


Article

Deep Deterministic Policy Gradient-Based Actor–Critic Reinforcement Learning for Torque Ripple Minimization in Switched Reluctance Motors

Divya Ramasamy ^{1,*}  and Sundaram Maruthachalam ²

¹ Department of Electrical and Electronics Engineering, PSG Institute of Technology and Applied Research, Neelambur, Coimbatore 641062, Tamilnadu, India

² Department of Electrical and Electronics Engineering, PSG College of Technology, Peelamedu, Coimbatore 641004, Tamilnadu, India; msm.eee@psgtech.ac.in

* Correspondence: divya@psgitech.ac.in

Abstract

The aim of this research is to investigate and reduce the torque ripple in Switched Reluctance Motor (SRM) drives, which is one of the major barriers to their acceptance for electric vehicle propulsion applications despite the advantages of robustness, efficiency, and wide operating range. High torque ripple not only deteriorates drive smoothness but also contributes to noise and vibration, demanding an advanced control strategy beyond traditional current-shaping and switching-based approaches. In this context, this work proposes a DDPG (Deep Deterministic Policy Gradient) Actor–Critic Neural Network-based reinforcement learning control framework that learns the optimal firing angle offsets dynamically to ensure less ripple electromagnetic torque under varying speeds and load conditions. The developed strategy has been designed and trained in MATLAB Simulink R2024b and then deployed in real time using an FPGA-based digital controller for validation on hardware. Comparative analysis with TSF (Torque Sharing Function) and DITC (Direct Instantaneous Torque Control) demonstrates that the reinforcement learning approach gives a much smoother torque response with better dynamic behavior over the operating range analyzed.

Keywords: switched reluctance motor; torque ripple minimization; deep deterministic policy gradient (DDPG); actor–critic neural network; reinforcement learning; Asymmetrical Half-Bridge converter; adaptive firing angle optimization; intelligent electric drive systems

1. Introduction

The growing global climate crisis has intensified the demand for sustainable and carbon-neutral energy solutions, with the transportation sector accounting for a significant share of global greenhouse gas emissions. Electric vehicles (EVs), with high energy efficiency, zero emissions, reduced noise, and improved performance compared to traditional internal combustion engine cars, have been considered a potential means of transportation in the future. The electric traction motor, which remains a pivotal part of an EV, significantly influences EV performance, efficiency, reliability, and driving comfort. The desired traction motor is characterized by high torque–power density, a broad constant torque speed range, compact size, high efficiency, robust construction, low cost, reduced noise, and minimized vibration effects [1]. The various configurations of electric machines that have been extensively researched for use in EV traction drives include Permanent Magnet Synchronous Motors (PMSMs), Squirrel Cage Induction Motors (SCIMs), Brush-Less DC



Academic Editor: Ahmed Abu-Siada

Received: 9 January 2026

Revised: 25 February 2026

Accepted: 11 March 2026

Published: 16 March 2026

Copyright: © 2026 by the authors.

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/).

Motors (BLDCs), Synchronous Reluctance Motors (SynRels), and Switched Reluctance Motors (SRMs). Among these, PMSMs have been extensively used in commercial EV developments owing to high torque density, high efficiency, and precise controllability [2]. Squirrel Cage Induction Motors (SCIMs), although non-magnetic, are less efficient, heavier, with a degraded power factor, especially when subjected to variable conditions [3]. Synchronous Reluctance Motors (SynRels), although a potential alternative solution to modern electric vehicles, are still limited by reduced torque density, with significant torque ripples, making them unsuitable for large-scale adoption in EVs [4]. Among the magnet-free motors, Switched Reluctance Motors (SRMs) have gained increasing attention for EV traction owing to their simple and robust construction, inherent fault tolerance, wide constant power speed range, and economical manufacturing [5–7]. Additionally, the concentration of losses in the stator facilitates effective thermal management. Despite advances in power electronics and control strategies, the major drawback of SRMs remains their high torque ripple caused by doubly salient geometry, nonlinear magnetic characteristics, and discrete phase commutation, which induces vibration and acoustic noise and limits their deployment in high-performance EV applications. A comparative assessment of major propulsion motors for EV applications is summarized in Table 1.

Table 1. Comparative analysis of propulsion motors for EV applications.

<i>Motor Parameters</i>	<i>SRM</i>	<i>PMSM</i>	<i>SCIM</i>
<i>Size</i>	Compact	Moderate	Moderate
<i>Weight</i>	Low	Moderate	Moderate
<i>Cost</i>	Low	High	Low
<i>Ruggedness</i>	High	Low	High
<i>Constant Torque/Speed Range</i>	Wide	Wide	Moderate
<i>Torque Ripple</i>	High	Low	Low
<i>Noise & Vibration</i>	High	Low	Low
<i>Power Converters</i>	Specific	Modular	Modular
<i>Permanent Magnets</i>	No	Yes	No
<i>Efficiency</i>	Moderate	High	Low

A number of studies have recently been published on the subject of the use of traditional methods in the design and operation of Switched Reluctance Motors (SRMs) to alleviate the problem of inherent torque ripple. One of the most common methods is that of Torque Sharing Functions (TSFs), where the reference torque is split between simultaneously providing torque from multiple phases using a determined profile, such as cubic, linear, sinusoidal, and partitioned functions. These methods were very successful at reducing torque ripple when SRMs operated at low and medium speeds, but as the speed of the motor increases, the impact of magnetic saturation on torque ripple, phase coupling effects, and nonlinear torque profiles tends to reduce the effectiveness of these algorithms [8–10]. During the past several years, both current profiling and current control methods were proposed as alternative means for shaping the phase currents of SRMs to generate smoother torque, but both types of schemes introduced higher switching losses and increased the complexity of the controller due to the need to accurately track phase current variation over wide operating ranges [11,12]. Additionally, the implementation of a Direct Torque Control (DTC) scheme directly controls the instantaneous torque and provides faster dynamical response; however, they are also sensitive to the accuracy of torque estimation and variations in machine parameters. This limits the robustness of DTC

schemes when applied to practical situations [13,14]. Recently, Model Predictive Control (MPC) and finite control-set predictive control have been applied to identify the optimal switching states in order to reduce the torque ripple of SRMs; however, although the MPC and finite control-set predictive control methods have proven to be successful, they will require large amounts of computational power and high-speed sensing capabilities, which limit their implementation in real time within automotive traction systems [15]. Classical techniques for minimizing torque ripple, such as rotor pole shaping and stator tooth modifications, have been explored in the areas of machine design as well as machine control. The use of these techniques provides some improvements in the overall structural design, but is limited due to manufacturing complexity, higher costs, and poor adaptability to changing machine operating conditions [16–18]. Despite the advances made in torque ripple minimization, the performance of traditional methods is still restricted by non-linear magnetic characteristics, machine operating point dynamic dependence, and added complexity of better control strategies, suggesting the need for adaptive control techniques.

Classical methods for minimizing the torque ripple in Switched Reluctance Motors (SRMs) typically use accurate magnetic models, fixed parameters, and offline tuning. This means the technique is less adaptable to changes in speed and load and has no internal mechanism for self-learning to accommodate changes in parameters. Many recent studies have focused on adopting artificial intelligence and machine learning based control techniques, with reinforcement learning (RL) being one of the most common since it offers a model-free algorithm that can learn optimal control actions by interacting with the motor in real-time. Within RL, the control policy is defined as a Markov Decision Process (MDP), where the controller observes the state of the system, makes a decision (action), and receives feedback from the system (reward). The goal of the RL algorithm is to maximize the expected sum of rewards as expressed, as in Equation (1):

$$\pi^* = \arg \mathbb{E} \left[\sum_{t=0}^{\infty} \gamma^t r_t \right] \quad (1)$$

where π denotes the control policy, r_t represents the instantaneous reward and $\gamma \in [0, 1]$ is the discount factor. This framework is particularly well-suited to Switched Reluctance Motors (SRMs) because torque ripple, speed variations, and current profiles can all be quantified and incorporated into the reward function without needing to develop complex models. In this work, an actor–critic architecture based on the Deep Deterministic Policy Gradient (DDPG) is developed for minimizing torque ripple in an SRM, powered by an Asymmetrical Half-Bridge (ASHB) converter, thus allowing continuous self-optimization through adaptive control actions.

2. Torque Ripple Phenomenon and Its Impact on SRM Performance

The Switched Reluctance Motor (SRM) is an electromagnetic energy-conversion device in which torque is produced by the tendency of its rotor to move toward a position of minimum reluctance. It consists of salient poles on both the stator and rotor, without permanent magnets or rotor windings. Each stator phase winding is concentrated around diametrically opposite poles and excited independently according to rotor position. SRM inherently exhibits torque ripple because of its doubly salient structure and nonlinear variation in phase inductance with rotor position. The instantaneous electromagnetic torque T_e can be defined as in Equation (2):

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

where i is the phase current and $L(\theta, i)$ is the position and current-dependent inductance profile of the phase winding. Torque is produced only in the region where $\frac{dL}{d\theta} > 0$; however, at commutation instants, when one phase current decays and the next builds up, discontinuities in the net torque lead to pulsations in the developed torque.

The torque ripple magnitude is generally quantified using the torque ripple factor (TRF), defined as in Equation (3):

$$\text{Torque Ripple (\%)} = \frac{T_{\max} - T_{\min}}{T_{\text{avg}}} \times 100 \quad (3)$$

where T_{\max} , T_{\min} and T_{avg} represent the maximum, minimum, and average torque within one electrical cycle, respectively.

A lower TRF indicates smoother torque production. For continuous analysis over time, the ripple can also be expressed in RMS form as in Equation (4),

$$\text{Torque Ripple (RMS)} = \frac{\sqrt{\frac{1}{T} \int_0^T (T_e(t) - T_{\text{avg}})^2 dt}}{T_{\text{avg}}} \quad (4)$$

where $T_e(t)$ is the instantaneous electromagnetic torque, and T is one electrical period. The primary sources of torque ripple in Switched Reluctance Motors arise from several inherent electromagnetic and control-related factors. Nonlinear magnetic saturation and the strong dependence of flux linkage on rotor position introduce variations in torque production. While the discrete nature of phase excitation and mutual coupling between phases also contribute to sudden torque variations, the asymmetry in magnetic reluctance of the rotor aligned and unaligned positions increases the torque pulsations. In addition, incorrect current profiling and poorly controlled phase commutating overlap can further enhance these ripples, which, altogether, result in significant torque ripple in SRM. These periodic torque ripples give rise to mechanical vibrations, acoustic noise, and resonance within the EV drivetrain. Further, oscillatory torque components lead to reduced efficiency, increased mechanical wear, and difficulties in smooth speed and position control. Thus, the reduction in torque ripple becomes a key concern for SRM-based electric vehicle drives to improve comfort, increase the mechanical life, and ensure real optimal efficiency. Current profiling, DITC, and TSF are the techniques widely employed. However, advanced intelligent methods are required to overcome inherent nonlinearity and obtain real-time torque smoothness under different load and speed conditions.

3. Selection of Converter Topology for SRM

Converter topology has a significant part in the performance, controllability, and efficiency of Switched Reluctance Motor (SRM) drives. Because SRMs require each phase to be excited independently according to rotor position, the converter should provide bidirectional current flow, fast turn-off capability, and efficient energy recovery. For controlled operation, the SRM needs an appropriate power converter; it should not be coupled directly to a direct-current or alternating current source. Its primary job is to properly energize and de-energize every phase winding to guarantee continual rotation. For SRM drives, a variety of converter types have been proposed. The type selected has a major impact on the cost, size, and performance of the drive. To enhance the performance of SRM drives, the power converter must satisfy a number of conditions to swiftly reach the reference phase current and raise the motor's base speed; the optimal SRM power converter should facilitate rapid magnetization while also ensuring rapid demagnetization to eliminate the current tail and prevent negative torque production. It must allow phase overlap, so that the incoming phase can be energized while the outgoing phase is de-energized. The capac-

ity to retain demagnetization energy for use in energizing other phases or return it to the source instead of dissipating it as thermal energy is crucial for high efficiency. The converter should also maintain minimum cost by using a reduced number of devices for switching, reducing gate-drivers' needs, and loss in switching. For critical applications, fault tolerance is important to ensure uninterrupted SRM operation even in the event of a phase failure, and the overall design should remain low in complexity with simple implementation and control requirements.

Several converter configurations have been proposed in the previous literature [19], including the Unipolar Converter, which is the simplest form, using a single switch and diode per phase, and is economical but suffers from slow demagnetization and lacks bidirectional current capability. The C-Dump Converter stores and reuses demagnetization energy through a common capacitor but introduces control complexity and subjects the dump capacitor to high voltage stress. In the case of the R-Dump Converter, the stored energy is dissipated through the resistor during demagnetization, which results in considerable energy loss and low drive efficiency. The Shared-Switch Converter decreases the overall number of switches but impairs independent phase control and increases commutation overlap, therefore giving rise to high torque ripple. Contrastingly, the ASHB Converter uses two controlled switches and two diodes per phase. It can provide fully independent phase control, quick current decay, and recover energy for the DC bus. Due to its excellent control features and strong alignment with high-performance control methodologies, the Asymmetric Half-Bridge (ASHB) converter is highly controllable. It allows independent phase excitation, so that each SRM phase can be switched on and off independently for precise torque control and accurate phase current shaping, suitable for several control methods in minimizing the torque ripple. The topology allows recovering energy efficiently during commutation by routing demagnetization current through the freewheeling diodes back to the supply, reducing thermal stress while improving energy efficiency. With its two-switch-per-phase configuration, fast current rise and decay are guaranteed, which is highly essential in minimizing torque ripple at high speeds. Additionally, the ASHB converter provides bidirectional current capability in providing regenerative braking with four-quadrant operational capabilities ideal for traction drive applications. Besides these, its fault tolerance is high, it allows the motor to continue its operation even in the event of one-phase failure, a very critical feature for safety-critical electric vehicle systems [20]. The converter's control structure is simple to implement and supports both hysteresis and PWM modulation techniques, enabling straightforward hardware realization using DSPs, FPGAs, or microcontrollers. Considering these advantages, the ASHB topology provides an optimal balance of hardware complexity, efficiency, and control flexibility and is therefore employed in this work to energize the SRM phases and implement torque ripple minimization through the proposed DDPG-based actor-critic reinforcement learning strategy.

Asymmetrical Half-Bridge (ASHB) Converter Topology

This subsection describes the widely used, popular SRM converter, the ASHB, operation, as shown in Figure 1, for a 4 ϕ SRM; this converter uses two semiconductor diodes and two controlled switches per phase.

Magnetization, free-wheeling, and demagnetization are the three operating states of the ASHB converter. Figures 2–4 show these stages, and the description that follows is for a particular phase, Ph_A , and Figure 5 shows the typical waveforms for voltage and current for the three states.

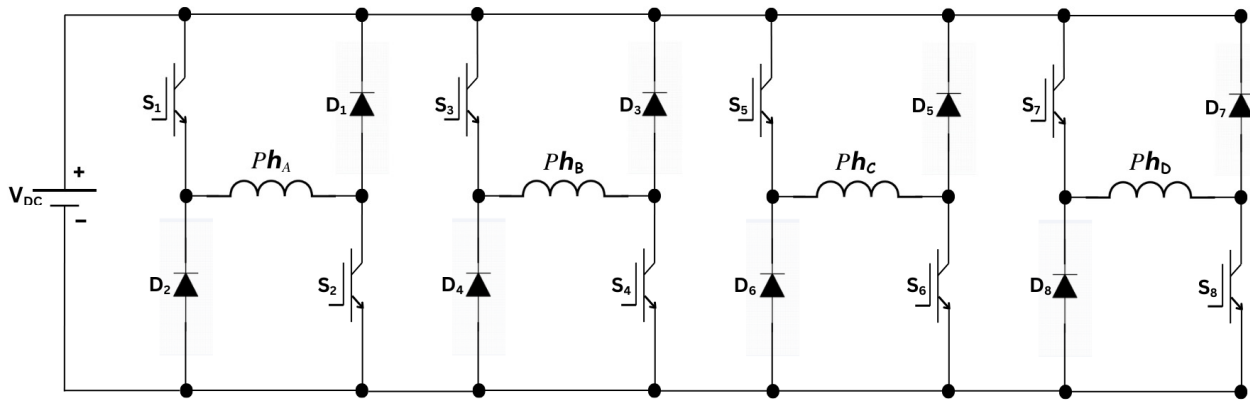


Figure 1. Asymmetrical Half-Bridge (ASHB) converter for a 4 ϕ Switched Reluctance Motor (SRM).

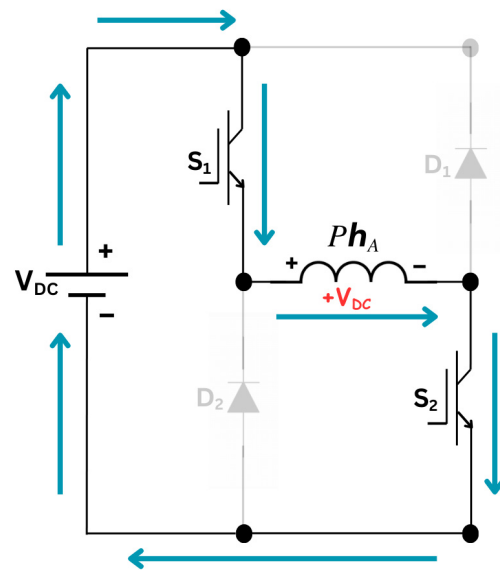


Figure 2. Magnetization state of the ASHB converter.

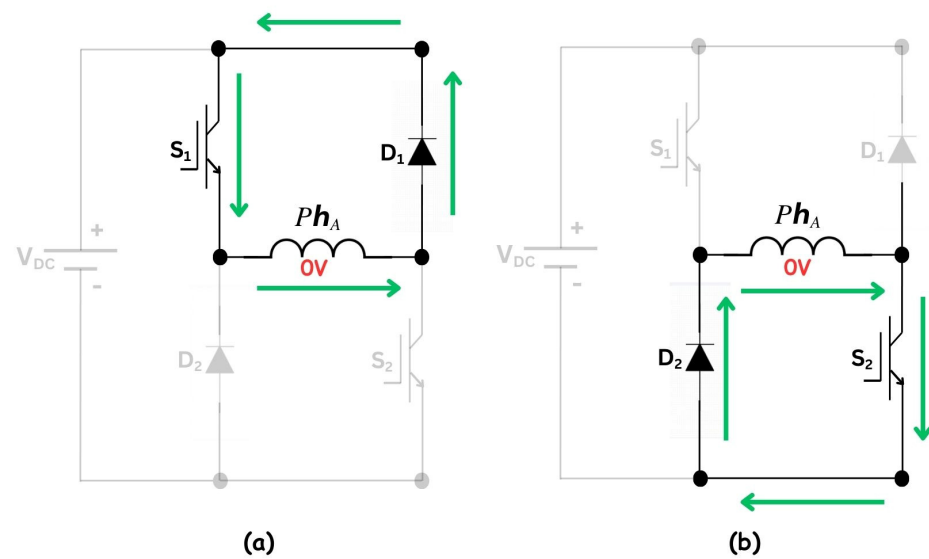


Figure 3. (a,b). Free-wheeling state of the ASHB converter.

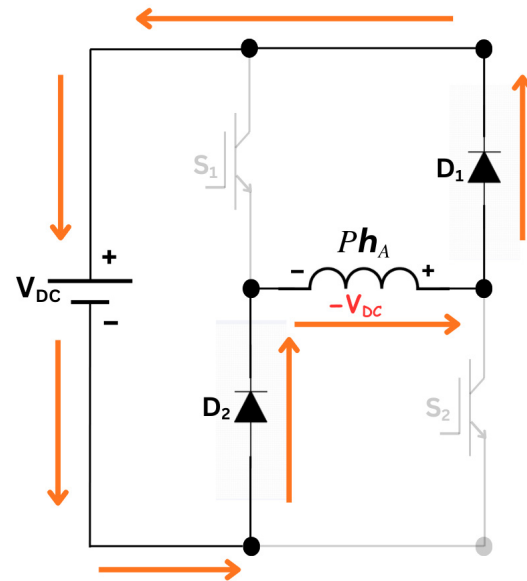


Figure 4. Demagnetization state of the ASHB converter.

