



Article

Machine Learning Structure for Controlling the Speed of Variable Reluctance Motor via Transitioning Policy Iteration Algorithm

Hamad Alharkan

Department of Electrical Engineering, College of Engineering, Qassim University, Buraydah 52571, Saudi Arabia; h.alharkan@qu.edu.sa

Abstract: This paper investigated a new speed regulator using an adaptive transitioning policy iteration learning technique for the variable reluctance motor (VRM) drive. A transitioning strategy is used in this unique scheme to handle the nonlinear behavior of the VRM by using a series of learning centers, each of which is an individual local learning controller at linear operational location that grows throughout the system's nonlinear domain. This improved control technique based on an adaptive dynamic programming algorithm is developed to derive the prime solution of the infinite horizon linear quadratic tracker (LQT) issue for an unidentified dynamical configuration with a VRM drive. By formulating a policy iteration algorithm for VRM applications, the speed of the motor shows inside the machine model, and therefore the local centers are directly affected by the speed. Hence, when the speed of the rotor changes, the parameters of the local centers grid would be updated and tuned. Additionally, a multivariate transition algorithm has been adopted to provide a seamless transition between the Q-centers. Finally, simulation and experimental results are presented to confirm the suggested control scheme's efficacy.

Keywords: machine learning method; variable reluctance motor; optimization problems; adaptive control systems



Citation: Alharkan, H. Machine Learning Structure for Controlling the Speed of Variable Reluctance Motor via Transitioning Policy Iteration Algorithm. *World Electr. Veh. J.* **2024**, *15*, 421. <https://doi.org/10.3390/wevj15090421>

Academic Editor: Joeri Van Mierlo

Received: 19 July 2024

Revised: 9 September 2024

Accepted: 12 September 2024

Published: 14 September 2024



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

1. Introduction

Variable reluctance motors (VRMs) are adaptable machines that can be utilized in various industrial applications due to their inherent resilience, simplicity, competitive price, and high efficiency. VRMs are currently used in high-performance applications, including hybrid electric vehicles (HEVs) and aircraft systems [1–4]. The most notable disadvantage of VRMs is their extreme nonlinearity, which creates a significant ripple in amperage [5,6]. This factor could contribute to vibration and acoustic noise during machine operation. Therefore, decreasing ripples is necessary to improve machine performance.

Two main strategies are used to mitigate current ripples: machine design optimization and dynamic current controller adoption to minimize current ripples without employing a high switching frequency. Such a current control method should generate an optimal phase voltage to achieve a rapid current response while preserving low current ripple [7]. Variations within the machine inductance cause back EMF, necessitating a high DC voltage from the voltage source. However, this can create substantial current ripples in traditional frequency-bounded switching methods, including delta-modulation and PI-based controllers. VRM control techniques can be classified into two categories. The first category comprises simple, standard techniques like hysteresis control, delta modulation, sliding mode methods, and fast PID controllers. The second category uses optimal techniques, including artificial neural networks, model predictive control (MPC), and neuro-fuzzy controller (NFC), to generate a duty cycle based on the back EMF of the machine [7–14]. There are multiple sophisticated approaches to appraise the inductance contour and achieve local linearization of the system, either through interpolation or advanced learning techniques.

One notable method involves applying Fourier series expansion to the inductance contour. This technique appraises the rotor positions: aligned, centrally located, and misaligned by determining the connection between the Fourier inductance profile factors and the rotor position angle. The Fourier analysis may be conducted using finite element analysis (FEA), providing a detailed mathematical representation of the inductance variation with rotor position and current. Another method for estimating the inductance contour entails enabling the controller to iteratively approximate the parameters of the model and monitor errors during each iteration. This adaptive approach requires the controller to actively learn the system's behavior, refining its model based on real-time feedback and error correction. Such an estimator is essential in updating the model, ensuring an accurate representation of the system's dynamics. This updated model can subsequently be employed in a model predictive controller (MPC), which is crucial for achieving precise tracking control. The policy iteration algorithm, a reinforcement learning technique, combines the adaptation and control tasks into a single cohesive framework. It enables the system to learn optimal control policies through interaction with the environment, making it well-suited for systems with nonlinearities. To leverage policy iteration while addressing the complexities of nonlinear systems, a scheduled local learning approach is proposed. This involves partitioning the state-space into manageable regions and applying policy iteration within each region to achieve effective current tracking control for VRMs. This method ensures robust performance across varying operating conditions, providing a comprehensive solution for current control in VRMs. The Q-learning transitioning controller is among the adaptive optimal control methods [15–17]. It employs reinforcement learning as a form of machine learning to tackle tracking problems. The infinite-horizon linear quadratic tracker (LQT) method was incorporated into the Q-learning algorithm to produce a controller that tracks the reference path. Since Q-learning only applies to linear systems, it cannot be applied directly to a nonlinear model, including a VRM. To resolve this limitation [16], the controller was incorporated into a transitioning strategy to enable nonlinear VRM control by transitioning linear Q controllers over the nonlinear VRM domain. A family of Q-cores is introduced on the system domain, each positioned at a local linear operational point where the Q-learning algorithm can be executed. Eventually, a Q-matrix grid should be trained for each iteration using data observed along the system trajectory to achieve tracking performance. The system can use locally affine Q-learning controllers to control a nonlinear system by traversing a path on this Q-grid.

This paper describes a speed regulator by transitioning policy iteration along with tridimensional Q-grid for a VRM. Unfortunately, using a bidimensional lookup grid does not include motor speed variations within the planned domain; thus, when a rotational speed is changed by adding mechanical load torque, the local nodes employed in the bidimensional Q-grid must be retrained to accommodate the speed change. Despite linear speed variations and Q-learning remaining stable throughout this adaptive process, a slow speed change response is observed due to this learning process. In other learning approaches [7], model-predictive control only learned the inductance profile. It did not require re-learning in response to the speed change since the speed can be fed into the model externally. This paper proposes a new speed regulator by presenting a tridimensional array containing rotor position, speed, and phase current, enabling the controller to train the Q parameters for any given speed. Additionally, a tridimensional interpolation mechanism is shown to manage the controller transitions over this tridimensional grid. Section 2 describes the Q-learning algorithm and illustrates the speed regulator of the VRM. Sections 3 and 4 demonstrate the speed regulator by offering simulation and experimental results.

2. Materials and Methods

Figure 1 depicts the suggested controller's principal layout, where an adaptive online policy iteration structure of variable reluctance motor regulates the speed level of variable switched reluctance motor. The next part goes into further detail on the internal structure of the introduced model.

