Optimal Current Balance Control of Three-Level Inverter under Grid Voltage Unbalance: An Adaptive Dynamic Programming Approach

Yunjun Yu 1,2,*, Zhongyang Wang 1 and Xiaofeng Wan 1

1 Information Engineering School, Nanchang University, Nanchang 330031, China
2 AI Institute, Nanchang University, Nanchang 330031, China

* Correspondence: yuyunjun@ncu.edu.cn; Tel.: +86-1387-080-9767

Received: 25 June 2019; Accepted: 20 July 2019; Published: 25 July 2019

Abstract: When the grid voltage is unbalanced, the positive and negative sequence components in the grid voltage cause grid current to be disordered. Under current balance control, proportional integral (PI) closed-loop control will increase the grid currents instantaneously, which affects the safety and reliability of the inverter operation, and PI parameters are difficult to select without the complete system mathematical model. This paper introduces an adaptive dynamic programming (ADP) approach to solve this problem. The best state feedback controller for the system is obtained by driving the ADP by the value iteration (VI) algorithm without the need for an accurate mathematical model. In the simulations, the ADP approach can improve the dynamic performance of the system, the current increase can be suppressed when the grid voltage is unbalanced, and the harmonic rate of output currents is reduced.

Keywords: grid voltage unbalance; adaptive dynamic programming; state feedback

1. Introduction

1.1. Motivation and Incitement

In a power electronic converter, the inverter can feed back the power to the grid and obtain a sinusoidal current waveform in the grid. Multi-level inverters are widely used in high-voltage and high-power fields compared to conventional inverters [1,2]. Among them, the three-level inverter is widely used because of its simple topology and low cost.

The proportional integral (PI) (note: PI is different from the following policy iterations (PI)) control principle is simple and easy to implement, but no error tracking can be achieved for the amount of communication. Some scholars have proposed grid feed-forward PI control. Although the influence of the grid voltage on the incoming current is eliminated, the complexity of the system is increased [3]. The work in [4] proposed a closed-loop control method for inner loop proportional control and external loop network side inductor current quasi-ratio resonance control. Although this method can effectively suppress static error, it lacks real-time performance; especially the network voltage is heavily distorted. When the dynamic response is slow, it is easy to cause system instability. The work in [5] proposed a control strategy for quasi-ratio resonance plus grid feed-forward. Although the control effect was improved, the system complexity increased.

Operational control strategy under three-phase conditions in grid-connected inverter unbalanced power grids should also be considered. It is necessary to consider how to obtain the three-phase balanced current under the condition of an unbalanced three-phase power grid. The software phase-locked loop based on the decoupling of the double-synchronous coordinate system has higher...
steady-state accuracy, but it relies on phase feedback. Once the phase of the grid is abrupt, there is a large overshoot in the transition process [6]. The work in [7,8] installed a low-pass filter between traditional phase-locked loops. This method can cause defects such as phase angle shift and slow response. The work in [9] introduced droop control of grid voltage unbalance. The work in [10] introduced the robust $H_\infty$ control of the autonomous microgrid. In the existing references, all work was basically based on a complete mathematical model.

Therefore, this paper takes the three-level inverter as the research object and proposes an adaptive dynamic programming method to make the three-phase current balance in the case of three-phase unbalance quickly. This method is an optimal algorithm that attempts to maximize or minimize performance metrics under certain constraints. Dynamic programming is essentially the optimal decision process, but it is difficult to solve dynamic programming due to dimensionality disasters [11,12]. In order to solve this problem, ADP has made a lot of progress [13–18]. Currently, there are two algorithms for solving ADP, which are policy iterations (PI) and value iteration (VI) [19–22]. The main difference between the two methods is that VI is usually implemented offline and PI is usually implemented online [23].

The linear optimal output regulation problem (loopR) can improve system transient performance [24–26]. Under certain conditions, System trace output can be converted to adjustment output [27]. For the discrete control system, $x(k + 1) = F(x(k), u(k))$. In general, for finite time optimal control, its performance indicators are usually written as $J = \sum_{k}^{\infty} (x^T(k)Qx(k) + u^T(k)Ru(k))$. However, for infinite time, this performance indicator function is no longer applicable, since the control track $u(k)$ depends on the change of the tracking track $\eta(k)$. By calculating $J$, it is possible to get infinite results. We can define the performance indicator function for time-varying optimal tracking control and write it as $J = \sum_{k}^{\infty} (\Delta x^T(k)Q \Delta x(k) + \Delta u^T(k)R \Delta u(k))$ [28,29]. We need to solve the riccati equation to find the optimal controller to maximize or minimize the performance index function [30]. However, in that work, the proposed control strategy was based on the complete mathematical model of the system. This method is not suitable when the system is unknown or unmodeled. The work in [31] was an optimal control strategy without the complete system mathematical model. A data-driven approach through input and output data was designed.

1.2. Contribution and Paper Structure

In past works, current control loops were controlled by PI control. The PI controller: $v_d = (k_p + \frac{k_i}{s})(i_d^* - i_d) - \omega L i_q^* - e_d$. $k_p$ and $k_i$ are closely related to the system model. However, in many cases, we are unable to get all the models of the system. The P and I parameters are difficult to select. In this paper, the ADP method was used instead of PI control. The ADP closed-loop control achieved good performance without knowing the complete mathematical model of the system.

