



HAL
open science

Fault-Tolerant Predictive Control With Deep-Reinforcement-Learning-Based Torque Distribution for Four In-Wheel Motor Drive Electric Vehicles

Huifan Deng, Youqun Zhao, Tran Anh-Tu Nguyen, Chao Huang

► **To cite this version:**

Huifan Deng, Youqun Zhao, Tran Anh-Tu Nguyen, Chao Huang. Fault-Tolerant Predictive Control With Deep-Reinforcement-Learning-Based Torque Distribution for Four In-Wheel Motor Drive Electric Vehicles. IEEE/ASME Transactions on Mechatronics, 2023, 28 (2), pp.668-680. 10.1109/TMECH.2022.3233705 . hal-04278821

HAL Id: hal-04278821

<https://uphf.hal.science/hal-04278821>

Submitted on 25 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/367220637>

Fault-Tolerant Predictive Control With Deep-Reinforcement-Learning-Based Torque Distribution for Four In-Wheel Motor Drive Electric Vehicles

Article in *IEEE/ASME Transactions on Mechatronics* · January 2023

DOI: 10.1109/TMECH.2022.3233705

CITATIONS

12

READS

275

4 authors:



Huifan Deng

Nanjing University of Aeronautics & Astronautics

21 PUBLICATIONS 134 CITATIONS

SEE PROFILE



You Qun Zhao

Nanjing University of Aeronautics & Astronautics

113 PUBLICATIONS 1,047 CITATIONS

SEE PROFILE



Anh-Tu Nguyen

Université Polytechnique Hauts-de-France

140 PUBLICATIONS 2,084 CITATIONS

SEE PROFILE



Chao Huang

The Hong Kong Polytechnic University

126 PUBLICATIONS 1,813 CITATIONS

SEE PROFILE

Fault-Tolerant Predictive Control with Deep Reinforcement Learning-Based Torque Distribution for Four In-Wheel Motor Drive Electric Vehicles

Huifan Deng, Youqun Zhao, Anh-Tu Nguyen*, *Senior Member, IEEE*, Chao Huang

Abstract—This paper proposes a fault-tolerant control (FTC) method for four in-wheel motor drive electric vehicles considering both vehicle stability and motor power consumption. First, a seven degrees-of-freedom vehicle nonlinear model integrating motor faults is built to design a hierarchical FTC control scheme. The control structure is composed of two levels: an upper-level nonlinear model predictive controller and a lower-level fault-tolerant coordinated controller. The upper-level controller provides an appropriate reference in terms of additional yaw moment and vehicle longitudinal force, required for vehicle stability control, to the lower-level controller. This latter aims at distributing the four-wheel torques taking into account both vehicle stability and power consumption. Specifically, the weighting factor involved in the optimization-based design of the lower-level controller is determined online by the randomized ensembled double Q -learning reinforcement learning algorithm to achieve an optimal control strategy for the whole vehicle operating range. Moreover, the tradeoff between vehicle stability and power consumption is analyzed, and the necessity of using reinforcement learning is discussed. Numerical experiments are performed under various driving scenarios with a high-fidelity CarSim vehicle model to demonstrate the effectiveness of the proposed control method. Via a comparative study, we highlight the advantages of the new FTC control method over many related existing control results in terms of improving the vehicle stability and driver comfort, as well as reducing the power consumption.

Index Terms—Electric ground vehicles, vehicle motion dynamics, fault-tolerant control, reinforcement learning, torque vectoring, in-wheel/hub motor.

I. INTRODUCTION

With the rapid development of automotive industry, energy issues have become more prominent. In addition, automobile exhaust pollution is another major concern of automotive industry. Carbon monoxide, nitrogen oxides and particulate matter are the main causes of environment pollution and global warming. Electric vehicles (EVs) have been considered as an effective solution for these issues [1], [2]. Four in-wheel motor drive (4IWMD) electric vehicles are the ultimate structure of future EVs, which have become a hot research application [3]. The major advantage of 4IWMD electric vehicles is that each in-wheel motor can be

independently controlled. Moreover, the motor torques and speeds can be measured in real-time, which makes the whole vehicle dynamics control more convenient and flexible to improve the vehicle stability [4]. Since the actuators of EVs are redundant, *i.e.*, over-actuated systems, how to distribute the motor torques to achieve the best vehicle control performance is a crucial research problem. In particular, due to the presence of several actuators and the common use of wire-controlled technology, actuator faults may frequently occur, which can seriously affect the driving safety [5]. Hence, fault-tolerant control (FTC) is essential for 4IWMD EVs.

Numerous fault-tolerant control results have been reported in the literature for electric vehicles, including sliding mode control [6]–[8], linear parameter-varying control [9], Takagi-Sugeno fuzzy model-based control [10], [11], observer-based non-fragile control [12]. An adaptive sliding mode FTC considering mismatched nonlinear disturbances has been proposed in [13] with hardware-in-the-loop validations. A nonlinear disturbance observer has been used to estimate non-matching disturbances and to attenuate the parameter perturbation. The authors in [6] have developed an adaptive sliding-mode fault-tolerant coordinated control method, which allows to reduce the chattering by an adaptive variable reaching law. To increase the applicability range, the FTC control method in [14] considers the loss-of-effectiveness fault and the bias fault in the vehicle modeling. A three-step FTC scheme taking into account uncertain parameters, external disturbances and actuator faults has been proposed in [15], where the issues on parametric uncertainty and faults are mitigated using Lyapunov techniques. Despite these great advances in EVs control, the power consumption of EVs has not been well addressed for the control design. Indeed, these control results have mainly focused on the vehicle stability in case of faulty motors. However, when the motor faults are not serious, FTC design with power consumption saving should be considered. The authors in [16]–[18] have developed torque distribution algorithms for 4IWMD EVs with healthy motor actuators by taking into account the motor power in the objective function of the quadratic optimization-based design problem to reduce the vehicle power consumption. In particular, fuzzy logic has been used in [18] to adjust the tradeoff between the power consumption and the vehicle stability. However, due to the strong coupling between stability and power consumption, the existing torque distribution methods can only provide the optimal solution for the current operating moment, which can be insignificant for the overall operating range of the vehicle. Model predictive control (MPC) technique can be considered as an alternative solution to deal with this tradeoff. Unfortunately, MPC prediction is usually available over a time window, not for the whole vehicle travel. In particular, if the MPC prediction horizon increases, then the computational burden of the involved online optimization problem becomes too heavy for real-time vehicle control purposes.

Reinforcement learning (RL) brings some new inspiration to control methods since it can interact with the environment and keep

Huifan Deng and Youqun Zhao are with the Department of Vehicle Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China. Huifan Deng is also a visiting PhD student at LAMIH laboratory UMR CNRS 8201, Université Polytechnique Hauts-de-France, F-59313 Valenciennes, France. (e-mail: denghuifan163@163.com, yqzhao@nuaa.edu.cn).

