Steady-State Error Reduction of Reinforcement Learning based Indirect Current Control of Permanent Magnet Synchronous Machines

Tobias Schindler[1,2], Lara Broghammer[1], Dennis Hufnagel[1], Nina Diringer[1], Benedikt Hofmann[1], Armin Dietz[1], Petros Karamanakos[3], Ralph Kennel[2]

[1] Technische Hochschule Nuremberg, Germany
[2] Technical University of Munich, Germany
[3] Tampere University, Finland

Corresponding author: Tobias Schindler, tobias.schindler@th-nuernberg.de
Speaker: Tobias Schindler, tobias.schindler@th-nuernberg.de

Abstract

Deep reinforcement learning (DRL) can achieve favorable dynamic performance compared to conventional control methods. However, steady-state errors are often present. This paper investigates the reduction of steady-state error in DRL-based current control of permanent magnet synchronous machines (PMSMs) by augmenting the integrated tracking error to the observation vector. More specifically, this paper assesses the performance of a DRL-based method under nominal and adverse operating conditions by considering PMSMs with linear and nonlinear magnetic circuits, which exhibit saturation, cross-coupling, and spatial harmonics. The latter include parameter mismatches between the training model and the physical system and misalignment of the dq-frame with respect to the identified position of the d-axis. As shown with the presented experimental results, the DRL-based control method can successfully operate the drive system under all operating conditions, with the steady-state and dynamic performance being similar to that of field-oriented control.

1 Introduction

Field-oriented control (FOC) is the state-of-the-art control scheme for permanent magnet synchronous machines (PMSMs). While FOC performs reasonably well in many applications, tuning the controller can be cumbersome. This is particularly true for drive systems with PMSMs that exhibit a non-linear magnetic circuit, e.g., highly utilized PMSMs for traction applications [1], more-electric aircraft, or machines without rare-earth magnets [2]. Controlling these more challenging machine types requires additional effort, e.g., implementing gain scheduling and decoupling based on flux maps [1], [3].

Deep reinforcement learning (DRL) control algorithms offer an alternative to FOC, potentially improving transient performance and reducing the tuning effort. On the one hand, several publications indicate that deep deterministic policy gradient (DDPG) [4] is a promising DRL algorithm for current control of PMSMs [5], [6], [7]. On the other hand, a persistent steady-state error (SSE) is frequently present in DRL-based control concepts and also in applications other than electrical machine control. In [8], SSE is present in some operating points when controlling plasma in a tokamak. Using recurrent neural networks in the actor to compensate for the SSE is suggested [8]. In [7], a DDPG-based current controller is trained in a simulation environment using a linear machine model. Training continues on the test bench, where the used PMSM exhibits significant saturation. [7] reports an average normalized SSE of $\approx 1\%$ and concurs with [8] on the use of recurrent neural networks. A method to deal with SSE in DRL-based controllers is to add a dedicated integral action, as shown in [9]. The proposed concept features an additional control output, i.e., one action per control signal is fed to an integrator and added to the existing action. Previous work of the authors, presented in [6], found a persisting SSE in DDPG-based indirect current control (DDPG-CC) of a PMSM with linear magnetic circuit. The SSE is compensated by augmenting the observation of the DDPG agent with a

discrete-time integrator for the tracking error.
This paper examines whether DDPG-CC can effectively address SSE when controlling more challenging PMSMs. Furthermore, adverse operating conditions are considered in which the assumptions of the training model are not applicable.
Two scenarios for adverse operating conditions are carried out. In the first scenario, additional resistors are added to the machine. The resistance varies within production batches [10], and motor cables of different lengths may be used in the installation. These effects are typically not accounted for when tuning the controller, and FOC is robust regarding these slight deviations. The effective stator resistance also depends on the frequency of the stator currents (skin & proximity effect) and the rotor temperature. Both effects can be modeled but are often neglected. Therefore, the ability of DDPG-CC to control PMSM when the stator resistance does not match the model value is investigated.
FOC and DDPG-CC control the stator currents in the rotating dq-frame, which has to be aligned with the α-axis of the machine. Sophisticated algorithms to perform the alignment exist [11], but the alignment is always imperfect due to measurement uncertainty. Additionally, the precision and bandwidth of the rotor position measurement are limited during operation. Therefore, some degree of misalignment, i.e., a difference between the measured position of the d-axis used by the controller and the d-axis of the machine, is always present in the control system. Thus, the second scenario investigates the effect of an intentionally induced misalignment of the dq-frame used by DDPG-CC.
A case study is carried out on two machines. One machine has a linear magnetic circuit, and the other shows significant saturation and spatial harmonics. DDPG-CC agents are trained on machine models with different degrees of fidelity to account for these effects. Control performance is evaluated for adverse scenarios and nominal operating conditions, i.e., the absence of adverse scenarios.

2 Model

The stator currents of the three-phase PMSM in the abc-frame are given by $\boldsymbol{i}_\mathrm{s}^{abc} = [i_\mathrm{s}^a \; i_\mathrm{s}^b \; i_\mathrm{s}^c]^\top$. The reduced, amplitude-invariant Clarke transformation

$$\boldsymbol{T}_\mathrm{c} = \frac{2}{3}\begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix} \tag{1}$$

is used to transform the phase quantities of the stator variables into the stator-fixed $\alpha\beta$-frame, e.g., the stator currents $\boldsymbol{i}_\mathrm{s}^{\alpha\beta} = \boldsymbol{T}_\mathrm{c}\boldsymbol{i}_\mathrm{s}^{abc}$. Note that the system is assumed to be balanced. Therefore, the zero component i_s^0 is neglected. Applying the Park transformation

$$\boldsymbol{T}_\mathrm{P}(\varphi) = \begin{bmatrix} \cos(\varphi) & \sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{bmatrix} \tag{2}$$

yields the quantities in the rotating dq-frame, e.g., the stator currents $\boldsymbol{i}_\mathrm{s}^{dq} = \begin{bmatrix} i_\mathrm{s}^d & i_\mathrm{s}^q \end{bmatrix}^\top$ using $\boldsymbol{i}_\mathrm{s}^{dq} = \boldsymbol{T}_\mathrm{P}(\varphi)\boldsymbol{i}_\mathrm{s}^{\alpha\beta}$. Herein, the rotor angle φ is defined between the α- and the d-axis of the machine. With (1) and (2), the generic PMSM model in the rotating dq-frame is given by

$$\boldsymbol{v}_\mathrm{s}^{dq} = R_\mathrm{s}\boldsymbol{i}_\mathrm{s}^{dq} + \omega_\mathrm{el}\boldsymbol{J}\boldsymbol{\psi}_\mathrm{s}^{dq}(\boldsymbol{i}_\mathrm{s}^{dq}, \varphi) + \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\psi}_\mathrm{s}^{dq}(\boldsymbol{i}_\mathrm{s}^{dq}, \varphi)\,, \tag{3}$$

with the stator voltages $\boldsymbol{v}_\mathrm{s}^{dq}$, the stator currents $\boldsymbol{i}_\mathrm{s}^{dq}$, the per-phase stator resistance R_s, the flux-linkage $\boldsymbol{\psi}_\mathrm{s}^{dq}$, and $\boldsymbol{J} = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$ [3], [10], [12]. The electrical angular speed $\omega_\mathrm{el} = \omega_\mathrm{m} p$ is given by the rotational angular speed ω_m and the number of pole pairs p.

