EAI Endorsed Transactions
on Energy Web

Research Article EAI.EU

Reinforcement Learning Data-Driven Optimal Load-Frequency Control for Power Systems

Yi Zhao

College of Engineering, Northeast Agricultural University, Harbin 150000, China

Abstract

INTRODUCTION: Power systems are complex due to their time-varying and uncertain parameters, challenging conventional control methods.
OBJECTIVES: This study proposes an adaptive dynamic programming (ADP) controller to address this limitation. The ADP controller eliminates the need for pre-existing knowledge of the system dynamics, a significant advantage in real-world applications.
METHODS: By iteratively solving the Riccati equation using only system state and input data, the controller learns an approximate optimal control strategy. In this study, we use an iterative computational approach with an online adaptive optimal controller designed for unknown power system dynamics.
RESULTS: Utilizing real-time collected system states and input information, even in the absence of knowledge about the power system matrix, we achieve iterative solutions for the algebraic Riccati equation, enabling the computation of an optimal controller. Simulation results demonstrate the ease of implementation of this approach in power system load frequency control (LFC).
CONCLUSION: The proposed ADP controller exhibits good control performance of grid stability, making it a valuable reference for LFC, especially in scenarios with unknown system parameters.

Keywords: Power system, reinforcement learning, data-driven, dynamic programming, Load frequency control

Received on 24 October 2024, accepted on 19 February 2025, published on 04 March 2025

doi: 10.4108/ew.7500

*Corresponding author. Email: Yi11Zhao@outlook.com

1. Introduction

The power system represents a dynamic and extensive system with substantial dimensions, exhibiting pronounced non-linearity, time variations, and unidentified parameters. It encompasses numerous unmodeled dynamic components, making efficient control a formidable challenge [1] . Nowadays, the expanding power system, characterized by large-capacity units, high-voltage grids, AC and DC interconnections and the utilization of new energy sources, such as solar energy, has exacerbated challenges in terms of security and stabilization. The power system's failure or random fluctuations, if not appropriately controlled, could precipitate system collapse, severely impacting daily life and potentially resulting in substantial economic losses. Consequently, effective control of the power system to ensure the stability and safety of its operation is of practical importance in the current[2][3][4][5][6].

Voltage frequency stands out as one of the most crucial indices of the power system. For the purpose of solving those problems effectively and provide reliable and stable power, the LFC design of power system is the key[7][8]. In this regard, the current more traditional method is based on the classical controller to curtail the regional control fault of the power system and improve the frequency response. Numerous scholars have delved into research on it[9][10][11][12]. However, for increasingly complex power systems, the traditional control methods have exposed problems such as slow response speed, low fault tolerance of the actuator and lack of robustness to parameter uncertainties. As a result, the optimal control method of load frequency has been spawned, capturing the interest of numerous scholars in recent years[13][14].

Optimal control, alternatively termed dynamic or process optimization, addresses the fundamental question of determining the optimal control law or control strategy depended on the dynamic characteristics of the controlled system within defined constraints. This aims to enable the system to operate in accordance with specific technical requirements, ensuring it attains the optimal value as per the specified performance index or objective function[15][16][17].

Aiming at the optimal control of load frequency, Vahid Gholamrezaie proposes an optimal control way for the dynamic load frequency control model, and uses PSO algorithm to accelerate the response time of the model[18] .The optimal regulator theory for the LFC is used in[19]. LFC in an economic dispatch perspective is applied to[20]. Liu F combines economic indicators and stability indicators to build an optimal load control framework for restructuring power systems[21].The optimal robust regulator for system LFC issue was proposed by Rahmani M by using a two-level strategy[22].

However, in the case of solving the optimal control issue of the dynamic system, the above method requires real-time observation of the dynamic state of the system, and the traditional dynamic programming method is often plagued by the ‘ dimension disaster ‘ problem, which makes its application limited[23][24]. Meanwhile, the common dynamic programming algorithm assumes that the parameters of the system are completely known. For the power system, this assumption is not convenient to apply in real life. In order to overcome this challenge, adaptive dynamic programming technology[25]. came into being. It is a new control technology that integrates dynamic programming, reinforcement learning, adaptive control, optimal control and other theories and methods. It builds a more practical theoretical system that provides a broader perspective on the solution of such problems. Its core idea is to obtain the system state and input data in real time, and dynamically adjust the control strategy iteration process to adapt to the dynamic changes of the system, thereby enhancing the robustness and performance.

Recently, the ADP method for discrete-time systems has garnered significant interest among numerous researchers, as follows : WeiQinglai(2010)[26], linxiaofeng(2011)[27], hamingmig(2023)[28], Ni, He, Zhong, and Prokhorov(2015)[29] and so on. Since many systems in practice can be modeled and represented as a continuous time system, adaptive dynamic programming algorithms for such systems have been widely studied, such: A model applicable to control parameters over a range is proposed in 2022, bridging the local optimization problem[30], there are also scholars who combine dynamic planning and heuristic optimization according to the characteristics of continuous-time systems to improve the overall performance of control and so on[31][32].

