

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

Mohammed Al-Saffar\*, *Member, IEEE*, Petr Musilek\*<sup>†</sup>, *Senior Member, IEEE*

\*Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada

<sup>†</sup>Department of Cybernetics, University of Hradec Kralove, Hradec Kralove, Czech Republic

**Abstract**—High levels of penetration of distributed photovoltaic generators can cause serious overvoltage issues, especially during periods of high power generation and light loads. 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 work develops a cooperative 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 the practical medium size distribution system in Alberta Canada.

**Index Terms**—Cooperative control, state of charge, PV penetration, over-voltage, model predictive control, energy storage, coordination, Monte Carlo tree search, reinforcement learning

## NOMENCLATURE

$a$	RL agent action.
$B_s$	BESS capacity kW.
$C_1$	Grid energy cost \$/kWh.
$C_2$	BESS energy cost \$/kWh.
$C_{purch}$	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.
$J_{Batt}$	BESS time-of-use (TOU) cost objective function \$/kWh.
$J_g$	Grid purchase cost objective function \$/kWh.
$k$	Prediction horizon samples minute.
$N$	Time horizon minute.
$N_{reg}$	The total number of buses in one of the impacted regions.
$P_{Batt}$	BESS power generation kW.
$P_L$	Load power kW

$P_{PV}$	PV power generation kW.
$P_a^{rev}$	Reverse power of the assisting buses kW.
$P_{ch}$	BESS charging power kW.
$P_{dis}$	BESS discharging power kW.
$P_g$	Grid power generation kW.
$P_i^{rev}$	Reverse power of the impacted buses kW.
$Q$	The action-value function .
$r$	Reward value.
$s$	System state.
$SOC$	BESS state of charge.
$T_s$	Sample time minute.
$\hat{u}_b$	State space controller for the impacted and assisting regions kW.
$V(s)$	Value function.
$V^*(s)$	The optimal value function.
$y_{Batt}$	BESS energy cost space system output \$/kWh.
$y_g$	Grid energy cost for the state space system output \$/kWh.
$\alpha_b$	BESS deployment rate of without CSOCC scenario.
$\alpha_i$	BESS deployment rate of the impacted region at the proposed CSOCC scenario.
$\alpha_a$	BESS deployment rate of the participation of assisting region at the proposed CSOCC scenario.
$\eta_{ch}$	Charging efficiency of the BESS.
$\eta_{dis}$	Discharging efficiency of the BESS.
$\pi(s)$	RL agent policy.
$\theta_i$	Reverse power threshold of the impacted buses kW.
$\theta_a$	Reverse power threshold of the assisting buses kW.

## I. INTRODUCTION

**I**NTEGRATION 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 [1], [2]. 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 buses susceptible 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 characteristics and loading conditions [2], [3]. It 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 [1].

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 [4]. 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 [5]. 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 [6], [7].

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 [8], [9]. 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 capacity. To address these issues, it has been proposed to use available storage capacity of neighboring BESS units [10]–[14]. This way, the charge/discharge power in small radial distribution networks can be maximized using common battery scheduling. 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 known 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 the 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 reinforcement learning (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 estimated reward of the new action using a transition probability. The purpose of Monte Carlo tree search (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 article, 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. The proposed algorithm and other compared algorithms are implemented and tested using Python, MATLAB, and CYME software.

This paper 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, RL-MCTS, 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.

## II. 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 [15]. The second, practical case is the distribution feeder of Lloydminster city 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 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 [16] and will be shown in section. V.

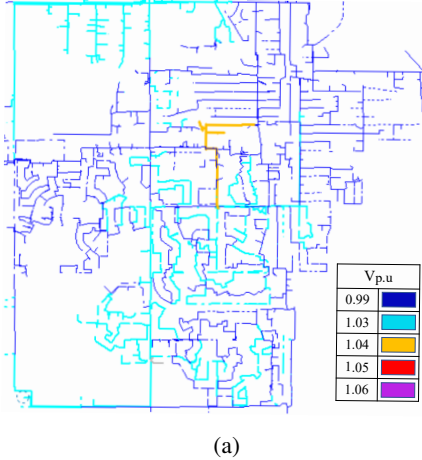


Fig. 1: A circuit diagram of the Lloydminster distribution system showing voltage levels at the full load.

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 [17]. 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) [18]. The impact of PV systems on a distribution circuit can 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 [3].

### III. 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 is either 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 [19], [20]. In centralized control, a central controller aggregates measurements from the entire network and determines the charging/discharging control tasks for each

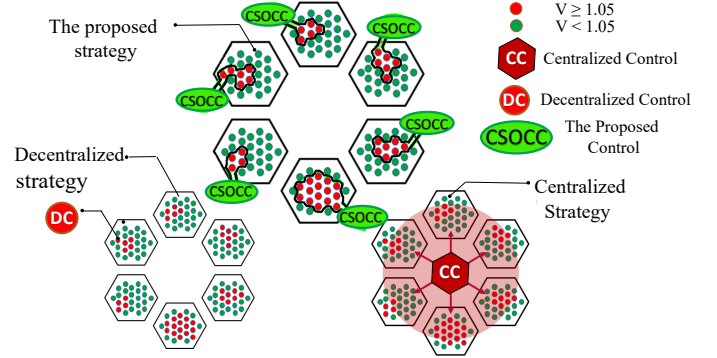


Fig. 2: Conceptual comparison of the proposed distributed control strategy against the conventional centralized and decentralized strategies.

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, BESS units may not respond to the central controller [21], [22], and the entire system may collapse.

