Joint Online Adaptive Optimal Tracking Control and Frequency-Response Method for Speed of PMSM and DC-Link Voltage Peak Controller Design in Bi-directional Quasi Z-Source Inverter

Cong-Thanh Pham(✉)

Faculty of Electrical and Electronic, Vietnam Aviation Academy, Ho Chi Minh, Vietnam
thanhpc@vaa.edu.vn

Abstract. This paper proposes a control strategy with three joint controllers to improve the speed response of a permanent magnet synchronous motor (PMSM) in an electric vehicle (EV) system. The first controller is an online adaptive optimal tracking controller (OTC) to enhance the tracking response of the speed of the PMSM. This controller is derived based on reinforcement learning (RL) method by approximating the solutions of the Hamilton-Jacobi-Isaacs (HJI) equations while still ensuring the stability of the closed-loop system of PMSM. Three components are approximated simultaneously online: one performance index (one critic), one control law (the first actor), and one worst disturbance law (the second actor). On the other hand, the DC-link voltage peak (DVP) the Bi-directional Quasi Z-Source Inverter (BZI) needs to be regulated so that the PMSM can be operational in the high-speed zone. Therefore, the next two controllers are applied based on a frequency-response method to improve the responses of the inductor current and the DVP in the closed-loop control system of the BZI with unknown dynamics of the PMSM and BZI (PZI). The simulation results of the speed response with the PZI model show the effectiveness and robustness of the proposed control strategy.

Keywords: Bi-directional Quasi-Z-source inverter · Permanent magnet synchronous motors · Optimal control · Reinforcement learning

1 Introduction

In practical electric vehicle (EV) systems, the supply voltage level performs poorly in the high speed range of the permanent magnet synchronous motor (PMSM). Therefore, many automotive applications add a DC-to-DC converter to the structure of the frequency converter to regulate the supplying DC-link voltage (DCV). The quality of interconnection between the speed of the PMSM and

N.-S. Vo et al. (Eds.): INISCOM 2024, LNICST 595, pp. 28–45, 2024.
https://doi.org/10.1007/978-3-031-67357-3_3

the DCV has drawn a signification attention from the experts and researchers in both industrial and academic sectors. A vector control method called field-oriented control (FOC) is implemented to control the speed of the PMSM. This FOC framework has three proportional plus integral (PI) controllers coupled together in a cascade structure. The DCV control system has two PI controllers to regulate the supply DCV required by the inverter [1,2]. Thus, there are five PI controllers in total for such a control framework. The problem of tuning the parameters of the PI controllers introduces a major challenge with continuous changes in value of the speed, the DC-input voltage, the resistance of the PMSM and the load torque.

The work in [3] proposed a control algorithm with self-tuning ability for the speed of the PMSM against the variations of the load torque. In this algorithm, the speed of the PMSM and the current are controlled in a cascade structure; however, the mathematical model must be known accurately for the algorithm to operate effectively. To handle the tuning problem of the many parameters of the traditional PI controllers, particle swarm optimization (PSO) and genetic algorithms (GA) can be used [4]. However, the micro-controller rate and the convergence rate of these control strategies are the major concerns [5].

In recent years, the GA and the PSO methods have been alternated with reinforcement learning (RL) to obtain the optimal control. With RL, the Hamilton-Jacobi-Isaacs (HJB) solution [6] and the HJI equations [7] are approximated to derive the optimal control law. Therefore, this paper proposes an online adaptive optimal tracking controller (OTC) based on RL method for the speed of the PMSM. The core idea of this method is to emulate the two-player zero-sum game problem in solving the solutions of the HJI equations in a real-time fashion. Specifically, this method applies the policy iteration principle to learn these solutions online. This way can improve the robustness and quality of the speed response of the PMSM.

In addition to developing the tracking controller for the speed, the DCV which is the supplying voltage for the inverter in the BZI scheme is required to be regulated. The PMSM can only operate above the rated speed when the DCV is boosted greater than the rated voltage. In case of this condition fails, the space vector pulse width modulation (SVM) falls into the over-modulation region [4]. On the other hand, the DCV is still affected by ripples when the DC-input voltage, the load torque, and the PMSM speed change continuously [1,8].

Based on the aforementioned discussions, there is a need for improving the response of the DCV so that the system can operate adaptively when the DC-input voltage decreases over time. It is essential that the DCV is able to follow the DC voltage standard because it affects the modulation quality of the inverter. If this voltage drops below a certain threshold, the performance of the SVM reduces drastically in the over-modulation region. This also leads to a reduction in the quality of rotor speed control. The DCV of the BZI is a pulse waveform, so it cannot be controlled directly. Instead, its maximum value is equal to the sum of the voltages across the two capacitors C_1 and C_2 [1], both of which can be

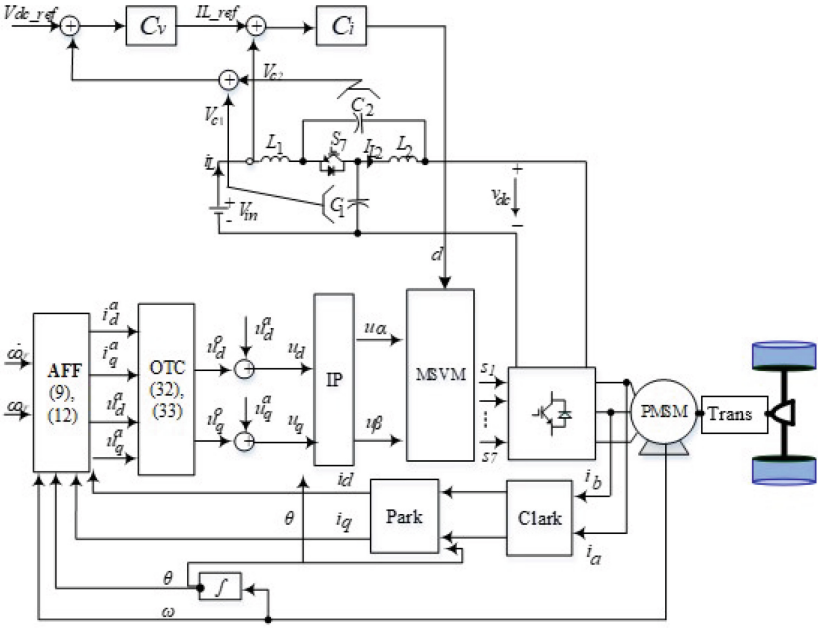

Fig. 1. Online adaptive optimal tracking control speed of PMSM fed by a BZI scheme.

controlled. Considering the nonlinear relationship between the DC-link voltage peak (DVP) and the duty d [1], if the DVP controller is not properly designed, it provides the correct voltage to the inverter, causing the performance of the PMSM system unsatisfactory. Therefore, this paper also proposes a frequency-response method to design the gains of the DVP controller in the BZI scheme (DC side).

This paper is organized as follows. Section 2 introduces the PZI model and the overall of the design control laws. Online adaptive tracking control for the PMSM speed is explained in Sect. 3. In Sect. 4, we present and discuss the simulation results of the proposed PZI control framework. Finally, the conclusion of this paper is given in Sect. 5.