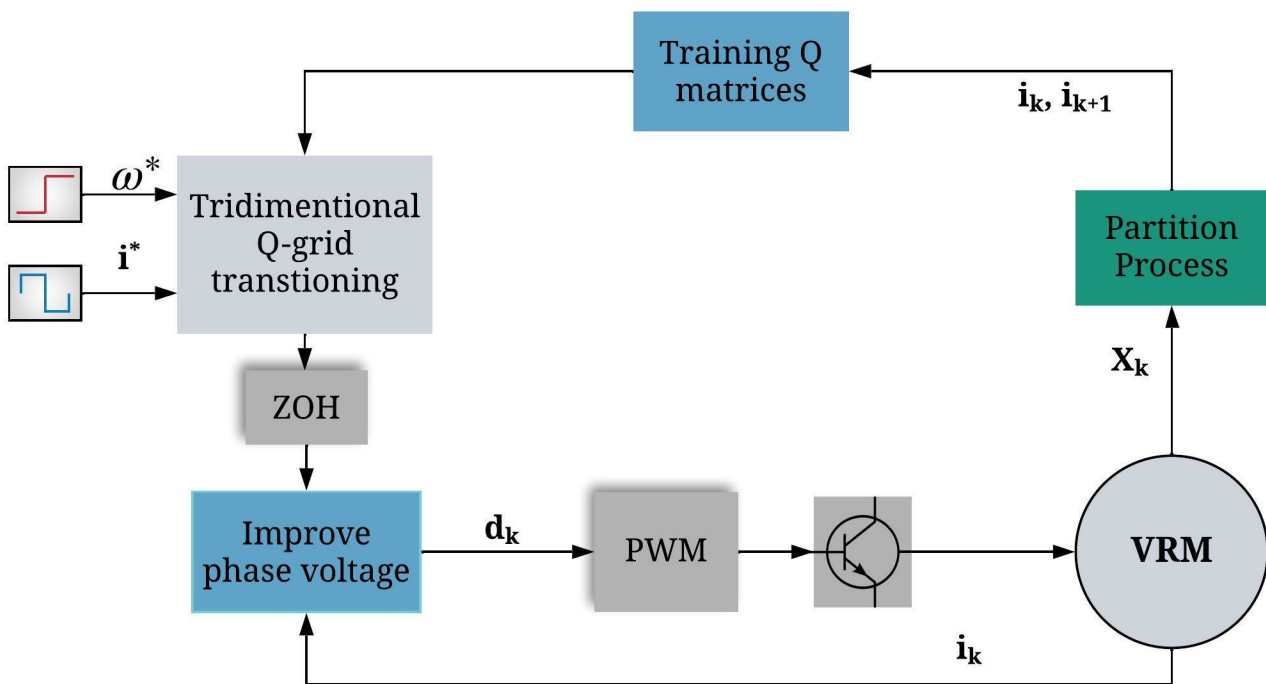


Figure 1. Block diagram of a tridimensional Q-grid learning control. ω^* and i^* represents the optimal trajectory for the speed and the current respectively.

2.1. Reinforcement Tracking Structure of VRM

The linear quadratic tracker (LQT) has become an increasingly vital design tool for tracking controls. Obtaining the optimal LQT solution enables the tracking of a predetermined reference signal by minimizing the cost function and the difference between the reference and output currents [18]. Typically, LQT solutions are derived by independently acquiring the solutions for the feedforward and feedback sections. The primary disadvantage of this solution is that it is computed offline in conjunction with the model’s parameters [19]. In this context, the Reinforcement Q-learning technique, which is a type of adaptive dynamic programming, provides an online solution to LQT problems without using VRM model information. This section presents the derivation of the LQT-augmented system. Additionally, the optimal solutions for LQT using the Bellman equation and Q-learning are included.

2.1.1. Cost Function Formulation of VRM

The primary purpose of VRM tracking control is to identify the optimal phase voltage to enable the VRM drive current to follow the trajectory of the reference current. To reduce the controller’s computing load and cost, the VRM model is developed without considering mutual coupling between phases. In order to develop such a controller, an initial assumption was made for the LQT problem involving the generation of a reference current from an independent generator model, as shown in [20]:

$$r_{k+1} = Fr_k \tag{1}$$

where r_k donates to the ideal current track, and F is the ideal current source that capable to produce a smooth train of pulses, which is the desired waveform for the VRM drive current. The generator model is integrated into the VRM model. The DC voltage supplied to the coil in the per-phase model combines the resistive voltage drop and the flux linkage across the coils over time. Therefore, the augmented VRM drive model can be formulated as follows:

$$X_{k+1} = \begin{bmatrix} A & 0 \\ 0 & F \end{bmatrix} \begin{bmatrix} x_k \\ r_k \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u_k \equiv A_a X_k + B_b u_k$$

$$Y_k = [C \ 0]x_k \equiv C_c X_k \quad (2)$$

where $A = 1 - TR/L_k$, $B = T/L_k$, x_k is the current injected each phase, u_k is the control input per phase, R is the internal resistance of the winding, and Y_k is the output current per phase. L_k is the inductance that varies regarding the stator current injected and rotor angle. T indicates to the switching time sample. Similar to the conventional performance index of LQR, the developed augmented model's cost function can be generated as follows:

$$V(x_k) = \frac{1}{2} \sum_{i=k}^{\infty} \gamma^{i-k} [X_i^T Q_q X_i + u_i^T R u_i] \quad (3)$$

where $Q_q = [C \ -I]^T Q [C \ -I]$ and Q and R are weight matrices that have been specified for the stator winding current and the applied DC voltage of the VRM drive, respectively, whereas $0 < \gamma \leq 1$ donates to a discount factor. For a fixed control input, the cost function may be rewritten in a quadratic style for a symmetric P matrix as $V(x_k) = \frac{1}{2} X_k^T P X_k$. The complete quadratic form derivation of the LQT cost function has been provided in [21]. By adopting a fixed policy, the infinite sum in Equation (3) of the augmented system can be expressed as:

$$V(X_k) = \frac{1}{2} \left[(X_k)^T Q_q (X_k) + u_k^T R u_k \right] + \gamma V(X_{k+1}) \quad (4)$$

2.1.2. Reinforcement of Policy Iteration Structure for Solving the Problem

Before designing the algorithm for addressing LQT problems, it is essential to derive the Bellman equation of the LQT. The Bellman equation can be used to solve problems involving voltage stabilization. In order to accomplish this, the policy iteration is applied to the LQT Lyapunov equation. However, this approach requires the system model's complete information [22]. The reference for the Bellman equation of LQT, which is dependent on the augmented model, is as follows:

$$X_k^T P X_k = x_i^T Q_q x_i + u_k^T R u_k + \gamma X_{k+1}^T P X_{k+1} \quad (5)$$

where P is the Algebraic Riccati equation solution, which fulfilled the optimal solution of LQT. By satisfying the stationary condition and utilizing the Hamiltonian equation, the tracking problem's optimal P matrix can be derived as follows:

$$P = Q_q + \gamma A_a^T P A_a - \gamma^2 A_a^T P B_b (R + \gamma B_b^T P B_b)^{-1} B_b^T P A_a \quad (6)$$

At this stage, the Bellman equation employing the policy iteration approach and the existence of the machine's model parameters can be implemented so that the numbers of the P matrix converge to their optimal values. On the other hand, training the Q function that comprises the cost function and the augmented system in addition to the reference current model yields impressive results. Moreover, it solves the ARE online without needing knowledge of the machine's model [23]. The Q -function might be formulated as a matrix using substitutions and definitions as follows:

$$Q(X_k, u_k) = \frac{1}{2} \begin{bmatrix} X_k \\ u_k \end{bmatrix}^T \begin{bmatrix} Q_q + \gamma A_a^T P A_a & \gamma A_a^T P B_b \\ \gamma B_b^T P A_a & R + \gamma B_b^T P B_b \end{bmatrix} \begin{bmatrix} X_k \\ u_k \end{bmatrix} \quad (7)$$