Anh-Tu Nguyen is with LAMIH laboratory UMR CNRS 8201, Université Polytechnique Hauts-de-France, F-59313 Valenciennes, France (e-mail: tnguyen@uphf.fr).

Chao Huang is with the Department of Industrial and System Engineering, The Hong Kong Polytechnic University, Hong Kong (e-mail: hchao.huang@polyu.edu.hk).

*Corresponding author: Anh-Tu Nguyen.

adapting to achieve the best control performance. RL has been applied to several fields, *e.g.*, battery energy management [19], robotic control [20], motion planning of autonomous vehicles [21]. However, due to the strict requirement on vehicle stability, very few studies have been reported on using RL for torque distribution of EVs. A direct torque distribution algorithm based on a deep RL algorithm has been proposed in [20] to improve the EVs stability while reducing the motor power consumption. However, the motor torques of EVs are directly output by the proposed RL algorithm, which may yield an unstable closed-loop vehicle system. Then, the RL control strategy in [20] may lead to some risky driving situations. RL algorithms have been also exploited for fault-tolerant control design for complex systems. Zhang et al. [22] have used RL to obtain implicit models as references for solving the FTC problem where explicit reference models are not available. An RL-based FTC method has been proposed in [23] to mitigate the impacts of modeling uncertainties and unexpected faults.

Motivated by the above technical issues, this paper investigates the FTC design for 4IWMD electric vehicles considering both power consumption and vehicle stability, which has not been addressed in the literature. The proposed fault-tolerant control structure is composed of two hierarchical levels: upper control level and lower control level. Nonlinear model predictive control (NMPC) technique is used to design the upper-level controller, which aims at providing appropriate references on additional yaw moment and longitudinal force to guarantee the vehicle stability. Using these vehicle input references, the lower-level controller distributes the four-wheel torques taking into account the trade-off between vehicle stability and power consumption. The main contributions of this paper are summarized as follows.

- We propose an RL-based FTC control method for 4IWMD EVs, which allows to improve the vehicle stability while globally reducing the power consumption for the whole operating range.
- The deep RL-based torque distribution algorithm can be adapted in function of vehicle operating conditions to effectively manage the stability-consumption tradeoff.
- Exploiting the randomized ensembled double Q -learning (REDQ) algorithm to EVs control, the proposed strategy can improve the safety issue compared to existing methods using RL algorithms to directly control the motor torques.

Several numerical experiments and comparative studies are performed with a high-fidelity CarSim vehicle model to demonstrate the effectiveness of the proposed control method. The necessity for a learning-based torque distribution to globally improve the vehicle stability and driver workload while reducing the power consumption is also highlighted.

II. PROBLEM FORMULATION

This section first introduces the overall hierarchical control structure for 4IWMD EVs. Then, after the vehicle modeling, the control problem is formulated. The vehicle nomenclature is given in Table I, where the indices $ij \in \{fl, fr, rl, rr\}$ respectively represent the left front wheel, the right front wheel, the left rear wheel, and the right rear wheel.

A. Overall Hierarchical Control Architecture

Fig. 1 depicts the overall architecture of the proposed hierarchical control scheme for 4IWMD EVs. The main components include a vehicle reference model, an upper-level controller, a lower-level controller, and a deep RL algorithm. The vehicle reference model

provides the desired yaw rate and sideslip angle to the upper-level controller. Based on NMPC technique and a seven degrees-of-freedom vehicle model, this latter is designed to generate appropriate references on yaw moment and longitudinal force to guarantee the vehicle stability. Using these information, the lower-level controller aims at distributing the four-wheel torques for fault-tolerant control purposes. A fault diagnosis (FD) scheme is used to provide the estimates of the fault factors. To manage the vehicle actuator faults, these estimates are input to the REDQ algorithm as the deep RL states. The fault factor estimates are also used to design the lower-level fault-tolerant controller. Moreover, the REDQ algorithm based on the risk assessment dynamically adjusts the tradeoff between the electric power consumption and the vehicle stability. Then, fault-tolerant coordination controller provides torque inputs to EVs while achieving the best consumption-stability tradeoff.

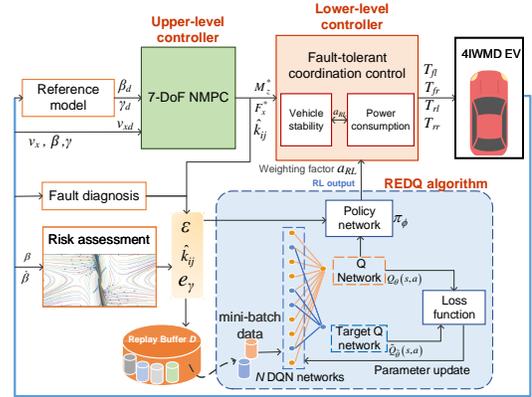


Fig. 1. Overall hierarchical control architecture for 4IWMD electric vehicles.

TABLE I
VEHICLE NOMENCLATURE.

Symbol	Description
m	Total mass of the vehicle
v_x	Longitudinal velocity
v_y	Lateral velocity
β	Sideslip angle of the vehicle center of gravity (CG)
γ	Vehicle yaw rate
δ_f	Steering angle of front wheels
a	Distance from the front wheel axles to the CG
b	Distance from the rear wheel axles to the CG
d_w	Track width
I_z	Yaw mass moment of inertia
R_w	Wheel rolling radius
C_f	Front tires cornering stiffness
C_r	Rear tires cornering stiffness
F_{xij}	Longitudinal force of the wheel
F_{yij}	Lateral force of the wheel
ω_{ij}	Wheel rotation rate
$I_{\omega_{ij}}$	Wheel moment of inertia
T_{ij}	Output torque of the in-wheel motor
g	Acceleration of gravity
a_y	Lateral acceleration

B. Vehicle Reference Model

To guarantee the vehicle stability, the tracking control of the vehicle yaw rate and the sideslip angle must be performed [24]. To this end, the desired value of the sideslip angle can be conservatively set as $\beta_d = 0$ as discussed in [25], while that of

the yaw rate $\dot{\gamma}_d$ is determined from the following two degrees-of-freedom (DoF) vehicle model [26]:

$$\begin{aligned} m\dot{v}_x(\dot{\beta} + \gamma) &= 2C_f \left(\delta_f - \beta - \frac{a\gamma}{v_x} \right) + 2C_r \left(\frac{b\gamma}{v_x} - \beta \right) \\ I_z \dot{\gamma} &= 2aC_f \left(\delta_f - \beta - \frac{a}{v_x} \right) - 2bC_r \left(\frac{b}{v_x} - \beta \right). \end{aligned} \quad (1)$$

Taking into account the road adhesion coefficient μ , the lateral acceleration $a_y = \gamma v_x$ is limited by $\gamma v_x \leq \mu g$,

where g is the acceleration of gravity. Then, the reference yaw rate can be deduced from (1) as

$$\gamma_d = \min \left\{ \left| \frac{v_x}{(l_f + l_r)(1 + Kv_x^2)} \delta \right|, \left| \frac{\mu g}{v_x} \right| \right\} \text{sign}(\delta),$$

with $K = \frac{m(aC_f - bC_r)}{2C_f C_r (a+b)^2}$.

