IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 18, 2021, accepted June 15, 2021, date of publication June 21, 2021, date of current version June 29, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3090815

Adaptive Optimal Control of CVCF Inverters With Uncertain Load: An Adaptive Dynamic Programming Approach

ZHONGYANG WANG[1] AND YUNJUN YU[2,3], (Member, IEEE)
[1]School of Applied Science and Engineering, Fuzhou Institute of Technology, Fuzhou 350506, China
[2]School of Information Engineering, Nanchang University, Nanchang 330031, China
[3]AI Institute, Nanchang University, Nanchang 330031, China

Corresponding author: Yunjun Yu (yuyunjun@ncu.edu.cn)

This work was supported in part by the Fujian Province Young and Middle-Aged Teacher Education Research Project under Grant JAT200901, in part by the Fuzhou Institute of Technology Research Fund Project under Grant FTKY022, in part by the National Natural Science Foundation of China (NSFC) under Grant 61563034, and in part by the International S&T Cooperation Program of China (ISTCP) under Grant 2014DFG72240.

ABSTRACT This paper proposed a data-driven adaptive optimal control approach for CVCF (constant voltage, constant frequency) inverter based on reinforcement learning and adaptive dynamic programming (ADP). Different from existing literature, the load is treated as a dynamic uncertainty and a robust optimal state-feedback controller is proposed. The stability of the inverter-load system has been strictly analyzed. In order to obtain accurate output current differential signal, this paper designs a tracking differentiator. It is ensured that the tracking error asymptotically converges to zero through the proposed output-feedback controllers. A standard proportional integral controller and linear active disturbance rejection control strategy are also designed for the purpose of comparison. The simulation results show that the proposed controller has inherent robustness and does not require retuning with different applications.

INDEX TERMS Adaptive optimal control, CVCF inverter, reinforcement learning, adaptive dynamic programming.

I. INTRODUCTION

With the reduction of fossil energy reserves, the increasing environmental pollution has aroused people's attention to new energy [1]. Therefore, smarter controllers are needed to balance consumption and power generation. In the public grid, batteries exchange energy in the form of direct current (DC) and CVCF inverter are required to convert DC voltage to AC voltage [2]–[4]. CVCF inverter are widely used in the industry, such as distributed power generation reactive power compensator, electric aircraft power system and uninterruptible power supplies [5], [6]. In this case, consider using a CVCF inverter because the amplitude and frequency of the output voltage will not change over time. The main goal of control technology is to make the output voltage signal track a given reference voltage. The most commonly used techniques in CVCF inverter applications are repetitive control [7] and resonance control [8]–[10]. In order to reduce high-frequency switching harmonics, an LC filter is only installed on the output side of the inverter. In addition, the CVCF inverter should be able to adapt to various loads (e.g., load step-change, unbalanced load, and nonlinear load) [13]. In this case, the controller should be carefully designed because the kinetic parameters of the LC circuit have peaks [11]. Inappropriate voltage controller may make the whole system unstable [12].

In the standard V/f control method, the *dq* axis voltage is controlled separately in the synchronous rotating reference frame. By using coordinate transformation, the AC value is converted to DC value. Under synchronous rotation reference coordinates, the dynamics model of CVCF inverter behaves as a linear time invariant (LTI) system. Therefore, linear control technology can be used to analyze and design inverter systems [24], such as proportional integral (PI) control, pole placement control and linear quadratic regulator (LQR) control [15]. However, these controllers require

The associate editor coordinating the review of this manuscript and approving it for publication was Shichao Liu.

precise mathematical models to design. Extensive research has been conducted on the control methods of uncertain problems and their impact on system performance. Based on disturbance estimation and compensation control method, such as disturbance observer based control [16]–[18] and the active disturbance rejection control [19]. Sliding mode control can reduce the sensitivity of uncertain models [20], [23]. However, these control guard parameters are difficult to solve and are not optimal parameters. In addition, the model predictive control method has been applied to CVCF inverter as an optimal control approach [25], [26]. One of the disadvantages of this method is that it may increase the computational burden. In the literature [38], a voltage source inverter control method based on the principle of internal model is proposed. However, this method is only suitable for linear loads and inverters whose output is sinusoidal voltage.

For control systems, the traditional controller design is based on accurate mathematical models, but in many cases, the internal model parameters of the system are uncertain and the external disturbance dynamics are unknown. In order to solve the problem, we invoke reinforcement learning theory [27] and ADP for non-model-based, data-driven adaptive optimal control design. In general, an reinforcement learning problem requires the existence of an agent, that can interact with some unknown environment by taking actions, and receiving a reward from it. In literature [27], the author defines the reinforcement learning ratio as how to map context to action to maximize the digital reward signal. The goal of reinforcement learning is to learn an optimal policy to solve the effect index of the optimal algorithm of maximization or minimization under certain constraints [28], [29]. For uncertain or unknown systems, ADP has made a lot of progress [30], [33], [34], [37]. Currently, there are two main solutions for ADP, which are on-policy learning and off-policy learning. For on-policy learning approach, every time a new control strategy represented by the gain matrix is obtained, the strategy must be implemented to generate a new solution for the closed loop system. These new solutions are then used to assess current costs and find new strategies. Although strategy iteration is very similar to biological learning, the entire adaptation process can be slow and requires continuous collection of online data until certain convergence criteria are met. In engineering applications, sometimes we are more interested in obtaining approximate optimal solutions by making full use of some limited data. This motivates us to develop a off-policy learning strategy in which we apply the initial control strategy to the system within a limited time interval and collect online measurement results. Then, all iterations are performed by repeatedly using the same online data.

For adaptive optimal output regulation, relevant scholars have proposed specific implementation plans [30]–[32]. It is worth noting that in these documents, the external interference is independent, and the system output and external interference have the same internal model. But for the CVCF inverter system, the external interference is the output current, which is not only related to the load dynamics, but also related to the output voltage. If the inverter has a non-linear load, then the interference of the system is in a non-linear state. The traditional ADP isn't able to solve the problem. Based on the traditional ADP, this paper develops a strict feedback nonlinear system ADP control method [35], [36]. This method does not need to understand the external disturbance dynamics of the system, and can approximate the relevant parameters of the system. Simulation shows that the control method proposed in this paper can adapt to different types of loads.

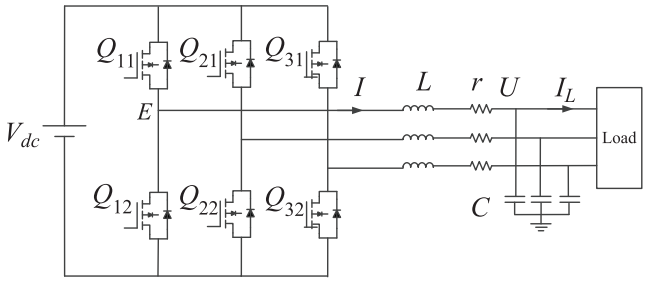


FIGURE 1. Three-phase CVCF inverter topology.

The rest of this paper is organized as follows: In Section II, the mathematical model of three-phase CVCF inverter is introduced. In Section III, we have designed a tracking differentiator to find the differential signal of the output current. In section IV, We have developed an off-policy-based ADP control method for strict feedback nonlinear systems. In section V, we designed a linear active disturbance rejection control (LADRC) strategy for comparison. Section VI gives the simulation results and discussion and section VI is conclusions.

Notations. $\otimes$ indicates the Kronecker product operator. For a symmetric matrix $P \in \mathbb{R}^{m\times m}$, an asymmetric matrix $Y \in \mathbb{R}^{n\times m}$ and a column vector $v \in \mathbb{R}^n$, operator vecs, vec and vecv denote $\text{vecs}(P) = [p_{11}, 2p_{12}, \cdots, 2p_{m-1,m}, p_{mm}]^T \in \mathbb{R}^{\frac{1}{2}m(m+1)}$, $\text{vec}(Y) = [y_1^T, y_2^T, \cdots, y_m^T]^T \in \mathbb{R}^{mn}$ with $y_i \in R^n$, $\text{vecv}(v) = [v_1^2, v_1v_2, \cdots, v_1v_n, v_2^2, v_2v_3, \cdots, v_{n-1}v_n, v_n^2]^T \in \mathbb{R}^{\frac{1}{2}n(n+1)}$.

