critic weight; (w), $w \in \mathbb{R}^m$
Initialize step size parameters: $\alpha > 0,\ \beta > 0$.
for *$t = 1$ to max episode length* **do**
  done = False
  while *not done* **do**
    for *agents $i = 1$ to N* **do**
      Choose action a_{t+1} based on probability:
      $\log \pi_\theta(a_t, s_t)$
      Receive observation $(s_{t+1}, s_t, r_{t+1}, done)$
    end
    Compute the TD error:
    $\delta \leftarrow (r_{t+1} + \gamma\hat{v}(s_{t+1}, w_{t+1}) - \hat{v}(s_t, w_t))$
    Compute the loss gradient for each agent by eq. (6.15)
    Update policy parameters for the actor networks:
    $\theta_{t+1} \leftarrow \theta_t + \alpha\nabla_\theta J^{\mathcal{E}}_{\pi\theta}$
    Update policy parameters for the critic network:
    $w_{t+1} \leftarrow w_t + \beta\delta\hat{v}_{(s,w)}$
  end
end

6.2.3 Deep RL-based Linear Programming

A power distribution network can be modeled by a directed tree graph $T(\Omega_G, \Omega_L) \subseteq (N, L)$. The nodes of the graph Ω_G, a subset of all network buses N, are linked by a set of distribution lines Ω_L, a subset of all network lines L. Node 0 is the starting point of the tree search, referred to as the root node j. In general, the root node can represent either a buyer or a seller of energy. Under the scenario considered in this study, the root node is specified as a buyer looking for the best seller(s). The remaining nodes are referred to as branch nodes. Each pair of adjacent branch nodes is connected by a branch line $l \in \Omega_L$. All nodes (except the leaf nodes) in this tree are parent nodes since they have a set of child nodes C_i linked by the branch lines. In addition, the child nodes may have connected DER units. i is the index of all buses that link to load bus j. Each line in L has an impedance $z_i = r_i + x_i$. Power injection from node i to node j is calculated using equation (6.3).

To minimize the cost of DER generation dispatch under system constraints, we propose a DOPF algorithm called Deep Reinforcement Learning-based Lin-

ear Programming (DRL-LP). In this model, LP is accelerated by selecting the optimal DERs through MLFACC, within the feasible region determined by MCTS-RL. Chazelle and Matousek [109] have analysed and estimated the computational complexity that describes the amount of time it takes to run LP by counting the number of input variables x and $g(x)$ constraints as follows

$$O(x)^{7x}(\log x)^{x} g(x), \tag{6.16}$$

where $O(.)$ denotes the time complexity. Thus the behavior of the LP complexity can be reduced by reducing the size of the input.

Based on the linearized power flow model described in section 6.1, the DRL-LP problem can be formulated as follows

$$\min \sum_{x \in \Omega_k} f_i(C_i^P), \tag{6.17}$$

$$g^{\text{eq}}(x) = \begin{cases} P_i - d_j = \sum_{i \in \Omega_k} P_{ij}^{\text{flow}} + P_{ij}^{\text{loss}} \, \forall j \in N, & (6.18) \\ \theta_{\text{ref}} = 0, & (6.19) \end{cases}$$

$$g^{\text{ineq}}(x) = \begin{cases} \underline{P}_{\text{DER}} \leqslant P_{\text{DER}} \leqslant \overline{P}_{\text{DER}} \forall \text{DER} \in \Omega_G, & (6.20) \\ \underline{P^{\text{flow}}}_{ij} \leqslant P_{ij}^{\text{flow}} \leqslant \overline{P_{ij}^{\text{flow}}} \, \forall l_{ij} \in L, & (6.21) \\ \underline{P^{\text{loss}}}_{ij} \leqslant P_{ij}^{\text{loss}} \leqslant \overline{P_{ij}^{\text{loss}}} \, \forall l_{ij} \in L, & (6.22) \\ \underline{V}_i \leqslant V_i \leqslant \overline{V}_i \quad \forall x \in N, & (6.23) \\ \text{LF}_{ij} \leqslant \phi \qquad \forall l_{ij} \in \Omega_L, & (6.24) \end{cases}$$

In this optimization problem, C_i^P is the optimal DER that is determined by MLFACC. It belongs to a node in the tree graph Ω_k delineated by MCTS as a feasible subset of the entire distribution network. The objective function aims to minimize the generation cost at node i, and implicitly minimizes the losses P_{ij}^{loss} of the line connecting nodes i and j. The set of all DER units within the feasible region is denoted Ω_G. Functions $g^{\text{eq}}(x)$ and $g^{\text{ineq}}(x)$ express, respectively, the equality and inequality constraints. The nodal balance power flow is restricted by constraint (6.18), where d_j is the power demand, while equation (6.19) holds the reference bus voltage angle at zero. Inequalities (6.20)–(6.22) express the upper and lower bounds of the power output of DER units, the power flows in the branches, and the power losses in the branches, respectively. The coupling constraint between region Ω_G and its neighboring regions is denoted ϕ. Based on

the concept of electric distance [97], ϕ represents the threshold value of LF of the line impedance between the load bus j and the cross-border buses separating region Ω_G from its neighbors. It can be considered a means to specify the borders of a feasible space of the DOPF problem among multiple regions. However, the network constraints must still be observed and communicated among clusters. Thus, there is a need for a central coordinator that can efficiently manage the information for multi-region systems. Form the implementation perspective, all information is sent to the central coordinator after the distributed optimizers are instantiated and load bus i spans solutions of its own OPF subproblems. The coordinator reconciles system state information for multiple clusters Ω_G.

6.3 Real-Time Balancing Electricity Market

To illustrate application of the proposed DOPF method using DRL-LP, we construct a distribution electricity market framework to facilitate effective integration of DERs into the electricity system. A central role in this framework is assumed by the distribution system operator (DSO) who facilitates DERs integration and delivers location services. It also provides real-time power balancing through dispatch of stochastic DERs and bidding of flexible loads.

The algorithm for balancing electricity market is executed every minute to accommodate (near) real-time power imbalances. Distribution locational marginal price (DLMP) differs from location to location due to the limits of the node voltages, line capacity, and network losses. This facilitates mitigation of over/under voltage and line congestion, and compensation of location-dependent network losses.

The main goal of the DSO is to maximize its economic benefit while providing the amount of power required by the balancing electricity market. The individual entities of the distribution system respond to specific price signals derived from the following DLMP equation

$$\text{DLMP}_i = \lambda_0^p + \lambda_0^p \cdot \sum_{i=1}^{N} \frac{\partial P_i^{\text{flow}}}{\partial P_i^{\text{DER}}} + \lambda_0^p \cdot \sum_{i=1}^{N} \frac{\partial P_i^{\text{loss}}}{\partial P_i^{\text{DER}}} + \lambda_0^p \cdot \sum_{i=1}^{N} \frac{\partial V_i}{\partial P_i^{\text{DER}}} \tag{6.25}$$

where λ_0^p is the active power exchange or the reference price. This is a known parameter that can be adjusted by the DSO. The three sum terms $\sum_{i=1}^{N} \partial P_i^{\text{flow}} / \partial P_i^{\text{DER}}$, $\sum_{i=1}^{N} \partial P_i^{\text{loss}} / \partial P_i^{\text{DER}}$, and $\sum_{i=1}^{N} \partial V_i / \partial P_i^{\text{DER}}$ are the total line flow factor, total system losses factor, and voltage deviation factor, respectively. All three factors are cal-

culated with respect to DER power injection, P_i^{DER}, from the constraint equations (6.21), (6.22), and (6.23). DLMP works as a price coordinator to ensure that any power imbalance in the system can be fully offset and objectives of all participating entities can be optimized simultaneously. This coordinated operation model is designed to include all required objective functions and system constraints.