2 The PZI Model and Design Control Laws

The PZI structure is detailed in Fig. 1 which includes two important controllers: one for the speed of the PMSM and the other for the DVP. It is a challenge to exactly determine the mathematical model of this structure due to the variations of the speed of the PMSM, the load torque, and the DC-input voltage. In this section, a mathematical model of the PMSM is formulated and used to design an augmented feed-forward control law for its speed. On the other hand, the model of the BZI scheme is also defined in (2); while a frequency-response method is utilized to develop the DVP controller in the BZI scheme as shown in Fig. 1.

2.1 The Mathematical Model of PMSM and BZI Scheme

The mathematical model of the PMSM system is defined in [9], given by

$$\begin{cases} \dot{\omega} = -\frac{B}{J}.\omega - \frac{T_L}{J} + \frac{1{,}5n_p.\Psi_f}{J}.i_q + \frac{n_\tau}{J}, \\ \dot{i}_d = -\frac{R_s}{L_d}.i_d + n_p.\omega.i_q + \frac{1}{L_d}.u_d + n_d, \\ \dot{i}_q = -n_p.\omega.i_d - \frac{R_s}{L_d}.i_q - n_p.\frac{\Psi_f}{L_q}.\omega + \frac{1}{L_q}.u_q + n_q, \end{cases} \tag{1}$$

where ω is the speed of PMSM; u_d and u_q are two voltages on the dq-axes, respectively; i_d and i_q are two stator currents on the dq-axes, respectively; $|u_d|$ and $|u_q|$ are smaller than λ, λ is the voltage limit on dq-axis, ($\lambda < 1$); L_d and L_q are the stator inductances of the dq-axes ($L_d = L_q$); Ψ_f is the permanent magnet flux linkage, R_s is the stator resistance, J is the moment of inertia,T_L is the load torque, B is the viscous friction coefficient, and n_p is the number of pole pairs.

The mathematical model of the BZI system is defined as [1,10,11]

$$\dot{x} = A.x + B.d, \tag{2}$$

where $x = \left[i_{L_1}\ i_{L_2}\ v_{c_1}\ v_{c_2} \right]^T$ is the state variable, d is the duty cycle [1], and A and B are respectively given by

$$A = \begin{bmatrix} -\frac{r_L+R_c}{L} & 0 & \frac{\overline{D}-1}{L} & \frac{\overline{D}}{L} \\ 0 & -\frac{r_L+R_c}{L} & \frac{\overline{D}}{L} & \frac{\overline{D}-1}{L} \\ \frac{1-\overline{D}}{C} & -\frac{\overline{D}}{C} & 0 & 0 \\ -\frac{\overline{D}}{C} & \frac{1-\overline{D}}{C} & 0 & 0 \end{bmatrix}, B = \begin{bmatrix} \frac{\overline{V}_{c_1}+\overline{V}_{c_2}-R_c.\overline{I}_{Load}}{L} \\ \frac{\overline{V}_{c_1}+\overline{V}_{c_2}-R_c.\overline{I}_{Load}}{L} \\ \frac{\overline{I}_{Load}-\overline{I}_{L_1}+\overline{I}_{L_2}}{C} \\ \frac{\overline{I}_{Load}-\overline{I}_{L_1}+\overline{I}_{L_2}}{C} \end{bmatrix}.$$

The bi-directional quasi z-source network is also included in the Fig. 1, the components include: two capacitors C_1, C_2 ($C_1 = C_2 = C$) and two inductors L_1, L_2, ($L_1 = L_2 = L$); a resistance series of r_L and L, and a resistance series of R_c and C; a load current i_{load}; a DC-input voltage v_{in}. Here, the DVP equals the total voltage on the two capacitors ($v_{dc} = v_{c_1} + v_{c_2}$); v_{c_1} and v_{c_2} are the small signal value of voltage on the two capacitors C_1 and C_2, respectively; i_{L_1} and i_{L_2} are the small signal value of inductor currents L_1 and L_2, respectively [1,11]; d is the small signal value of duty cycle and $\overline{D}$ is the average value of duty cycle; $\overline{V}_{c_1}$ is the average value of capacitor voltage C_1; $\overline{V}_{c_2}$ is the average value of capacitor voltage C_2; $\overline{I}_{L_1}$ and $\overline{I}_{L_2}$ are the average value of two inductor currents through L_1 and L_2 in BZI scheme, respectively; and $\overline{I}_{Load}$ is the average value of load current.

2.2 Design of DVP Controller in BZI

From the state equation (2), the two transfer functions of $G_{id}(s)$ and $G_{vd}(s)$ can be determined. $G_{id}(s)$ is the ratio between the inductor current i_{L_1} and the

duty circle d. Meanwhile, $G_{vd}(s)$ is ratio between the capacitor voltage and the shoot-through duty cycle. They are formulated as

$$G_{id}(s) = \frac{\overline{V}.C.s + \left(2\overline{D} - 1\right).\overline{I}}{L.C.s^2 + C(r_L + R_c)s + \left(1 - -2\overline{D}\right)^2}, \tag{3}$$

where $\overline{V} = \overline{V}_{c_1} + \overline{V}_{c_2} - R_c.\overline{I}_{Load}$ and $\overline{I} = \overline{I}_{Load} - \overline{I}_{L_1} - \overline{I}_{L_2}$,

$$G_{vd}(s) = \frac{L.s.\overline{I} + \left[\left(1 - -2\overline{D}\right).\overline{V} + \overline{I}.(r_L + R_c)\right]}{L.C.s^2 + C(r_L + R_c)s + \left(1 - -2\overline{D}\right)^2}. \tag{4}$$

Similarly, the transfer function of inverter is determined as [2]

$$G_{inv}(s) = \frac{1}{T_{sf}s + 1}. \tag{5}$$

Figure 2 explains that the integration of the closed-loop of the current inductor i_{L_1} with the closed-loop of the DVP v_{dc}, namely the DVP closed-loop control system of the BZI (DC side).

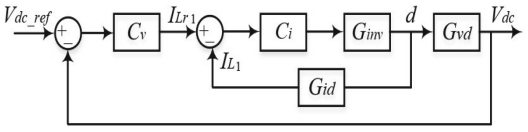

Fig. 2. The block diagram of [2].

In practice, the DCV is a pulse square; therefore, it cannot be controlled directly. In order to influence or control this DCV, its peak voltage (DVP) must be controlled instead [1]. In this paper, the "sisotool" of Matlab is used to simulate the DVP controller. The sisotool enables us to import the transfer functions of $G_{i2d}(s)$, $G_{inv}(s)$ and $G_{v2d}(s)$. Using the functions of "root locus editor for loop transfer" and "bode editor for loop transfer" provided by this tool, the controllers $C_i(s)$ and $C_v(s)$ are then designed based on the following two steps:

- Step 1: The transfer function of the controller $C_i(s)$ is designed so that the inductor current I_{L_1} tracks to the inductor current reference $I_{L_{ref_1}}$. Figure 2 and Fig. 3a) display the "root locus editor" for the loop transfer of $C_i(s)$ in which the poles (pink color) and the zeros are positioned so that the inductor current response I_{L_1} follows the reference signal. The closed-loop control system of the inductor current response I_{L_1} is stable as shown in Fig. 3b) and Fig. 3c). Then, the inductor current controller $C_i(s)$ is extracted from the "control system designer" tab.
- Step 2: Similar to the step 1, the transfer function of the DVP controller $C_v(s)$ is also designed using the "control system design" tab. The goal of this controller is to make the DVP response V_{dc} track its reference $V_{dc_{ref}}$ which is

also displayed in Fig. 2. Following these two steps, the controllers $C_i(s)$ and $C_v(s)$ have the corresponding forms:

$$\begin{cases} C_i\,(s) = K_{p_i} + \frac{K_{i_i}}{s} = 0.333 + \frac{5}{s} \\ C_v\,(s) = K_{p_v} + \frac{K_{i_v}}{s} = 0.445 + \frac{30}{s} \end{cases} \tag{6}$$

