

Article

Optimal Voltage Recovery Learning Control for Microgrids with N-Distributed Generations via Hybrid Iteration Algorithm

Lüeshi Li ^{1,2}, Ruizhuo Song ^{3,*} and Shuqing Dong ⁴

¹ School of Electrical Engineering, Northeast Electric Power University, Jilin 132012, China; lueshi.li.2022@uni.strath.ac.uk

² School of Engineering, University of Strathclyde, Glasgow G1 1XJ, Scotland, UK

³ School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China

⁴ North China Electric Power Experimental Research Institute, China Datang Group Science and Technology Research Institute Co., Ltd., Beijing 100097, China; dongshuqing@cdt-kxjs.com

* Correspondence: ruizhuosong@ustb.edu.cn

Abstract: Considering that the nonlinearity and uncertainty of the microgrid model complicate the derivation and design of the optimal controller, an adaptive dynamic programming (ADP) algorithm is designed to solve the model-free non-zero-sum game. By combining the advantages of policy iteration and value iteration, an optimal learning control scheme based on hybrid iteration is constructed to provide stringent real power sharing for the nonlinear and coupled microgrid systems with N-distributed generations. First, using non-zero-sum differential game strategy, a novel distributed secondary voltage recovery consensus optimal control protocol is built using a hybrid iteration method to realize the voltage recovery of microgrids. Then, the data of the system state and input are gathered along a dynamic system trajectory and a data-driven optimal controller learns the game solution without microgrid physics information, enhancing convenience and efficiency in practical applications. Furthermore, the convergence analysis is given in detail, and it is proved that the control protocol can converge to the optimal solution so as to ensure the stability of the voltage recovery of the microgrid system. Convergence analysis proves the convergence of the the protocol to the optimal solution, ensuring voltage recovery stability. Simulation results validate the feasibility and effectiveness of the proposed scheme.

Keywords: adaptive dynamic programming; nonzero-sum games; microgrid; hybrid iteration



Citation: Li, L.; Song, R.; Dong, S. Optimal Voltage Recovery Learning Control for Microgrids with N-Distributed Generations via Hybrid Iteration Algorithm. *Electronics* **2024**, *13*, 2093. <https://doi.org/10.3390/electronics13112093>

Academic Editor: Maciej Ławryńczuk

Received: 4 May 2024

Revised: 21 May 2024

Accepted: 23 May 2024

Published: 28 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years the ecological and environmental foundation of the world has been deteriorating and the demand for energy has become increasingly urgent. The utilization of renewable clean energy and the improvement of energy efficiency have become crucial for environmental protection, yet the structure and operational mode of traditional power distribution systems have proven inadequate in meeting the demands of large-scale renewable power integration [1–3]. In order to meet this challenge, it is imperative to transform and upgrade the traditional power distribution system. Furthermore, the advancement of smart technology has led to the gradual personalization of energy demand and electricity consumption patterns, accelerating the transition of the power system towards a smart grid. This transition has rendered the traditional large grid increasingly incapable of satisfying diversified power supply needs [4–6]. To address these challenges, the microgrid has emerged as a viable solution, playing a pivotal role in maintaining the smooth operation of power grids. A smart grid, as a new power system model, aims to achieve energy conversion and efficient use through advanced communication technology and control means. The goal is to improve the safety, reliability, economy and efficiency of the operation of the power system to meet the growing demand for energy while reducing the environmental

impact of energy consumption. As an integral component of the smart grid, the microgrid represents a small-scale power system that effectively integrates distributed generators [7,8]. Traditional microgrid voltage restoration methods often rely on fixed control strategies and parameter settings, which is difficult to adapt to the changes of different microgrid structures and operating conditions. These methods often show poor adaptability and robustness in the face of complex and changeable power system environments. Therefore, the development of an appropriate frequency and voltage control strategy is crucial for ensuring the stable operation of the microgrid. Research on this topic holds considerable practical significance. This paper aims to address these challenges by exploring innovative control strategies tailored for the unique characteristics of the microgrid. By doing so, the advancement of smart grid technology and enhancements for the overall performance and reliability of power systems worldwide are given in this paper.

Over the past few years, numerous scholars have begun to explore innovative control strategies for microgrids, aiming to achieve stable operation and efficient utilization of microgrids by precisely controlling energy conversion, charge and discharge of energy storage devices and load management [9–11]. Shuai et al. conducted a comprehensive review of research on microgrid stability [12]. Specifically, ref. [13] highlighted the increasing replacement of distributed generators with diesel generators in standalone microgrids. To address this, a decoupled controller for frequency and voltage was introduced, enabling the maintenance of grid frequency and voltage magnitude stability. The high-order full drive model of a distributed generator microgrid is explored by using the large signal model of a microgrid, and the complex relationship and interaction between voltage and current are effectively captured in [14]. In Ref. [15] a moving target defense mechanism was proposed, which can effectively limit network attacks on smart microgrid systems. Furthermore, ref. [16] employed a robust passive control strategy to stabilize and regulate the DC bus voltage. These results represent significant contributions to the field, providing valuable insights into voltage/frequency recovery techniques for generation systems. However, despite these advancements, there remains a need for further research and development in this area, particularly in addressing the unique challenges posed by microgrids and their integration into smart grid systems. This paper wants to contribute to this ongoing effort by exploring innovative control strategies tailored to the specific requirements of microgrids, with a focus on enhancing their stability and performance.

The implementation of a nonlinear disturbance observer is proposed to enhance the stability of the system and mitigate steady-state errors resulting from CPL variations. In Ref. [17], mixed iterative adaptive dynamic programming is explored, providing a solution to the optimal control problem associated with battery energy management in smart residential microgrid systems. This method comprises two iterative processes: P-iteration and V-iteration, both grounded in reinforcement learning techniques. The P-iteration process follows a policy iteration approach, with iterative updates to the value function based on the sequence of iterative control laws. Conversely, the V-iteration process adheres to value iteration, enabling the derivation of iterative control laws in each iterative cycle. It is noteworthy that traditional centralized control strategies rely heavily on extensive communication systems to gather and process substantial amounts of information from distributed generators. These strategies primarily concentrate on data generation [18], energy management [19], state estimation [20], optimal scheduling [21] and reliability evaluation [22]. The nonlinear disturbance observer offers an innovative approach to enhancing system stability and eliminating steady-state errors, presenting a promising alternative to traditional centralized control strategies. This advancement holds significant potential for improving the performance and reliability of smart residential microgrid systems.