In the above literature review, we note that the application of adaptive dynamic programming in power system load frequency control is relatively limited. Such limitations and challenges may lead to poor performance of the system in dealing with complex low-frequency disturbances, and affect the stable operation and robustness of the system. In view of the existing situation, this paper aims to study the application of ADP in LFC of power system.

The primary contributions of this paper are, first, the successful elimination of the dependence of traditional dynamic programming algorithms on a priori information about the system dynamics, paving the way for the application of the ADP algorithm. Second, real-time information about the states and inputs is incorporated into the iterative solution process of the Riccati formulation, thus eliminating the need for prior knowledge of the system matrix. Third, compared with the traditional method, the method can reuse the same state and input data in iterations at fixed time intervals, thus avoiding the difficult problem of dimensional catastrophe, making it more applicable to real power system states, and improving the algorithm’s computational adaptability and practicality. The innovation of this study primarily lies in the novelty of the algorithm itself. It is to be noted that the algorithms in this paper are model-free based, unlike traditional control methods that must require the system model to be known, and thus do not lend themselves to direct comparisons with these methods. Simulation results show that the algorithm is suitable for the complex LFC domain in power systems. In the context of the increasing complexity of power systems, this study is of great significance for understanding and applying more generalized ADP algorithms.

The document’s structure is outlined as outlined below. In Section 2, a succinct introduction is provided, the policy iteration approach is applied to the conventional linear optimal control problem relevant to continuous-time power systems. Transitioning to Section 3, a new computational ADP approach is introduced, showcasing its convergence. Furthermore, an online algorithm with practical implications is presented. Section 4 delves into the application of the developed method to the optimal design of the regulator challenge in a power system. In Section 5, we conclude with some concluding remarks and discuss potential extensions.

Notation

We utilize R representing the real numbers set, $\otimes$ is the tensor product of matrices, $\|\bullet\|$ is the norm. Z^+ stands for the positive integers and zero, and $vec(A) = [a_1^T a_1^T \cdots a_m^T]^T$ where $a_i \in \mathbb{R}^n$ are the columns of A.

2. Materials and Methods

This paper will study a power system[33],which is well-documented in existing literature. In the situation of normal power system operation, disturbances are generally of small magnitude and their impact on the state of the system is relatively limited. In view of this, the above system can be generalized as shown below. The model block diagram is shown in Fig.1:

$$\Delta \dot{P}_g = -\frac{1}{T_T}\Delta P_g + \frac{1}{T_T}\Delta X_g \tag{1}$$

$$\Delta \dot{f} = -\frac{1}{T_p}\Delta f + \frac{K_p}{T_p}\Delta P_g \tag{2}$$

$$\Delta\dot{E} = K_E \Delta f \quad (3)$$

$$\Delta\dot{X}_g = -\frac{1}{R_\tau T_G}\Delta f - \frac{1}{T_G}\Delta X_g - \frac{1}{T_G}\Delta E + \frac{1}{T_G}u \quad (4)$$

The above symbols are described as follows, ΔP_g is the change of the generator output, Δf represents the frequency deviation, ΔE means the integral control, ΔX_g indicates the change of the regulator position, T_T is the time constant of turbine, T_G is the time constant of governor , T_p is the time constant of the model, R_τ represents the speed regulation. K_E and K_p represent the corresponding gain. For the convenience of readers, Table 1 provides the symbols used in the manuscript and their corresponding definitions for quick reference.

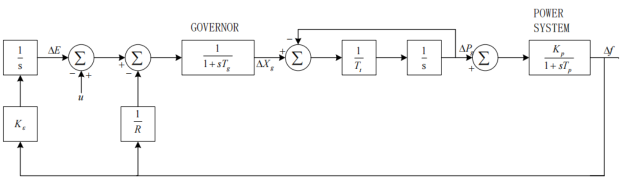

Fig. 1. Power system model sketch

The above model (1)-(4) can be denoted as Eq. (5):

$$\dot{x} = Ax + Bu \quad (5)$$

where, $x = [\Delta f, \Delta P_g, \Delta X_g, \Delta E]^T$, $B \in \mathbb{R}^{n\times m}$, $A \in \mathbb{R}^{n\times n}$ are unknown matrix. $u \in \mathbb{R}^m$ as the control data. extraordinary, this paper assumes that the system is stable.

$$A = \begin{bmatrix} -\frac{1}{T_p} & \frac{K_p}{T_p} & 0 & 0 \\ 0 & -\frac{1}{T_T} & -\frac{1}{T_T} & 0 \\ -\frac{1}{RT_G} & 0 & -\frac{1}{T_G} & -\frac{1}{T_G} \\ K_E & 0 & 0 & 0 \end{bmatrix}$$

$$B^T = [0 \quad 0 \quad \frac{1}{T_G} \quad 0]$$

The main objectives of this paper are to address the specified controller:

$$u = -Nx \quad (6)$$

Make the following evaluation index function minimum

$$J = \int_0^\infty (x^T Qx + u^T Ru)dt \quad (7)$$

where $Q^T = Q \geq 0$, $R^T = R > 0$ with $(A, Q^{1/2})$ observable.