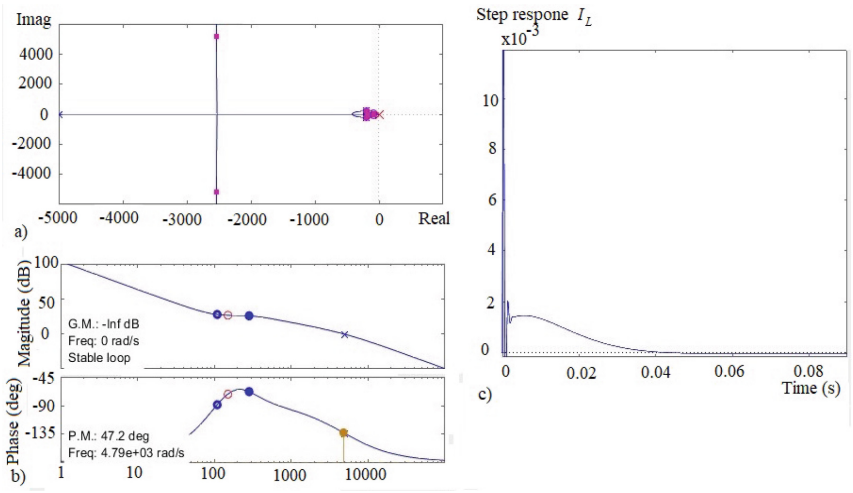

Fig. 3. Design controller $C_i\,(s)$, Root locus Editor a) Bode Editor b) Step response c).

2.3 PZI in Affine System and Augmented Feedforward Control law Designed

The system (1) is a nonlinear continuous-time system in a strict-feedback form [7]. In PMSM control system, the feedforward controller is designed to convert from optimal tracking control problem (1) in strict-feedback system into optimal tracking control problem in affine system. It is rewritten in terms of the error of speed PMSM and currents on dq-axis.

Define tracking error for speed PMSM and virtual control inputs of currents on *dq* axis. From (1), tracking error for speed PMSM is defined by difference of feedback signal and its reference signal. The currents on the dq-axis are considered as virtual control inputs of dq-axis voltages controller [1], which are shown as Fig. 1, the speed PMSM and the curents dq-axis tracking errors are defined as

$$\begin{cases} z_\omega = \omega - \omega_r, \\ z_{i_d} = i_d - i_d^r, \\ z_{i_q} = i_q - i_q^r, \end{cases} \tag{7}$$

where ω_r is the reference speed of PMSM; i_d^r, i_q^r are two reference curents on dq-axis, respectively; z_ω, z_{i_d}, z_{i_q} are the tracking errors which are the differences of the speed PMSM, two curents on dq-axis and its references. Using the mathematical model of the PMSM in (1), the tracking error of speed PMSM are arranged in the form of

$$\dot{z}_\omega = -\dot{\omega}_r - \frac{B}{J}.\omega - \frac{T_L}{J} + \frac{1.5n_p\Psi_f}{J}z_{i_q} + \frac{1.5n_p\Psi_f}{J}i_q^r + \frac{n_\tau}{J}. \tag{8}$$

Based on Fig. 1, the currents i_d^r and i_q^r are considered as the virtual control inputs of voltage on dq-axis. It includes two parts: the feedback online adaptive optimal control i_d^o, i_q^o and the augmented feedforward control inputs i_d^a, i_q^a on dq-axis, respectively [9]. These virtual control inputs are defined as

$$\begin{cases} i_d^r = i_d^o + i_d^a \\ i_q^r = i_q^o + i_q^a \end{cases} \tag{9}$$

These augmented feedforward control inputs are designed as

$$\begin{cases} i_d^a = 0, \\ i_q^a = \frac{J}{1.5n_p\Psi_f}\left(\dot{\omega}_r - k_{\omega_I}\int_0^t z_\omega(\tau)\,d\tau - k_{\omega_P}.z_\omega\right) - k_i z_{i_q}, \end{cases} \tag{10}$$

where $0 \le k_i \le 1$, $k_{\omega_I} \ge 0$, and $k_{\omega_P} \ge 0$ are design factors. Replacing (10) and (9) to (8), so the tracking error of the speed PMSM (8) is rewritten as

$$\begin{aligned} \dot{z}_\omega = -\dot{\omega}_r - \frac{B}{J}.\omega - \frac{T_L}{J} + \frac{1.5n_p\Psi_f}{J}z_{i_q} + \frac{1.5n_p\Psi_f}{J}i_q^o + \\ \frac{1.5n_p\Psi_f}{J}\left[\frac{J}{1.5n_p\Psi_f}\left(\dot{\omega}_r - k_{\omega_I}\int_0^t z_\omega(\tau)\,d\tau - k_{\omega_P}.z_\omega\right) - k_i z_{i_q}\right] + \frac{n_\tau}{J} \\ \Rightarrow \begin{array}{l} \dot{z}_\omega = -\frac{B}{J}.\omega - \frac{T_L}{J} + \frac{1.5n_p\Psi_f}{J}.(1-k_i)\,z_{i_q} - k_{\omega_P}.z_\omega \\ \qquad -k_{\omega_I}\int_0^t z_\omega(\tau)\,d\tau + \frac{1.5n_p\Psi_f}{J}i_q^o + \frac{n_\tau}{J} \end{array} \end{aligned} \tag{11}$$

Define tracking error of currents on on *dq* axis and virtual control inputs of of voltages on *dq* axis. The voltage control signals u_d and u_q on dq-axis of (1) are the total of feedback online adaptive optimal control signals (u_d^o, u_q^o) and augmented feedforward control signals (u_d^a, u_q^a) [6,7,13], these signals are shown in Fig. 1 and written as below

$$\begin{cases} u_d = u_d^o + u_d^a, \\ u_q = u_q^o + u_q^a, \end{cases} \tag{12}$$

where the augmented feedforward control signals u_d^a, u_q^a are designed as

$$\begin{cases} u_d^a & = L_d\dot{i}_d^r - k_{d_I}\int_0^t z_{i_d}(\tau)\,d\tau - k_{d_P}.z_{i_d}, \\ u_q^a & = L_q\dot{i}_q^r - \frac{1.5n_p\Psi_f}{J}(1-k_i)\,z_\omega, \\ & \quad -k_{q_I}\int_0^t z_{i_q}(\tau)\,d\tau - k_{q_I}.z_{i_q}, \end{cases} \tag{13}$$

where $k_{d_I} > 0$, $k_{d_P} > 0$, $k_{q_I} > 0$, and $k_{q_P} > 0$ are the design factors. Replacing (13) to (12) and substituting (12) to (1), and noted from (7), the tracking error of currents on dq-axis is rewritten as

$$\begin{cases} \dot{z}_{i_d} = -\frac{R_s}{L_d}.i_d + n_p.\omega.i_q \\ -\frac{1}{L_d}.\left(k_{d_i}.\int_0^t z_{i_d}(\tau)\,d\tau + k_{d_P}.z_{i_d}\right) + \frac{1}{L_d}.u_d^o + n_d \\ \dot{z}_{i_q} = -n_p.\omega.i_d - \frac{R_s}{L_q}.i_q - n_p.\frac{\Psi_f}{L_q}.\omega - \frac{1.5n_p\Psi_f}{J}.(1-k_i).z_\omega \\ -\frac{1}{L_q}.\left(k_{q_I}.\int_0^t z_{i_q}(\tau)\,d\tau + k_{q_p}.z_{i_q}\right) + \frac{1}{L_q}.u_q^o + n_q \end{cases} \tag{14}$$

