Reinforcement Learning Control of Six-Phase Permanent Magnet Synchronous Machines

Lara Broghammer[†], Dennis Hufnagel[†], Tobias Schindler[*†], Michael Hoerner[*†],
Petros Karamanakos[‡], Armin Dietz[†], and Ralph Kennel[*]
[*]Chair of Electrical Drive Systems and Power Electronics, Technical University of Munich, Germany
[†]Institute ELSYS, Technische Hochschule Nuremberg, Germany
[‡]Faculty of Information Technology and Communication Sciences, Tampere University, Finland
Email: lara.broghammer@th-nuernberg.de

Abstract—**Control of multi-phase machines is a challenging topic due to the high number of controlled variables. Conventional control methods, such as field-oriented control (FOC), address this issue by introducing more control loops. This, however, increases the controller design complexity, while the tuning process can become cumbersome. To tackle the above, this paper proposes a deep deterministic policy gradient algorithm based controller that fulfills all the control objectives in one computational stage. More specifically, the proposed approach aims to learn a suitable current control policy for six-phase permanent magnet synchronous machines to simplify the commissioning of the drive system. In doing so, physical limitations of the drive system can be accounted for, while the compensation of imbalances between the two three-phase subsystems is rendered possible. After validating the training results in a controller-in-the-loop environment, test bench measurements are provided to demonstrate the effectiveness of the proposed controller. As shown, favorable steady-state and dynamic performance is achieved that is comparable to that of FOC. Therefore, as indicated by the presented results, reinforcement learning-based control approaches for multi-phase machines is a promising research area.**

Index Terms—**Multi-phase machines, current control, permanent magnet synchronous machine (PMSM), power electronics, deep reinforcement learning, deep deterministic policy gradient (DDPG)**

I. INTRODUCTION

Permanent magnet synchronous machines (PMSMs) employing a multi-phase winding configuration have attracted considerable attention due to their fault-tolerant operating capabilities. As a result, they are increasingly used in applications where reliability is crucial, such as autonomous driving or more-electric aircraft.

However, the control of multi-phase PMSMs is considerably more complex than that of conventional three-phase machines. Vector space decomposition (VSD) is a common approach for modeling the machine [1], resulting in three subsystems for a six-phase PMSM. Based on the developed system model, a model-based control, such as field-oriented control (FOC), is a typical current control approach for six-phase PMSMs. Such a control scheme, however, requires additional control loops to account for imbalances or to compensate for the

dominant harmonics [2], [3]. As a result, the controller design complexity increases, and the tuning procedure becomes more cumbersome, thus frequently making the commissioning process challenging.

Model predictive control (MPC) is an alternative to the aforementioned control method for six-phase PMSMs [4], [5]. As MPC can tackle all relevant control objectives in one control loop—by means of an objective function that quantifies the control tasks—it greatly simplifies the controller design procedure. At the same time, the tuning effort is also alleviated. However, despite the promising results, the associated computational complexity of MPC makes its real-time implementation particularly challenging, as high-end control platforms are typically required to meet the increased computational requirements. Moreover, its dependency on the mathematical model of the plant implies that a precise model of the drive system is crucial. Failing to meet this, the system performance can significantly deteriorate.

Reinforcement learning (RL) based algorithms are another emerging method for controlling power electronic systems [6]. Tested applications range from grid-tied inverters [7] to electric machines that are controlled by means of direct torque control [8] or indirect current control [9]. RL offers the opportunity to reduce the complexity of the control structure and streamline the commissioning process of the drive system.

In this paper, the current control concept of [10] based on the deep deterministic policy gradient (DDPG)—introduced in [11] as a continuous control approach—is extended to six-phase machines. More specifically, a DDPG agent independently controls the stator currents of a six-phase PMSM obtained by the VSD transformation. This control approach is referred to as DDPG-CC hereinafter. Moreover, as with FOC, imbalances of the three-phase subsets can be compensated for by controlling the currents in the additional subspaces. The presented results demonstrate the advantages of the proposed control approach. Finally, to the best knowledge of the authors, the presented work is the first to apply DDPG-based control to multi-phase machines.

The structure of this paper is as follows. The succeeding section presents the concept of the proposed DDPG-CC. Following, the training of the DDPG agent, which is conducted exclusively in simulation, is described. Subsequently, details
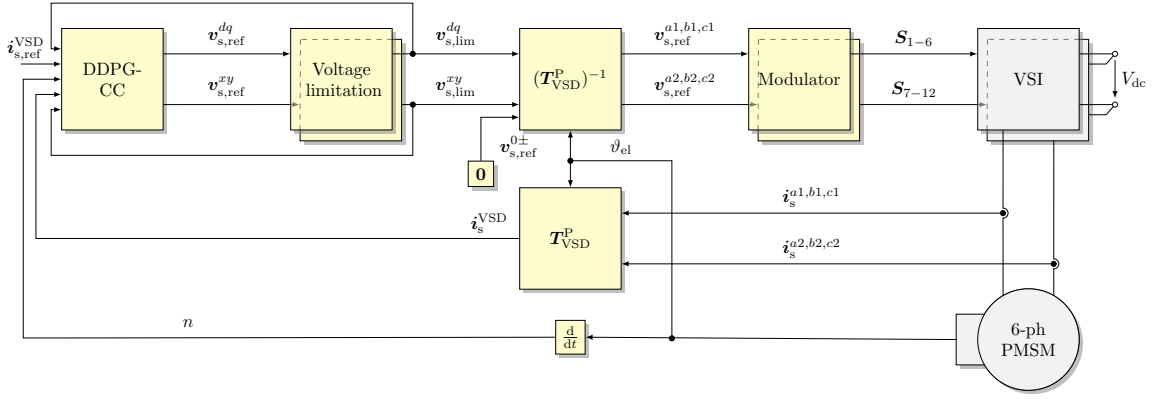
Fig. 1: Block diagram of the proposed DDPG-CC for six-phase PMSM.

regarding the experimental setup for transferring the trained agent from the simulation to the test bench are provided. Finally, the performance of the proposed control approach is assessed. Specifically, measurements on the test bench are discussed, and the control performance is compared with that of FOC.

II. CONTROL CONCEPT

The block diagram of the proposed DDPG-CC concept is illustrated in Fig. 1. The DDPG-CC manipulates the reference voltages. Its setup is described in Section II-B. Additionally, the control loop comprises two separate voltage limitations, modulators and voltage source inverters (VSIs), the VSD and inverse VSD transformations, and a six-phase PMSM. The implementation of the voltage limitation is detailed in Section II-C. Carrier-based pulse width modulation (CB-PWM) is employed for the voltage modulation. The outputs of both modulators, the gate signals S_{1-12}, are fed into two independent three-phase VSIs, which share a common dc-link voltage source. The six-phase PMSM model, the corresponding VSD, and inverse VSD transformation are defined in Section II-A.