The proposed CSOCC strategy forms multiple distributed sections within an entire network [23], [24]. 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 approach, 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 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 [25], [26]. 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

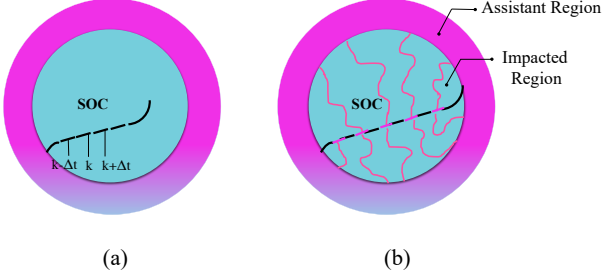


Fig. 3: Principle of the proposed CSOCC approach: BESS in the assisting regions effectively stretch the available storage capacity to avoid its premature saturation or depletion.

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 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 using the centralized controller. The energy transferred from other regions is represented by the wavy lines in Figure 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.

#### A. MCTS-RL

Methods of machine learning and artificial intelligence 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) [27] [28], [29] 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 paper, 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., it learns from the results) to further optimize its own decisions in an unsupervised fashion. This approach uses temporal difference RL methods such as SARSA [30]. 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 [30]. It can be described using the following equation:

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

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

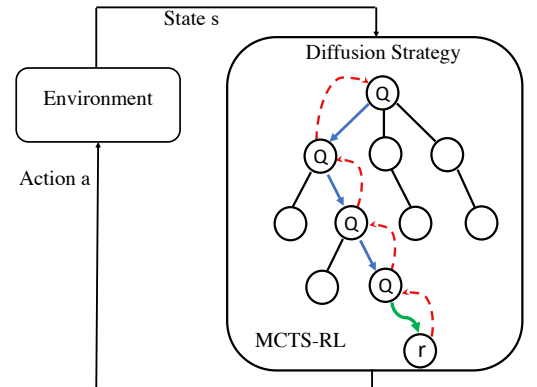


Fig. 4: The proposed algorithm based on MCTS-RL method.

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. The MCTS-RL model receives the state from the environment, transforms it into the transition probability, computes  $Q(s, a)$  values, and selects the 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.

### B. Battery modeling

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

$$SOC(k+1) = SOC(k) + \frac{T_s}{B_s} (\eta_{ch} P_{ch}(k) - \eta_{dis} P_{dis}(k)), \quad (2)$$

where  $SOC$  is the state variable (state of charge),  $\eta_{ch}$ ,  $\eta_{dis}$  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 [31].

The  $SOC$  of the battery bank is subject to the upper and lower boundaries [32]

$$SOC_{min} \leq SOC(k) \leq SOC_{max}. \quad (3)$$

The boundary constraints for the charging and discharging power can be represented as

$$0 \leq P_{ch}(k) \leq P_{max}, \quad (4)$$

$$0 \leq P_{dis}(k) \leq P_{max}, \quad (5)$$

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

$$P_{ch}(k) \cdot P_{dis}(k) = 0. \quad (6)$$

### C. MPC Design

This section develops MPC controller design procedure [33] that implements the proposed method. The dynamic behavior of SOC (2) can be simplified in terms of the state space model

$$SOC(k+1) = SOC(k) + B_b(k) \cdot u_b(k). \quad (7)$$

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

$$\hat{u}_b(k) = [0, (P_{ch,i}(k) + P_{ch,a}(k)), (P_{dis,i}(k) + P_{dis,a}(k))]^T, \quad (8)$$

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

$$SOC(k+1) = SOC(k) + B_b(k) \cdot \hat{u}_b(k). \quad (9)$$

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

$$J_e = SOC(k+1) - (SOC(k) + B_b(k) \cdot \hat{u}_b(k)). \quad (10)$$

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 [34]. The energy balance within the system is maintained through the following equality constraint

$$P_g(k) = P_L(k) - (P_{PV}(k) + P_{Batt}(k)), \quad (11)$$

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  $C_2$  [35], the goal of the MPC assignment can be expressed as minimization of the following two variables [36]

$$y_g(k) = C_1(k) P_g(k), \quad (12)$$

$$y_{Batt}(k) = C_2(k) (-P_{ch}(k) + P_{dis}(k)). \quad (13)$$

The augmented system state can be expressed as

$$x(k+1) = [SOC(k+1), y_g(k), y_{Batt}(k)]^T, \quad (14)$$

and transformed into state-space model [37]