Inspired by the distributed and consistent control methods for nonlinear multi-agent systems, ref. [23] introduced the concept of distributed cooperative control and employed a hybrid iteration technique to design a nonlinear model controller for microgrids. This approach effectively bridges the performance gap between the well-known policy iteration

and value iteration methods. In policy iteration, derived from the Newton–Raphson method, convergence to the optimal value occurs at a rate of quadratic convergence [24]. Conversely, value iteration differs as it does not require an initial admissible control policy, as is the case in policy iteration [25]. However, hybrid iteration combines the strengths of both methods, eliminating the need for the initial stability control strategy required in policy iteration while achieving a faster convergence rate than traditional value iteration. In recent years, the application of hybrid iterative technology in modern microgrids has gradually highlighted its importance and potential. In Ref. [26], an optimal output regulation method for islanded modern microgrids was presented. This method demonstrates that the hybrid iterative approach significantly reduces the convergence time of the deployed CPU, minimizes the number of learning iteration cycles and eliminates the need for an initial stability control strategy. However, although the application of hybrid iteration in modern microgrids has made remarkable progress, there remains a lack of research on the acquisition and analysis of optimal feedback control protocols for nonlinear N-distributed generations in microgrids. This research gap not only limits the further application of hybrid iteration in the field of microgrids, but also affects the further improvement of the overall performance of microgrids. By analyzing the nonlinear N-distributed generation in microgrids and combining it with the characteristics of hybrid iteration, this study introduces a novel feedback control protocol, which can not only effectively deal with the nonlinear problems, but also realizes the precise adjustment of the output voltage of a microgrid, thereby contributing to the enhancement of operational efficiency and stability.

With the emergence of novel optimization algorithms, the intelligent control optimization of microgrid systems has seen significant progress. The inherent randomness, dynamics and nonlinearity of microgrid operations pose challenges to traditional model-driven control strategies, limiting their accuracy and efficiency. Adaptive dynamic programming (ADP), an intelligent control approach, exhibits strong self-learning and adaptive capabilities, independent of prior system knowledge or models. ADP leverages iteration to derive control strategies, encompassing policy iteration and value iteration. Utilizing neural network approximations, ADP employs critic and action networks to approximate optimal performance indices and control strategies, respectively. Through continuous information feedback and transmission, the critic network evaluates and updates strategies, enabling the system to converge to an optimal control strategy and its corresponding value function [27]. Game theory, analyzing behavior prediction among competitive individuals and optimizing strategies for multi-agent decision-making, provides a theoretical foundation for the microgrid system's behavior trajectory. Key elements in game theory include players, strategies and cost functions. Ref. [28] transformed the microgrid system with external interference into a zero-sum differential game model to design a robust control scheme. Ref. [29] simplified the distributed secondary voltage recovery consistency control to a general distributed zero-sum differential game and then solved it. Players aim to minimize their cost functions through strategic choices, aligning with the objectives of this study.

This study aims to explore the voltage recovery control problem of microgrid systems with N-distributed generations and proposes an innovative method for this purpose. The main contributions are as follows: (1) voltage recovery control in microgrid systems involves the interaction between multiple distributed power sources and loads, so this paper transforms it into a non-zero-sum game problem to effectively deal with the cooperative relationship between multiple participants. (2) An ADP method is adopted to obtain the optimal control scheme by iteratively solving the Hamilton–Jacobi (HJ) equation. The proposed hybrid iteration combines the advantages of value iteration and policy iteration and adaptively adjusts the step size to ensure the convergence and stability of the algorithm. A strict convergence proof is given to prove the effectiveness of the proposed method. (3) In practical applications, the mathematical models of microgrid systems are often uncertain and complex, making it difficult to model accurately. In order to meet this challenge, the

algorithm proposed in this paper does not rely on mathematical models and gradually approaches the optimal solution through iteration and learning.

This paper addresses the optimal control problem for nonlinear N-distributed generations, introducing a novel hybrid iteration technique to derive an optimal policy without relying on the system model. Initially, the optimal control strategy and cost function value are defined, providing a foundation for subsequent analysis. Subsequently, the effectiveness of the proposed hybrid iteration algorithm is demonstrated in learning optimal controller laws for N-distributed generations. Notably, this strategy does not require a system model, relying solely on online state and input data. This approach is particularly useful for nonlinear microgrid systems, where establishing a precise dynamics model can be challenging. By leveraging the loci of these systems, our method offers a practical solution for real-time optimal control.

The remainder of this paper is structured into six sections. Section 2 introduces the optimal control problem statement and formulation for microgrids with N-distributed generations. Section 3 presents both model-based and data-driven hybrid iteration methods. Section 4 provides a rigorous convergence proof for the proposed methods. Section 5 presents simulation results for a microgrid system with two distributed generations, along with a technical discussion. Finally, Section 6 concludes the paper, summarizing the key findings and contributions.

2. Problem Formulation

The N-generation distributed microgrid system refers to a multi-level complex power grid system formed by multiple microgrids. Each generation of microgrids has a certain degree of autonomy and independence. Consider the microgrid system with N-distributed generations, let the state space variables of i th generation be $\dot{s}_i = \begin{bmatrix} \dot{s}_{i,1} \\ \dot{s}_{i,2} \end{bmatrix} \in \mathbb{R}^{n_1+n_2}$, which contains the key information of the generation of the whole microgrid. $s_{i,1} \in \mathbb{R}^{n_1}$ shows the magnitude of $v_{o,i}$ and represents i th generation output voltage value and $s_{i,2} \in \mathbb{R}^{n_2}$ is the derivative of the magnitude of $v_{o,i}$.

To capture the continuous voltage changes and understand the dynamic behavior of the microgrids in the voltage recovery process more deeply, the state-space model pertaining to the voltage recovery layer of microgrids is formulated in the following differential form:

$$\dot{s}_i = \begin{bmatrix} g_{i,1}(s_i)s_{i,2} \\ f_{i,2}(s_i) \end{bmatrix} + \begin{bmatrix} 0 \\ g_{i,2}(s_i) \end{bmatrix} w_i \quad (1)$$

where $g_{i,1} \in \mathbb{R}^{n_1 \times n_1}$ is the nonlinear smooth function, $w_i \in \mathbb{R}^q$ is the control input and $f_{i,2} \in \mathbb{R}^{n_2}$ and $g_{i,2} \in \mathbb{R}^{n_2 \times q}$ are nonlinear uncertain smooth functions. For ease of analysis, let $f_i(s_i) = \begin{bmatrix} g_{i,1}(s_i)s_{i,2} \\ f_{i,2}(s_i) \end{bmatrix}$, $g_i(s_i) = \begin{bmatrix} 0 \\ g_{i,2}(s_i) \end{bmatrix}$.