In this paper, the inverter output voltage was used as the control variable and the grid current as the state variable. The grid voltage is the external measurable variable of the system, and the optimal control is obtained by minimizing some secondary performance indicators. The closed loop system of the strategy has good transient performance. This paper mainly contributes as follows: The first part introduces the modeling of the inverter on the dq coordinate system [32]. The second part reviews the linear quadratic regulator (LQR) problem, derives the system error equation, sets the performance index function, and rescues the riccati equation. We then show the tracking of the extended state equation of the problem through a linear system to express the information of the system and design the VI method to find the optimal control trajectory. The third part introduces the main results. The fourth part gives the relevant conclusions.
2. Three-Level Inverter

2.1. Mathematical Modeling in the Coordinate System

The main circuit diagram of the three-level inverter is shown in Figure 1.

![Three-level inverter topology](image)

Each phase consisted of four power switches. Take the Phase A switch as an example. When \(T_{11}\) and \(T_{12}\) are turned on at the same time, the output level is \(\frac{V_{dc}}{2}\), denoted as: \(S_a = 1\). When \(T_{12}\) and \(T_{13}\) are turned on at the same time, the output level is zero, denoted as: \(S_a = 0\). When \(T_{13}\) and \(T_{14}\) are turned on at the same time, the output level is \(-\frac{V_{dc}}{2}\), denoted as: \(S_a = -1\). The voltage of the output phase voltage relative to the point \(O\) of the power supply can be represented by the DC side voltage \(V_{dc}\) and the switching function \(S_k\):

\[
u_{ko} = \frac{V_{dc}}{2} S_k
\]

We can get the output voltage of the inverter:

\[
u_k = (S_k - \frac{1}{3} \sum_{k=a,b,c} S_k) V_{dc}
\]

According to Kirchhoff’s law of voltage and current:

\[
L \frac{di_k}{dt} + r_i + e_k = \nu_k
\]

where \(\nu_k\) is the angular frequency.

According to the transformation matrix, we can get the dynamic mathematical model under the \(dq\) coordinate system.

\[
\begin{align*}
L \frac{di_k}{dt} + r_i + e_k &= \nu_k \\
L \frac{di_k}{dt} + r_i + e_k &= \nu_k
\end{align*}
\]

\[
T_{\alpha\beta/dq} = \begin{bmatrix}
cos\varphi & sin\varphi \\
-sin\varphi & cos\varphi
\end{bmatrix}
\]

where \(\varphi\) is the angle between the \(d\) and \(\alpha\)-axis.

According to the transformation matrix, we can get the dynamic mathematical model under the \(dq\) coordinate system.

\[
\begin{align*}
L \frac{di_d}{dt} + r_i + e_d &= \omega Li_d + v_d \\
L \frac{di_q}{dt} + r_i + e_q &= -\omega Li_d + v_q
\end{align*}
\]

where \(\omega\) is the angular frequency.
2.2. Mathematical Modeling under Grid Voltage Unbalance

The grid voltage $e$ consists of positive sequence $e^+$, negative sequence $e^-$, and zero sequence component $e^0$ with the grid voltage unbalanced. For a neutral-free three-phase system, the influence of the zero sequence voltage is often ignored, i.e.,

$$ e = e^+ + e^- $$

In the coordinate system, the grid voltage vector can be expressed as:

$$ e = e^{j\omega t} e^+ + e^{-j\omega t} e^- $$

For the same reason:

$$ i = e^{j\omega t} i^+ + e^{-j\omega t} i^- $$

$$ v = e^{j\omega t} v^+ + e^{-j\omega t} v^- $$

By (4) and the transformation matrix, we can obtain:

$$ \begin{align*}
    e_{dq}^+ &= v_{dq}^+ + L \frac{di_{dq}^+}{dt} + r i_{dq}^+ - jwLi_{dq}^+ \\
    e_{dq}^- &= v_{dq}^- + L \frac{di_{dq}^-}{dt} + r i_{dq}^- - jwLi_{dq}^-
\end{align*} $$

(6)

Positive and negative sequence quantities are decomposed as follows on the $dq$ coordinate axis:

$$ \begin{align*}
    e_d^+ &= e_d^+ + je_q^+ \\
    e_d^- &= e_d^- + je_q^-
\end{align*} $$

(7)

By (6) and (7), we have:

$$ \begin{align*}
    e_d^+ &= v_d^+ + L \frac{di_d^+}{dt} + ri_d^+ + wLi_d^+ \\
    e_q^+ &= v_q^+ + L \frac{di_q^+}{dt} + ri_q^+ - wLi_d^+ \\
    e_d^- &= v_d^- + L \frac{di_d^-}{dt} + ri_d^- + wLi_d^- \\
    e_q^- &= v_q^- + L \frac{di_q^-}{dt} + ri_q^- - wLi_d^-
\end{align*} $$

(8)

By (8), we can obtain:

$$ \begin{align*}
    \dot{x} &= Ax + Bu + D\sigma \\
    y &= Cx \\
    \dot{x}' &= A'x' + B'u' + D'\sigma' \\
    y' &= C'x'
\end{align*} $$

(9)

where $x = [i_d^+; i_q^+], \quad x' = [i_d^-; i_q^-], \quad u = [u_d^+; u_q^+], \quad u' = [u_d^-; u_q^-], \quad \sigma = [e_d^+; e_q^+], \quad \sigma' = [e_d^-; e_q^-], \quad A = A' = \begin{bmatrix} -\frac{1}{\tau} & -w \\
    w & -\frac{1}{\tau} \end{bmatrix}, \quad B = B' = \begin{bmatrix} -\frac{1}{\tau} & 0 \\
    0 & -\frac{1}{\tau} \end{bmatrix}, \quad D = D' = \begin{bmatrix} 1 & 0 \\
    0 & 1 \end{bmatrix}, \quad \text{and} \quad C = C' = \begin{bmatrix} 1 & 0 \\
    0 & 1 \end{bmatrix}$.