Also, this could be defined as

$$Q(X_k, u_k) = \frac{1}{2} \begin{bmatrix} X_k \\ u_k \end{bmatrix}^T \begin{bmatrix} G_{XX} & G_{Xu} \\ G_{uX} & G_{uu} \end{bmatrix} \begin{bmatrix} X_k \\ u_k \end{bmatrix} \quad (8)$$

The Q-grid algorithm can be trained via either policy or value iteration RL techniques. The proposed algorithm consists of two processes, i.e., (i) Policy Evaluation and (ii) Policy Improvement. In the policy evaluation process, the Q -matrix is trained using the machine

operation data, which includes the state of the current value, the upcoming current state and the reference current (x_k, x_{k+1} and r_k) [23]. In this initial step, either the standard or recursive least square approaches may be employed. Furthermore, the applied voltage that achieves tracking performance is adjusted in the policy improvement stage. M is denoted as $M = [X_k \ u_k]^T$. These two processes are reiterated until the online Q-grid training algorithm used to track the VRM drive's current is as shown in Algorithm 1:

Algorithm 1: Online Training of Q-matrix by adopting voltage iteration scheme

Initialization: Start the algorithm with a steady voltage input. Reiterate and refine the subsequent pair of processes up to the point of confluence:

(i) Policy Evaluation:

$$M_k^T G^{i+1} M_k = (X_k^T) Q_q(X_k) + (u_k^i)^T R (u_k)^i + \gamma M_{k+1}^T G^{i+1} M_{k+1} \quad (9)$$

(ii) Policy Improvement:

$$u_k^{i+1} = - (G_{uu}^{-1})^{i+1} G_{uX}^{i+1} X_k \quad (10)$$

2.2. Regulating the Speed of the VRM Drive Using Reinforcement Structure

VRM is a type of salient machine in which the stator and rotor have a different number of poles. The value of the magnetic flux's reluctance varies due to the VRM's rotor rotation. Reluctance is lowest when the stator and rotor are perfectly aligned [24,25]. This causes the machine's inductance profile to have its maximum inductance value. When the stator position and rotor angle are entirely misaligned, the reluctance is at its highest value, resulting in the lowest inductance on the inductance profile. High magnetic saturation levels should be incorporated into the VRM's design to enable adequate electromechanical energy conversion. Since the magnetic nature of VRM changes between the aligned and unaligned positions of the stator and rotor, the inductance per phase can be varied at any instant current. The small air gap between the stator and rotor causes a substantial variation in the inductance of the aligned position with respect to the current value [26]. On the other hand, a large air gap in the unaligned position produces a slight variation in inductance. For the ideal case, the inductance profile resembles the trapezoidal waveforms. However, in reality, the trapezoidal wave is rounded at the corners due to saturation, resulting in a sinusoid with a slightly changing frequency. As shown in Figure 2, the current and rotor position influence the actual 12/8 VRM's inductance profile [27].

The traditional VRM drive current controllers, such as the hysteresis controller, enable the motor speed to be independent of the primary model. Motor speed can be accessed and injected into the model in this type of controller. Therefore, such controllers' VRM drive control system consists of a current and speed controller. The speed controller observes the actual motor speed and subtracts it from the reference speed to inject the desired current into the current controller [28,29]. By determining the rotor position and converting it to speed $\omega = d\theta/dt$, it is possible to observe the rotational speed. Additionally, the current controller does not require the speed information as it is model-independent [30]. In this study, the Q-node on the inductance surface is used to regulate the current of the VRM drive. Due to the machine's nonlinearity, LQT Q-learning is insufficient to control the current. Therefore, it is necessary to supplement the proposed method with crucial transitioning techniques to use Q-learning in nonlinear equations. The nonlinearity surface is subdivided into a sufficient number of Q-cores, with each Q-core functioning as a locally linearized region where the linear quadratic equation can be applied. Each cycle uses data tuples collected along the system trajectory to train the grid of Q-matrices positioned upon the system domain to its optimal values [29]. In order to apply the Q-grid after training all Q-matrices, the algorithm will detect the current and rotor angle value to determine the interpolative Q-matrix to be transmitted to the policy improvement step.

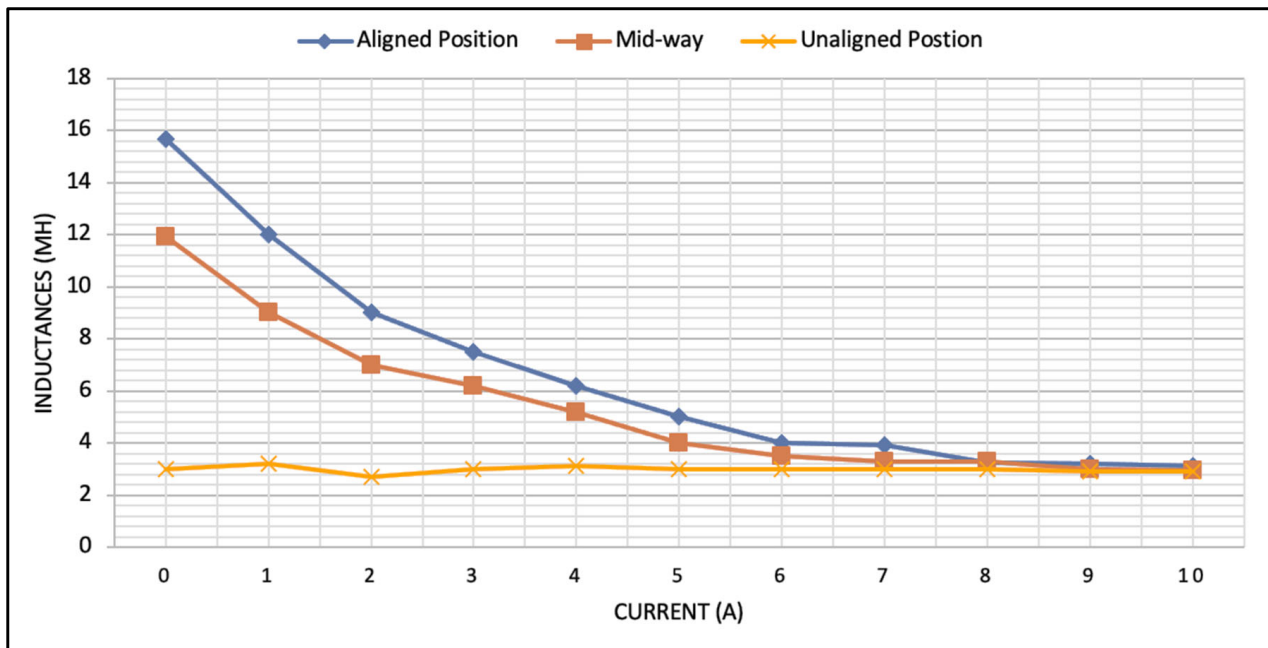


Figure 2. The nonlinear inductance profile of a VRM.