A. Model of the Six-Phase PMSM

For the considered non-fault-tolerant operation, VSD is a suitable modeling approach for multi-phase machines, which can also be derived from the theory of symmetrical components [12]. It represents a generalized form of the Clarke transformation for multi-phase machines. The VSD transformation matrix $\boldsymbol{T}_{\text{VSD}}$ is based on the number of phases of the machine and the geometric distribution of the windings. The stator currents are decomposed into several orthogonal subspaces that allow independent control of each component. For the six-phase machine under consideration (Fig. 2), the VSD matrix is given as

$$\boldsymbol{T}_{\text{VSD}} = \frac{1}{3} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} & 0 \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} & \frac{1}{2} & \frac{1}{2} & -1 \\ 1 & -\frac{1}{2} & -\frac{1}{2} & -\frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & -\frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} & \frac{1}{2} & \frac{1}{2} & -1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}. \quad (1)$$
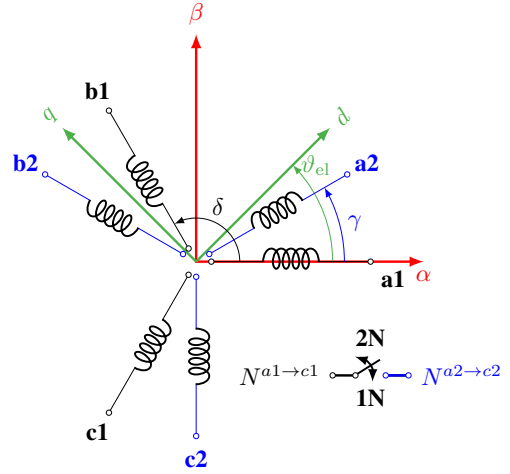

Fig. 2: Winding configuration in the stator- and rotor-fixed coordinate systems with $\delta = \frac{2}{3}\pi$ and $\gamma = \frac{\pi}{6}$. 1N and 2N refer to the two possible neutral point configurations, namely 1N stands for the case where the neutral points of both three-phase winding subsets $N^{a1 \rightarrow c1}$ and $N^{a2 \rightarrow c2}$ are connected, while 2N for the case where the neutral points are not connected.

The natural phase quantities of the stator currents are defined as

$$\boldsymbol{i}_{\text{s}}^{a1 \rightarrow c2} = [i_{\text{s}}^{a1} \ i_{\text{s}}^{b1} \ i_{\text{s}}^{c1} \ i_{\text{s}}^{a2} \ i_{\text{s}}^{b2} \ i_{\text{s}}^{c2}]^T. \quad (2)$$

From (1) and (2), the decomposed stator-fixed VSD current components are calculated by

$$\boldsymbol{i}_{\text{s}}^{\text{VSD}'} = \boldsymbol{T}_{\text{VSD}} \boldsymbol{i}_{\text{s}}^{a1 \rightarrow c2}, \quad (3)$$

where $\boldsymbol{i}_{\text{s}}^{\text{VSD}'} = [i_{\text{s}}^{\alpha} \ i_{\text{s}}^{\beta} \ i_{\text{s}}^{X} \ i_{\text{s}}^{Y} \ i_{\text{s}}^{0^+} \ i_{\text{s}}^{0^-}]^T$. Since the machine model shall be defined in rotating reference frames, the Park transformation, given by

$$\boldsymbol{T}_{\text{P}}(\vartheta_{\text{el}}) = \begin{bmatrix} \cos(\vartheta_{\text{el}}) & \sin(\vartheta_{\text{el}}) \\ -\sin(\vartheta_{\text{el}}) & \cos(\vartheta_{\text{el}}) \end{bmatrix}, \quad (4)$$

is employed. To this aim, with the electrical rotor position ϑ_{el} for the $\alpha\beta$-components and $-\vartheta_{\text{el}}$ for the XY-subspace, as

in [3], a transformation matrix is defined as

$$\boldsymbol{T}_{\mathrm{P}}^{\alpha\beta XY} = \begin{bmatrix} \boldsymbol{T}_{\mathrm{P}}(\vartheta_{\mathrm{el}}) & \boldsymbol{0}_2 & \boldsymbol{0}_2 \\ \boldsymbol{0}_2 & \boldsymbol{T}_{\mathrm{P}}(-\vartheta_{\mathrm{el}}) & \boldsymbol{0}_2 \\ \boldsymbol{0}_2 & \boldsymbol{0}_2 & \boldsymbol{I}_2 \end{bmatrix} . \tag{5}$$

Combining (1) and (5) yields the transformation matrix

$$\boldsymbol{T}_{\mathrm{VSD}}^{\mathrm{P}} = \boldsymbol{T}_{\mathrm{P}}^{\alpha\beta XY} \boldsymbol{T}_{\mathrm{VSD}} . \tag{6}$$

With (6), the stator current vector in rotating reference frames

$$\boldsymbol{i}_{\mathrm{s}}^{\mathrm{VSD}} = [i_{\mathrm{s}}^d \ i_{\mathrm{s}}^q \ i_{\mathrm{s}}^x \ i_{\mathrm{s}}^y \ i_{\mathrm{s}}^{0^+} \ i_{\mathrm{s}}^{0^-}]^T \tag{7}$$

can be calculated by

$$\boldsymbol{i}_{\mathrm{s}}^{\mathrm{VSD}} = \boldsymbol{T}_{\mathrm{VSD}}^{\mathrm{P}} \boldsymbol{i}_{\mathrm{s}}^{a1\to c2} . \tag{8}$$