Equation (9) is discretized in the following form:

$$ \begin{align*}
    x(k+1) &= G(T)x(k) + H(T)u(k) + M(T)\sigma(k) \\
    y(k) &= Cx(k)
\end{align*} $$

(10)
where $T$ is the sampling period, $G(T) = TA + I$, $H(T) = TB$, and $M(T) = TD$.

### 3. Linear Optimal Output Adjustment Problem

#### 3.1. Problem Description

We consider the following system:

$$
\begin{align*}
    x(k+1) &= Ax(k) + Bu(k) + D\sigma(k) \\
    y(k) &= Cx(k)
\end{align*}
$$

(11)

where $x(k) \in \mathbb{R}^n$ is the state variable of this system. $u(k) \in \mathbb{R}^m$ is the control input. $\sigma(k) \in \mathbb{R}^m$ is the external disturbance of the measurable system, and $y(k) \in \mathbb{R}^n$ is the system output. $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $D \in \mathbb{R}^{n \times m}$, and $C \in \mathbb{R}^{n \times n}$ are certain constant matrices. We need to find an optimal control strategy $u$, so that the system (11) output tracking $\eta(k) \in \mathbb{R}^n$.

**Assumption 1.** For controllable $u$, the system (11) is controllable. The inverter controllability proved in the Appendix A;

**Assumption 2.** There is a one-to-one correspondence between $x(k)$ and $\eta(k)$ ($\eta(k)$ is the ideal state).

We write the ideal control system as follows:

$$
\eta(k+1) = A\eta(k) + Bu^*(k) + D\sigma(k)
$$

(12)

where $\eta(k)$ is the ideal state, $u^*(k)$ is the ideal control, and $\sigma(k)$ is defined as above.

The ideal control input can be written as:

$$
u^*(k) = B^{-1}(\eta(k+1) - A\eta(k) - D\sigma(k))
$$

(13)

Then, from (11) and (12), we can get the error system:

$$
\begin{align*}
    e(k+1) &= Ae(k) + Bv(k)
\end{align*}
$$

(14)

which minimizes the following performance index:

$$
J_e(k) = \sum_0^\infty (e^T(k)Qe(k) + v^T(k)Rv(k))
$$

(15)

where $Q = Q^T \geq 0$, $R = R^T \geq 0$, $e(k) = x(k) - \eta(k)$ is the output error of the system. $v(k) = u(k) - u^*(k)$ is the input error of the system.

According to the above mathematical transformation. The problem of System (11) changes to the optimal adjustment problem of System (14). The optimal controller is:

$$
u(k) = u^*(k) + v(k)
$$

(16)

#### 3.2. Linear Quadratic Regulator

According to the linear optimal theory, the feedback controller can be written as:

$$
v(k) = -Ke(k)
$$

(17)
Since in the LQR problem, each value is under the action of a positive definite symmetric matrix, the performance index can be written as a quadratic form:

\[ J_e(k) = e^T(k)P_e(k) \]  

where \( P^* = (P^*)^T > 0 \) is the only solution for the discrete time algebraic Riccati equation (ARE) below:

\[ P = A^TPA + Q - A^TPB(R + B^TPB)^{-1}B^TPA \]  

Then, in this setup, the optimal feedback control gain is:

\[ K = (R + B^TPB)^{-1}B^TPA \]  

Then, by (14) and (17), we have:

\[ e(k) = (A - BK)e(k - 1) = (A - BK)^k e(0) \]  

For the error system, \( k \to \infty, e(k) \to 0 \).

We note that \( P \) in Equation (19) is non-linear, and the exact value is usually difficult to find. Below, we use the iterative method to find the approximate solution of \( P \). The specific steps are as follows.

**Theorem 1.** The VI Algorithm 1 is able to approximate the solution to (19) with assured convergence. Consider \( P_j \) and \( K_{j+1} \) defined in (22) and (23). Then, for all \( j = 0, 1, \cdots \):

1. \( A - BK_j \) satisfies \( |\lambda_i| < 1 \);
2. \( j \to \infty, K_j \to K^*, \) and \( P_j \to P^* \).

The proof is in Appendix C.

Algorithm 1 VI algorithm [23].