Combine (8) and (14), the equations group are rewritten as

$$\begin{cases} \dot{z}_\omega = -\frac{B}{J}.\omega - \frac{T_L}{J} - k_{\omega_I}.\int_0^t z_\omega(\tau)\,d\tau - k_{\omega_P}.z_\omega \\ +\frac{1.5n_p\Psi_f}{J}.i_q^o + \frac{1}{J}.n_\tau \\ \dot{z}_{i_d} = -\frac{R_s}{L_d}.i_d + n_p.\omega.i_q \\ -\frac{1}{L_d}.\left(k_{d_i}.\int_0^t z_{i_d}(\tau)\,d\tau + k_{d_P}.z_{i_d}\right) + \frac{1}{L_d}.u_d^o + n_d \\ \dot{z}_{i_q} = -n_p.\omega.i_d - \frac{R_s}{L_d}.i_q - n_p.\frac{\Psi_f}{L_q}.\omega \\ -\frac{1}{L_q}\left(k_{q_I}.\int_0^t z_{i_q}(\tau)\,d\tau + k_{q_p}.z_{i_q}\right) + \frac{1}{L_q}.u_q^o + n_q \end{cases} \tag{15}$$

Theorem: The PZI scheme in strict-feedback form is converted into affine form, which is shown in (15). Together with (9), (11) and (12), where the augmented feedforward control inputs of the current q axis (10) and the augmented feedforward control inputs of the voltage q-axis are designed as (13). Here u^o is adaptive optimal control signals vector, i.e., $u^o = \left[i_q^o\ u_d^o\ u_q^o\right]^T$. Assume that u^o is designed to stabilize closed-loop control system of speed PMSM as shown in Fig. 1, n is the disturbances signals. This system is summarized as

$$\dot{z} = F(z) + G(z).u^o + K(z).n, \tag{16}$$

where,

$z = \left[z_\omega\ z_{id}\ z_{iq}\right]^T$, $F = \left[F_1\ F_2\ F_3\right]^T$,

$F_1 = -\frac{B}{J}.\omega - \frac{T_L}{J} - k_{\omega_I}.\int_0^t z_\omega(\tau)\,d\tau - k_{\omega_P}.z_\omega$,

$F_2 = -\frac{R_s}{L_d}.i_d + n_p.\omega.i_q - \frac{1}{L_d}\left(k_{d_i}.\int_0^t z_{i_d}(\tau)\,d\tau + k_{d_P}.z_{i_d}\right)$,

$F_3 = -n_p.\omega.i_d - \frac{R_s}{L_d}.i_q - n_p.\frac{\Psi_f}{L_q}.\omega - \frac{1}{L_q}.\left(k_{q_I}.\int_0^t z_{i_q}(\tau)\,d\tau + k_{q_p}.z_{i_q}\right)$,

$G = \left[G_1\ G_2\ G_3\right]^T$, $G_1 = \frac{1.5n_p\Psi_f}{J}$,$G_2 = \frac{1}{L_d}$, $G_3 = \frac{1}{L_q}$

$K = \begin{bmatrix} \frac{1}{J} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $n = \left[n_\tau\ n_d\ n_q\right]^T$,

where F_1 is unknown accurately, the above theorem can be proven that the online adaptive optimal tracking control problem as (1) of strict-feedback system is converted to that of an affine system as (16). This affine system must satisfy the requirements of Boundedness and Assumption defined as

Boundedness: Based on the physical properties nature of PZI system, there exist $||J|| \leq J_{max}$,$||G_i|| \leq ||G_{i_{max}}||$, $||F_i|| \leq ||F_{i_{max}}||$, $i = 1$ to 3, where J_{max}, $G_{i_{max}}$, $F_{i_{max}}$ are the unknown positive scalars, J is the performance index, and J_{max} is the performance index max.

Assumption: The reference speed of PMSM ω_r is bounded and smooth [9,12]. With unknown dynamics F, the adaptive optimal tracking control problem, which is presented in (1), is transformed to the adaptive optimal tracking control problem of (16), the adaptive optimal control signals u^o and the disturbances signals n. This result is proved in [7].

Proof: Lyapunov function is chosen for system (1) as

$$L_1 = \frac{1}{2} z_\omega^T . z_\omega + \frac{1}{2} z_{i_d}^T . z_{i_d} + \frac{1}{2} z_{i_q}^T . z_{i_q}. \tag{17}$$

The derivative of L_1 with respect to time is given by

$$\begin{aligned}
\dot{L}_1 &= z_\omega^T . \dot{z}_\omega + z_{i_d}^T . \dot{z}_{i_d} + z_{i_q}^T . \dot{z}_{i_q} + z_{v_c}^T . \dot{z}_{v_c} + z_{i_L}^T . \dot{z}_{i_L} \\
&= z_\omega^T . \left(-\frac{B}{J} . \omega - \frac{T_L}{J} - k_{\omega_I} \int_0^t z_\omega(\tau)\, d\tau - k_{\omega_P} . z_\omega \right) + \boldsymbol{z_\omega^T . \frac{1.5 n_p \Psi_f}{J} (1 - k_i) z_{i_q}} \\
&\quad + z_{i_d}^T . \left(-\frac{R_s}{L_d} i_d + n_p \omega i_q - \frac{1}{L_d} \left(k_{d_i} . \int_0^t z_{i_d}(\tau)\, d\tau + k_{d_P} . z_{i_d} \right) + \frac{1}{L_d} u_d^o + n_d \right) + \\
&\quad z_{i_q}^T . \left(-n_p \omega i_d - \frac{R_s}{L_d} i_q - n_p \frac{\Psi_f}{L_q} \omega - \frac{1}{L_q} \left(k_{q_I} . \int_0^t z_{i_q}(\tau)\, d\tau + k_{q_p} . z_{i_q} \right) + \frac{1}{L_q} u_q^o + n_q \right) \\
&\quad \boldsymbol{- z_{i_q}^T . \frac{1.5 n_p \Psi_f}{J} (1 - k_i) z_\omega}
\end{aligned} \tag{18}$$

We can see that $z_\omega^T = z_\omega$ and $z_{i_q}^T = z_{i_q}$and $z_\omega^T . \frac{1.5 n_p \Psi_f}{J} (1 - k_i) z_{i_q} = z_{i_q}^T . \frac{1.5 n_p \Psi_f}{J} (1 - k_i) z_\omega$ so the (18) is rewritten as

$$\dot{L}_1 = z^T . (F + G.u^o + K.n) = z^T . \dot{z}. \tag{19}$$

The derivative of Lyapunov function $\dot{L}_2$ is chosen for strict feedback system (16) with respect to time

$$\dot{L}_2 = z^T . \dot{z}. \tag{20}$$

Based on Lyapunov stable theory, if u^o makes the strict feedback system (16) stable then $\dot{L}_2$ must be negative. From (19) and (20), it can be seen that $\dot{L}_1 = \dot{L}_2$ so $\dot{L}_1 < 0$ and the tracking error of speed PMSM in (15) is UUB [7]. Therefore, the speed of PMSM control system is proposed by (1) which is converted to the adaptive optimal tracking control speed of PMSM as (16). This proof is completed.