Table 1: Notation and Corresponding Definitions

Notation	Definition
ΔP_g	the change of the generator output
Δf	frequency deviation
ΔE	integral control
ΔX_g	the change of the regulator position
T_T	the time constant of turbine
T_G	the time constant of governor
T_p	the time constant of the model
R_τ	the speed regulation
K_E	the corresponding gain
K_p	the corresponding gain

In accordance with the linear optimal control theory proposed by Lewis and Symos, when the accurate information of matrices B and A is available, solving the renowned ARE provides the solution to the problem.

$$A^T M + MA + Q - MBR^{-1}B^T M = 0 \quad (8)$$

where M^* is a matrix of positive definite symmetry, consequently, an optimal feedback matrix N^* can be obtained by using the following equation

$$N^* = R^{-1}B^T M^* \quad (9)$$

Since M in the above Riccati equation is nonlinear, directly solving M* can be challenging, particularly for matrices of substantial size. Nonetheless, numerous effective algorithms have been devised for the numerical approximation of the solution to equation (8). An example of such an algorithm is the one developed in the 70's., which is elucidated below:

Given $N_0 \in R^{m\times n}$ is stable feedback matrix, and if a positive definite symmetric matrix M_i is a solution of the following equation[34].

$$(A - BN_i)M_i + M_i(A - BN_i) + Q + N_i^T RN_i \quad (10)$$

Where, N_i , i =1,2, …, is updated by:

$$N_i = R^{-1}B^T M_{i-1} \quad (11)$$

Among them, the subsequent characteristics are established:

$A - BN_i$ is Hurwitz, $M_i \geq M_{i+1} \geq M^*$ and $\lim_{i\to\infty} N_i = N^*, \lim_{i\to\infty} M_i = M^*$

Based on the fact that it is very difficult to solve Eq. (10) without explicitly defining A, Vrabie et al. researchers used an approach through online information collection as shown in Eq.(12), which successfully reduces the difficulty of the problem and eases the designer's workload.

$$x^T(t)M_i x(t) - x^T(t+\Delta t)M_i x(t+\Delta t) = \int_t^{t+\Delta t}(x^T Qx + u_i^T Ru_i)d\tau \tag{12}$$

where $u_i = -N_i x$ represents the system's manipulated command during $[t, t+\Delta t]$.

Owing to the online measurement of both x and u_i ,a distinctive symmetric solution, denoted as M_i , can be uniquely ascertained in the presence of a persistent excitation (PE) condition. But, as evident from (11), Accurate understanding of the system matrix B remains essential for the iterative process. Furthermore, ensuring the PE condition may require resetting the state at each iteration, potentially posing challenges for stability. Another approach involves incorporating measurement noise, where the control input u_i is given by $u_i = -N_i x + e$, with e representing measurement noise serves as the actual input data in equation (12). Consequently, the solution M_i obtained from (12) differs from that obtained from (10). Additionally, following each revision of the control policy, it is necessary to gather information on both the state and input to facilitate the subsequent iteration This process may impede the learning pace, particularly in higher dimensional system.

3. Results

In the previous chapter, we have mentioned that even when the matrices A and B are completely known, solving Eq. (10) is also a big job and difficult to complete. This section proposes a data-driven solution to Eq. (10) that does not depend on the matrix pair (A, B) information.

First, we assume that the stable N0 is known. And for $i \in Z+$, the matrix M_i satisfying (10) is obtained, and the $N_{i+1} \in \mathbb{R}^{m\times n}$ is obtained by $N_{i+1} = R^{-1}B^T M_i$ iterative updating.