2.1 Linear Case

Assuming a linear magnetic circuit, the stator flux-linkages in d- and q-axis are given by

$$\psi_\mathrm{s}^d = \psi_\mathrm{PM}^d + L_\mathrm{s}^d i_\mathrm{s}^d\,, \tag{4}$$
$$\psi_\mathrm{s}^q = L_\mathrm{s}^q i_\mathrm{s}^q\,, \tag{5}$$

and the machine model is described using

$$v_\mathrm{s}^d = R_\mathrm{s} i_\mathrm{s}^d + L_\mathrm{s}^d \frac{\mathrm{d}i_\mathrm{s}^d}{\mathrm{d}t} - \omega_\mathrm{el}\psi_\mathrm{s}^q\,, \tag{6a}$$
$$v_\mathrm{s}^q = R_\mathrm{s} i_\mathrm{s}^q + L_\mathrm{s}^q \frac{\mathrm{d}i_\mathrm{s}^q}{\mathrm{d}t} + \omega_\mathrm{el}\psi_\mathrm{s}^d\,, \tag{6b}$$

where ψ_PM^d is the flux-linkage of the permanent magnets, L_s^d is the absolute stator inductance in the direct axis, and L_s^q is the absolute stator inductance in the quadrature axis.

2.2 Linear Case including Iron Losses

The iron losses are modeled by an additional parallel resistance in the equivalent circuit as depicted in Fig. 1 [11], [14], [13]. The machine model, including

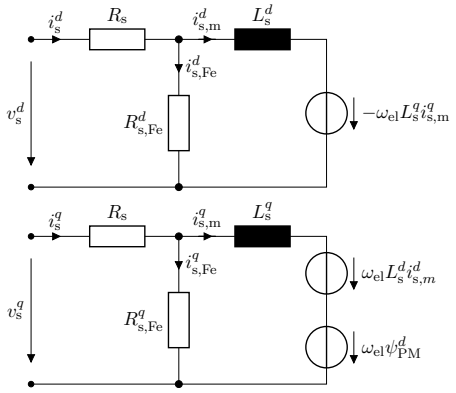

Fig. 1: Equivalent circuit of fundamental wave model including fixed iron loss resistance [11], [13].

iron losses [14] is given by

$$v_\mathrm{s}^d = i_\mathrm{s}^d \left(R_\mathrm{s} + \frac{L_\mathrm{s}^d L_\mathrm{s}^q \omega_\mathrm{el}^2}{R_\mathrm{s,Fe}^q} \right) - \omega_\mathrm{el} L_\mathrm{s}^q i_\mathrm{s}^q + L_\mathrm{s}^d \frac{\mathrm{d}i_\mathrm{s}^d}{\mathrm{d}t} + \frac{L_\mathrm{s}^q \omega_\mathrm{el}^2 \psi_\mathrm{PM}^d}{R_\mathrm{s,Fe}^q}\,, \tag{7a}$$

$$v_\mathrm{s}^q = i_\mathrm{s}^q \left(R_\mathrm{s} + \frac{L_\mathrm{s}^d L_\mathrm{s}^q \omega_\mathrm{el}^2}{R_\mathrm{s,Fe}^d} \right) + \omega_\mathrm{el} L_\mathrm{s}^d i_\mathrm{s}^d + L_\mathrm{s}^q \frac{\mathrm{d}i_\mathrm{s}^q}{\mathrm{d}t} + \omega_\mathrm{el} \psi_\mathrm{PM}^d\,, \tag{7b}$$

with the equivalent iron loss resistances $R_\mathrm{s,Fe}^d$ and $R_\mathrm{s,Fe}^q$ in the d- and q-axis, respectively. The currents $i_\mathrm{s,Fe}^d$ and $i_\mathrm{s,Fe}^q$ account for the power loss due to iron losses [11]

$$P_{Fe} = \frac{3}{2} \left(R_\mathrm{s,Fe}^d (i_\mathrm{s,Fe}^d)^2 + R_\mathrm{s,Fe}^q (i_\mathrm{s,Fe}^q)^2 \right). \tag{8}$$

2.3 Saturation, Cross-coupling, and Spatial Harmonics

A model that accounts for saturation, cross-coupling, and the dependency on the rotor angle is presented in [3], [12], [15]. The stator flux in d- and q-axis is assumed to be a function of the stator current $\boldsymbol{i}_\mathrm{s}^{dq}$ and the rotor angle φ [3] [12] described by

$$\psi_\mathrm{s}^d = f_1(\boldsymbol{i}_\mathrm{s}^{dq}, \varphi)\,, \tag{9a}$$

$$\psi_\mathrm{s}^q = g_1(\boldsymbol{i}_\mathrm{s}^{dq}, \varphi)\,. \tag{9b}$$

The differential inductances

$$\frac{\partial \psi_\mathrm{s}^d(\boldsymbol{i}_\mathrm{s}^{dq}, \varphi)}{\partial i_\mathrm{s}^d} = L_\mathrm{s}^{dd}(\boldsymbol{i}_\mathrm{s}^{dq}, \varphi)\,, \tag{10a}$$

$$\frac{\partial \psi_\mathrm{s}^q(\boldsymbol{i}_\mathrm{s}^{dq}, \varphi)}{\partial i_\mathrm{s}^q} = L_\mathrm{s}^{qq}(\boldsymbol{i}_\mathrm{s}^{dq}, \varphi)\,, \tag{10b}$$

$$\frac{\partial \psi_\mathrm{s}^d(\boldsymbol{i}_\mathrm{s}^{dq}, \varphi)}{\partial i_\mathrm{s}^q} = L_\mathrm{s}^{dq}(\boldsymbol{i}_\mathrm{s}^{dq}, \varphi)\,, \tag{10c}$$

$$\frac{\partial \psi_\mathrm{s}^q(\boldsymbol{i}_\mathrm{s}^{dq}, \varphi)}{\partial i_\mathrm{s}^d} = L_\mathrm{s}^{qd}(\boldsymbol{i}_\mathrm{s}^{dq}, \varphi) \tag{10d}$$

are defined by the partial derivative of the stator flux $\boldsymbol{\psi}_\mathrm{s}^{dq}(\boldsymbol{i}_\mathrm{s}^{dq}, \varphi)$ with respect to the stator currents $\boldsymbol{i}_\mathrm{s}^{dq}$. In addition, the partial derivative of the stator flux-linkage $\boldsymbol{\psi}_\mathrm{s}^{dq}(\boldsymbol{i}_\mathrm{s}^{dq}, \varphi)$ with respect to the rotor angle φ is given by

$$\frac{\partial \psi_\mathrm{s}^d(\boldsymbol{i}_\mathrm{s}^{dq}, \varphi)}{\partial \varphi} = \Lambda_\mathrm{s}^d(\boldsymbol{i}_\mathrm{s}^{dq}, \varphi)\,, \tag{11a}$$

$$\frac{\partial \psi_\mathrm{s}^q(\boldsymbol{i}_\mathrm{s}^{dq}, \varphi)}{\partial \varphi} = \Lambda_\mathrm{s}^q(\boldsymbol{i}_\mathrm{s}^{dq}, \varphi)\,. \tag{11b}$$