3 Online Adaptive Optimal Tracking Control for the Speed of PMSM

In one restropective [13], online adaptive optimal tracking control problem can be considered as a zero-sum game theory with two players in which the adaptive optimal control signals u^o and the disturbances signals n are arguments (players) to minimize or maximize the value function. The structure of online adaptive optimal tracking control speed of PMSM is shown in Fig. 1. Based on [7], the online adaptive optimal control laws u^o of the PMSM system (16) are developed and implemented in the following section.

3.1 Online Adaptive Optimal Tracking Control Theory

As mentioned previously, the information about the internal dynamics F is difficult to accurately determine. However, the proposed control diagram can update the dynamics model F automatically; therefore, the procedure for determining the model F can be omitted. In this section, the design method of control laws for current on q-axis and control law of voltages on dq-axis are called control law u. The disturbance law n is explained in details. These control law and disturbance law are considered as two actors in RL technique, is shown in Fig. 1 [6,14,15].

From (16), the new system is rewritten as: $\dot{z} = F(z) + G(z).u + K(z).n$; $z = \left[z_\omega \; z_{id} \; z_{iq}\right]^T$; $u = \left[i_q \; u_d \; u_q\right]^T$; $n = \left[n_\tau \; n_d \; n_q\right]^T$. F and G are presented in (16), where $z \in R^{3\times 1}$, $F(z) \in R^{3\times 1}$, $G(z) \in R^{3\times 1}$, control laws $u \in R^{3\times 1}$, $K(z) \in R^{3\times 1}$ and disturbance $n \in R^{3\times 1}$. $F(z)$ is assumed Lipschitz locally; where $F(0) = 0$, $z = 0$ is the equilibrium point of the new system (16).

From [14], the performance index is defined as

$$J(z(0), u, n) = \int_0^\infty r(z, u, n)\, dt, \tag{21}$$

where $r(z, u, n)$ is the RL signals, $r(z, u, n) = Q(z) + u^T R u - \gamma^2 ||n||^2$. $Q(z) > 0$, $R > 0$, $R^T > 0$, $R = R^T$ and $0 \le \gamma^* \le \gamma$ where γ^* is the smallest and γ is selected to stabilize control system. From [14], the cost value function $V(z, u, n)$ of system (16) is called as a critic of the RL and it is defined as

$$V(z, u, n) = \int_t^\infty (r(z, u, n))\, d\tau. \tag{22}$$

In (22), $V(z, u, n)$ is finite, $V(0) = 0$. The derivative of both sides of (22) is a Bellman equation form,

$$\begin{gathered} r(z, u, n) + \frac{\partial V^T(z)}{\partial z}\dot{z} = 0 \\ \Longleftrightarrow r(z, u, n) + \nabla V^T(z).(F(z) + G(z).u + K(z).n) = 0 \end{gathered} \tag{23}$$

Based on theory control [6], the Hamiltonian function is defined as

$$H\left(z,u,n,\nabla V\left(z\right)\right)=r\left(z,u,n\right)+\nabla V^{T}\left(z\right).\left(F\left(z\right)+G\left(z\right).u+K\left(z\right).n\right) \quad (24)$$

Via the two-players zero-sum game studied in [6], the optimal value function is defined as

$$V^{o}\left(z\left(0\right)\right)=\underset{u}{min}\ \underset{n}{max}J\left(z\left(0\right),u,n\right). \quad (25)$$

From system (16), we can see that u is the first player to minimize $V(z,u,n)$ of and n is the second player to maximize $V(z,u,n)$. If this game theory exists saddle point [6], the adaptive optimal control problem exist one unique solution [6,14]. The HJI equation is solved to determine the optimal value function $V^{o}\left(z\left(0\right)\right)$.

$$r\left(z,u^{o},n^{o}\right)+\nabla V^{oT}\left(z\right).\left(F+G.u^{o}+K.n^{o}\right)=0 \quad (26)$$

With the stationary condition, $\frac{\partial H(z,u,n,\nabla V(z))}{\partial u}=0$ and $\frac{\partial H(z,u,n,\nabla V(z))}{\partial n}=0$, the saturated adaptive optimal control law u^{o} and the disturbance law n^{o} are determined as

$$u^{o}=\underset{u}{argmin}H\left(z,u,n^{o},\nabla V^{o}\left(z\right)\right)=-\frac{1}{2}R^{-1}.G\left(z\right)^{T}.\nabla V^{o}\left(z\right) \quad (27)$$

$$n^{o}=\underset{n}{argmin}H\left(z,u^{o},n,\nabla V^{o}\left(z\right)\right)=\frac{1}{2\gamma^{2}}K^{T}.\nabla V^{o}\left(z\right) \quad (28)$$

3.2 Design Control Laws and Disturbance Laws of Online Adaptive Optimal Tracking by Approximate Solution of HJI

The HJI equation is very difficult to solve using any analytical method. To solve this equation, RL techniques are used to approximate the equation online and find the optimal solutions. This paper proposes an approximation that includes one critic and two actors. The critic is the cost value function $V(z)$; the first actor is an online adaptive optimal control law u and the second actor is the disturbance law n. The cost value function $V(z)$ is approximated as

$$V(z)=W_{c}^{T}.\Theta_{c}\left(z\right)+\epsilon\left(z\right), \quad (29)$$

where $\Theta_{c}(z)$ is the active function of neuron network, $\epsilon\left(z\right)$ is errors of cost value function and W_{c} is critic weights. The $\hat{W}_{c}$ is the current approximation weights, the current critic $\hat{V}(z)$ is defined as

$$\hat{V}(z)=\hat{W}_{c}^{T}.\Theta_{c}\left(z\right) \quad (30)$$

Replace (30) to (26), the HJI equation approximation is rewritten as

$$\begin{aligned}\hat{H}\left(z,u,n,\nabla V\left(z\right)\right)=\hat{W}_{c}^{T}.\nabla\Theta_{c}.\left(F\left(z\right)+G\left(z\right).\hat{u}+K\left(z\right).\hat{n}\right)+Q\left(z\right)\\+\hat{u}^{T}.R.u-\gamma^{2}.||n||^{2}=e_{1},\end{aligned} \quad (31)$$

where e_1 is the approximation error of the current HJI equation, the current approximation of the adaptive optimal control laws $\hat{u}$ and disturbances $\hat{n}$ are determined as

$$\hat{u} = -\frac{1}{2\rho} R^{-1}.G\left(z\right)^{T}.\frac{\partial \Theta_c}{\partial z}\left(z\right)^{T}.\hat{W}_u, \forall z, \tag{32}$$

$$\hat{n} = \frac{1}{2\gamma^2}.K^{T}.\frac{\partial \Theta_c}{\partial z}\left(z\right)^{T}.\hat{W}_n, \forall n, \tag{33}$$

where $\hat{u} = \left[\hat{i}_q\ \hat{u}_d\ \hat{u}_q\right]^T$; $\hat{n} = \left[\hat{n}_\tau\ \hat{n}_d\ \hat{n}_q\right]^T$