6.4 Results and Discussion

To demonstrate the proposed DRL-LP algorithm, the IEEE 69-bus test system [110] is considered as a case study. To examine the leverage provided by the RL agents in the distributed optimization sub-problems, a search tree is progressively build using MCTS-RL. This tree is a randomly biased sequence of actions applied by an RL agent to a given series of states until a predefined coupling constraint ϕ is reached. This way, the feasible region suitable for power transactions is obtained based on the concept of electric distance. To exclude buses with high losses (4 and 5), the feasible region A is selected as shown in Figure 6.2. The DER units within this region are connected to 9 buses. MLFACC determines the optimal DERs, consequently reducing the number of variables and their conditional statements to accelerate the LP process. As shown in Figure 6.3, MLFACC obtains the best DERs as: 54, 56, and 66 by agents 1, 2, and 3. respectively. It can be seen that most agents converged after about 2500 episodes

The following sections analyze the optimization process for two scenarios: an unconstrained system with a single multi-agent system (cluster A), and a constrained system with a group of multi-agent systems (clusters A and B).

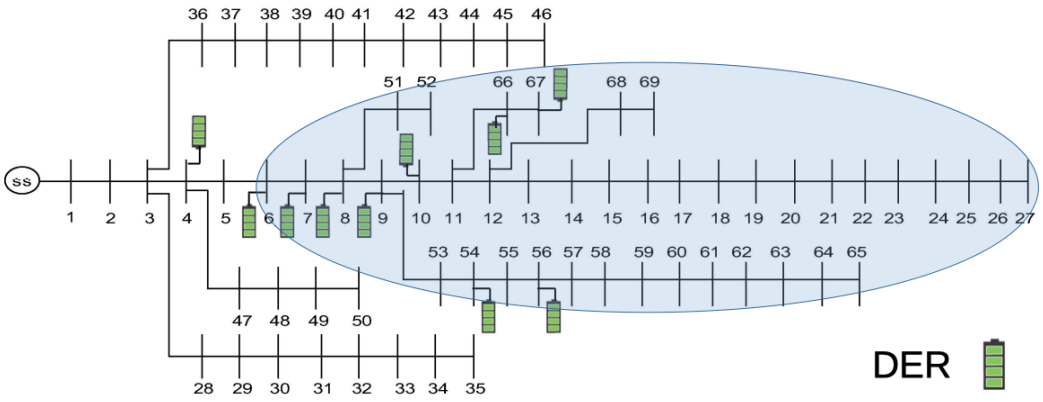


Figure 6.2: IEEE 69 bus distribution system with region A.

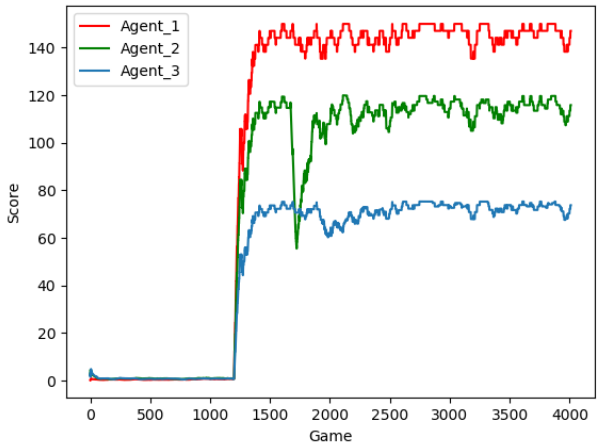


Figure 6.3: The learning simulation of the MLFACC for region A buses.

6.4.1 Single Cluster Architecture

For simplicity, this analysis considers a single dynamic real-time load located at bus 15 (see Figure. 6.4). This load is also considered an extra load to be balanced by the generation of the main feeder. The load data has been extracted from a residential community in Edmonton, Alberta, Canada, and scaled to the transformer level. Loads of the remaining buses of the circuit are based on the original static load data. Real-time optimization of the DER generation in region A is performed on a 1-minute basis after the MLFACC obtains the best DERs on buses 54, 56, and 66, as shown in Figure 6.5. It can be seen that, when DER at bus 56 reached its limit (30kW), the algorithm switched to DER at bus 54. Should demand be concentrated on a few DERs, the corresponding segments of distribution lines can become overloaded. To mitigate the occurrence of overloaded lines due to DER optimization, the algorithm also tracks its impact on the distribution circuit. The optimization process mainly relies on the variation of DER benefits stemming from the reduction of the active power flow losses within the circuit. Based on their location, the power the DERs generate usually flows in the direction opposite to the main power stream of the feeder. This causes a reduction of the total losses of the feeder power flow. The variation of losses also causes considerable differences of DLMP, especially when the DER generation levels approach their maximum energy export capacity. More details about the DLMP pricing are provided in the next section describing the multi-agent optimization in a multi-cluster architecture.

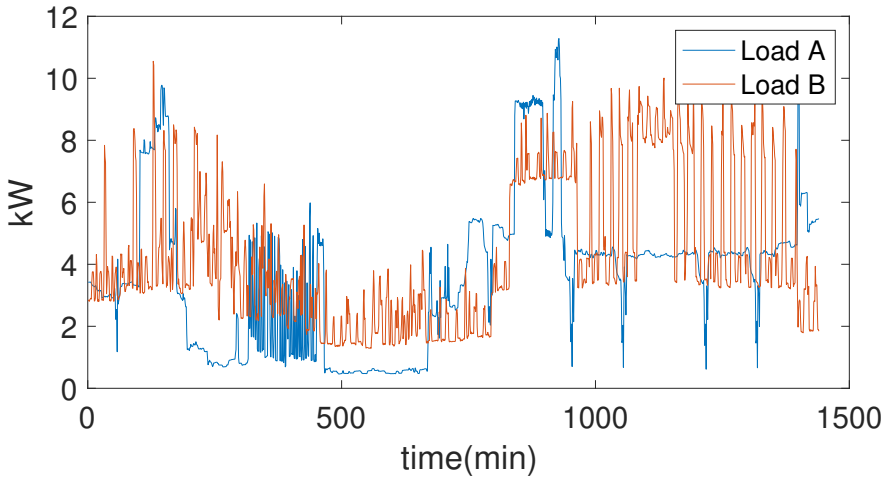


Figure 6.4: The real-time load profiles at bus 15 and 67 at region A and B respectively.

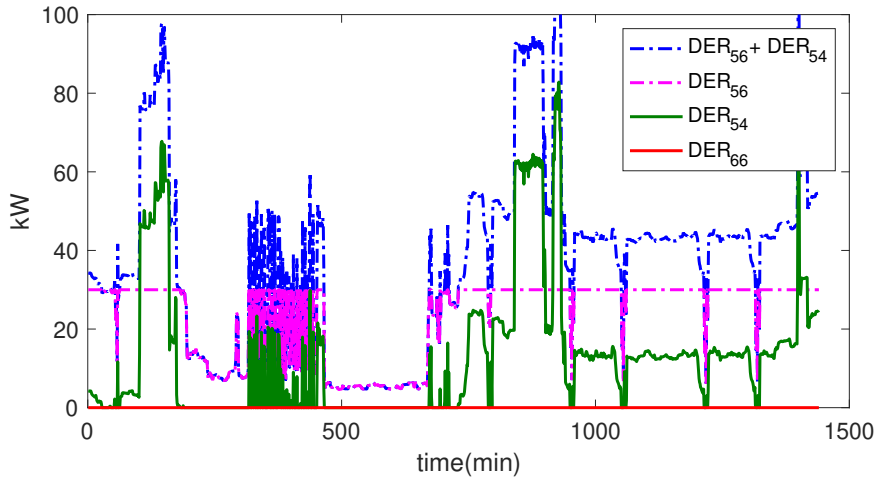


Figure 6.5: The real-time optimization of DER generation within region A, controlled by agent A, without considering the contingency constraint.