II. MODELING OF SYSTEM

Consider the three-phase CVCF inverter with unknown parameters is shown in Fig. 1. Assume that the input DC voltage is constant. V_{dc} is the DC side voltage, the output LC filter is composed of L and C, r is a small series resistance. The angular frequency of the system is ω.

The inverter side output voltage can be written as

$$E(t) = \begin{cases} V_{dc} & Q_{11} = Q_{22} = 1, Q_{12} = Q_{21} = 0 \\ -V_{dc} & Q_{11} = Q_{22} = 0, Q_{12} = Q_{21} = 1, \end{cases} \quad (1)$$

where Q_{ij} is the state of the switch tube ($Q_{ij} = 1$ means on, and $Q_{ij} = 0$ means off).

CVCF inverter mathematical model can be described by the following set of differential equations:

$$\begin{cases} L\dot{I}(t) = E(t) - U(t) - rI(t) \\ C\dot{U}(t) = I(t) - I_L(t), \end{cases} \tag{2}$$

where $I(t)$ is the inductor current, $U(t)$ is the output voltage, $I_L(t)$ is the load current, and $E(t)$ is the control input voltage.

We reach the following state-space form

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) + Dw(t) \\ y(t) = x_2(t), \end{cases} \tag{3}$$

where $x(t) = [I(t), U(t)]^T$, $u(t) = E(t)$, $y(t) = U(t)$, $w(t) = I_L(t)$,

$$A = \begin{bmatrix} -\frac{r}{L} & -\frac{1}{L} \\ \frac{1}{C} & 0 \end{bmatrix}, \quad B = \begin{bmatrix} \frac{1}{L} \\ 0 \end{bmatrix}, \text{ and } \quad D = \begin{bmatrix} 0 \\ -\frac{1}{C} \end{bmatrix}.$$

Define the tracking error of the system:

$$\dot{e}(t) = Ae(t) + Bv(t), \tag{4}$$

where $e(t) = x(t) - x^*(t)$ and $v(t) = u(t) - u^*(t)$, $*$ represents the reference value. The reference value can be obtained by the following formula:

$$\begin{cases} I^*(t) = C\dot{U}^*(t) + I_L(t) \\ E^*(t) = L\dot{I}^*(t) + rI^*(t) + U^*(t). \end{cases} \tag{5}$$

III. TRACKING DIFFERENTIATOR

In this paper, the differential signal of the load current needs to be used and will use the tracking differentiator to obtain the differential signal.

The traditional differential signal is obtained by the following differential link:

$$\dot{v} = G(s)v = \frac{s}{Ts+1}v = \frac{1}{T}(v - \frac{v}{Ts+1}), \tag{6}$$

where T is the sample time.

The biggest problem of the differentiator is that it is very sensitive to noise signals. There are harmonics in the output current of the inverter. The differential signal obtained by the traditional differentiator will amplify the noise signal.

In order to eliminate or weaken the effect of noise amplification, the differential approximation formula is written as follows:

$$\begin{aligned} \dot{v} &\approx \frac{v(t-\tau_1) - v(t-\tau_2)}{\tau_2 - \tau_1} \\ &= \frac{1}{\tau_2 - \tau_1}(\frac{1}{\tau_1 s + 1} - \frac{1}{\tau_2 s + 1})v \\ &= \frac{s}{\tau_1\tau_2 s^2 + (\tau_1 + \tau_2)s + 1}v, \end{aligned} \tag{7}$$

where $0 < \tau_1 < \tau_2$.

According to the above transfer function, we can write its differential form

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -\frac{1}{\tau_1\tau_2}(x_1 - v) - \frac{\tau_1 + \tau_2}{\tau_1\tau_2}x_2 \\ \dot{v} = x_2 \end{cases} \tag{8}$$

The discrete form of (8) with a sample period T can be written as

$$\begin{cases} \dot{x}_{1(k+1)} = x_{1k} + Tx_{2k} \\ \dot{x}_{2(k+1)} = -\frac{T}{\tau_1\tau_2}x_{1k} + (1 + \frac{\tau_1 + \tau_2}{\tau_1\tau_2})x_{2k} + \frac{T}{\tau_1\tau_2}v_k \\ \dot{v}_k = x_{2k}, \end{cases} \tag{9}$$

where $T = 10^{-6}$, $\tau_1 = 10^{-6}$, and $\tau_2 = 2 \times 10^{-6}$.

IV. ADAPTIVE OPTIMAL CONTROL OF CVCF INVERTER

A. LINEAR QUADRATIC REGULATOR

The control objective finds the optimal feedback controller $v = -Ke$ to solve the following constraint minimization Problem 1:

$$\min_v \int_0^\infty (e^T Qe + v^T Rv)d\tau, \tag{10}$$

where $Q = Q^T \geq 0$, $R = R^T \geq 0$, with $(A, \sqrt{Q})$ observable.

Problem 1 is a linear quadratic regulator problem. Through linear optimal control theory, the feedback gain is

$$K = R^{-1}B^T P, \tag{11}$$

where

$$P = \int_0^\infty M^T(e^T Qe + v^T Rv)Md\tau \tag{12}$$

with $M = \exp((A - BK)\tau)$, and $P = P^T > 0$ can be found by solving the following algebraic Riccati equation (also known as ARE):

$$A^T P + PA + Q - PBR^{-1}B^T P = 0. \tag{13}$$

Algorithm 1 Iterative Algorithm

1.Let K_0 be any stabilizing feedback gain matrix($A - BK_0$ is a *Hurwitz* matrix), let $j = 1$ and $\tau > 0$.

2.Solve *Lyapunov* equation:

$$A_j^T P + PA_j + Q + K_j R K_j = 0. \tag{14}$$

3. Update the feedback gain matrix by

$$K_{j+1} = R^{-1}B^T P_j. \tag{15}$$

4.If $|P_{j+1} - P_j| \leq \tau$, repeat step (5), else $j = j + 1$, repeat step (2).

5.Use $u = -K_{j^*}(x - x^*) + u^*$ as controller.

The Algorithm 1 is able to approximate the solution to (13) with assured convergence, and the following properties hold.

1) $A - BK_j$ is a *Hurwitz* matrix;

2) $P_{j^*} \leq P_{j+1} \leq P_j$;

3) $\lim_{j\to\infty} K_j = K_{j^*}$ and $\lim_{j\to\infty} P_j = P_{j^*}$.

B. ADAPTIVE OPTIMAL CONTROL DESIGN

The system parameters are uncertain and the reference value of the state quantity cannot be directly calculated.We assume:

$$\bar{x}_1 = x_1 - m \quad \text{and} \quad \bar{x} = [\bar{x}_1, x_2]^T \tag{16}$$

where m is a constant.

Consider the following control policy:

$$u = u_0 \tag{17}$$

Under the control policy, the original system can be rewritten as:

$$\dot{\bar{x}} = A_j\bar{x} + B(K_j\bar{x} + u_0) + Dw + Nm \tag{18}$$

where $N = [-r/L, 0]^T$.

Then, taking the time derivative of $\bar{x}^T P_j \bar{x}$ and results in

$$\begin{aligned}\frac{d}{dt}\bar{x}^T P\bar{x} = {} & \bar{x}^T(A_j^T P_j + P_j A_j)\bar{x} \\ & +2(K_j\bar{x} + u_0)^T B^T P_j\bar{x} + 2w^T D^T P_j\bar{x} \\ & +2m^T N^T P_j\bar{x}\end{aligned} \tag{19}$$

During the time period $[t, t+\delta t]$, rewrite (19) as

$$\begin{aligned}\bar{x}^T P_j\bar{x}|_t^{t+\delta t} = {} & -\int_t^{t+\delta t} \bar{x}^T Q_j\bar{x} d\tau \\ & +2\int_t^{t+\delta t} (K_j\bar{x} + u_0)^T RK_{j+1}\bar{x} d\tau \\ & +2\int_t^{t+\delta t} w^T D^T P_j\bar{x} d\tau \\ & +2\int_t^{t+\delta t} m^T N^T P_j\bar{x} d\tau\end{aligned} \tag{20}$$

where $Q_j = Q + K_j^T RK_j = -(A_j^T P_j + P_j A_j)$ and $RK_{j+1} = B^T P_j$.