$$x(k+1) = Ax(k) + B \cdot \hat{u}_b(k) \quad (15)$$

to calculate the output

$$y(k) = Cx(k), \quad (16)$$

where the model matrices are

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

$$C = [0_{2 \times 1} \quad I_{2 \times 2}].$$

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

$$J_g(k) = \min \sum_k^{k+N_p} C_1(k) P_g(k), \quad (17)$$

and

$$J_{Batt}(k) = \min \sum_k^{k+N_p} C_2(k) P_{Batt}(k), \quad (18)$$

where the battery TOU cost is

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

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



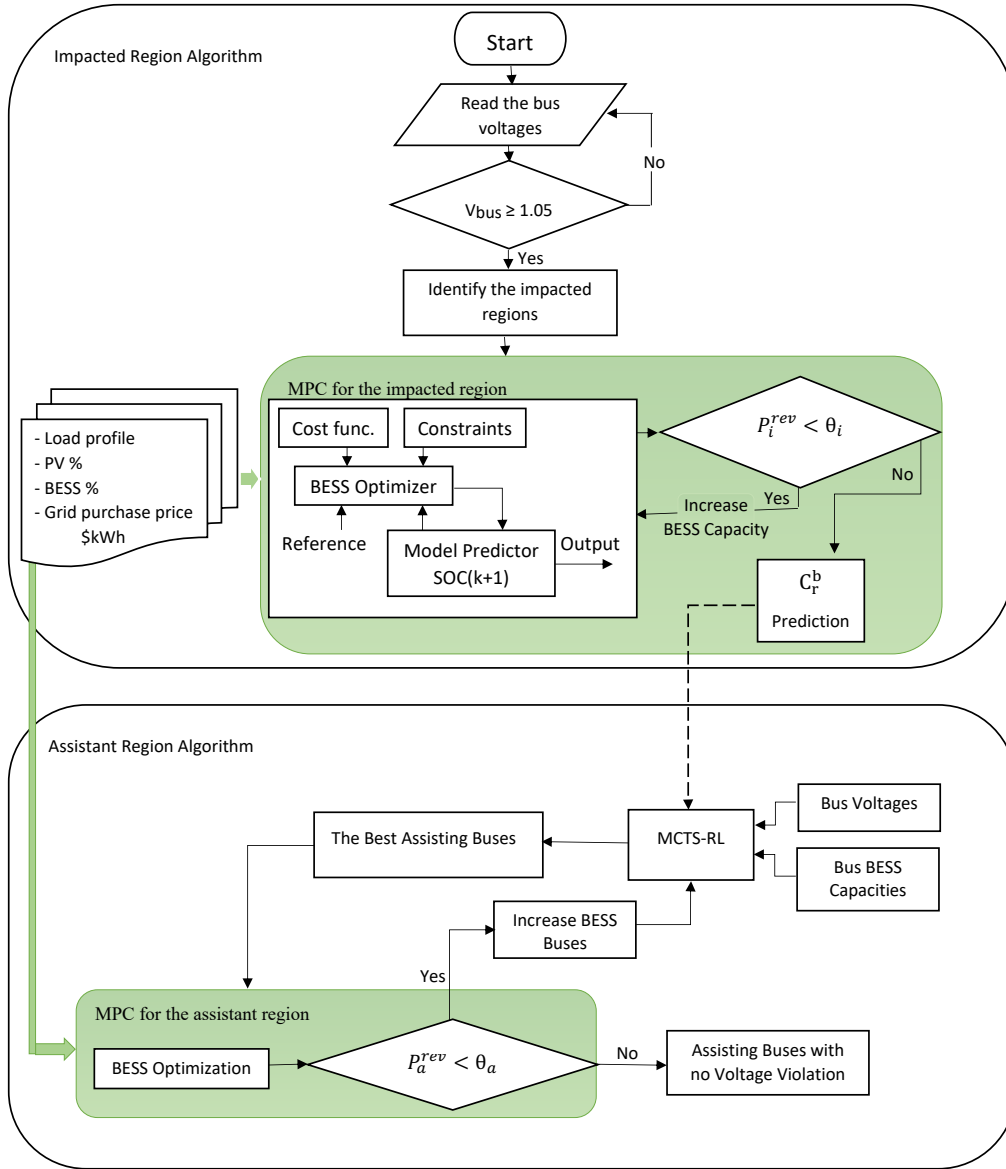


Fig. 5: A flowchart of the proposed algorithm.

The state-value  $V(s)$ , under the optimal policy  $\pi$  for given state  $s$  and the state-action-value  $Q(s, a)$  for taking action  $a$ , is calculated as follows [28]

$$V^\pi(s) = \max_{a \in A} Q^\pi(s, a), \quad (25)$$

where  $\pi$  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  $\alpha_a$

$$\alpha_a = 1/D_a * \left( \max(P_{ch,a,without CSOCC}) - \max(P_{ch,a,with CSOCC}) \right), \quad (26)$$

where  $D_a$  is the full load demand of the assisting bus (subscript  $a$  denotes variables related to the assisting region). The value of  $\alpha_a$  may be reduced if the reverse power of an assisting bus  $P_a^{rev}$  exceeds a predetermined threshold  $\theta_a$ . This threshold is determined in the same manner as  $\theta_i$  of the impacted buses as described above, i.e.,  $\theta_a$  also equals to zero. The resulting shortage of power injection capacity of the assisting buses is compensated by adding more BESS units.

## V. RESULTS AND DISCUSSION

### A. Stochastic framework for PV deployment modelling

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 [17]. 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) [18]. The impact of PV systems

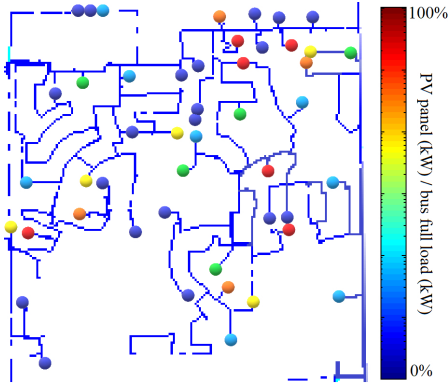


Fig. 6: An illustration of the stochastic PV deployment scenario.

on a distribution circuit can be determined using a stochastic deployment framework implemented by a probabilistic power flow-based Monte Carlo simulations (PPF-MCS) [3], [40]. 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 [3]. The probability distributions of bus voltages are compared against 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 on 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 (e.g., 100) are considered, each with a random distribution of locations and sizes of PV generators at the individual points of supply. An illustration of PV deployment scenario is shown for a small section of the circuit in Figure 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 prescribed reliability indices.
- 4) When PV deployment scenarios modelling 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 [41] 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) [13], [42]. In simulations, the load power factor (PF) of 1.0 is used, that corresponds to the worst-case scenario.