6.4.2 Multi-Cluster Architecture

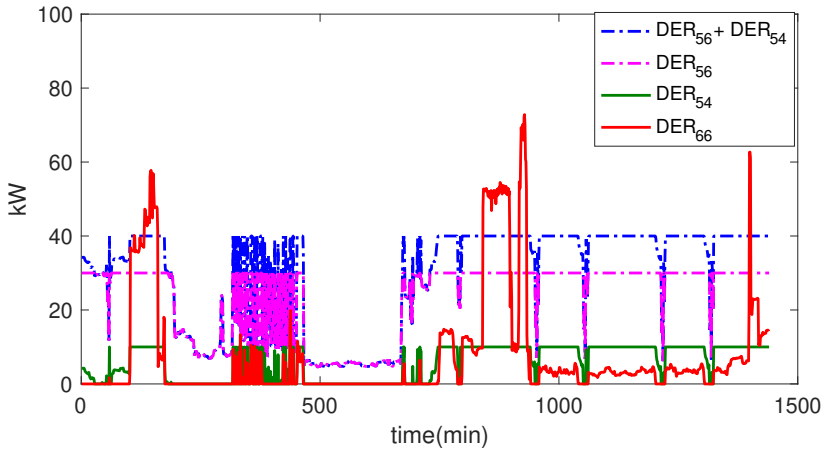
To demonstrate the distributed optimization framework, the multi-cluster interaction is analyzed in this section. Assume that another region (B) is settled and managed by a second independent group of agents at cluster B. This region has only two buses, 66 and 67, both with DERs. The DER connected at bus 66 is shared between the two regions A and B (DER AB). Similar to region A, region B also has one dynamic load, located at bus 67. Each cluster attempts to opti-

mize the DER generation levels within its region while tracking its own impact on the circuit. Communicating through the central coordinator, each cluster can also track the impact of the other cluster(s). This way, they can update the estimates of the next RL states and, accordingly, they can change their optimization policies to maintain the required security level of the circuit.

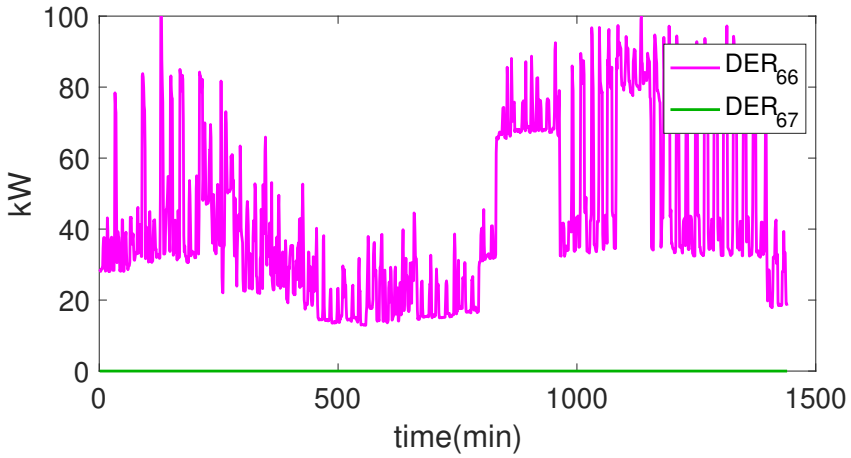
The simulation of the above scenario using the proposed DRL-LP model can be demonstrated as follows. Line connecting buses 10-11, with an original flow of 740 kW, is assumed to have two cases of flow limits: 780 kW and 740 kW. In the first case (I), the limit of 780 kW allows $\Delta P = 40$ kW of extra power to be carried by this line. When the load of bus 15 increases such that the flow on the line 10-11 exceeds 40 kW, the flow congestion occurs at this line, as shown in Figure 6.6a. In such scenario, the amount of generation of the DERs that compromise network security is reduced. Instead, another DER that does not influence system security while offering an acceptable price is called by cluster A. Hence, the new marginal DER is the common generator at bus 66. In addition, in this case, the generation limit of the common DER is sufficient for both regions A and B. Cluster B uses only one DER, as shown in Figure 6.6b.

The primary goal of this step is to change the system flow to prevent congestion, while providing energy to the load at an acceptable price. However, this leads to a price step-change. The price change to avoid 1 kW of line congestion is called congestion DLMP. Similarly, the change to avoid line losses is called losses DLMP. The daily values of congestion and losses DLMP in region A are 4¢ and -31.36¢, respectively. In region B, these values are 0¢ and -30¢, respectively. The negative sign signifies the reduction of losses due to DER generation. The DLMP values for both regions A an B are shown in Figure 6.6c.

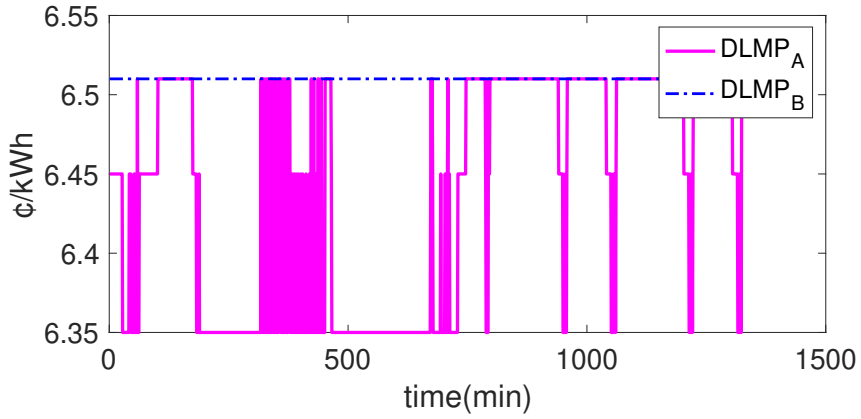
In the second case (II), shown in Figure 6.7a, the flow limit is 740 kW. In this case, the common DER (connected at bus 66) reaches its generation limit. As shown in Figure 6.7b, agent B calls the second DER with a relatively higher price, connected at bus 67, to compensate the generation demand. The daily values of the congestion and losses DLMP in region A are, respectively, 494.23¢ and -5¢. In region B, these values are 0¢ and 29.8¢, respectively. The DLMP values for both regions A an B are shown in Figure 6.7c. Finally, the total cost of the system generation for case II is shown in Figure 6.8. It should be noted that the common DER (66) feeds both regions and is affected by interactions of both clusters. The interactions are terminated when the clusters reach agreement on the amounts and prices of power supplied by the common DER. This agreement is known as consensus dynamics [111].