Furthermore, this approach necessitates knowledge of the four nearest Q-matrices for computation of the transitioned Q-matrix, which will be utilized to update the policy during policy improvement. It implies that only the Q-grid matrices used to model the system's operational state will be trained and utilized. Since the reference current is injected into the Q-grid, the Q-matrices must be retrained whenever the reference current changes. However, the speed is not considered in this method as a component of the nonlinearity. Although the model is not nonlinear in terms of speed, speed does exist in the model. Therefore, adopting a bidimensional Q-grid will necessitate retraining of the local Q-cores as the motor speed varies. A bidimensional grid requires substantially less memory than a tridimensional grid that contains speed as an axis. However, the tridimensional grid presented in this research will result in considerably faster dynamics and a better response. Designing speed regulator can be explored using two- and three-dimensional Q-grid techniques by using the policy iteration Q-grid learning algorithm.

2.2.1. Speed Regulator for the VRM Drive Using a Bidimensional Q-Grid

The stator current and rotor position constitute this approach's only two dimensions of the Q-grid. A three-phase 12/8 VRM is examined to demonstrate the bidimensional Q-grid for simulation outcomes. As shown in Figure 3, the phase shift between the phases is 30° , and the inductance configuration of the motor is periodic for every 22.5° . In the Q-grid, the rotor angle dimension is sampled every 2.5° , resulting in a total of 10 registered rotor positions and a sequence of stator current values with 2 A time steps on the current dimension. The angular velocity is not introduced as a dimension to the Q-grid by using the bidimensional Q-grid transitioning approach. The speed is not accessible as the Q-grid does not include knowledge about the speed. When all Q-matrices have been trained, the motor speed will finally adapt to the desired value. It indicates that altering the speed at any cycle forces the Q-grid to restart learning, which will create a transitory response due to the system's learning process.

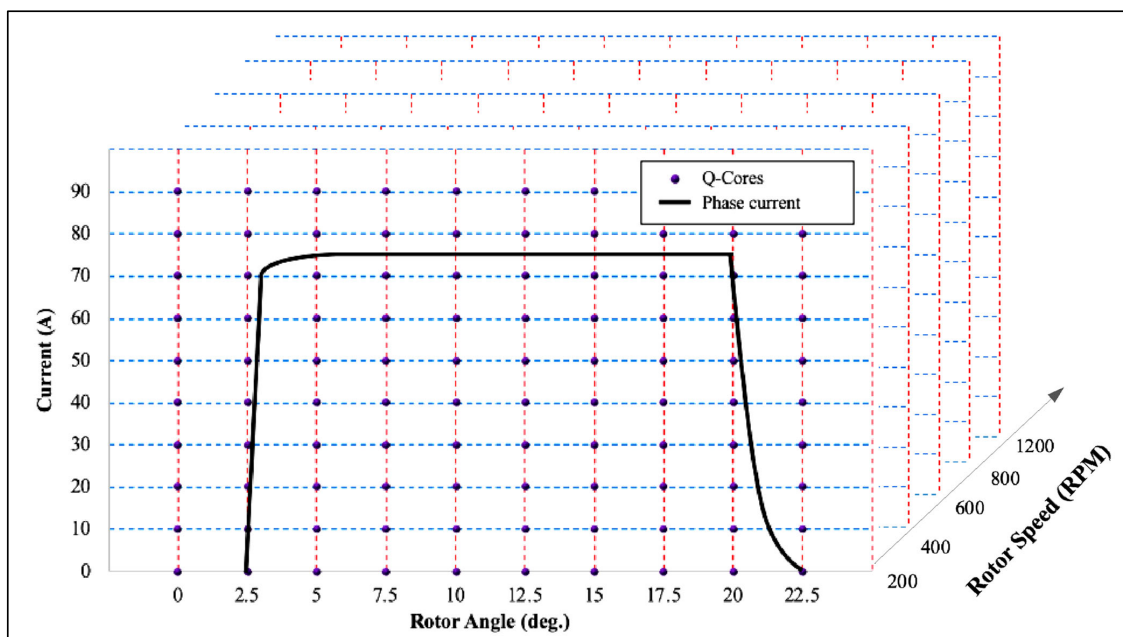


Figure 3. The current pulse on the tridimensional Q-grid.

2.2.2. Speed Regulator for the VRM Drive Using a Tridimensional Q-Grid

The previous section described the bidimensional Q-grid problem, i.e., the forced retraining of the Q matrices whenever the rotational speed changes. To address this problem, a tridimensional Q-grid has been created using the bidimensional Q-grid and the motor speed axis. Accordingly, the rotor position, phase current, and speed are the axes of the tridimensional Q-grid. Using a tridimensional Q-grid will make it possible to access the motor speed, allowing for a quicker response to changes in speed without the previously mentioned retraining process. The tridimensional Q-grid method includes three processes: partitioning, extraction, and transitioning. Firstly, the VRM's nonlinear surface domain must be partitioned and filled with sufficient sample points to create the tridimensional Q-grid. As shown in Figure 4, the tridimensional Q-grid is created by an equal step sampling along each Q-core surface axis. In order to extract the eight Q-matrices situated at the closest phase current, the extraction method aims to locate the Q-matrix at a predetermined speed level. Finally, the interpolative Q-matrix is computed during transitioning using the input phase current, speed signal, and eight extracted Q-matrices. Figure 4 depicts the steps necessary to apply the tridimensional Q-grid.

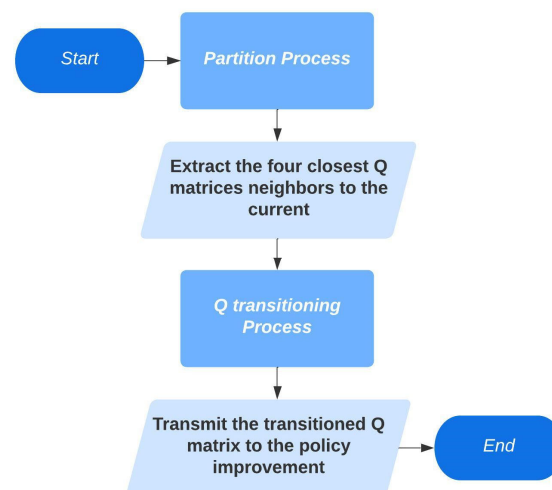


Figure 4. Flowchart of implementing tridimensional Q-grid algorithm.

Now, in order to ensure a smooth transition over the surface of this quantized tridimensional domain, the tridimensional interpolation mechanism must be introduced to this grid. The tridimensional grid's transitioning function for obtaining the interpolative Q-matrix (Q_s) can be expressed as follows:

$$Q_s = Q^T \mathcal{H}^T \mathcal{M} \tag{11}$$

where Q is the eight nearest Q-matrices to the Q_s ,

$$Q = [Q_1 \ Q_2 \ Q_3 \ Q_4 \ Q_5 \ Q_6 \ Q_7 \ Q_8] \tag{12}$$

\mathcal{H} is a predefined constant matrix,

$$\mathcal{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & -1 & 0 & 1 & 0 \\ 1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & -1 & 1 & 0 & 0 \\ -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 \end{bmatrix} \tag{13}$$

and \mathcal{M} is the vector of distances related to the three axes,

$$\mathcal{M} = [1 \ \Delta i \ \Delta \theta \ \Delta \omega \ \Delta i \Delta \theta \ \Delta \theta \Delta \omega \ \Delta \omega \Delta i \ \Delta i \Delta \theta \ \Delta \omega], \tag{14}$$

where $\Delta i = (i - i_0 / i_1 - i_0)$, $\Delta \theta = (\theta - \theta_0 / \theta_1 - \theta_0)$, and $\Delta \omega = (\omega - \omega_0 / \omega_1 - \omega_0)$. The parameters of tridimensional grid transitioning are illustrated in Figure 5.