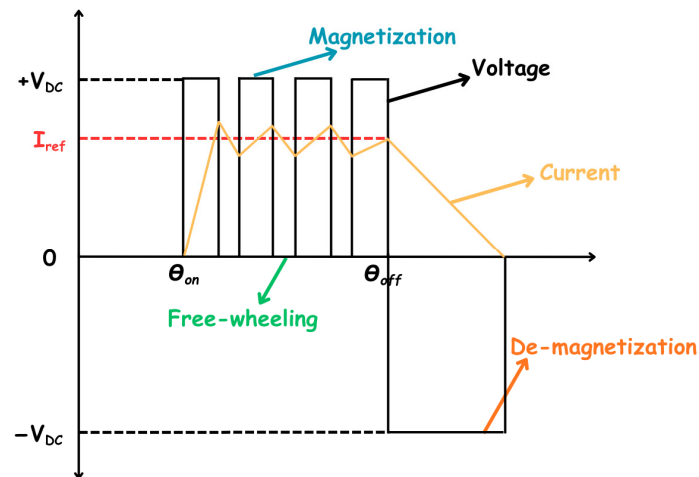


Figure 5. Phase voltage and current waveforms characteristics associated with an ASHB.

Magnetization ($+V_{DC}$): As shown in Figure 2, switches S_1 and S_2 of the phase to be energized are activated in this state. By applying the entire DC link voltage to the phase, this operation causes the current to increase.

Free-wheeling (0 V): For this voltage level, two switch patterns are feasible. The first configuration has S_1 turned on and S_2 turned off, whereas the second configuration has S_2 turned on and S_1 turned off. The patterns of zero voltage being placed across the phase winding are depicted in Figures 3a and 3b, respectively. Bridge losses and overheating can be balanced by switching between these two (0 V) states.

Demagnetization ($-V_{DC}$): As shown in Figure 4, switches S_1 and S_2 are simultaneously off in this state. Before the phase reaches the area of negative torque production, the demagnetization current from the winding of the motor is pumped back to the DC link. The motor's winding current is then decreased to zero.

This converter offers several advantages that make it suitable for SRM applications. It permits each step to be independently controlled and provides three distinct voltage levels, namely plus V_{DC} , 0 V, and $-V_{DC}$. The circuit has low complexity while maintaining high fault tolerance, since there is no possibility of a shoot-through condition in the link switches.

It also achieves high efficiency as the magnetic energy stored in the phase winding is returned to the DC link during demagnetization.

4. Proposed System

The proposed system introduces an adaptive, data-driven torque ripple minimization strategy for an 8/6 Switched Reluctance Motor (SRM) drive using a Deep Deterministic Policy Gradient (DDPG)-based Reinforcement Learning (RL) framework. Unlike conventional current profiling or firing angle lookup methods, this proposed controller autonomously learns the optimal phase excitation strategy by interacting directly with the SRM environment. The architecture employs a dual neural network framework, the actor and the critic, that cooperatively learn a continuous control policy capable of reducing torque ripple under varying speeds and load conditions.

Figure 6 illustrates the overall system architecture, which consists of three primary subsystems: the SRM drive module, the converter and gating controller, and the reinforcement learning agent. The SRM drive module represents the electromagnetic behavior of the 8/6 machine, capturing nonlinear magnetization, mutual inductance, and rotor position-dependent inductance profiles. An Asymmetrical Half-Bridge (ASHB) converter is used to energize each phase independently, allowing flexible excitation control through gating signals.

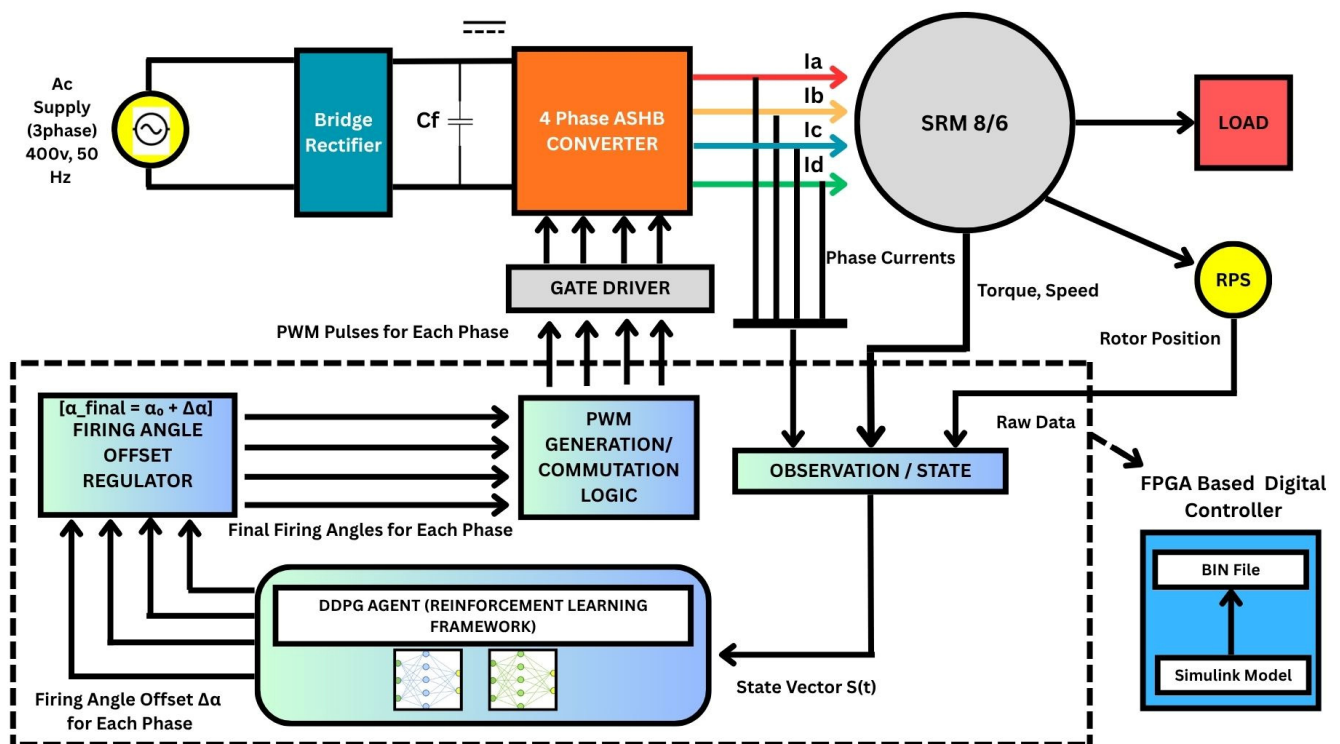


Figure 6. Overall system architecture of the proposed DDPG-based reinforcement learning framework for a torque ripple minimization scheme for an 8/6 SRM drive.

The block diagram illustrates the complete control architecture for an SRM drive system using a reinforcement learning based approach. The three-phase AC supply is rectified and fed to a four-phase Asymmetrical Half-Bridge converter, which drives the SRM; the gate driver generates PWM pulses for each phase of the converter. Rotor position, torque speed, phase currents, and flux linkages are sensed and fed to an observation block, which forms the state vector for the DDPG agent. The agent computes firing angle offsets, which are combined with base angles to generate the final switching signals for controlling

the converter; the entire control system is implemented using an FPGA-based Waveact digital controller.

4.1. Detailed Mathematical Model and Derivations

4.1.1. Electromagnetic Torque of SRM

For an 8/6 SRM with four independently excited phases $i \in \{A, B, C, D\}$, the electromagnetic torque produced by phase i is given by the standard energy-derived expression as in Equation (5):

$$T_{e,i}(\theta, t) = \frac{1}{2} i_i^2(t) \frac{dL_i(\theta)}{d\theta} \quad (5)$$

where $i_i(t)$ denotes the instantaneous phase current (A), $L_i(\theta)$ represents the phase inductance as a function of the rotor electrical angle θ and the derivative $\frac{dL_i}{d\theta}$ captures the change in inductance with respect to rotor position, characterizing the position-dependent conversion of magnetic energy into torque.

The net electromagnetic torque is obtained by superimposing the individual torque contributions of all excited phases and is given by Equation (6):

$$T_e(\theta, t) = \sum_{i \in \{A, B, C, D\}} T_{e,i}(\theta, t) = \frac{1}{2} \sum_i i_i^2(t) \frac{dL_i(\theta)}{d\theta} \quad (6)$$

Equation (5) is a standard relation of the SRM. It shows that torque depends quadratically on current and linearly on the slope of inductance with respect to rotor angle.

4.1.2. Phase Current Dynamics During Conduction

Each phase circuit can be modeled as in Equation (7):

$$V_s(t) = R_i i_i(t) + \frac{d}{dt} (L_i(\theta(t)) i_i(t)) \quad (7)$$

where $V_s(t)$ is the applied phase voltage during conduction (either supply voltage or zero, depending on gating) and R_i is the phase resistance. Rewriting the derivative, as in Equation (8),

$$\frac{d}{dt} (L_i(\theta(t)) i_i(t)) = L_i(\theta) \frac{di_i(t)}{dt} + i_i(t) \dot{\theta} \frac{dL_i(\theta)}{d\theta} \quad (8)$$

Substituting (8) in (7) gives Equation (9):

$$V_s(t) = R_i i_i(t) + L_i(\theta) \frac{di_i(t)}{dt} + i_i(t) \dot{\theta} \frac{dL_i(\theta)}{d\theta} \quad (9)$$

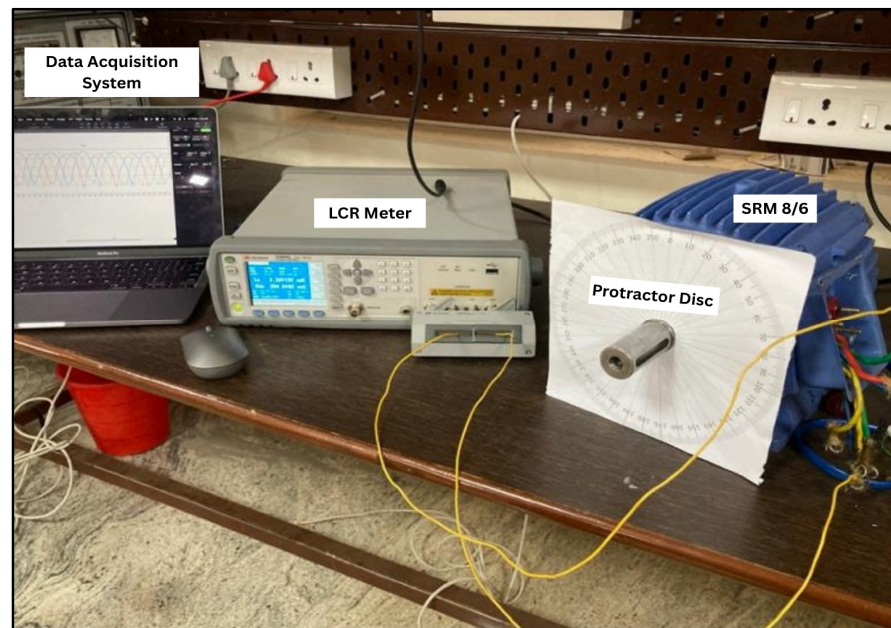
Equation (9) is the electrical differential equation of each phase that governs current evolution while voltage is applied and the rotor moves. The θ term couples mechanical motion into the electrical circuit.

4.2. Role of Inductance Profile in RL-Based Optimal Phase Excitation

The inductance profile of an 8/6 Switched Reluctance Motor (SRM) plays a critical role in the proposed Reinforcement Learning (RL)-based optimal phase excitation strategy. In this work, the phase inductance variation with rotor position is experimentally obtained and incorporated into the RL control framework to enable informed and adaptive commutation decisions by the DDPG agent.

The experimental arrangement used for inductance measurement is shown in Figure 7a, which includes a precision LCR meter, the SRM under test, and a manual rotor indexing mechanism. During the measurement process, each phase winding of the SRM is individually connected to the LCR meter, while the remaining phases are left unex-

cited. The rotor shaft is manually rotated in discrete angular steps of 10 degrees, and the corresponding phase inductance values are recorded at each position. The inductance and resistance values displayed on the LCR meter during measurement are shown in Figure 7b.



(a)



(b)



(c)

Figure 7. (a) Overall experimental setup for inductance profile 8/6 Switched Reluctance Motor. (b) LCR meter display showing the measured phase inductance and winding resistance of the 8/6 SRM at a fixed rotor position during experimental characterization. (c) Experimental setup for encoder response measurement of the 8/6 Switched Reluctance Motor.

Accurate rotor position reference during inductance measurement is provided using an incremental encoder mounted at the non-driving end of the motor shaft. The encoder is powered by a regulated 5 V DC supply, and the encoder output voltage varies with rotor position as shown in Figure 7c. The encoder response is recorded simultaneously with the inductance measurements to establish a precise mapping between rotor position and phase inductance.

The experimentally obtained encoder voltage responses are shown in Figure 8a. These results clearly identify the aligned and unaligned inductance regions for each phase and demonstrate the periodic variation in inductance with rotor position. The detailed phase-wise inductance profile over one complete electrical cycle, obtained through discrete rotor stepping, is presented in Figure 8b.

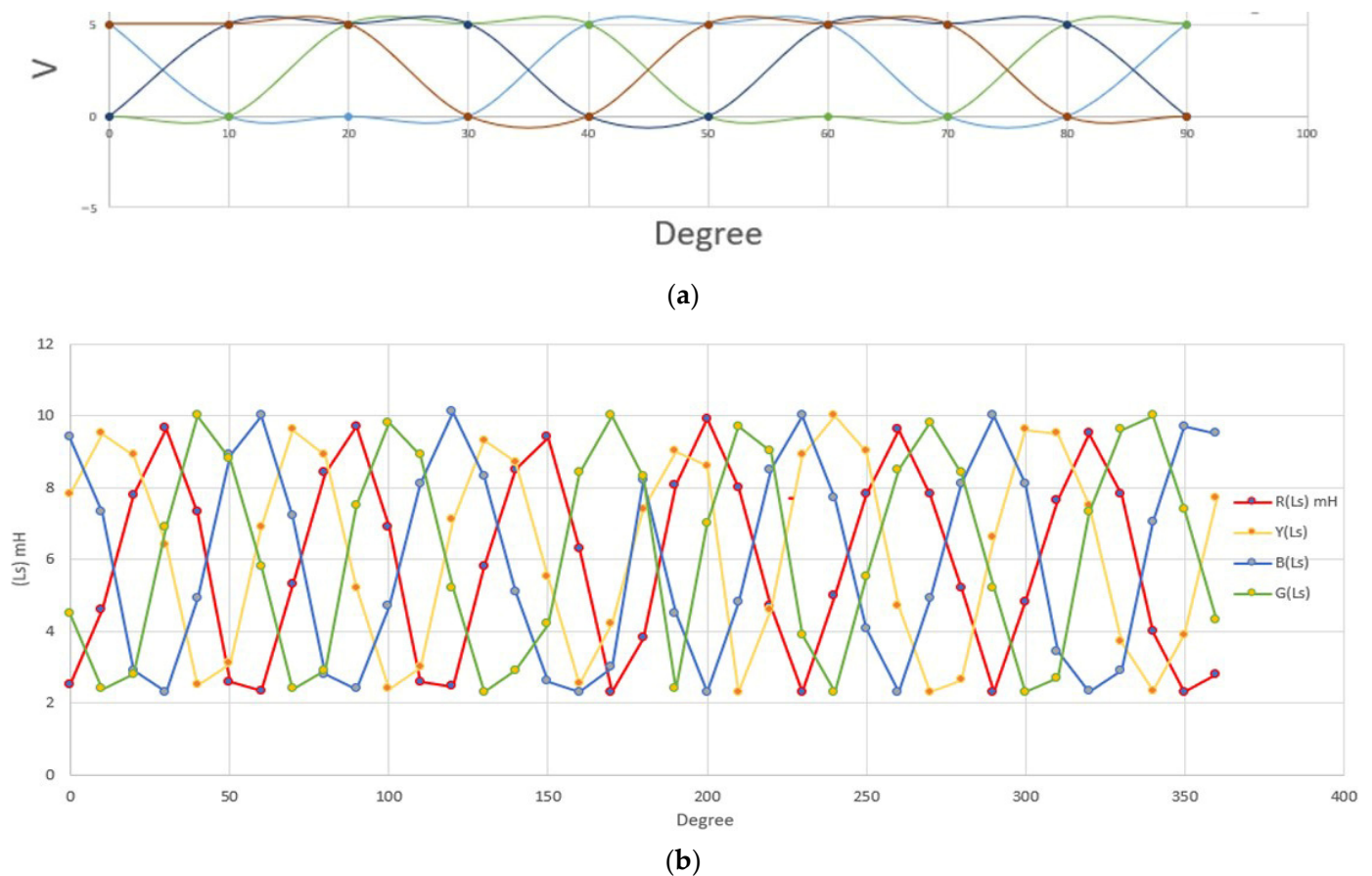


Figure 8. (a) Experimentally obtained encoder voltage responses as a function of rotor position. (b) Measured inductance variation in SRM phases over a full electrical cycle obtained through discrete rotor position stepping.

The measured inductance profiles are incorporated into the reinforcement learning environment as a reference for identifying high-torque operating regions. In SRMs, electromagnetic torque is maximized when phase current is applied in the region of increasing inductance (positive $\frac{dL(\theta)}{d\theta}$) and minimized when current excitation is avoided in the decreasing inductance region. By utilizing the experimentally obtained inductance map along with encoder feedback, the DDPG-based actor–critic controller learns optimal firing angle offsets and dwell durations, ensuring that current excitation occurs predominantly in the rising inductance region for different speeds and loads.

As a result, the inductance profile serves as a key enabler of the proposed RL-based control strategy, allowing the agent to implicitly learn the electromagnetic characteristics of

the SRM without relying on complex analytical torque models. This approach results in smooth torque production, effective torque ripple reduction, and improved overall drive performance, thereby demonstrating the practical relevance of incorporating experimentally measured inductance data into the control framework.

5. Reinforcement Learning Framework for Torque Ripple Minimization in SRM

5.1. Overview of Reinforcement Learning

A subfield of machine learning called reinforcement learning (RL) allows an agent to interact with its surroundings and discover the best control plan of action [21]. In contrast to supervised learning, RL does not require pre-labeled data; instead, the agent explores various actions and learns from rewards received as feedback. This makes RL particularly suitable for systems with complex, nonlinear, and poorly modeled dynamics. In the torque ripple minimization, the RL agent learns how to adjust the firing angle offsets of each SRM phase in real time by observing the system state (current, torque, rotor speed, rotor position, and flux linkages) and optimizing the long-term cumulative reward associated with smoother torque generation.

An RL process is formally modeled as a Markov Decision Process (MDP), represented by the tuple as given in Equation (10):

$$\mathcal{M} = (\mathcal{S}, \mathcal{A}, P, R, \gamma) \quad (10)$$

The Markov Decision Process (MDP) \mathcal{M} formulation for the SRM control problem is defined as follows. The state space \mathcal{S} represents the set of all possible motor conditions, including phase currents i_i , electromagnetic torque T_e , flux linkage ψ , rotor speed ω , and rotor position (two scalars). The action space \mathcal{A} consists of continuous control inputs, expressed as phase firing angle offsets $\Delta\theta = [\Delta\theta_A, \Delta\theta_B, \Delta\theta_C, \Delta\theta_D]$. The state transition model $P(s_{k+1} | s_k, a_k)$ describes the probabilistic evolution of the motor states under a given control action. The reward function $R(s_k, a_k)$ assesses the immediate performance, which is formulated to penalize torque ripple or current distortion. A discount factor $\gamma \in (0, 1)$ is used to weigh future rewards relative to present outcomes, ensuring stable long-term optimization.

The objective of RL is to determine an optimal policy π^* that maximizes the expected cumulative discounted reward, as represented in Equation (11):

$$J(\pi) = \mathbb{E}_{\pi} \left[\sum_{k=0}^{\infty} \gamma^k r(s_k, a_k) \right] \quad (11)$$

The Q-function or state-action value function is the expected cumulative reward starting from the state s_k , followed by the action a_k and thereafter following policy π , as given in Equation (12):

$$Q^{\pi}(s_k, a_k) = \mathbb{E}_{\pi} \left[\sum_{t=k}^{\infty} \gamma^{t-k} r(s_t, a_t) \mid s_k, a_k \right] \quad (12)$$

The optimal policy satisfies the Bellman optimality condition shown in Equation (13):

$$Q^*(s, a) = R(s, a) + \gamma \max_{a'} Q^*(s', a') \quad (13)$$

In torque ripple minimization, $R(s, a)$ penalizes deviations from desired torque smoothness, hence the agent learns to select actions (firing angle offsets) that reduce ripple amplitude in the long run.