Therefore, we write system (5) as follows:

$$\dot{x} = A_i x + B(N_i x + u) \tag{13}$$

where, $A_i = A - BN_i$

Then, according to the second section, the solution of equation (9) can be expressed as

$$\begin{aligned} x(t)M_i x(t) - x(t+\Delta t)^T M_i x(t+\Delta t) &= -\int_t^{t+\Delta t}[2(u+N_i x)^T B^T M_i x + x^T(A_i^T M_i + M_i A_i)x]d\tau \\ &= -2\int_t^{t+\Delta t}(u+N_i x)^T RN_{i+1}x]d\tau + \int_t^{t+\Delta t} x^T Q_i x d\tau \end{aligned} \tag{14}$$

where $Q_i = Q + N_i^T RN_i$

Remark 1. In Eq.(14), the term involving the unknown matrices A and B, denoted as $x^T(A_i^T M_i + M_i A_i)x$ is replaced by a term that can be obtained by online state measurements. Also, the term $B^T M_n$ is substituted with RN_{i+1}, in which N_{i+1} will be solved for later as an unknown quantity through an iterative formula containing M_i . Hence, Eq. (14) eliminates the need for system matrices A and B for controller parameters solving. Besides, the state and input data of the system will participate in the subsequent iterations of the solution.

Remark 2. It's worth noting that in Eq. (14), equation is consistently maintained when both M_i, N_{i+1} adhere to the conditions specified in (10) and (11). and x for system (13) with any u allows us to utilize $u = -N_0 x + e$ as the input data for learning, where e represents the exploration noise. Importantly, this choice does not compromise the astringency of the learning procedure.

Subsequently, we demonstrate that given a stabilizing N_i , matrices $(M_i \quad N_{i+1})$, with $M_i = N_{i+1}$ meeting (10) and (11) can be only ascertained without knowledge of A or B.

To achieve this, this paper introduces the following two operators:

$$M \in \mathbb{R}^{n\times n} \to \hat{M} \in \mathbb{R}^{\frac{1}{2}n\times(n+1)}, \text{ and } x \in \mathbb{R}^n \to \bar{x} \in \mathbb{R}^{\frac{1}{2}n(n+1)}$$

where

$$\hat{M} = [M_{11}, 2M_{12}, \ldots, 2M_{1n}, M_{22}, 2M_{23}, \ldots, 2M_{n-1}, n, M_{nn}]^T$$

$$\bar{x} = [x_1^2, x_1x_2, \ldots, x_1x_n, x_2^2, x_2x_3, \ldots, x_{n-1}x_n, x_n^2]^T$$

Additionally, through the representation using the tensor product, we obtain $x^T Q_k x = (x^T \otimes x^T)vec(Q_k)$ and

$$(u+N_i x)^T RN_{i+1}x = [(x^T \otimes x^T)(I_n \otimes N_i^T R) + (x^T \otimes u^T)(I_n \otimes R)vec(N_{i+1})]$$

Further, for positive integer l, we define matrices $\Delta_{xx} \in \mathbb{R}^{l\times\frac{1}{2}n\times(n+1)}$, $E_{xx} \in \mathbb{R}^{l\times n^2}$, $E_{xu} \in \mathbb{R}^{l\times mn}$,such that

$$\Delta_{xx} = [\bar{x}(t_1) - \bar{x}(t_0), \bar{x}(t_2) - \bar{x}(t_1), \ldots, \bar{x}(t_l) - \bar{x}(t_{l-1})]^T,$$

$$E_{xx} = [\int_{t_0}^{t_1} x \otimes x d\tau, \int_{t_1}^{t_2} x \otimes x d\tau, \ldots \int_{t_{l-1}}^{t_l} x \otimes x d\tau]^T,$$

$$E_{xu} = [\int_{t_0}^{t_1} x \otimes u d\tau, \int_{t_1}^{t_2} x \otimes u d\tau, \ldots \int_{t_{l-1}}^{t_l} x \otimes u d\tau]^T,$$

where $0 \le t_0 < t_1 < \cdots < t_l$

Later, for any N_i , (14) leads to the matrix form of the following linear equations:

$$\Theta_i \begin{bmatrix} \hat{M}_k \\ vec(N_{i+1}) \end{bmatrix} = \Xi_i \qquad (15)$$

where, $\Theta_i = [\Delta_{xx}, -2E_{xx}(I_n \otimes N_i^T R) - 2E_{xu}(I_n \otimes R)]$, $\Xi_i = -E_{xx} vec(Q_i)$.

It is noteworthy, if matrix Θ_i possesses column full rank, the solution to (15) could be obtained by:

$$\begin{bmatrix} \hat{M}_i \\ vec(N_{i+1}) \end{bmatrix} = (\Theta_i^T \Theta_i)^{-1} \Theta_i^T \Xi_i \qquad (16)$$

Next, we will specifically introduce this ADP algorithm that solves iteratively without relying on the system parameters, and give the operation flow of the algorithm as follows Fig. 2. The paper also includes relevant discussions on the stability and convergence analysis to further validate the robustness and effectiveness of the proposed method.