For each generation, the utility function is designed as

$$r_i(s_i, w_i, w_{-i}) = Q_i(s_i) + w_i^T R_i w_i \quad (2)$$

where $w_{-i} = \{w_1, \dots, w_{i-1}, w_{i+1}, \dots, w_N\}$ denotes the control input set which does not contain w_i . $Q_i(s_i) = s_i^T Q_i s_i$ is positive definite, in which $Q_i > 0$. The weight association symmetric matrix $R_i > 0$.

Then, the cost function of the i th generation is

$$J_i(s_i, w_i, w_{-i}) = \int_0^\infty r_i(s_i, w_i, w_{-i}) dt \quad (3)$$

The optimal control problem addressed in this paper aims to determine the control strategy that minimizes the performance index value for the microgrid system. Based on (3), the optimal cost function P^* is presented by

$$P_i^*(s_i) = \min_{w_i} J_i(s_i, w_i, w_{-i}). \tag{4}$$

The optimization problem of the microgrid system, which is difficult to obtain the model for, is transformed into the solution of a non-zero-sum game problem in this paper. Considering the interaction of various components in the system, the stable operation of the system is achieved by optimizing the overall performance. The ADP algorithm is used to adapt to the uncertainties in the system, adjust the control strategy flexibly and improve the operating efficiency and reliability of the system.

Based on the stationary conditions and utilizing the optimal cost function (4), the optimal feedback control protocol can be formulated as follows [24]

$$w_i^* = -\frac{1}{2} R_i^T g_i^T(s_i) \nabla P_i^* \tag{5}$$

The aforementioned control law aims to achieve the Nash equilibrium solution of the game model, targeting the equilibrium of interests among multiple entities within the microgrid system. This is not only an important guarantee for system stability and reliability, but also the key to the efficient and sustainable operation of microgrid systems. Additionally, it is evident from the expression that the control law output of each player in the microgrid system is unaffected by other players, which aligns with the concept of a non-zero-sum game. Each player attains concordant interests through mutual promotion and coordination, thereby minimizing the designed value function.

In order to solve (4) and (5), the Hamiltonian function is defined as

$$\mathfrak{H}(s_i, \nabla P_i, w_i) = \nabla P_i^T (f_i + g_i w_i) + Q(s_i) + w_i^T R_{i,i} w_i \tag{6}$$

According to optimal control theory, it has

$$\begin{cases} \inf_{w_i} \mathfrak{H}(s_i, \nabla P_i, w_i) = 0 \\ P_i(0) = 0 \end{cases} \tag{7}$$

From (4) and (5), the coupled equation of HJ can be expressed by

$$\begin{aligned} 0 = & \frac{1}{4} \sum_{i=1}^N \nabla P_i^{*T} g_{i,2}(s_i) R_i^{-1} R_{i,i} R_i^{-1} g_{i,2}^T \nabla P_i^* + \nabla P_i^{*T} f_{i,2}(s_i) \\ & + Q_i(s_i) - \frac{1}{2} \nabla P_i^{*T} \sum_{i=1}^N g_{i,2}(s_i) R_i^{-1} g_{i,2}^T \nabla P_i^* \end{aligned} \tag{8}$$

with the initial condition $P_i^*(0) = 0$.

In the next section, the aforementioned coupled HJ equation is approximated using the ADP algorithm in a hybrid iterative manner. According to the above preparation, it becomes evident that each generation of the coupled HJ equation, as presented in (8), has a connection with the other control laws. Consequently, the multiple coupled generations are not entirely autonomous. Therefore, this paper aims to establish an adaptive and distributed consistent voltage recovery control strategy for the N-distributed game system (1). To tackle this problem, there have been studies on designing near-optimal control methods, including policy iteration and value iteration, which are all online learning techniques.

The implementation of policy iteration requires certain initial conditions. Specifically, the system must commence iteration with an admissible control strategy. Subsequently, the iteration performance index is approximated, the control law of the current iteration step is evaluated and then the strategy is updated. These evaluation and update steps are repeated until a termination condition, similar to the inequality in (12), is satisfied.

Obviously, the advantage of this iterative method lies in its high computational efficiency, as policy iteration based on an initial control law can enable the system to quickly obtain the optimal control law. In contrast to policy iteration, the initial value of the iteration value function in value iteration can be arbitrary. In order to overcome these limitations, a hybrid iterative algorithm is used to solve the optimal control strategy. This algorithm combines the advantages of policy iteration and value iteration and not only does it not require the initially allowed control strategy, but it also has a high convergence speed.

3. Hybrid Iteration for the Microgrid Systems

In this part, by applying the hybrid iteration strategy and the reinforcement learning method, a recovery control protocol for adaptive distributed voltage is constructed. First, a hybrid iteration approach based on the system model is proposed and a convergence analysis is provided. Then, the implementation process of the hybrid iteration approach based on data is presented, along with proof of convergence.

The hybrid iteration is based on the execution of policy iteration and the hybrid iteration algorithm is presented as follows. First, choose positive semidefinite $\tilde{P}_i^{[0]} \in C(R^n)$, select $\hat{Q}_i(x)$ satisfying $\hat{Q}_i(x) - Q_i(x) > 0$. It has the iterative update rule of the control policy, as follows

$$\tilde{w}_i^{[n+1]} = -\frac{1}{2}R_{i,i}^{-1}g_i^T \nabla \tilde{P}_i^{[n]} \tag{9}$$

The value function is renewed by

$$\tilde{P}_i^{[n+1]} = \varepsilon_i^{[n]} \nabla \tilde{P}_i^{[n]T} f_i(s_i) + \hat{Q}_i(x) - \tilde{w}_i^{[n+1]T} R_{i,i} \tilde{w}_i^{[n+1]} + \tilde{P}_i^{[n]} \tag{10}$$

where $\varepsilon_i^{[n]}$ is a deterministic sequence, which satisfies

$$\begin{cases} \varepsilon_i^{[n]} > 0 \\ \sum_{n=0}^{\infty} \varepsilon_i^{[n]} = \infty \\ \lim_{n \rightarrow \infty} \varepsilon_i^{[n]} = 0 \end{cases} \tag{11}$$

The cyclic search process (9) and (10) of the admissible control policy will stop until

$$\tilde{P}_i^{[n+1]} - \tilde{P}_i^{[n]} \leq \varepsilon_i^{[n]} (\hat{Q}_i(x) - Q_i(x)) \tag{12}$$

Then, based on the admissible control law (9), it obtains the exploring process of the optimal control policy based on the following two iterative equations. The iteration cost function $P_i^{[n]}$ is solved from

$$\mathfrak{H}(s_i, \nabla P_i^{[n]}, w_i^{[n]}) = 0 \tag{13}$$

and the control policy can be updated according to the following equation

$$w_i^{[n+1]} = -\frac{1}{2}R_{i,i}^{-1}g_i^T \nabla P_i^{[n]} \tag{14}$$

Through the above two steps of policy evaluation and policy improvement, the optimal policy sequence and its corresponding value function are finally reached or converge.