Traditional torque ripple reduction methods, such as current profiling and PI-based angle tuning, require precise mathematical modeling of inductance variation and magnetic saturation. However, SRM characteristics are highly nonlinear and vary with operating conditions such as temperature, speed, and load. Reinforcement learning overcomes this by learning directly from experience [22,23], mapping observations to actions without requiring analytical models [24]. Figure 9 illustrates the Reinforcement learning framework proposed for the SRM control scheme. The SRM and its power converter form the environment, while a DDPG-based RL agent functions as the controller. The agent receives motor feedback states, including phase currents, torque, flux linkage, rotor speed, and rotor position (two scalars), and then, the agent generates control actions in the form of firing angle adjustments or gate signal commands. The reward function evaluates the quality of the control action based on torque ripple and updates the policy accordingly. Using the feedback and reward information, the RL agent continuously adjusts the firing angles to achieve reduced torque ripple and improved drive performance.

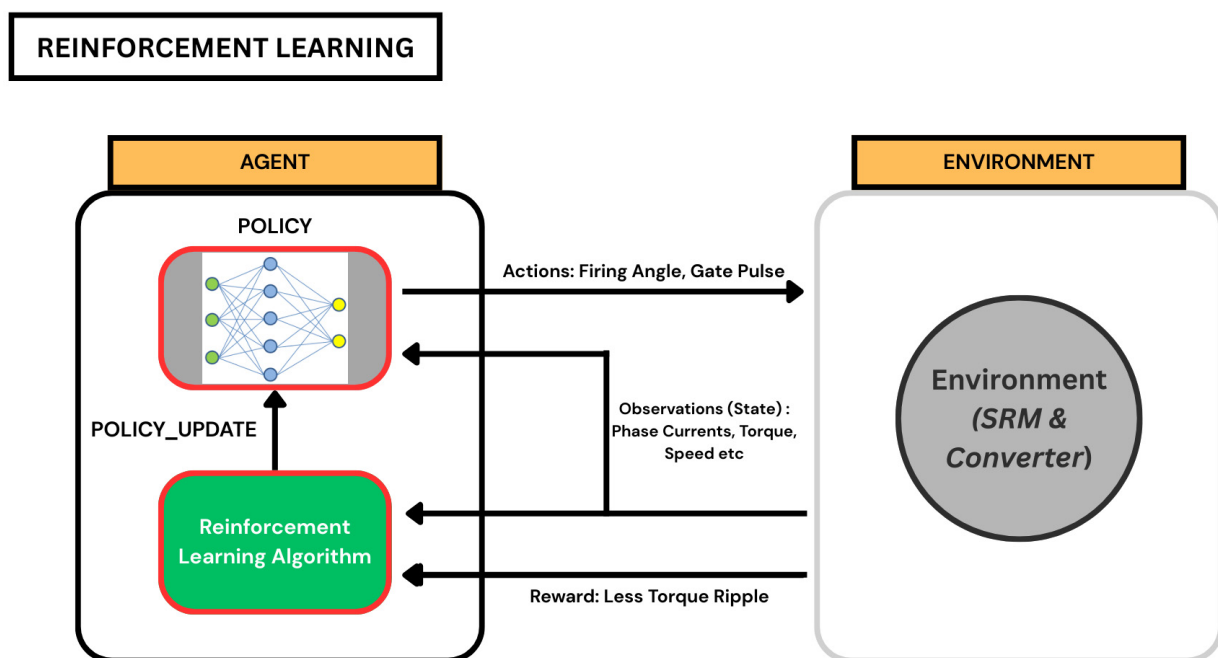


Figure 9. Reinforcement learning framework for the proposed SRM control scheme.

Through repeated interaction, the RL agent learns to shape excitation intervals such that the sum of individual phase torques yields a smooth total torque waveform.

5.2. Deep Deterministic Policy Gradient (DDPG) Framework

A model-free, off-policy reinforcement learning technique for continuous action spaces is the Deep Deterministic Policy Gradient (DDPG) algorithm. Using deep neural networks to approximate the policy and value functions, it combines the ideas of Deterministic Policy Gradient (DPG) and Deep Q-Learning, unlike discrete-action algorithms (e.g., Q-Learning), which are unsuitable for fine-grained control tasks like angle tuning. DDPG can output continuous values, such as firing angle shifts with sub-degree precision. The Deep Deterministic Policy Gradient (DDPG) framework consists of several key components. The actor network $\mu(s | \theta^\mu)$ outputs deterministic continuous control actions based on the current system states and the critic network $Q(s, a | \theta^Q)$ evaluates these actions by estimating the expected long-term reward for every state-action pair. For stable learning, DDPG utilizes the target networks μ' and Q' , which are slowly updated versions of the actor and critic networks, respectively. Further, a replay buffer \mathcal{D} stores previously experienced

transitions (s_k, a_k, r_k, s_{k+1}) , allowing for random experience sampling and diminishing temporal correlations while training.

Figure 10 shows the internal learning mechanism of the DDPG algorithm; it merges policy-based and value-based reinforcement learning for the control of continuous systems, such as the SRM drive. The actor network receives the current system state s and produces an action a , which in this case corresponds to the firing angle adjustment applied to the converter. The critic network evaluates how good this action is in terms of performance by computing the Q-value, which represents the expected reward (lower torque ripple). Both networks have corresponding target networks (target actor and target critic) that are updated gradually through soft updates to prevent instability and oscillation during learning. The critic network is trained using a loss function computed from the difference between predicted and target Q-values, while the actor network is updated via policy gradients to improve its decision-making ability. Over time, this structure allows the agent to learn optimal firing angle strategies that yield minimum torque ripple in the SRM.

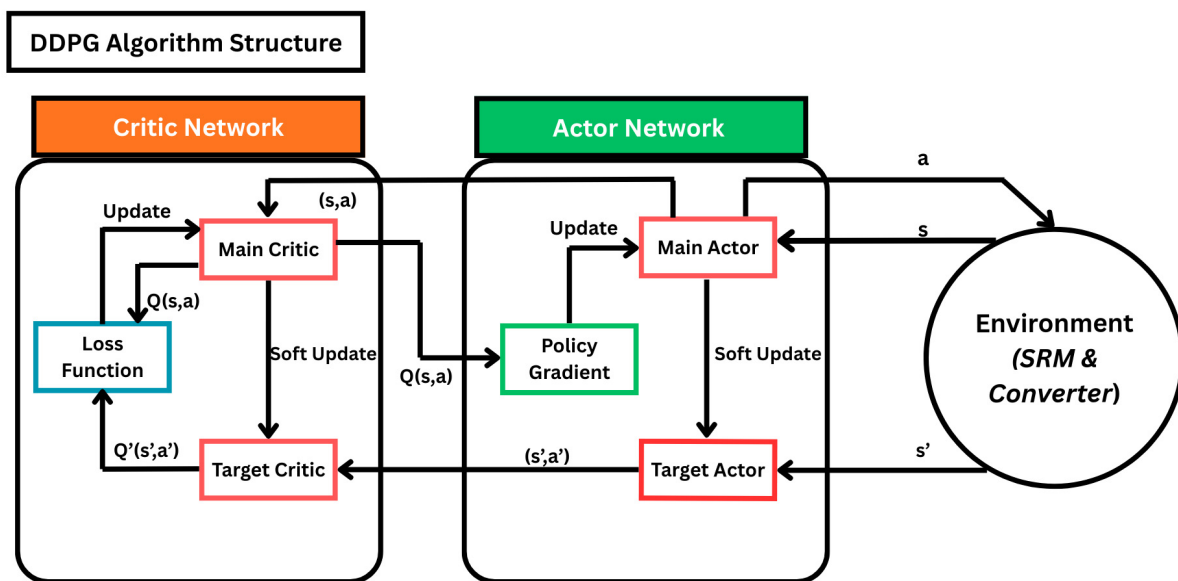


Figure 10. Structural representation of the Deep Deterministic Policy Gradient (DDPG) algorithm used in the proposed SRM control system.

The deterministic actor policy generates a continuous control action according to Equation (14):

$$a_k = \mu(s_k | \theta^\mu) \tag{14}$$

The critic network is trained using the Bellman target, defined in Equation (15):

$$y_k = r_k + \gamma Q'(s_{k+1}, \mu'(s_{k+1} | \theta^{\mu'}) | \theta^{Q'}) \tag{15}$$

The actor network is updated using the Deterministic Policy Gradient (DPG) theorem, where the policy gradient is computed as in Equation (16):

$$\nabla_{\theta^\mu} J \approx \frac{1}{N} \sum_{k=1}^N \left[\nabla_a Q(s, a | \theta^Q) \Big|_{a=\mu(s)} \cdot \nabla_{\theta^\mu} \mu(s | \theta^\mu) \right] \tag{16}$$

The critic network parameters are updated by minimizing the mean-squared Temporal Difference (TD) error using mini-batch samples randomly drawn from the replay buffer. To improve training stability, slowly updated target networks are employed for both the actor

and critic. These target networks are updated using soft updates with factor τ , following the standard DDPG framework.

The Deep Deterministic Policy Gradient (DDPG) is effective in solving complex non-linear control problems in real time [25–27], like torque ripple minimization in SRMs, due to the following reasons: it allows continuous control by managing continuous SRM phase excitation angles, its model-free learning completely removes the need for detailed magnetic or torque models it incorporates temporally correlated noise via the Ornstein–Uhlenbeck process for efficient exploration, but in a way to generate smooth movement within the control space, its dual-network stability due to actor–critic separation ensures stable convergence even in the presence of noisy torque feedback and finally, the learned policy exhibits strong adaptability, generalizing effectively across a wide range of load and speed variations [28,29].

5.3. Critic Neural Network

The critic network, as a value estimator, evaluates the quality of action taken [30,31]. It learns the function $Q(s, a | \theta^Q)$, which predicts the expected cumulative reward, effectively quantifying how each firing angle configuration impacts long-term torque ripple. In the SRM, the critic observes the following, as represented in Equation (17):

$$s = [i_A, i_B, i_C, i_D, T_e, \omega, \theta, \psi, \text{lastAction}] \quad (17)$$

and evaluates the predicted reward $Q(s, a)$ for an action vector $a = [\Delta\theta_A, \Delta\theta_B, \Delta\theta_C, \Delta\theta_D]$.

The critic network is designed to approximate the Bellman optimality, as shown in Equation (18):

$$Q^*(s, a) = \mathbb{E}_{r, s'} [r(s, a) + \gamma \max_{a'} Q^*(s', a')] \quad (18)$$

In DDPG, the true Q-function is approximated using a parameterized critic network $Q(s, a | \theta^Q)$; Equation (19) defines the mean-squared Temporal Difference (TD) error, which is minimized to update the parameters.

$$L_Q(\theta^Q) = \mathbb{E}_{(s, a, r, s') \sim D} \left[\left(Q(s, a | \theta^Q) - y \right)^2 \right] \quad (19)$$

where the target value y is computed using the slowly updated target networks, as expressed in Equation (20):

$$y = r + \gamma Q'(s', \mu'(s' | \theta^{\mu'}) | \theta^{Q'}) \quad (20)$$

These formulations ensure stable critic learning while maintaining consistency with the Bellman equation.

The critic network uses a neural architecture in which the input layer receives the concatenated motor state vector (approximately 13 dimensions) and the corresponding four-dimensional action vector. Following this is a structure of two fully connected hidden layers comprising about 128–256 neurons each, using ReLU for the introduction of nonlinearity. The forward propagation through each layer is defined by the linear transformation given in Equation (21):

$$z^{(l)} = W^{(l)} x^{(l-1)} + b^{(l)} \quad (21)$$

where $W^{(l)}$ and $b^{(l)}$ denote the weight matrix and bias vector of the l^{th} layer, respectively. The output of each hidden layer is then obtained by applying a nonlinear activation function, as expressed in Equation (22):

$$x^{(l)} = f(z^{(l)}) \quad (22)$$

with $f(\cdot)$ being the ReLU activation for the hidden layers. The final output layer is linear and produces the scalar estimate $\hat{Q}(s, a)$, corresponding to the network’s approximation of the Q-function, expressed in Equation (23):

$$\hat{Q}(s, a) = f_L(f_{L-1}(\dots f_1([s, a]) \dots)) \tag{23}$$

Figure 11 shows that the critic takes both the 13-dimensional state vector and the 4-dimensional action vector, which are concatenated into a 17-dimensional input. Then, these concatenated features are fed into two ReLU-activated hidden layers comprising 256 and 128 neurons, respectively. The output layer is a single linear scalar $Q(s, a)$, for predicting the long-term reward to be received by applying an action a in the state s . This value will provide the gradient information required to update the firing angle policy toward minimum torque ripple.

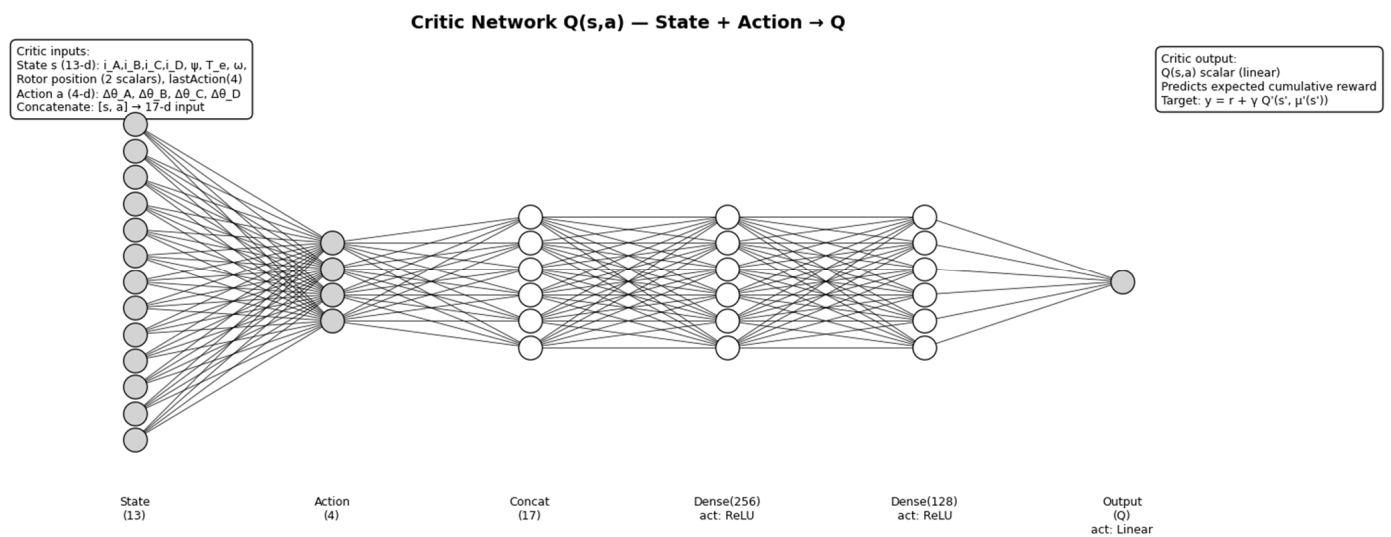


Figure 11. Critic neural network architecture, employed in the DDPG-based SRM torque ripple minimization controller.

Through this structure, the critic learns incrementally to estimate the long-term value of each action under varying operating conditions of the SRM. Actions that result in smoother torque and reduced torque ripple yield higher predicted Q-values, thereby encouraging the actor to generate similar firing angle offsets. The critic constructs a coherent representation of the torque smoothness landscape by minimizing the error between the predicted and target Q-values, thereby associating firing decisions with torque ripple performance and providing the gradient information required for actor optimization.

5.4. Actor Neural Network

The actor network serves as the policy generator with continuous-valued actions corresponding to the firing angle offsets that try to maximize the expected Q-value estimated by the critic. Formally, the deterministic policy is defined in Equation (24):

$$a = \mu(s | \theta^\mu) \tag{24}$$

where $a \in \mathbb{R}^4$ represents the four phase-offset actions and $s \in \mathbb{R}^{13}$ denotes the motor state vector. The actor is trained to maximize the critic’s evaluation of the selected actions. Accordingly, the policy objective is expressed as in Equation (25):

$$J(\theta^\mu) = \mathbb{E}_{s \sim D} [Q(s, \mu(s | \theta^\mu) | \theta^Q)] \tag{25}$$

Using the chain rule, the policy gradient needed to update the actor parameters is provided in Equation (26):

$$\nabla_{\theta^\mu} J = \mathbb{E}_{s \sim D} \left[\nabla_a Q(s, a | \theta^Q) \Big|_{a=\mu(s)} \nabla_{\theta^\mu} \mu(s | \theta^\mu) \right] \quad (26)$$

This is the gradient that drives the actor to produce actions that lead to higher predicted Q-values, meaning smoother torque with reduced ripple.

The state vector $s = [i_A, i_B, i_C, i_D, T_e, \omega, \theta, \psi, \text{lastAction}]$ is fed to the actor network architecture, then two fully connected hidden layers are applied with 128–256 neurons and ReLU activation. The output layer contains four neurons, one for each phase firing offset, using a tanh activation in order to bound the values within $[-1, 1]$. The phase offsets, after such a normalization, are mapped into physically allowable angles using Equation (27):

$$\Delta\theta_i = \Delta\theta_{\max} \cdot \frac{a_i + 1}{2} \quad (27)$$

where $\Delta\theta_{\max}$ is the maximum permissible deviation for each phase. The activation of each neuron in layer l is computed according to Equation (28):

$$x_j^{(l)} = f\left(\sum_i W_{ji}^{(l)} x_i^{(l-1)} + b_j^{(l)}\right) \quad (28)$$

and the corresponding weights and biases are updated with backpropagation driven by the policy gradient in Equation (26).

Figure 12 shows that the actor receives a 13-dimensional state vector consisting of the phase currents (i_A, i_B, i_C, i_D), flux linkage (ψ), electromagnetic torque (T_e), rotor speed (ω), rotor position represented using two scalars, and the previous firing angle offsets ($\Delta\theta_{A,B,C,D}^{\text{prev}}$). The network consists of three fully connected hidden layers (256, 128, and 64 neurons) using ReLU activation. The output layer is composed of four neurons with tanh activation, creating the normalized firing angle offsets $[\Delta\theta_A, \Delta\theta_B, \Delta\theta_C, \Delta\theta_D]$, which are then mapped to the physically admissible conduction-angle limits of the SRM.

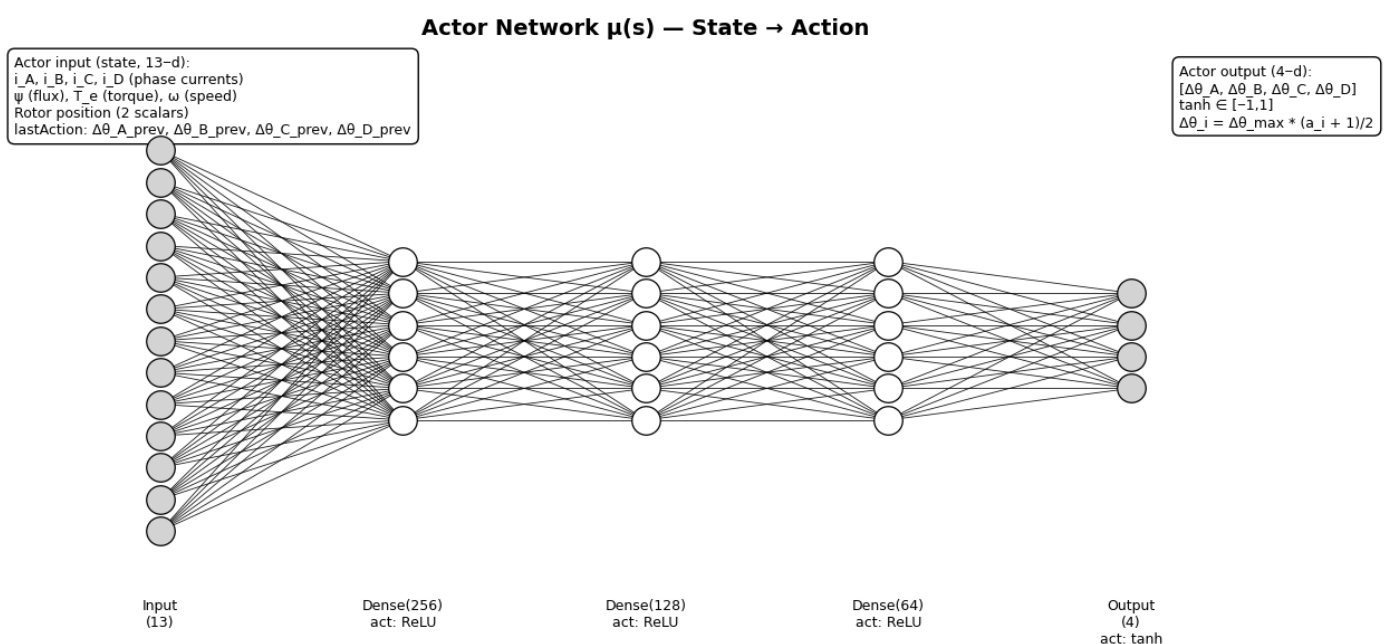


Figure 12. Actor neural network architecture, employed in the DDPG-based SRM torque ripple minimization controller.