Note that the dependency of the stator flux $\boldsymbol{\psi}_\mathrm{s}^{dq}(\boldsymbol{i}_\mathrm{s}^{dq}, \varphi)$ (and other quantities such as differential inductances) on the stator current and rotor angle is omitted in the following notation for readability. Rewriting (3) using (10) and (11) yields

$$v_\mathrm{s}^d = R_\mathrm{s} i_\mathrm{s}^d + L_\mathrm{s}^{dd} \frac{\mathrm{d}i_\mathrm{s}^d}{\mathrm{d}t} + L_\mathrm{s}^{dq} \frac{\mathrm{d}i_\mathrm{s}^q}{\mathrm{d}t} + \omega_\mathrm{el} (\Lambda_\mathrm{s}^d - \psi_\mathrm{s}^q)\,, \tag{12a}$$

$$v_\mathrm{s}^q = R_\mathrm{s} i_\mathrm{s}^q + L_\mathrm{s}^{qq} \frac{\mathrm{d}i_\mathrm{s}^q}{\mathrm{d}t} + L_\mathrm{s}^{qd} \frac{\mathrm{d}i_\mathrm{s}^d}{\mathrm{d}t} + \omega_\mathrm{el} (\Lambda_\mathrm{s}^q + \psi_\mathrm{s}^d)\,. \tag{12b}$$

2.4 Saturation and Cross-coupling

If the spatial harmonics are neglected, i.e., if the flux-linkage $\boldsymbol{\psi}_\mathrm{s}^{dq}(\boldsymbol{i}_\mathrm{s}^{dq})$ is assumed to be independent of the rotor angle ($\Lambda_\mathrm{s}^{dq} = 0$), the model described in (12) is simplified to

$$v_\mathrm{s}^d = R_\mathrm{s} i_\mathrm{s}^d + L_\mathrm{s}^{dd} \frac{\mathrm{d}i_\mathrm{s}^d}{\mathrm{d}t} + L_\mathrm{s}^{dq} \frac{\mathrm{d}i_\mathrm{s}^q}{\mathrm{d}t} - \omega_\mathrm{el} \psi_\mathrm{s}^q\,, \tag{13a}$$

$$v_\mathrm{s}^q = R_\mathrm{s} i_\mathrm{s}^q + L_\mathrm{s}^{qq} \frac{\mathrm{d}i_\mathrm{s}^q}{\mathrm{d}t} + L_\mathrm{s}^{qd} \frac{\mathrm{d}i_\mathrm{s}^d}{\mathrm{d}t} + \omega_\mathrm{el} \psi_\mathrm{s}^d\,. \tag{13b}$$

3 Control Algorithm

This section briefly summarizes the control concept of DDPG-CC, as introduced in [6], for drive systems consisting of a three-phase PMSM and a two-level voltage source inverter (VSI). Additionally, the used reference FOC is described.

3.1 DDPG-CC

The stator currents $\boldsymbol{i}_\mathrm{s}^{dq}$ of the PMSM are controlled in the rotating dq-plane to track the reference currents $\boldsymbol{i}_{\mathrm{s,ref}}^{dq} = [i_{\mathrm{s,ref}}^d \; i_{\mathrm{s,ref}}^q]^\top$ and minimize the tracking error $\boldsymbol{e}_\mathrm{s}^{dq} = \boldsymbol{i}_{\mathrm{s,ref}}^{dq} - \boldsymbol{i}_\mathrm{s}^{dq}$. The control output (action) of DDPG-CC is defined as $\boldsymbol{v}_{\mathrm{s,ref}}^{dq} = [v_{\mathrm{s,ref}}^d \; v_{\mathrm{s,ref}}^q]^\top$. Since the available output voltage of the VSI is limited by the available dc-link V_dc, the applied action is limited to the linear region of a space vector modulation (SVM) $\boldsymbol{v}_{\mathrm{s,lim}}^{dq} = [v_{\mathrm{s,lim}}^d \; v_{\mathrm{s,lim}}^q]^\top$. Therefore, the modulation index is $m \leq 2/\sqrt{3}$ and $\|\boldsymbol{v}_{dq}^*\|_2 \leq V_\mathrm{dc}/\sqrt{3}$ applies. The limitation is implemented according to [16].

Similar to [6], two different observation vectors are considered. The observation

$$\boldsymbol{o}_1(k) = \begin{bmatrix} \boldsymbol{e}_\mathrm{s}^{dq}(k) \\ \int \boldsymbol{e}_\mathrm{s}^{dq}(k) \cdot f_\mathrm{c} \\ \boldsymbol{i}_\mathrm{s}^{dq}(k) \\ \boldsymbol{v}_{\mathrm{s,lim}}^{dq}(k-1) \\ n(k) \end{bmatrix} \tag{14}$$

features the integral of the tracking error using Euler-forward discretization, scaled by the control frequency f_c and the rotational speed n. The observation

$$\boldsymbol{o}_2(k) = \begin{bmatrix} \boldsymbol{e}_\mathrm{s}^{dq}(k) \\ \boldsymbol{i}_\mathrm{s}^{dq}(k) \\ \boldsymbol{v}_{\mathrm{s,lim}}^{dq}(k-1) \\ n(k) \end{bmatrix} \tag{15}$$

does not feature the integral over the tracking error. Note that all quantities are normalized by their rated values.

The investigated agents are trained using the reward

$$r_1(k) = \begin{cases} -\|\boldsymbol{e}_\mathrm{s}^{dq}(k)\|_1, & \text{for} \quad i_1(k) \leq I_\mathrm{max} \\ -\|\boldsymbol{e}_\mathrm{s}^{dq}(k)\|_1 - i_1(k), & \text{for} \quad i_1(k) > I_\mathrm{max} \end{cases} \tag{16}$$

where an additional penalty is given to the agent if the amplitude of the current $i_1(k) = \|\boldsymbol{i}_\mathrm{s}^{dq}\|_2$ exceeds the maximum allowed current I_max.