C. Vehicle Dynamics with Motor Actuator Faults

The vehicle model (1) only represents the basic lateral dynamics, without any information on the nonlinear tires characteristics. Moreover, the left wheels and the right wheels are assumed to be the same. Hence, motor actuator faults cannot be considered with model (1). To avoid these drawbacks, a seven DoF vehicle model, as depicted in Fig. 2, can be used. The corresponding vehicle dynamics can be described by [27]

$$\begin{aligned} m\dot{v}_x &= (F_{xfl} + F_{xfr}) \cos \delta_f - (F_{yfl} + F_{yfr}) \sin \delta_f \\ &\quad + F_{xrl} + F_{xrr} + mv_x \gamma \\ m\dot{v}_y &= (F_{xfl} + F_{xfr}) \sin \delta_f + (F_{yfl} + F_{yfr}) \cos \delta_f \\ &\quad + F_{yrl} + F_{yrr} - mv_x \gamma \\ I_z \dot{\gamma} &= a(F_{yfl} + F_{yfr}) \cos \delta_f - b(F_{yrl} - F_{yrr}) \\ &\quad + \frac{d_w}{2}(F_{yfl} - F_{yfr}) \sin \delta_f + M_z \\ I_{\omega_{ij}} \dot{\omega}_{ij} &= -F_{xij} R_{\omega} + T_{ij}. \end{aligned} \quad (2)$$

The vehicle control inputs, *i.e.*, longitudinal force F_x and yaw moment M_z , are expressed as follows [17]:

$$\begin{aligned} F_x &= F_{xfl} + F_{xfr} + F_{xrl} + F_{xrr}. \\ M_z &= F_{xfl} \left(a \sin \delta_f - \frac{d_w}{2} \cos \delta_f \right) + \frac{d_w}{2} (F_{xrr} - F_{xrl}) \\ &\quad + F_{xfr} \left(a \sin \delta_f + \frac{d_w}{2} \cos \delta_f \right). \end{aligned} \quad (3)$$

The nonlinear characteristics of the tires are represented by the well-known Magic formula as [28]

$$\begin{aligned} F_{xij}^0(\lambda_{ij}) &= D_x \sin \{ C_x \text{atan} [B_x (\lambda_{ij} + S_{hx}) (1 - E_x) \\ &\quad + E_x \text{atan} (B_x (\lambda_{ij} + S_{hx}))] \} + S_{vx}, \\ F_{yij}^0(\alpha_{ij}) &= D_y \sin \{ C_y \text{arctan} [B_y (\alpha_{ij} + S_{hy}) (1 - E_y) \\ &\quad + E_y \text{atan} (B_y (\alpha_{ij} + S_{hy}))] \} + S_{vy}, \end{aligned} \quad (4)$$

where F_{xij}^0 and F_{yij}^0 are respectively the longitudinal forces and lateral forces of the tires, for $ij \in \{fl, fr, rl, rr\}$, and $B_x, B_y, C_x, C_y, D_x, D_y, E_x, E_y, S_{hx}, S_{hy}, S_{vx}, S_{vy}$ are the characteristic parameters of the Magic formula. The tire slip angles are expressed by

$$\begin{aligned} \alpha_{fl} &= \text{atan} \left(\frac{v_y + a\gamma}{v_x - \frac{1}{2}d_w\gamma} \right) - \delta_f, \quad \alpha_{rl} = \text{atan} \left(\frac{v_y - b\gamma}{v_x - \frac{1}{2}d_w\gamma} \right) \\ \alpha_{fr} &= \text{atan} \left(\frac{v_y + a\gamma}{v_x + \frac{1}{2}d_w\gamma} \right) - \delta_f, \quad \alpha_{rr} = \text{atan} \left(\frac{v_y - b\gamma}{v_x + \frac{1}{2}d_w\gamma} \right) \end{aligned} \quad (5)$$

The tire longitudinal slip ratios are defined as

$$\lambda_{ij} = \frac{v_{ij} - \omega_{ij} R_{\omega}}{v_{ij}}, \quad ij \in \{fl, fr, rl, rr\}, \quad (6)$$

where the vehicle speeds at the four wheels are expressed by

$$\begin{aligned} v_{fl} &= (v_y + a\gamma) \sin \delta_f + \left(v_x - \frac{\gamma d_w}{2} \right) \cos \delta_f, \\ v_{fr} &= (v_y + a\gamma) \sin \delta_f + \left(v_x + \frac{\gamma d_w}{2} \right) \cos \delta_f, \\ v_{rl} &= v_x - \frac{\gamma d_w}{2}, \quad v_{rr} = v_x + \frac{\gamma d_w}{2}. \end{aligned}$$

Since the tire forces must satisfy the adhesion ellipse, the longitudinal forces and the lateral forces of the tires need to be corrected as follows [17]:

$$F_{xij} = \frac{F_{xij}^0 \psi_{xij}}{\psi_{ij}}, \quad F_{yij} = \frac{F_{yij}^0 \psi_{xij}}{\psi_{ij}}, \quad (7)$$

with $\psi_{ij} = \sqrt{\psi_{xij}^2 + \psi_{yij}^2}$, $\psi_{xij} = -\frac{\lambda_{ij}}{1 + \lambda_{ij}}$ and $\psi_{yij} = -\frac{\tan(\alpha_{ij})}{1 + \lambda_{ij}}$, for $ij \in \{fl, fr, rl, rr\}$.

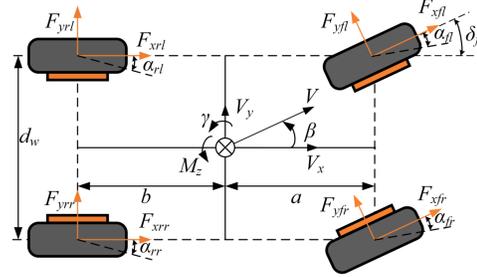


Fig. 2. Schematic of a seven degrees-of-freedom nonlinear vehicle model.