Following this, the machine model is described by

$$\boldsymbol{v}_{\mathrm{s}}^{dq} = R_{\mathrm{s}} \boldsymbol{i}_{\mathrm{s}}^{dq} + \omega_{\mathrm{el}} \boldsymbol{J} \boldsymbol{\psi}_{\mathrm{s}}^{dq} + \frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{\psi}_{\mathrm{s}}^{dq} , \tag{9a}$$

$$\boldsymbol{v}_{\mathrm{s}}^{xy} = R_{\mathrm{s}} \boldsymbol{i}_{\mathrm{s}}^{xy} - \omega_{\mathrm{el}} \boldsymbol{J} \boldsymbol{\psi}_{\mathrm{s}}^{xy} + \frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{\psi}_{\mathrm{s}}^{xy} , \tag{9b}$$

$$\boldsymbol{v}_{\mathrm{s}}^{0\pm} = R_{\mathrm{s}} \boldsymbol{i}_{\mathrm{s}}^{0\pm} + \frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{\psi}_{\mathrm{s}}^{0\pm} , \tag{9c}$$

where $\boldsymbol{v}_{\mathrm{s}}^m$ are the decomposed stator voltages, R_{s} is the per-phase stator resistance, $\boldsymbol{J} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, ω_{el} is the electrical rotational speed, $\boldsymbol{\psi}_{\mathrm{s}}^m$ are the decomposed flux-linkages, and $m \in \{dq, xy, 0\pm\}$. Since only the 2N topology is considered for this work, (9c) is neglected. Finally, assuming $\boldsymbol{\psi}_{\mathrm{s}}^m$ to be linear, the flux-linkage is modeled using absolute inductances, i.e.,

$$\psi_{\mathrm{s}}^d = \psi_{\mathrm{PM}} + L_{\mathrm{s}}^d i_{\mathrm{s}}^d , \tag{10a}$$

$$\psi_{\mathrm{s}}^q = L_{\mathrm{s}}^q i_{\mathrm{s}}^q , \tag{10b}$$

$$\psi_{\mathrm{s}}^x = L_{\mathrm{s}}^x i_{\mathrm{s}}^x , \tag{10c}$$

$$\psi_{\mathrm{s}}^y = L_{\mathrm{s}}^y i_{\mathrm{s}}^y , \tag{10d}$$

where ψ_{PM} is the flux-linkage of the permanent magnets and L_{s}^m is the stator inductance on the m coordinate.

B. DDPG-Based Current Controller

In the control setup shown in Fig. 1, the goal of the DDPG-CC is to minimize the difference

$$\boldsymbol{e}_{\mathrm{s}}^{\mathrm{VSD}} = \boldsymbol{i}_{\mathrm{s,ref}}^{\mathrm{VSD}} - \boldsymbol{i}_{\mathrm{s}}^{\mathrm{VSD}} \tag{11}$$

between the reference currents

$$\boldsymbol{i}_{\mathrm{s,ref}}^{\mathrm{VSD}} = [i_{\mathrm{s,ref}}^d \ i_{\mathrm{s,ref}}^q \ i_{\mathrm{s,ref}}^x \ i_{\mathrm{s,ref}}^y]^T \tag{12}$$

and the actual currents

$$\boldsymbol{i}_{\mathrm{s}}^{\mathrm{VSD}} = [i_{\mathrm{s}}^d \ i_{\mathrm{s}}^q \ i_{\mathrm{s}}^x \ i_{\mathrm{s}}^y]^T . \tag{13}$$

In RL, this is represented by the reward function $r(k)$, which the agent maximizes in the training. In this work, the reward function

$$r(k) = \begin{cases} -\|\boldsymbol{e}_{\mathrm{s}}^{\mathrm{VSD}}(k)\|_1, & \text{for} \quad i_1(k) \leq I_{\max} \\ -\|\boldsymbol{e}_{\mathrm{s}}^{\mathrm{VSD}}(k)\|_1 - i_1(k), & \text{for} \quad i_1(k) > I_{\max} \end{cases} \tag{14}$$

is defined based on the ℓ_1-norm of (11). Additionally, to stay within the operating limits of the system, the reward function introduces a penalty when the amplitude of the phase current $i_1(k) = \|\boldsymbol{i}_{\mathrm{s}}^{\mathrm{VSD}}\|_2$ exceeds its maximum allowed value $I_{\max}$.

Regarding the observation vector, as per [10], it is defined as

$$\boldsymbol{o}(k) = \begin{bmatrix} \boldsymbol{e}_{\mathrm{s}}^{\mathrm{VSD}}(k) \\ \int \boldsymbol{e}_{\mathrm{s}}^{\mathrm{VSD}}(k) \cdot f_{\mathrm{c}} \\ \boldsymbol{i}_{\mathrm{s}}^{\mathrm{VSD}}(k) \\ \boldsymbol{v}_{\mathrm{s,lim}}^{\mathrm{VSD}}(k-1) \\ n(k) \end{bmatrix} \tag{15}$$

to account for a six-phase PMSM. As can be seen in (15), the proposed observation vector contains the reference current tracking error $\boldsymbol{e}_{\mathrm{s}}^{\mathrm{VSD}}$, the integrated reference current tracking error scaled by the control frequency f_{c}, the actual currents $\boldsymbol{i}_{\mathrm{s}}^{\mathrm{VSD}}$, the limited reference voltages $\boldsymbol{v}_{\mathrm{s,lim}}^{\mathrm{VSD}}$ applied in the previous time step, and the mechanical speed of the PMSM n. To ensure that all observations are within the same order of magnitude, all inputs of the observation vector are normalized by their respective rated or maximum values, i.e., rated current I_{r}, rated speed n_{r}, and maximum stator voltage $V_{\max}$.

The DDPG-CC outputs the action vector

$$\boldsymbol{a}(k) = \begin{bmatrix} v_{\mathrm{s,ref}}^d(k) & v_{\mathrm{s,ref}}^q(k) & v_{\mathrm{s,ref}}^x(k) & v_{\mathrm{s,ref}}^y(k) \end{bmatrix}^T \tag{16}$$

that contains the reference voltages as the manipulated variables of the system. As stated in Section II-A, the zero sequence subspace is omitted due to the usage of the 2N topology. Therefore, the respective voltage reference is set to zero $\boldsymbol{v}_{\mathrm{s,ref}}^{0\pm} = \boldsymbol{0}$.