3.2 Field-oriented control

The reference FOC, including a decoupling network, is tuned according to the modulus optimum method [1]. The small time constants of the system are lumped into the time constant $\tau_\sigma \approx 1.5 f_\mathrm{c}^{-1}$ with the control frequency f_c. For machines with a linear magnetic circuit, the controller parameters

$$K_\mathrm{I}^d = \frac{R_\mathrm{s}}{2\tau_\sigma}\,, \qquad K_\mathrm{I}^q = \frac{R_\mathrm{s}}{2\tau_\sigma}\,, \tag{17a}$$

$$K_\mathrm{P}^d = \frac{L_\mathrm{s}^d}{2\tau_\sigma}\,, \qquad K_\mathrm{P}^q = \frac{L_\mathrm{s}^q}{2\tau_\sigma} \tag{17b}$$

are used. For non-linear machines, i.e., machines that exhibit saturation behavior, a gain scheduling approach [1] is used for the parallel PI controllers with

$$K_\mathrm{P}^d = \frac{\psi_\mathrm{s}^d(i_{\mathrm{s,ref}}^d, i_\mathrm{s}^q) - \psi_\mathrm{s}^d(i_\mathrm{s}^d, i_\mathrm{s}^q)}{2\tau_\sigma e_\mathrm{s}^d}\,, \tag{18a}$$

$$K_\mathrm{P}^q = \frac{\psi_\mathrm{s}^q(i_\mathrm{s}^d, i_{\mathrm{s,ref}}^q) - \psi_\mathrm{s}^q(i_\mathrm{s}^d, i_\mathrm{s}^q)}{2\tau_\sigma e_\mathrm{s}^q}\,. \tag{18b}$$

Decoupling is achieved by adding

$$v_{\mathrm{s,decoup}}^d = -\omega_\mathrm{el}\psi_\mathrm{s}^q(\boldsymbol{i}_\mathrm{s}^{dq})\,, \tag{19a}$$

$$v_{\mathrm{s,decoup}}^q = \omega_\mathrm{el}\psi_\mathrm{s}^d(\boldsymbol{i}_\mathrm{s}^{dq}) \tag{19b}$$

to the outputs of the PI controllers.

Tab. 1: Drive parameters for M1 (Heidrive HMD06-005) and M2 (prototype).

Parameter	Symbol	Unit	M1	M2
d-axis inductance	L_s^d	mH	1.13	0.44
q-axis inductance	L_s^q	mH	1.42	2.45
Stator resistance	R_s	mΩ	543	249
PM flux linkage	ψ_PM^d	mVs	16.9	20
Pole pairs	p		3	4
Rated dc-link	$V_{\mathrm{dc,r}}$	V	48	48
Rated speed	n_r	min^{-1}	3000	1000
Rated current	I_r	A	4.2	15
Maximum current	I_max	A	10.8	20
Control frequency	f_c	kHz	10	10

Tab. 2: Trained agents, the used observation vector, the training model, and the machine.

Parameter	A1	A2	A3	A4	A5
Observation	(14)	(15)	(14)	(15)	(14)
Integrator	✓	✗	✓	✗	✓
Model	(6)	(6)	(13)	(13)	(12)
Spatial harmonics	✗	✗	✗	✗	✓
Saturation	✗	✗	✓	✓	✓
Cross-coupling	✗	✗	✓	✓	✓
Machine	M1	M1	M2	M2	M2

4 Experimental Setup

Two different machines in three operating conditions, see Sec. 4.1, are used to evaluate the performance of DDPG-CC. Table 1 lists the nominal

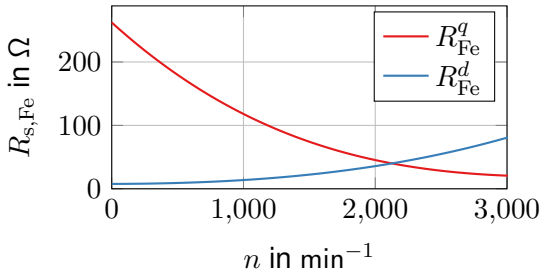

Fig. 2: Identified equivalent iron loss resistance of M1.

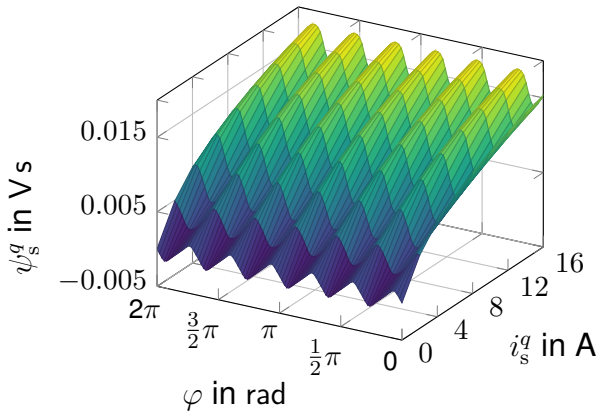

Fig. 3: Flux map ψ^q_{s} at $i^d_{\mathrm{s}} = 0\,\mathrm{A}$ with spatial harmonics of M2 (FEA data).

values of the machines. Machine M1 is a commercial off-the-shelf PMSM with low anisotropy. It is also used in [6].

The linear model (6) for M1 is parameterized based on the values presented in Table 1. To consider iron losses as in (7), the equivalent iron loss resistance $\boldsymbol{R}^{dq}_{\mathrm{s},Fe}$ is identified according to [11]. For simulation, the equivalent iron loss resistance is assumed to be only dependent on the rotational speed n, i.e., the dependency on $\boldsymbol{i}^{dq}_{\mathrm{s}}$ is neglected. The identification result is depicted in Fig. 2.

The prototype machine M2, which serves as a benchmark, is intentionally designed to exhibit significant saturation and spatial harmonics, see Fig. 3. Thus, this machine is modeled based on flux maps accounting for spatial harmonics by (12) using $\boldsymbol{\psi}^{dq}_{\mathrm{s}}(\boldsymbol{i}^{dq}_{\mathrm{s}}, \varphi)$ based on FEA data. The model neglecting spatial harmonics in (13) is parametrized by $\boldsymbol{\psi}^{dq}_{\mathrm{s}}(\boldsymbol{i}^{dq}_{\mathrm{s}})$ using FEA data as well. A detailed description of M2 is presented in [10].

Table 2 shows the evaluated agents, along with their corresponding training models and observation vectors. Agents A1 and A2 control machine M1 and are taken from [6] without modification. Agents A3 - A5 are trained to control M2 according to Sec. 4.2. The agents A1, A3, and A5 use the observation vector with the augmented integral of the tracking error.

DDPG-CC is compared to FOC in Sec. 5. FOC is tuned according to (17) for M1 based on the machine parameters of Table 1. To control M2, gain scheduling and decoupling according to (18) and (19) is employed. For this, the controller requires the flux maps of $\boldsymbol{\psi}^{dq}_{\mathrm{s}}(\boldsymbol{i}^{dq}_{\mathrm{s}})$ during runtime. The flux maps are implemented using the analytical approximation approach of [2].

DDPG-CC and FOC are implemented on the R5 processor of the UltraZohm, see [6], [17].

4.1 Operating conditions

The performance of all controllers is investigated in three different operating conditions:

1. Nominal operating condition. The d-axis of the controller is aligned to the α-axis of the machine according to the principle outlined in [11], which accounts for iron losses in the machine. No changes to the drive system.

2. Increased stator resistance by adding $100\,\mathrm{m\Omega}$ resistors to each phase of the machines. This corresponds to an increase of $20\,\%$ and $40\,\%$ for M1 and M2, respectively, compared to Tab. 1.