(a)

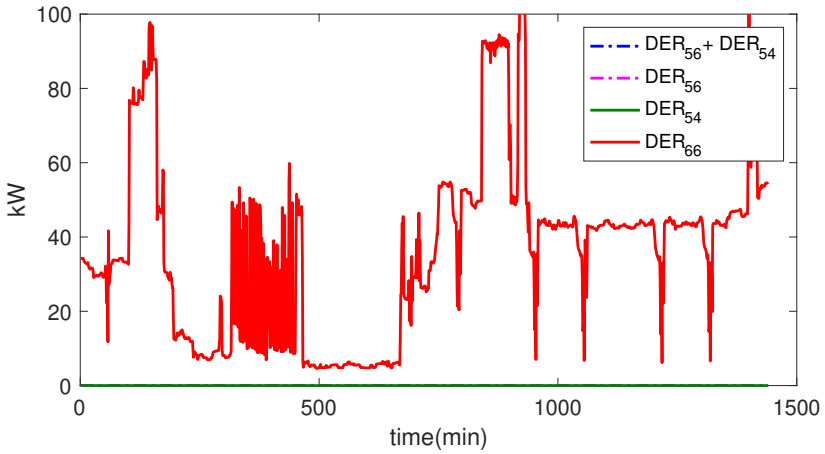


(b)

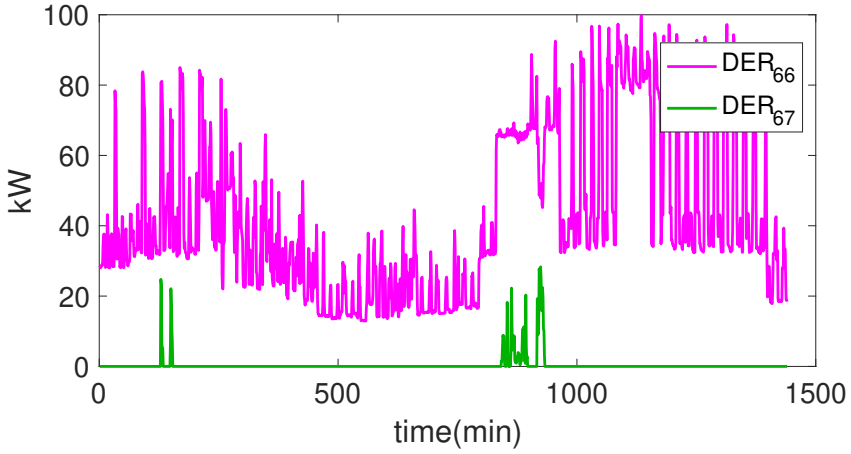


(c)

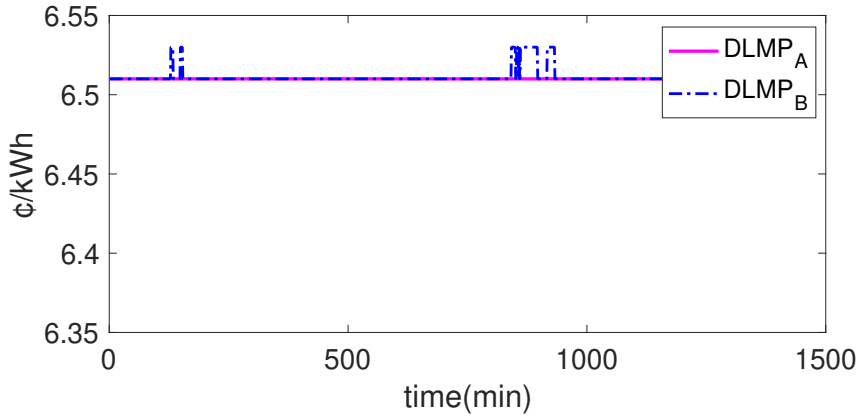
Figure 6.6: The real-time optimization of the DER generations, with the case I contingency constraint: (a) Region A, (b) Region B, (c) The total DLMP prices per 1 kW of both regions A and B.



(a)



(b)



(c)

Figure 6.7: The real-time optimization of the DER generations, with the case II contingency constraint: (a) Region A, (b) Region B, (c) The total DLMP prices per 1 kW of both regions A and B.

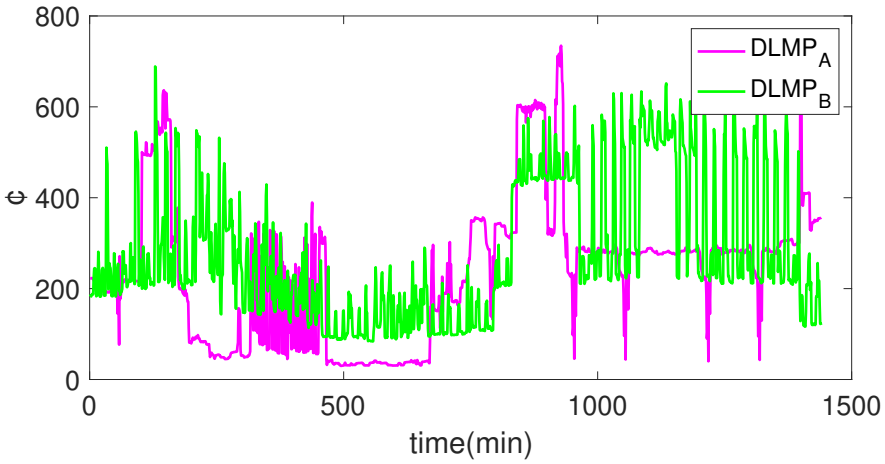


Figure 6.8: The total generation costs for both regions A and B.

The two scenarios confirm that the proposed model can successfully consider the system constraints and provide an effective power market solution without jeopardizing system security. The IEEE 69-bus system considered in this study experiences only minor voltage control issues. This is due to the nature of the load connected at nodes 15 and 67. The losses are also relatively low as they are proportional to the network size.

6.5 Final Remarks

This research introduces a novel approach for OPF in distribution systems with high penetration of DERs. Using modern methods of artificial intelligence, the proposed approach facilitates accurate OPF calculation while reducing its computational burden. The proposed method is based on an effective combination of Monte Carlo tree search (MCTS) and reinforcement learning (RL). Through navigation steps of a diffusion strategy, MCTS-RL generates a set of optimal paths to the most suitable nodes in the grid. Only the selected paths are then considered by the optimizer applied on the linearized problem, thus guaranteeing convergence. This way, the combined algorithm mitigates the challenges associated with stochasticity of DERs while addressing the problem of dimensionality faced by conventional optimization techniques using the proposed deep learning-based approach (MLFACC). Furthermore, the multi-agent nature of the proposed approach allows direct application of DOPF in systems with multiple interacting entities. The presented simulation results clearly demonstrate the ef-

fectiveness of the proposed method to solve the distributed economic dispatch problem while maintaining the system security limits. In addition, the proposed methodology is suitable for large and complex networks.

Chapter 7

Conclusions

This thesis highlights the necessity for identifying to impact of the distributed DER units on the distribution systems, also developing intelligent techniques to resolve these issues in the most cost-effectively way. The presented research is primarily concerned with power distribution networks with high penetration of stochastic DER units. In addition to presenting a methodology for assessment of DER-related issues in distribution networks, this thesis successfully resolves two critical problems: mitigation of the voltage rise over the specified limit and design of a fully distributed OPF algorithm.

The overvoltage mitigation problem is solved using a newly designed algorithm that coordinates and controls the state of charge of multiple BESS using techniques of Monte Carlo tree search, reinforcement learning (RL), and model predictive control (MPC). In addition, this coordination control strategy also optimizes the battery schedules from the perspective of energy prices and losses.