C. Voltage Limitation

The control action of DDPG-CC is bounded to ensure that the physical limitation of the stator voltage is taken into account in the training process and the subsequent controller behavior. The maximum available voltage depends on the modulation scheme and the dc-link voltage V_{dc}. As CB-PWM is used for each three-phase subsystem, a maximum stator voltage of

$$V_{\max} = \frac{V_{\mathrm{dc}}}{2} \tag{17}$$

can be realized. As detailed in [1], the three subspaces of the used VSD are orthogonal to each other. Therefore, the amplitude of the reference voltage that accounts for all subspaces can be calculated with the ℓ_2-norm of the reference voltages in each axis. Considering (17), this amplitude $\|\boldsymbol{a}(k)\|_2$ is bound by

$$\|\boldsymbol{a}(k)\|_2 \overset{!}{\leq} V_{\max} . \tag{18}$$

In this work, the reference voltages in the dq-subspace $\boldsymbol{v}_{\mathrm{s,ref}}^{dq}(k) = [v_{\mathrm{s,ref}}^d(k) \ v_{\mathrm{s,ref}}^q(k)]^T$ and xy-subspace $\boldsymbol{v}_{\mathrm{s,ref}}^{xy}(k) = [v_{\mathrm{s,ref}}^x(k) \ v_{\mathrm{s,ref}}^y(k)]^T$ are separately bounded, albeit by considering the same maximum voltage V_{lim}. Consequently, the voltage constraints

$$\|\boldsymbol{v}_{\mathrm{s,ref}}^{dq}(k)\|_2 \leq V_{\mathrm{lim}} \tag{19a}$$

$$\|\boldsymbol{v}_{\mathrm{s,ref}}^{xy}(k)\|_2 \leq V_{\mathrm{lim}} \tag{19b}$$

have to be met at each discrete time instance. With (18) and (19), the maximum stator voltage is defined by

$$V_{\max} = \sqrt{2(V_{\lim})^2} \,. \tag{20}$$

Hence, the maximum voltage of the dq-and xy-subspaces $V_{\lim}$ is calculated by

$$V_{\lim} = \frac{V_{\max}}{\sqrt{2}} = \frac{V_{\mathrm{dc}}}{2\sqrt{2}} \,. \tag{21}$$

The voltage limitation is implemented within the subspaces based on [13]. The concept described in [13] for the limitation of the voltages in the dq-system of a three-phase PMSM is adopted for the dq-subspace of the six-phase PMSM and transferred to the xy-subspace in the same way. Specifically, when either of the limitations in (19) is violated, the voltages are limited to (21) while prioritizing the d- and the x-direction respectively and maintaining a safety margin of 5%, leading to the limited reference voltage vector

$$\boldsymbol{v}_{\mathrm{s,lim}}^{\mathrm{VSD}}(k) = \left[v_{\mathrm{s,lim}}^d(k) \; v_{\mathrm{s,lim}}^q(k) \; v_{\mathrm{s,lim}}^x(k) \; v_{\mathrm{s,lim}}^y(k) \right]^T \,. \tag{22}$$

If the voltage limitation is active in either the dq- or the xy-subspace, the integration of the reference current tracking error in the observation vector $\boldsymbol{o}(k)$ is stopped to prevent integrator windup (clamping).

One disadvantage of the implemented approach for the voltage limitation is that the dc-link voltage cannot be utilized to its full extent. Furthermore, a preference of either the x- or the y-axis is inconclusive compared to the dq-subspace. Different approaches to handling the voltage limitation for multi-phase systems have been presented in the literature, e.g., [14], [15]. Since this paper focuses on extending the DDPG-CC to six-phase machines, these approaches are omitted to simplify the control concept. Nevertheless, a voltage limitation concept with a higher utilization for the presented DDPG-CC will be investigated in future work.

III. TRAINING OF THE DDPG AGENT

The training of the DDPG-CC is conducted exclusively in a simulation environment in this work. For this purpose, the Matlab 2023a Reinforcement Learning Toolbox [16] is used. Compared to the real-world control setup shown in Fig. 1, the simulation model omits the VSI, the modulation, and the VSD transformation. Therefore, a sinusoidal voltage is assumed to be applied to the PMSM, which is modeled according to (9) and (10). The voltage limitation is implemented as defined in Section II-C.

In each training episode, uniformly distributed random values for all reference currents $\boldsymbol{i}_{\mathrm{s,ref}}^{\mathrm{VSD}}$ and the speed n are determined and kept constant for the entire episode. The duration of one training episode is set to $\tau_q = L_{\mathrm{s}}^q/R_{\mathrm{s}}$.

Suitable hyperparameters for the DDPG agent are determined by random search, which is chosen as a compromise between finding agents with sufficient control performance and manageable computational effort. Only a subset of hyperparameters is considered in the random search. Specifically, the learn rates of the actor and critic are varied, as well as the

TABLE I: Hyperparameters used in the training.

Parameter	Value
Training samples	700,000
Minibatch size	64
Experience buffer length	700,000
L2 regularization actor & critic	0.01
Target network update frequency actor & critic	1
Target smooth factor	$1 \cdot 10^{-3}$
Discount factor	0.9
Learn rate critic	$1 \cdot 10^{-6}$ - $1 \cdot 10^{-3}$
Learn rate actor	$1 \cdot 10^{-6}$ - $1 \cdot 10^{-3}$
Exploration standard deviation	$1\% - 10\% \cdot V_{\max}$
Exploration decay rate	$0.1\% - 20\%$ of samples
Actor hidden layer	1
Actor neurons in hidden layer	64
Actor hidden layer activation function	ReLU
Actor output layer activation function	tanh
Critic hidden layer	3
Critic neurons in hidden layer	64/64/64
Critic hidden layer activation function	ReLU
Critic output layer activation function	linear

standard deviation and decay rate of the Ornstein-Uhlenbeck noise, which is used for exploration in DDPG [11]. The remaining parameters are set to fixed values based on the literature [10]. The resulting values or value ranges of the hyperparameters are listed in Table I. Based on Table I, 1024 agents are randomly sampled within the ranges and trained.

IV. EXPERIMENTAL SETUP

Among the trained agents, the best one—determined according to the criteria explained in Section V-A—is implemented on the control platform. Before applying the DDPG-CC to a real-world setting, it is tested in a controller-in-the-loop (CIL) environment. Moreover, a conventional control method, i.e., FOC, is designed in this section to serve as a benchmark for evaluating the performance of the proposed DDPG-CC scheme.

A. Implementation on the Control System

The UltraZohm, an open-source rapid control prototyping system, is used as the control platform [17]. It uses a Xilinx Zynq UltraScale+ MPSoC, which offers two ARM-R5 processors (RPU), four ARM-A53 processors (APU), and programmable logic (PL). The actor neural network of the DDPG-CC is implemented in the RPU of the UltraZohm. The parameters of the actor neural network, i.e., the weights and biases, are exported from Matlab after the training and implemented in the UltraZohm software; no further training on the target platform is conducted. The actor neural network is calculated in real-time with a control frequency of $f_{\mathrm{c}} = 10\,\mathrm{kHz}$. The modulator and inverter handling are realized in the PL.