In the proposed torque ripple minimization framework, the actor receives real-time torque and phase current information as its state input and generates optimized firing angle timing offsets for the four SRM phases. These offsets are incorporated into the gating signals of the ASHB converter to reshape the phase current profiles. Through continuous interaction with the environment over successive training episodes, the actor progressively converges to firing patterns that improve torque smoothness and effectively reduce torque ripple.

The actor fosters exploration at the beginning of training by generating stochastic actions through noise injection. Mathematically, this can be expressed as in Equation (29):

$$a_t = \mu(s_t | \theta^\mu) + \mathcal{N}_t \tag{29}$$

where \mathcal{N}_t is Ornstein–Uhlenbeck noise, typically used in DDPG for temporally correlated exploration in continuous actions. As training progresses and the critic network becomes more accurate, the magnitude of exploration noise is gradually reduced, allowing the actor to converge toward a stable, deterministic policy. Figure 13 illustrates the learning process of the DDPG agent, which enables the proposed SRM controller to autonomously determine the optimal firing angle offset. The current state data s_t is sent to the actor network and outputs a continuous action a_t , which represents the firing angle adjustment. To encourage exploration during training, small random noise is added to this action.

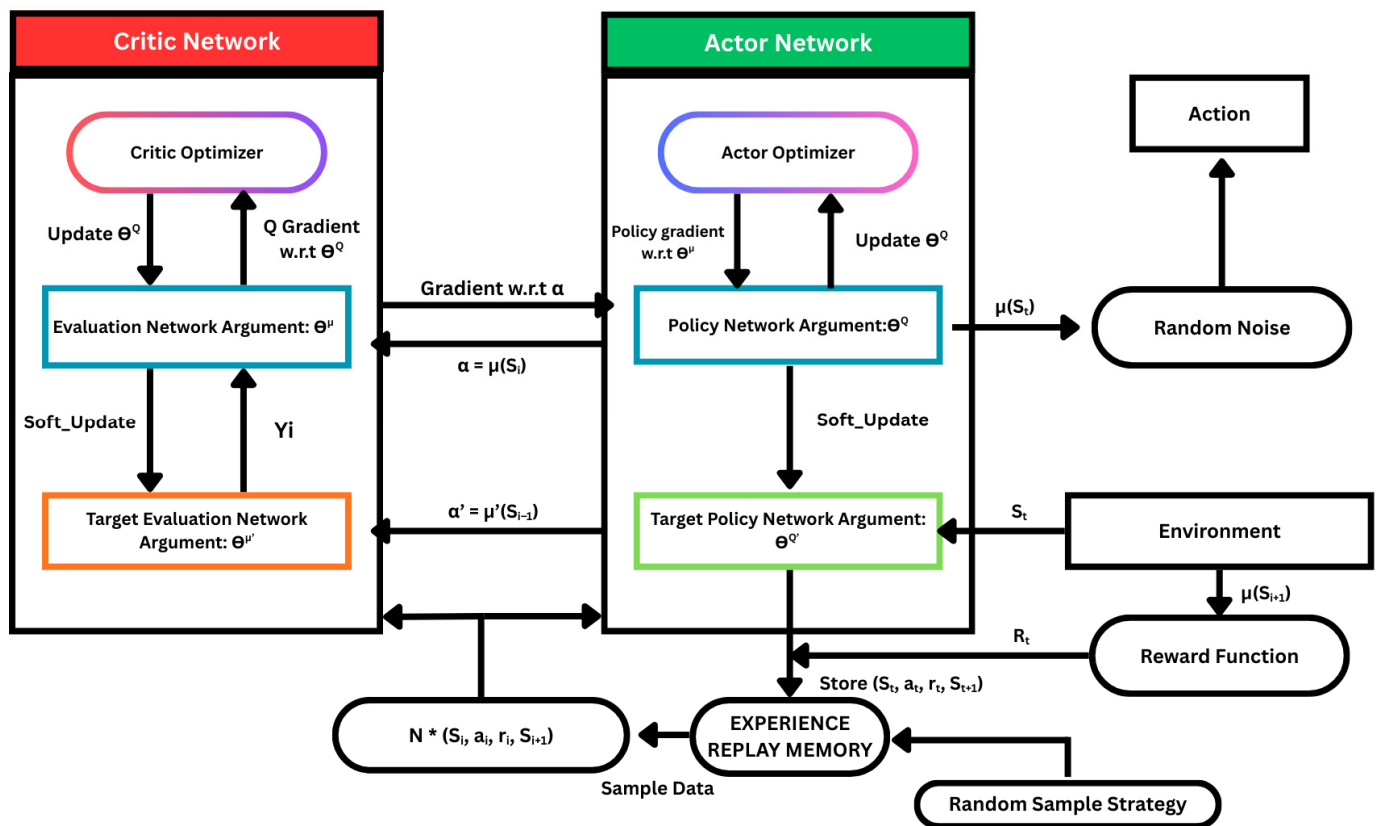


Figure 13. Learning policy flow of the Deep Deterministic Policy Gradient (DDPG) algorithm.

The environment, consisting of the SRM and its power converter, responds to the applied action and generates the next state S_{t+1} and the reward R_t , which reflects the torque ripple performance. The resulting experience tuples (S_t, a_t, R_t, S_{t+1}) are stored in an experience replay memory, from which the agent samples batches to update the learning networks. The critic network evaluates the effectiveness of the selected action by estimating

the Q-value, while the actor network is updated using policy gradients derived from the Critic's evaluation. To ensure stable learning, target actor and target critic networks are employed and updated through soft parameter updates using slow, incremental blending. Through repeated interaction and learning, the DDPG agent refines its control strategy by minimizing torque ripple, even in the absence of any explicit mathematical model of the SRM. In this manner, the actor becomes an adaptive timing generator. It continuously optimizes the phase excitation angles based on the observed torque feedback. This actor, while experiencing an oscillatory torque waveform at the rotor, learns to provide phase advancement or delay to specific phase currents in order to achieve an optimal interlocking between adjacent phase torque profiles. This learned synchronization of phase torque reduces the net torque ripple without using analytical current profiling or lookup tables.

Table 2 shows the heuristic learning behavior interpretations used to guide the calibration of training hyperparameters in the DDPG-based SRM control framework. The properties summarized in this table provide qualitative insight into the interaction between actor–critic learning dynamics, time-scale separation, and torque ripple reduction. These interpretations support the selection of hyperparameters and help explain the observed bounded and smooth learning behavior. The effectiveness of these choices is validated through simulation and experimental results.

Table 2. Heuristic learning-behavior interpretations used for calibration of the DDPG-based SRM control framework.

<i>Property</i>	<i>Heuristic Interpretation/Calibration Rationale</i>	<i>Practical Meaning in SRM Control</i>
<i>Critic Learning Behavior</i>	A high discount factor ($\gamma = 0.99$) and a moderate critic learning rate (2×10^{-4}) are selected to promote numerically stable Bellman updates and bounded Q-function estimation. A large replay buffer (10^6 samples) and mini-batch training (batch size = 128) reduce temporal correlation in torque ripple data.	Prevents divergence in value estimation and stabilizes the learning of the torque ripple cost landscape during training.
<i>Actor Learning Smoothness</i>	A smaller actor learning rate (1×10^{-4}) relative to the critic, together with a physically bounded action space for firing angle offsets, promotes gradual policy updates.	Ensures smooth adaptation of phase firing angle offsets, thereby avoiding excessive torque ripple caused by abrupt or aggressive commutation.
<i>Actor–critic Time-Scale Separation</i>	A two-time-scale learning strategy is achieved by faster critic updates relative to actor updates, reinforced through a soft target network update rate ($\tau = 0.005$).	Maintains stable interaction between value estimation and policy improvement throughout the learning process.
<i>Torque Ripple Energy Reduction (Empirical Indicator)</i>	Reward shaping penalizes peak-to-peak torque ripple, while bounded exploration using Ornstein–Uhlenbeck noise ($\mu = 0$, $\sigma = 0.25$) encourages local exploration.	Produces progressively smoother electromagnetic torque and reduced torque ripple in closed-loop SRM operation.

Table 2 provides heuristic interpretations used to guide the calibration of training hyperparameters and does not imply formal stability or convergence guarantees.

6. Simulation Setup and Training Process

6.1. Simulation Setup of DDPG-Based Actor–Critic Neural Network Reinforcement Learning Framework for Minimizing the Torque Ripple in SRM

The proposed Deep Deterministic Policy Gradient (DDPG)-based reinforcement learning control strategy was implemented and trained in the MATLAB Simulink environment. The complete simulation architecture is shown in Figure 14. The setup integrates a four-

phase Asymmetrical Half-Bridge (ASHB) converter, an 8/6 Switched Reluctance Motor (SRM), and a DDPG-based reinforcement learning agent that governs the gate pulse generation through an adaptive Firing Angle Offset Regulator. The DDPG framework consists of dual neural networks (actor and critic) that interact with the environment in a continuous control setting. The actor network maps observed motor states such as phase currents, torque, flux linkage, rotor speed, and rotor position (two scalars) to optimal action values corresponding to firing angles and current references, while the critic network evaluates the expected return (reward) associated with each state-action pair. The reward function is defined to penalize high torque ripple, large control variations, and deviations from reference torque, while providing positive reinforcement for smooth, stable operation and low ripple percentages.

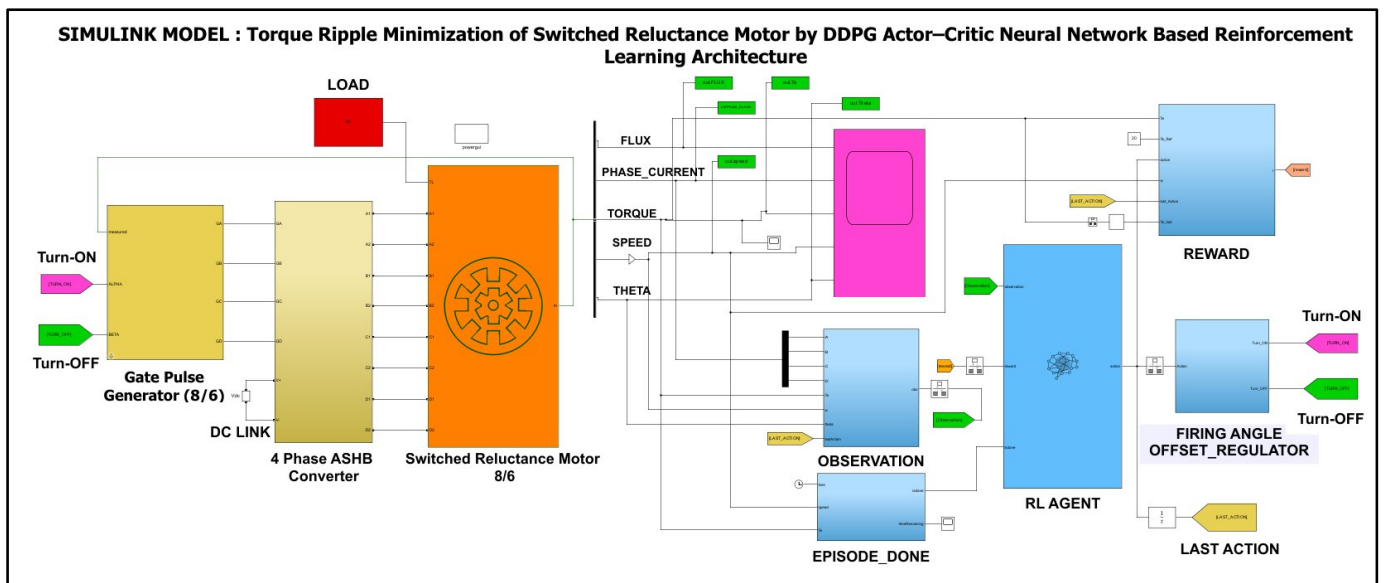


Figure 14. Simulink model of the proposed DDPG-based reinforcement learning control architecture for torque ripple minimization in a Switched Reluctance Motor.

Table 3 summarizes the SRM drive configuration, which was modeled as an 8/6 machine operating with a 220 V DC link supplied through an ASHB converter in MATLAB Simulink. Each phase is excited sequentially using a gate pulse generator, whose switching instants are adaptively adjusted by the trained RL agent. The observation block collects real-time parameters, T_e (electromagnetic torque), ω (rotor speed), rotor position (two scalars), ψ (flux linkage), $\Delta\theta_{A,B,C,D}^{prev}$ (previous firing angle offsets), and i_A, i_B, i_C, i_D (phase currents). These signals are processed to construct the state vector used for updating the agent. The reward block computes the instantaneous reward based on torque ripple and tracking error, establishing a direct relationship in which smoother torque production results in higher reward values.

Table 3. Simulation parameters of the 8/6 Switched Reluctance Motor drive system in MATLAB SIMULINK.

Parameter	Value/Range	Unit	Description
SRM Type	8/6 (4-Phase)	-	Switched Reluctance Motor Configuration
Rated Power	75	kW	Nominal Power Output
DC Link Voltage	220	V	Input DC Bus Voltage to ASHB Converter

Table 3. *Cont.*

<i>Parameter</i>	<i>Value/Range</i>	<i>Unit</i>	<i>Description</i>
<i>Simulated Speed Range</i>	1000–3000	RPM	Speed Range simulated in Simulink
<i>Load Torque</i>	20	Nm	Applied Load Torque
<i>Simulated Phase Current</i>	5–10	A	Current range used in simulation
<i>Aligned Inductance</i>	55	mH	Inductance at Aligned Position
<i>Unaligned Inductance</i>	12	mH	Inductance at Unaligned Position
<i>Maximum Flux Linkage</i>	1.5	Wb	Per-Phase Maximum Flux Linkage

6.2. Training Process

Training was performed in an episodic manner, where each episode corresponds to one complete operational cycle of the SRM. In the initial phase of exploration, the agent executes random actions within the safe operating limits to learn about the nonlinear torque-current dynamics of the SRM. Over successive episodes, the actor–critic pair updates their policy and value function via gradient descent, guided by experience replay and soft target updates. The network was thus trained until convergence, where the torque ripple and cumulative reward attained a steady level at the plateau, showing consistent learning and stable torque regulation. The robustness of the learned policy was tested under different load torque conditions. The system ensured near-constant torque and smooth flux linkage profiles. Thus, the proposed approach using RL-based control can effectively learn the complex switching and magnetic saturation dynamics of the SRM without explicit current profiling or pre-defined lookup tables. The DDPG-based control architecture training was performed in the MATLAB Simulink environment, within the hyperparameter settings as given in Table 4. These parameters include the learning rates of actor and critic networks, the discount factor, target network update rate, replay buffer size, mini-batch size, exploration noise model, and training episode configuration.

Table 4. Training hyperparameters adopted in the DDPG-based actor–critic learning framework for torque ripple minimization.

<i>Parameter</i>	<i>Value/Setting</i>
<i>RL Algorithm</i>	DDPG (Actor–Critic)
<i>Learning Rate (Actor Network)</i>	1×10^{-4}
<i>Learning Rate (Critic Network)</i>	2×10^{-4}
<i>Discount Factor (γ)</i>	0.99
<i>Target Network Update Rate (τ)</i>	0.005
<i>Replay Buffer Size</i>	1×10^6 samples
<i>Mini-Batch Size</i>	128
<i>Exploration Noise Model</i>	Ornstein–Uhlenbeck Noise ($\mu = 0, \sigma = 0.25$)
<i>Maximum Training Episodes</i>	500 Episodes
<i>Episode Length</i>	1 s (Simulation Time)

The structural configuration of the Actor–Critic Neural Networks adopted in the proposed RL framework is summarized in Table 5. This table details the layer dimensions, activation functions, and output specifications used for both the actor and critic pathways.

Table 5. Actor–Critic Neural Network configuration used for continuous control learning.

<i>Network</i>	<i>Layer Structure</i>	<i>Activation Function</i>	<i>Output Dimension</i>
<i>Actor Network</i>	Input Layer → Dense (256) → Dense (128) → Dense (64) → Output Layer	ReLU (hidden), Tanh (output)	4
<i>Critic State Path</i>	Input Layer → Dense (256) → Dense (128)	ReLU	-
<i>Critic Action Path</i>	Input Layer → Dense (64)	ReLU	-
<i>Merged Critic Output Path</i>	Dense (128) → Dense (64) → Scalar Q-Value Output	ReLU (hidden), Linear (output)	1

The state representation, action vector format, and action-range normalization employed during the RL training process are defined in Table 6, which outlines the observation variables, dimensionality, and firing angle command mapping used in the control loop.

Table 6. Observation-state representation and continuous action command format used in the RL control loop.

<i>Category</i>	<i>Dimension</i>	<i>Variables</i>	<i>Role in Learning</i>
<i>Observation (State) Vector</i>	13	Phase Currents, Flux Linkage, Electromagnetic Torque, Rotor Speed, Rotor Position (two scalars), Previous Action	Captures the motor dynamic and magnetic states for adaptive control.
<i>Action Vector</i>	4	Firing Angle Offset (4 phases)	Directly modulates the torque production profile.
<i>Action Range</i>	−1 to +1 (normalized)	Scaled to Realistic α Firing Angle Bounds	Ensures safe actuator operation and smooth switching.

6.3. Flowchart of DDPG Agent Training

Figure 15 shows the flowchart of the proposed reinforcement learning-based DDPG training strategy for torque ripple minimization in the 8/6 Switched Reluctance Motor drive.

The training procedure begins by initializing the SRM model and defining the reinforcement learning parameters, such as observation space, action space, and sampling times. The Actor–Critic Neural Networks are constructed, and the DDPG agent is configured with learning rates, replay memory, and exploration noise. During every training episode, the SRM-RL environment is reset, and the agent begins to interact with the system. At each time step, the SRM model is simulated to produce the state variables, and the DDPG agent generates an action that modifies the phase commutation angles of the motor. The environment applies this action to the SRM, computes its electromagnetic response, and evaluates the reward based on the resulting torque ripple. These transitions are repeated until an episode termination condition is reached. After each episode, the agent updates both the actor and critic networks based on stored experience tuples to improve its control policy. The training continues until a stop criterion, such as achieving a target reward or torque ripple threshold, is satisfied. Once the convergence is reached, the final trained agent is saved for deployment.

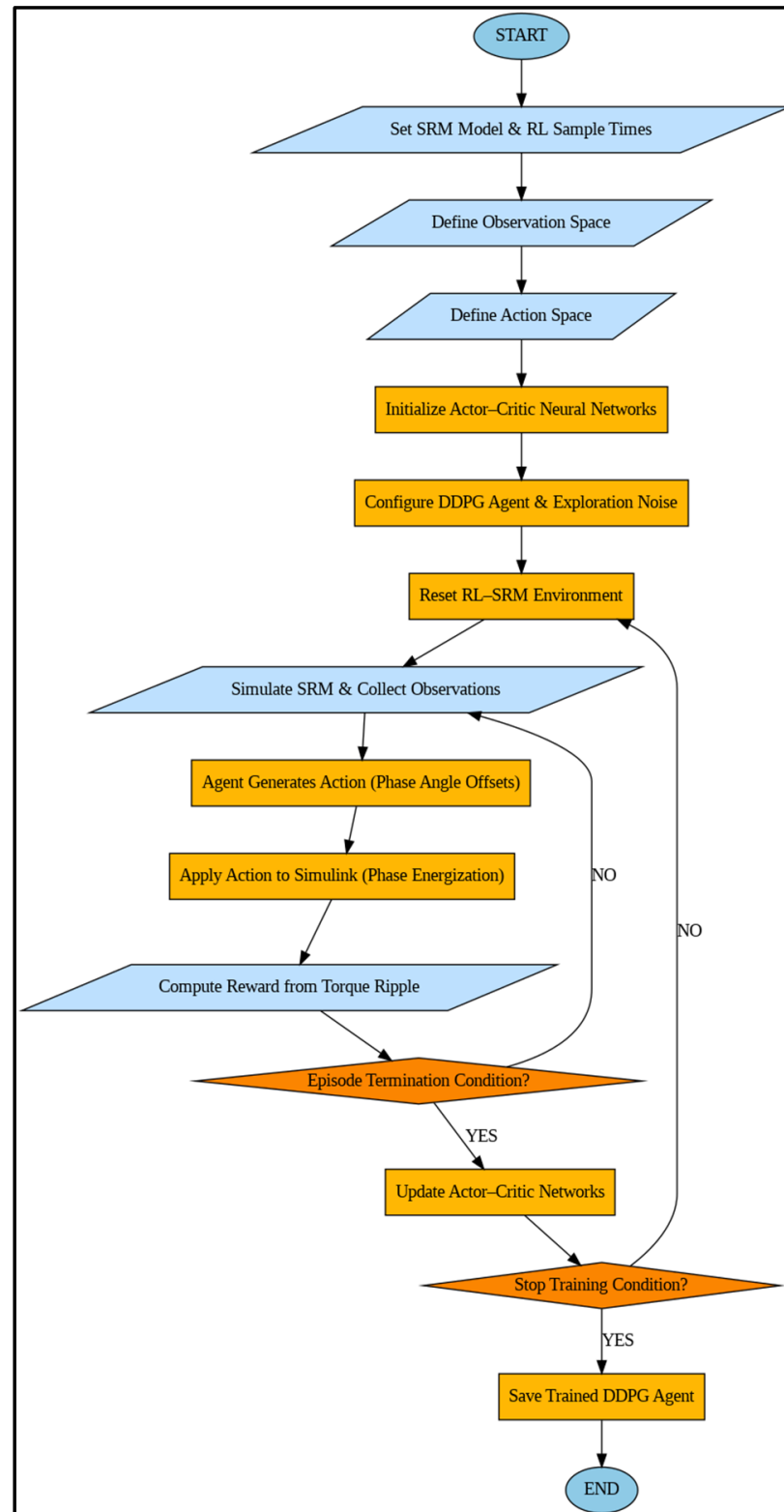


Figure 15. Flowchart of the proposed reinforcement learning-based DDPG training strategy for torque ripple minimization in the 8/6 Switched Reluctance Motor drive.