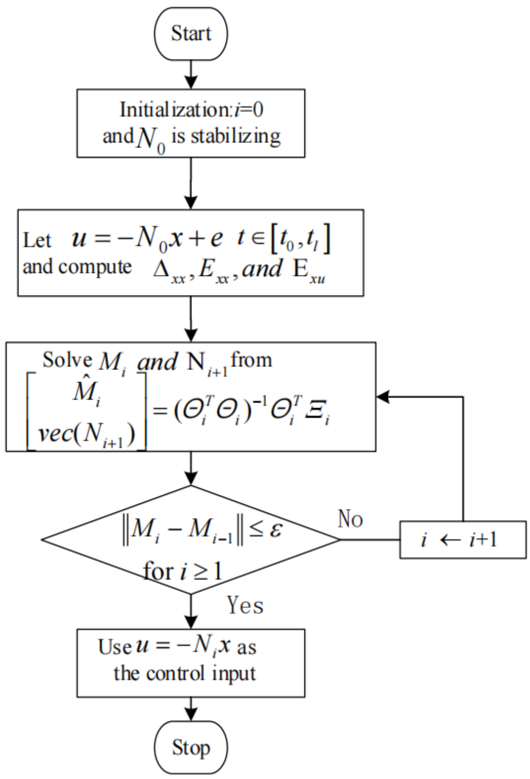

Fig. 2. algorithm flow chart

A. Algorithm

1. Utilizes $u = -N_0 x + e$,where N_0 is steady and e denotes the noise. Iterate the computation of $\Delta_{xx}, E_{xx}, E_{xu}$ until the rank condition specified in (17) below is met. Set the initial value of i to 0.

2. Calculate M_i and N_{i+1} from (16).

3. Let i ← i+1, and return to Step 2 until $\|M_i - M_{i-1}\| \le \varepsilon$ for i≥1, where the constant ε > 0 is the small threshold is already defined.

4. Utilize $u = -N_i x$ as the approximated optimal control approach.

Remark 3. The primary computational load in executing Algorithm is associated with the computation of matrices E_{xx} and E_{xu}, which can be computed employing $\frac{1}{2} n(n+1) + mn$ integrator within the flow path for the accumulation of information regarding the input and state

Remark 4. Practically, numerical error might arise during the computation of E_{xx} and E_{xu}. Consequently, the solution to (15) might not be feasible. In such instances, the key to (16) could be interpreted as the least-squares solution to (15)

Subsequently, we demonstrate that Algorithm converges given a certain rank condition

Lemma 1. Provided an integer $b_0 > 0$ exists where for all $b > b_0$

$$R([E_{xx}, E_{xu}]) = \frac{n(n+1)}{2} + mn \qquad (17)$$

then Θ_i has full column rank for all $N \in \mathbb{R}^{+}$.

Theorem 1. Given a positive definite gain $N_0 \in \mathbb{R}^{m \times n}$, According to Lemma 1,both $\{M_\varsigma\}_{\varsigma=0}^{\infty}$ and $\{N_\upsilon\}_{\upsilon=1}^{\infty}$ will be obtained from (16) and will eventually approximate to the optimal values M^* and N^* individually.

Proof. Given a stabilized feedback matrix N_i. According to the iterative update formula (11) for N, the value of N_{i+1} will be uniquely determined if the solution to Eq. (10) is correctly obtained. By means of Eq. (14), we can determine that M_i and N_{i+1} can be constructed in the form of matrices as follows.

$$\Theta_i \begin{bmatrix} \hat{\mathrm{M}} \\ vec(N) \end{bmatrix} = \Xi_i$$

where, $M = M^T \in \mathbb{R}^{n \times n}$ and $N \in \mathbb{R}^{m \times n}$

Subsequently, we promptly obtain $\hat{M} = \hat{M}_i$ and $vec(N) = vec(N_{i+1})$. In accordance with Lemma 1, N and $M = M^T$ have the only values. Additionally, $M_i = M$ and $N_{i+1} = N$ are distinctly confirmed

Hence, policy iteration using (16) is equivalent to the solution of (10) and (11). According to Theorem 1, convergence is thus demonstrated. Due to the length limitations of the article, more proofs of stability can be found in this reference[35].

Remark 5. It is evident that Algorithm comprises distinct stages. Initially, Control input data contains noise for stability and data is documented in Δxx E_{xx} and E_{xu} until the fulfillment of the rank condition in (17). Subsequently, devoid of the need for extra system details, the matrices Δxx , E_{xx} and E_{xu} are iteratively utilized to execute the process. This results in a series of controllers converging towards the optimal control policy.

4. Discuss

We investigate the design of this power system controller. This system is a fourth-order linear system in continuous time. A and B of the system directly derived from[33], as shown below:

$$A = \begin{bmatrix} -0.0665 & 8 & 0 & 0 \\ 0 & -3.663 & 3.663 & 0 \\ -6.86 & 0 & -13.736 & -13.736 \\ 0.6 & 0 & 0 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & 0 & 13.736 & 0 \end{bmatrix}^T$$

To showcase the validity of the controller in the section 3, here, the design of the optimal controller does not rely on the exact knowledge of A and B. Given the stability of the physical system, the initial N0 = 0.