Note that the hybrid iteration (9)–(14) still faces a strict requirement, since it still requires precise knowledge of system dynamics models. For many complex practical systems, it is not easy to acquire accurate knowledge of system dynamics models. Even when some degree of knowledge of system dynamics models can be obtained in some cases, it is a challenge to effectively apply this knowledge to hybrid iteration because, in practice, the model needs to be simplified or approximated in order to be able to perform effective

iterative calculations. However, such simplifications or approximations may introduce errors that affect the accuracy of the iterative results. Hence, to avoid dependence on the system model, the hybrid iteration (9)–(14) will be improved, and the hybrid iteration based on state and input information will be developed and proposed.

Taking the derivative of the cost function P_i has

$$\dot{V}_i(s_i) = \nabla P_i \dot{s}_i = \mathfrak{H}(s_i, \nabla P_i, w_i) - Q(s_i) + \hat{Q}(s_i) - \hat{Q}(s_i) - w_i^T R_{i,i} w_i \quad (15)$$

Here, define $\bar{\mathfrak{H}}(s_i, \nabla P_i, w_i) = \mathfrak{H}(s_i, \nabla P_i, w_i) - Q(s_i) + \hat{Q}(s_i)$, then (15) is

$$\dot{V}_i(s_i) = \bar{\mathfrak{H}}(s_i, \nabla P_i, w_i) - \hat{Q}(s_i) - w_i^T R_{i,i} w_i \quad (16)$$

Since the system lacks a dynamical model, the system data are used to construct the approximated functions. Without loss of generality, we take the i th generation as an example, the approximated functions $\hat{V}_i(s_i)$ and $\hat{\mathfrak{H}}(s_i, \nabla P_i, w_i)$ are given by

$$\hat{V}_i(s_i) = \sum_{j=1}^M \hat{W}_j \phi_j(s_i) \quad (17)$$

and

$$\hat{\mathfrak{H}}(s_i, \nabla P_i^{[n]}, w_i^{[n]}) = \sum_{j=1}^{M_0} \hat{\eta}_j \varphi_j(s_i) + \sum_{j=1}^{M_1} \hat{\gamma}_j \psi_j(s_i) W_i + w_i^T R_{i,i} w_i \quad (18)$$

where $\phi_j(s_i)$, $\varphi_j(s_i)$ and $\psi_j(s_i)$ are linearly independent basis functions and \hat{W}_j , $\hat{\eta}_j$ and $\hat{\gamma}_j$ are weight matrices. The differential equations are used to update the weight matrices.

$$\frac{d}{ds} \hat{W} = \int_0^{t_f} \Phi \Phi^T dt \int_0^{t_f} \Phi \hat{\mathfrak{H}}(s_i, \nabla P_i^{[n]}, \hat{w}_i^{[n]}) dt \quad (19)$$

where $\hat{w}_i^{[n]} = -\frac{1}{2} R_{i,i}^{-1} \left(\sum_{j=1}^{M_1} \hat{\gamma}_j \psi_j(s_i) \right)^T$ and $\Phi = [\phi_1 \quad \phi_2 \quad \phi_M]^T$. Over the procedure, an essentially bounded input such as exploration noise is used to update the weights during a time interval. The iterative process from (17)–(19) are operated until an admissible control is gained. Based on the admissible control, the system (1) is rewritten as

$$\dot{s}_i = f_i(s_i) + g_i(s_i) w_i^{[n]} + g_i(s_i) (w_i - w_i^{[n]}) \quad (20)$$

Then, along the solutions (20) based on (13) and (14), it has

$$\begin{aligned} \dot{P}_i^{[n]}(s_i) &= \nabla P_i^{[n]T} (f_i + g_i w_i^{[n]}) + \nabla P_i^{[n]T} (g_i(s_i) (w_i - w_i^{[n]})) \\ &= -Q(s_i) - w_i^{[n]T} R_{i,i} w_i^{[n]} - 2w_i^{[n+1]T} R_{i,i} (w_i - w_i^{[n]}) \end{aligned} \quad (21)$$

Over any time interval $[t_k, t_{k+1}]$, we integrate both sides of (21) and it has

$$\begin{aligned} &P_i^{[n]}(s_i(t_k)) - P_i^{[n]}(s_i(t_{k-1})) \\ &= \int_{t_{k-1}}^{t_k} -Q(s_i) - w_i^{[n]T} R_{i,i} w_i^{[n]} - 2w_i^{[n+1]T} R_{i,i} (w_i - w_i^{[n]}) dt \end{aligned} \quad (22)$$

and

$$\hat{w}_i^{[n]} = -\frac{1}{2} R_{i,i}^{-1} \left(\sum_{j=1}^{M_1} \hat{\gamma}_j \psi_j(s_i) \right)^T \quad (23)$$

From the expression of the above formula, it can be seen that the hybrid iteration scheme in terms of ADP is based on data and does not require accurate parameters of the microgrid model to ensure that the microgrid system reaches the Nash equilibrium situation.

Merge (22) with (23) and obtain

$$\begin{aligned} & \sum_{j=1}^M \hat{w}_j^{[n]} (\phi_j(s_i(t_k)) - \phi_j(s_i(t_{k-1}))) \\ &= \int_{t_{k-1}}^{t_k} (-Q(s_i) - w_i^{[n]T} R_{i,i} w_i^{[n]}) dt \\ & \quad - \int_{t_{k-1}}^{t_k} 2 \left(\sum_{j=1}^{M_1} \hat{\gamma}_j \hat{\psi}_j(s_i) \right)^T R_{i,i} (w_i - w_i^{[n]}) dt + e_{i,k}^{[n]} \end{aligned} \tag{24}$$

where $e_{i,k}^{[n]}$ is the error generated in the approximation process.

It can be seen that $\hat{w}_j^{[n]}$ and $\hat{\gamma}_j$ are obtained from (24) in a least squares sense, which is making the $\sum_{j=1}^l e_{i,k}^2$ minimal.

For the microgrids with N-distributed generations, voltage recovery control is very important. In order to maintain the stability of the power system, it is necessary to adopt appropriate control algorithms to ensure that the voltage reaches the best situation. The hybrid iteration algorithm based on ADP proposed in this paper is an efficient learning control strategy for voltage recovery. It gradually adjusts the power output of the microgrid in an iterative manner to achieve optimal voltage recovery. Next, the learning process of the algorithm is demonstrated intuitively through Algorithm 1.