The fault diagnosis and fault factor estimation for the motors of 4IWM electric vehicles have been widely studied in the literature, see for instance [7], [29], [30] and related references. Here, we mainly focus on the torque distribution for FTC purposes. Then, it is assumed that the fault diagnosis scheme proposed in [7] is given to estimate the fault factors k_{ij} , for $ij \in \{fl, fr, rl, rr\}$, for FTC design. Moreover, to better represent the fault factor estimation, the effects of time delay and disturbances are also taken into account as

$$k_{ij} = \frac{T_{aij}}{T_{ij}}, \quad \hat{k}_{ij} = \frac{1}{t_p s + 1} k_{ij} + \Delta d, \quad (8)$$

where T_{aij} is the actual vehicle torque input, k_{ij} is the fault factor, \hat{k}_{ij} is its estimate, t_p is the delay time, Δd is the disturbance, and s is the Laplace variable. Considering the fault factors in (8), and the relationship between the motor torques and the longitudinal forces $F_{xij} = T_{ij}$, for $ij \in \{fl, fr, rl, rr\}$, the longitudinal force F_x and the yaw moment M_z in (3) can be represented by

$$F_x = A_{F_x} u_c, \quad M_z = A_{M_z} u_c, \quad (9)$$

where the vector of the four-wheel motor torques $u_c = [T_{fl} \ T_{fr} \ T_{rl} \ T_{rr}]^T$ is the vehicle control input, and

$$\begin{aligned} A_{F_x} &= \frac{1}{R_{\omega}} [\hat{k}_{fl} \cos \delta_f \quad \hat{k}_{fr} \cos \delta_f \quad \hat{k}_{rl} \quad \hat{k}_{rr}], \\ A_{M_z} &= \frac{1}{R_{\omega}} [\zeta_1 \hat{k}_{fl} \quad \zeta_2 \hat{k}_{fr} \quad -\frac{d_w}{2} \hat{k}_{rl} \quad \frac{d_w}{2} \hat{k}_{rr}], \end{aligned}$$

with $\zeta_1 = l_f \sin \delta_f - \frac{d_w}{2} \cos \delta_f$ and $\zeta_2 = \frac{d_w}{2} \cos \delta_f + l_f \sin \delta_f$.

D. Problem Formulation

The nonlinear seven DoF vehicle dynamics can be rewritten from (2), (4), (5), (6) and (7), in form

$$\dot{x} = f(x, u), \quad (10)$$

where $x = [v_x \ \beta \ \gamma]^\top$ is the vehicle state, and $u = [F_x \ M_z]^\top$ is the vehicle control input. The explicit form of the nonlinear function $f(\cdot)$ in (10) can be easily obtained, which is omitted for brevity. This paper develops a multiobjective control strategy for the 4IWMD electric vehicle (10), which is typically an over-actuated system as shown in (9). The motor torques are distributed to guarantee both safety and comfort issues while globally reducing the power consumption of the motors over the whole vehicle operating range. In particular, the electric vehicle can also safely operate despite the occurrence of motor faults. To this end, we propose a fault-tolerant control scheme with two hierarchical levels: upper-level NMPC control and lower-level fault-tolerant torque distribution, as detailed in Section III.

III. HIERARCHICAL FAULT-TOLERANT CONTROL DESIGN

This section provides the technical details on both upper and lower control levels of the proposed hierarchical FTC scheme.

A. Upper-Level NMPC Control

Based on the seven DoF vehicle model (10) and NMPC technique, the upper control level aims at generating the ideal total longitudinal force F_x^* and the additional yaw moment M_z^* , i.e., $u^* = [F_x^* \ M_z^*]^\top$, to track the vehicle state reference for stability guarantee while taking into account the comfort and the energy saving issues.

Considering the nonlinear tire characteristics, when the tires operate within an unstable region, the actual lateral tire force cannot be well represented by a linear tire force model [26]. Then, a controller, whose design is based on the linear tire force model, cannot meet the performance requirements under such limited operating conditions. To improve the control performance under limited driving scenarios, the NMPC technique is used together with a nonlinear vehicle model taking into account the nonlinear tire characteristics. As a result, the NMPC controller design for the vehicle model (10) can be formulated as the following nonlinear optimization problem:

$$\begin{aligned} \min_{\mathcal{X}, \mathcal{U}} \quad & \sum_{k=0}^{N_p-1} \left(e_k^\top W_Q e_k + u_k^\top W_R u_k + \Delta u_k^\top W_S \Delta u_k \right) \\ \text{s.t.} \quad & x_{k+1} = f_d(x_k, u_k), \quad k = 0, \dots, N_p - 2, \\ & x_{\min} \leq x_k \leq x_{\max}, \quad k = 0, \dots, N_p - 1, \\ & u_{\min} \leq u_k \leq u_{\max}, \quad k = 0, \dots, N_p - 1, \end{aligned} \quad (11)$$

where $\Delta u = u_k - u_{k-1}$, $e_k = x_k - x_k^d$, $x_d = [v_{xd} \ \beta_d \ \gamma_d]^\top$ is the vehicle state reference, $f_d(x_k, u_k)$ is the discrete-time counterpart of the nonlinear function $f(x, u)$ in (10), x_{\min} and x_{\max} are the upper and lower bounds of the vehicle states, u_{\min} and u_{\max} are the upper and lower bounds of the vehicle inputs, $\mathcal{X} = [x_0, \dots, x_{N_p-1}]$ and $\mathcal{U} = [u_0, \dots, u_{N_c-1}]$ are the state and the input sequences in the prediction horizon N_p and the control horizon N_c , respectively. The weighting matrices W_Q , W_R and W_S are predefined to track the ideal vehicle state quickly and smoothly, and to guarantee that the 4IWMD electric vehicle can achieve a good dynamic performance.

The nonlinear optimization problem (11) can be solved using the sequential quadratic programming algorithm [31]. After obtaining the optimal control solution sequence \mathcal{U} , only its first element is implemented into the lower control level and the predicted time horizon is moved backward by one step to continue solving the new nonlinear programming problem.

B. Lower-Level Fault-Tolerant Torque Distribution

The lower control level aims at distribute four-wheel torque considering motor fault, with torque control inputs u_c as (9). Such that the total longitudinal force F_x^* and the additional yaw moment M_z^* , given by the upper control level, can be achieved. For fault-tolerant torque distribution, based on a deep RL algorithm, the tradeoff between the vehicle stability and the power consumption can be dynamically adapted according to the current state of the vehicle. Moreover, to maintain the longitudinal velocity for lateral control purposes, the lower-level controller must satisfy the condition.

$$F_x^* = A_{F_x} u_c, \quad (12)$$

where A_{F_x} and u_c are defined in (9).

1) *Integrated Cost Function*: The fault-tolerant coordinated control design with energy saving consideration can be reformulated as a multiobjective optimization problem, whose three main objectives are described below.

a) *Guarantee the Vehicle Stability despite of*: Motor Faults Since the longitudinal velocity is controlled by (12), the four-wheel torques should satisfy the ideal additional yaw moment input M_z^* from the upper-level controller. This can be represented by the following cost function:

$$\mathcal{J}_1 = (A_{M_z} u_c - M_z^*)^\top W_{M_z} (A_{M_z} u_c - M_z^*), \quad (13)$$

where A_{M_z} and u_c are defined in (9), and W_{M_z} is a weighting matrix.