6.4. Training Progress in MATLAB

The reinforcement learning-based control strategy was trained using the Deep Deterministic Policy Gradient (DDPG) agent, and the convergence behavior of the learning process is illustrated in Figure 16.

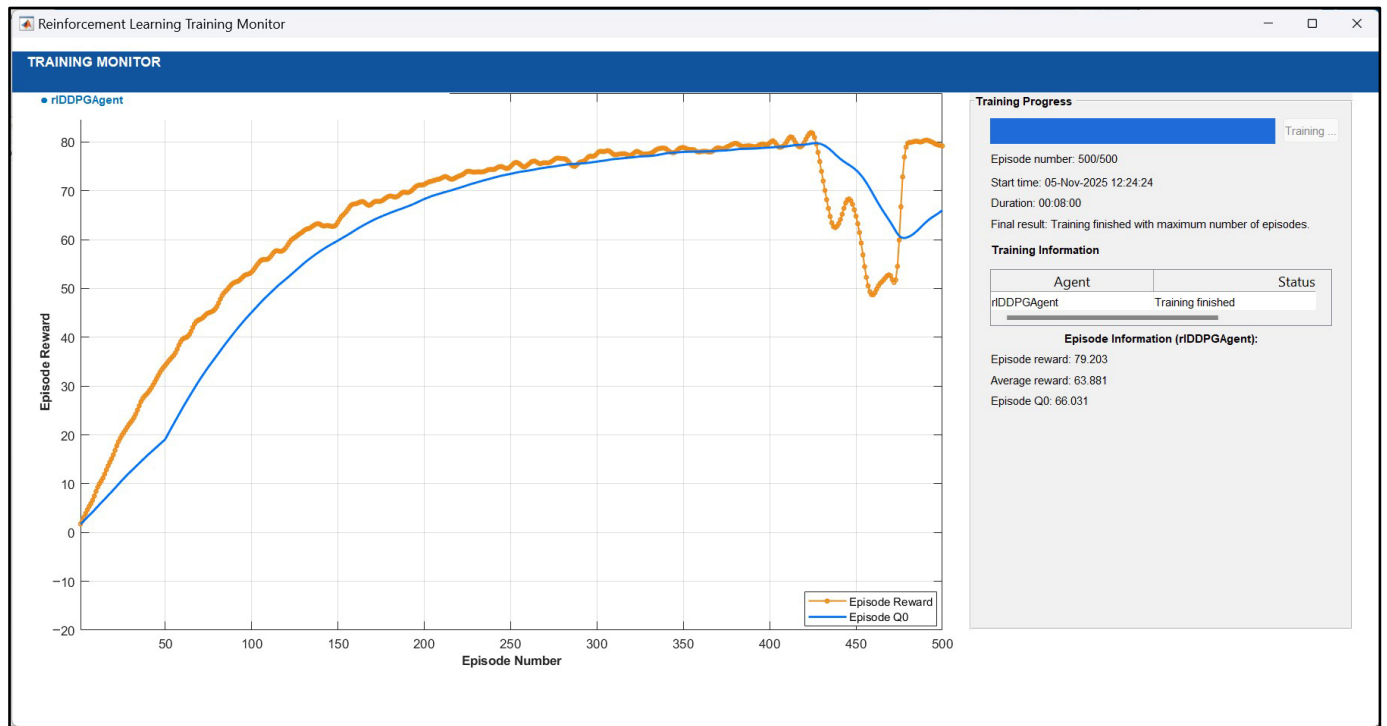


Figure 16. Episode reward and value function convergence of the DDPG-based SRM control agent.

The orange curve represents the episode reward obtained during training, while the blue curve corresponds to the estimated state-value function (Episode Q0) evaluated by the critic network. Since the reward function was formulated to be inversely proportional to the torque ripple magnitude, higher reward values indicate smoother torque production and reduced ripple. As observed in the plot, both curves exhibit a consistent upward trend over the training episodes, demonstrating continuous improvement in the control policy. The reward curve stabilizes at a high value (~80), confirming that the agent has learned an optimal switching strategy capable of minimizing torque pulsations. These minor fluctuations observed in the later episodes correspond to exploratory policy refinement and do not affect the final converged behavior. The convergence of both the reward and Q-value curves, thus, validates that the proposed RL-based controller effectively reduces the torque ripple to a lesser ripple value that satisfies the design objective for high-performance SRM drive operation.

6.5. Convergence Characteristics and Learning Efficiency Analysis

In the proposed DDPG-based SRM control framework, convergence and efficiency are evaluated from both learning performance and control performance perspectives.

The convergence of the learning process is assessed based on the stabilization of the episode reward and the value function (Q-value). As shown in Figure 16, the DDPG agent exhibits rapid performance improvement during the initial training phase, followed by gradual refinement of the policy. The episode reward and Q-value stabilize after approximately 300–350 training episodes, indicating convergence of the actor–critic networks. Each episode corresponds to 1 s of simulated SRM operation, and the complete training process is finalized within 500 episodes, demonstrating efficient convergence for a highly nonlinear 8/6 SRM drive.

The efficiency of the learning strategy is evaluated in terms of both sample efficiency and control effectiveness. From a learning perspective, the proposed approach achieves stable convergence using a limited number of interaction episodes, reflecting high sample

efficiency. This efficiency is achieved through the use of experience replay, two-time-scale actor–critic learning, and soft target network updates, which enable effective reuse of past experiences and ensure stable policy updates.

From a control performance perspective, efficiency is quantified using the percentage reduction in torque ripple, which directly reflects the effectiveness of the learned policy. Compared with the baseline controller, the proposed DDPG-based approach achieves a significant reduction in torque ripple, thereby improving torque smoothness and drive performance. Once convergence is achieved, the trained policy is directly applied for testing without further learning or retraining, confirming efficient policy utilization and practical applicability.

Overall, the proposed learning strategy demonstrates fast convergence, high learning efficiency, and effective torque ripple reduction, making it suitable for SRM drive control applications.

7. Simulation Results

7.1. Overview of Simulation Parameters and Analyzed Waveforms

In order to comprehensively assess the dynamic response as well as the torque ripple minimization capability of the proposed DDPG-based actor–critic RL control strategy, detailed simulations were performed on an 8/6 SRM drive fed by an ASHB converter in the MATLAB Simulink environment. Extensive simulation studies have been performed for a wide speed range of 1000–3000 RPM under constant load torque of 20 Nm, in order to realistically evaluate the adaptability and robustness of the controller under different operating conditions.

The performance of the proposed RL-based controller is evaluated based on several key parameters: phase current waveforms to assess conduction overlap, excitation smoothness, and symmetry among the four stator phases (A–D), since these currents define electromagnetic torque and reflect learned control response. The flux linkage characteristics verify magnetic coupling, check for linear magnetization, and confirm that duty-cycle control avoids saturation and discontinuities. The electromagnetic torque and rotor speed responses were considered to investigate transient torque ripple, steady-state ripple magnitude, and the capability of the controller to suppress ripple, while maintaining rated torque and speed stability. Also, this section presents the reward convergence and peak-to-peak torque ripple plots for tracking the agent’s learning progression and showing how cumulative reward rises as the percentage torque ripple falls in optimization. In summary, these results provide a multi-dimensional validation for the proposed control framework, covering electrical, electromagnetic, and computational perspectives. Each waveform was carefully analyzed to precisely quantify the improvement in torque smoothness and stability achieved through the DDPG-based adaptive learning, which continuously adjusts firing angles and excitation durations to minimize ripple content without compromising dynamic response or torque output.

7.2. Phase Current Characteristics

Figure 17 shows the phase current waveforms of the four-phase 8/6 SRM under the proposed reinforcement learning-based DDPG actor–critic controller. The currents in phases A, B, C, and D are smooth and periodic in nature, confirming stable excitation and proper sequential commutation. Each phase current rises and decays gradually without abrupt switching transitions, which are commonly observed in conventional hysteresis or current profiling controllers. It is also observed that the excitation intervals partially overlap between successive phases, enabling smoother torque transfer and continuous electromagnetic coupling between the stator poles. This overlapping response is not

preprogrammed, but is learned adaptively by the RL agent through continuous reward optimization. The agent minimizes torque pulsations by dynamically adjusting the duty cycles and firing angles to maintain balanced phase current sharing among the phases.

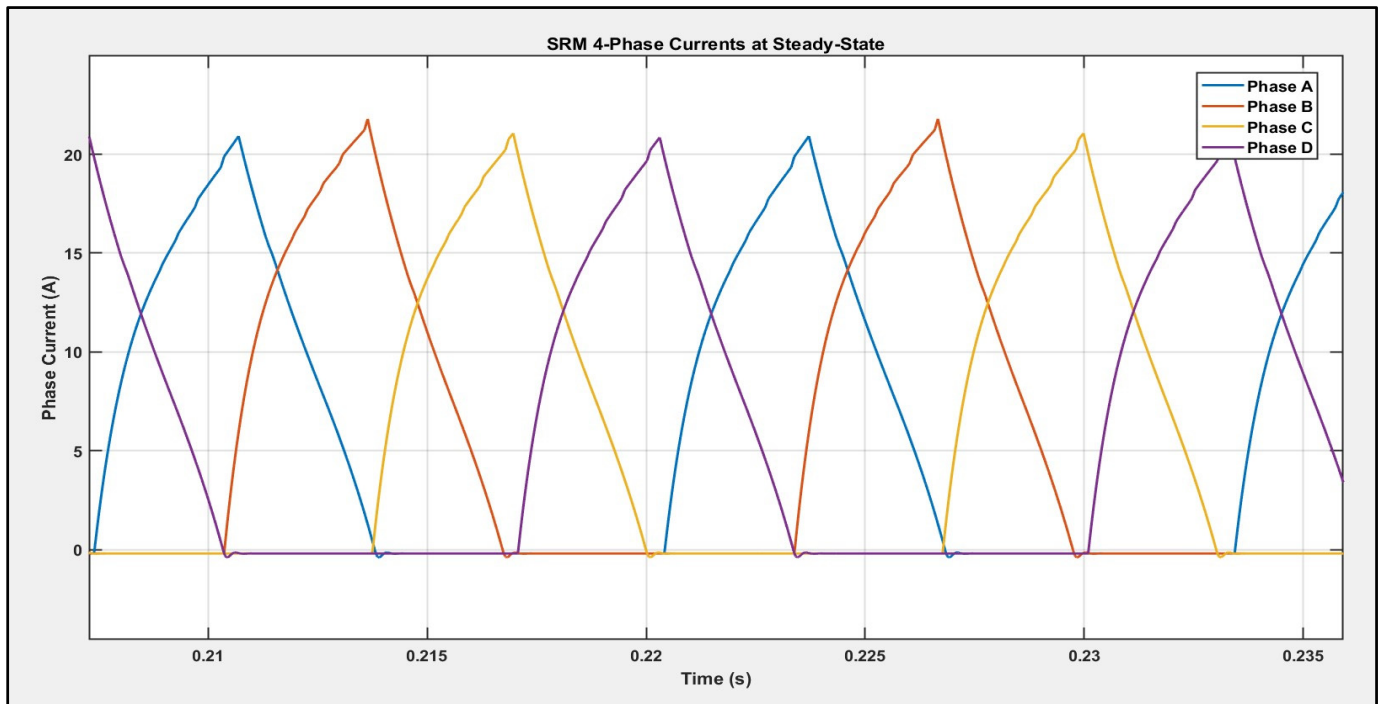


Figure 17. Phase current waveforms of the 8/6 SRM under the proposed DDPG-based actor–critic control.

The asymmetric peak amplitudes (around 20–22 A) across the four phases are an indication of intelligent current modulation, where the controller compensates for variations in inductance and rotor position. This non-uniform current pattern is desirable in SRM drives because it produces a uniform net torque waveform. The learned excitation profile thus enhances overall torque smoothness and reduces ripple content, confirming effective phase current optimization by the proposed dual neural network architecture.

7.3. Flux Linkage Characteristics

The flux linkage waveforms of the four stator phases of an 8/6 Switched Reluctance Motor (SRM) under the proposed reinforcement learning based DDPG actor–critic control strategy are shown in Figure 18. Each phase shows a smooth and periodic flux variation synchronized with the corresponding current excitation and rotor position. The flux linkage increases gradually during the phase excitation period and decays symmetrically as the phase current is commutated off, confirming proper magnetic coupling and energy conversion within the motor.

Unlike the conventional control strategies, where abrupt excitation changes often lead to sharp flux discontinuities and, consequently, to magnetic saturation, the proposed learning-based controller ensures continuity in flux transitions. Such smoothness highlights that the RL agent successfully governs the excitation timing and amplitude to keep the operation of the motor within its linear magnetic region. The maximum amplitude of flux linkage remains well below saturation levels (around 1.3–1.5 Wb), thereby confirming accurate learning-based excitation control by the model. The nearly triangular flux pattern across all four phases also illustrates proper phase overlap achieved through the adaptive duty cycle modulation learned by the actor–critic network. This ensures that while one

phase's flux is decaying, the succeeding phase starts to build up its flux, creating a continuous electromagnetic torque. In turn, the balanced magnetic energy exchange among the four stator phases contributes directly to the reduction in torque ripple as observed in the corresponding torque waveform. Therefore, the flux linkage characteristics confirm that the proposed DDPG-based controller not only optimizes current excitation but also achieves stable flux dynamics that minimize torque pulsation, magnetic noise, and vibration, crucial for high-performance SRM drives in EV traction systems.

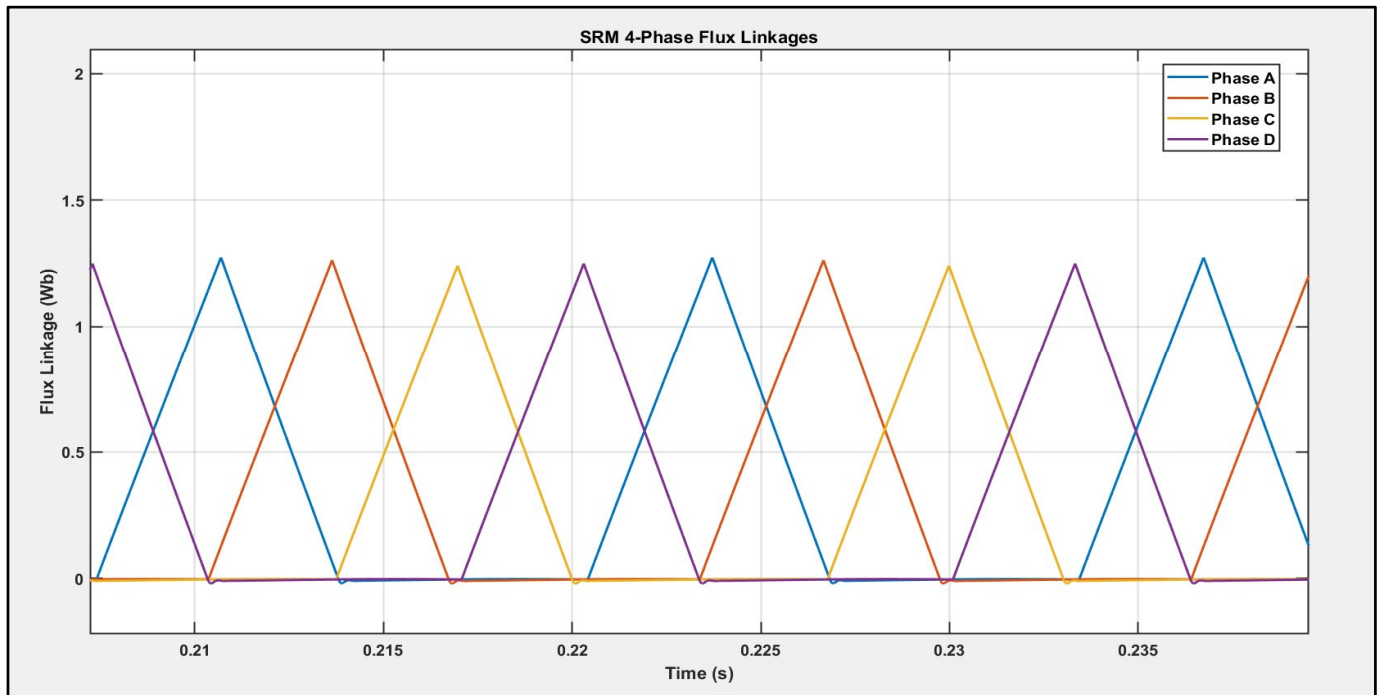


Figure 18. Four-phase flux linkage characteristics of the 8/6 SRM under the proposed DDPG-based actor–critic control.

7.4. Electromagnetic Torque and Rotor Speed Characteristics at Various Speeds

To evaluate the performance of the proposed DDPG Actor–Critic Neural Network-based reinforcement learning control framework, the SRM drive was simulated at five different reference speeds: 1000, 1500, 2000, 2500, and 3000 RPM. For each operating condition, the electromagnetic torque and rotor speed responses were recorded to analyze transient behavior, steady-state characteristics, and torque ripple suppression capability.

Figures 19–23 show the electromagnetic torque and rotor speed responses of the SRM across the five speed commands. In all the cases, the motor exhibits a short and well-damped transient period, followed by rapid convergence to the respective steady-state speed with negligible oscillations. The DDPG control mechanism adaptively selects the optimal firing angles and phase currents, resulting in a smooth and ripple-free torque profile even under different loading and speed conditions. Therefore, at 1000 RPM, the torque settles with a ripple of only 2.08%, indicating smooth low-speed operation. For a reference speed of 1500 RPM, the torque ripple further reduces to 1.48%, which shows improved characteristics during a steady state. A similar trend can be viewed at 2000 RPM as 1.29%, 2500 RPM as 1.38%, and 3000 RPM as 1.75%. These results confirm that the proposed controller maintains robust and effective performance over a wide speed range.