1. Let \( K_0 \) be any stabilizing feedback gain matrix (\( |\lambda_i| < 1 \)), and \( \lambda_i \) are eigenvalues of the matrix \( A - BK_0 \). The proof is in Appendix B). Repeat the following steps for \( j = 0, 1, \cdots \) and \( \tau > 0 \)

2. Solve Equation (19):

\[ P_j = A^TP_jA + Q - K_j^T(R + B^TP_jB)K_j \]  

3. Update control gain by:

\[ K_{j+1} = (R + B^TP_jB)^{-1}B^TP_jA \]  

4. Until \( |P_{j+1} - P_j| \leq \tau \)

However, this optimal controller is completely dependent on the model and requires a perfect understanding of the model. In many cases, we are unable to get all the models of the system. In the next section, we will show how to use the VI method to design the controller.
3.3. Solve ADP by the VI Method

In this section, we design the VI method to find the optimal controller without an accurate system model. We consider the error system (14). The dynamics can be written over the time range \([k-q,k]\) as the unfolded state and the output equation. The new equation is:

\[
e(k) = A^q e(k-q) + B_q \bar{v}(k-1,k-q)
\]

\[
\bar{v}(k-1,k-q) = C_q e(k-q) + T_q \bar{v}(k-1,k-q)
\]

where:

\[
B_q = \begin{bmatrix} B & AB & \cdots & A^{q-1}B \end{bmatrix}
\]

\[
C_q = \begin{bmatrix} A^{q-1} & A^{q-2} & \cdots & I \end{bmatrix}^T
\]

\[
\bar{v}(k-1,k-q) = \begin{bmatrix} e(k-1) \\ e(k-2) \\ \vdots \\ e(k-q) \end{bmatrix}, \bar{v}(k-1,k-q) = \begin{bmatrix} v(k-1) \\ v(k-2) \\ \vdots \\ v(k-q) \end{bmatrix}
\]

and:

\[
T_q = \begin{bmatrix} 0 & B & AB & \cdots & A^{q-2}B \\ 0 & 0 & AB & \cdots & A^{q-3}B \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}
\]

By (24), we can obtain:

\[
e(k) = Mz(k) = \begin{bmatrix} M_e & M_v \end{bmatrix} z(k)
\]

with \(M = [A^q C_q^+, B_q - A^q C_q^+ T_q] \), \(C_q^+ \) is the generalized inverse matrix of \(C_q \), and \(z(k) = \begin{bmatrix} \bar{v}(k-1,k-q) \\ \bar{v}(k-1,k-q) \end{bmatrix} \).

Let us assume \(E = Q + A^T P A, F = A^T P B, \) and \(G = R + B^T P B. \) Then, \(K = G^{-1} F^T \) and \(G^T = G. \)

By (19), we can obtain:

\[
e^T(k) Pe(k) = e^T(k)(E - FG^{-1} F^T)e(k)
\]

and:

\[
e^T(k) FG^{-1} F^T e(k) = -v^T(k) F^T e(k)
\]

\[
= -e^T(k) F v(k)
\]

\[
= v^T(k) G v(k)
\]

\[
e^T(k) F e(k) - e^T(k) F v(k) - v^T(k) G v(k)
\]

By (26) and (27), we have:

\[
e^T(k) Pe(k) = \begin{bmatrix} e(k) \\ v(k) \end{bmatrix}^T \begin{bmatrix} E & F \\ F^T & G \end{bmatrix} \begin{bmatrix} e(k) \\ v(k) \end{bmatrix}
\]

By (25), we have:

\[
e^T(k) Pe(k) = \begin{bmatrix} z(k) \\ v(k) \end{bmatrix}^T \begin{bmatrix} E' & F' \\ F'^T & G' \end{bmatrix} \begin{bmatrix} z(k) \\ v(k) \end{bmatrix}
\]
Then, the following properties hold \cite{33}:
\[ e^T(k)Pe(k) = e^T(k)Qe(k) + v^T(k)Re(k) + e^T(k + 1)Pe(k + 1) \] (28)

and:
\[ e^T(k + 1)Pe(k + 1) = \begin{bmatrix} z(k + 1) \\ v(k + 1) \end{bmatrix}^T \begin{bmatrix} E' \\ F'^T \end{bmatrix} \begin{bmatrix} z(k + 1) \\ v(k + 1) \end{bmatrix} \]
\[ + \begin{bmatrix} z(k + 1) \\ -K'z(k + 1) \end{bmatrix}^T \begin{bmatrix} E' \\ F'^T \end{bmatrix} \begin{bmatrix} z(k + 1) \\ v(k + 1) \end{bmatrix} \]

where \( v(k + 1) = -Ke(k) = -KMz(k) \). Then, \( K' = KM = G'F'T \).

Change Equation (19) to the following equation:
\[ \begin{bmatrix} z(k) \\ v(k) \end{bmatrix}^T \begin{bmatrix} E' \\ F'^T \end{bmatrix} \begin{bmatrix} z(k + 1) \\ v(k + 1) \end{bmatrix} - \begin{bmatrix} z(k + 1) \\ -K'z(k + 1) \end{bmatrix}^T \begin{bmatrix} E' \\ F'^T \end{bmatrix} \begin{bmatrix} z(k + 1) \\ v(k + 1) \end{bmatrix} = e^T(k)Qe(k) + v^T(k)Re(k) \]

From the previous, we know that \( p \) is only related to \( M_s, M_v, \) and \( P_j, \) \( \lim_{j \to \infty} P_j = P^* \). \( P^* \) is only related to the system characteristics and has nothing to do with the input and output of the system. For different control signals \( v(k) \), the \( P \) value is determined. When \( j \to \infty, E'_j \to E'_s, F'_j \to F'_s, G'_j \to G'_s, \) and \( K'_j \to K'_s \).

For a symmetric matrix \( P \in \mathbb{R}^{m \times m} \) and a column vector \( v \in \mathbb{R}^n \), we assume 
\[ \text{vecs}(P) = [P_{11}, P_{12}, \ldots, P_{m-1,m}, P_{mn}]^T \in \mathbb{R}^{\frac{1}{2}m(n+1)} \] and 
\[ \text{vec}(v) = [v_1^T, v_2^T, \ldots, v_n^T]^T \in \mathbb{R}^{1 \times n} \].