b) *Minimize the Attachment Utilization Ratio*: To describe the tire stability margin, the following adhesion utilization ratio, i.e., the ratio between the actual road adhesion and the maximum possible adhesion, is used:

$$\lambda_r = \sum_{i=1}^4 \frac{F_{xij}^2 + F_{yij}^2}{(\mu F_{zij})^2}, \quad ij \in \{fl, fr, rl, rr\},$$

where μ is the road adhesion coefficient. The expressions of the vertical loads of the four vehicle wheels F_{zij} , for $ij \in \{fl, fr, rl, rr\}$, are detailed in [17]. Note that the tire lateral forces are uncontrollable, then we only consider the tire longitudinal forces for the torque control allocation. Hence, the following cost function is used to guarantee the maximum tire stability margin:

$$\mathcal{J}_2 = u_c^\top W_a u_c, \quad (14)$$

with $W_a = \text{diag} \left(\frac{1}{(\mu R_\omega F_{zij})^2} \right)_{4 \times 4}$.

c) *Minimize the Motor Power Consumption*: The motor power consumption is computed as

$$\mathcal{P}_{ij} = \frac{1}{\eta(T_{ij}, \omega_{ij})} T_{ij} \omega_{ij}, \quad ij \in \{fl, fr, rl, rr\}, \quad (15)$$

where the motor efficiency $\eta(T_{ij}, \omega_{ij})$, depending on the motor torque and speed, is obtained from a look-up-table as depicted in Fig. 3. Then, considering that the values of the motor torques can

be both negative and positive, the following cost function is used to minimize the motor power consumption:

$$\mathcal{J}_3 = \sum_{ij \in \{fl, fr, rl, rr\}} \mathcal{P}_{ij}^2 = u_c^\top W_p u_c, \quad (16)$$

where the weighting matrix W_p is computed from (15) as $W_p = \text{diag} \left(\frac{\omega_{ij}^2}{\eta^2 (T_{ij}^*, \omega_{ij})} \right)_{4 \times 4}$, for $ij \in \{fl, fr, rl, rr\}$.

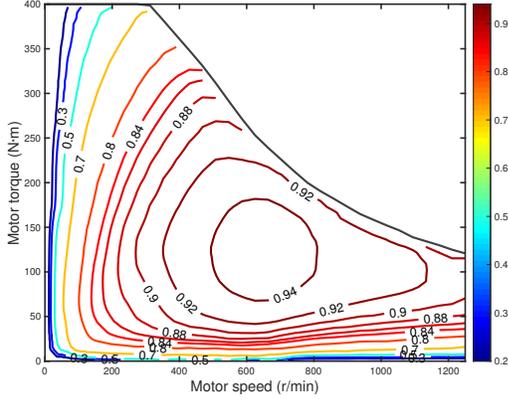


Fig. 3. Electric motor efficiency map.

The integrated cost function, taking into account the specifications on both vehicle stability and motor power consumption, can be reformulated from (13), (14) and (16) as

$$\mathcal{J} = \alpha_1 \mathcal{J}_1 + \alpha_2 \mathcal{J}_2 + \alpha_3 \mathcal{J}_3. \quad (17)$$

The weighting factors α_1 , α_2 and α_3 are defined as

$$\alpha_1 = \alpha_0 + (1 - \alpha_0)\alpha_{RL}, \quad \alpha_2 = \alpha_{RL}, \quad \alpha_3 = 1 - \alpha_{RL}, \quad (18)$$

where $\alpha_{RL} \in [0, 1]$ is the preference factor representing the tradeoff between power consumption and vehicle stability for the motor torque distribution. This factor α_{RL} is obtained from a deep RL algorithm, described in Section III-C.

Remark 1. Since safety is a primary requirement for vehicle driving in any situation, the parameter $\alpha_0 \in [0, 1]$ is introduced in (18) such that \mathcal{J}_1 is always involved in the integrated cost function \mathcal{J} defined in (17). This allows avoiding the vehicle instability issue that may occur with an unstable RL agent obtained from a training process. After several tests, we select $\alpha_0 = 0.2$ to design the RL-based torque distribution algorithm.

From (17) and (18), the integrated cost function can be rewritten as follows:

$$\mathcal{J} = u_c^\top H u_c + G^\top u_c + C, \quad (19)$$

with

$$\begin{aligned} H &= (\alpha_0 + (1 - \alpha_0)\alpha_{RL}) A_{M_z}^\top W_a A_{M_z} \\ &\quad + \alpha_{RL} W_a + (1 - \alpha_{RL}) W_p, \\ G^\top &= (\alpha_0 + (1 - \alpha_0)\alpha_{RL}) (-2M_z^\top W_{M_z} A_{M_z}), \\ C &= (\alpha_0 + (1 - \alpha_0)\alpha_{RL}) M_z^\top W_{M_z} M_z. \end{aligned}$$

Since u_c is not involved in C , this term has no impact on the control optimization problem. Hence, it follows from (19) that the final integrated cost function can be defined as

$$\mathcal{J} = u_c^\top H u_c + G^\top u_c. \quad (20)$$

2) *Optimization-Based Torque Distribution:* Note that the actual motor torque outputs must be constrained by the maximum adhesion force of the road surface and the external characteristics of the motors as

$$\begin{aligned} -\sqrt{(\mu F_{z_{ij}})^2 - F_{y_{ij}}^2} R\omega &\leq T_{ij} \leq \sqrt{(\mu F_{z_{ij}})^2 - F_{y_{ij}}^2} R\omega, \\ T_{ij} &\leq T_{ij}^*(\omega_{ij}), \end{aligned} \quad (21)$$

where the peak torques $T_{ij}^*(\omega_{ij})$, for $ij \in \{fl, fr, rl, rr\}$, of the electric motors are obtained from the motor external characteristics, e.g., the motor efficiency map.

From the definition of the integrated cost function \mathcal{J} in (20), the torque equality constraint (12) and the torque limitation constraint (21), the torque distribution problem can be reformulated as the following quadratic optimization problem:

$$\begin{aligned} \min_{u_c} \quad & \mathcal{J} = u_c^\top H u_c + G^\top u_c \\ \text{s.t.} \quad & A_{F_x} u_c = F_x^*, \\ & \underline{u}_c \leq u_c \leq \bar{u}_c, \end{aligned} \quad (22)$$

where the inequality operator in (22) is element-wise, and the vectors \underline{u}_c and \bar{u}_c are respectively defined as

$$\begin{aligned} \underline{u}_{ij} &= \max \left(-T_{ij}^*(\omega_{ij}), -\sqrt{(\mu F_{z_{ij}})^2 - F_{y_{ij}}^2} R\omega \right), \\ \bar{u}_{ij} &= \min \left(T_{ij}^*(\omega_{ij}), \sqrt{(\mu F_{z_{ij}})^2 - F_{y_{ij}}^2} R\omega \right), \end{aligned}$$

for $ij \in \{fl, fr, rl, rr\}$. Here, the quadratic optimization problem (22) is solved using the Gurobi solver [32] to obtain the four-wheel torques.