The distributed OPF procedure accelerates the optimization process for large-scale systems where the use of centralized fashion is not possible. The algorithm identifies clusters of related distribution lines and coordinates partial OPF solutions through agent-based approach. It results in an extension of conventional linear programming (LP) algorithm into it RL-based counterpart for effective balancing of power transactions in large distribution circuits.

7.1 Contributions

Referring to the objectives stated in section 1.1, the contributions of this thesis can be organized in the following three groups:

1. Voltage violation assessment in large distribution systems with high DER

penetration:

- An assessment of three circuits of varying size and complexity has been conducted. The PVHC of the three distribution networks (Fort McMurray, Lloydminster, and Drumheller) was determined as 10%, 60%, and 70%, respectively. The developed Monte Carlo simulation-based probabilistic power flow is effective, even when dealing with large and complex networks.
- The sensitivity of distribution system PVHC has been analyzed, and relevant influential factors identified: circuit load conditions, location and size of PV installations, reactive power of the circuit, and short circuit level of the system.

2. DER impact mitigation using energy management and coordination methods:

   - A new method, called cooperative state of charge control (CSOCC), has been designed to deal with large and complex distribution networks. It divides an entire network into several smaller networks to facilitate indetification of regions impacted by overvoltage and regions that can assist in its mitigation. The developed network structure facilitates sharing of energy between neighboring BESS units to prevent the voltage problems in real-time; this would be either impossible or very expensive using independent BESS controllers.
   - The proposed intelligent method, based on MCTS-RL, builds a list of feasible regions. This reduces the computational burden associated with a large real network. It also coordinates the distributed batteries from the other areas to provide voltage regulation and adapt to any changes.
   - A decentralized controller is designed using MPC. It not only controls the battery SOC to mitigate the overvoltage, but also minimizes the system operation cost through optimal battery scheduling.

3. Distributed optimal power flow for distribution systems with stochastic DER units:

   - A fully distributed OPF has been proposed to capture all physical realities of distribution networks and achieve the optimal power trans-

actions between network entities. This algorithm is based on the intelligent method of MCTS-RL that navigates through the circuit to determine feasible regions based on the concept of electric distance.

- Reinforcement learning is used to effectively solve stochastic sequential decision-making problem without relying on prior system information. It determines the most suitable DERs in a complex environment. As a result, it accelerates the linear programming optimization method by reducing the number of its arithmetic operators and their conditional statements.
- The proposed distributed OPF is designed to achieve interactions between independent entities cooperatively, while maintaining the system security constraints such as the congested lines, without upgrading the capacity of grid infrastructure.

7.2 Future Work

There are many opportunities to further research topics that have been addressed in this thesis. The studies proposed in this thesis are bounded to particular methodologies and parameterizations that help to manage and further extend the scope of the thesis research.

High penetration levels of DERs, such as PV systems, into the existing low voltage (LV) distribution grids can cause a number of issues. The main problem, considered in this research is the the change of grid voltage levels due to reverse power flow in distribution feeders, especially at light loads. However, due to distinct cable types that are used in a feeder's circuit, the impact on line loading is notably more clustered then the impact on the bus voltages. This bring concerns about the line thermal limits, especially when the analysis is applied to larger systems where high levels of reverse power flow are likely to occur.

For most of residential and commercial clients, single-phase connections of grid-tied inverters are used. In this type of connection, PV systems are connected randomly along with the phases. However, due to the PV systems are relatively large unbalanced customer loading and voltages of the phases are the most pronounced. Electric vehicle chargers also fall under this category. Unbalanced phase loading and voltages may result in unwanted tripping of the converters. Moreover, they significantly affect the power quality of the distribution network. Therefore, issues related to power system protection and power

quality are promising topics for future research.

In the proposed DOPF methodology, a medium-sized circuit is considered. It would be interesting to apply the methodology to large, complex networks where the voltage control issues and high power losses caused by DER penetration are more noticeable. Finally, the agent-based DOPF can be extended to a multi-agent architecture. This would improve resiliency of the calculations when communication fault occur, or when individual local controllers go offline.

Bibliography

[1] M. Rylander, J. Smith, D. Lewis, and S. Steffel, "Voltage impacts from distributed photovoltaics on two distribution feeders," in *2013 IEEE Power Energy Society General Meeting*, July 2013, pp. 1–5.

[2] Z. Zhang, K. Wang, Q. Zhao, Y. Tian, and L. Li, "Positive sequence current phase-based improved reverse-power protection and a pv hosting capacity assessment method for spot networks," *IEEE Access*, vol. 8, pp. 42 529–42 539, 2020.

[3] A. Navarro-Espinosa and L. F. Ochoa, "Increasing the pv hosting capacity of lv networks: Oltc-fitted transformers vs. reinforcements," in *2015 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, 2015, pp. 1–5.

[4] S. Mohan, S. Hasan, Y. Gebremariam, and R. K. Varma, "Increasing hosting capacity of pv solar systems using smart inverter volt-var control," in *2018 20th National Power Systems Conference (NPSC)*, 2018, pp. 1–6.

[5] C. Long and L. F. Ochoa, "Voltage control of pv-rich lv networks: Oltc-fitted transformer and capacitor banks," *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 4016–4025, 2016.

[6] A. Ballanti, F. Pilo, A. Navarro-Espinosa, and L. F. Ochoa, "Assessing the benefits of pv var absorption on the hosting capacity of lv feeders," in *IEEE PES ISGT Europe 2013*, Oct 2013, pp. 1–5.

[7] J. Seuss, M. J. Reno, R. J. Broderick, and S. Grijalva, "Improving distribution network pv hosting capacity via smart inverter reactive power support," in *2015 IEEE Power Energy Society General Meeting*, July 2015, pp. 1–5.

[8] S. Hashemi, J. Ostergaard, and G. Yang, "Effect of reactive power management of pv inverters on need for energy storage," in *2013 IEEE 39th Photovoltaic Specialists Conference (PVSC)*, June 2013, pp. 2304–2308.

[9] O. C. Rascon, B. Schachler, J. Bühler, M. Resch, and A. Sumper, "Increasing the hosting capacity of distribution grids by implementing residential pv storage systems and reactive power control," in *2016 13th International Conference on the European Energy Market (EEM)*, June 2016, pp. 1–5.

[10] S. Hashemi and J. Ostergaard, "Efficient control of energy storage for increasing the pv hosting capacity of lv grids," in *2017 IEEE Manchester PowerTech*, June 2017, pp. 1–1.

[11] M. N. Kabir, Y. Mishra, G. Ledwich, Z. Y. Dong, and K. P. Wong, "Coordinated control of grid-connected photovoltaic reactive power and battery energy storage systems to improve the voltage profile of a residential distribution feeder," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 2, pp. 967–977, May 2014.

[12] L. Wang, D. H. Liang, A. F. Crossland, P. C. Taylor, D. Jones, and N. S. Wade, "Coordination of multiple energy storage units in a low-voltage distribution network," *IEEE Transactions on Smart Grid*, vol. 6, no. 6, pp. 2906–2918, Nov 2015.

[13] Y. Wang, K. T. Tan, X. Y. Peng, and P. L. So, "Coordinated control of distributed energy-storage systems for voltage regulation in distribution networks," *IEEE Transactions on Power Delivery*, vol. 31, no. 3, pp. 1132–1141, June 2016.

[14] M. Zeraati, M. E. H. Golshan, and J. Guerrero, "Distributed control of battery energy storage systems for voltage regulation in distribution networks with high pv penetration," *IEEE Transactions on Smart Grid*, vol. PP, no. 99, pp. 1–1, 2016.