Given a positive integer \( k_0 \) and \( q < k_0 \) and a sufficiently large positive integer \( s \) and given an arbitrary control policy \( v(k) \) on \( [0, k_0 + s] \), let:
\[ \psi^l(k) = e^T(k)Qe(k) + v^T(k)Re(k) + \text{vec}(a(k))^T\text{vecs}( \begin{bmatrix} E' \\ F'_j \end{bmatrix} ) \]

and:
\[ \Theta = \begin{bmatrix} \text{vec}(\beta(k_0))^T \\ \text{vec}(\beta(k_0 + 1))^T \\ \vdots \\ \text{vec}(\beta(k_0 + s))^T \end{bmatrix} \]

where \( a(k) = [z(k + 1), -K'z(k + 1)] \) and \( \beta(k) = [z(k), v(k)] \). For \( j = 0, 1, 2, \ldots \), we define:
\[ \Psi^l = [\psi^l(k_0), \psi^l(k_0 + 1), \ldots, \psi^l(k_0 + s)]^T \] (29)

Then:
\[ \Theta\text{vecs}( \begin{bmatrix} E'_{j+1} & F'_{j+1} \\ F'^T_{j+1} & G'_{j+1} \end{bmatrix} ) = \Psi^l \] (30)

**Remark 1.** As \( j = 0 \), matrices \( P_0 = 0 \) and \( K_0 = 0 \) indicate that 
\[ \psi^0(k) = e^T(k)Qe(k) + v^T(k)Re(k) + e^T(k + 1)Qe(k + 1) \]. Therefore, when implementing Algorithm 1, we do not need the initial values \( K'_0 \) and \( Y_0 \).

Then, the following properties hold \cite{33}:

1. \( j \to \infty, Y_j \to Y^* \),
Hence, Algorithm 2 converges, and the closed-loop system is stable. The VI-based measurement feedback ADP flowchart is shown in Figure 2.

Algorithm 2 ADP algorithm by the VI method.

1. Select an error \( \tau > 0 \) and \( j = 0 \).
2. Calculate \( K'_j \):

\[
\Theta_{\text{vecs}}(Y_{j+1}) = \Psi^j
\]
\[
K'_{j+1} = (Y_{j+1}^{22})^{-1}Y_{j+1}^{21}
\]

where
\[
Y_{j+1} = \begin{bmatrix}
E'_{j+1} & F'_{j+1} \\
E'^T_{j+1} & G'_{j+1}
\end{bmatrix}
\]

3. Until \( |Y_{j+1} - Y_j| < \tau \)

Get the approximated optimal control policy:

\[
v(k) = -K'_j z(k)
\]  

(32)

Theorem 2. The original system optimal controller is:

\[
u(k) = v(k) + u^*(k)
\]  

(33)
where $u^*(k) = -\tilde{R}'z(k)$ and $z(k) = \begin{bmatrix} \tilde{y}(k-1,k-q) \\ u^*(k-1,k-q) \\ \tilde{\sigma}(k-1,k-q) \end{bmatrix}$. The proof is in Appendix D.

The optimal feedback controller is:

$$u(k) = -K'_r z(k) - \tilde{R}'z(k)$$

(34)

4. Simulations

In this paper, $r = 0.1 \Omega$, $L = 1.5 \text{ mH}$, $V_{dc} = 1000 \text{ V}$, $T_s = 10^{-5} \text{ s}$, and $Q = R = \text{eye}(3)$. Given $i_d^* = 10$, $i_q^* = 0$, and $q = 1$, the inverter filter structure diagram and the comparison of the two control methods are shown in Figure 3. The inverter structure and $dq$ component separation are shown in Figure 3a.

![Figure 3. Inverter block diagram (a). PI closed-loop control block diagram (b). ADP closed-loop control block diagram (c).](image)

PI control is shown in Figure 3b. The PI regulation decoupling equation is:

$$\begin{cases} v_d = (k_p + \frac{k_i}{2})(i_d^* - i_d) - wLi_q^* - e_d \\ v_q = (k_p + \frac{k_i}{2})(i_q^* - i_q) + wLi_d^* - e_q \end{cases}$$

(35)
where $K_p = 0.02$ and $K_I = 500$ [3]. ADP control is shown in Figure 3c. Calculate $v$ and $u^*$ according to the current error and the command current, respectively. In Figure 3c, we can obtain ADP by Algorithm 1:

$$v_{dq}(k) = -K'_{s} \begin{bmatrix} e_{12}(k-1) \\ \psi_{12}(k-1) \end{bmatrix} - \bar{K}'_{s} \begin{bmatrix} i_{dq}(k-1) \\ \psi_{dq}(k-1) \end{bmatrix}$$

where:

$$K'_{s} = \begin{bmatrix} -2.2762 & 0.0144 & 0.0153 & -0.0000 \\ 0.0144 & -2.2762 & 0.0000 & 0.0153 \end{bmatrix}$$

and:

$$\bar{K}'_{s} = \begin{bmatrix} -2.2762 & 0.0144 & 0.0153 & -0.0000 & -0.0153 & 0.0000 \\ 0.0144 & -2.2762 & 0.0000 & 0.0153 & -0.0000 & -0.0153 \end{bmatrix}$$

4.1. Grid Voltage Balance

Grid voltage is shown in Figure 4a. It can be seen that the grid voltage is three-phase symmetrical with an amplitude of 310 V.