### B. Test Case 1: IEEE 33 Bus System

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 90% PV penetration at 3 critical buses 16, 17, and 18 where voltages exceed 1.045 p.u.

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 decentralized controller using MPC without a central coordinator. To make this test case consistent with the real case (circuit in Lloydminster, Alberta, Canada) described next, the hourly averaged electricity pool prices [43], [44] 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 7.

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

$$\alpha_b = \max(P_{ch,i})/D_i, \quad (27)$$

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 8, have  $\max(P_{ch,i}) = 18$  kW. After dividing this value by the full load  $D_i = 90$  kW (see Figure 7), the value of deployment rate is obtained as  $\alpha_b = 20\%$ . Figure 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).

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



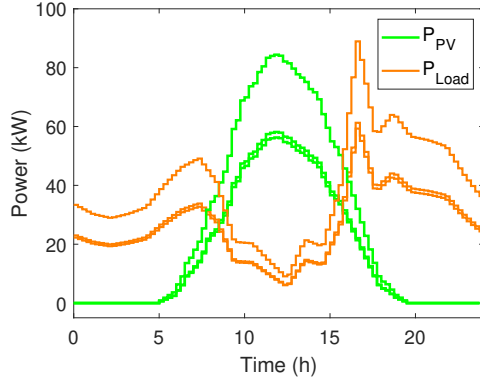


Fig. 7: The load demand and PV power generation of the 3 critical buses 16, 17, and 18.

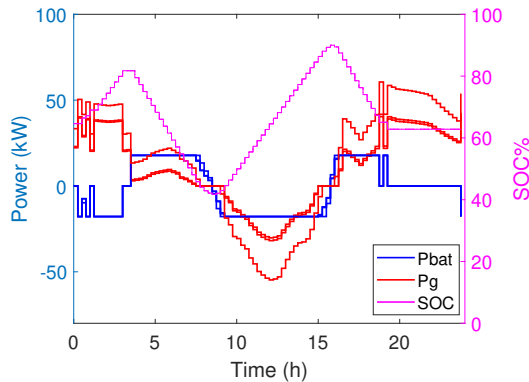


Fig. 8: BESS power controlled by MPC at the critical buses and corresponding grid terminal power flows and SOC plots (without CSOCC).

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 17).

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 represented by so called Q matrix [45] illustrated in Table I. 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 the MCTS terminology). An action is represented by a random selection of one of the children. As shown in Table I, the algorithm starts from a group of

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.

TABLE I: State-action matrix

State	Action	
	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
⋮	⋮	⋮

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 (24). 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 further 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 9. It can be seen that different Q-values are assigned to different buses. Bus 8 has 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. 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  $\alpha_b=20\%$  of the full load using equation (27). With CSOCC, the BESS deployment rate in the impacted region is increased to a new value  $\alpha_i$  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 II shows bus labels (2-33) in the search space. MCTS uses two objective functions: the line

TABLE II: 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	-

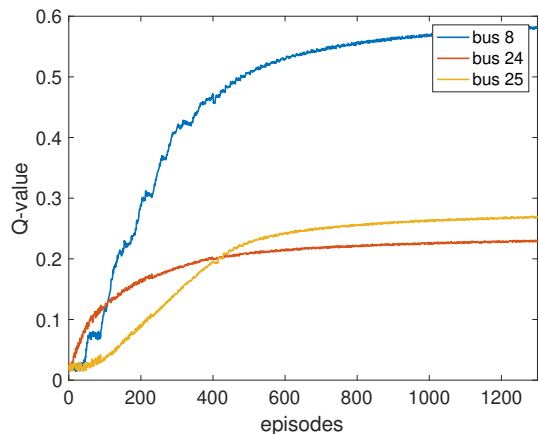


Fig. 9: Plots of Q-values of buses 8, 24, and 25

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 [41] 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.

The required BESS capacities resulting from the MPC optimization process correspond to the energy differences between the two scenarios shown in Figures 8 and 10, over the time period when SOC changes from its minimum to maximum value. These capacities are determined based on the optimal BESS cost (13) while using the MPC controller (8). Deployment rate of the critical buses in the impacted region,  $\alpha_i$ , is calculated through (27) 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 7 and 10, the deployment rate is  $\alpha_i = 72\text{kW}/90\text{kW} = 80\%$ . 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

$$C_r^b \approx \sum_{n_{\text{bus}}=1}^{N_{\text{reg}}} \int_{t_{\text{soc},\min}}^{t_{\text{soc},\max}} D_i(t)(\alpha_i(t) - \alpha_b(t))dt, \quad (28)$$

where  $N_{\text{reg}}$  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 10). The total required BESS capacity is  $C_r^b = 462.8\text{kWh}$ , and the individual required capacities for buses 16, 17, and 18 are 116.2 kWh, 118.4 kWh, and 228.2 kWh, respectively. Figure 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 II column 5) is  $\alpha_a=10\%$  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 the level of participation of BESS installed in the assisting region towards the impacted region).

Figure 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.