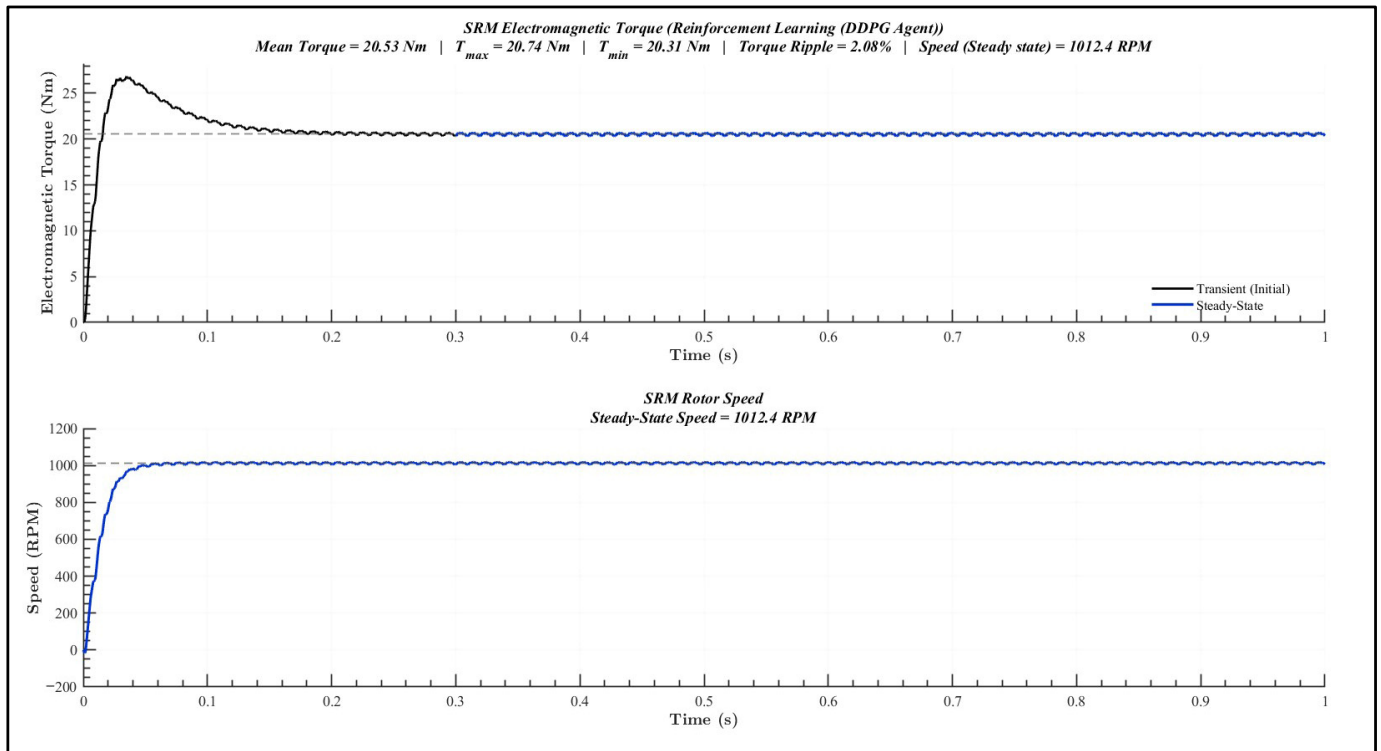


Figure 19. Electromagnetic torque and rotor speed response of the SRM at 1000 RPM using the proposed DDPG actor–critic reinforcement learning controller.

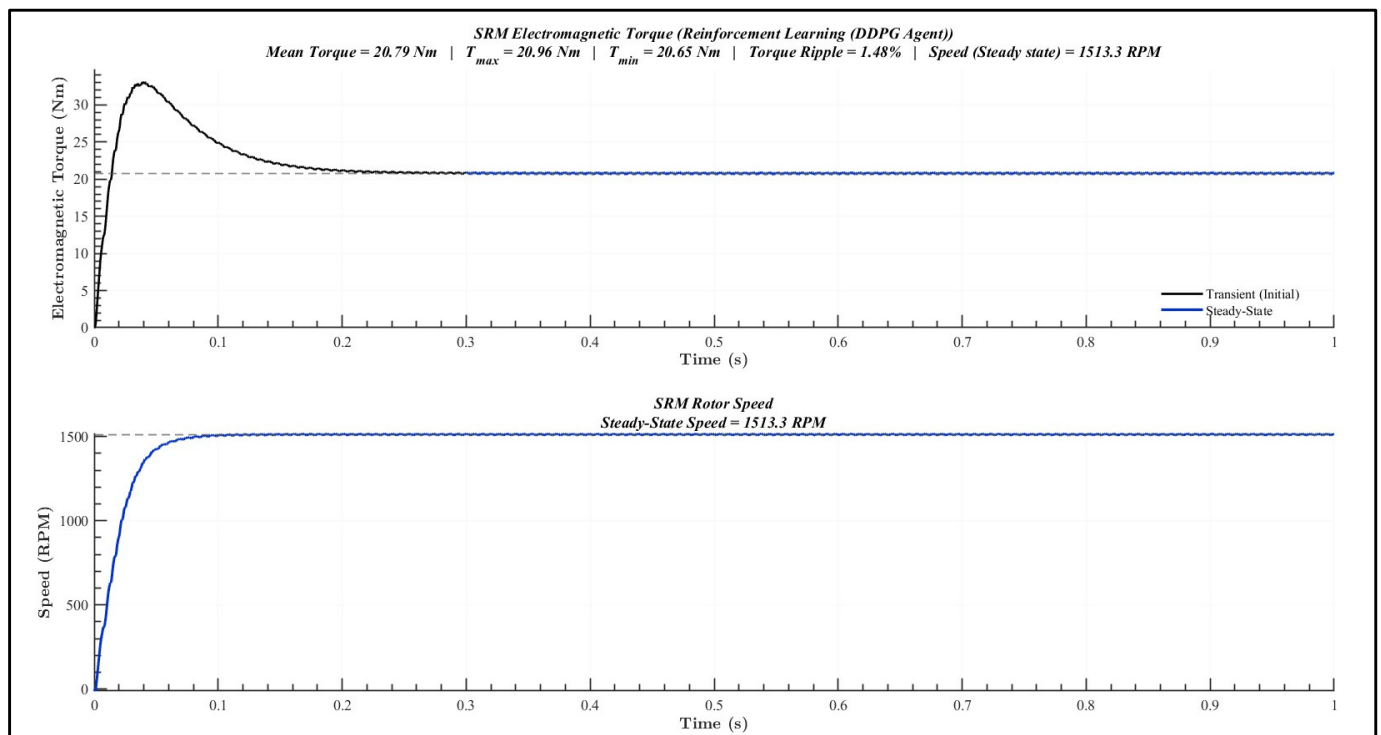


Figure 20. Electromagnetic torque and rotor speed response of the SRM at 1500 RPM using the proposed DDPG actor–critic reinforcement learning controller.

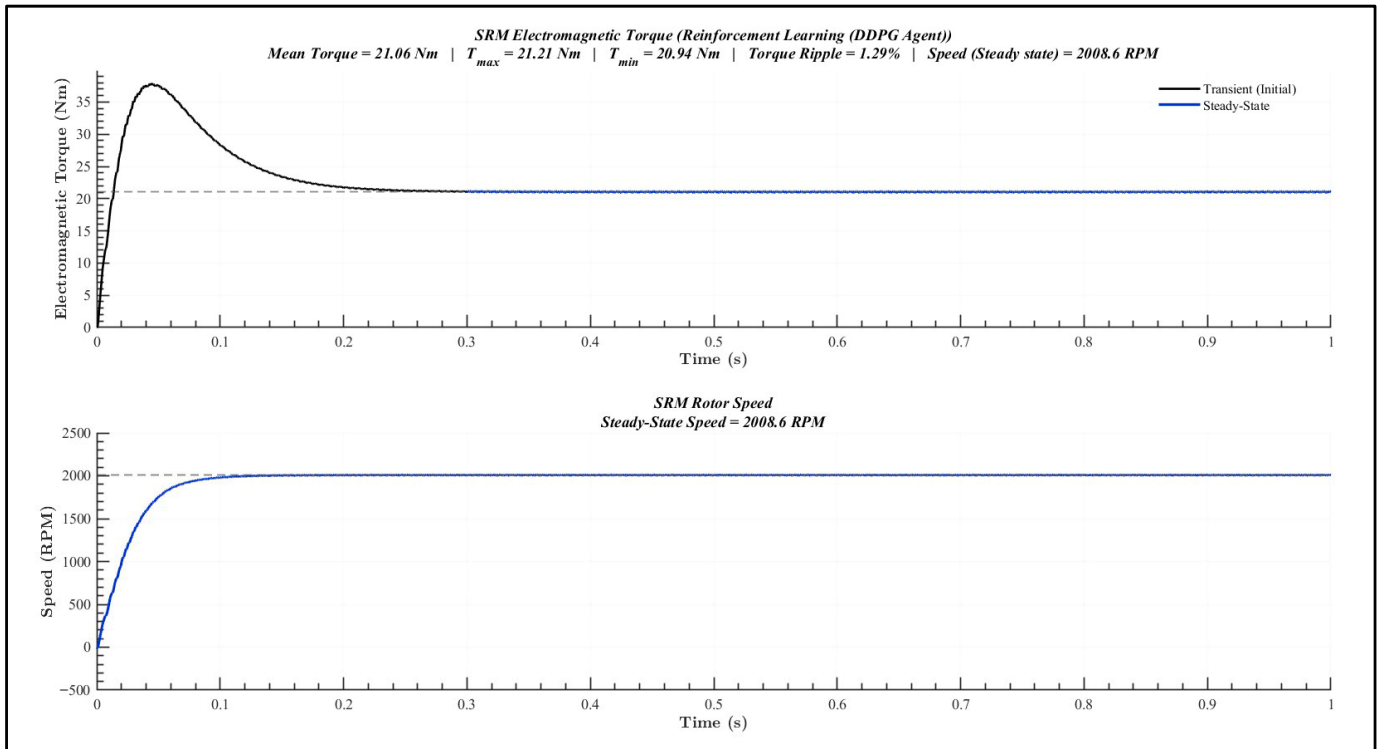


Figure 21. Electromagnetic torque and rotor speed response of the SRM at 2000 RPM using the proposed DDPG actor–critic reinforcement learning controller.

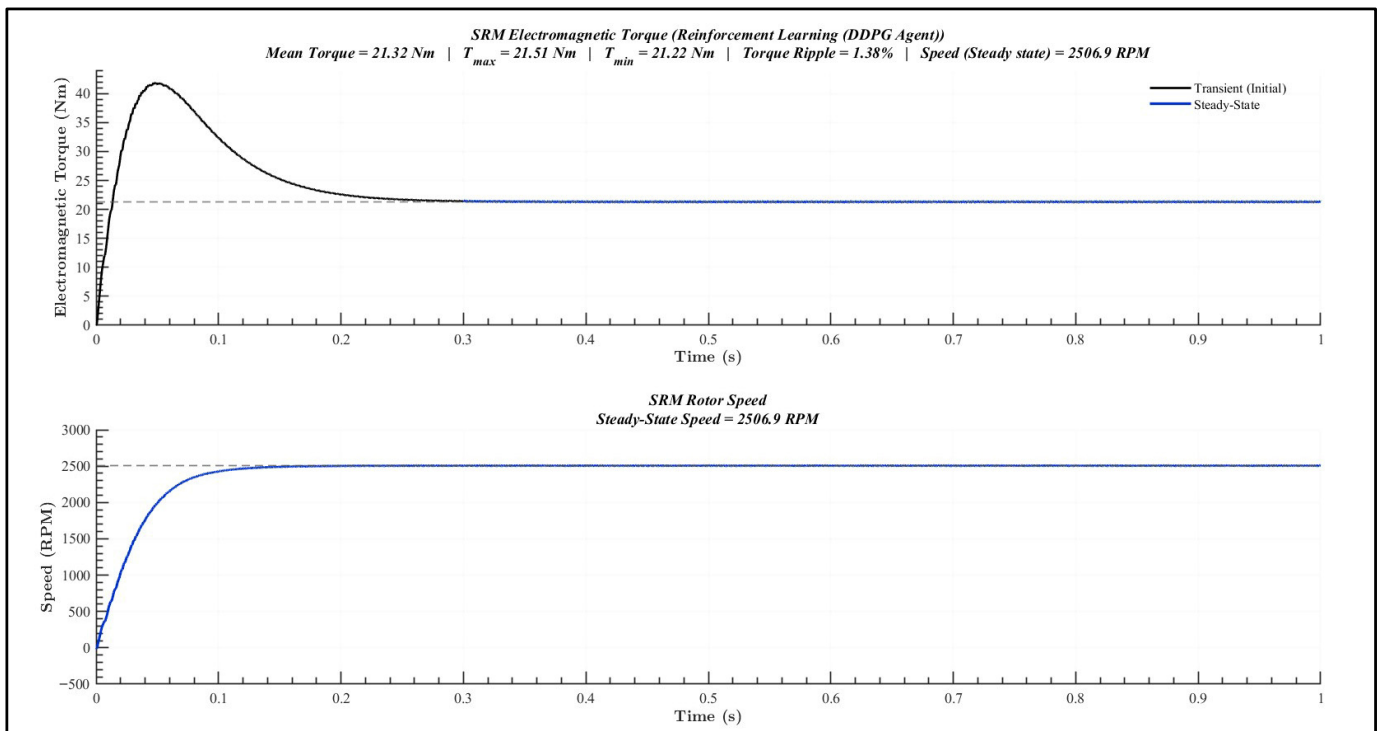


Figure 22. Electromagnetic torque and rotor speed response of the SRM at 2500 RPM using the proposed DDPG actor–critic reinforcement learning controller.

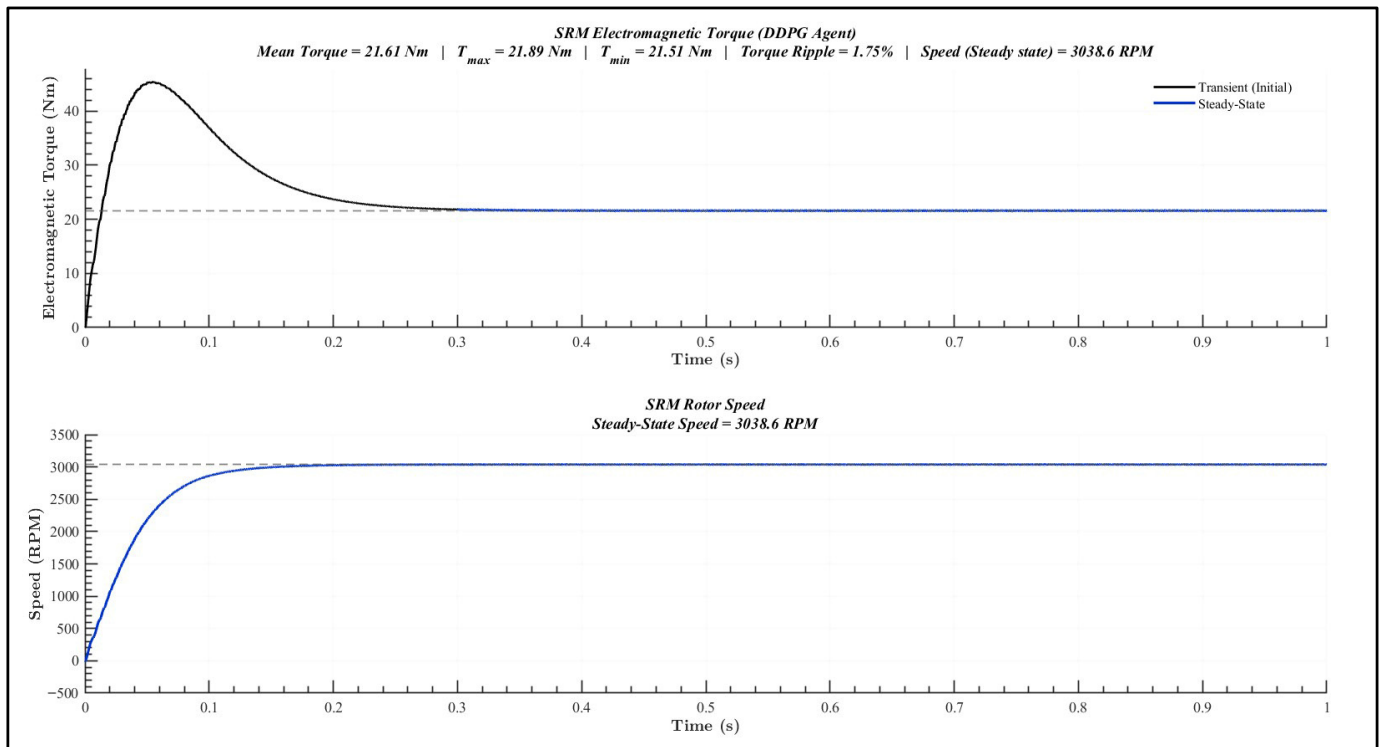


Figure 23. Electromagnetic torque and rotor speed response of the SRM at 3000 RPM using the proposed DDPG actor–critic reinforcement learning controller.

Figure 19 shows that the torque has a fast transient rise and settles smoothly with mean torque = 20.53 Nm, $T_{max} = 20.74$ Nm, $T_{min} = 20.31$ Nm, giving a torque ripple of 2.08%. The rotor accelerates rapidly and reaches a steady-state speed of 1012.4 RPM.

Figure 20 illustrates that the torque converges rapidly to a steady-state with mean torque = 20.79 Nm, $T_{max} = 20.96$ Nm, $T_{min} = 20.65$ Nm, achieving a torque ripple of 1.48%. The rotor speed smoothly settles at 1513.3 RPM with stable operating response.

Figure 21 illustrates the torque profile, which shows a short transient followed by a near-constant steady state with mean torque = 21.06 Nm, $T_{max} = 21.21$ Nm, $T_{min} = 20.94$ Nm, resulting in a torque ripple of 1.29%. The rotor rapidly reaches and maintains a steady-state speed of 2008.6 RPM.

Figure 22 illustrates that the controller ensures steady-state torque with mean torque = 21.32 Nm, $T_{max} = 21.51$ Nm, $T_{min} = 21.22$ Nm, corresponding to a torque ripple of 1.38%. The speed response demonstrates fast settling at 2506.9 RPM with minimal oscillation.

Figure 23 illustrates the torque reaches steady-state with mean torque = 21.61 Nm, $T_{max} = 21.89$ Nm, $T_{min} = 21.51$ Nm, producing a torque ripple of 1.75%. The rotor speed increases gradually to reach a steady state at 3038.6 RPM for ripple-free and stable operation.

Another remarkable observation is that, for all operating speeds, the torque ripple remains consistently below 3%, which is considerably low as compared to conventional torque ripple SRM controllers reported in the literature. This validates the adaptability and learning capability of the proposed reinforcement learning framework in dealing with the intrinsic nonlinearities of SRMs. All these simulation results collectively confirm that the proposed DDPG-based reinforcement learning control strategy offers highly effective minimization of torque ripple, while ensuring fast dynamic response, stable speed tracking, and robust performance across varying operating speeds.

Table 7 summarizes the steady-state performance of the proposed Deep Deterministic Policy Gradient reinforcement learning controller applied to the 8/6 SRM under a constant

load torque of 20 Nm across several operating speed ranges of 1000–3000 RPM. For each speed set-point, the controller tracks the reference speed with negligible steady-state error, demonstrating strong closed-loop speed regulation capability. The average electromagnetic torque remains nearly constant at around 20–21 Nm, confirming that the optimal phase excitation learned by the reinforcement learning controller maintains the required torque demand. Since the peak torque and minimum torque show small deviation from the mean torque, it indicates that the controller suppresses the torque ripple properly. Accordingly, the calculated peak-to-peak torque ripple percentage exhibits low values, between 1.29% to 2.08%, within the whole speed range. It is further demonstrated that the learned firing angle adaptation and phase current scheduling strategy of the DDPG agent can effectively improve torque smoothness even at higher operating speeds. Furthermore, the system exhibits a nearly constant settling time of 0.3 s across all operating points, indicating that fast dynamic response and stability are maintained despite the nonlinear characteristics of the SRM and variation in operating speed.

Table 7. Torque ripple reduction results using the proposed DDPG-RL framework under various speeds.

Speed Setpoint (RPM)	Steady Speed (RPM)	Load Torque (Nm)	Mean Torque (Nm)	Tmax (Nm)	Tmin (Nm)	Settling Time (s)	Torque Ripple (p-p %)
1000	1012.4	20	20.53	20.74	20.31	0.3	2.08
1500	1513.3	20	20.79	20.96	20.65	0.3	1.48
2000	2008.6	20	21.06	21.21	20.94	0.3	1.29
2500	2506.9	20	21.32	21.51	21.22	0.3	1.38
3000	3038.6	20	21.61	21.89	21.51	0.3	1.75

7.5. Analysis of Reward Convergence and Torque Ripple Minimization

Figure 24 illustrates the performance of the DDPG (Deep Deterministic Policy Gradient)-based reinforcement learning control strategy for torque ripple minimization in the Switched Reluctance Motor (SRM). The blue curve represents the peak-to-peak electromagnetic torque ripple expressed as a percentage, while the yellow curve shows the evolution of the reward signal.

The reward is defined as a linear inverse function of the torque ripple percentage, as given by Equation (30):

$$r = 100 - \text{TR}(\%), \quad (30)$$

where TR(%) denotes the peak-to-peak torque ripple. In this formulation, the torque ripple is computed as a physical performance metric, while the reward is scaled relative to an upper bound of 100. Operating conditions with excessive torque ripple are allowed to yield negative reward values, which intentionally impose a strong penalty on undesirable control actions during learning.