The grid current of PI closed-loop control and ADP closed-loop control is shown in Figure 4b,c, respectively. The figure shows in the case of grid voltage balance that the grid current tracking command value when the system reaches stability. From Figure 4b, we can see that the grid current reaches the steady state value at $t = 0.005$ s. From Figure 4c, we can see that the grid current reaches the steady state value at $t = 0.035$ s. This shows that the grid current of ADP closed-loop approach had a faster response speed than the PI closed-loop approach.

![Figure 4](image)

**Figure 4.** Grid voltages (a). Grid currents of PI closed-loop control (b). Grid currents of ADP closed-loop control (c).

4.2. Grid Voltage Unbalance

The grid voltage unbalance is shown in Figure 5a; at $t = 0.3$ s, the amplitude of the phase A voltage dropped by 80%.

The grid current of the PI closed-loop approach and the ADP closed-loop approach is shown in Figure 5b,c, respectively. As shown in Figure 5b, after the Phase A voltage dropped, the Phase A current rose sharply, and the highest amplitude reached 100 A. The Phase B and C currents also rose and finally stabilized at about 60 A. The current increased several times compared to the normal
operation, which can damage the entire system. As shown in Figure 5c, after the Phase A voltage dropped, the phase A current saw almost no change, and the Phase B and C currents were reduced.

![Figure 5](image_url)

**Figure 5.** Grid voltages (a). Grid currents of PI closed-loop control (b). Grid currents of ADP closed-loop control (c).

### 4.3. Grid Current Balance Control

The grid voltage unbalance had positive and negative sequence components. Therefore, the positive and negative sequence components should be controlled separately. To suppress the negative sequence components, the current command is:

\[
\begin{align*}
    i_d^{+} &= 10 \\
    i_q^{+} &= 0 \\
    i_d^{-} &= 0 \\
    i_q^{-} &= 0
\end{align*}
\]  

(37)

By (9) and (10), positive and negative sequence component control had the same mathematical model. Therefore, the positive and negative sequences had the same control loop. By (35), the positive and negative sequence PI adjustment decoupling equations are:

\[
\begin{align*}
    v_d^+ &= (k_p + \frac{k_i}{s})(i_d^{+} - i_d) - wLi_d^{+} - e_d^+ \\
    v_q^+ &= (k_p + \frac{k_i}{s})(i_q^{+} - i_q) + wLi_q^{+} - e_q^+
\end{align*}
\]  

(38)

and:

\[
\begin{align*}
    v_d^- &= (k_p + \frac{k_i}{s})(i_d^{-} - i_d) - wLi_d^{-} - e_d^- \\
    v_q^- &= (k_p + \frac{k_i}{s})(i_q^{-} - i_q) + wLi_q^{-} - e_q^-
\end{align*}
\]  

(39)

Then, negative and positive sequences had the same control process. By (5) and (6), the output voltage of the inverter is:

\[
v_{dq} = e^{j\omega t}v_{dq}^+ + e^{-j\omega t}v_{dq}^-
\]  

(40)
In this paper, the positive and negative sequence components were separated by modulus detection. By (6), in the $\alpha\beta$ coordinate system, the grid voltage positive sequence vector modulo $|e^+|$, the angular frequency $\omega$ rotated counterclockwise, the grid voltage negative sequence vector modulo $|e^-|$, and angular frequency $\omega$ rotated clockwise are shown in Figure 6.

Both had the same angular velocity and the opposite direction of rotation. $|e|$ was the largest when two vectors were in the same direction, and $|e|$ was the smallest when in the opposite direction.

\[
\begin{align*}
\max |e| &= |e^+| + |e^-| \\
\min |e| &= |e^+| - |e^-|
\end{align*}
\]

(41)

The modulus of the positive and negative sequence components is:

\[
\begin{align*}
|e^+| &= \frac{1}{2}(\max |e| + \min |e|) \\
|e^-| &= \frac{1}{2}(\max |e| - \min |e|)
\end{align*}
\]

(42)

We can get the positive and negative order $dq$ components respectively:

\[
\begin{align*}
e_d^+ &= |e^+| \cos \phi \\
e_q^+ &= |e^+| \sin \phi
\end{align*}
\]

and:

\[
\begin{align*}
e_d^- &= |e^-| \cos \phi \\
e_q^- &= |e^-| \sin \phi
\end{align*}
\]

The grid voltage unbalance is shown in Figure 7a; at t = 0.3 s, the amplitude of the Phase A voltage dropped by 80%.

The grid voltage $dq$ axis positive and negative sequence components are shown in Figure 7b.

The grid current of PI closed-loop control and ADP closed-loop control is shown in Figure 7c,d, respectively. As shown in Figure 7c, after the Phase A voltage dropped, the grid currents increased instantaneously, and the maximum was 40 A. After a brief rise, the current gradually approached stability, and when t = 0.345 s, the grid currents were stable. As shown in Figure 7d, after the Phase A voltage dropped, the grid current had almost no fluctuations in this process, keeping the command value. From this, we can see that the balanced current strategy using ADP closed-loop control not only could suppress the sharp rise of current, but also made the current respond quickly to the steady state.
The current overshoot of two control methods is shown in Table 1. Under PI closed-loop control, \( \text{Overshoot}(A) < \frac{40-10}{10} = 300\% \), \( \text{Overshoot}(B) = \frac{40-10}{10} = 300\% \), and \( \text{Overshoot}(C) < \frac{40+10}{10} = -300\% \). Under the ADP closed-loop control, the grid currents were not overshot.