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

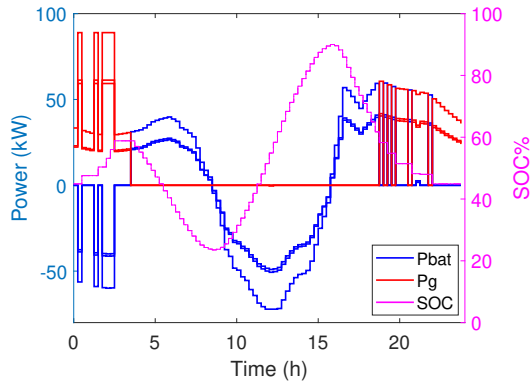


Fig. 10: BESS power controlled by MPC at the critical buses and corresponding grid terminal power flows and SOC plots (with CSOCC).

### C. Test Case 2: Lloydminster Circuit

1) *PV hosting capacity*: Similarly, to previous test case, the load flow analysis of the distribution circuit was simulated by PPF-MCS, this time using CYME power engineering software. The same assumptions about the reactive power and the power factor as the previous case were considered. Certain penetration levels cause the occurrence of impacted regions (regions with marginal, but still acceptable voltage increase) shown in orange in Figure 12. The voltage violations start at 60% PV penetration at 14 buses called critical buses they are located at three regions A, B, and C, shown in red, called the impacted buses, where voltages exceed 1.05 p.u.

2) *Battery implementation without CSOCC*: The load demand and PV power generation profiles of the 14 critical buses are shown in Figure 13, grouped into the same three regions as in Figure 12. The load profiles are obtained as averages of the residential and industrial load curves from the circuit data. In this case study, only one region is sufficient to explain the simulation. For demonstration, area C has been selected as in this area the assisting buses fittingly surround the impacted buses where they will be needed in the next scenario (with CSOCC). The worst bus in this area is bus 4 that causes the highest voltage violation. Based on the worst bus, the MPC optimization results, shown in Figure 14, have  $\max(P_{ch,i}) = 5.34$  kW. and  $D_i = 62$  kW (see Figure 13); the value of deployment rate is  $\alpha_i = 9\%$ .

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 17).

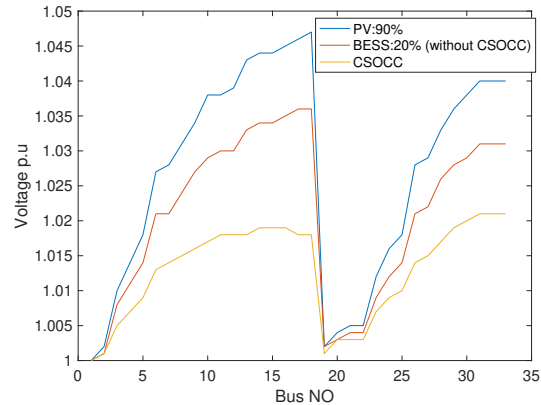


Fig. 11: 33-bus system voltages in three cases: i) 90% PV penetration with no storage; ii) 90% PV penetration with 20% BESS penetration but without CSOCC; and iii) 90% PV penetration with 20% BESS penetration and with CSOCC.

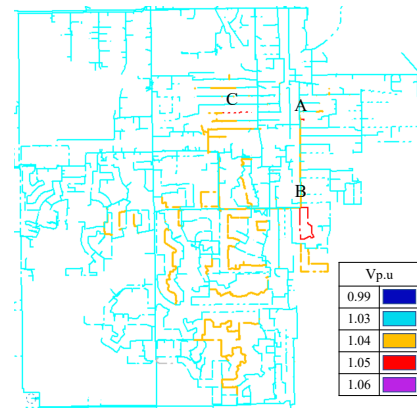


Fig. 12: The impacted regions (orange) and critical buses (red); the impacted regions contain the following buses: a) bus 1 at (A), b) buses 2-8 at (B), c) buses 9-14 at (C).

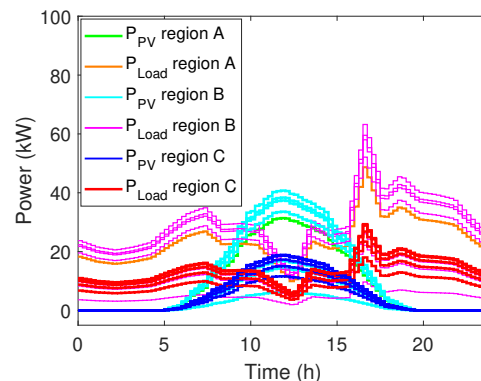


Fig. 13: The load demand and PV power generation of the 14 critical buses in the three regions A, B, and C.

3) *Battery implementation with CSOCC*: As mentioned in the previous scenario, area C is considered in this test case. Figure 15d shows the buses in the search space that are selected as the best candidates by following the two objective

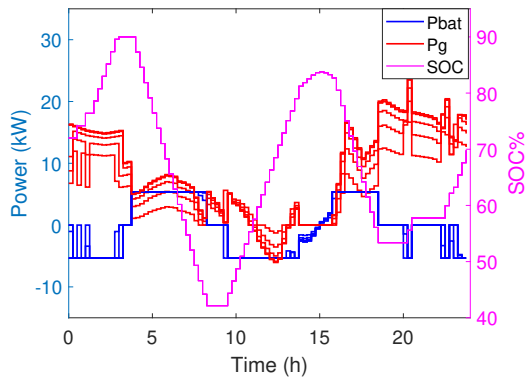


Fig. 14: Terminal power flows at the critical buses and corresponding BESS flows (without CSOCC).

functions of line losses and the BESS capacities. Their values are shown through Figures 15a and b respectively. The bus voltages that are checked against the voltage constraint are also shown in Figure 15c. The reverse powers if they occur in the assisting buses are subject to  $\theta_a$ . Similarly to the previous 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 using 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 14 and 16, over the time period when SOC transitions from the maximum to minimum value.