C. Deep Reinforcement Learning for Torque Distribution

The proposed torque distribution is based on the result of the optimization problem (22). Hence, the weighting factor α_{RL} involved in the integrated cost function \mathcal{J} plays a key role to achieve the best tradeoff between vehicle stability and motor power consumption. It is challenging to adapt this factor in function of each specific driving situation to manage such a stability-consumption tradeoff, especially in the presence of motor faults. In contrast to traditional control methods, RL can take into account the rewards for the whole operating range of EVs. Hence, RL can offer a better performance when the future system states should be considered to appropriately execute tasks. Therefore, to achieve the best performance possible from a global viewpoint with the whole vehicle operating range, we propose to compute the value of α_{RL} based on an RL algorithm. The proposed RL algorithm can constantly interact with the driving environment to deduce the most appropriate value of α_{RL} for the whole range by maximizing a predefined reward. RL algorithms can be classified into on-policy and off-policy algorithms. Off-policy learning allows the use of older samples during the training process, which can improve the sample efficiency. It has been shown that the soft actor-critic (SAC) algorithm is one of the best performing off-policy RL algorithms [33]. However, the REDQ algorithm can outperform the SAC algorithm, and can be even better than model-based RL algorithms for multiple tasks [34]. Therefore, here we select the integrated learning REDQ framework to develop our RL-based torque distribution algorithm.

The RL action space and state space are selected as follows:

$$\mathcal{A} = \{\alpha_{RL}\}, \quad \mathcal{S} = \{e_\gamma, \hat{k}_{ij}, \varepsilon\}, \quad (23)$$

where $e_\gamma = \gamma - \gamma_d$ is the yaw rate deviation. The stability index ε in the state space \mathcal{S} in (23) can be obtained from the vehicle phase portrait [35], which is defined as

$$\varepsilon = \left| \frac{1}{B_2} \beta + \frac{B_1}{B_2} \dot{\beta} \right|,$$

where the parameters B_1 and B_2 are related to the road surface adhesion coefficient, and their values can be found in [36].

Since we focus on ensuring the stability while improving the power consumption and the driving handling of 4IWMD electric vehicles, the reward function \mathcal{R} is defined as

$$\mathcal{R} = \mathcal{R}_{stability} + \mathcal{R}_{economy} + \mathcal{R}_{driver}, \quad (24)$$

with

$$\begin{aligned} \mathcal{R}_{stability} &= -W_s \int_0^t \left((\beta - \beta_d)^2 + (\gamma - \gamma_d)^2 \right) d\tau, \\ \mathcal{R}_{economy} &= -W_e \int_0^t \frac{T_{ij} \omega_{ij}}{\eta(T_{ij}, \omega_{ij})} d\tau, \\ \mathcal{R}_{driver} &= -W_d \int_0^t \left(\dot{\delta}_{sw}^2 + W_{a\delta} a_x^2 \right) d\tau, \end{aligned} \quad (25)$$

where W_s , W_e and W_d are respectively the penalty factors corresponding to the vehicle stability, the motor power consumption, and the driver workload. The weighting factor $W_{a\delta}$ represents the tradeoff between the rate of the driver's steering angle δ_{sw} and the vehicle longitudinal acceleration a_x .

The RL sample process is stated as follows. At step t , the state $s_t \in \mathcal{S}$ receives a reward $r_t = \mathcal{R}(s_t)$ as defined in (24). Then, the policy π generates an action $a_t \in \mathcal{A}$, which interacts with the environment to generate the next state s_{t+1} . As a measure of the randomness of a policy, the entropy is inclined in the objective of the REDQ algorithm to make the generated policy more robust and generalizable. Then, the optimal policy π^* is expressed as

$$\pi^* = \arg \max_{\pi} \mathbb{E}_{\tau \sim \pi} \left[\gamma_f (r_t + \xi \mathcal{H}(\pi(\cdot|s_{t+1}))) \right], \quad (26)$$

where $\mathbb{E}_{\tau \sim \pi} [\cdot]$ is the expectation corresponding to the trajectory distribution τ of (s_t, a_t) followed by the policy π , ξ is a hyperparameter to determine the weight of policy entropy in the optimization objective, $\gamma_f \in [0, 1]$ is a discount factor, and $\mathcal{H}(\cdot)$ is the entropy corresponding to the policy π .

The optimal state action-value function $Q^\pi(s_t, a_t)$ reflects the performance of the control system, *i.e.*, a higher value of $Q^\pi(s_t, a_t)$ represents a higher reward for the executing action a_t in the state s_t . According to the Bellman equation, the objective of the Q -value is expressed as

$$Q^\pi(s_t, a_t) = \mathbb{E}_{a_{t+1} \sim \pi} [\psi_t], \quad (27)$$

with $\psi_t = r_t + \gamma_f \max_{a_{t+1}} Q^\pi(s_{t+1}, a_{t+1}) + \xi \mathcal{H}(\pi(\cdot|s_{t+1}))$.

Due to the high dimensionality of the state space \mathcal{S} , we use a deep neural network to represent the Q -value function (27) and the policy π . Then, the evaluate Q -value function is represented as $Q_\theta(s_t, a_t)$ by the network with the hyperparameter θ , and the policy π is represented as π_ϕ by the network with the hyperparameter ϕ . To reduce the estimation error of the Q -function, we include a target Q -network $\tilde{Q}_{\tilde{\theta}}(s_t, a_t)$ with $\tilde{\theta}$ as the network hyperparameter. For the REDQ algorithm, N double DQN networks are built, and we compute the target value of Q -function by randomly selecting M networks $\tilde{Q}_{\tilde{\theta}}(s_t, a_t)$, with $M \leq N$, among these N DQN networks as

$$y^* = r_t + \gamma_f \left(\min_{m \in \mathcal{I}_M} \tilde{Q}_m(s_{t+1}, a_{t+1}, \tilde{\theta}_i) + \xi \mathcal{H}(\pi(\cdot|s_{t+1})) \right) \quad (28)$$

where the subset \mathcal{I}_M of M random elements is such that $\mathcal{I}_M \subseteq \{1, 2, \dots, N\}$. The hyperparameters of the evaluate Q -networks can be obtained by minimizing the loss function

$$\mathcal{L}_\theta(\theta_i) = \mathbb{E}_{\zeta_t} \left[(y^* - Q_{\theta_i}(s_t, a_t))^2 \right], \quad i = 1, 2, \dots, N, \quad (29)$$

with $\zeta_t = (s_t, a_t, r_t, s_{t+1})$. During the training process, the evaluation network hyperparameters θ_i are used to periodically update the target network hyperparameters $\tilde{\theta}_i$ as follows:

$$\tilde{\theta}_i \leftarrow \rho \tilde{\theta}_i + (1 - \rho) \theta_i, \quad i = 1, 2, \dots, N, \quad (30)$$

where ρ is an updating factor. Moreover, we can train the hyperparameter ϕ of the policy π_ϕ by minimizing the following loss function:

$$\mathcal{L}_\phi(\phi) = \mathbb{E}_{\zeta_t} \left[\frac{1}{N} \sum_{i=1}^N Q_{\theta_i}(s_t, a_t) + \xi \mathcal{H}(\pi(a_\phi(s)|s_t)) \right], \quad (31)$$

where $a_\phi(s) \sim \pi_\phi(\cdot|s)$ represents the action according to the policy π_ϕ . The REDQ algorithm generates a Gaussian distribution with a mean value $\mu_\phi(s_t)$ and a standard deviation value $\sigma_\phi(s_t)$ such that the actual value of $\alpha_{RL}(s_t)$ is computed as [33]

$$\alpha_{RL}(s_t) = \frac{1}{2} \tanh(\mu_\phi(s_t) + \sigma_\phi(s_t) \odot \varphi) + \frac{1}{2}, \quad (32)$$

where \odot denotes the Hadamard product, and $\varphi \sim N(0, I)$ is an independent noise. The $\mu_\phi(s_t)$ network and the $\sigma_\phi(s_t)$ network are initialized with random values when they are untrained.