Algorithm 1 The hybrid iteration algorithm for microgrids with N-distributed generations

Initialization:

- Select an initial system state $s_i(0)$;
- Select a positive semidefinite $\tilde{P}_i^{[0]} \in \mathcal{C}(R^n)$;
- Select an iteration precision $\epsilon_e > 0$;

Main loop:

- 1: **while** $\|w_i^{n*} - w_i^*\| > \epsilon_e$ or $\|P_i^{n*} - P_i^*\| > \epsilon_e$ **do**
 - 2: Update the control policy in Equation (23)
 - 3: Update the cost function in Equation (17)
 - 4: **end while**
 - 5: **return** w_i^* and P_i^*
-

At the beginning, the hybrid iteration algorithm is initialized and the relevant parameters are set. The iterative process is then entered and in each iteration, the algorithm calculates the power output scheme of each microgrid according to the current parameters and the real-time operating state of the microgrid. These schemes are improved by the proposed optimization algorithm to ensure the effect of voltage recovery. As the number of iterations increases, algorithm 1 will continuously adjust the power output scheme of the microgrid to achieve voltage recovery. At the same time, algorithm 1 dynamically adjusts parameters and strategies according to the real-time operating state of the microgrid and changes in the external environment. When the algorithm meets the convergence condition, the final voltage recovery scheme is output and applied to the microgrid system to achieve the optimal voltage recovery situation.

4. Convergence of the Hybrid Iteration

In this section, the convergence of (9)–(14) is analyzed first. Then, we rigorously prove that the property of hybrid iteration with no need for the information with respect to the system physics.

Theorem 1. For the state-space model (1), the value function $P_i^{[n]}$ and the control policy $w_i^{[n+1]}$ are given as in (13) and (14), then $P_i^{[n]}$ is convergent to P_i^* at a quadratic convergence rate and $w_i^{[n]}$ is convergent to w_i^* .

Proof. In [23] and [30], the proof of $\sup_{s_i} \|P_i^{[n]} - P_i^*\| \leq \mathcal{E}$ and $\sup_{s_i} \|w_i^{[n]} - w_i^*\| \leq \mathcal{E}$ has been given as $i \rightarrow \infty$. Note that the conditions of (12) are sufficient to make the following two inequations hold.

$$\tilde{P}_i^{[n]} > 0 \tag{25}$$

and

$$\nabla P_i^{[n-1]T}(f_i + g_i w_i^{[n]}) \leq -u_i^{[n]T} R_{i,i} u_i^{[n]} - Q_i < 0 \tag{26}$$

Therefore, $\exists n_e$, such that an admissible control input law $w_i^{[n_e]}$ is obtained from (9)–(12) with respect to the microgrids with an N-distributed generation system (1). Let the system (1) be driven by the control input $w_i^{[n_e]}$, then, according to (3), integrate (26) with respect to time and you can obtain

$$P_i^{[n-1]}(s_i) \geq J(s_i, w_i^{[n_e]}) \tag{27}$$

That means $w_i^{[n_e]}$ is an admissible control. Then, based on [31], it has $\{w_i^{[n]}\}_{n=n_e}^\infty$ converge point-wise to w_i^* , as $n \rightarrow \infty$, and $\{P_i^{[n]}\}_{n=n_e}^\infty$ converges on P_i^* , as $n \rightarrow \infty$.

Furthermore, according to (13) and (14), the above formula has

$$(\nabla P_i^{[n+1]} - \nabla P_i^{[n]})^T(f_i + g_i \vartheta_i^{[n]}) + (\nabla P_i^{[n]})^T(f_i + g_i \vartheta_i^{[n]}) + r_i(s_i, \vartheta_i) = 0 \tag{28}$$

where $\vartheta_i = \arg \inf_{w_i} \{\nabla P_i^{[n]T}(f_i + g_i w_i) + r_i(s_i, w_i)\}$. It can be seen that (28) is a Newton–Raphson method and the target is to obtain ∇P_i from the nonlinear equation $\inf_{w_i} \{\nabla P_i^{[n]T}(f_i + g_i w_i) + r_i(s_i, w_i)\}$. This explains that the $P_i^{[n]}$ converges in a quadratic rate. This completes the proof. \square

Theorem 2. For the state-space model (1), $\exists \varepsilon_e > 0$, from (17)–(19), there exist N^* , M_0^* and M_1^* , if $N > N^*$, $M_0 > M_0^*$ and $M_1 > M_1^*$, then the control policy obtained by (23) satisfies $\|w_i^{n^*} - w_i^*\| \leq \varepsilon_e$ and the cost function showed in (17) satisfies $\|P_i^{n^*} - P_i^*\| \leq \varepsilon_e$.

Proof. It has been proved in [23] that one can select a large t_f which could make (19) have the uniqueness of the solution. Since $\hat{w}_i^{[n_e]}$ is an admissible control law, and on the basis of [31], the optimal control policy and the corresponding value function can be obtained approximately as n , M_0 and M_1 going to infinity. \square

The non-zero-sum hybrid iteration mode devised in this paper, which is independent of the system model, does not need the system dynamics parameters at all and uses the collected data for training and learning to describe the optimal behavior trajectory under the interaction between individuals. Based on the rigorous analysis of the two theorems, it can be concluded that the control strategy attained through the model-free hybrid iteration scheme is convergent and can make the microgrid system stable.

5. Simulation Results

In this section, the microgrid system with two distributed generations is given to show the effectiveness of the proposed control method. The dynamics of the system are given as [32]

$$f_i(s_i) = \begin{bmatrix} \cos(s_{i,1})s_{i,2} \\ 0.8 \sin(s_{i,1}) \end{bmatrix} \tag{29}$$

and

$$g_i(s_i) = \begin{bmatrix} 0 \\ \cos(s_{i,2}) \end{bmatrix} \quad (30)$$

In the simulation experiment, the initial states of the above system are given as $s_1(0) = [0.6, -1.1]^T$ and $s_2(0) = [-1, 1]^T$. In the networks, the activation functions are selected as $\phi_j = \varphi_j = \psi_j = [s_{i,1}^2, s_{i,1}s_{i,2}, s_{i,2}^2]^T$. To indicate the effectiveness and feasibility of the presented hybrid iteration method, the optimal control protocols obtained by the value iteration and policy iteration are also considered. Then, the number costed by the learning iterations and the average CPU time are analyzed and summarized for the hybrid iteration as well as for the value iteration and policy iteration as in Table 1.

Table 1. Comparison results of hybrid iteration, value iteration and policy iteration.