Given time series $t_0 \le t_1 \le t_2 \cdots \le t_s$, define the following matrices:

$$\begin{aligned}I_{\bar{x},\bar{x}} &= \left[\int_{t_0}^{t_1}\bar{x}\otimes\bar{x}d\tau, \cdots, \int_{t_{s-1}}^{t_s}\bar{x}\otimes\bar{x}d\tau\right]^T, \\ I_{\bar{x},u_0} &= \left[\int_{t_0}^{t_1}\bar{x}\otimes u_0 d\tau, \cdots, \int_{t_{s-1}}^{t_s}\bar{x}\otimes u_0 d\tau\right]^T, \\ I_{\bar{x},w} &= \left[\int_{t_0}^{t_1}\bar{x}\otimes w d\tau, \cdots, \int_{t_{s-1}}^{t_s}\bar{x}\otimes w d\tau\right]^T, \\ I_{\bar{x},m} &= \left[\int_{t_0}^{t_1}\bar{x}\otimes m d\tau, \cdots, \int_{t_{s-1}}^{t_s}\bar{x}\otimes m d\tau\right]^T, \\ \Gamma_{\bar{x},\bar{x}} &= \left[\text{vecv}(\bar{x}(\tau))|_{t_0}^{t_1}, \cdots, \text{vecv}(\bar{x}(\tau))|_{t_{s-1}}^{t_s}\right]^T.\end{aligned}$$

Then, (20) represents the following matrix linear equation form, i.e.,

$$\Phi_j \begin{bmatrix}\text{vecs}(P_j) \\ \text{vec}(K_{j+1}) \\ \text{vec}(D^T P_j) \\ \text{vec}(N^T P_j)\end{bmatrix} = \Psi_j \tag{21}$$

where $\Phi_j = [\Gamma_{\bar{x},\bar{x}}, -2I_{\bar{x},\bar{x}}(I_n \otimes K_j^T R) - 2I_{\bar{x},u_0}(I_n \otimes R), -2I_{\bar{x},w}, -2I_{\bar{x},m}]$ and $\Psi_j = -I_{\bar{x},\bar{x}}\text{vec}(Q_j)$

A sufficient condition that ensures the uniqueness of the last equation is that the column Φ_j of the matrix is full rank as detailed in (22).

$$\text{rank}(\Phi_j) = 9. \tag{22}$$

Algorithm 2 Adaptive Optimal Controller Design

1. Let K_0 be any stabilizing feedback gain matrix($A - BK_0$ is a *Hurwitz* matrix), $j = 0$ and $\tau > 0$
2. Online data collection: Use $u = u_0$ as the control policy, and compute Ψ_j until the rank condition (21) is satisfied.
3. Policy evaluation and improvement:

$$\begin{bmatrix}\text{vecs}(P_j) \\ \text{vec}(K_{j+1}) \\ \text{vec}(D^T P_j) \\ \text{vec}(N^T P_j)\end{bmatrix} = (\Phi_j^T\Phi_j)^{-1}\Phi_j\Psi_j \tag{23}$$

4. Let $j \leftarrow j+1$, if $|P_j - P_{j-1}| > \tau$, repeat Step (3), otherwise let $j^* \leftarrow j$, go to step 5
5. Apply $u = -K_{j^*}e + u^*$ as the control

The convergence of Algorithm 2 is guaranteed under the rank condition (22) (see [37]).

Then, the system parameters are calculated by the ADP algorithm as

$$\begin{cases}C = -\frac{1}{D_{(2)}} = -\dfrac{1}{(P_j^{-1}P_j D)_{(2)}} \\ L = \frac{1}{B_{(1)}} = \dfrac{1}{(P_j^{-1}K_{j+1}R)_{(1)}} \\ r = -LN_{(1)} = -\dfrac{(P_j^{-1}P_j N_{(1)})_{(1)}}{(P_j^{-1}K_{j+1}R)_{(1)}},\end{cases} \tag{24}$$

where the subscript (1) indicates the first element of the column vector, and the subscript (2) indicates the second element of the column vector.

Block diagram of the ADP-based CVCF inverter control system is shown in Fig. 2.

C. INITIAL CONTROL GAIN AND STABILITY ANALYSIS

In this paper, a stable initial control gain needs to be selected. We choose $K_0 = [0, 0]$ as the initial control gain, then verify whether $A - BK_0$ is a *Hurwitz* matrix.

$$\begin{aligned}|\lambda I - (A - BK_0)| &= \begin{vmatrix}\lambda + \dfrac{r}{L} & \dfrac{1}{L} \\ -\dfrac{1}{C} & \lambda\end{vmatrix} \\ &= \lambda^2 + \frac{r}{L}\lambda + \frac{1}{LC} \\ &= 0\end{aligned} \tag{25}$$

Then

$$\lambda_{1,2} = -\frac{r}{2L} \pm \sqrt{\frac{r^2}{4L^2} - \frac{1}{C}} \in \mathbb{C}^- \tag{26}$$

For CVCF inverter, $K_0 = [0, 0]$ is a stable feedback gain.

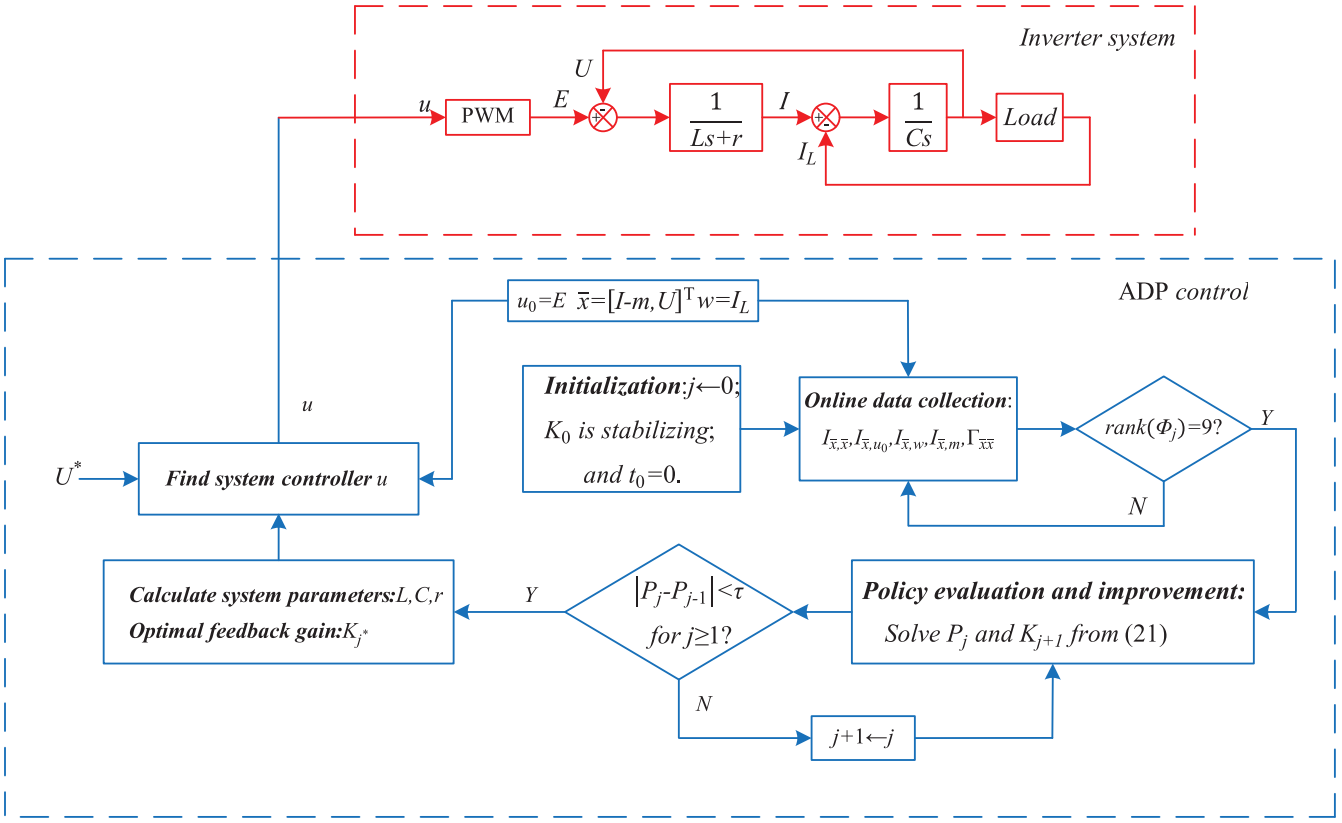


FIGURE 2. Block diagram of the ADP-based CVCF inverter control system.