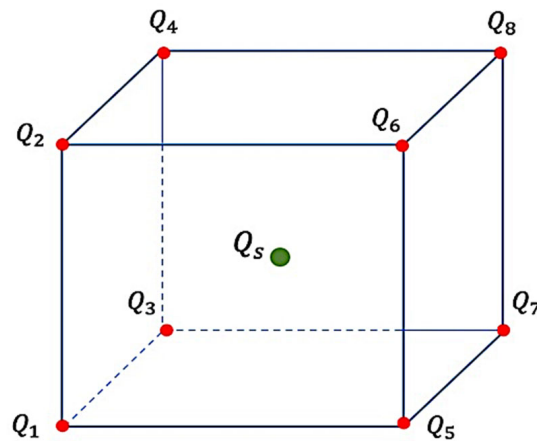


Figure 5. The definition of tridimensional transitioning parameters.

3. Simulation Results

The speed change responsiveness of the VRM drive has been demonstrated using a simulation of the introduced control algorithm. Figure 1 illustrates the block scheme of the controller. ω^* and i^* represents the optimal trajectory for the speed and the current respectively. There are two primary components of the controller. The first component trains the Q-matrices to evaluate the policy, while the second modifies the control input to improve the policy. Table 1 below shows the simulation results based on the machine's specifications. The equivalent model of the VRM may be expressed as

$$V = R_s i + \frac{d\lambda(\theta, i)}{dt} \quad (15)$$

$$T = \frac{1}{2} i^2 \frac{dL(\theta, i)}{d\theta} \quad (16)$$

where R_s represents the internal resistance of the winding and λ represents leakage flux for each phase derived by $\lambda = L(\theta, i)i$. L donates the inductance variations which depends on the rotor angle (θ) and the stator current (i). The control's sampling period has been set at 100 μ s. A stabilizing voltage signal is used to initialize the algorithm. The initial augmented state and voltage gains are denoted as $X_0 = [0 \ 0]^T$ and $K_0 = [1 \ -1]^T$, respectively. The performance index is calculated with Q determined to be 100 and R determined to be 0.001 as the parameters. The discount factor is chosen to be 0.9. The reference current model generates a series of pulses having a maximum amplitude of 4 A. The algorithm gathers 10 data tuples from the system trajectory at each Q-core in order to train the Q-matrix. The algorithms have been evaluated for both bidimensional and tridimensional Q-grids to illustrate the distinctions between the two.

Table 1. The specification of the VRM.

Parameters	Amount
Phase	3
Stator-poles/Rotor-poles	12/8
Rated power	0.7 HP
Stator resistance	2 Ω
Maximum inductance	16.6 mH
Minimum inductance	6 mH

3.1. Speed Regulator of Tridimensional Q-Grid Algorithm

The current control was initially tested to ensure the output current traced the reference current (Figure 6). The speed was set to 60 RPM and exhibited a transient response until it reached the appropriate level due to the learning process. Moreover, the speed was altered to 100 RPM at 0.55 s after a specific cycle. As a result, the Q-matrices must be retrained in reaction to that change. Figure 7 exhibits the behavior of the stator current during the speed change.

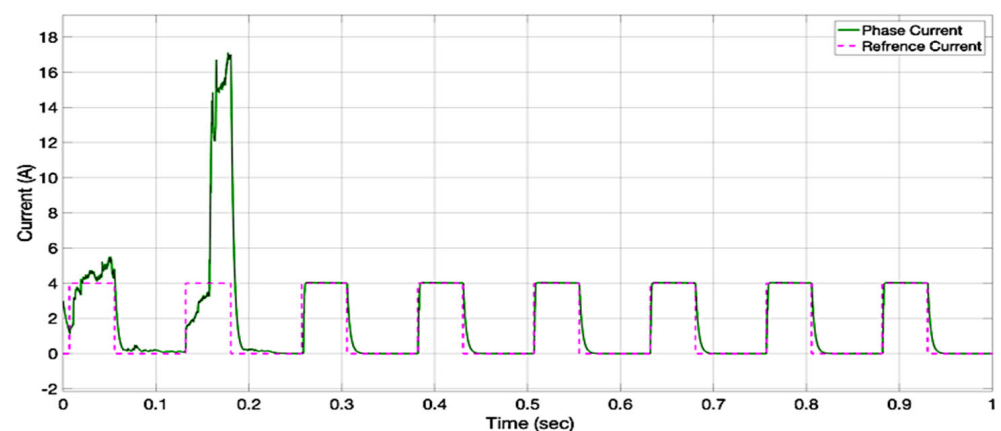


Figure 6. The current trajectory of the proposed control is comparing with the ideal current.

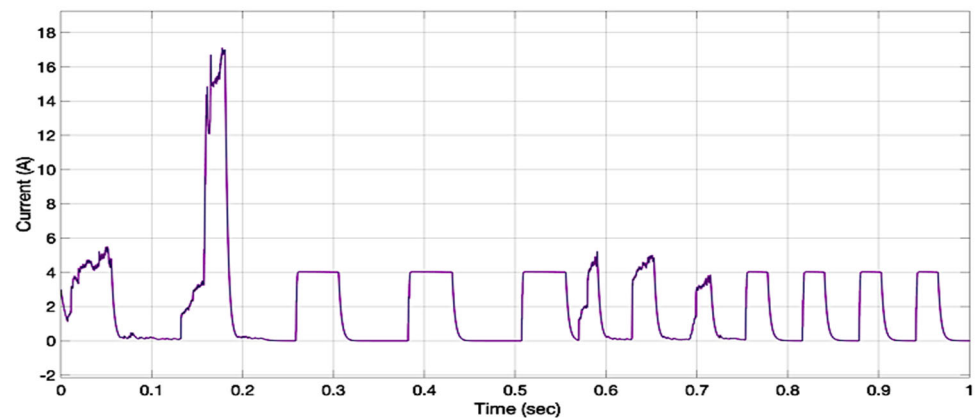


Figure 7. The nature of the current when the speed is altered using a bidimensional Q-grid.

3.2. Speed Regulator Using Tridimensional Q-Grid Learning Algorithm

The tridimensional Q-grid is implemented in this test. During training, when the rotational speed is increased from 60 to 100 RPM, the speed responds quickly and remains stable. Figure 8 depicts the conduct of the current at the moment that the motor speed modified. Due to the presence of the speed axis in the Q-grid, it is unnecessary to retrain the Q-matrices when the speed changes every 0.6 s. After the algorithm has been initialized and the Q-grid has adapted to the plant, the controller will use the plane corresponding to the recorded speed. Consequently, at any given speed, the controller will use the optimal Q values with no requirement of re-adaptation. Figure 9 depicts the behavior of the input voltage as the speed varies.