B. CIL Environment

The correct implementation of the DDPG-CC in the RPU is ensured by utilizing the CIL environment offered by the
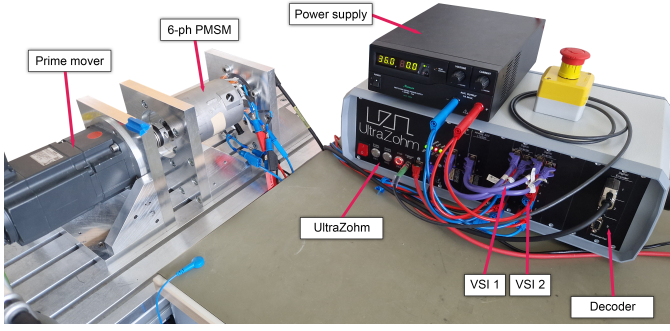
Fig. 3: Test bench setup.

TABLE II: Parameters of the PMSM, inverter, and controller.

Parameter	Symbol	Value	Unit
d-axis inductance	L_s^d	125	μH
q-axis inductance	L_s^q	126	μH
x-axis inductance	L_s^x	39	μH
y-axis inductance	L_s^y	35	μH
Stator resistance	R_s	64.3	mΩ
PM flux linkage	ψ_PM	4.7	mVs
Pole pairs	p	5	-
Rated dc-link	$V_\mathrm{dc,r}$	36	V
Rated current	I_r	18	A
Maximum current	I_max	30	A
Rated speed	n_r	1,100	min^{-1}
Control frequency	f_c	10	kHz
Switching frequency	f_PWM	10	kHz

UltraZohm. Subsequently, instead of commissioning the algorithm on the real system, a six-phase PMSM model is simulated in the PL. The machine model is formulated according to (9) and is thus identical to the model used in training, i.e., a correct implementation of the control algorithm must lead to consistent results between CIL and simulation. This procedure is beneficial for complex control algorithms, such as DDPG-CC. To this extent, the correctness and the real-time capability of the implemented algorithm on the target platform are tested over the entire operating range without the risks that commissioning on the test bench entails. Without the intermediate CIL step, discrepancies between the test bench and simulation results cannot conclusively be attributed to whether they stem from deviations between the simulation model and reality or a faulty implementation. Further details about the CIL environment are provided in [18].

C. Test Bench Setup

Fig. 3 shows the laboratory setup used for the experimental tests. The prime mover is coupled with the device-under-test (DUT), a six-phase PMSM. The prime mover sets different fixed rotational speeds, while the current control of the DUT is tested using validation profiles for the dq- and the xy-subspace. An incremental encoder is coupled with the DUT and fed into the digital incremental decoder card in the UltraZohm. The six phases of the PMSM are connected with two independent three-phase VSIs, specifically designed to be used in conjunction with the UltraZohm. They share a common dc-link voltage source. The analog measurement signals are
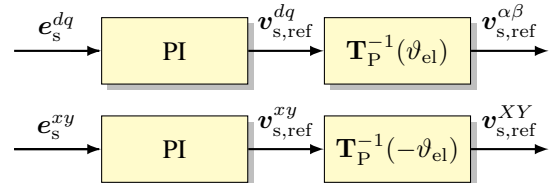

Fig. 4: FOC structure for the dq- and xy-subspace.

transmitted from both VSIs to two analog-to-digital adapter boards (LTC2311-16) inside the UltraZohm.

The parameters of the test bench setup are listed in Table II. The simulation and CIL model are built based on these parameters.

D. Conventional Current Control of Multi-Phase Machines

In the context of multi-phase machines, extensions of conventional FOC algorithms are employed to account for the additional subspaces and the occurring harmonic currents. As shown in [2] and [3], besides the proportional-integral (PI) controllers for the d- and q-currents, additional PI controllers for the x- and y-currents are used to compensate for imbalances between the three-phase subsets. Therefore, the x- and y-currents are controlled to zero. Moreover, to improve the current quality, as quantified by the current total harmonic distortion, resonant (R) controllers are added to the dq- and xy-subspace. Since the 2N topology is used in this work, no R or P controllers are required for the $0\pm$-subspace.

For a fair comparison in the context of this work, the R controllers are omitted since the DDPG-based current control approach presented in Section II-B does not aim to compensate for harmonic currents, but it only accounts for imbalances. The resulting FOC structure is shown in Fig. 4. The controllers are tuned according to the modulus optimum method.

V. PERFORMANCE ASSESSMENT

This section assesses the performance of the DDPG-CC for a six-phase PMSM. To this end, the performance of the best agent according to Section V-A is evaluated based on a validation profile of reference changes and compared with that of FOC. Test bench measurements and CIL experiments are reported hereafter. In addition, the capability of the DDPG-CC to handle imbalances between the two three-phase subsystems is discussed.

A. Evaluation of the Trained Agents

The trained agents are evaluated according to the integral absolut error (IAE) over a series of defined simulation runs. In each run, the same set of arbitrary reference-point changes for the d-, q-, x- and y-currents is simulated at a fixed speed $n \in \{0, 0.25, 0.5, 0.75, 1\} \cdot n_\mathrm{r}$. The simulation time is set to $t_\mathrm{sim} = 15\tau_q$, and the reference change occurs at $t = 5\tau_q$. The IAE criterion

$$Q_\mathrm{IAE} = \frac{1}{4} \int_0^{t_\mathrm{end}} \|e_\mathrm{s}^\mathrm{VSD}(t)\|_1 \, \mathrm{d}t \qquad (23)$$