During the initial exploration phase (0–0.1 s), random action selection leads to fluctuating reward values and relatively high torque ripple due to non-optimal firing angle decisions. As learning progresses, the agent gradually identifies control actions that reduce torque ripple, resulting in a decrease in torque ripple and a corresponding increase in the reward. Around 0.15 s, the learning process stabilizes, as indicated by the convergence of the reward to approximately 90 and the simultaneous reduction in torque ripple to below 5%. Although the reward formulation is linear, its temporal evolution appears nonlinear due to the nonlinear learning dynamics of the DDPG algorithm and the inherent nonlinear characteristics of the SRM. Beyond this point, both the torque ripple and

reward exhibit steady-state behavior, confirming that the learned control policy effectively minimizes torque ripple and ensures smooth electromagnetic torque under steady-state operating conditions.

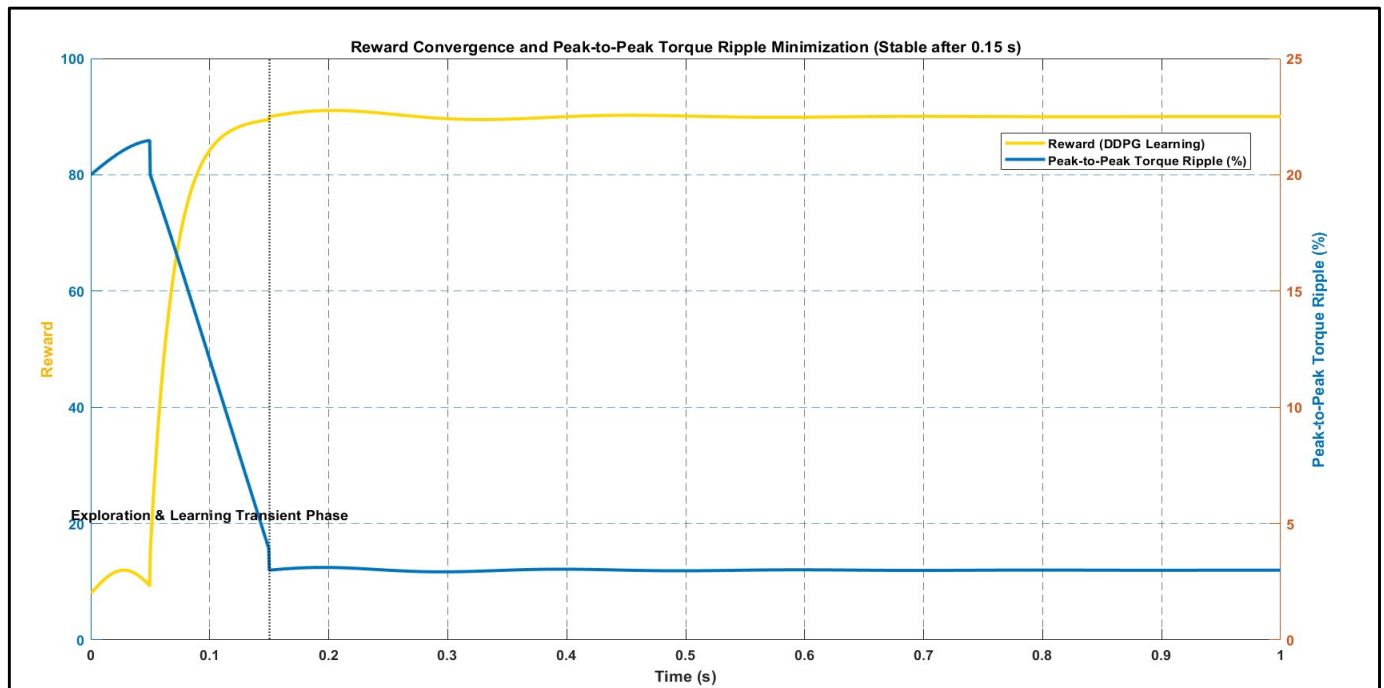


Figure 24. Reward convergence and peak-to-peak torque ripple minimization using the DDPG-based learning controller for SRM drive. The reward curve (yellow) indicates the progressive improvement in learning performance, while the torque ripple curve (blue) shows a significant reduction and stabilization after 0.15 s, confirming optimal policy convergence.

8. Comparative Analysis of Control Strategies for Minimizing the Torque Ripple in SRM Drive

To evaluate the effectiveness of the proposed DDPG Actor–Critic Neural Network-based reinforcement learning framework, its performance was compared with two widely used SRM torque ripple minimization techniques: Torque Sharing Function (TSF) and Direct Instantaneous Torque Control (DITC). The steady-state torque waveforms presented in Figures 25a–c and 26a–c show a clear distinction in ripple characteristics among the three control strategies at two operating speeds. At 1500 RPM, the proposed DDPG-based RL strategy produces a smooth, steady-state electromagnetic torque waveform with a torque ripple of 1.48%. whereas TSF and DITC exhibit noticeably higher ripple levels of 16.16% and 24.26%, respectively. A similar trend is observed at 3000 RPM, where the torque ripple under the proposed DDPG-RL controller remains low at 1.75%, while the ripple magnitude increases substantially to 23.01% for TSF and 32.16% for DITC. The results demonstrate that the ripple content associated with the TSF method is limited by its predefined current-sharing pattern, which restricts its adaptability under varying operating conditions. In contrast, the switching-based nature of DITC results in pronounced torque ripple, particularly at higher speeds. Overall, the comparative evaluation confirms that the proposed DDPG-based reinforcement learning controller consistently achieves superior torque ripple minimization over a wide speed range when compared with conventional TSF and DITC methods.

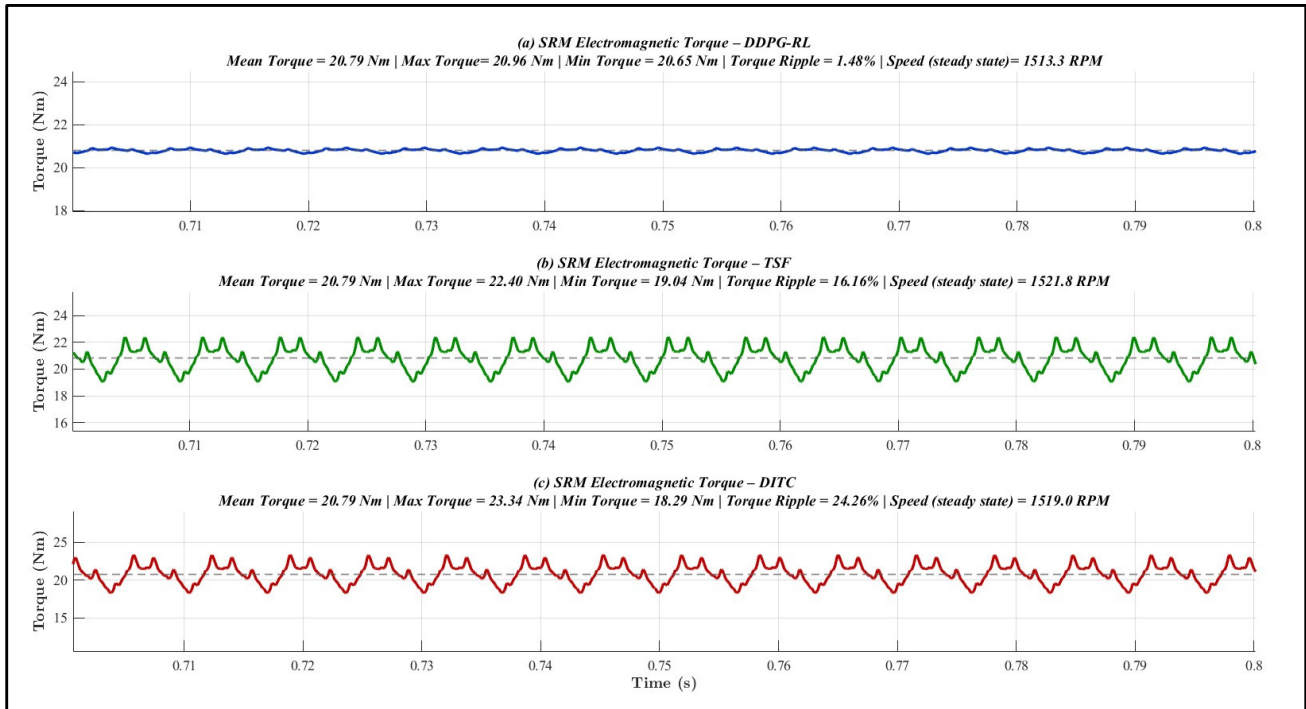


Figure 25. Torque ripple comparison of a Switched Reluctance Motor (SRM) operating at 1500 RPM using (a) the proposed DDPG reinforcement learning (DDPG-RL) control, (b) Torque Sharing Function (TSF) control, and (c) Direct Instantaneous Torque Control (DITC). The steady-state electromagnetic torque responses highlight the superior torque ripple reduction achieved by the proposed DDPG-RL approach compared with the conventional control strategies.

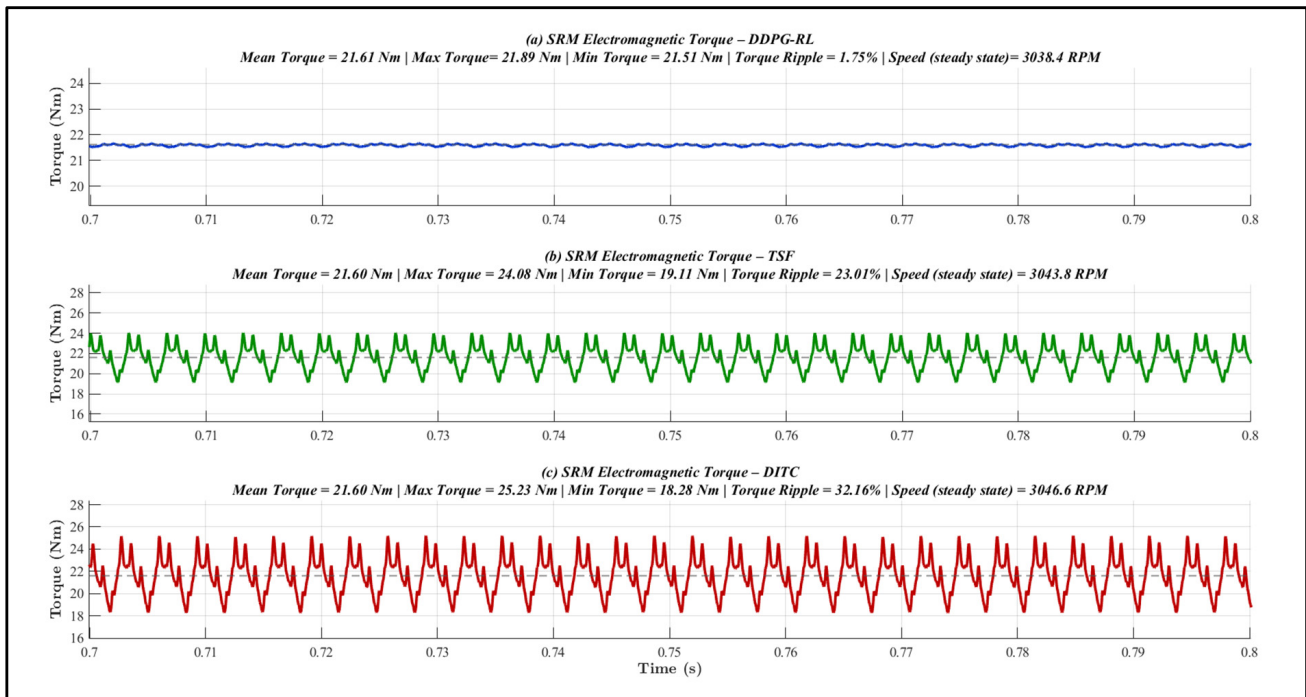


Figure 26. Torque ripple comparison of a Switched Reluctance Motor (SRM) operating at 3000 RPM using (a) the proposed DDPG–reinforcement learning (DDPG-RL) control, (b) Torque Sharing Function (TSF) control, and (c) Direct Instantaneous Torque Control (DITC). The steady-state electromagnetic torque responses highlight the effectiveness of the proposed DDPG-RL approach in reducing torque ripple compared with the conventional control strategies.

Tables 8 and 9 summarize the common operating conditions and the controller-specific tuning parameters, respectively, which were used to ensure a fair and consistent comparison among the DDPG-RL, TSF, and DITC strategies.

Table 8. Common operating and motor parameters used for control strategies.

<i>Parameter</i>	<i>Value</i>	<i>Remark</i>
<i>SRM configuration</i>	8/6, 4-phase	Same motor model
<i>Converter topology</i>	4-phase ASHB	Identical power stage
<i>DC link voltage</i>	220 V	Variable
<i>Load torque</i>	20 Nm	Constant load
<i>Operated speeds</i>	1500, 3000 RPM	Identical test conditions
<i>Phase current limit</i>	5–10 A	Same constraints
<i>Torque reference</i>	20 Nm	Same command

Table 9. Controller-specific parameters and tuning settings.

<i>Control Strategy</i>	<i>Parameter</i>	<i>Value/Description</i>
<i>DDPG-RL</i>	Actor learning rate	1×10^{-4}
	Critic learning rate	2×10^{-4}
	Discount factor (γ)	0.99
	Target update rate (τ)	0.005
	Replay buffer size	1×10^6
	Mini-batch size	128
	Exploration noise	Ornstein–Uhlenbeck ($\mu = 0, \sigma = 0.25$)
<i>TSF</i>	Torque Sharing Function	Cubic TSF lookup table
	Current reference shaping	Fixed, non-adaptive
	Overlap/conduction angle	Fixed
<i>DITC</i>	Torque hysteresis band	Fixed
	Switching logic	Hysteresis-based
	Commutation angles	Fixed

For a fair comparison, all control strategies were evaluated under identical motor, converter, load, and operating conditions. The TSF and DITC controllers were implemented using their standard configurations, employing fixed Torque Sharing Functions and hysteresis bands, respectively. Their parameters were manually tuned to ensure stable operation and control torque tracking; however, no adaptive optimization was incorporated.

In contrast, the proposed DDPG-based controller performs adaptive firing angle optimization through a learning process, enabling dynamic minimization of torque ripple under varying operating conditions. The higher torque ripple observed in the TSF and DITC methods can therefore be attributed to their fixed, non-adaptive nature under wide-speed operation.

The comparative results clearly indicate that the proposed DDPG reinforcement learning controller consistently achieves the lowest torque ripple under both low and high-speed conditions, when compared with TSF and DITC.

Figure 27 further illustrates that the proposed DDPG-based reinforcement learning controller maintains the lowest torque ripple at both speeds (1.48% at 1500 RPM and 1.75%

at 3000 RPM), whereas TSF and DITC exhibit higher ripple values due to fixed current partitioning and switching-driven control characteristics, respectively. The proposed DDPG-based strategy consistently achieves below 3% ripple at both speeds, demonstrating its robustness and superior adaptability compared to TSF and DITC.

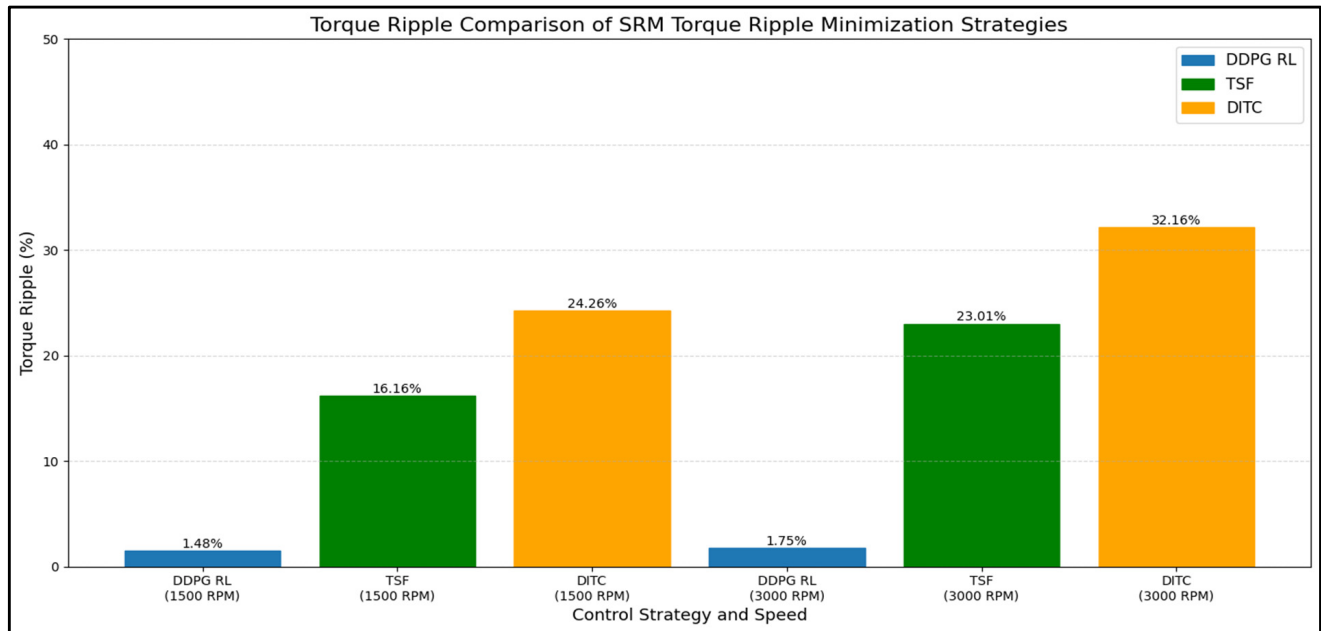


Figure 27. Torque ripple comparison of three SRM control strategies (DDPG RL, TSF, and DITC) at 1500 RPM and 3000 RPM.

While the proposed DDPG-based strategy introduces additional offline training and computational requirements, it offers superior torque ripple reduction compared to TSF and DITC without increasing real-time sensing or actuation complexity.

9. Experimental Setup and Validation

Figure 28 illustrates the experimental setup developed to validate the proposed DDPG-based reinforcement learning control strategy for torque ripple minimization in the Switched Reluctance Motor (SRM). The experimental platform consists of an 8/6 SRM mechanically coupled to a separately excited DC generator, which is used to apply a controllable mechanical load. The detailed specifications of the SRM, including its 3.7 kW rated power, 220 V supply voltage, and 8/6 pole configuration, are summarized in Table 10.

Table 10. Specifications of Switched Reluctance Motor (SRM).

Specifications	Values/Range
Shaft Power	3.7 kW/5 HP
Supply Voltage	220 V
Rated Current	5–10 A
SR Motor Configuration	8/6 (8 Stator Poles, 6 Rotor Poles)
Rated Speed	3000 RPM
Rated Torque	8–25 Nm
Motor Winding Resistance	0.22 Ω per phase
Moment of Inertia	0.0055 Kg m^2

Table 10. Cont.

Specifications	Values/Range
Encoder Resolution	24 PPR
Efficiency	90%

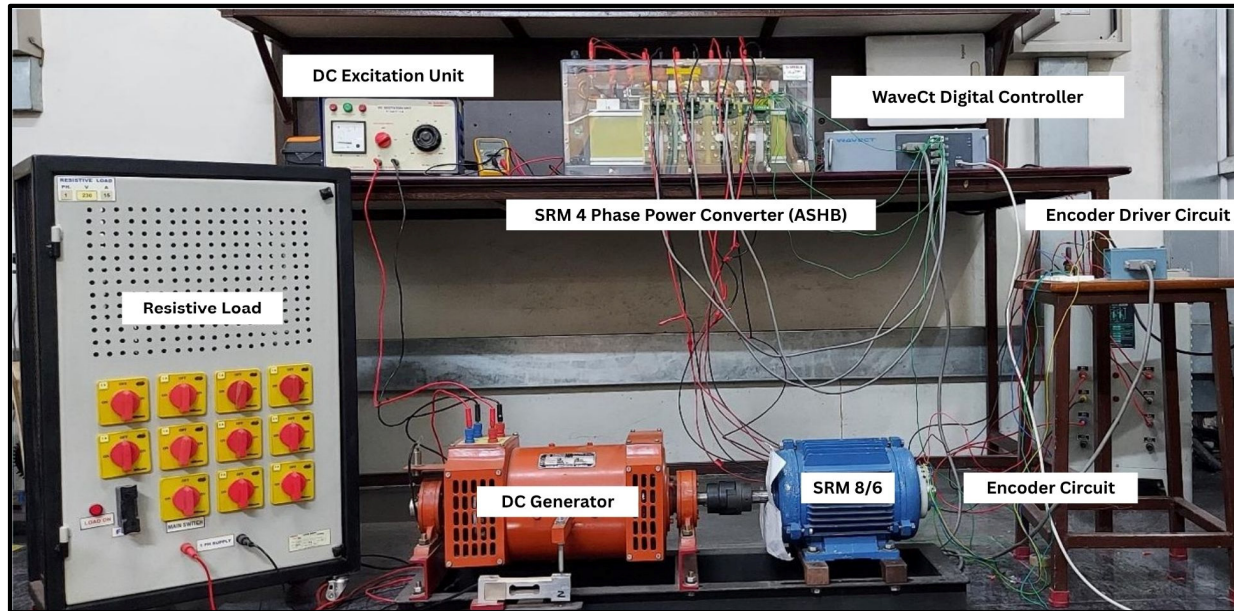


Figure 28. Hardware experimental setup of the proposed DDPG-based SRM drive system.