	Average CPU-Time Elapsed for Convergence	Number of Learning Iterations
Hybrid-iteration	0.00142	13%
Value-iteration	0.00815	390%
Policy-iteration	0.00221	14%

According to the results of Table 1, it can be seen that hybrid iteration spends less CPU time and iterations to obtain the same simulation result. Therefore, hybrid iteration is a more powerful tool, it uses less numbers for learning iterations required for convergence and costs less average CPU time until convergence. Furthermore, the previous information of the admissible control law is dispensable to initiate hybrid iteration.

The state trajectories of each generation are given in Figures 1 and 2. The system state described in Figure 1 mainly represents the output voltage value of the microgrid system, and the stability of the output voltage is directly related to the normal operation of each device in the microgrid system and the user's power consumption experience. The system state described in Figure 2 is the derivative of the amplitude, reflecting the rate of change of the system state, which is of great significance for the analysis of the dynamic performance and stability of the system. It can be seen that under the action of the algorithm designed in this paper, the system state converges to zero at a faster speed and finally reaches a steady situation. The results also show that the mathematical model established in this paper is very effective for the microgrid system. Based on the control strategy achieved through hybrid iteration, the model can guarantee the stability and realize the reliability of the microgrid system with relatively low computational costs.

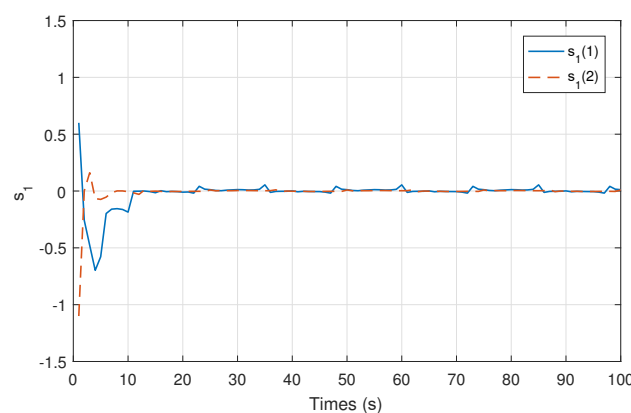


Figure 1. The state s_1 .

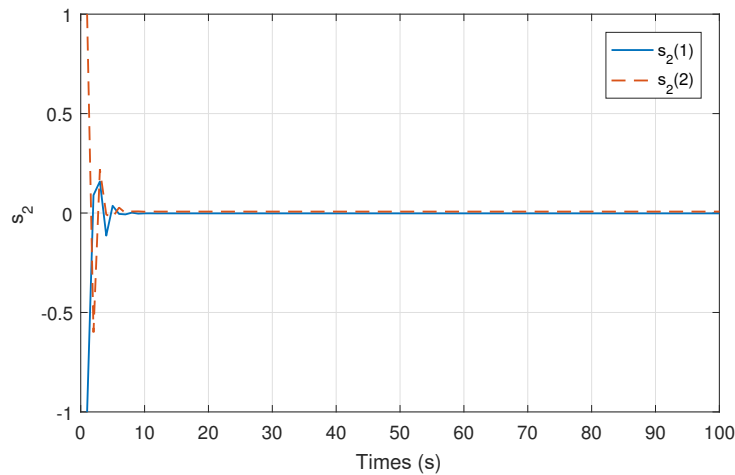


Figure 2. The state s_2 .

The voltage recovery control protocol input trajectories by hybrid iteration are displayed in Figure 3. The designed controller can generate a continuous control input to the microgrid system, avoid chattering in the system state and make the system run stably. It is worth noting that the value function obtained by the convergent optimal control strategy applied to the microgrid system is also optimal. Combining Figures 1–3, it can be found that the proposed method saves a lot of learning iterations required for convergence.

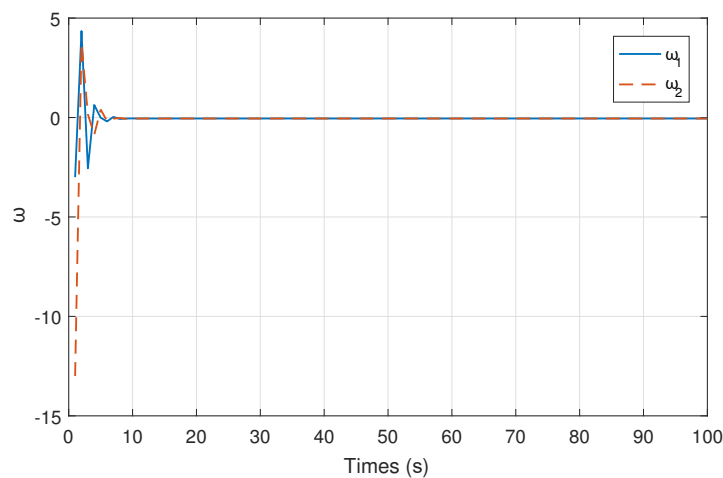


Figure 3. The control input w .

Figures 4 and 5 depict the value function surfaces according to the hybrid iteration-based control law designed by this paper. These surfaces not only reflect the dynamic characteristics of the control law, but also show the evolution of the performance index function during the control process. The performance index function is affected by the punishment or reward generated by the controlled object in different stages and updates the parameters based on the principle of optimality. The critic network guides the action network to approximate the optimal control strategy; in other words, the approximation of the control law is carried out on the basis of the estimation of the performance index function, so the dynamic change of the value function is very important. It can be intuitively seen from these two 3D graphs that both sets of value functions are monotone and non-increasing, which means that after the finite times of hybrid iteration and strategy update designed in this paper, the value function gradually converges to the optimal.

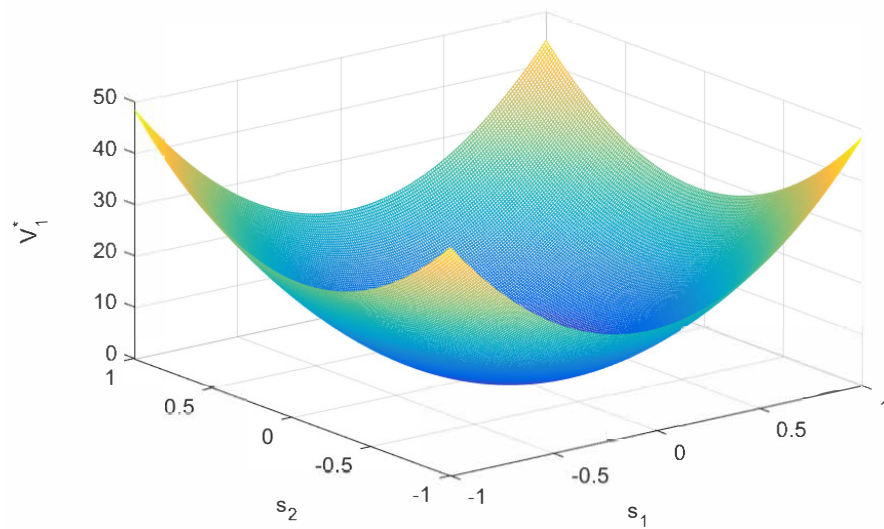


Figure 4. The optimal value function V_1^* .