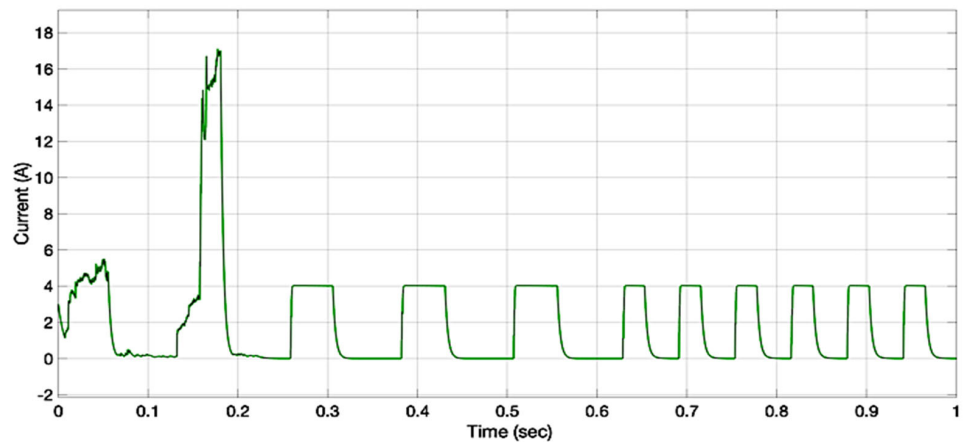


Figure 8. The nature of the current when regulating the speed using a tridimensional Q-grid.

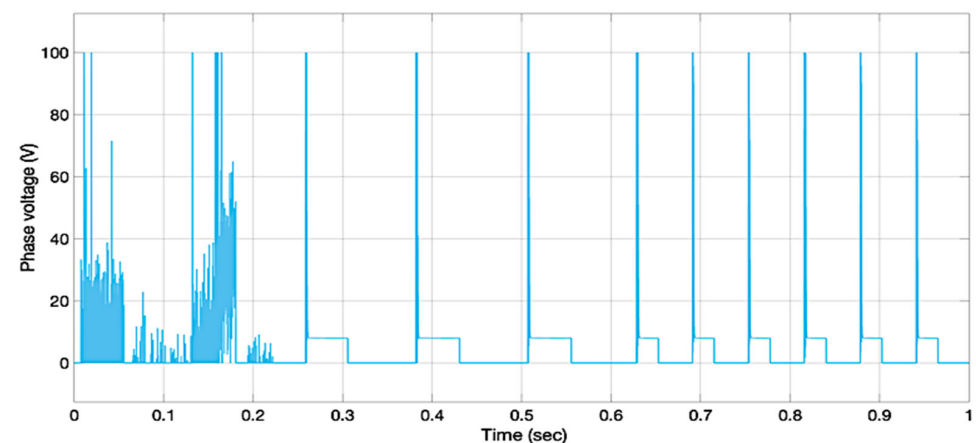


Figure 9. The optimal applied voltage when the speed changes using a tridimensional Q-grid.

4. Experimental Results

In this section, an experiment to study the proposed speed regulator while using the policy iteration learning control method is presented. The structure and the components of the experiment are shown in Figure 10. After the algorithm's code is tested, it is connected to the control board, which contains a reliable microcontroller to handle the rapid switching frequency. The parameters of SRM used in the experiment are the same as used in the simulation results (Table 1) with a DC voltage of 100 V. An asymmetric bridge topology, which involves two transistor and two diodes per phase, has been designed for this experiment. In addition, the DC machine is connected to the SRM as a mechanical load. In this test, initially, when the rotational speed has been changed, the speed regulator using the bidimensional method along with current was not stable because of training procedure (Figure 11a). When the Q-matrices are completely trained after certain time steps, the speed is considered "learned" and becomes stable at a constant value. The zoomed version of the proposed algorithm, which uses the tridimensional scheme, is illustrated in Figure 11b. The policy iteration learning algorithm along with tridimensional lookup grid includes motor speed variations within the planned domain; thus, when a rotational speed is changed by adding mechanical load torque, the speed responds very fast compared to bidimensional algorithm, since it does not need further steps of learning to do so. For the purpose of demonstrating the speed regulator between the policy iteration learning transitioning regulator and the conventional regulator, hysteresis control has been added to this test. To showcase the uniqueness of the proposed method, a model-free policy iteration control can effectively adjust the speed in merely four cycles, typically under 15 milliseconds. This rapid response highlights the efficiency of the method. Furthermore, the technique significantly reduces pulsations compared to the traditional hysteresis method, ensuring smoother operation. It also eliminates the need for retraining when speed changes, which enhances its practicality in dynamic environments. Additionally, the controller requires no tuning, even if the VRM parameters change due to factors such as aging or variations in air gaps. The continuous online training ensures that the system remains adaptive and robust, maintaining optimal performance without manual adjustments.

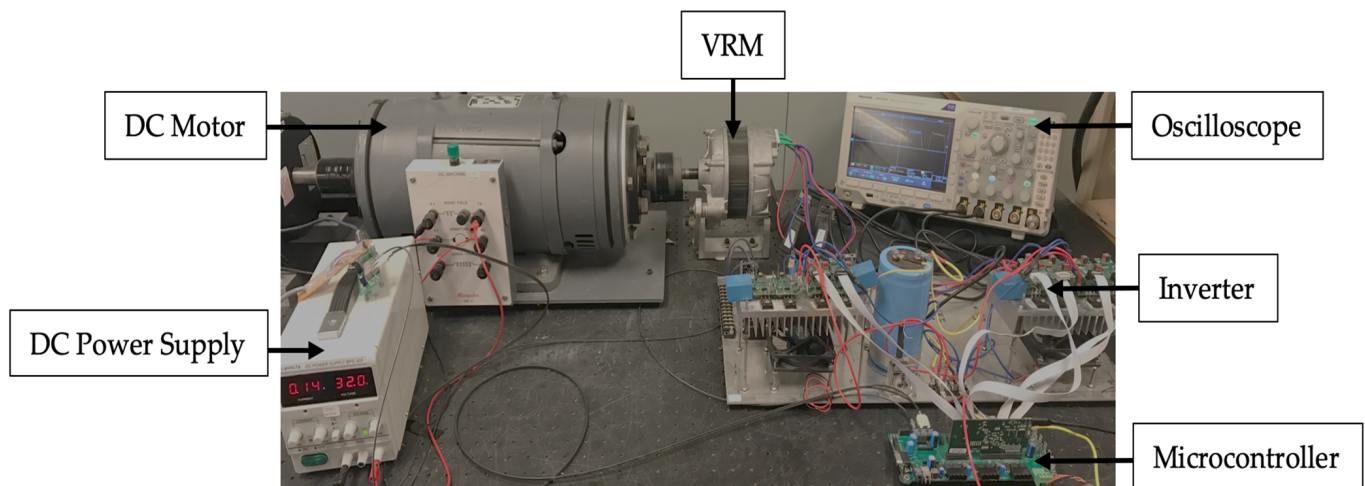


Figure 10. The structure of the experiment.