3. Induced misalignment of the dq-frame by 5° compared to nominal operating conditions. Effectively, the Park transformation $\boldsymbol{T}_{\mathrm{P}}\big(\varphi(k)+5^\circ\big)$ and the inverse Park transformation $\boldsymbol{T}^{-1}_{\mathrm{P}}\big(\varphi(k)+5^\circ\big)$ are used by the controller.

Note that all reported measurement and simulation data are the sampled data of the controller, which is misaligned in operating condition 3.

4.2 Training

The DDPG-CC agents A3 - A5 are trained using the Matlab/Simulink Reinforcement Learning Toolbox 2023a [18]. The machine models (12) and (13) are used for M2 as per Table 2. The VSI is assumed to be ideal in the training environment. For each agent, a set of 60 hyperparameter combinations is trained using random search to find appropriate values for the discount factor, critic and actor learn rates and exploration standard deviation. Table 3 provides a summary of the hyperparameters.

The reference currents $\boldsymbol{i}^{dq}_{\mathrm{s,ref}}$ and rotational speed n are set to uniformly distributed random values within the machine's operating range during each

Tab. 3: Hyperparameter of agents A3 - A5.

Parameter	Value
Training samples	1,8 Mio.
Minibatch size	64
Experience buffer length	900,000
L2 regularization actor & critic	0.01
Target network update frequency actor & critic	1
Target smooth factor	$1 \cdot 10^{-3}$
Discount factor	$0.95 - 0.999$
Learn rate critic	$1 \cdot 10^{-6} - 1 \cdot 10^{-4}$
Learn rate actor	$1 \cdot 10^{-6} - 1 \cdot 10^{-4}$
Exploration standard deviation	$0.1\,\% - 1\,\% V_{\max}$
Exploration decay rate	$10\,\%$ of samples
Actor neurons in hidden layer	64
Actor hidden layer	1
Actor hidden layer activation function	ReLU
Actor output layer activation function	tanh
Critic neurons in hidden layer	256
Critic hidden layer	5
Critic hidden layer activation function	ReLU
Critic output layer activation function	linear

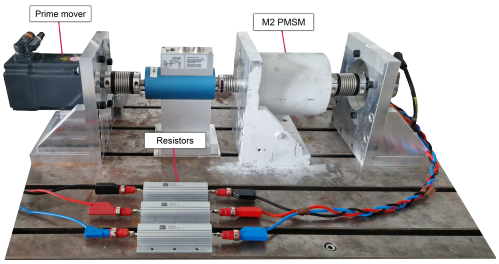

Fig. 4: Test bench with M2 and additional resistors.

training episode. These values are then kept constant for the duration of the episode. The duration of each training episode is set to $7 \cdot L_\mathrm{s}^q / R_\mathrm{s}$. Training 60 hyperparameter combinations for 1,800,000 samples in parallel takes 14 h on a workstation with an AMD Ryzen Threadripper Pro 3995WX.

5 Performance evaluation

Performance evaluation of the trained agents is done on the test bench in nominal and adverse operating conditions according to Sec. 4.1. Measurements are conducted at different rotating speeds set by a speed-controlled prime mover. Fig. 4 shows the test bench setup with M2, including the additional phase resistance. The test bench setup for M1 can be found in [6].

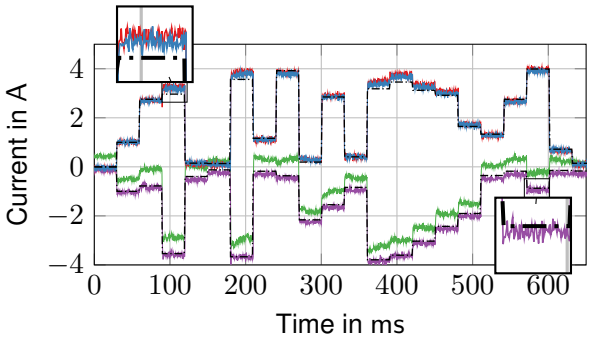

Fig. 5: Measured currents of i_s^d (purple) and i_s^q (blue) with sampling delay compensation for φ and without compensation i_s^d (green) and i_s^q (red) of A2 controlling M1 at $n = 3000\,\mathrm{min}^{-1}$. Reference values are black.

An evaluation profile with arbitrarily chosen reference values for $\boldsymbol{i}_{\mathrm{s,ref}}^{dq}$ within the machine's operating range is used. For machine M1, the evaluation profile is depicted in Fig. 5 and matches [6]. For machine M2, the profile is scaled by the rated current of M2, and the duration between reference changes is increased from $30\,\mathrm{ms}$ to $1.1\,\mathrm{s}$. This allows for the calculation of the Fast Fourier Transform (FFT) in Sec. 5.4.
The steady-state performance is quantified by the mean tracking-error

$$\bar{Q}_\mathrm{SSE} = \frac{1}{22} \sum_{m=1}^{22} \frac{100\,\%}{2 I_\mathrm{r}} \| \bar{\boldsymbol{e}}_\mathrm{s}^{dq}(m) \|_2^2 \qquad (20)$$

over all $m = 22$ reference values $\boldsymbol{i}_{\mathrm{s,ref}}^{dq}$ in the evaluation profile. The mean tracking-error $\bar{\boldsymbol{e}}_\mathrm{s}^{dq}$ is averaged over $20\,\mathrm{ms}$ at each operating point.
Transient performance is evaluated using the average of the integral absolute error (IAE) criterion

$$\bar{Q}_\mathrm{IAE} = \frac{1}{22} \sum_{m=1}^{22} \int_{t_\mathrm{start}}^{t_\mathrm{end}} \| \bar{\boldsymbol{e}}_\mathrm{s}^{dq}(m) \|_2^2 \, \mathrm{d}t \,. \qquad (21)$$

Here, t_start is set to each reference change in the validation profile and $t_\mathrm{end} = t_\mathrm{start} + 5\,\mathrm{ms}$ for M1 and $t_\mathrm{end} = t_\mathrm{start} + 10\,\mathrm{ms}$ for M2. Similar to $\bar{Q}_\mathrm{SSE}$, the mean IAE $\bar{Q}_\mathrm{IAE}$ over all operating points is reported.