[15] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep reinforcement learning: A brief survey," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 26–38, Nov 2017.

[16] A. R. Sharma and P. Kaushik, "Literature survey of statistical, deep and reinforcement learning in natural language processing," in *2017 Interna-*

tional Conference on Computing, Communication and Automation (ICCCA), May 2017, pp. 350–354.

[17] R. S. Sutton and A. G. Barto, *Temporal-Difference Learning*. MIT Press, 2018. [Online]. Available: https://www.andrew.cmu.edu/course/10-703/textbook/BartoSutton.pdf

[18] W. Liu, P. Zhuang, H. Liang, J. Peng, and Z. Huang, "Distributed economic dispatch in microgrids based on cooperative reinforcement learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 6, pp. 2192–2203, June 2018.

[19] P. Hasanpor Divshali and L. Söder, "Improving hosting capacity of rooftop pvs by quadratic control of an lv-central bss," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 919–927, Jan 2019.

[20] C. Lin and S. Lin, "Distributed optimal power flow with discrete control variables of large distributed power systems," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1383–1392, 2008.

[21] R. Tonkoski, D. Turcotte, and T. H. M. EL-Fouly, "Impact of high pv penetration on voltage profiles in residential neighborhoods," *IEEE Transactions on Sustainable Energy*, vol. 3, no. 3, pp. 518–527, July 2012.

[22] K. Coogan, M. J. Reno, S. Grijalva, and R. J. Broderick, "Locational dependence of pv hosting capacity correlated with feeder load," in *2014 IEEE PES T D Conference and Exposition*, April 2014, pp. 1–5.

[23] M. Rylander, J. Smith, and W. Sunderman, "Streamlined method for determining distribution system hosting capacity," *IEEE Transactions on Industry Applications*, vol. 52, no. 1, pp. 105–111, Jan 2016.

[24] A. Dubey, S. Santoso, and A. Maitra, "Understanding photovoltaic hosting capacity of distribution circuits," in *2015 IEEE Power Energy Society General Meeting*, July 2015, pp. 1–5.

[25] S. Hashemi, J. Ostergaard, T. Degner, R. Brandl, and W. Heckmann, "Efficient control of active transformers for increasing the pv hosting capacity of lv grids," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 1, pp. 270–277, Feb 2017.

[26] T. Stetz, F. Marten, and M. Braun, "Improved low voltage grid-integration of photovoltaic systems in germany," in *2013 IEEE Power Energy Society General Meeting*, July 2013, pp. 1–1.

[27] H. Xia, Q. Li, R. Xu, T. Chen, J. Wang, M. A. S. Hassan, and M. Chen, "Distributed control method for economic dispatch in islanded microgrids with renewable energy sources," *IEEE Access*, vol. 6, pp. 21 802–21 811, 2018.

[28] Q. Li, C. Peng, M. Wang, M. Chen, J. M. Guerrero, and D. Abbott, "Distributed secondary control and management of islanded microgrids via dynamic weights," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 2196–2207, March 2019.

[29] Q. Li, F. Chen, M. Chen, J. M. Guerrero, and D. Abbott, "Agent-based decentralized control method for islanded microgrids," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 637–649, March 2016.

[30] F. Guo, C. Wen, J. Mao, and Y. Song, "Distributed economic dispatch for smart grids with random wind power," *IEEE Transactions on Smart Grid*, vol. 7, no. 3, pp. 1572–1583, May 2016.

[31] M. Střelec, K. Macek, and A. Abate, "Modeling and simulation of a microgrid as a stochastic hybrid system," in *2012 3rd IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*, Oct 2012, pp. 1–9.

[32] J. Wu and X. Guan, "Coordinated multi-microgrids optimal control algorithm for smart distribution management system," *IEEE Transactions on Smart Grid*, vol. 4, no. 4, pp. 2174–2181, Dec 2013.

[33] G. Obando, N. Quijano, and N. Rakoto-Ravalontsalama, "Distributed resource allocation over stochastic networks: An application in smart grids," in *2015 IEEE 2nd Colombian Conference on Automatic Control (CCAC)*, Oct 2015, pp. 1–6.

[34] A. Saad, T. Youssef, A. T. Elsayed, A. Amin, O. H. Abdalla, and O. Mohammed, "Data-centric hierarchical distributed model predictive control for smart grid energy management," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 7, pp. 4086–4098, July 2019.

[35] E. R. Stephens, D. B. Smith, and A. Mahanti, "Game theoretic model predictive control for distributed energy demand-side management," *IEEE Transactions on Smart Grid*, vol. 6, no. 3, pp. 1394–1402, May 2015.

[36] Y. Zheng, S. Li, and R. Tan, "Distributed model predictive control for on-connected microgrid power management," *IEEE Transactions on Control Systems Technology*, vol. 26, no. 3, pp. 1028–1039, May 2018.

[37] I. Ranaweera and O. M. Midtgard, "Centralized control of energy storages for voltage support in low-voltage distribution grids," in *2016 IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC)*, June 2016, pp. 1–6.

[38] J. Cai, C. Chen, S. Duan, and D. Yang, "Centralized control of large capacity parallel connected power conditioning system for battery energy storage system in microgrid," in *2014 IEEE Energy Conversion Congress and Exposition (ECCE)*, Sept 2014, pp. 409–413.

[39] K. Worthmann, C. M. Kellett, P. Braun, L. Grüne, and S. R. Weller, "Distributed and decentralized control of residential energy systems incorporating battery storage," *IEEE Transactions on Smart Grid*, vol. 6, no. 4, pp. 1914–1923, July 2015.

[40] Q. Xu, J. Xiao, P. Wang, X. Pan, and C. Wen, "A decentralized control strategy for autonomous transient power sharing and state-of-charge recovery in hybrid energy storage systems," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 4, pp. 1443–1452, Oct 2017.

[41] T. Masuta and A. Yokoyama, "Atc enhancement considering transient stability based on optimal power flow control by upfc," in *2006 International Conference on Power System Technology*, 2006, pp. 1–6.

[42] Akihiro Wada and Teruhisa Kumano, "Fast estimation of transient stability constrained atc by relevance vector machine," in *2008 IEEE 2nd International Power and Energy Conference*, 2008, pp. 89–93.

[43] B. Marisekar and P. L. Somasundaram, "Computation of available transfer capability(atc) in the open access transmission system(oats) for various uncertainity conditions," in *2015 International Conference on Circuits, Power and Computing Technologies [ICCPCT-2015]*, 2015, pp. 1–5.

[44] H. Mokhlis, A. Shahriari, and J. Laghari, "Fast and accurate second order load flow method based on fixed jacobian matrix," *Applied Mathematics and Computation*, vol. 269, pp. 584 – 593, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0096300315009947

[45] C. Fu, X. Ren, Y. Yang, Y. Xia, and W. Deng, "An interval precise integration method for transient unbalance response analysis of rotor system with uncertainty," *Mechanical Systems and Signal Processing*, vol. 107, pp. 137 – 148, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0888327018300396

[46] S. N. K and A. Kumar, "Distribution system nodal pricing analysis with realistic zip load and variable wind power source," in *2015 Annual IEEE India Conference (INDICON)*, Dec 2015, pp. 1–6.

[47] E. A. Farsani, H. A. Abyaneh, M. Abedi, and S. H. Hosseinian, "A novel policy for lmp calculation in distribution networks based on loss and emission reduction allocation using nucleolus theory," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 143–152, Jan 2016.