![Figure 7](image)

**Figure 7.** Grid voltages (a). Positive and negative sequence \( dq \) component of grid voltages (b). Grid currents of PI closed-loop control (c). Grid currents of ADP closed-loop control (d).

<table>
<thead>
<tr>
<th>Control Method</th>
<th>Phase A Current</th>
<th>Phase B Current</th>
<th>Phase C Current</th>
</tr>
</thead>
<tbody>
<tr>
<td>PI closed loop control</td>
<td>&lt;300%</td>
<td>300%</td>
<td>−300%</td>
</tr>
<tr>
<td>ADP closed loop control</td>
<td>No overshoot</td>
<td>No overshoot</td>
<td>No overshoot</td>
</tr>
</tbody>
</table>

### 4.4. Grid Current Balance Control of Command Power

By [34], the command current calculation formula is:

\[
i_{\alpha \beta}^* = \frac{2}{3} \left\{ \frac{1}{e^{j\omega t}} - e^{-j\omega t} \right\} \left[ (p^* - q^*)e^{j\omega t} - (p^* + jq^*)e^{-j\omega t} \right]
\]

In this paper, we assumed \( p^* = 6000W \) and \( q^* = 2000W \). The command current \( \alpha \) and \( \beta \) components are shown in Figure 8a. The grid currents of PI closed-loop control and ADP closed-loop control are shown in Figure 8b,c, respectively. We can see that both control methods stabilized the current. Two control methods’ Phase A current comparison is shown in Figure 8d. We can see that the current of the ADP closed-loop control method was smoother at steady state. The THD of grid currents are shown in Table 2. We can see that the three-phase grid currents \( THD_{PI} \) were 2.04\%, 1.87\%, and 1.97\%, and \( THD_{ADP} \) were 1.30\%, 1.29\%, 1.29\%. This shows that the ADP closed-loop control method had better steady state performance.
4.5. Discussions

The main purpose of this paper is to introduce the application of adaptive dynamic programming in grid-connected inverters. The simulation experiment was performed in the MATLAB/Simulink environment. Although the verification was for the Phase A voltage drop of 80%, this can be verified for other situations.

5. Conclusions

In this paper, an ADP algorithm was proposed instead of the original PI algorithm. The optimal tracking problem was solved without the complete system mathematical model. In the control strategy proposed in this paper, the grid currents can track the command currents. Compared with PI closed-loop control, ADP closed-loop control reduces current oscillation and improved response speed.

In this paper, we only considered current as the indicator. However, in the actual microgrid system, there are many indicators that affect system performance. In the hierarchical droop-based control strategies [35, 36], we can set different performance indicators under different requirements, such as $\Delta P$, $\Delta Q$ ($\Delta P = P - P^*$, and $P$ and $P^*$ represent output active power and command active power, respectively. $\Delta Q = Q - Q^*$, and $Q$ and $Q^*$ represent output reactive power and command reactive power, respectively). In island mode, we can set $\Delta w$ and $\Delta U (\Delta w = w - w^*$, and $w$ and $w^*$ represent the output angular frequency and the given angular frequency, respectively. $\Delta U = U - U^*$, and $U$ and $U^*$ represent the output voltage and the given voltage, respectively) as performance indicators. Then, we can set the performance indicator function according to the selected indicators and find the optimal controller by the ADP approach. This reduces system oscillations and improves performance.

Author Contributions: Y.Y., Z.W., and X.W. proposed the idea. Y.Y., Z.W., and X.W. built the MATLAB/Simulink simulation platform. Z.W. wrote the first draft of the manuscript. All authors contributed to and have approved the final manuscript.
Funding: This work was supported in part by the National Science Foundation of China (NSFC: 61563034) and the International S&F Cooperation Program of China (ISTCP:2014DFG72240).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

Variables

- $S_k$: Switching variable
- $u_k$: Inverter output voltage
- $i_k$: Grid current
- $i_{dq}$: Grid current dq-axis component
- $i_{dq}^+$: Grid current positive sequence dq-axis component
- $i_{dq}^-$: Grid current negative sequence dq-axis component
- $e_k$: Grid voltage
- $e_{dq}$: Grid voltage dq-axis component
- $e_{dq}^+$: Grid voltage positive sequence dq-axis component
- $e_{dq}^-$: Grid voltage negative sequence dq-axis component
- $\phi$: Angle between the d-axis and the a-axis
- $w$: Angular frequency
- $J_e(k)$: Performance
- $P$: Solution of ARE
- $K$: Error system feedback control gain
- $e(k)$: Current error
- $v(k)$: Error system optimal controller
- $u^*(k)$: Reference control track
- $u(k)$: Original system controller
- $K'_e$: Error system optimal feedback control gain
- $K'_o$: Original system optimal feedback control gain

Acronyms

- PI: Proportional integral
- ADP: Adaptive dynamic programming
- PI: Policy iteration
- VI: Value iteration
- LQR: Linear quadratic regulator
- ARE: Algebra Riccati equation
- DT: Discrete-time