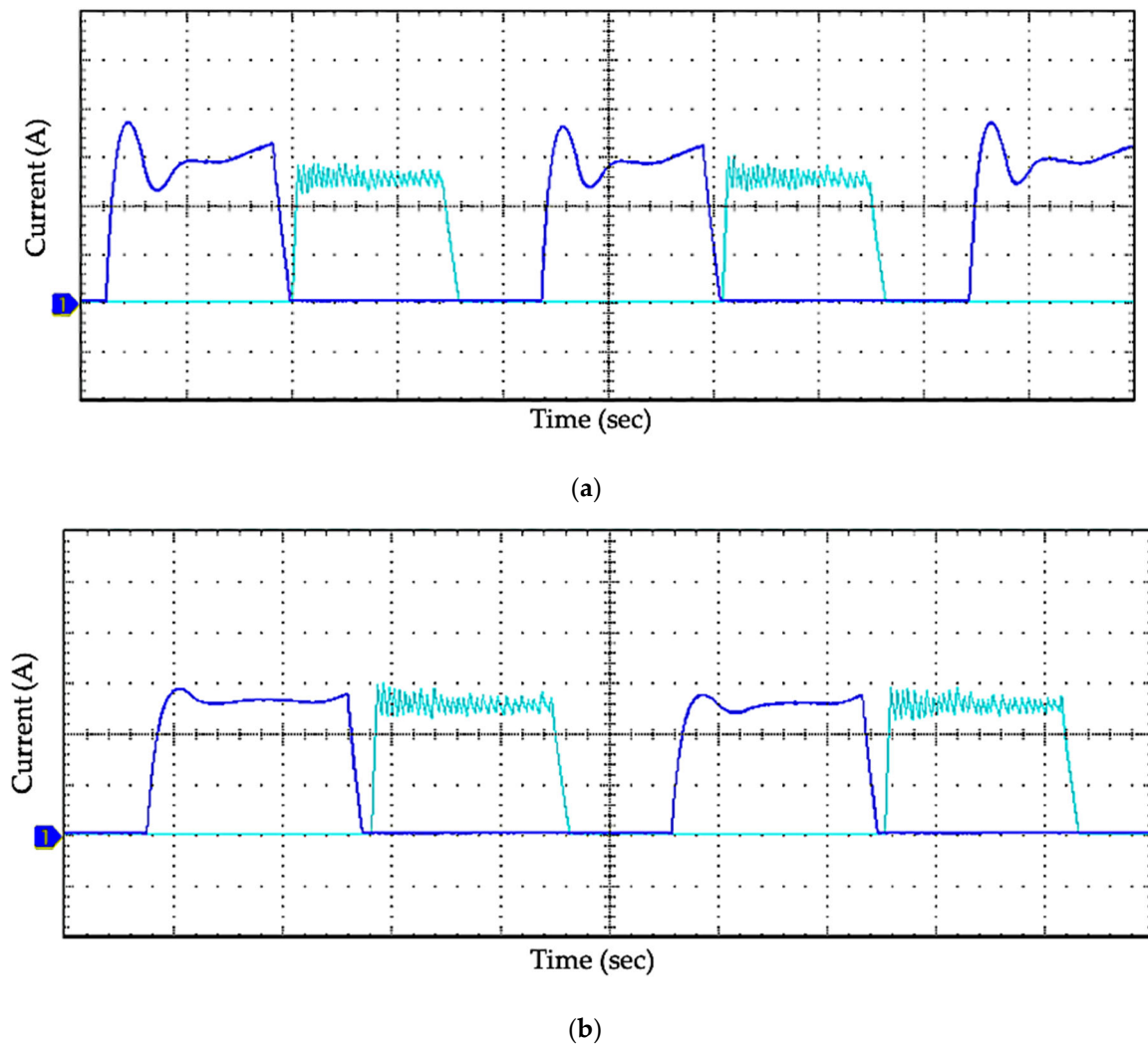


Figure 11. The behavior of the speed regulator using both algorithms, (a) using a bidimensional grid and untrained Q-grid (b) using a tridimensional trained Q-grid.

5. Conclusions

This research introduced a novel speed regulator using an adaptive online transitioned policy iteration technique for the variable reluctance motor drive. After presenting the learning algorithm, a novel interpolation transitioning technique has been incorporated into the proposed controller to implement a linear controller in a highly nonlinear system. This control technique gives exceptional tracking performance for the VRM. The primary disadvantage of employing a bidimensional algorithm for the proposed controller was its response to speed changes. Since the controller's motor speed was not included in the inductance surface model of the machine, any change in speed necessitated a forced retraining of the local Q-cores. In this research, a three-dimensional grid has been used to address this issue. The tridimensional algorithm enables access to the motor speed and its modification without retraining. In addition, a tridimensional interpolation was added to this tridimensional grid to smooth out the transitions of the controller across the local learning centers. Lastly, the simulation and experimental results illustrated the behavior of the VRM's speed when employing tridimensional policy iteration learning algorithm. The proposed algorithm successfully regulated the machine's speed and significantly reduced current oscillations without the need for additional procedures to handle the model's inherent nonlinearity. By eliminating the necessity for complex compensatory mechanisms, the algorithm demonstrates a streamlined and efficient approach to speed control. This

innovation simplifies the control process, making it more robust and adaptable to varying operational conditions.

Funding: This research was funded by the Qassim University, under grant number QU-APC-2024-9/1.

Data Availability Statement: The original data presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Acknowledgments: The Researchers would like to thank the Deanship of Graduate Studies and Scientific Research at Qassim University for financial support (QU-APC-2024-9/1). The researcher would also like to express gratitude to the colleagues at the G2 Power Lab, Department of Electrical Engineering, Missouri University of Science and Technology, for their assistance in preparing the experimental setup for this paper.

Conflicts of Interest: The author declares no conflict of interest.