Consider the Lyapunov equation (13) with $j = 0$. Since $A - BK_0$ is Hurwitz, by (12) we know P_0 is finite and positive definite. In addition, by (12) and (14) we have

$$P_0 - P_1 = \int_0^\infty \exp(A_1^T \tau)(K_0 - K_1)^T \times R(K_0 - K_1)\exp(A_1 \tau) d\tau \geq 0 \quad (27)$$

Similarly, by (12) and (13) we obtain

$$P_1 - P^* = \int_0^\infty \exp(A_1^T \tau)(K_1 - K^*)^T \times R(K_1 - K^*)\exp(A_1 \tau) d\tau \geq 0 \quad (28)$$

Then, $P^* \leq P_1 \leq P_0$ can be obtained. Since P^* is positive definite and P_0 is finite, P_1 must be finite and positive definite. This implies that $A - BK_1$ is Hurwitz. Repeating the above analysis for $j = 1, 2, \cdots$, we have $A - BK_j$ is Hurwitz.

For error system

$$\begin{aligned} \dot{e} &= Ax + B(-K_j e + u^*) + Dw - \dot{x}^* \\ &= (A - BK_j)e \end{aligned} \quad (29)$$

We have $\lim_{t\to\infty} e = 0$ and the original system is stable.

V. COMPARE CONTROLLER

In order to verify the superiority of the control algorithm proposed in this paper, this paper designs standard proportional integral (PI) control and linear active disturbance rejection control (LADRC).

Consider the following second-order system

$$\ddot{y} = -a_1 \dot{y} - a_2 y + bu + w \quad (30)$$

where y is system output, u is system control, w is interference. Assume $x_1 = y$, $x_2 = \dot{y}$, and $x_3 = -a_1\dot{y} - a_2 y + (b - b_0)u + w$. Then, a new state equation can be written as

$$\begin{cases} \dot{x} = Ax + Bu + E\dot{x}_3 \\ y = Cx \end{cases} \quad (31)$$

where $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 \\ b_0 \\ 0 \end{bmatrix}$, $E = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, and $C = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$.

Construct the corresponding continuous linear extended state observer [21], [22]:

$$\begin{cases} \dot{z} = Az + Bu + L(y - Cz) \\ \bar{y} = Cz \end{cases} \quad (32)$$

where $L = [3\omega_o, 3\omega_o^2, \omega_o^3]^T$. The PD controllers that LADRC can use are as follows

$$u_0 = k_p(y^* - z_1) - k_d z_2 \quad (33)$$

where $k_p = \omega_c^2$ and $k_d = 2\omega_c$. In this paper, we choose $\omega_c = 5000$ and $\omega_o = 10000$ [39].

Then, the controller of the original system is

$$u = u_0 - \frac{z_3}{b_0}. \quad (34)$$

TABLE 1. System parameters.

Parameter	Description	Value
V_{dc}	DC Voltage	$600V$
L	Inverter side inductance	$2mH$
r	Inverter side resistance	0.05Ω
C	Filter capacitor	$35\mu F$
ω	Power frequency	$314.159rad/s$

TABLE 2. System parameters comparison.

Reference value(x^*)	Calculated value(x)	Error rate
$C^* = 35\mu F$	$C_{j^*} = 34.887\mu F$	0.32%
$L^* = 2mH$	$L_{j^*} = 1.988mH$	0.6%
$r^* = 0.05\Omega$	$r_{j^*} = 0.049\Omega$	2%
$K^* = [66.5955, 30.6386]$	$K_{j^*} = [66.6, 30.64]$	0.023%

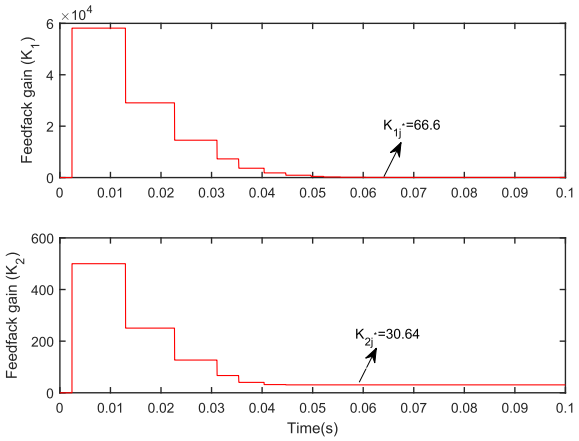


FIGURE 3. Profile of the feedback gains during the simulation.

VI. SIMULATION RESULTS

In order to verify the effectiveness of the algorithm, we built a Simulink simulation model. System parameters are shown in

The reference voltage of phase A of the system is

$$U_a^*(t) = 311 \sin \omega t$$

and

$$u_0 = \text{PWM}(311 \sin(\omega t)) \tag{35}$$

where PWM(311 sin(ωt)) represents the inverter side voltage signal obtained after the three-phase voltage signal is modulated by PWM.

The system parameters calculated by ADP are shown in Tab. 2, and online data collection under the initial controller are shown in Tab. 3.

It can be seen from Tab. 2 that the ADP algorithm can approximate the various parameters of the system.

The improvement of the feedback gains is shown in Fig. 3. ADP approach is able to find the approximate optimal control gain.

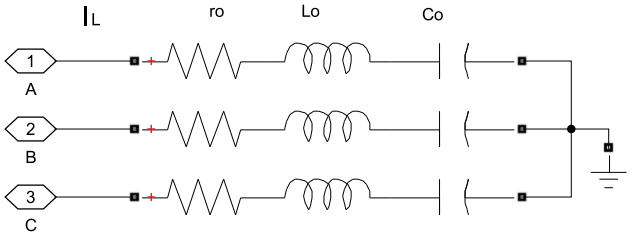


FIGURE 4. Load connections: *RLC* load.

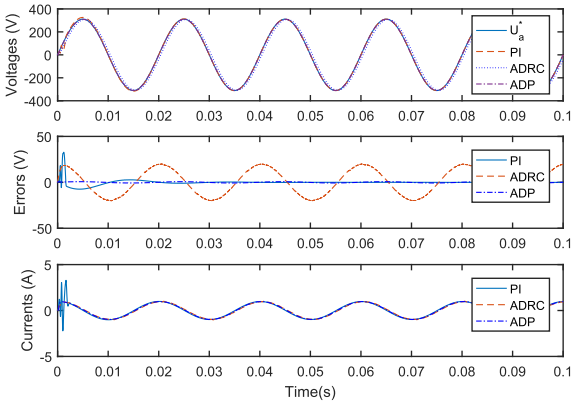


FIGURE 5. Profile of the output voltages, errors, and currents during the simulation under $L = 2mH$, $C = 35\mu F$.

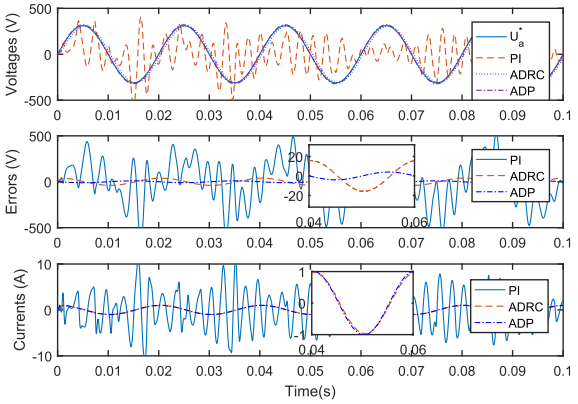


FIGURE 6. Profile of the output voltages, errors, and currents during the simulation under $L = 3mH$, $C = 40\mu F$.