To solve the problem of low data utilization, we introduce the update-to-data (UTD) ratio U to control the number of data reuse. During the training process, we periodically get multiple samples from the replay buffer \mathcal{D} to improve the data utilization [34]. Algorithm 1 shows the update process of the proposed deep RL algorithm, for which the soft update method is used to update the Q -networks. The REDQ algorithm can provide the optimal policy π^* in (26). Hence, we can compute the most appropriate value of α_{RL} for the optimization problem (22) according to the current state as in (32). Moreover, note that the trained RL algorithm does not induce any additional computational burden to the 4IWMD electric vehicle control problem.

Algorithm 1 Randomized Ensembled Double Q -Learning

- 1: Initialize N hyperparameters θ_i of $Q_\theta(s_t, a_t)$
Set target hyperparameters $\tilde{\theta}_i \leftarrow \theta_i$, $i \in \{1, 2, \dots, N\}$
Empty the replay buffer \mathcal{D}
 - 2: **while** agent interacting with the environment **do**
 - 3: Select an action a_t based on $\pi_\phi(\cdot|s_t)$
 - 4: Collect r_t and s_{t+1} after taking the action a_t
 - 5: Store the data $\{s_t, a_t, r_t, s_{t+1}\}$ in the replay buffer \mathcal{D}
 - 6: **for** U updates **do**
 - 7: Sample a mini-batch $\{(s_t, a_t, r_t, s_{t+1})\}$ from \mathcal{D}
 - 8: Select randomly a subset $\mathcal{I}_M \subseteq \{1, 2, \dots, N\}$
 - 9: Compute the target value y^* with expression (28)
 - 10: **for** $i = 1, 2, \dots, N$ **do**
 - 11: Update θ_i using $\mathcal{L}_\theta(\theta_i)$ in (29) and gradient descent
 - 12: Update the hyperparameters of $\tilde{Q}_{\tilde{\theta}}(s_t, a_t)$ with (30)
 - 13: **end for**
 - 14: **end for**
 - 15: Update ϕ using $\mathcal{L}_\phi(\phi)$ in (31) and gradient descent
 - 16: **end while**
-

IV. ILLUSTRATIVE RESULTS AND EVALUATIONS

This section first describes the test environment and some performance indicators. Then, several case studies are presented to demonstrate the effectiveness of the proposed control method. Due to the over-actuation feature, 4IWMD electric vehicles can normally perform the driving tasks if one single motor or two motors on different sides are damaged. However, driving with two simultaneously damaged electric motors is not common in practice. Hence, here we focus on the torque distribution in case of partial or complete damage of a single electric motor. Note that when the RL algorithm is trained, the computational burden of the proposed FTC method mainly depends on the online NMPC optimization problem (11), which has been experimentally shown to be compatible with real-time control implementation for 4IWMD EVs in [18].

A. Simulation Setting

For validation purposes, a simulation platform is set up as shown in Fig. 1. A high-fidelity 4IWMD electric vehicle model is built using CarSim platform, the control algorithms are implemented in Matlab/Simulink, and the RL algorithm is implemented in Python. The signals can be transmitted and communicated between CarSim, Matlab/Simulink and Python for co-simulations. For test scenarios, we set $t_p = 0.1$, and Δd is a uniformly distributed random number with a sampling time of 0.05 and an amplitude of 0.03 for the fault factor estimation (8). The nonlinear vehicle model (10) is discretized using the multiple shooting discretization method with a 4th-order Runge Kutta solver [37]. The NMPC algorithm implemented at the upper control level is solved using the ACADO toolkit [31]. The numerical efficiency of this NMPC control implementation for real-time EVs control has been shown in [4], [16], [18]. To verify the effectiveness of the proposed control method, we select the following five representative control strategies for comparisons and validations.

1) *RL-FTC*: This is the proposed fault-tolerant hierarchical control scheme with the REDQ RL algorithm.

2) *SAC-FTC*: This control scheme has the same structure as RL-FTC. However, a commonly used Soft Actor-Critic RL algorithm [33] is used in place of the REDQ RL algorithm as for RL-FTC control.

3) *CO-FTC*: To our knowledge, FTC for 4IWMD electric vehicles considering both power consumption and vehicle stability has not been available in the literature. Hence, the control method in [18] is adapted taking into account the motor faults in the control design. The stability-consumption tradeoff of this control method is managed using fuzzy logic.

4) *CO w/o FTC*: The control strategy is similar as CO-FTC without taking into account the motor faults, *i.e.*, the control results in [18].

5) *w/o control*: There is no vehicle lateral control, the steering is controlled by the driver, and the speed tracking is performed by a tuned PID controller available in CarSim.

The following evaluation indicators are proposed to analyze the quality of the torque distribution algorithms in terms of handling stability, driver workload, motor load, additional yaw moment, and velocity tracking:

- Handling stability indicator $\mathcal{E}_s = \int_0^t (e_\beta^2 + e_\gamma^2) d\tau$, where $e_\beta = \beta - \beta_d$ and $e_\gamma = \gamma - \gamma_d$.
- Driver workload indicator $\mathcal{E}_d = -\mathcal{R}_{driver}$, where \mathcal{R}_{driver} is defined in (25).
- Motor load indicator $\mathcal{E}_m = \int_0^t \sum_{ij \in \{fl, fr, rl, rr\}} \Delta T_{ij}^2 d\tau$.
- Additional yaw moment indicator $\mathcal{E}_{M_z} = \int_0^t M_z^2 d\tau$.
- Velocity tracking indicator $\mathcal{E}_{v_x} = \int_0^t (v_{xd} - v_x)^2 d\tau$.

B. Test 1: Analysis of Stability-Consumption Tradeoff

For this test, we examine the role of different weights α_{RL} to analyze the tradeoff between vehicle stability and power consumption. To this end, the electric vehicle performs a double lane change (DLC) with $\mu = 0.5$ and $v_{xd} = 72$ (km/h). Moreover, the motor fault occurs at 7s with $k_{rl} = 0.5$.