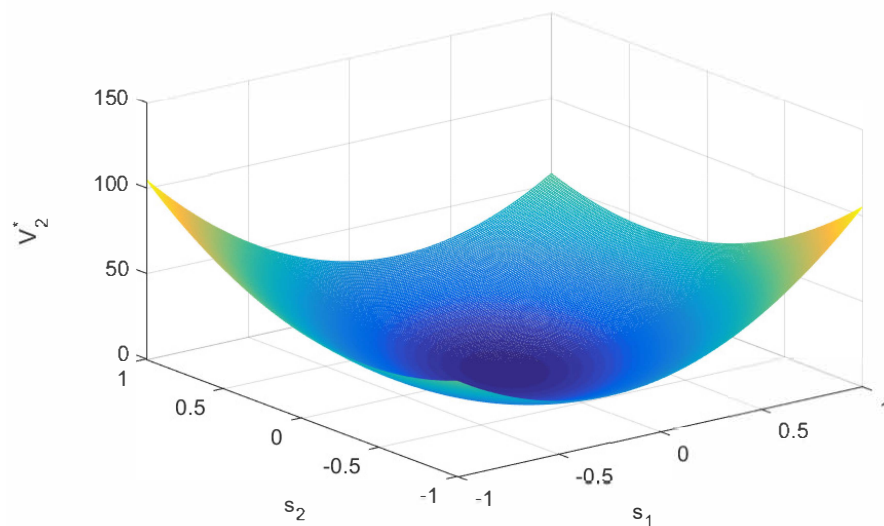


Figure 5. The optimal value function V_2^* .

The data statistics and simulation results in this section show that the ADP algorithm based on hybrid iteration designed in this paper can make the microgrid system reach the control target with fewer iterations and has the advantages of fast computing speed and low implementation cost. In general, the proposed optimal voltage recovery learning scheme can be applied to microgrid systems and the control effect is satisfactory.

6. Conclusions

In view of the complexity of the actual microgrid system structure and the sensitivity to external environmental interference, the uncertainty contained in it is often difficult to predict. This paper solves the optimal control problem of the microgrids with N-distributed generations system with unknown dynamics by means of a non-zero-sum differential game and a novel optimal control law has been put forward for dynamic optimization and adaptive regulation of microgrid systems by introducing an adaptive hybrid iteration algorithm. It is worth noting that the control strategy in this paper is not based on the unknown physical characteristics of the system, but is obtained in real time according to the system state and input information data. This makes the control strategy more

targeted and practical, and it can be dynamically adjusted according to the actual operation. In the simulation result, it can be seen that the hybrid iteration algorithm has a much faster convergence process to the optimal solution than value iteration. Furthermore, it requires tremendously less consumption of time of CPU and less iterations number in the course of learning. In the future, we will further optimize the control strategy and improve the convergence speed and computational efficiency of the algorithm in order to solve more uncertainty problems that may be encountered in the actual operation of the microgrid system.

Author Contributions: Conceptualization, L.L.; methodology, R.S.; software, S.D.; validation, L.L., R.S. and S.D.; writing—original draft preparation, R.S.; writing—review and editing, L.L.; visualization, S.D.; supervision, L.L.; funding acquisition, R.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation grant number 62273036.

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: Author Shuqing Dong was employed by the company China Datang Group Science and Technology Research Institute Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Mu, C.; Zhang, Y.; Jia, H.; He, H. Energy-Storage-Based Intelligent Frequency Control of Microgrid with Stochastic Model Uncertainties. *IEEE Trans. Smart Grid* **2020**, *11*, 1748–1758. [[CrossRef](#)]
2. Massenio, P.R.; Naso, D.; Lewis, F.L.; Davoudi, A. Assistive Power Buffer Control via Adaptive Dynamic Programming. *IEEE Trans. Energy Convers.* **2020**, *35*, 1534–1546. [[CrossRef](#)]
3. Zhang, B.; Dou, C.; Yue, D.; Zhang, Z.; Zhang, T. Hierarchical Control Strategy for Networked DC Microgrid Based on Adaptive Dynamic Program and Event-Triggered Consensus Algorithm Considering Economy and Actuator Fault. *J. Frankl. Inst.* **2020**, *357*, 8631–8656. [[CrossRef](#)]
4. Yin, L.; Yu, T.; Yang, B.; Zang, X. Adaptive Deep Dynamic Programming for Integrated Frequency Control of Multi-Area Multi-Microgrid Systems. *Neurocomputing* **2019**, *344*, 49–60. [[CrossRef](#)]
5. Hu, J.; Sun, Q.; Wang, R.; Wang, Y. An Improved Privacy-Preserving Consensus Strategy for AC Microgrids Based on Output Mask Approach and Node Decomposition Mechanism. *IEEE Trans. Autom. Sci. Eng.* **2024**, *21*, 642–651. [[CrossRef](#)]
6. Wang, R.; Li, J.; Sun, Q.; Wang, P. Current Edge-Control Strategy for Multiple Energy Routers Based on Cyber-Energy Dual Modulations. *IEEE Trans. Ind. Electron.* **2024**, *71*, 4079–4088. [[CrossRef](#)]
7. Hirase, Y.; Abe, K.; Sugimoto, K.; Sakimoto, K.; Bevrani, H.; Ise, T. A Novel Control Approach For Virtual Synchronous Generators To Suppress Frequency And Voltage Fluctuations In Microgrids. *Appl. Energy* **2018**, *210*, 699–710. [[CrossRef](#)]
8. Ma, D.; Cao, X.; Sun, C.; Wang, R.; Sun, Q.; Xie, X.; Peng, W. Dual-Predictive Control With Adaptive Error Correction Strategy For AC Microgrids. *IEEE Trans. Power Deliv.* **2021**, *Early Access*. [[CrossRef](#)]
9. Mehrizi-Sani, A.; Iravani, R. Online Setpoint Adjustment For Trajectory Shaping In Microgrid Applications. *IEEE Trans. Power Syst.* **2012**, *27*, 216–223. [[CrossRef](#)]
10. Huang, B.; Zhao, T.; Yue, M.; Wang, J. Bi-Level Adaptive Storage Expansion Strategy for Microgrids Using Deep Reinforcement Learning. *IEEE Trans. Smart Grid* **2024**, *15*, 1362–1375. [[CrossRef](#)]
11. Lee, S.; Seon, J.; Sun, Y.G.; Kim, S.H.; Kyeong, C.; Kim, D.I.; Kim, J.Y. Novel Architecture of Energy Management Systems Based on Deep Reinforcement Learning in Microgrid. *IEEE Trans. Smart Grid* **2024**, *15*, 1646–1658. [[CrossRef](#)]
12. Shuai, Z.; Sun, Y.; Shen, Z.J.; Tian, W.; Tu, C.; Li, Y.; Yin, X. Microgrid stability: Classification and a review. *Renew. Sustain. Energy Rev.* **2016**, *58*, 167–179. [[CrossRef](#)]
13. Joung, K.; Kim, T.; Park, J. Decoupled Frequency And Voltage Control For Stand-Alone Microgrid With High Renewable Penetration. *IEEE Trans. Ind. Appl.* **2019**, *55*, 122–133. [[CrossRef](#)]
14. Yu, Y.; Liu, G.; Huang, Y.; Guerrero, J.M. Coordinated Predictive Secondary Control for DC Microgrids Based on High-Order Fully Actuated System Approaches. *IEEE Trans. Smart Grid* **2023**, *15*, 19–33. [[CrossRef](#)]
15. Naderi, E.; Asrari, A.; Ramos, B. Moving Target Defense Strategy to Protect a PV/Wind Lab-Scale Microgrid Against False Data Injection Cyberattacks: Experimental Validation. In Proceedings of the 2023 IEEE Power & Energy Society General Meeting (PESGM), Orlando, FL, USA, 16–20 July 2023; pp. 1–5.
16. Taghizadeh, S.; Karimi-Ghartemani, M.; Hossain, M.; Lu, J. A Fast and Robust DC-Bus Voltage Control Method for Single-Phase Voltage-Source DC/AC Converters. *IEEE Trans. Smart Grid* **2019**, *34*, 9202–9212. [[CrossRef](#)]