A. LINEAR LOAD

The linear load adopts *RLC* load, the topological diagram is shown in Fig. 4. Values of the components are $r_o = 10\Omega$, $L_o = 1mH$, and $C_o = 10\mu F$.

Profile of the output voltages, errors, and currents during the simulation under $L = 2mH$, $C = 35\mu F$ are shown in Fig. 5. Profile of the output voltages, errors, and currents during the simulation under $L = 3mH$, $C = 40\mu F$ are shown in Fig. 6. Profile of the output voltages, errors, and currents during the simulation under $L = 2.5mH$, $C = 30\mu F$ are shown in Fig. 7. THD of the system output under linear load are shown in Tab. 4. It can be seen that under different system parameters, ADP control has better output, low error and low harmonics. PI control is sensitive to parameter changes, and unsatisfactory system parameters will cause

TABLE 3. Online data collection under the initial controller.

Time(s)	$u_0(t)$(V)	$x_1(t)$(A)	$x_2(t)$(V)	$v(t)$(A)	Time(s)	$u_0(t)$(V)	$x_1(t)$(A)	$x_2(t)$(V)	$v(t)$(A)
0	0	0	0	0	0.031	199.3162	17.10048	196.2411	30.96974
0.001	-199.287	-46.5176	-166.486	-35.3966	0.032	-199.303	-30.7345	89.71838	0.002692
0.002	199.3151	18.22021	-175.086	-0.00525	0.033	-199.289	-43.8384	-176.033	-28.173
0.003	199.2863	47.00095	185.9353	42.95634	0.034	-199.295	-38.5267	-231.391	-27.3922
0.004	199.289	44.36777	201.1887	38.60073	0.035	-199.297	-35.8364	-238.185	-38.8581
0.005	199.2893	44.05564	214.2528	43.07769	0.036	-398.607	-60.0221	-416.261	-97.2279
0.006	398.602	64.66882	426.7507	90.98227	0.037	-398.612	-54.8213	-299.251	-67.0087
0.007	398.6121	54.57204	329.8431	63.62871	0.038	-398.605	-62.0934	-322.566	-43.7272
0.008	398.6059	60.79498	340.1474	47.40194	0.039	-199.302	-31.6058	-252.138	-15.0395
0.009	199.3024	30.97413	250.0311	16.96156	0.04	-199.313	-20.824	-269.173	-28.0293
0.01	199.3133	19.98254	261.1803	27.22563	0.041	-199.316	-17.1	-196.264	-30.9786
0.011	199.3166	16.77269	193.0438	28.90974	0.042	199.3026	30.73519	-89.7156	-0.00269
0.012	-199.302	-30.972	90.08432	0.002703	0.043	199.2895	43.83801	176.0114	28.16956
0.013	-199.289	-43.9846	-179.963	-28.9468	0.044	199.2948	38.5273	231.3931	27.3879
0.014	-199.295	-38.4947	-230.73	-28.2205	0.045	199.2975	35.83591	238.1991	38.85668
0.015	-199.297	-36.028	-235.518	-39.1661	0.046	398.6066	60.02042	416.2683	97.23112
0.016	-398.606	-60.3674	-415.184	-96.6537	0.047	398.6118	54.82063	299.2388	67.01097
0.017	-398.612	-54.9301	-301.428	-66.5817	0.048	398.6046	62.09379	322.5542	43.72617
0.018	-398.605	-62.0231	-324.703	-43.9088	0.049	199.3017	31.60614	252.1382	15.0386
0.019	-199.302	-31.5516	-252.037	-15.2154	0.05	199.3125	20.82415	269.1765	28.02936
0.02	-199.313	-20.7832	-268.495	-28.0261	0.051	199.3162	17.10002	196.2661	30.97925
0.021	-199.316	-17.1024	-195.916	-30.8497	0.052	-199.303	-30.7352	89.71545	0.002691
0.022	199.3026	30.72703	-89.7419	-0.00269	0.053	-199.289	-43.838	-176.01	-28.1693
0.023	199.2895	43.84219	176.3204	28.21578	0.054	-199.295	-38.5273	-231.393	-27.3876
0.024	199.2948	38.51803	231.3702	27.44803	0.055	-199.297	-35.8359	-238.2	-38.8566
0.025	199.2975	35.84098	237.9939	38.87642	0.056	-398.606	-60.209	-385.582	-96.5456
0.026	398.6066	60.04447	416.1649	97.18411	0.057	-398.611	-56.1191	-289.918	-61.22
0.027	398.6118	54.83083	299.4052	66.97858	0.058	-398.604	-63.1518	-339.494	-41.0892
0.028	398.6046	62.08866	322.7237	43.74005	0.059	-199.301	-31.892	-258.59	-17.2158
0.029	199.3017	31.60176	252.1277	15.05148	0.06	-199.312	-21.0284	-265.554	-29.8663
0.03	199.3125	20.8215	269.125	28.02919					

TABLE 4. THD of the system output under linear load.

		Voltage			Current		
L	C	PI	ADRC	ADP	PI	ADRC	ADP
$2mH$	$35\mu F$	0.04%	0.16%	0.02%	0.10%	1.45%	0.06%
$3mH$	$40\mu F$	561%	0.10%	0.02%	374.64%	0.89%	0.04%
$2.5mH$	$30\mu F$	0.05%	0.15%	0.01%	0.08%	1.34%	0.04%

system instability. ADRC can ensure the stability of the system, but the error is a sinusoidal signal and there is a phase deviation.

B. NON-LINEAR LOAD-1

Non-linear load uses uncontrollable diode rectifier current load, the topological diagram is as shown in Fig. 8. Values of the components are $r_o = 10\Omega$, $L_o = 1.35mH$, and $C_o = 47\mu F$.

TABLE 5. THD of the system output under non-linear load-1.

		Voltage			Current		
L	C	PI	ADRC	ADP	PI	ADRC	ADP
$2mH$	$35\mu F$	5.92%	3.67%	2.47%	29.02%	49.56%	26.52%
$3mH$	$40\mu F$	11.15%	4.32%	4.09%	27.99%	46.70%	26.60%
$2.5mH$	$30\mu F$	8.80%	4.06%	3.26%	28.30%	49.38%	26.38%

Profile of the output voltages, errors, and currents during the simulation under $L = 2mH$, $C = 35\mu F$ are shown in Fig. 9. Profile of the output voltages, errors, and currents during the simulation under $L = 3mH$, $C = 40\mu F$ are shown in Fig. 10. Profile of the output voltages, errors, and currents during the simulation under $L = 2.5mH$, $C = 30\mu F$ are shown in Fig. 11. THD of the system output under non-linear load-1 are shown in Tab. 5. It can be seen that under different system parameters, ADP control has better output, low error

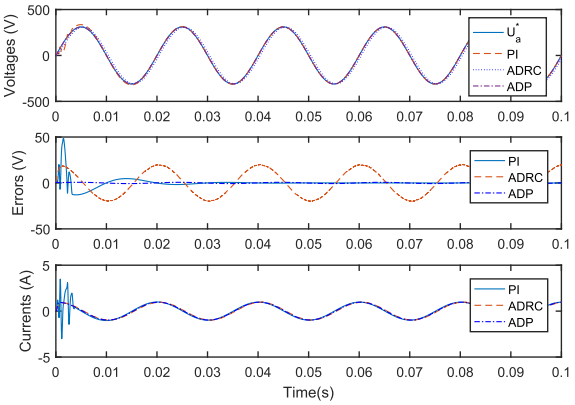


FIGURE 7. Profile of the output voltages, errors, and currents during the simulation under $L = 2.5mH$, $C = 30\mu F$.

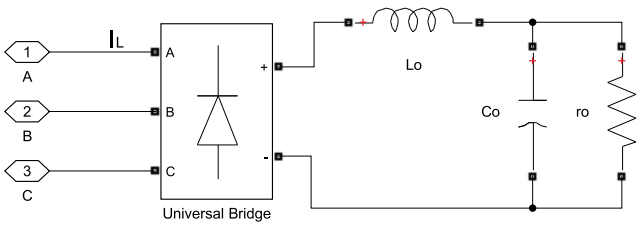


FIGURE 8. Load connections: uncontrollable diode rectifier current load.