References

- Valdivia, V.; Todd, R.; Bryan, F.J.; Barrado, A.; Lázaro, A.; Forsyth, A.J. Behavioral Modeling of a Switched Reluctance Generator for Aircraft Power Systems. *IEEE Trans. Ind. Electron.* **2013**, *61*, 2690–2699. [\[CrossRef\]](#)
- Kawa, M.; Kiyota, K.; Furqani, J.; Chiba, A. Acoustic Noise Reduction of a High-Efficiency Switched Reluctance Motor for Hybrid Electric Vehicles with Novel Current Waveform. *IEEE Trans. Ind. Appl.* **2018**, *55*, 2519–2528. [\[CrossRef\]](#)
- Chen, H.; Yan, W.; Gu, J.J.; Sun, M. Multiobjective Optimization Design of a Switched Reluctance Motor for Low-Speed Electric Vehicles with a Taguchi–CSO Algorithm. *IEEE/ASME Trans. Mechatron.* **2018**, *23*, 1762–1774. [\[CrossRef\]](#)
- Kiyota, K.; Chiba, A. Design of Switched Reluctance Motor Competitive to 60-KW IPMSM in Third-Generation Hybrid Electric Vehicle. *IEEE Trans. Ind. Appl.* **2012**, *48*, 2303–2309. [\[CrossRef\]](#)
- Yan, N.; Cao, X.; Deng, Z. Direct Torque Control for Switched Reluctance Motor to Obtain High Torque–Ampere Ratio. *IEEE Trans. Ind. Electron.* **2019**, *66*, 5144–5152. [\[CrossRef\]](#)
- Isfahani, A.H.; Fahimi, B. Comparison of Mechanical Vibration between a Double-Stator Switched Reluctance Machine and a Conventional Switched Reluctance Machine. *IEEE Trans. Magn.* **2014**, *50*, 293–296. [\[CrossRef\]](#)
- Jia, C.; He, H.; Zhou, J.; Li, J.; Wei, Z.; Li, K. Learning-Based Model Predictive Energy Management for Fuel Cell Hybrid Electric Bus with Health-Aware Control. *Appl. Energy* **2024**, *355*, 122228. [\[CrossRef\]](#)
- Gobbi, R.; Ramar, K. Optimisation Techniques for a Hysteresis Current Controller to Minimise Torque Ripple in Switched Reluctance Motors. *IET Electr. Power Appl.* **2009**, *3*, 453. [\[CrossRef\]](#)
- Shao, B.; Emadi, A. A Digital PWM Control for Switched Reluctance Motor Drives. In Proceedings of the 2010 IEEE Vehicle Power and Propulsion Conference, Lille, France, 1–3 September 2010; pp. 1–6.
- Schulz, S.E.; Rahman, K.M. High-Performance Digital PI Current Regulator for EV Switched Reluctance Motor Drives. *IEEE Trans. Ind. Appl.* **2003**, *39*, 1118–1126. [\[CrossRef\]](#)
- Ye, J.; Malysz, P.; Emadi, A. A Fixed-Switching-Frequency Integral Sliding Mode Current Controller for Switched Reluctance Motor Drives. *IEEE J. Emerg. Sel. Top. Power Electron.* **2014**, *3*, 381–394. [\[CrossRef\]](#)
- Lukic, S.M.; Emadi, A. State-Switching Control Technique for Switched Reluctance Motor Drives: Theory and Implementation. *IEEE Trans. Ind. Electron.* **2010**, *57*, 2932–2938. [\[CrossRef\]](#)
- Lin, Z.; Reay, D.; Williams, B.; He, X. High-Performance Current Control for Switched Reluctance Motors Based on on-Line Estimated Parameters. *IET Electr. Power Appl.* **2010**, *4*, 67–74. [\[CrossRef\]](#)
- Akcayol, M.A. Application of Adaptive Neuro-Fuzzy Controller for VRM. *Adv. Eng. Softw.* **2004**, *35*, 129–137. [\[CrossRef\]](#)
- Alharkan, H.; Shamsi, P.; Saadatmand, S.; Ferdowsi, M. Q-Learning Scheduling for Tracking Current Control of Switched Reluctance Motor Drives. In Proceedings of the 2020 IEEE Power and Energy Conference at Illinois (PECI), Champaign, IL, USA, 27–28 February 2020; pp. 1–6.
- Alharkan, H.; Saadatmand, S.; Ferdowsi, M.; Shamsi, P. Optimal Tracking Current Control of Switched Reluctance Motor Drives Using Reinforcement Q-Learning Scheduling. *IEEE Access* **2021**, *9*, 9926–9936. [\[CrossRef\]](#)
- Alharkan, H. *Adaptive Dynamic Programming Methods for Tracking Current Control of Switched Reluctance Motor Drive*; Scholars Mine: Rolla, MI, USA, 2021.
- Lewis, F.L.; Vrabie, D.; Syrmos, V.L. *Optimal Control*; John Wiley & Sons: Hoboken, NJ, USA, 2012; ISBN 1118122720.
- Kiumarsi, B.; Lewis, F.L.; Modares, H.; Karimpour, A.; Naghibi-Sistani, M.-B. Reinforcement Q-Learning for Optimal Tracking Control of Linear Discrete-Time Systems with Unknown Dynamics. *Automatica* **2014**, *50*, 1167–1175. [\[CrossRef\]](#)
- Kiumarsi-Khomartash, B.; Lewis, F.L.; Naghibi-Sistani, M.-B.; Karimpour, A. Optimal Tracking Control for Linear Discrete-Time Systems Using Reinforcement Learning. In Proceedings of the 52nd IEEE Conference on Decision and Control, Firenze, Italy, 10–13 December 2013; pp. 3845–3850.
- Lewis, F.L.; Vrabie, D. Reinforcement Learning and Adaptive Dynamic Programming for Feedback Control. *IEEE Circuits Syst. Mag.* **2009**, *9*, 32–50. [\[CrossRef\]](#)

22. Reinforcement Learning and Feedback Control: Using Natural Decision Methods to Design Optimal Adaptive Controllers. *IEEE Control Syst.* **2012**, *32*, 76–105. [[CrossRef](#)]
23. Werbos, P.J.; Miller, W.T.; Sutton, R.S. A Menu of Designs for Reinforcement Learning over Time. In *Neural Networks of Control*; MIT Press: Cambridge, MA, USA, 1995; pp. 67–95.
24. Liu, D.; Lewis, F.L.; Wei, Q. Editorial Special Issue on Adaptive Dynamic Programming and Reinforcement Learning. *IEEE Trans. Syst. Man. Cybern. Syst.* **2020**, *50*, 3944–3947. [[CrossRef](#)]
25. Matwankar, C.S.; Pramanick, S.; Singh, B. Position Sensorless Torque Ripple Control of Switched Reluctance Motor Drive Using B-Spline Neural Network. In Proceedings of the IECON 2021—47th Annual Conference of the IEEE Industrial Electronics Society, Toronto, ON, Canada, 13–16 October 2021; pp. 1–6.
26. Kiumarsi, B.; Lewis, F.L. Actor–Critic-Based Optimal Tracking for Partially Unknown Nonlinear Discrete-Time Systems. *IEEE Trans. Neural Netw. Learn. Syst.* **2015**, *26*, 140–151. [[CrossRef](#)]
27. Buriakovskiy, S.; Maslii, A.; Tyshchenko, A. Synthesis of the Speed Controller of the Switched Reluctance Motor. In *Systems, Decision and Control in Energy V*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 179–193.
28. Sun, X.; Xiong, Y.; Yao, M.; Tang, X. A Hybrid Control Strategy for Multimode Switched Reluctance Motors. *IEEE/ASME Trans. Mechatron.* **2022**, *27*, 5605–5614. [[CrossRef](#)]
29. Feng, L.; Sun, X.; Yang, Z.; Diao, K. Optimal Torque Sharing Function Control for Switched Reluctance Motors Based on Active Disturbance Rejection Controller. *IEEE/ASME Trans. Mechatron.* **2023**, *28*, 2600–2608. [[CrossRef](#)]
30. Jackiewicz, K.; Straś, A.; Bałkowiec, T.; Kaszewski, A.; Ufnalski, B. Novel Dual Thread Angle Sampled Multioscillatory-Based Control for Speed Ripple Reduction in a Switched Reluctance Machine-Based Drive. *ISA Trans.* **2023**, *139*, 724–738. [[CrossRef](#)] [[PubMed](#)]

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