(a) $n = 100\,\mathrm{min}^{-1}$.   (b) $n = 500\,\mathrm{min}^{-1}$.   (c) $n = 1100\,\mathrm{min}^{-1}$.
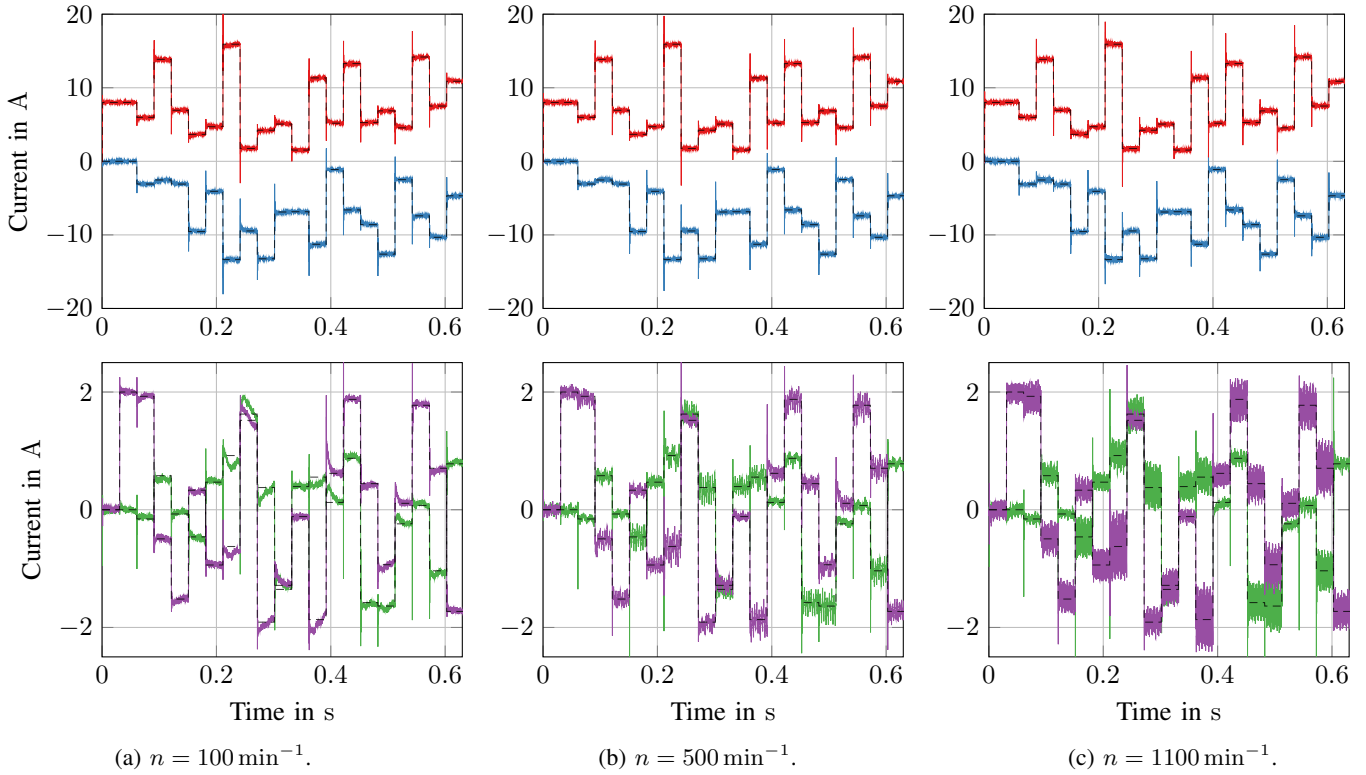
Fig. 5: Experimental results for DDPG-CC with arbitrary changes in the references of i_s^d (blue), i_s^q (red), i_s^x (green), i_s^y (purple) at different rotational speeds.

is calculated for each simulation with $t_{\mathrm{end}} = t_{\mathrm{sim}}$. The best-performing agent is determined using the arithmetic mean $\bar{Q}_{\mathrm{IAE}}$ over all simulation runs.

B. Control Performance in the dq- and xy-Subspace

A validation profile with arbitrary changes of $\boldsymbol{i}_{\mathrm{s,ref}}^{\mathrm{VSD}}$ is used to determine the control performance of the DDPG-CC algorithm over a wide range of operating points. The range of the set-point values for the x- and y-currents is smaller than that for the d- and q-currents to represent a real-world scenario. Fig. 5 shows the measurement results of the DDPG-CC for three different rotational speeds. The DDPG-CC algorithm tracks the reference values $\boldsymbol{i}_{\mathrm{s,ref}}^{\mathrm{VSD}}$ in the dq- and the xy-axis without a steady-state error for rotational speeds within the rated operation area.

Current harmonics are present in all subspaces. However, as can be observed, some harmonics increase with an increasing rotational speed while being more pronounced in the xy-subspace. The root cause of this is the presence of 5th and 7th harmonics in the induced phase voltage of the real machine. Since the control approaches under consideration are not designed to compensate for their resulting currents, the excitation of harmonic currents of corresponding orders increases with higher rotational speeds. This phenomenon is more pronounced in the xy-subspace since the current limiting impedance within this subspace is relatively small compared to the dq-subspace. As mentioned in Section III, these effects are not modeled in the training. Therefore, DDPG-CC cannot

suppress these harmonics. However, the ability to control all currents of the subspaces individually is not hindered by the fact that these effects are not modeled in the training.

C. Control of Imbalances

As described in Section IV-D, controlling the dc-component in the xy-subspace compensates for imbalances between the two three-phase subsystems. The DDPG-CC allows independent control of all subspaces (see Fig. 5). Thus, the DDPG-CC can compensate for the aforementioned imbalances. Fig. 6 shows the operation of the six-phase PMSM with a change in the reference of the y-axis from $i_{\mathrm{s,ref}}^y = 0\,\mathrm{A}$ to $i_{\mathrm{s,ref}}^y = 2\,\mathrm{A}$. While $\boldsymbol{i}_s^{xy} = \boldsymbol{0}\,\mathrm{A}$, the two three-phase subsystems are balanced. An asymmetry between the subsystems is created by specifying $i_{\mathrm{s,ref}}^y \neq 0$ as indicated by the phase currents in Fig. 6.

D. Comparison to FOC

To better assess the performance of the proposed controller, the DDPG-CC is benchmarked against the FOC scheme designed in Section IV-D. The transient behavior of the evaluated DDPG-CC agent achieves faster rise times than FOC in most operating points. This, however, occurs at the expense of higher overshoots, as illustrated in Fig. 7 for one operating point. In addition to controlling the changing reference for i_s^q, the influence on the other axes is comparable between the DDPG-CC and FOC. With regards to the steady-state behavior, as previously mentioned, the current harmonics produced by
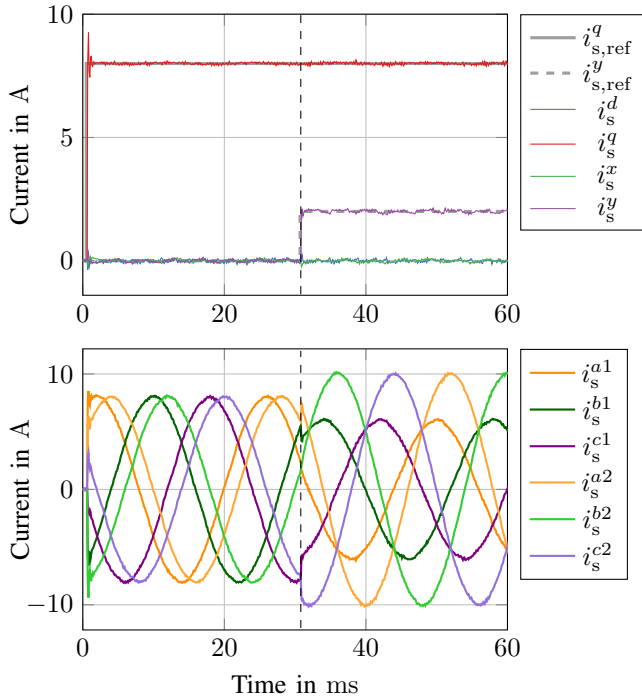
Fig. 6: Experimental results for DDPG-CC with stator currents in rotating reference frame (top) and natural reference frame (bottom) at $n = 500\,\text{min}^{-1}$. Reference $i_{\text{s,ref}}^y \neq 0$ allows control of imbalances between the two three-phase subsystems.
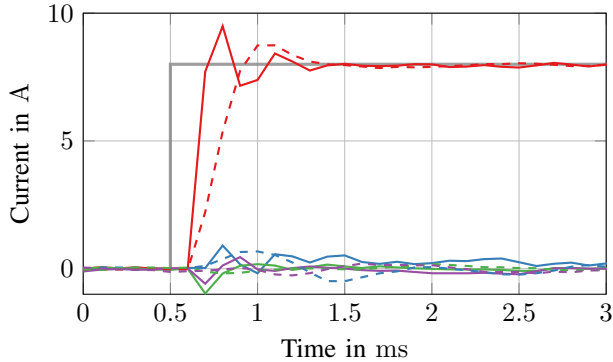