Now, it is desired to select $\hat{W}_c$ such that the squared residual error $E_1 = \frac{1}{2}e_1^T e_1$ is minimized. From [14], derivative of current critic weights $\hat{W}_c$ is given as

$$\begin{aligned}\dot{\hat{W}}_c &= -\alpha_c.\frac{\partial E_1}{\partial \hat{W}_c} \\ &= -\alpha_c.m.\left[\sigma.\hat{W}_c + Q\left(z\right) + \hat{u}^T.R.u - \gamma^2.||\hat{n}||^2\right],\end{aligned} \tag{34}$$

where α_c is the convergence gain of critic; $\overline{\sigma} = \sigma\left(\sigma^T.\sigma + 1\right)$;

$\sigma = \nabla\Theta_c.\left(F\left(z\right) + G\left(z\right).\hat{u} + K\left(z\right).\hat{n}\right)$; $m = \frac{\sigma}{\left(\sigma^T.\sigma+1\right)^2}$. The current weights $\hat{W}_u$ of the actor $\hat{u}$, This current weights $\hat{W}_u$ must be chosen to stabilize closed-loop control system. So, derivative of the current optimal control weights $\hat{W}_u$ is obtained as

$$\begin{aligned}\dot{\hat{W}}_u = -\alpha_u.\{&\left[\hat{W}_u - \overline{\sigma}^T.\hat{W}_c\right] \\ -\frac{1}{4}&\left[\nabla\Theta_c.G\left(z\right).R^{-1}.G^T\left(z\right).\nabla\Theta_c^T\right].\hat{W}_u.m^T.\hat{W}_c\},\end{aligned} \tag{35}$$

where α_u is the convergence gain of actor $\hat{u}$. Derivative of the current disturbance weights $\hat{W}_n$ are obtained as

$$\dot{\hat{W}}_n = -\alpha_n.\left\{\left(\hat{W}_n - \overline{\sigma}^T.\hat{W}_c\right) + \frac{1}{4\gamma^2}.\left(\nabla\Theta_c.K.K^T.\nabla\Theta_c^T\right).\hat{W}_n.m^T.\hat{W}_c\right\} \tag{36}$$

Similarly, α_n is the convergence gain of actor $\hat{n}$. Using the Lyapunov's theory and UUB, this problem converge is given in [6,14] to stabilize closed-loop control system of the speed PMSM.

Algorithm adaptive optimal tracking control Start with initial

$\hat{W}_c \leftarrow 0$, $\hat{W}_u \leftarrow 0$, $\hat{W}_n \leftarrow 0$, $\hat{u} \leftarrow 0$,$\hat{n} \leftarrow 0$, $\hat{V} \leftarrow 0$, $\epsilon_c = \epsilon_u = \epsilon_n = 0.0001$,

Chosen Θ_c; $\alpha_u > 0$,$\alpha_n > 0$, $\gamma > 0$, $Q > 0$, $R > 0$, $l \leftarrow 0$, Stop step: l_{max}

Repeat
$\hat{V}^{(l)} \leftarrow \hat{W}_c^{(l)T}.\Theta_c$
$\hat{u}^{(l)} \leftarrow -\frac{1}{2}R^{-1}.G\left(z\right)^T.\nabla\Theta_c.\left(z\right)^T.\hat{W}_u^{(l)}$
$\hat{n}^{(l)} \leftarrow \frac{1}{2\gamma^2}.K^T.\nabla\Theta_c.\left(z\right)^T.\hat{W}_n^{(l)}$
$\hat{u}^{(l)} \leftarrow \hat{u}^{(l)} + u^a$
$z \leftarrow$Update $\hat{u}^{(l)}$ and $\hat{n}^{(l)}$ to system (16)
$\hat{W}_c^{(l+1)} \leftarrow$ Update weights of value function (34)
$\hat{W}_u^{(l+1)} \leftarrow$ Update weights of optimal control laws (35)
$\hat{W}_n^{(l+1)} \leftarrow$ Update weights of optimal control laws (36)
$l \leftarrow l+1$
Until $(l > l_{max})$ or $\hat{W}_c^{(l)} - \hat{W}_c^{(l-1)} < \varepsilon_c$ and $\hat{W}_u^{(l)} - \hat{W}_u^{(l-1)} < \epsilon_c$ and $\hat{W}_n^{(l)} - \hat{W}_n^{(l-1)} < \epsilon_c$
End

4 Simulation of PZI Control System

4.1 The DPV of BZI Scheme Simulated Results

If the PZI system is to operate with a high performance, the DC-link supply inverter must be stabilized firstly. Therefore, the speed controller design of PMSM and DVP must be synchronized with each other. The integrating structure of the control systems is shown in Fig. 1. The parameters of BZI are used to simulate the DVP [1] as: $\overline{I}_l$ is average value of load current, $\overline{I}_l = 0.8A$; f_{sf} is switching frequency, $f_{sf} = 5kHz$; $\overline{I}_{L_1}$and $\overline{I}_{L_2}$ are average value of inductor currents through L_1 and L_2, respectively, $\overline{I}_{L_1}$=$\overline{I}_{L_2}$= 0.6A. $\overline{D}$ is average value of duty cycle, $\overline{D}$= 0.4; V_{in} is input voltage, $V_{in} = 40V$.L_1 and L_2 are two inductors of BZI, respectively, $L_1 = L_2 = L = 0.6 * 10^{-4}H$. C_1 and C_2 are two capacitors of BZI, respectively, $C_1 = C_2 = C = 0.6 * 10^{-4}F$. L_l is stator inductance of PMSM, $L_l\left(L_l = L_d = L_q\right) = 8.5mH$; R_l is stator resistance of PMSM, $R_l\left(R_l = R_s\right) = 2.7\Omega$.

The parameters of DVP, currrent controller gains are designed as above part 2.1: K_{pv} is the DVP controller proportional gain, $K_{pv} = 0.445$; K_{iv} is is the DVP controller integral gain, $K_{iv} = 30$; K_{pi} is the current controller proportional gain, $K_{pi} = 0.33$; K_{ii} is the current controller integral gain, $K_{ii} = 5$.

The DC voltage input (V_{in}) is used to supply voltage for the PZI system. Hence, the DVP must be controlled to boost or buck voltage to the system requirements under the disturbances and the changes of the load torque and the speed of the PMSM. If V_{in} is dropped (or load torque increases), the DCV must be controlled to increase its value, avoid over-modulation in MSVM [1]. As shown in Fig. 4, the DCV has its peak value as the total voltages of the two capacitors. During the time from 2.0001 s to 2.002 s, the DVP was dropped according to the decrease of V_{in} at time 2 s and 3 s as seen in Fig. 5a).