In this test case, the deployment rate for the impacted buses is  $\alpha_i = 11.5\text{kW}/62\text{kW} = 18\%$  (obtained using information from Figures 13 and 16, as in the previous case).

$C_r^b$  values for impacted regions A, B, and C are determined using equation (28) as 42.2 kWh, 197.75 kWh, and 124.3 kWh, respectively. 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 penetration level requirement can be considered an additional advantage of a robust power management provided by 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 17. The semitransparent blue plane represents the voltage limit of 1.05 p.u.

Finally, Table III shows the system costs calculated for region C using the following three scenarios: conventional method (without MPC), with MPC but without CSOCC and with MPC and CSOCC. The conventional method is explained in [10], [12], and [13], where the BESS charges at the time period of high PV production and 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

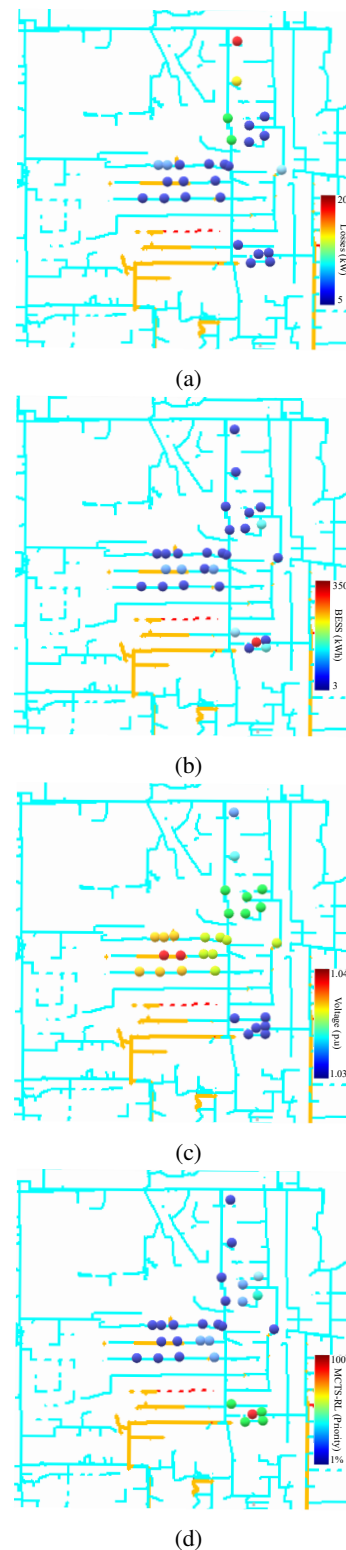


Fig. 15: States at the assisting regions: a) line losses, b) bus BESS capacities, c) bus voltages, d) the optimized assisting buses obtained using MCTS-RL d) the final results of the assisting buses corresponding to the critical region required energy

row for each scenario. It is obvious that the total costs in the two scenarios without and with CSOCC are significantly

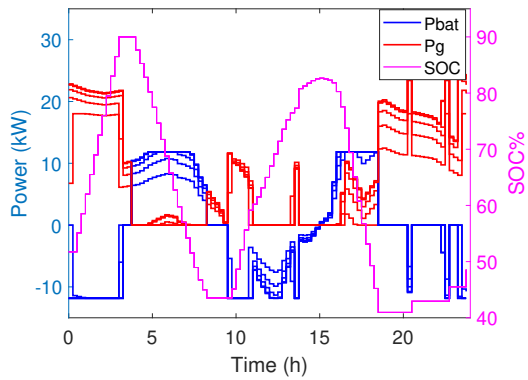


Fig. 16: Terminal power flows at the critical buses and corresponding BESS flows (with CSOCC) .

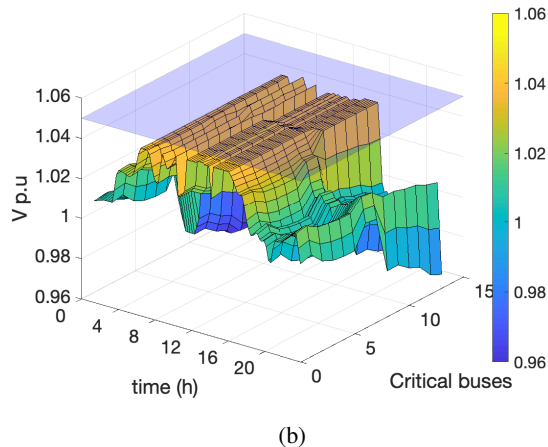
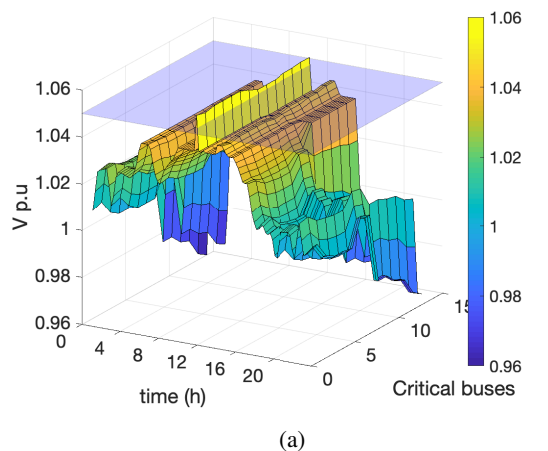