[48] C. B. Browne, E. Powley, D. Whitehouse, S. M. Lucas, P. I. Cowling, P. Rohlfshagen, S. Tavener, D. Perez, S. Samothrakis, and S. Colton, "A survey of monte carlo tree search methods," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 4, no. 1, pp. 1–43, March 2012.

[49] M. Rylander, J. Smith, D. Lewis, and S. Steffel, "Voltage impacts from distributed photovoltaics on two distribution feeders," in *2013 IEEE Power Energy Society General Meeting*, July 2013, pp. 1–5.

[50] N. C. Tang and G. W. Chang, "A stochastic approach for determining pv hosting capacity of a distribution feeder considering voltage quality constraints," in *2018 18th International Conference on Harmonics and Quality of Power (ICHQP)*, May 2018, pp. 1–5.

[51] S. K. Singh and N. Bansal, "Output impedance as figure of merit to predict transient performance for embedded linear voltage regulators," in *2014 27th International Conference on VLSI Design and 2014 13th International Conference on Embedded Systems*, Jan 2014, pp. 516–521.

[52] M. Rylander, J. Smith, W. Sunderman, D. Smith, and J. Glass, "Application of new method for distribution-wide assessment of distributed energy resources," in *2016 IEEE/PES Transmission and Distribution Conference and Exposition (T D)*, May 2016, pp. 1–5.

[53] M. J. Reno, K. Coogan, S. Grijalva, R. J. Broderick, and J. E. Quiroz, "Pv interconnection risk analysis through distribution system impact signatures

and feeder zones," in *2014 IEEE PES General Meeting | Conference Exposition*, July 2014, pp. 1–5.

[54] J. E. Quiroz, M. J. Reno, and R. J. Broderick, "Pv-induced low voltage and mitigation options," in *2015 IEEE 42nd Photovoltaic Specialist Conference (PVSC)*, June 2015, pp. 1–6.

[55] A. Ballanti and L. F. Ochoa, "On the integrated pv hosting capacity of mv and lv distribution networks," in *2015 IEEE PES Innovative Smart Grid Technologies Latin America (ISGT LATAM)*, Oct 2015, pp. 366–370.

[56] Y. Liu, Y. Tai, C. Huang, H. Su, P. Lan, and M. Hsieh, "Assessment of the pv hosting capacity for the medium-voltage 11.4 kv distribution feeder," in *2018 IEEE International Conference on Applied System Invention (ICASI)*, April 2018, pp. 381–384.

[57] Z. Ren, K. Wang, W. Li, L. Jin, and Y. Dai, "Probabilistic power flow analysis of power systems incorporating tidal current generation," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 3, pp. 1195–1203, July 2017.

[58] A. Kazemdehdashti, M. Mohammadi, and A. R. Seifi, "The generalized cross-entropy method in probabilistic optimal power flow," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 5738–5748, Sep. 2018.

[59] J. Li, B. Zhang, and Y. Liu, "Data mining in nonlinear probabilistic load flow based on monte carlo simulation," in *2009 First International Conference on Information Science and Engineering*, Dec 2009, pp. 833–836.

[60] D. Cheng, B. A. Mather, R. Seguin, J. Hambrick, and R. P. Broadwater, "Photovoltaic (pv) impact assessment for very high penetration levels," *IEEE Journal of Photovoltaics*, vol. 6, no. 1, pp. 295–300, Jan 2016.

[61] E. Mulenga and M. H. J. Bollen, "Impact of service and feeder cable upgrade on hosting capacity for single phase connected photovoltaics," in *2018 18th International Conference on Harmonics and Quality of Power (ICHQP)*, May 2018, pp. 1–6.

[62] Y. Mi, C. Ma, S. Wang, C. Bi, Y. Zhu, and H. Zhang, "Reactive power control of an isolated wind-diesel hybrid power system based on svc by using sliding mode control," in *2016 35th Chinese Control Conference (CCC)*, July 2016, pp. 9882–9887.

[63] "Distribution Test 33-Bus Feeder," http://www.ece.ubc.ca/~hameda/download_files/.

[64] C. I. T. Inc, "Distribution System Analysis," www.cyme.com/software/cymdist/, 2017.

[65] S. Lakshmi and S. Ganguly, "Modelling and allocation planning of voltage-sourced converters to improve the rooftop pv hosting capacity and energy efficiency of distribution networks," *IET Generation, Transmission Distribution*, vol. 12, no. 20, pp. 4462–4471, 2018.

[66] P. Wang, D. H. Liang, J. Yi, P. F. Lyons, P. J. Davison, and P. C. Taylor, "Integrating electrical energy storage into coordinated voltage control schemes for distribution networks," *IEEE Transactions on Smart Grid*, vol. 5, no. 2, pp. 1018–1032, March 2014.

[67] H. Xin, Y. Liu, Z. Qu, and D. Gan, "Distributed control and generation estimation method for integrating high-density photovoltaic systems," *IEEE Transactions on Energy Conversion*, vol. 29, no. 4, pp. 988–996, Dec 2014.

[68] T. Morstyn, M. Momayyezan, B. Hredzak, and V. G. Agelidis, "Distributed control for state-of-charge balancing between the modules of a reconfigurable battery energy storage system," *IEEE Transactions on Power Electronics*, vol. 31, no. 11, pp. 7986–7995, Nov 2016.

[69] Y. Xu, W. Zhang, G. Hug, S. Kar, and Z. Li, "Cooperative control of distributed energy storage systems in a microgrid," *IEEE Transactions on Smart Grid*, vol. 6, no. 1, pp. 238–248, Jan 2015.

[70] S. K. Kim, J. Y. Kim, K. H. Cho, and G. Byeon, "Optimal operation control for multiple besss of a large-scale customer under time-based pricing," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 803–816, Jan 2018.

[71] I. Lampropoulos, P. Garoufalis, P. P. J. van den Bosch, and W. L. Kling, "Hierarchical predictive control scheme for distributed energy storage integrated with residential demand and photovoltaic generation," *IET Generation, Transmission Distribution*, vol. 9, no. 15, pp. 2319–2327, 2015.

[72] E. Mayhorn, L. Xie, and K. Butler-Purry, "Multi-time scale coordination of distributed energy resources in isolated power systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 2, pp. 998–1005, March 2017.

[73] A. K. Verma, H. B. Gooi, A. Ukil, N. R. Tummuru, and S. K. Kollimalla, "Microgrid frequency stabilization using model predictive controller," in *2016 IEEE PES Transmission Distribution Conference and Exposition-Latin America (PES T D-LA)*, Sept 2016, pp. 1–6.

[74] C. Ju and P. Wang, "Energy management system for microgrids including batteries with degradation costs," in *2016 IEEE International Conference on Power System Technology (POWERCON)*, Sept 2016, pp. 1–6.

[75] B. Zhu, H. Tazvinga, and X. Xia, "Switched model predictive control for energy dispatching of a photovoltaic-diesel-battery hybrid power system," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 3, pp. 1229–1236, 2015.

[76] O. Gomozov, J. P. F. Trovão, X. Kestelyn, and M. R. Dubois, "Adaptive energy management system based on a real-time model predictive control with nonuniform sampling time for multiple energy storage electric vehicle," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 7, pp. 5520–5530, July 2017.

[77] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, "Short-term residential load forecasting based on lstm recurrent neural network," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 841–851, Jan 2019.

[78] S. Jothibasu and S. Santoso, "Sensitivity analysis of photovoltaic hosting capacity of distribution circuits," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, July 2016, pp. 1–5.