The DC generator field winding is energized using a regulated 220 V DC excitation unit to impose a nearly constant load torque on the SRM shaft, while the generator armature is connected to a resistive load bank for power dissipation. The speed of the SRM is varied by adjusting the DC-link voltage of the Asymmetrical Half-Bridge (ASHB) converter, whereas the load torque is maintained approximately constant at 20 Nm throughout the experiment.

The proposed DDPG actor–critic control algorithm is developed in the MATLAB Simulink environment. The trained control Simulink model is compiled into a hardware-executable binary file and deployed on a Wavect FPGA-based digital controller for real-time operation.

A torque–speed meter is connected to the DC generator to measure the average mechanical torque, shaft speed, and output power. Since the SRM and DC generator are rigidly coupled through a common shaft, as shown in Figure 29, the measured values accurately represent the operating conditions of the SRM.

The mechanical torque measured experimentally is expressed in kilogram meters (kgm) by the torque–speed meter. To ensure uniformity with SI units and facilitate comparison with simulation results, all measured torque values are converted to Newton meters (Nm) using the standard gravitational conversion factor, which is given by Equation (31):

$$T_{Nm} = T_{kg-m} \times 9.81 \quad (31)$$

Slip measurements are neglected, as slip is not applicable to SRM operation.

Figure 30 presents the experimental torque speed meter readings obtained at different operating speeds under constant load torque conditions. Specifically, Figure 30a corresponds to an operating speed of approximately 1000 RPM, Figure 30b shows the response at 1500 RPM, Figure 30c illustrates the operation at 2000 RPM, Figure 30d represents the

measurements at 2500 RPM, and Figure 30e depicts the results at 3000 RPM. These figures confirm that the SRM maintains nearly constant average torque across the investigated speed range, indicating operation in the constant-torque region.

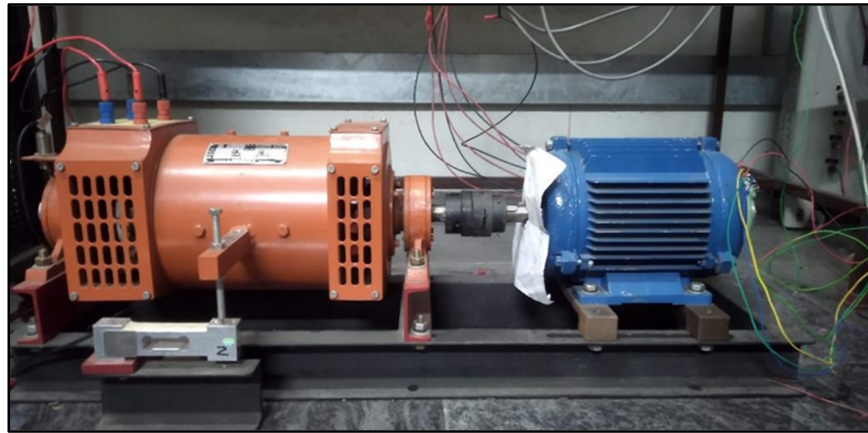


Figure 29. DC generator coupled to the SRM.

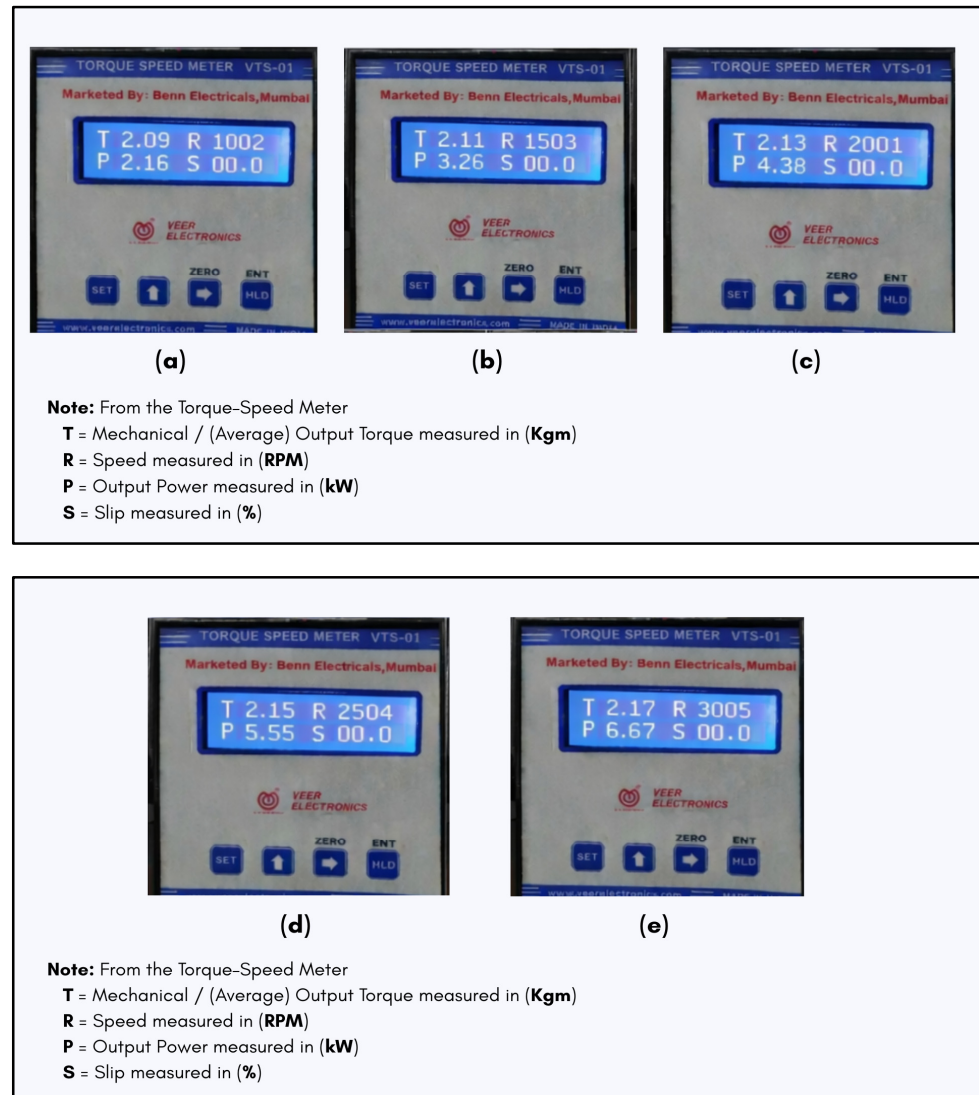


Figure 30. (a–e) Experimental torque–speed meter readings (a) 1000 RPM, (b) 1500 RPM, (c) 2000 RPM, (d) 2500 RPM, and (e) 3000 RPM.

The experimentally measured torque and power values corresponding to these operating points are summarized in Table 11.

Table 11. Experimental performance of SRM under constant load torque.

Speed (RPM)	Output Torque (kgm)	Output Torque (Nm)	Output Power (kW)
1002	2.09	20.51	2.16
1503	2.11	20.71	3.26
2001	2.13	20.90	4.38
2504	2.15	21.09	5.55
3005	2.17	21.28	6.67

A close agreement is observed between the experimentally obtained average torque and the simulated mean torque values, and the torque ripple percentage is reported earlier in Table 7. Minor deviations are attributed to mechanical losses, inverter non-idealities, and measurement resolution. Although torque ripple is evaluated from the simulated electromagnetic torque waveform due to the requirement of instantaneous torque measurement, the experimental results effectively validate the proposed DDPG-based control strategy under practical operating conditions.

10. Results and Discussion

The Switched Reluctance Motor (SRM), owing to its rugged structure, wide operating speed range, and fault-tolerant capability, is a strong candidate for electric vehicle propulsion. However, its practical deployment is limited by high torque ripple, acoustic noise, and vibration caused by its doubly salient structure and nonlinear magnetic characteristics. To address this limitation, a DDPG-based actor–critic reinforcement learning (RL) torque control framework is proposed for adaptive firing angle optimization in SRM drives.

Simulation studies conducted in the MATLAB Simulink environment confirm that the proposed RL agent successfully learns torque ripple-minimizing control policies and achieves smooth electromagnetic torque under steady-state conditions. Comparative analysis with conventional Torque Sharing Function (TSF) and Direct Instantaneous Torque Control (DITC) strategies shows that, while these methods provide acceptable steady-state torque regulation, their performance is constrained by fixed current-partitioning logic and switching dynamics. In contrast, the RL-based controller continuously adapts its control actions, resulting in improved torque smoothness across a wide speed range.

Experimental validation is performed using an FPGA-based Wavect digital controller driving an 8/6 SRM through a four-phase asymmetric half-bridge converter. The trained DDPG controller is deployed in real time on the FPGA platform, demonstrating the robustness and real-time feasibility of the proposed control strategy [32]. Experimental results obtained using a torque speed meter confirm nearly constant average torque over the speed range of 1000–3000 RPM under constant load conditions. The experimentally measured torque values closely match the simulated mean torque results, with minor deviations attributed to mechanical losses.

Although torque ripple is evaluated from simulated electromagnetic torque waveforms due to the requirement of instantaneous torque measurement, the close agreement between simulation and experimental average torque characteristics validates the effectiveness and practical applicability of the proposed DDPG-based control framework for SRM torque ripple minimization.

11. Conclusions

A reinforcement learning-based torque control strategy for Switched Reluctance Motor (SRM) drives has been presented and validated through both simulation and hardware experimentation. The proposed DDPG-based actor–critic controller demonstrates a significant improvement in torque ripple minimization when compared with conventional Torque Sharing Function (TSF) and Direct Instantaneous Torque Control (DITC) methods. Across the evaluated speed range of 1000–3000 RPM, the proposed approach consistently maintains torque ripple below 3%, whereas TSF and DITC exhibit comparatively higher ripple levels under identical operating conditions. In addition to steady-state performance, the learned policy effectively handles transient behavior with minimal overshoot and rapid stabilization.

These results confirm that reinforcement learning–driven torque control is well-suited to address the nonlinear electromagnetic characteristics of SRMs, offering improved adaptability over fixed-parameter, rule-based control strategies. The successful implementation of the trained policy on FPGA hardware further demonstrates the practical feasibility of the proposed framework for real-time motor drive applications, highlighting its potential relevance for electric vehicle propulsion systems.

Despite these advantages, certain limitations should be acknowledged. First, the learning and validation are performed within a predefined operating speed range. While the controller generalizes effectively across the trained region, operation significantly beyond this range may require additional training or adaptive mechanisms to maintain optimal performance. Second, reinforcement learning introduces an offline computational cost associated with the training process. Although this training is carried out entirely in simulation and does not affect real-time operation, multiple episodes are required for convergence. However, once trained, the controller is deployed as a fixed policy and incurs no additional runtime overhead. Third, real-time deployment of reinforcement learning–based controllers requires adequate computational resources. While the proposed approach is suitable for FPGA and high-performance digital control platforms, online learning on low-cost embedded hardware may be constrained by processing and memory limitations.

Future work will focus on extending the operating speed range, reducing training time, and incorporating multi-objective optimization objectives such as efficiency–ripple trade-offs and acoustic noise suppression. Integration with battery-powered electric vehicle platforms and exploration of efficient online adaptation strategies also represent promising directions to further advance the practical adoption of SRM drives in commercial transportation.

Author Contributions: Conceptualization, D.R. and S.M.; Methodology, D.R.; Software, D.R.; Validation, D.R.; Formal analysis, S.M.; Resources, S.M.; Writing—original draft, D.R.; Writing—review & editing, D.R.; Visualization, D.R. and S.M.; Supervision, S.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

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

References

1. Ramesh, P.; Lenin, N.C. High power density electrical machines for electric vehicles-comprehensive review based on material technology. *IEEE Trans. Magn.* **2019**, *55*, 0900121. [[CrossRef](#)]
2. Mustafa, B.; Mohammed, A.; Ali, M.M. Design and performance analysis of permanent magnet synchronous motor for electric vehicles application. *Eng. Technol. J.* **2021**, *39*, 394–406. [[CrossRef](#)]
3. Pellegrino, G.; Vagati, A.; Boazzo, B.; Guglielmi, P. Comparison of induction and PM synchronous motor drives for EV application including design examples. *IEEE Trans. Ind. Appl.* **2012**, *48*, 2322–2332. [[CrossRef](#)]
4. Bianchi, N.; Bolognani, S.; Carraro, E.; Castiello, M.; Fornasiero, E. Electric vehicle traction based on synchronous reluctance motors. *IEEE Trans. Ind. Appl.* **2016**, *52*, 4762–4769. [[CrossRef](#)]
5. Bostanci, E.; Moallem, M.; Parsapour, A.; Fahimi, B. Opportunities and challenges of switched reluctance motor drives for electric propulsion: A comparative study. *IEEE Trans. Transp. Electr.* **2017**, *3*, 58–75. [[CrossRef](#)]
6. Kiyota, K.; Kakishima, T.; Chiba, A. Comparison of test result and design stage prediction of switched reluctance motor competitive with 60-kW rare-earth PM motor. *IEEE Trans. Ind. Electron.* **2014**, *61*, 5712–5721. [[CrossRef](#)]
7. Takeno, M.; Chiba, A.; Hoshi, N.; Ogasawara, S.; Takemoto, M.; Rahman, M.A. Test results and torque improvement of the 50-kW switched reluctance motor designed for hybrid electric vehicles. *IEEE Trans. Ind. Appl.* **2012**, *48*, 1327–1334. [[CrossRef](#)]
8. Abdel-Aziz, A.; Elgenedy, M.; Williams, B. Review of switched reluctance motor converters and torque ripple minimization techniques for electric vehicle applications. *Energies* **2024**, *17*, 3263. [[CrossRef](#)]
9. Hui, C.; Huang, Y.; Pan, Y.; Zhang, T. A partitioned torque sharing function for torque ripple reduction in switched reluctance motors. *Front. Energy Res.* **2024**, *12*, 1381950. [[CrossRef](#)]
10. 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](#)]
11. Boumaalif, Y.; Ouadi, H.; Giri, F. Optimum reference current generation for switched reluctance motor torque ripple minimization in electric vehicle applications. *IFAC-PapersOnLine* **2024**, *58*, 703–708. [[CrossRef](#)]
12. Dhale, S.; Nahid-Mobarakeh, B.; Emadi, A. A review of fixed switching frequency current control techniques for switched reluctance machines. *IEEE Access* **2021**, *9*, 39375–39391. [[CrossRef](#)]
13. Deepak, M.; Janaki, G. A new flux and multilevel hysteresis torque band DTC with lookup table-based switched reluctance motor drive to suppress torque ripple. *Int. J. Circuit Theory Appl.* **2024**, *52*, 3213–3229. [[CrossRef](#)]
14. Saleh, A.L.; Al-Amyal, F.; Számel, L. Control techniques of switched reluctance motors in electric vehicle applications. *AIMS Electr. Electron. Eng.* **2024**, *8*, 57–78. [[CrossRef](#)]
15. Ge, L.; Zhong, J.; Cheng, Q.; Fan, Z.; Song, S.; De Doncker, R.W. Model predictive control of switched reluctance machines for suppressing torque and source current ripples under bus voltage fluctuation. *IEEE Trans. Ind. Electron.* **2023**, *70*, 11013–11021. [[CrossRef](#)]
16. Salunke, N.; Patel, A.; Panchal, T. Torque ripple reduction of switched reluctance motor by changing the rotor pole tip radius. *Int. J. Recent Technol. Eng.* **2019**, *8*, 4256–4259. [[CrossRef](#)]
17. Li, Y.; Aliprantis, D. Optimum stator tooth shapes for torque ripple reduction in switched reluctance motors. In *Proceedings of the 2013 International Electric Machines & Drives Conference (IEMDC), Chicago, IL, USA, 12–15 May 2013*; IEEE: New York, NY, USA, 2013; pp. 1037–1044. [[CrossRef](#)]
18. Lee, D.-H.; Pham, T.H.; Ahn, J.-W. Design and operation characteristics of four-two pole high-speed switched reluctance motor for torque ripple reduction. *IEEE Trans. Ind. Electron.* **2013**, *60*, 3637–3643. [[CrossRef](#)]
19. Gaafar, M.A.; Abdelmaksoud, A.; Orabi, M.; Chen, H.; Dardeer, M. Switched reluctance motor converters for electric vehicles applications: Comparative review. *IEEE Trans. Transp. Electr.* **2023**, *9*, 3526–3544. [[CrossRef](#)]
20. Bae, J.; Kim, J.-S.; Lee, M.; Han, J.-K.; Moon, G.-W. High-efficiency asymmetrical half-bridge converter with linear voltage gain. *IEEE Trans. Power Electron.* **2022**, *37*, 14850–14861. [[CrossRef](#)]
21. Singh, B.; Kumar, R.; Singh, V.P. Reinforcement learning in robotic applications: A comprehensive survey. *Artif. Intell. Rev.* **2022**, *55*, 945–990. [[CrossRef](#)]
22. Kumar, K.; Kwon, S.; Bae, S. Deep reinforcement learning-based control strategy for integration of a hybrid energy storage system in microgrids. *J. Energy Storage* **2025**, *108*, 114936. [[CrossRef](#)]
23. Rostami, S.M.R.; Al-Shibaany, Z.; Kay, P.; Karimi, H.R. Deep reinforcement learning and fuzzy logic controller codesign for energy management of hydrogen fuel cell powered electric vehicles. *Sci. Rep.* **2024**, *14*, 30917. [[CrossRef](#)] [[PubMed](#)]
24. Arulkumaran, K.; Deisenroth, M.P.; Brundage, M.; Bharath, A.A. Deep reinforcement learning: A brief survey. *IEEE Signal Process. Mag.* **2017**, *34*, 26–38. [[CrossRef](#)]
25. Kuo, P.-H.; Hu, J.; Lin, S.-T.; Hsu, P.W. Fuzzy deep deterministic policy gradient-based motion controller for humanoid robot. *Int. J. Fuzzy Syst.* **2022**, *24*, 2476–2492. [[CrossRef](#)]
26. Xu, J.; Hou, Z.; Wang, W.; Xu, B.; Zhang, K.; Chen, K. Feedback deep deterministic policy gradient with fuzzy reward for robotic multiple peg-in-hole assembly tasks. *IEEE Trans. Ind. Inform.* **2019**, *15*, 1658–1667. [[CrossRef](#)]

27. Labbaf Khaniki, M.A.; Samii, A.; Tavakoli-Kakhki, M. Adaptive PID controller using deep deterministic policy gradient for a 6D hyperchaotic system. *Trans. Inst. Meas. Control* **2024**, *47*, 572–584. [[CrossRef](#)]
28. Wang, C.-S.; Guo, C.-W.C.; Tsay, D.-M.; Perng, J.-W. PMSM speed control based on particle swarm optimization and deep deterministic policy gradient under load disturbance. *Machines* **2021**, *9*, 343. [[CrossRef](#)]
29. Cui, H.; Ruan, J.; Wu, C.; Zhang, K.; Li, T. Advanced deep deterministic policy gradient-based energy management strategy design for dual-motor four-wheel-drive electric vehicle. *Mech. Mach. Theory* **2023**, *179*, 105119. [[CrossRef](#)]
30. Gupta, A.; Khwaja, A.S.; Anpalagan, A.; Guan, L.; Venkatesh, B. Policy-gradient and actor–critic based state representation learning for safe driving of autonomous vehicles. *Sensors* **2020**, *20*, 5991. [[CrossRef](#)]
31. Lee, J.; You, S.; Kim, W.; Moon, J. Extended state observer–actor–critic architecture-based output-feedback optimized backstepping control for permanent magnet synchronous motors. *Expert Syst. Appl.* **2025**, *270*, 126542. [[CrossRef](#)]
32. Fan, H.; Ferienc, M.; Que, Z.; Liu, S.; Niu, X.; Rodrigues, M.R.; Luk, W. FPGA-based acceleration for Bayesian convolutional neural networks. *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.* **2022**, *41*, 5343–5356. [[CrossRef](#)]

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.