Fig. 17: Voltages levels at the critical buses during a 24 hour operation without CSOCC (a), and with CSOCC (b).

reduced in comparison to the conventional scenario. The use of CSOCC reduces the impacted region costs even further, as also shown in plots a) and b) of Figure 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

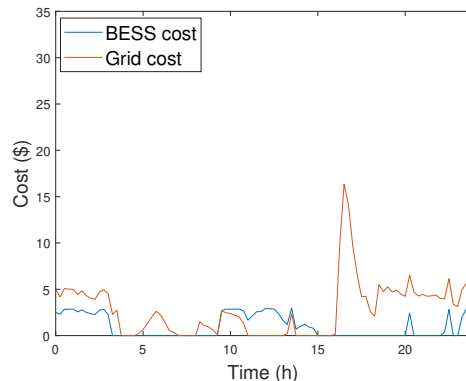
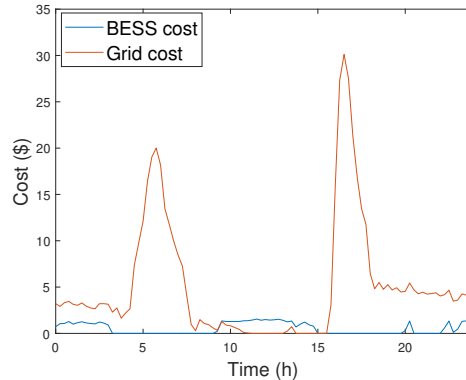


Fig. 18: The critical buses costs: a) without CSOCC, b) with CSOCC.

small because this algorithm only coordinates the BESS units in the two regions without adding new energy to the system.

TABLE III: System scenario 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	

There are other published works that also use coordination strategies to share energy among neighbouring BESS units [10]–[14]. However, they do not consider the stochasticity of DER. Therefore, to facilitate coordination, these strategies 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 a 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 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 rather than that examining the entire search space.

## VI. CONCLUSION

This paper introduces a novel method for distributed BESS control to mitigate voltage rise in power distribution networks caused by high penetration of residential PV generators. The proposed method, dubbed coordinated state of charge control (CSOCC), combines MCTS-RL and co-operative 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 be extended to incorporate other practical factors such as thermal constraints or ability to handle unbalanced three-phase systems.

## ACKNOWLEDGMENT

The research described in this paper has been supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) grant RGPIN-2017-05866, by the Canada First Research Excellence Fund (CFREF) under the Future Energy Systems research initiative at the University of Alberta, and by the Faculty of Science, University of Hradec Kralove. The authors would like to thank Dr. Shahed Mortazavian for her helpful comments on the initial draft of this paper. We would also like to acknowledge the support provided by ATCO Electric during this research, namely Dr. Alexandre Nassif, Tom Greenwood-Madsen, and Matt Sveinbjornson.

## REFERENCES

- [1] 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, DOI 10.1109/TSTE.2012.2191425, no. 3, pp. 518–527, Jul. 2012.
- [2] 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*, DOI 10.1109/TDC.2014.6863515, pp. 1–5, Apr. 2014.
- [3] A. Dubey, S. Santoso, and A. Maitra, "Understanding photovoltaic hosting capacity of distribution circuits," in *2015 IEEE Power Energy Society General Meeting*, DOI 10.1109/PESGM.2015.7286510, pp. 1–5, Jul. 2015.
- [4] 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*, DOI 10.1109/ISGTEurope.2013.6695423, pp. 1–5, Oct. 2013.
- [5] 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*, DOI 10.1109/PESGM.2015.7286523, pp. 1–5, Jul. 2015.
- [6] 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, DOI 10.1109/TII.2016.2619065, no. 1, pp. 270–277, Feb. 2017.
- [7] 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*, DOI 10.1109/PESGM.2013.6672092, pp. 1–1, Jul. 2013.
- [8] 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)*, DOI 10.1109/EEM.2016.7521338, pp. 1–5, Jun. 2016.
- [9] S. Hashemi and J. Ostergaard, "Efficient control of energy storage for increasing the pv hosting capacity of lv grids," in *2017 IEEE Manchester PowerTech*, DOI 10.1109/PTC.2017.7981148, pp. 1–1, Jun. 2017.
- [10] 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, DOI 10.1109/TII.2014.2299336, no. 2, pp. 967–977, May. 2014.
- [11] 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, DOI 10.1109/TSG.2015.2452579, no. 6, pp. 2906–2918, Nov. 2015.