5.1 Sample Delay Compensation

The SSE of agent A2, which does not use the augmented integral of the tracking error, increases with the rotational speed of the machine in [6]. The measured SSE is much more significant compared to simulation results. Due to faulty implementation in [6], the sample delay of acquiring $\varphi(k)$ is

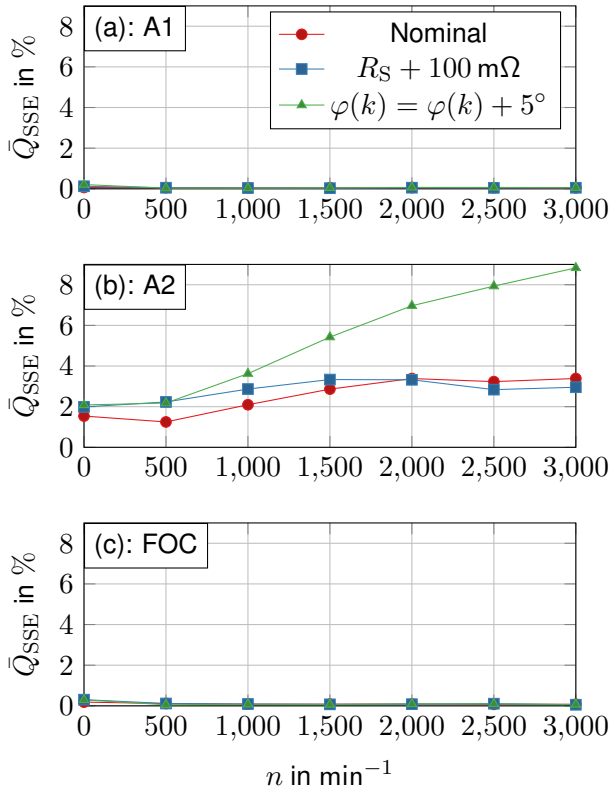

Fig. 6: Measured average SSE of controllers for M1 in different operation conditions.

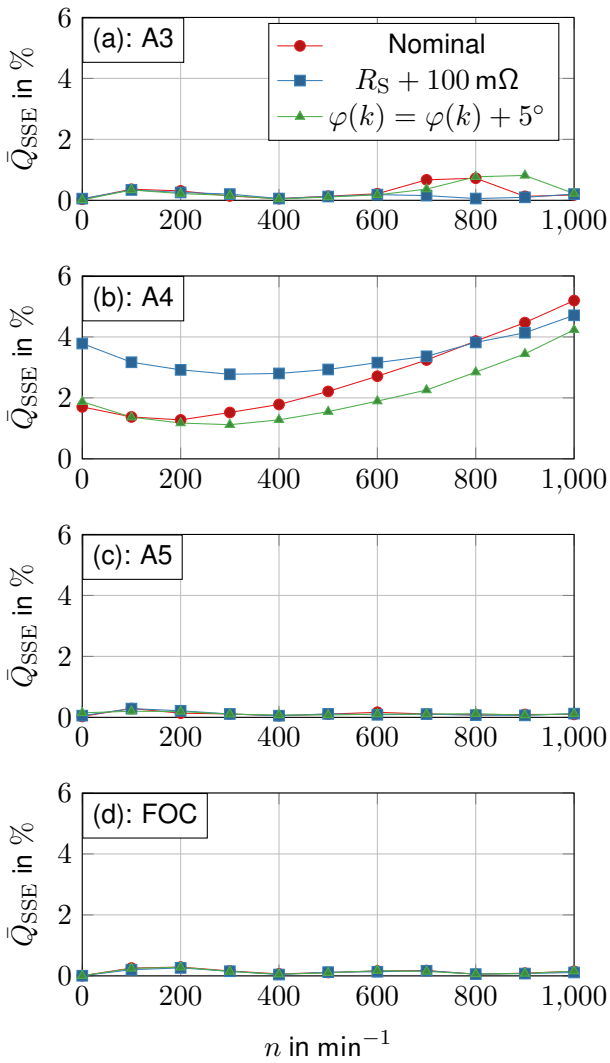

Fig. 7: Measured average SSE of controllers for M2 in different operation conditions.

not compensated in the inverse dq-transformation $\boldsymbol{T}_\mathrm{P}^{-1}(\varphi(k))$. Compensation is required since the rotor position of the machine changes between the discrete sample instance and the application of the voltage by the VSI using SVM. The angle φ changes linearly between two concurrent sampling instances, assuming that the machine speed only changes marginally compared to the sampling time [19]. Therefore, the delay of φ is compensated by adding

$$\Delta\varphi(k) = \omega_\mathrm{el}(k)\tau_\sigma \tag{22}$$

to the inverse dq-transformation $\boldsymbol{T}_\mathrm{P}^{-1}\big(\varphi(k)+\Delta\varphi(k)\big)$ [19]. Fig. 5 shows the measurement results of A2 with and without proper compensation of the sample delay. The SSE, especially in the d-axis, is significantly reduced when using (22). Note that this effect is distinct from the intentionally induced misalignment of Sec. 4.1. The compensation of the sample delay (22) is used in all subsequent experiments.

5.2 Steady-state Performance

The steady-state performance of all controllers for machine M1 and M2 are shown in Fig. 6 and Fig. 7, respectively, in nominal and adverse operating conditions. Agents A1, A3, and A5 using the observation (14) with integrator compensate the SSE. Agents A1 and A5 match the performance of FOC in all nominal and adverse operating conditions. Agent A3 shows increased SSE from $n = 600\,\mathrm{min}^{-1}$ to $900\,\mathrm{min}^{-1}$. The adverse operating conditions do not significantly degrade the steady-state performance of DDPG-CC with augmented integration state in the measurements.

The agent A2 does not compensate the SSE in nominal operation (Fig. 6-(b)). Additionally, misalignment of the d-axis further deteriorates the steady-state performance. The increase in stator resistance results in additional SSE at slow rotational speeds but reduces the SSE slightly above $n = 2000\,\mathrm{min}^{-1}$. A possible explanation for the improvement is that agent A2 does show significant SSE, and a mismatch of the parameters compared to the training environment potentially moves the

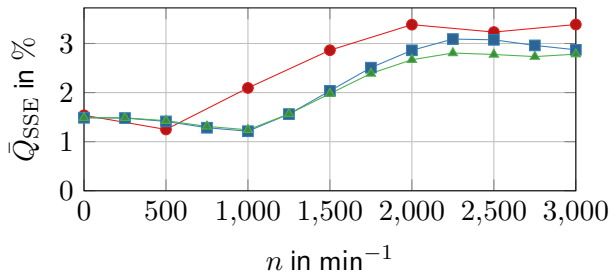

Fig. 8: Measured SSE of A2 (red) compared to to simulation based on (7) (including iron losses) (blue) and without iron losses modelled based on (6) (green).

control action in the right direction due to effects not accounted for by the model. This behavior is consistent for agent A4. However, contrary to A2, the SSE of A4 improves when misalignment of the dq-frame is present.

Fig. 8 compares the measured SSE of agent A2 to simulation results obtained using the linear machine model that includes iron losses (7). As shown, the SSE of agent A2 is generally increased in the measurement compared to simulation. Significant SSE is observed even when the agent controls the machine in ideal conditions, i.e., in simulation where the evaluation plant matches the training plant. The mismatch between simulation and measurement is attributed to neglected effects in the simulation, e.g., the inverter is assumed to be ideal in simulation, and parameter deviations.

5.3 Transient Performance

Fig. 9 and Fig. 10 depict the transient performance of M1 and M2, respectively, in nominal and adverse operating conditions. FOC (M1 & M2) as well as A1 is relatively robust regarding the adverse operating conditions. The transient response of FOC is decreased when the additional resistors are added to M1 and M2. The decrease is more severe for A4 and A5, especially at low rotational speed. On the contrary, the transient performance of agent A1 improves slightly when additional resistors are added. [6] indicates that agent A1 is prone to overshoot, which might be reduced in this operating condition. All control algorithms with integrator show no significant impact when misalignment of the dq-frame is induced.