The weight matrix design is as follows:

$$Q = diag[30, 0.1, 0.1, 1]$$

$$R = I_1 * 200$$

The simulation platform required for the experiment is MATLAB2022b, and the CPU is Inter Core i5-11400H with 2.70 GHz. Throughout the experiment, initial parameter of the state variables is arbitrarily determined in close proximity to the origin. The system input is characterized by the following exploration noise during the time interval from t = 0 to 5 s.

$$e = 100 \sum_{j=1}^{100} \sin(\omega_j t) \quad (18)$$

where ω_j , with j = 1, . . ., 100, are randomly chosen from [-500,500].

At runtime, the algorithm automatically collects the system's state and input parameter information every 0.01s, and the noisy data is removed after 5s. It converges after 16 iterations, and the end of the program satisfies this condition $\|M_i - M_{i-1}\| \le 4.2e^{-10}$.

Fig. 3 illustrates the norm changes of the state during the process time.

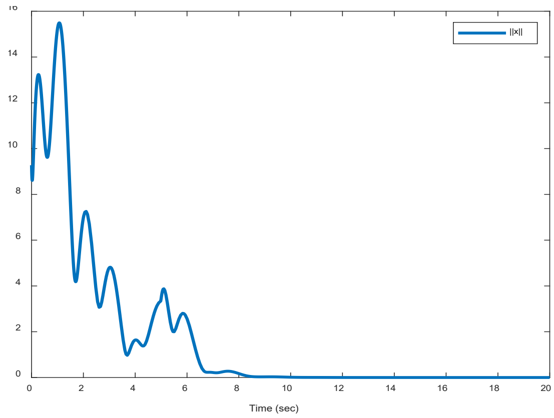

Fig. 3. The norm changes of the State during the process time.

In Fig. 4, for the first 5 s, the changes in the system state parameters are more pronounced due to the presence of the detection noise. after 5 s, the noise is removed and the optimal controller is obtained, under which the changes in the four parameters of the power system finally converge to 0, proving that the control is effective

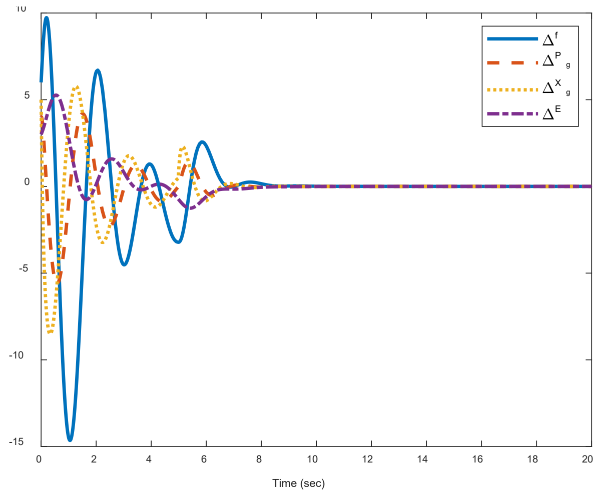

Fig. 4. Power system state parameter changes over time

Fig. 5 and Fig. 6 show the convergence of M_i and N_i to the optimum, respectively, and it can be seen that the paradigms of their differences from the optimum all converge to 0 at the end of the program, demonstrating the effectiveness and excellence of the algorithm