17. Wei, Q.; Liu, D.; Lewis, F.L.; Liu, Y.; Zhang, J. Mixed Iterative Adaptive Dynamic Programming for Optimal Battery Energy Control in Smart Residential Microgrids. *IEEE Trans. Ind. Electron.* **2017**, *64*, 4110–4120. [[CrossRef](#)]
18. Zhang, H.; Hu, X.; Ma, D.; Wang, R.; Xie, X. Insufficient Data Generative Model for Pipeline Network Leak Detection Using Generative Adversarial Networks. *IEEE Trans. Cybern.* **2022**, *52*, 7107–7120. [[CrossRef](#)]
19. Wei, Q.; Liao, Z.; Shi, G. Generalized Actor-Critic Learning Optimal Control in Smart Home Energy Management. *IEEE Trans. On Industrial Inform.* **2021**, *17*, 6614–6623. [[CrossRef](#)]
20. Chen, L.; Li, Y.; Huang, M.; Hui, X.; Gu, S. Robust Dynamic State Estimator of Integrated Energy Systems Based on Natural Gas Partial Differential Equations. *IEEE Trans. Ind. Appl.* **2022**, *58*, 3303–3312. [[CrossRef](#)]
21. Dong, W.; Zhang, L.; Zhang, D.; Cui, Q. Hierarchical Multi-objective Planning for Integrated Energy Systems in Smart Parks Considering Operational Characteristics. *CSEE J. Power Energy Syst.* **2022**, *8*, 1760–1771.
22. Wang, S.; Zhai, J.; Hui, H.; Ding, Y.; Song, Y. Operational Reliability of Integrated Energy Systems Considering Gas Flow Dynamics and Demand-Side Flexibilities. *IEEE Trans. Ind. Inform.* **2023**, *20*, 1360–1373. [[CrossRef](#)]
23. Qasem, O.; Gao, W.; Vamvoudakis, K.G. Adaptive Optimal Control Of Continuous-Time Nonlinear Affine Systems Via Hybrid Iteration. *Automatica* **2023**, *157*, 111261. [[CrossRef](#)]
24. Bertsekas, D.P. Value and Policy Iterations in Optimal Control and Adaptive Dynamic Programming. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, *28*, 500–509. [[CrossRef](#)] [[PubMed](#)]
25. Bian, T.; Jiang, Z.-P. Value Iteration And Adaptive Dynamic Programming For Data-Driven Adaptive Optimal Control Design. *Automatica* **2016**, *71*, 348–360. [[CrossRef](#)]
26. Qasem, O.; Davari, M.; Gao, W.; Kirk, D.R.; Chai, T. Hybrid Iteration ADP Algorithm to Solve Cooperative, Optimal Output Regulation Problem for Continuous-Time, Linear, Multiagent Systems: Theory and Application in Islanded Modern Microgrids With IBRs. *IEEE Trans. Ind. Electron.* **2024**, *71*, 834–845. [[CrossRef](#)]
27. Wei, Q.; Zhu, L.; Song, R.; Zhang, P.; Liu, D.; Xiao, J. Model-Free Adaptive Optimal Control for Unknown Nonlinear Multiplayer Nonzero-Sum Game. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *33*, 879–892. [[CrossRef](#)] [[PubMed](#)]
28. Tong, X.; Ma, D.; Wang, R.; Xie, X.; Zhang, H. Dynamic Event-Triggered-Based Integral Reinforcement Learning Algorithm for Frequency Control of Microgrid With Stochastic Uncertainty. *IEEE Trans. Consum. Electron.* **2023**, *69*, 321–330. [[CrossRef](#)]
29. Liu, G.; Zou, L.; Sun, Q.; Wang, R.; Sui, Z. Multiagent-Based Consensus Optimal Control for Microgrid With External Disturbances via Zero-Sum Game Strategy. *IEEE Trans. Control Netw. Syst.* **2023**, *10*, 659–670. [[CrossRef](#)]
30. Bian, T.; Jiang, Z. Reinforcement Learning and Adaptive Optimal Control for Continuous-Time Nonlinear Systems: A Value Iteration Approach. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *33*, 2781–2790. [[CrossRef](#)]
31. Jiang, Y.; Jiang, Z. Robust Adaptive Dynamic Programming And Feedback Stabilization Of Nonlinear Systems. *IEEE Trans. Neural Netw. Learn. Syst.* **2014**, *25*, 882–893. [[CrossRef](#)]
32. Liu, G.; Sun, Q.; Wang, R.; Hu, X. Nonzero-Sum Game-Based Voltage Recovery Consensus Optimal Control for Nonlinear Microgrids System. *IEEE Trans. Neural Netw. Learn. Syst.* **2023**, *34*, 8617–8629. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.