Fig. 7: Experimental results for DDPG-CC with a change in the reference $i_{\text{s,ref}}^q$ (gray) with i_s^d (blue), i_s^q (red), i_s^x (green), i_s^y (purple) and FOC (dashed) at $n = 1100\,\text{min}^{-1}$.

Fig. 8: Experimental results for the FFT of the phase currents, FOC (blue) and DDPG-CC (red) at $n = 1100\,\text{min}^{-1}$ and with $i_{\text{s,ref}}^d = -11.3\,\text{A}$, $i_{\text{s,ref}}^q = 11.3\,\text{A}$, $i_{\text{s,ref}}^x = 0.55\,\text{A}$ and $i_{\text{s,ref}}^y = -1.87\,\text{A}$.

Fig. 9: Comparison of i_s^q between FOC (blue) and DDPG-CC (red) in CIL (dashed) and measurement (solid line) at $n = 500\,\text{min}^{-1}$.

DDPG-CC increase at higher rotational speeds due to the increased back-electromotive force harmonics, see Fig. 5. Since both DDPG-CC and FOC do not actively control the current harmonics, i.e., FOC does not use R controllers and DDPG-CC is trained on a machine model without current harmonics, they show similar behavior regarding the current ripples. This is illustrated in the Fast Fourier Transform (FFT) in Fig. 8 for one exemplary operating point of the validation profile at rated speed. Overall, DDPG-CC is able to decrease the current harmonics compared to FOC. A possible explanation for this is the faster transient behavior of DDPG-CC, as indicated by
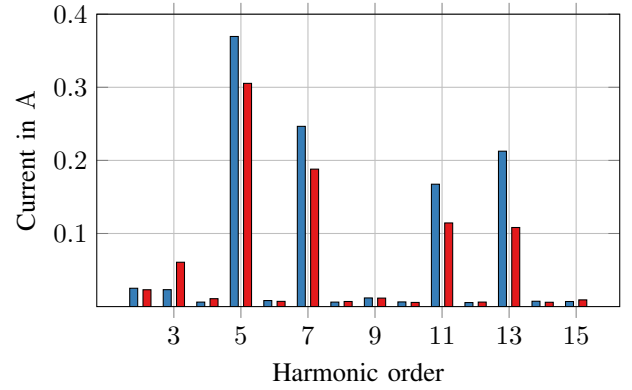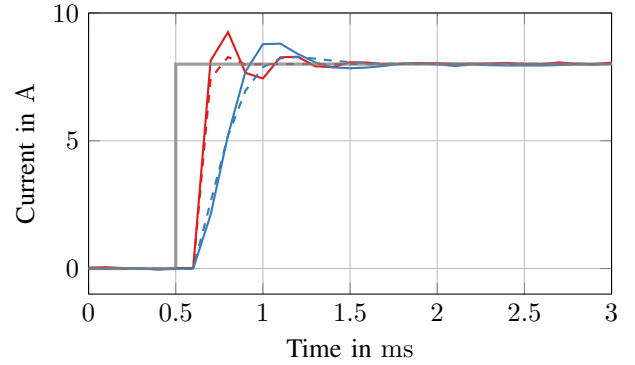
the short rise time of DDPG-CC in Fig. 7.

Following, the validity of the testing procedure is demonstrated. To this aim, Fig. 9 shows measurements of the test bench and CIL environment for the control strategies in question. Both DDPG-CC and FOC exhibit a higher overshoot on the test bench compared to the CIL environment. This indicates a discrepancy between the identified machine parameters in Table II, which are used for the training, tuning of FOC, and CIL, and the real machine. These deviations are expected due to measurement uncertainties and effects not accounted for in the model, e.g., saturation and current harmonics.

The IAE is calculated as a performance metric according to (23) using the validation profile for DDPG-CC and FOC on the test bench as well as in the CIL environment. The full validation profile of Fig. 5 is used, i.e., $t_{\text{end}} = 0.632\,\text{s}$. Fig. 10 shows the resulting IAE values at each measured rotational speed. The performance of DDPG-CC is comparable to that of FOC over the entire operating range. One noteworthy effect of DDPG-CC is the dependency of the IAE on the speed in the CIL environment. While Q_{IAE} of FOC remains nearly constant in CIL, DDPG-CC performs considerably better in
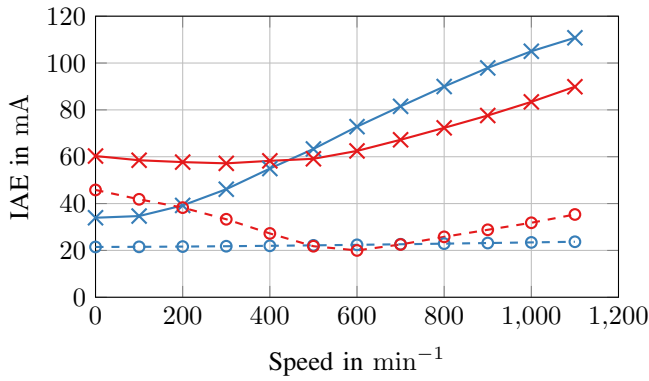
Fig. 10: Q_{IAE} of the validation profile for DDPG-CC (red) and FOC (blue) in CIL (dashed, circle) and measurement (solid line, cross) at different rotational speeds.

the region of $n_{\mathrm{r}}/2$. This indicates insufficient training at the boundaries of the rotational speed. A possible explanation is the uniform distribution of random speed values between standstill and rated speed used in the training, which skews the training towards the median of $550\,\mathrm{min}^{-1}$.