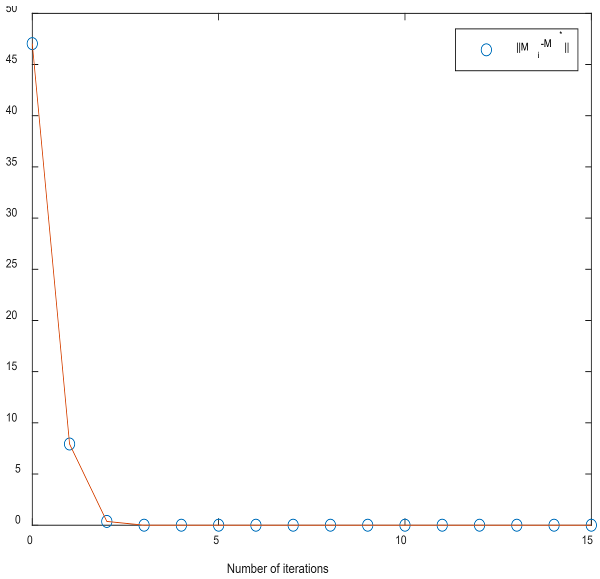

Fig. 5. The convergence process of M_i in the algorithm

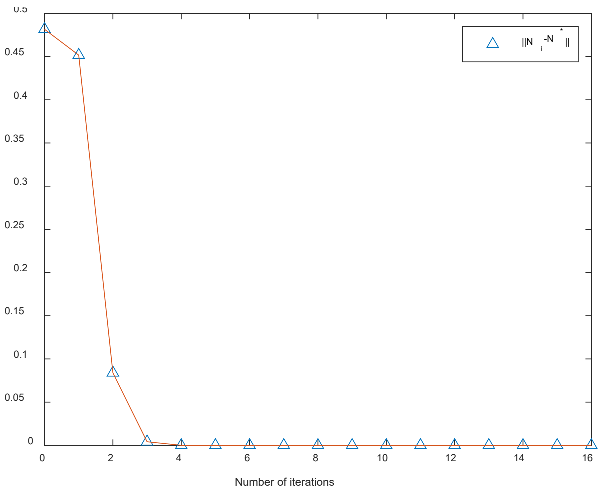

Fig. 6. The convergence process of N_i in the algorithm

5. Conclusions

On the whole, this study addresses the LFC issue in power systems by employing a computational strategy iteration approach with an online adaptive optimal controller designed for unknown power system dynamics. Using real-time system data, even without knowledge of the power system matrix, we achieve iterative solutions for the algebraic Riccati equation, resulting in an optimal control strategy. Experimental simulations validate the practicality and potential value of this method in LFC for power systems. This approach offers a viable solution to the online adaptive optimal control challenge in the presence of unknown system dynamics, presenting a novel concept for adaptive control in power system load frequency regulation. Meanwhile, the control strategy based on reinforcement learning and data-driven has been applied in other practical industrial scenarios such as autonomous driving, robot control, etc. In future research, we will continue to explore the applicability of this method in larger-scale power systems and multi-regional scenarios. In addition, we plan to further study its robustness in combination with other algorithms to enhance its performance and ensure its scalability in more complex environments.

References

[1] Wu Jie, Liu Yongqiang, Chen Wei. Application of modern control technology in power system control (1). Chinese Journal of Electrical Engineering, 1998, (06): 2-7.

[2] Cheng Liang. Research on the application of adaptive dynamic programming in power system. Huazhong University of Science and Technology, 2014.

[3] Wu F. F., Moslehi K. and Bose A., "Power System Control Centers: Past, Present, and Future," in Proceedings of the IEEE, vol. 93, no. 11, pp. 1890-1908, Nov. 2005, doi: 10.1109/JPROC.2005.857499.

[4] Kundur P. Power system stability. Power system stability and control, 2007, 10: 7-1.

[5] Kundur P., Paserba J. and Vitet, S. "Overview on definition and classification of power system stability," CIGRE/IEEE PES International Symposium Quality and Security of Electric Power Delivery Systems, 2003. CIGRE/PES 2003., Montreal, QC, Canada, 2003, pp. 1-4, doi: 10.1109/QSEPDS.2003.159786.

[6] Peddakapu K, Mohamed M R, Srinivasarao P, et al. A state-of-the-art review on modern and future developments of AGC/LFC of conventional and renewable energy-based power systems. Renewable energy focus, 2022.

[7] Rangi S, Jain S, Arya Y. SSA-optimized cascade optimal-PIDN controller for multi-area power system with RFB under deregulated environment. Optim Control Appl Meth. 2023; 44(4): 1972–1994.

[8] Sandeep Rangi, Sheilza Jain, Yogendra Arya, Utilization of energy storage devices with optimal controller for multi-area hydro-hydro power system under deregulated environment, Sustainable Energy Technologies and Assessments, Volume 52, Part C,2022,102191, ISSN 2213-1388.

[9] Jagatheesan, Kaliannan et al. "Evolutionary Algorithm Based LFC of Single Area Thermal Power System with Different Steam Configurations and Non-linearity." Advances in intelligent systems and computing 435 (2016): 185-194.

[10] Khooban M H, Niknam T. A new intelligent online fuzzy tuning approach for multi-area load frequency control: Self Adaptive Modified Bat Algorithm.International Journal of Electrical Power & Energy Systems, 2015, 71(oct.):254-261. DOI:10.1016/j.ijepes.2015.03.017.

[11] Bechert T. E. and Chen Nanming, "Area automatic generation control by multi-pass dynamic programming," in IEEE Transactions on Power Apparatus and Systems, vol. 96, no. 5, pp. 1460-1469, Sept. 1977, doi: 10.1109/T-PAS.1977.32474.

[12] Vrdoljak K., Perić N. and Šepac D., "Optimal distribution of load-frequency control signal to hydro power plants," 2010 IEEE International Symposium on Industrial Electronics, Bari, Italy, 2010, pp. 286-291, doi: 10.1109/ISIE.2010.5637553.

[13] Hain, Y, Kulessky, et al. Identification-Based Power Unit Model for Load-Frequency Control Purposes. IEEE Transactions on Power Systems, 2000.