Constants

- $r$: Filter resistor
- $L$: Filter inductor
- $V_{dc}$: DC side voltage
- $T_s$: Sampling period
- $Q$: State error weighting matrix
- $R$: Control error weighting matrix
- $q$: Number of online datasets
- $i^*_{dq}$: Command current dq-axis component
- $i^+_{dq}$: Command current positive sequence dq-axis component
- $i^-_{dq}$: Command current negative sequence dq-axis component
- $K_P$: Proportional parameter
- $K_I$: Integral parameter
Appendix A

Proof. For a linear discrete system:

\[ x(k + 1) = Gx(k) + Hu(k) \]  \hspace{1cm} (A1)

where \( x \) is an n-dimensional state vector and \( u \) is an n-dimensional input vector. By (A1), we can obtain:

\[ x(k) = G^k x(0) + \sum_{i=0}^{k-1} G^{k-1-i} Hu(i) \]  \hspace{1cm} (A2)

We assume \( x(n) = 0 \) when \( k = n \), then:

\[ x(0) = - \sum_{i=0}^{n-1} G^{-1-i} Hu(i) = -[G^{-1} Hu(0) + G^{-2} Hu(1) + \cdots + G^{-n} Hu(n-1)] \]

\[ = -S \begin{bmatrix} u(0) \\ u(1) \\ \vdots \\ u(n-1) \end{bmatrix} \]  \hspace{1cm} (A3)

where \( S = \begin{bmatrix} G^{-1} & G^{-2} & \cdots & G^{-n} \end{bmatrix} \in \mathbb{R}^{2n \times n} \). For any initial value \( x(0) \), the necessary and sufficient condition for the solution of Equation (A3) is:

\[ \text{rank}(S) = n \]  \hspace{1cm} (A4)

In this paper, \( r = 0.1\Omega \), \( L = 1.5 \text{ mH} \), \( w = 314 \text{ rad/s} \), \( A = \begin{bmatrix} -\frac{r}{L} & -\frac{w}{L} \\ \frac{w}{L} & \frac{1}{L} \end{bmatrix} \), \( B = \begin{bmatrix} -\frac{1}{L} & 0 \\ \frac{r}{L} & -\frac{1}{L} \end{bmatrix} \), \( G = TA + I \), and \( H = TB \). Then:

\[ x(k + 1) = Gx(k) + Hu(k) = \begin{bmatrix} 0.9993 & -0.0031 \\ 0.0031 & 0.9993 \end{bmatrix} x(k) + \begin{bmatrix} -0.0067 & 0 \\ 0 & -0.0067 \end{bmatrix} u(k) \]  \hspace{1cm} (A5)

Then:

\[ S = \begin{bmatrix} G^{-1} & G^{-2} \end{bmatrix} = \begin{bmatrix} -0.0067 & 0 \\ 0 & -0.0067 \end{bmatrix} \]  \hspace{1cm} (A6)

\[ \text{rank}(S) = 2 \]  \hspace{1cm} (A7)

The proof is thus complete. \( \square \)
Appendix B

Proof. For a matrix $A - BK_0$, let us assume that it has $n$ unequal eigenvalues, which are $\lambda_1, \lambda_2, \lambda_3, \cdots, \lambda_n$. The feature matrix of $A - BK_0$ is $T$. We have:

$$
e(k) = (A - BK_0)^k e(0) = (T \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_n \end{bmatrix} T^{-1})^k e(0)$$

$$= T \begin{bmatrix} \lambda_1^k & & \\ & \lambda_2^k & \\ & & \ddots \\ & & & \lambda_n^k \end{bmatrix} T^{-1} e(0) \quad (A8)$$

Since $T, e(0)$ are bounded and $|\lambda_i| < 1$. Then, $k \to \infty$, and we have $|\lambda_i|^k \to 0$ and $e(k) \to 0$. If matrix $A - BK_0$ has heavy roots, we have the same proof. □

Appendix C

Proof of Theorem 1. By (15), (18), and (21):

$$P = \sum_{k=0}^{\infty} (A - BK)^T (Q + K^T R K) (A - BK)^k \quad (A9)$$

When $j = 0$, by (22) and (A9), we have:

$$P_0 - P_1 = \sum_{k=0}^{\infty} A_1^T (K_0 - K_1)^T R (K_0 - K_1) A_1^k \geq 0 \quad (A10)$$

where $A_j = (A - BK_j)^k$. Then, by (19) and (A9), we obtain:

$$P_1 - P^* = \sum_{k=0}^{\infty} A_1^T (K_1 - K^*)^T R (K_1 - K^*) A_1^k \geq 0 \quad (A11)$$

Therefore, we have $P_0 \leq P_1 \leq P^*$. For $j = 0, 1, \cdots$, we can obtain that $P_j$ is the existence of a unique value. The proof is thus complete. □

Appendix D

Proof of Theorem 2. By (12), (24), and (25), we can obtain:

$$\eta(k) = \overline{M} z(k) = \begin{bmatrix} M_e & M_v & M_\sigma \end{bmatrix} z(k) \quad (A12)$$

where $M_\sigma = D_q - A^q C_q^{+} T_q$ and $D_q = \begin{bmatrix} D & AD & \cdots & A^{q-1}D \end{bmatrix}$.

The ideal control system and error control system have the same coefficient matrices $A$ and $B$. According to the above calculation process, we have:

$$\tilde{K}_j = \tilde{K}_j' \begin{bmatrix} M_e & M_v & M_\sigma \end{bmatrix} = \tilde{G}' T' \quad □$$
References