The V_{in} always decreases to time, it is shown as in Fig. 5. In period of time, from $0s$ to $2s$, the V_{in} has a value voltage is 40V, but at $2s$, the voltage drop

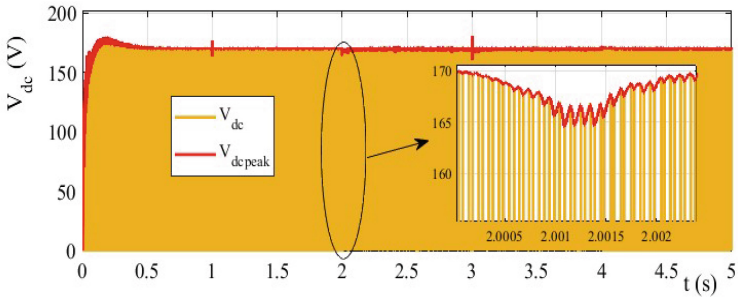

Fig. 4. The DCV V_{dc} and the DCV peak V_{dcpeak}

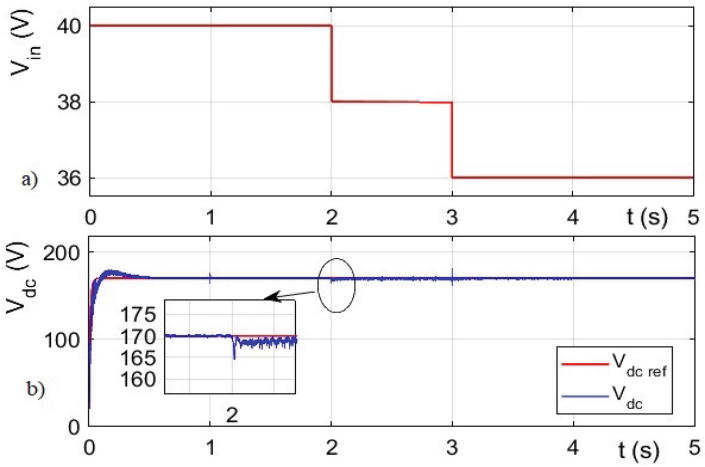

Fig. 5. a) The V_{in} b) The DVP V_{dc} ; the DVP reference $V_{dc_{ref}}$

of V_{in} is 38V. Continuously, from $2s$ to $5s$, the voltage drop of V_{in} is 36V that is shown in Fig. 5 a). The $C_i\,(s)$ and $C_v\,(s)$ controllers were designed by the frequency-response method as part 2.1 of this paper, the DVP follows always to the DVP reference 170V in Fig. 5 b). The DVP is boosted from 40V to 170V as in Fig. 5 b) and it is controlled to stabilize at 170V. If the V_{in} changes, the DVP still remains to stabilize at 170V by the DCV closed-loop control system, is shown in Fig. 2.

4.2 Simulation Results for the Speed of PMSM

The online adaptive optimal tracking control for the speed of PMSM is built in Matlab simulink. The parameters of OTC controller are chosen as: α_c is convergence gain of weights $\hat{W}_c$, $\alpha_c = 0.1$; α_u is convergence gain of weights $\hat{W}_u$, $\alpha_u = 0.2$; α_n is convergence gain of weights $\hat{W}_n$, $\alpha_n = 0.15$; γ is the coefficient of performance index J, γ= 5; ϵ_c,ϵ_u, ϵ_n are the stop gains of error weights

$\hat{W}_c, \hat{W}_u, \hat{W}_n$, respectively, $\epsilon_c = \epsilon_u = \epsilon_n = 1e^{-4}$; R is the coefficient of control laws $\hat{u}$, $R = 1$; ρ is the coefficient of OTC, $\rho = 0.8$.

The nominal parameters of PMSM drives are given as: rated power P_n, (P_n=400W); rated voltage U_n, (U_n=200V); rated speed ω, ($\omega = 3000rpm$); rate torque T_L, (T_L= 1.27$N.m$); permanent magnet flux Ψ_f, (Ψ_f =0.0615Wb); stator resistor R_s, ($R_s = 2.7\Omega$); stator inductor L_d, L_q, ($L_d = L_q = 8.5mH$); moment of inertia J, ($J = 31.69 \times 10^{-6} kg.m^2$); viscous friction coefficient B, ($B = 52.79 \times 10^{-6} Ns.m^{-2}$); pole pair n_p, ($n_p = 4pairs$).

$$\Theta_c\left(z_\omega, z_{i_d}, z_{i_d}\right) = \left[\, z_\omega^2 \;\; z_\omega.z_{i_d} \;\; z_\omega.z_{i_q} \;\; z_{i_d}^2 \;\; z_{i_d}.z_{i_q} \;\; z_{i_q}^2 \,\right]^T;$$

$$\frac{\partial \Theta_c}{\partial z} = \nabla \Theta_c\left(z_\omega, z_{i_d}, z_{i_d}\right) = \left[\, \frac{\partial \Theta_c}{\partial z_\omega} \;; \frac{\partial \Theta_c}{\partial z}; \; \frac{\partial \Theta_c}{\partial z} \,\right]$$

$$\hat{W}_u = \left[\, \hat{W}_{u_{iq}} \; \hat{W}_{u_{u_d}} \; \hat{W}_{u_{u_q}} \,\right]^T; \; \hat{W}_c = \left[\, \hat{W}_{iq} \; \hat{W}_{u_d} \; \hat{W}_{u_q} \,\right]^T;$$

$$\hat{W}_u = \left[\, \hat{W}_{u_{iq}} \; \hat{W}_{u_{u_d}} \; \hat{W}_{u_{u_q}} \,\right]^T$$

The load torque T_L is varied and its value is shown in the Fig. 6. In period of time, from 0s to 0.25s, the load torque is 0$N.m$. Next, from 0.25s to 1s, the load torque is 1.27/4N.m. From 1s to 2.5s, the load torque is 1.27/3N.m. Continuously, from 2.5s onwards the load torque is 1.27$N.m$.

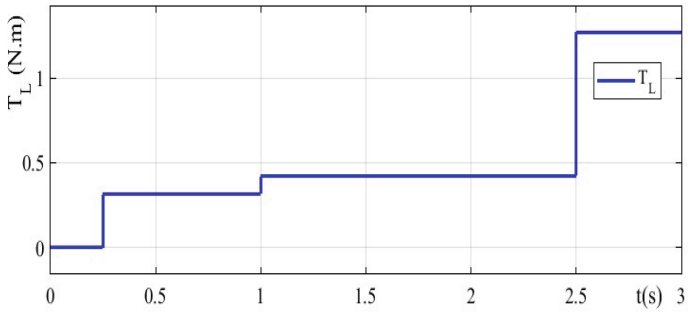

Fig. 6. Load torque (T_L) changes to time

From Fig. 7, at time 1 s and the load torque increased from 1.27/4 N.m to 1.27/3 N.m, the resulting speed of the PMSM from the OTC controller was able to track its reference robustly. Meanwhile, the result from the PI controller was unable to follow the input reference signal.

The results from Fig. 7 proved that the OTC controller outperforms the traditional PI controller in term of robustness. The control inputs are the voltage on dq-axis, u_d and u_q is shown in Fig. 1, $u_d \approx 0$, u_d tracks zeros, u_q changes to variations of speed PMSM, load torque and DVP in Fig. 8.