[14] Feliachi, A. Load frequency control using reduced order models and local observers. Int. J. Energy Syst.

[15] Gao Guige. Development and prospect of optimal control theory. Journal of Shanghai Dianji University, 2005, (03): 33-35 + 39.

[16] Todorov E. Optimal control theory. Bayesian brain: probabilistic approaches to neural coding, 2006: 268-298.

[17] Jordan B W, Polak E. Theory of a class of discrete optimal control systems. International Journal of Electronics, 1964, 17(6): 697-711.

[18] Gholamrezaie V., Dozein M. G., H. Monsef and B. Wu, "An Optimal Frequency Control Method Through a Dynamic Load Frequency Control (LFC) Model Incorporating Wind Farm," in IEEE Systems Journal, vol. 12, no. 1, pp. 392-401, March 2018, doi: 10.1109/JSYST.2016.2563979.

[19] Calovic M., "Linear Regulator Design for a Load and Frequency Control," in IEEE Transactions on Power Apparatus and Systems, vol. PAS-91, no. 6, pp. 2271-2285, Nov. 1972, doi: 10.1109/TPAS.1972.293383.

[20] Jia Y., Meng K., Wu K., Sun C. and Dong Z. Y., "Optimal Load Frequency Control for Networked Power Systems Based on Distributed Economic MPC," in IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 51, no. 4, pp. 2123-2133, April 2021, doi: 10.1109/TSMC.2020.3019444.

[21] Liu F, Song Y H, Ma J, et al. Optimal load-frequency control in restructured power systems. IEE Proceedings-Generation, Transmission and Distribution, 2003, 150(1): 87-95.

[22] Rahmani M, Sadati N. Hierarchical optimal robust load-frequency control for power systems. IET generation, transmission & distribution, 2012, 6(4): 303-312.

[23] Alhelou H H, Hamedani-Golshan M E, Zamani R, et al. Challenges and opportunities of load frequency control in conventional, modern and future smart power systems: A comprehensive review. Energies, 2018, 11(10): 2497.

[24] Bian T, Jiang Z P. Value iteration and adaptive dynamic programming for data-driven adaptive optimal control design. Automatica, 2016, 71: 348-360. DOI: 10.1016/j.automatica.2016.05.003.

[25] Werbos P.J. Advanced forecasting methods for global crisis warning and models of intelligence. General Systems Yearbook, 1977:25~38

[26] Optimal control strategy for a class of discrete-time nonlinear systems with time delay based on adaptive dynamic programming. Journal of Automation, 2010,36 (01): 121-129.

[27] Lin, Zhang, Song et al. Adaptive dynamic programming with ε error bound for nonlinear discrete-time systems. Control and decision, 2011,26 (10): 1586-1590 +1595. DOI: 10.13195 / j.cd.2011.10.149. linxf. 019

[28] Mingming HA. Integrated adaptive dynamic programming intelligent control for discrete-time dynamic systems. Beijing University of Science and Technology, 2023. DOI: 10.26945 / d.cnki.gbjku.2023.000441

[29] Ni Z, He H, Zhong X, et al. Model-Free Dual Heuristic Dynamic Programming. IEEE Transactions on Neural Networks & Learning Systems, 2017, 26(8): 1834-1839. DOI:10.1109/TNNLS.2015.2424971.

[30] Rizvi S. A. A. and Lin Z., "Adaptive Dynamic Programming for Model-Free Global Stabilization of Control Constrained Continuous-Time Systems," in IEEE Transactions on Cybernetics, vol. 52, no. 2, pp. 1048-1060, Feb. 2022, doi: 10.1109/TCYB.2020.2989419.

[31] Jiang Y. and Jiang Z. P., "Global Adaptive Dynamic Programming for Continuous-Time Nonlinear Systems," in IEEE Transactions on Automatic Control, vol. 60, no. 11, pp. 2917-2929, Nov. 2015, doi: 10.1109/TAC.2015.2414811.

[32] Mu C., Ni Z., Sun C. and He H., "Data-Driven Tracking Control with Adaptive Dynamic Programming for a Class of Continuous-Time Nonlinear Systems," in IEEE Transactions on Cybernetics, vol. 47, no. 6, pp. 1460-1470, June 2017, doi: 10.1109/TCYB.2016.2548941.

[33] Wang Y., Zhou R. and Wen C., "Robust controller design for power system load frequency control," [Proceedings 1992] The First IEEE Conference on Control Applications, Dayton, OH, USA, 1992, pp. 642-646 vol.2, doi: 10.1109/CCA.1992.269772.

[34] Kleinman D L. On an iterative technique for Riccati equation computations.IEEE Trans.a.c, 1968, 13(1):114-115.DOI:10.1109/TAC.1968.1098829.

[35] Jiang, Y. Z., et al. Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics. Automatica Oxford, 2012