- [12] 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, DOI 10.1109/TPWRD.2015.2462723, no. 3, pp. 1132–1141, Jun. 2016.
- [13] 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, DOI 10.1109/TSG.2016.2636217, no. 99, pp. 1–1, 2016.
- [14] 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, DOI 10.1109/TSG.2017.2754943, no. 1, pp. 919–927, Jan. 2019.
- [15] "Distribution Test 33-Bus Feeder," [http://www.ece.ubc.ca/~hameda/download\\_files/](http://www.ece.ubc.ca/~hameda/download_files/).
- [16] C. I. T. Inc, "Distribution System Analysis," [www.cyme.com/software/cymdist/](http://www.cyme.com/software/cymdist/), 2017.
- [17] 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, DOI 10.1109/JPHOTOV.2015.2481605, no. 1, pp. 295–300, Jan. 2016.
- [18] 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, DOI 10.1049/iet-gtd.2018.5692, no. 20, pp. 4462–4471, 2018.
- [19] K. Worthmann, C. M. Kelleit, 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, DOI 10.1109/TSG.2015.2392081, no. 4, pp. 1914–1923, Jul. 2015.
- [20] 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, DOI 10.1109/TSTE.2017.2688391, no. 4, pp. 1443–1452, Oct. 2017.
- [21] 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)*, DOI 10.1109/EEEIC.2016.7555488, pp. 1–6, Jun. 2016.
- [22] 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)*, DOI 10.1109/ECCE.2014.6953422, pp. 409–413, Sep. 2014.
- [23] 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, DOI 10.1109/TSG.2013.2292530, no. 2, pp. 1018–1032, Mar. 2014.
- [24] 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, DOI 10.1109/TEC.2014.2357689, no. 4, pp. 988–996, Dec. 2014.
- [25] T. Morstyn, M. Momayyezani, 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, DOI 10.1109/TPEL.2015.2513777, no. 11, pp. 7986–7995, Nov. 2016.
- [26] 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, DOI 10.1109/TSG.2014.2354033, no. 1, pp. 238–248, Jan. 2015.
- [27] 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, DOI 10.1109/TCI-AIG.2012.2186810, no. 1, pp. 1–43, Mar. 2012.
- [28] K. Arulkumar, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep reinforcement learning: A brief survey," *IEEE Signal Processing Magazine*, vol. 34, DOI 10.1109/MSP.2017.2743240, no. 6, pp. 26–38, Nov. 2017.
- [29] A. R. Sharma and P. Kaushik, "Literature survey of statistical, deep and reinforcement learning in natural language processing," in *2017 International Conference on Computing, Communication and Automation (ICCCA)*, DOI 10.1109/ICCCA.2017.8229841, pp. 350–354, May. 2017.
- [30] R. S. Sutton and A. G. Barto, *Temporal-Difference Learning*. MITP, 1998. [Online]. Available: <https://ieeexplore-ieee.org/login.ezproxy.library.ualberta.ca/xpl/articleDetails.jsp?arnumber=6282965>
- [31] 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, DOI 10.1109/TPWRS.2017.2696571, no. 1, pp. 803–816, Jan. 2018.
- [32] I. Lampropoulos, P. Garoufalos, 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, DOI 10.1049/iet-gtd.2014.0908, no. 15, pp. 2319–2327, 2015.
- [33] 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, DOI 10.1109/TSG.2016.2547342, no. 2, pp. 998–1005, Mar. 2017.
- [34] 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)*, DOI 10.1109/TDC-LA.2016.7805637, pp. 1–6, Sep. 2016.
- [35] 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)*, DOI 10.1109/POWERCON.2016.7754011, pp. 1–6, Sep. 2016.
- [36] 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, DOI 10.1109/TCST.2014.2361800, no. 3, pp. 1229–1236, May. 2015.
- [37] O. Gomofov, 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, DOI 10.1109/TVT.2016.2638912, no. 7, pp. 5520–5530, Jul. 2017.
- [38] 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, DOI 10.1109/TSG.2017.2753802, no. 1, pp. 841–851, Jan. 2019.
- [39] S. Jothibasu and S. Santoso, "Sensitivity analysis of photovoltaic hosting capacity of distribution circuits," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, DOI 10.1109/PESGM.2016.7741861, pp. 1–5, Jul. 2016.
- [40] 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*, DOI 10.1109/ICISE.2009.448, pp. 833–836, Dec. 2009.
- [41] 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, DOI 10.1109/TPWRS.2017.2735898, no. 1, pp. 1148–1150, Jan. 2018.
- [42] 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*, DOI 10.1109/ICPE.2011.5944614, pp. 585–592, May. 2011.
- [43] "Alberta Electric System Operator," <http://ets.aeso.ca/>, 2018.
- [44] 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, DOI 10.1109/JESTPE.2018.2828803, no. 4, pp. 2015–2030, Dec. 2018.
- [45] M. Pouyan, A. Mousavi, S. Golzari, and A. Hatam, "Improving the performance of q-learning using simultaneous q-values updating," in *2014 International Congress on Technology, Communication and Knowledge (ICTCK)*, DOI 10.1109/ICTCK.2014.7033528, pp. 1–6, Nov. 2014.



**Mohammed Al-Saffar** received the B.Sc. degree from the University of Technology, Baghdad, Iraq, and the M.Sc. degree from the University of Technology, Malaysia, both in electrical engineering, in 2004 and 2014, respectively. He is currently working toward the Ph.D. degree in electrical and computer engineering at the University of Alberta, Edmonton, AB,

Canada. His research is concerned with the optimization and management of the Distributed Energy Resources (DER) in power distribution systems. He develops distributed architectures that can interact with each other using intelligent techniques.



**Dr. Musilek** (M'99, SM'07) received his Ing degree (Electrical Engineering, with great distinction) in 1991, and Ph.D. (Cybernetics) in 1995 from the Military Academy in Brno, Czech Republic. In 1995, he was appointed Head of the Computer Applications Group at the Institute of Informatics, Military Medical Academy in Hradec Kralove, Czech Republic. Between 1997 and 1999, he was

a NATO Science Fellow at the Intelligent Systems Research Lab, University of Saskatchewan, Canada. In 1999 he joined the Department of Electrical and Computer Engineering at the University of Alberta, Canada, where he is currently a Full Professor. He was the Director of Computer Engineering Program (2016-2017), and currently serves as the Associate Chair (Research and Planning) of the ECE Department. His research interests include artificial intelligence and energy systems. He developed a number of innovative solutions in the areas of renewable energy systems, smart grids, wireless sensor networks, and environmental monitoring and modelling.