[79] S. Wang, Q. Liu, and X. Ji, "A fast sensitivity method for determining line loss and node voltages in active distribution network," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 1148–1150, Jan 2018.

[80] M. C. Kisacikoglu, B. Ozpineci, and L. M. Tolbert, "Reactive power operation analysis of a single-phase ev/phev bidirectional battery charger," in *8th International Conference on Power Electronics - ECCE Asia*, May 2011, pp. 585–592.

[81] "Alberta Electric System Operator," http://ets.aeso.ca/, 2018.

[82] M. S. Taha, H. H. Abdeltawab, and Y. A. I. Mohamed, "An online energy management system for a grid-connected hybrid energy source," *IEEE*

Journal of Emerging and Selected Topics in Power Electronics, vol. 6, no. 4, pp. 2015–2030, Dec 2018.

[83] M. Pouyan, A. Mousavi, S. Golzari, and A. Hatam, "Improving the performance of q-learning using simultanouse q-values updating," in *2014 International Congress on Technology, Communication and Knowledge (ICTCK)*, Nov 2014, pp. 1–6.

[84] N. C. Tang and G. W. Chang, "A stochastic approach for determining pv hosting capacity of a distribution feeder considering voltage quality constraints," in *2018 18th International Conference on Harmonics and Quality of Power (ICHQP)*, May 2018, pp. 1–5.

[85] S. K. Singh and N. Bansal, "Output impedance as figure of merit to predict transient performance for embedded linear voltage regulators," in *2014 27th International Conference on VLSI Design and 2014 13th International Conference on Embedded Systems*, Jan 2014, pp. 516–521.

[86] R. Tonkoski, D. Turcotte, and T. H. M. EL-Fouly, "Impact of high pv penetration on voltage profiles in residential neighborhoods," *IEEE Transactions on Sustainable Energy*, vol. 3, no. 3, pp. 518–527, July 2012.

[87] K. Coogan, M. J. Reno, S. Grijalva, and R. J. Broderick, "Locational dependence of pv hosting capacity correlated with feeder load," in *2014 IEEE PES T D Conference and Exposition*, April 2014, pp. 1–5.

[88] M. Rylander, J. Smith, and W. Sunderman, "Streamlined method for determining distribution system hosting capacity," *IEEE Transactions on Industry Applications*, vol. 52, no. 1, pp. 105–111, Jan 2016.

[89] M. Rylander, J. Smith, W. Sunderman, D. Smith, and J. Glass, "Application of new method for distribution-wide assessment of distributed energy resources," in *2016 IEEE/PES Transmission and Distribution Conference and Exposition (T D)*, May 2016, pp. 1–5.

[90] M. J. Reno, K. Coogan, S. Grijalva, R. J. Broderick, and J. E. Quiroz, "Pv interconnection risk analysis through distribution system impact signatures and feeder zones," in *2014 IEEE PES General Meeting | Conference Exposition*, July 2014, pp. 1–5.

[91] J. E. Quiroz, M. J. Reno, and R. J. Broderick, "Pv-induced low voltage and mitigation options," in *2015 IEEE 42nd Photovoltaic Specialist Conference (PVSC)*, June 2015, pp. 1–6.

[92] A. Ballanti and L. F. Ochoa, "On the integrated pv hosting capacity of mv and lv distribution networks," in *2015 IEEE PES Innovative Smart Grid Technologies Latin America (ISGT LATAM)*, Oct 2015, pp. 366–370.

[93] Y. Liu, Y. Tai, C. Huang, H. Su, P. Lan, and M. Hsieh, "Assessment of the pv hosting capacity for the medium-voltage 11.4 kv distribution feeder," in *2018 IEEE International Conference on Applied System Invention (ICASI)*, April 2018, pp. 381–384.

[94] Z. Ren, K. Wang, W. Li, L. Jin, and Y. Dai, "Probabilistic power flow analysis of power systems incorporating tidal current generation," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 3, pp. 1195–1203, July 2017.

[95] A. Kazemdehdashti, M. Mohammadi, and A. R. Seifi, "The generalized cross-entropy method in probabilistic optimal power flow," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 5738–5748, Sep. 2018.

[96] J. Li, B. Zhang, and Y. Liu, "Data mining in nonlinear probabilistic load flow based on monte carlo simulation," in *2009 First International Conference on Information Science and Engineering*, Dec 2009, pp. 833–836.

[97] E. Mulenga and M. H. J. Bollen, "Impact of service and feeder cable upgrade on hosting capacity for single phase connected photovoltaics," in *2018 18th International Conference on Harmonics and Quality of Power (ICHQP)*, May 2018, pp. 1–6.

[98] Y. Mi, C. Ma, S. Wang, C. Bi, Y. Zhu, and H. Zhang, "Reactive power control of an isolated wind-diesel hybrid power system based on svc by using sliding mode control," in *2016 35th Chinese Control Conference (CCC)*, July 2016, pp. 9882–9887.

[99] P. D. Achlerkar, S. Kewat, B. K. Panigrahi, and B. Singh, "Distributed control framework for autonomous seamless operation of electronically interfaced distributed generators in ac microgrid," in *2018 8th IEEE India International Conference on Power Electronics (IICPE)*, Dec 2018, pp. 1–6.

[100] M. Al-Saffar and P. Musilek, "Distributed optimal power flow for electric power systems with high penetration of distributed energy resources," in *2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE)*, 2019, pp. 1–5.

[101] D. Shchetinin, T. T. De Rubira, and G. Hug, "On the construction of linear approximations of line flow constraints for ac optimal power flow," *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1182–1192, 2019.

[102] J. Yang, N. Zhang, C. Kang, and Q. Xia, "A state-independent linear power flow model with accurate estimation of voltage magnitude," *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3607–3617, Sep. 2017.

[103] H. Yuan, F. Li, Y. Wei, and J. Zhu, "Novel linearized power flow and linearized opf models for active distribution networks with application in distribution lmp," *IEEE Transactions on Smart Grid*, vol. 9, no. 1, pp. 438–448, Jan 2018.

[104] R. S. Sutton and A. G. Barto, *Reinforcement learning: an introduction*. MITP, 2015. [Online]. Available: https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf

[105] M. Al-Saffar and P. Musilek, "Reinforcement learning-based distributed bess management for mitigating overvoltage issues in systems with high pv penetration," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 2980–2994, 2020.

[106] D. Lee and J. Lee, "Incremental receptive field weighted actor-critic," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 1, pp. 62–71, 2013.

[107] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," pp. 5998–6008, 2017.

[108] S. Yoon, Y. Choi, J. Park, and S. Bahk, "Stackelberg-game-based demand response for at-home electric vehicle charging," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 6, pp. 4172–4184, 2016.

[109] B. Chazelle and J. Matoušek, "On linear-time deterministic algorithms for optimization problems in fixed dimension," *Journal of Algorithms*, vol. 21, no. 3, pp. 579 – 597, 1996. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0196677496900607

[110] B. Venkatesh, S. Chandramohan, N. Kayalvizhi, and R. P. K. Devi, "Optimal reconfiguration of radial distribuion system using artificial intelligence methods," in *2009 IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH)*, Sep. 2009, pp. 660–665.

[111] Wei Ren, R. W. Beard, and E. M. Atkins, "A survey of consensus problems in multi-agent coordination," in *Proceedings of the 2005, American Control Conference, 2005.*, 2005, pp. 1859–1864 vol. 3.