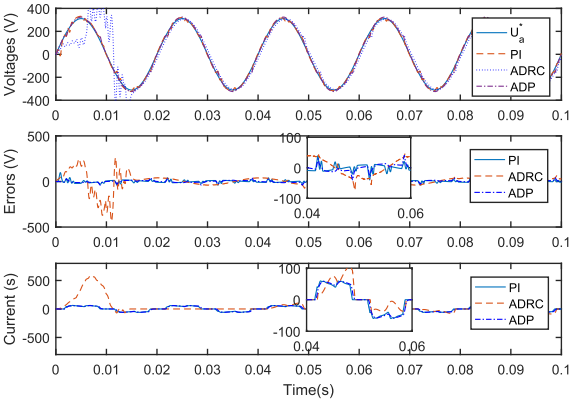


FIGURE 9. Profile of the output voltages, errors, and currents during the simulation under $L = 2mH$, $C = 35\mu F$.

and low harmonics. PI control has a certain steady-state error and large harmonics. ADRC dynamic process oscillates greatly, and there is phase deviation.

C. NON-LINEAR LOAD-2

The topological diagram is as shown in Fig. 8. Values of the components are $r_o = 10\Omega$, $L_o = 5mH$, and $C_o = 4700\mu F$.

Profile of the output voltages, errors, and currents during the simulation under $L = 2mH$, $C = 35\mu F$ are shown in Fig. 12. Profile of the output voltages, errors, and currents during the simulation under $L = 3mH$, $C = 40\mu F$ are shown in Fig. 13. Profile of the output voltages, errors, and currents during the simulation under $L = 2.5mH$, $C = 30\mu F$ are shown in Fig. 14. THD of the system output under non-linear load-2 are shown in Tab. 6. It can be seen that under different system parameters, ADP control has better output,

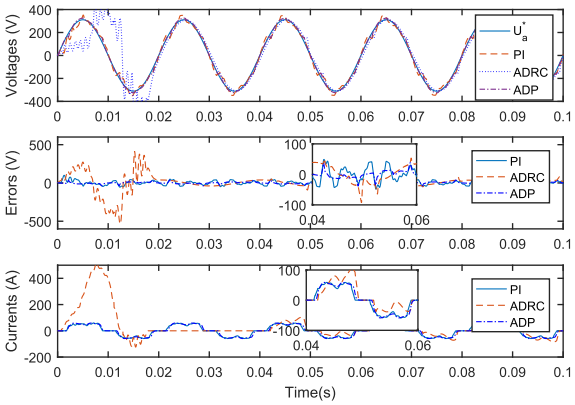


FIGURE 10. Profile of the output voltages, errors, and currents during the simulation under $L = 3mH$, $C = 40\mu F$.

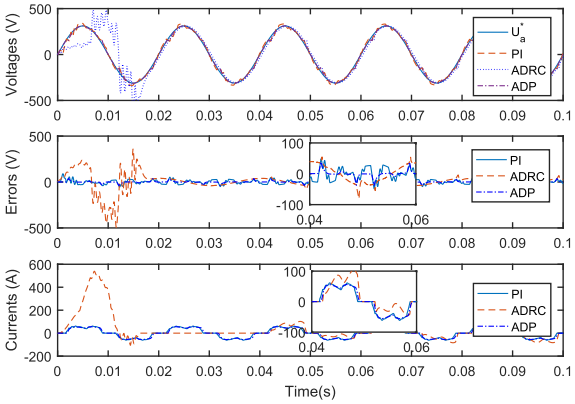


FIGURE 11. Profile of the output voltages, errors, and currents during the simulation under $L = 2.5mH$, $C = 30\mu F$.

TABLE 6. THD of the system output under non-linear load-2.

		Voltage			Current		
L	C	PI	ADRC	ADP	PI	ADRC	ADP
$2mH$	$35\mu F$	343.83%	5.13%	4.29%	731.41%	46.24%	35.01%
$3mH$	$40\mu F$	250.37%	6.83%	4.40%	327.78%	58.17%	28.36%
$2.5mH$	$30\mu F$	333.49%	5.91%	4.33%	621.94%	46.51%	34.48%

low error and low harmonics. PI control is sensitive to load parameters. When the load nonlinearity is high, PI control is unstable. ADRC dynamic process oscillates greatly, there is phase deviation, and load current harmonics are large.

D. LOAD IMBALANCE

The load is a three-phase unbalanced non-linear load, the topological diagram is as shown in Fig. 15.

Values of the components are $r_1 = 20\Omega$, $r_2 = 50\Omega$, $r_3 = 100\Omega$, $L_1 = 0.1mH$, $L_2 = 1mH$, $L_3 = 0mH$, and $C_1 = 200\mu F$, $C_3 = 0\mu F$, $C_1 = 100\mu F$.

Profile of the output voltages, errors, and currents during the simulation under $L = 2mH$, $C = 35\mu F$ are shown in Fig. 16. Profile of the output voltages, errors, and currents during the simulation under $L = 3mH$, $C = 40\mu F$ are shown in Fig. 17. Profile of the output voltages, errors, and currents during the simulation under $L = 2.5mH$, $C = 30\mu F$ are

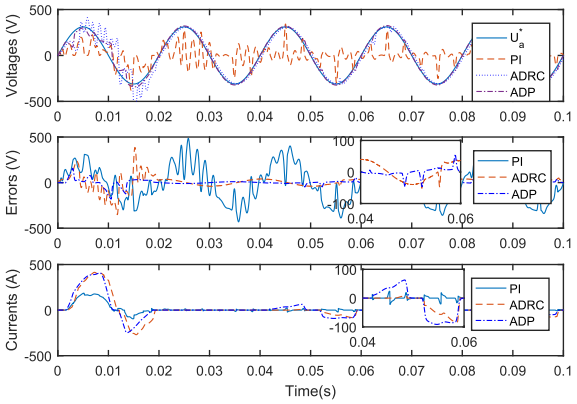


FIGURE 12. Profile of the output voltages, errors, and currents during the simulation under $L = 2mH$, $C = 35\mu F$.

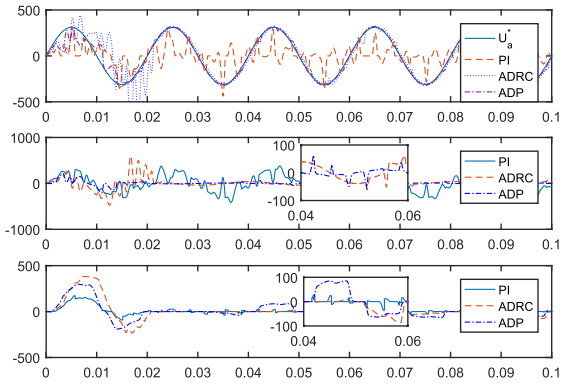


FIGURE 13. Profile of the output voltages, errors, and currents during the simulation under $L = 3mH$, $C = 40\mu F$.

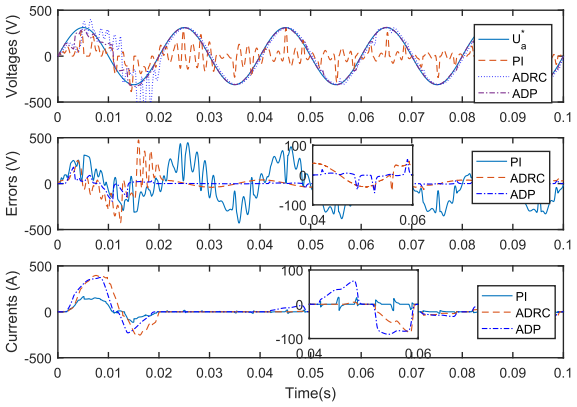


FIGURE 14. Profile of the output voltages, errors, and currents during the simulation under $L = 2.5mH$, $C = 30\mu F$.

TABLE 7. THD of the system output under load imbalance.