Similar to the behavior in steady-state, the transient response of agent A2 deteriorates noticeably at higher speeds when misalignment is present. Unexpectedly, $\bar{Q}_{\mathrm{IAE}}$ of agent A4 is not affected by misalignment in the conducted experiments.

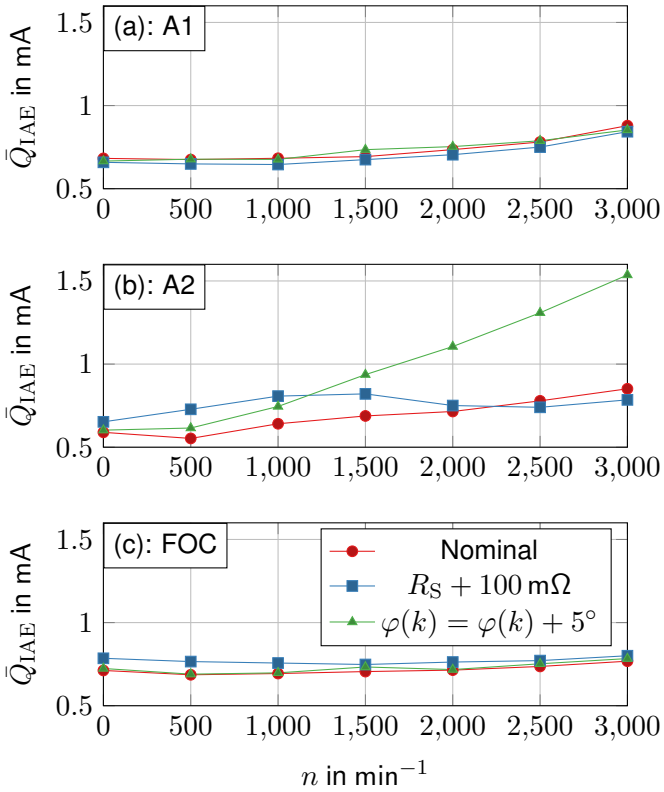

Fig. 9: Measured transient performance of controllers for M1 in different operating conditions.

The behavior is consistent for the machines M1 and M2 except A4. In general, the transient performance of the investigated agents is comparable to FOC at higher rotational speeds. However, Agent A5 exhibits subpar transient performance compared to the other controllers at low rotational speed. Furthermore, the additional resistors degrade the transient performance of A4 and A5 stronger compared to the other algorithms.

5.4 Current Harmonics

Fig. 11 depicts the phase currents of M2 for the first reference change in the evaluation profile. Noticeable distortion of the phase currents is visible due to the spatial harmonics of the machine. The harmonic components of the phase currents, based on FFT over 20 periods, is displayed in Fig. 12 at one operating point. The agent A5, trained on the model accounting for spatial harmonics (12), can compensate for the fourth and seventh-order harmonics. This results in a decrease of the total harmonic distortion (THD) to $11.6\,\%$ using A5 compared to $16.9\,\%$ using FOC.

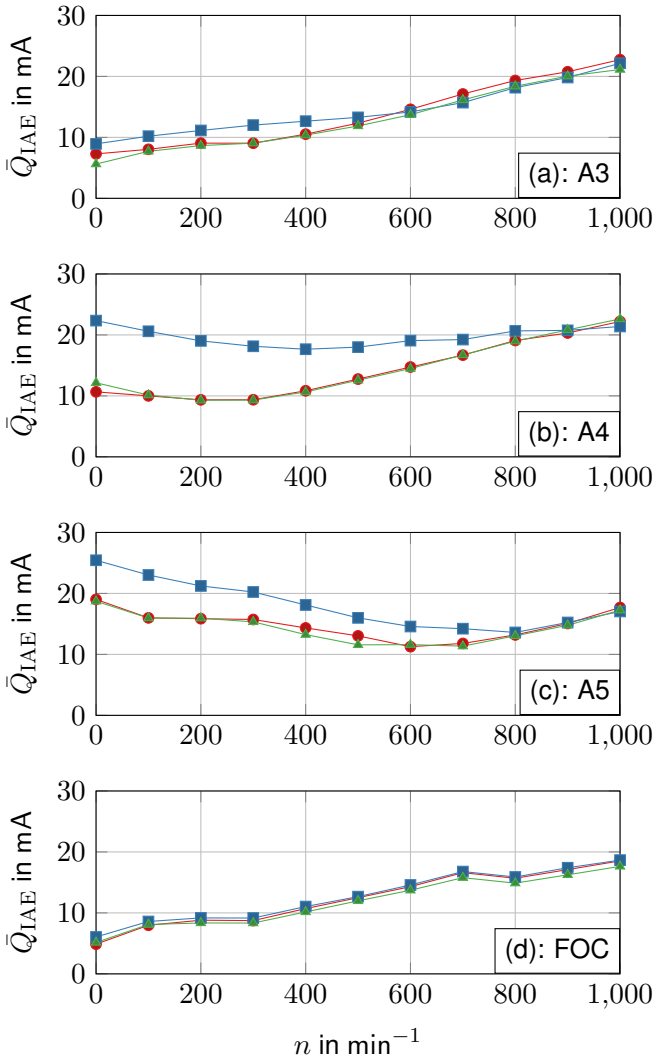

Fig. 10: Measured transient performance of controllers for M2 in nominal operation (red), additional resistors (blue), and misalignment (red).

6 Conclusion

This paper investigated the ability of of DDPG-CC to address the SSE that is present when controlling PMSMs with linear and nonlinear magnetic circuit. Nominal and adverse operating conditions were considered. A high-fidelity model was used in the training to control a PMSM that exhibits saturation, cross-coupling, and spatial harmonics. As demonstrated by the presented experimental results, when an integrator is not included in the observation vector, DDPG-CC is sensitive to parameter mismatches, and it becomes less effective when operating under adverse conditions. On the other hand, adding an integrating element to the observation vector significantly improves the steady-state performance of DDPG-CC as it removes any SSE under all operating conditions. Unfortunately, this results in a reduced dynamic performance. The latter could be improved by increasing the size of the neural network of the actor but at the expense of a higher computational burden. Finally, as shown, accounting for spatial harmonics allows DDPG-CC to produce currents of higher quality compared to FOC, albeit at the cost of slower dynamics.

Acknowledgment

The authors thank Philipp Dölger for his work on the flux-map approximation used in the FOC.

References

[1] T. Gemaßmer, “Effiziente und dynamische Drehmomenteinprägung in hoch ausgenutzten Synchronmaschinen mit eingebetteten Magneten,” German, Ph.D. dissertation, 2015. DOI: 10.5445/KSP/1000046666.

[2] S.-W. Su, C. M. Hackl, and R. Kennel, “Analytical prototype functions for flux linkage approximation in synchronous machines,” *IEEE Open Journal of the Industrial Electronics Society*, vol. 3, pp. 265–282, 2022. DOI: 10.1109/OJIES.2022.3162336.