As expected, the IAE of the CIL measurements are smaller compared to the test bench measurements due to noise, parameter mismatch, and current harmonics that adversely affect the latter. The mentioned effects depend on the rotational speed, which leads to increasing IAE values at higher speeds. Due to its ability to handle current harmonics more favorably than FOC, DDPG-CC has a lower IAE at higher speeds in the test bench measurements.

VI. CONCLUSION

This paper presented a novel deep RL current control scheme based on the DDPG algorithm for six-phase PMSMs. The DDPG agent was trained exclusively in a simulation environment and the best agent according to IAE was transferred to the test bench. Moreover, to account for the physical limitations of the drive system in the training process, a limitation of the action vector of the agent was introduced. As demonstrated with the presented experimental results, the proposed DDPG-CC scheme achieves independent control of the currents in the dq- and xy-subspace, implying that compensating imbalances between the two three-phase subsystems is possible. Furthermore, as shown, the control performance, as quantified by the IAE over a defined validation profile, is comparable to that of conventional control techniques, such as FOC.

Future work could explore the possibility of determining agents with improved control performance by applying more advanced hyperparameter optimization methods than random search, such as Bayesian optimization or evolutionary algorithms. In addition, the training model could be extended to account for current harmonics. Alternatively, training DDPG-CC on the test bench is an interesting research direction. This would allow consideration of current harmonics while eliminating the discrepancies between test bench measurements and

simulations caused by the challenging identification of the exact machine parameters. Finally, applying the control concept to machines with higher-rated power should be investigated.

REFERENCES

[1] Y. Zhao and T. Lipo, "Space vector PWM control of dual three-phase induction machine using vector space decomposition," vol. 31, no. 5, pp. 1100–1109, 1995.

[2] A. G. Yepes, J. Doval-Gandoy, F. Baneira, D. Pérez-Estévez, and O. López, "Current harmonic compensation for n -phase machines with asymmetrical winding arrangement and different neutral configurations," *IEEE Transactions on Industry Applications*, vol. 53, no. 6, pp. 5426–5439, 2017.

[3] H. M. Eldeeb, A. S. Abdel-Khalik, and C. M. Hackl, "Dynamic modeling of dual three-phase IPMSM drives with different neutral configurations," vol. 66, no. 1, pp. 141–151, Jan. 2019.

[4] M. Hoerner, S. Wendel, A. Dietz, P. Karamanakos, and R. Kennel, "Variable switching point predictive current control for multi-phase permanent magnet synchronous drives," in *2021 IEEE International Conference on Predictive Control of Electrical Drives and Power Electronics (PRECEDE)*, 2021, pp. 184–189.

[5] P. F. C. Gonçalves, S. M. A. Cruz, and A. M. S. Mendes, "Predictive current control of six-phase permanent magnet synchronous machines based on virtual vectors with optimal amplitude and phase," in *2019 International Conference on Smart Energy Systems and Technologies (SEST)*, 2019, pp. 1–6.

[6] S. Zhang, O. Wallscheid, and M. Porrmann, "Machine learning for the control and monitoring of electric machine drives: Advances and trends," *IEEE Open Journal of Industry Applications*, vol. 4, pp. 188–214, 2023.

[7] D. Weber, M. Schenke, and O. Wallscheid, "Steady-state error compensation for reinforcement learning-based control of power electronic systems," *IEEE Access*, pp. 1–1, 2023.

[8] M. Schenke and O. Wallscheid, "A deep Q-learning direct torque controller for permanent magnet synchronous motors," *IEEE Open Journal of the Industrial Electronics Society*, vol. 2, pp. 388–400, 2021.

[9] G. Book, A. Traue, P. Balakrishna, A. Brosch, M. Schenke, S. Hanke, W. Kirchgässner, and O. Wallscheid, "Transferring online reinforcement learning for electric motor control from simulation to real-world experiments," *IEEE Open Journal of Power Electronics*, vol. 2, pp. 187–201, 2021.

[10] T. Schindler, L. Broghammer, P. Karamanakos, A. Dietz, and R. Kennel, "Deep reinforcement learning current control of permanent magnet synchronous machines," in *2023 IEEE International Electric Machines & Drives Conference (IEMDC)*, 2023, pp. 1–7.

[11] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning." [Online]. Available: https://arxiv.org/pdf/1509.02971.pdf

[12] A. A. Rockhill and T. A. Lipo, "A generalized transformation methodology for polyphase electric machines and networks," Coeur d'Alene, ID, USA, May 2015, pp. 27–34.

[13] P. Q. Nguyen and D. Jörg-Andreas, *Vector control of three-phase AC Machines*. Springer, 2015.

[14] A. G. Yepes, J. Doval-Gandoy, and H. A. Toliyat, "Multifrequency current control for n-phase machines including antiwindup and distortion-free saturation with full dc-bus utilization," *IEEE Transactions on Power Electronics*, vol. 34, no. 10, pp. 9891–9905, 2019.

[15] E. Levi, D. Dujic, M. Jones, and G. Grandi, "Analytical determination of dc-bus utilization limits in multiphase vsi supplied ac drives," *IEEE Transactions on Energy Conversion*, vol. 23, no. 2, pp. 433–443, 2008.

[16] "Matlab documentation: deep deterministic policy gradient (DDPG) agents," https://de.mathworks.com/help/reinforcement-learning/ug/ddpg-agents.html, accessed: 2023-07-26.

[17] S. Wendel, A. Geiger, E. Liegmann, D. Arancibia, E. Durán, T. Kreppel, F. Rojas, F. Popp-Nowak, M. Diaz, A. Dietz, R. Kennel, and B. Wagner, "UltraZohm— a powerful real-time computation platform for MPC and multi-level inverters," Quanzhou, China, 2019, pp. 1–6.

[18] E. Liegmann, T. Schindler, P. Karamanakos, A. Dietz, and R. Kennel, "UltraZohm - an open-source rapid control prototyping platform for power electronic systems," in *2021 International Aegean Conference on Electrical Machines and Power Electronics (ACEMP) & 2021 International Conference on Optimization of Electrical and Electronic Equipment (OPTIM)*, 2021, pp. 445–450.