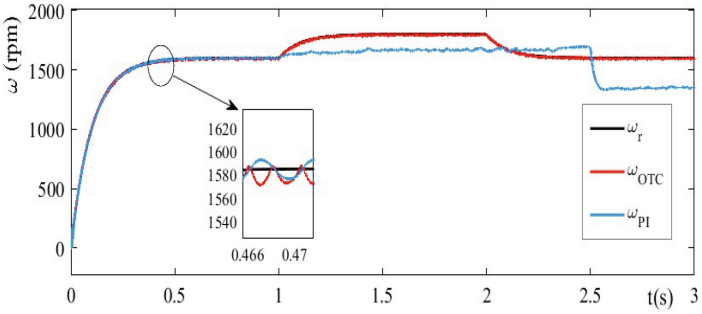

Fig. 7. The speed of PMSM response with PI, OTC

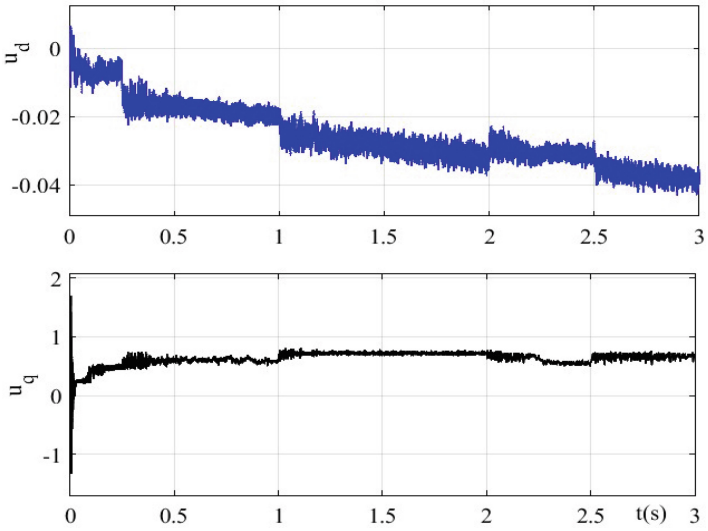

Fig. 8. The control inputs $u_{dq}-$axis with OTC controller

5 Conclusions

In this paper, we have proposed a control strategy with three joint controllers to improve the speed response of a permanent magnet synchronous motor (PMSM) in an electric vehicle (EV) system. In case of controlling speed of the PMSM, the simulation results show that the derived OTC control law is capable of handling the unknown dynamics of the PZI system. The control law is determined based on the solutions of the HJI equations for an approximated minimal cost function by the RL method. Utilizing the OTC controller in the simulation, the speed of the PMSM is able to robustly track the reference signal. Moreover, the corresponding tracking error from this controller is smaller than the error when using the traditional PI controller. In case of controlling the DVP, a frequency-response method is used to develop two controllers of the inductor current and

the DVP to form the close-loop control system of the BZI. The results indicate that the controlled DVP managed to follow its reference despite the variation from the DC input voltage and the speed of the PMSM. In summary, this paper proposes an OTC controller to handle the speed of the PMSM and a frequency-response method to design of two controllers in the BZI scheme. The operations of these controllers for the closed-loop PZI system show the effective results in the presence of noise and disturbances.

Acknowledgment. The authors would like to thank the Vietnam Aviation Academy for Science and Technology Development for the support in 2023–2024.

References

1. Pham, C.T., Huu, C.T.N, Tran, Q.K., Van Thien, T., Nguyen, D.T.H: Adaptive Backstepping sliding mode control for speed of PMSM and DC-link voltage in bidirectional quasi Z-source inverter. In: International Conference on Industrial Networks and Intelligent Systems, pp. 185–202 (2023). https://doi.org/10.1007/978-3-031-47359-3_14
2. Na, T., Zhang, Q., Zhou, C.: Modeling and design of quasi-Z-source inverter for PMSM drive system. In: 2016 UKACC 11th International Conference on Control (CONTROL), pp. 1–5 (2016)
3. Kim, S.K.: Robust adaptive speed regulator with self-tuning law for surfaced-mounted permanent magnet synchronous motor. Control Eng. Pract. **61**, 55–71 (2017)
4. Shen, A.W., Pham, C.T., Dzung, P.Q., Anh, N.B., Viet, L.H.: Using fuzzy logic self-tuning pi controller in z-source inverter for hybrid electric vehicles. In: 2012 World Conference on Science and Engineering, Hong Kong, China, August (2012)
5. Nguyen, T.N.A., Pham, D.C., Pham, C.T., Thanh, N.H.C.: D-axis stator current control methods applied to PMSG-based wind energy systems: a comparative study. WSEAS Trans. syst. Control. **14**, 239-239 (2019)
6. Vamvoudakis, K.G., Kokolakis, N.-M.T.: Synchronous reinforcement learning-based control for cognitive autonomy. Found. TrendsR Syst. Control **8**(1–2), 1–175 (2020)
7. Tan, L.N., Pham, T.C.: Optimal tracking control for PMSM with partially unknown dynamics, saturation voltages, torque, and voltage disturbances. IEEE Trans. Industr. Electron. **69**(4), 3481–3491 (2022). https://doi.org/10.1109/TIE.2021.3075892
8. Ayad, A., Karamanakos, P., Kennel, R., Rodríguez, J.: Direct Model Predictive Control of Bidirectional Quasi-Z-Source Inverters Fed PMSM Drives. In: 2017 11th IEEE International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG), pp. 671–676 (2017)
9. Tan, L.N., Cong, T.P., Cong, D.P.: Neural network observers and sensorless robust optimal control for partially unknown PMSM with disturbances and saturating voltages. IEEE Trans. Power Electron. **36**(10), 12045–12056 (2021). https://doi.org/10.1109/TPEL.2021.3071465
10. Ayad, A., Karamanakos, P., Kennel, R.: Direct Model Predictive Current Control Strategy of Quasi-Z-Source Inverters. IEEE Trans. Power Electron. **32**(7), 5786–5801 (2017)

11. Ayad, A.N.F.: Advanced Control Techniques of Impedance Source Inverters for Distributed Generation Applications. PhD project has been done at the Chair of Electrical Drive Systems and Power Electronics at Technischen Universität München (TUM), Munich, Germany (2017)
12. Zargarzadeh, H., Dierks, T., Jagannathan, S.: Optimal control of nonlinear continuous-time systems in strict-feedback form. IEEE Trans. Neural Netw. Learn. Syst. **26**(10), 2535–2549 (2015)
13. Kiumarsi, B., Vamvoudakis, K.G., Modares, H., Lewis, F.L.: Optimal and autonomous control using reinforcement learning: a Survey. IEEE Trans. Neural Netw. Learn. Syst. **29**(6), 2042–2062 (2018)
14. Vamvoudakis, K.G., Lewis, F.L.: Online solution of nonlinear two-player zero-sum games using synchronous policy iteration. In: 49th IEEE Conference on Decision and Control (CDC), Atlanta, GA, USA, 2010, pp. 3040-3047 (2010). https://doi.org/10.1109/CDC.2010.5717607.
15. Xu, Z., Kontoudis, G.P., Vamvoudakis K.G.: Online and robust intermittent motion planning in dynamic and changing environments. IEEE Trans. Neural Netw. Learn. Syst. (2023). https://doi.org/10.1109/TNNLS.2023.3303811.