		Voltage			Current		
L	C	PI	ADRC	ADP	PI	ADRC	ADP
$2mH$	$35\mu F$	3.73%	2.66%	2.26%	48.15%	46.38%	37.32%
$3mH$	$40\mu F$	8.23%	2.94%	2.03%	48.82%	50.92%	37.20%
$2.5mH$	$30\mu F$	8.87%	3.32%	2.66%	59.47%	49.94%	37.95%

shown in Fig. 18. THD of the system output under load imbalance are shown in Tab. 7. It can be seen that under different system parameters, ADP control has better output, low error and low harmonics. PI control has a certain steady-state error and large harmonics. ADRC current has large harmonics and phase deviation.

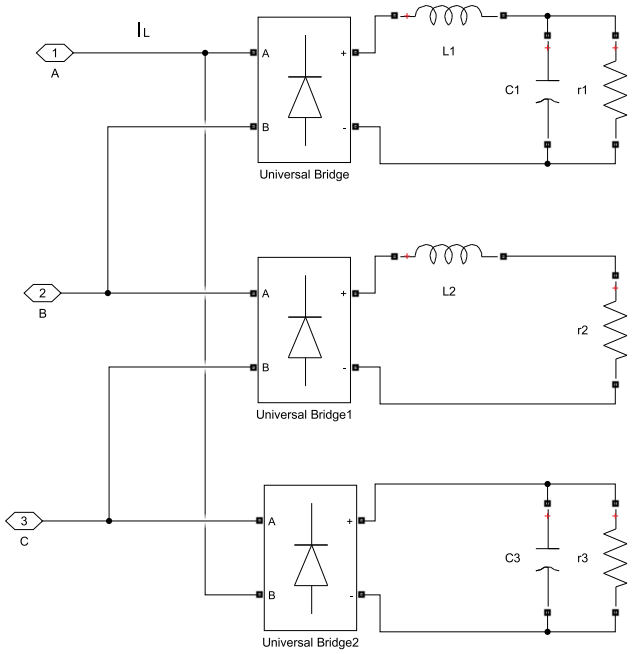


FIGURE 15. Three-phase unbalanced non-linear load.

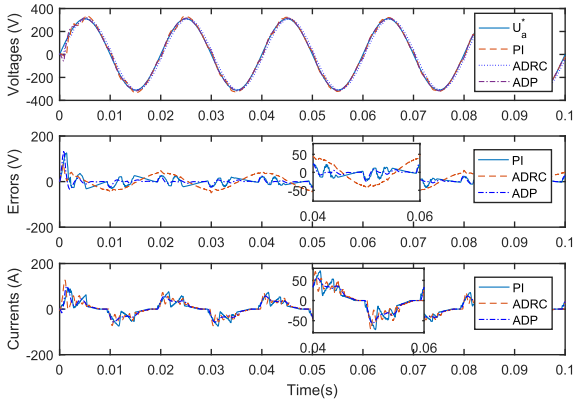


FIGURE 16. Profile of the output voltages, errors, and currents during the simulation under $L = 2mH$, $C = 35\mu F$.

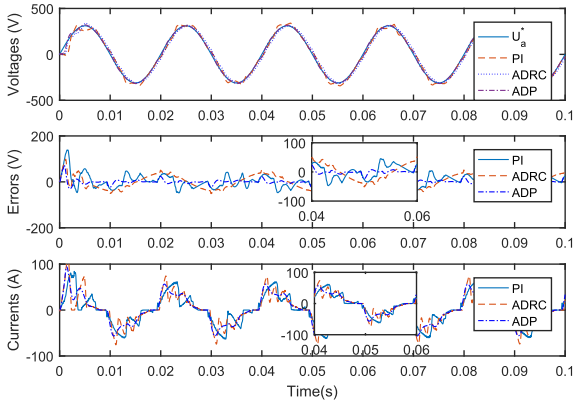


FIGURE 17. Profile of the output voltages, errors, and currents during the simulation under $L = 3mH$, $C = 40\mu F$.

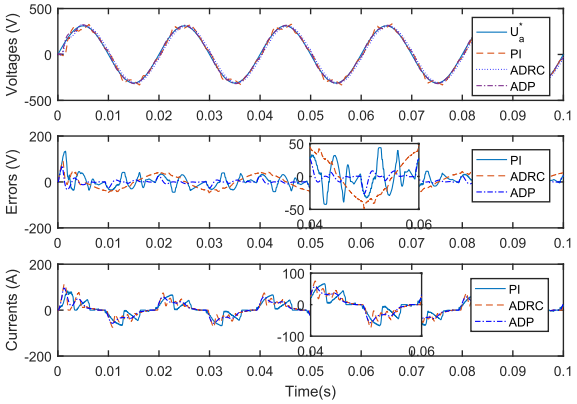


FIGURE 18. Profile of the output voltages, errors, and currents during the simulation under $L = 2.5mH$, $C = 30\mu F$.

VII. CONCLUSION

In this paper, an approximate optimal control approach is proposed to obtain the optimal controller of the CVCF inverter with unknown parameters. Through reinforcement learning and adaptive dynamic programming, a non-model-based scheme is proposed, and an adaptive optimal controller is designed. This paper uses an off-policy learning strategy. Online data is collected through the initial control strategy. Then, the approximate optimal controller is obtained by repeatedly using the same online data for all iterations. The simulation shows that the proposed control scheme can approximate the parameters of the system, and does not require retuning with different applications. The controller design method is generalized for grid-tie applications, since the grid can be regarded as a stable load system.

The controller designed in this paper is compared with the traditional PI controller and linear active disturbance rejection control Strategy. The simulation shows that the ADP controller has good performance for nonlinear loads with different parameters. For different load types, ADP control still has the ability to ensure stability, and reduces the harmonic rate of output voltages and currents.

REFERENCES

[1] S. Eren, M. Pahlevaninezhad, A. Bakhshai, and P. K. Jain, "Composite nonlinear feedback control and stability analysis of a grid-connected voltage source inverter with LCL filter," *IEEE Trans. Ind. Electron.*, vol. 60, no. 11, pp. 5059–5074, Oct. 2012.

[2] X. He, H. Geng, and S. Ma, "Transient stability analysis of grid-tied converters considering PLL's nonlinearity," *CPSS Trans. Power Electron. Appl.*, vol. 4, no. 1, pp. 40–49, Mar. 2019.

[3] Y. Gui, B. Wei, M. Li, J. M. Guerrero, and J. C. Vasquez, "Passivity-based coordinated control for islanded AC microgrid," *Appl. Energy*, vol. 229, pp. 551–561, Nov. 2018.

[4] X. Wang and F. Blaabjerg, "Harmonic stability in power electronic-based power systems: Concept, modeling, and analysis," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2858–2870, May 2019.

[5] X. Guo, Y. Yang, and X. Zhang, "Advanced control of grid-connected current source converter under unbalanced grid voltage conditions," *IEEE Trans. Ind. Electron.*, vol. 65, no. 12, pp. 9225–9233, Dec. 2018.

[6] X. Guo, "A novel CH5 inverter for single-phase transformerless photovoltaic system applications," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 64, no. 10, pp. 1197–1201, Oct. 2017.

[7] G. A. Ramos, R. I. Ruget, and R. Costa-Castelló, "Robust repetitive control of power inverters for standalone operation in DG systems," *IEEE Trans. Energy Convers.*, vol. 35, no. 1, pp. 237–247, Mar. 2020.

[8] O. Husev, C. Roncero-Clemente, E. Makovenko, S. P. Pimentel, D. Vinnikov, and J. Martins, "Optimization and implementation of the proportional-resonant controller for grid-connected inverter with significant computation delay," *IEEE Trans. Ind. Electron.*, vol. 67, no. 2, pp. 1201–1211, Feb. 2020.

[9] G. Keiel, J. V. Flores, C. Lorenzini, L. F. A. Pereira, and A. T. Salton, "Affine discretization methods for the digital resonant control of uninterruptible power supplies," *J. Franklin Inst.*, vol. 356, no. 15, pp. 8646–8664, Oct. 2019.