[3] J. Richter, *Modellbildung, Parameteridentifikation und Regelung hoch ausgenutzter Synchronmaschinen*. Karlsruhe: KIT Scientific Publishing, Oct. 2016. DOI: 10.5445/KSP/1000057097.

[4] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, *et al.*, *Continuous control with deep reinforcement learning*, 2015.

[5] S. Bhattacharjee, S. Halder, Y. Yan, A. Balamurali, L. V. Iyer, and N. C. Kar, “Real-time SIL validation of a novel PMSM control based on deep deterministic policy gradient scheme for electrified vehicles,” *IEEE Transactions on Power Electronics*, vol. 37, no. 8, pp. 9000–9011, 2022. DOI: 10.1109/TPEL.2022.3153845.

[6] 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. DOI: 10.1109/IEMDC55163.2023.10238988.

[7] G. Book, A. Traue, P. Balakrishna, A. Brosch, M. Schenke, *et al.*, “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. DOI: 10.1109/OJPEL.2021.3065877.

[8] J. Degrave, F. Felici, J. Buchli, M. Neunert, B. Tracey, *et al.*, *Magnetic control of tokamak plasmas through deep reinforcement learning*, 2022. DOI: 10.1038/s41586-021-04301-9.

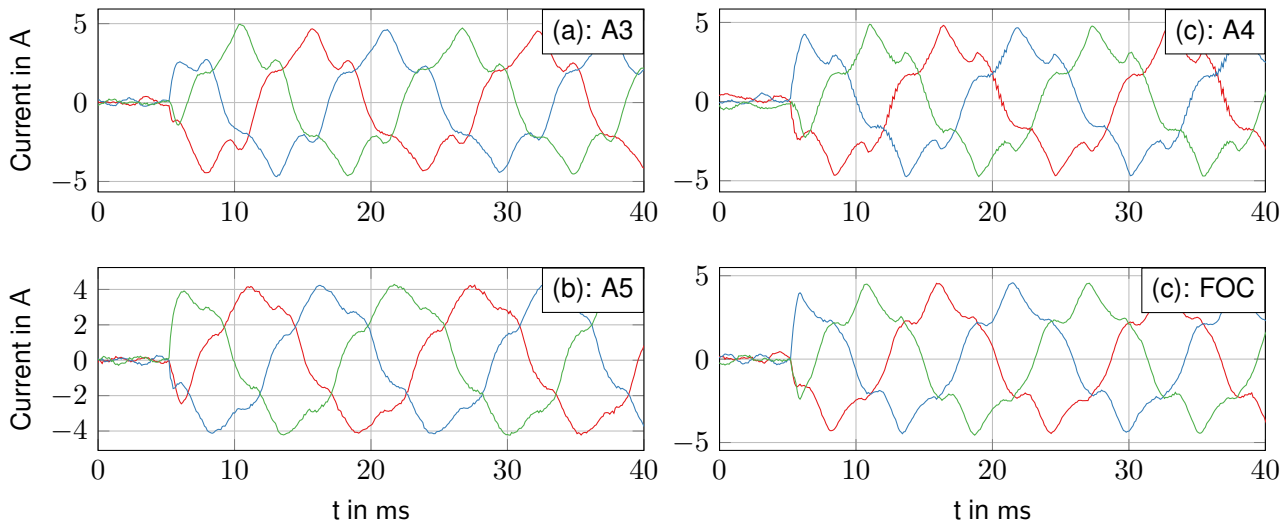

Fig. 11: Measured phase currents i_s^a (red), i_s^b (blue), and i_s^c (green) of M2 controlled by different algorithms at $n = 900\,\mathrm{min}^{-1}$.

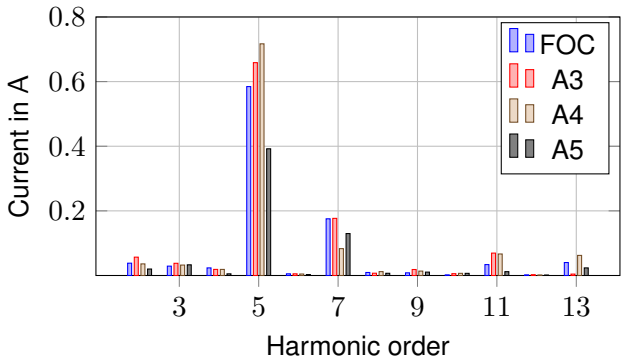

Fig. 12: Experimental results of the harmonic components of the phase currents of M2 calculated by FFT at $n = 900\,\mathrm{min}^{-1}$, $i_\mathrm{s}^q = 3.57\,\mathrm{A}$, and $i_\mathrm{s}^q = -1.0\,\mathrm{A}$.

[9] D. Weber, M. Schenke, and O. Wallscheid, "Steady-state error compensation for reinforcement learning-based control of power electronic systems," *IEEE Access*, vol. 11, pp. 76 524–76 536, 2023. DOI: 10.1109/ACCESS.2023.3297274.

[10] S. Wendel, "Model predictive control of small mechatronic drive systems with pm synchronous machines," Ph.D. dissertation, Technische Universität München, 2022.

[11] J. Richter, A. Dollinger, and M. Doppelbauer, "Iron loss and parameter measurement of permanent magnet synchronous machines," in *2014 International Conference on Electrical Machines (ICEM)*, 2014, pp. 1635–1641. DOI: 10.1109/ICELMACH.2014.6960401.

[12] X. Chen, J. Wang, B. Sen, P. Lazari, and T. Sun, "A high-fidelity and computationally efficient model for interior permanent-magnet machines considering the magnetic saturation, spatial harmonics, and iron loss effect," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 7, pp. 4044–4055, 2015. DOI: 10.1109/TIE.2014.2388200.

[13] D. Schroeder, *Elektrische Antriebe - Regelung von Antriebssystemen*. Springer, 2015.

[14] S. Kellner, "Parameteridentifikation bei Permanenterregten Synchronmaschinen.," Ph.D. dissertation, 2012.

[15] S. Li, D. Han, and B. Sarlioglu, "Modeling of interior permanent magnet machine considering saturation, cross coupling, spatial harmonics, and temperature effects," *IEEE Transactions on Transportation Electrification*, vol. 3, no. 3, pp. 682–693, 2017. DOI: 10.1109/TTE.2017.2679212.

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

[17] S. Wendel, A. Geiger, E. Liegmann, D. Arancibia, E. Durán, *et al.*, "UltraZohm— a powerful real-time computation platform for MPC and multi-level inverters," Quanzhou, China, 2019, pp. 1–6.

[18] *Matlab documentation: Deep deterministic policy gradient (DDPG) agents*, https://de.mathworks.com/help/reinforcement-learning/ug/ddpg-agents.html, Accessed: 2023-04-12.

[19] N. U and R. A, *Benötigt ein pulsweitenmoduliert betriebener Drehstromantrieb einen Stromzustandsregler?* Technischer Bericht - IAF-Report, Nr. 8, 2012.