[10] M. Iftikhar, M. Aamir, A. Waqar, F. B. Muslim, and I. Alam, "Line-interactive transformerless uninterruptible power supply (UPS) with a fuel cell as the primary source," *Energies*, vol. 11, no. 3, p. 542, Mar. 2018.

[11] M. Monfared, S. Golestan, and J. M. Guerrero, "Analysis, design, and experimental verification of a synchronous reference frame voltage control for single-phase inverters," *IEEE Trans. Ind. Electron.*, vol. 61, no. 1, pp. 258–269, Jan. 2014.

[12] Y. W. Li, "Control and resonance damping of voltage-source and current-source converters with *LC* filters," *IEEE Trans. Ind. Electron.*, vol. 56, no. 5, pp. 1511–1521, Nov. 2009.

[13] L. Li, T. Jin, and K. M. Smedley, "A new analog controller for three-phase voltage generation inverter," *IEEE Trans. Ind. Electron.*, vol. 55, no. 8, pp. 2894–2902, Aug. 2008.

[14] Z. Li, C. Zang, P. Zeng, H. Yu, S. Li, and J. Bian, "Control of a grid-forming inverter based on sliding-mode and mixed H_2/H_∞ control," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 3862–3872, Dec. 2017.

[15] Y. Xu and F. Li, "Adaptive PI control of STATCOM for voltage regulation," *IEEE Trans. Power Del.*, vol. 29, no. 3, pp. 1002–1011, Jun. 2014.

[16] H. Sun, Y. Li, G. Zong, and L. Hou, "Disturbance attenuation and rejection for stochastic Markovian jump system with partially known transition probabilities," *Automatica*, vol. 89, pp. 349–357, Mar. 2018.

[17] G. Zong, Y. Li, and H. Sun, "Composite anti-disturbance resilient control for Markovian jump nonlinear systems with general uncertain transition rate," *Sci. China Inf. Sci.*, vol. 62, no. 2, p. 22205, Feb. 2019.

[18] Y. Li, H. Sun, G. Zong, and L. Hou, "Composite anti-disturbance resilient control for Markovian jump nonlinear systems with partly unknown transition probabilities and multiple disturbances," *Int. J. Robust Nonlinear Control*, vol. 27, no. 14, pp. 2323–2337, Sep. 2017.

[19] Z.-L. Zhao and B.-Z. Guo, "On active disturbance rejection control for nonlinear systems using time-varying gain," *Eur. J. Control*, vol. 23, pp. 62–70, May 2015.

[20] Y. Wang, H. R. Karimi, H. Shen, Z. Fang, and M. Liu, "Fuzzy-model-based sliding mode control of nonlinear descriptor systems," *IEEE Trans. Cybern.*, vol. 49, no. 9, pp. 3409–3419, Jun. 2019.

[21] W. Deng and J. Yao, "Extended-state-observer-based adaptive control of electrohydraulic servomechanisms without velocity measurement," *IEEE/ASME Trans. Mechatronics*, vol. 25, no. 3, pp. 1151–1161, Jun. 2020.

[22] W. Deng, J. Yao, and D. Ma, "Time-varying input delay compensation for nonlinear systems with additive disturbance: An output feedback approach," *Int. J. Robust Nonlinear Control*, vol. 28, no. 1, pp. 31–52, Jan. 2018.

[23] B. Jiang, H. R. Karimi, Y. Kao, and C. Gao, "A novel robust fuzzy integral sliding mode control for nonlinear semi-Markovian jump T–S fuzzy systems," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 6, pp. 3594–3604, May 2018.

[24] Y. Cho and K.-B. Lee, "Virtual-flux-based predictive direct power control of three-phase PWM rectifiers with fast dynamic response," *IEEE Trans. Power Electron.*, vol. 31, no. 4, pp. 3348–3359, Apr. 2016.

[25] S. Vazquez, A. Marquez, R. Aguilera, D. Quevedo, J. I. Leon, and L. G. Franquelo, "Predictive optimal switching sequence direct power control for grid-connected power converters," *IEEE Trans. Ind. Electron.*, vol. 62, no. 4, pp. 2010–2020, Apr. 2015.

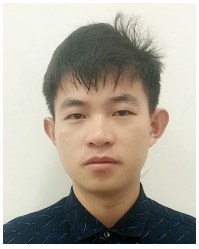
[26] Y. Yu and X. Wang, "Multi-step predictive current control for NPC grid-connected inverter," *IEEE Access*, vol. 7, pp. 157756–157765, 2019.

[27] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 2018.

[28] F. L. Lewis, D. Vrabie, and K. G. Vamvoudakis, "Reinforcement learning and feedback control: Using natural decision methods to design optimal adaptive controllers," *IEEE Control Syst.*, vol. 32, no. 6, pp. 76–105, Dec. 2012.

[29] F. L. Lewis and K. G. Vamvoudakis, "Reinforcement learning for partially observable dynamic processes: Adaptive dynamic programming using measured output data," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 41, no. 1, pp. 14–25, Feb. 2011.

[30] W. Gao and Z.-P. Jiang, "Adaptive dynamic programming and adaptive optimal output regulation of linear systems," *IEEE Trans. Autom. Control*, vol. 61, no. 12, pp. 4164–4169, Dec. 2016.
[31] W. Gao and Z.-P. Jiang, "Adaptive optimal output regulation of time-delay systems via measurement feedback," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 3, pp. 938–945, Mar. 2019.
[32] C. Chen, H. Modares, K. Xie, F. L. Lewis, Y. Wan, and S. Xie, "Reinforcement learning-based adaptive optimal exponential tracking control of linear systems with unknown dynamics," *IEEE Trans. Autom. Control*, vol. 64, no. 11, pp. 4423–4438, Nov. 2019.
[33] Y. Jiang and Z.-P. Jiang, "Robust adaptive dynamic programming with an application to power systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 7, pp. 1150–1156, Jul. 2013.
[34] D. Liu, D. Wang, D. Zhao, Q. Wei, and N. Jin, "Neural-network-based optimal control for a class of unknown discrete-time nonlinear systems using globalized dual heuristic programming," *IEEE Trans. Autom. Sci. Eng.*, vol. 9, no. 3, pp. 628–634, Jul. 2012.
[35] H. Zargarzadeh, T. Dierks, and S. Jagannathan, "Optimal control of nonlinear continuous-time systems in strict-feedback form," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 10, pp. 2535–2549, Oct. 2015.
[36] Y.-M. Li, X. Min, and S. Tong, "Adaptive fuzzy inverse optimal control for uncertain strict-feedback nonlinear systems," *IEEE Trans. Fuzzy Syst.*, vol. 28, no. 10, pp. 2363–2374, Oct. 2020.
[37] Y. Jiang and Z.-P. Jiang, "Global adaptive dynamic programming for continuous-time nonlinear systems," *IEEE Trans. Autom. Control*, vol. 60, no. 11, pp. 2917–2929, Nov. 2015.
[38] J. Teng, W. Gao, D. Czarkowski, and Z.-P. Jiang, "Optimal tracking with disturbance rejection of voltage source inverters," *IEEE Trans. Ind. Electron.*, vol. 67, no. 6, pp. 4957–4968, Jun. 2020.
[39] Y. Cao, Q. Zhao, Y. Ye, and Y. Xiong, "ADRC-based current control for grid-tied inverters: Design, analysis, and verification," *IEEE Trans. Ind. Electron.*, vol. 67, no. 10, pp. 8428–8437, Oct. 2020.



ZHONGYANG WANG was born in Chuzhou, China. He received the B.Eng. degree from the East China University of Technology, Nanchang, China, in 2017, and the M.Eng. degree from Nanchang University, Nanchang, in 2020.

YUNJUN YU (Member, IEEE) was born in Shangrao, China. He received the B.Sc. degree in automation and the M.Sc. degree in control theory and control engineering from Nanchang University, China, in 2000 and 2007, respectively, and the Ph.D. degree from the Chinese Academy of Sciences, in 2013. He is currently an Associate Professor with the Department of Electrical and Automation Engineering, School of Information Engineering, Nanchang University. His research interests include fault diagnosis, data-driven optimal control and it's applied in photovoltaic micro-grid systems, ADRC, and low carbon electricity technology.