The value of α_{RL} is set as a constant for several experiments with the proposed FTC method and the corresponding results are shown in Table II. Following the definition of the integrated objective function \mathcal{J} in (20), a smaller α_{RL} leads to a less motor power consumption \mathcal{E} (kJ). However, note from Table II that when $\alpha_{RL} = 0$ the motors consume 3.9% more energy than when $\alpha_{RL} = 0.25$. This is because the vehicle stability is degraded when the torque distribution focuses on the power consumption as shown by the values of \mathcal{E}_s . Specifically, when the power consumption becomes higher, the upper-level controller generates more additional yaw moment M_z to maintain the stability, which in turn requires an important torque to generate the additional yaw moment. The increasing torque affects the motor power consumption. At the lower control level, the torque distribution based on the quadratic optimization problem (22) can only get the optimal solution at the current moment. Then, the whole running state of the vehicle should be considered via RL to achieve the best stability-consumption tradeoff for the whole vehicle running state. Note that the weight allocation based on fuzzy logic [18] cannot allow to get the optimal solution under the whole running condition.

TABLE II
PERFORMANCE EVALUATION W.R.T. DIFFERENT RL WEIGHTS α_{RL} .

Weight	\mathcal{E}_s	\mathcal{E}_d	\mathcal{E}_m [N ² · m ² · s]	\mathcal{E}_{M_z} [N ² · m ² · s]	\mathcal{E} [kJ]
$\alpha_{RL} = 1$	0.3898	16.84	27290	742900	128.4
$\alpha_{RL} = 0.75$	0.3901	17.13	24410	751700	128.4
$\alpha_{RL} = 0.5$	0.4019	18.13	31760	753500	121.5
$\alpha_{RL} = 0.25$	0.4104	18.35	27680	766800	119.1
$\alpha_{RL} = 0$	0.4263	18.59	24700	801100	123.9

C. Test 2: Training with a Random Motor Fault

To train a well-adapted RL agent for multiple driving conditions, we select a DLC with straight and turning traffic conditions with $\mu = 0.3$ and $v_{xd} = 72$ (km/h), which can cover different levels of driving risk. Also, to simulate the motor faults, we set $k_{ij} = 0.5$, for $t \geq 7s$, where ij where is randomly selected from the set $\{fl, fr, rl, rr\}$.

For both RL-FTC and SAC-FTC controllers, the agents are trained under the same conditions. Observe from Fig. 4(a) that both control strategies converge to the highest rewards when increasing the number of training episodes. Also, Figs. 4(b) and (c) show that both controllers increase the vehicle stability while reducing the power consumption. Despite their effectiveness, we can see however that the RL-FTC controller achieves higher rewards and converges faster than the SAC-FTC controller, *i.e.*, the rewards of the RL-FTC controller begin to increase at episode 180 in place of episode 250 for the SAC-FTC controller.

For comparisons, we select the trained agents at different episodes to describe the variation of the agents under test conditions with an increasing training. Fig. 5 depicts the vehicle performance in case of a motor fault factor $k_{fr} = 0.5$ under 0-700 training episodes. As shown in Fig. 5(a), the vehicle can complete the DLC under all control results, which indicates that the

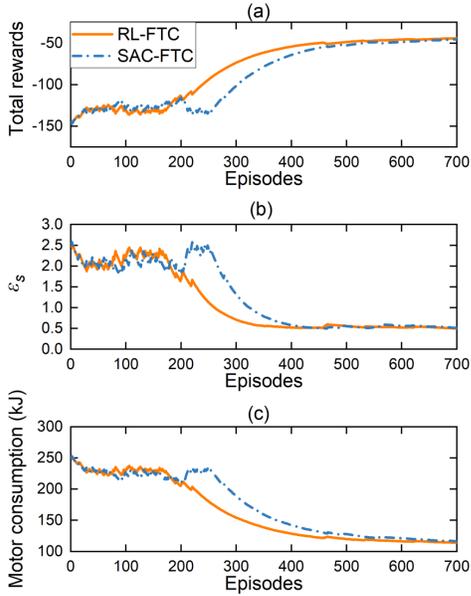


Fig. 4. RL algorithm training results. (a) Total rewards, (b) Performance indicator \mathcal{E}_s , (c) Motor power consumption \mathcal{E} .

RL-FTC controller without training can also properly ensure the considered driving task. Meanwhile, Figs. 5(b) and (c) show that the vehicle stability increases with the number of training episodes. In particular, the test performance improves more significantly with the 200-episodes and 500-episodes trained agents. This is because in these cases, the agent receives valid data from the replay buffer, which significantly decreases the loss of the estimated Q -network. This result is also consistent with the trend depicted in Fig. 4(b).

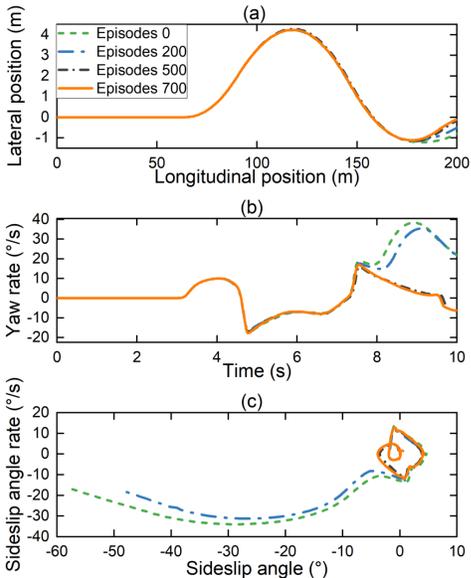


Fig. 5. Test results with different training episodes. (a) Vehicle trajectory, (b) Yaw rate, (c) Phase trajectory portrait.

Table III shows the results of the evaluation metrics according to the number of training episodes. Compared to the untrained agent, the stability indicator of the agent trained with 700 sets is increased by 93%, while the motor power consumption indicator

is decreased by 77%. The reduction of power consumption is also confirmed by the result shown in Fig. 6(a). We can see that in addition to an improved vehicle stability and a reduced power consumption, the control performance in terms of driver workload, motor load, additional yaw moment, and speed tracking ability are also improved. The behavior of α_{RL} according to the number of training episodes is presented in Fig. 6(b). Without training, the value of α_{RL} is generated based on a randomly generated action network. When the number of training sets reaches 200, the RL algorithm starts learning the behavior of α_{RL} to achieve a good reward. However, the action variation of α_{RL} is not yet significant. After 500 training episodes, the RL training results allow to achieve the best reward with some local tuning.

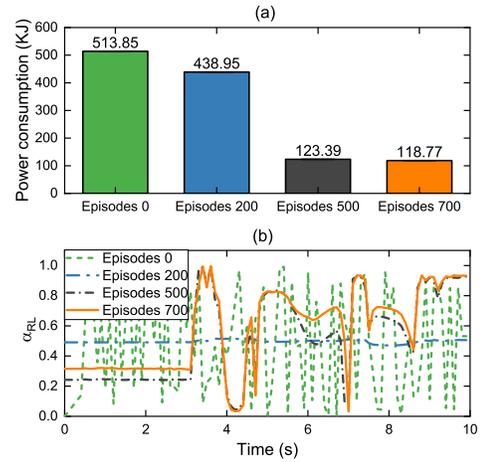


Fig. 6. Test results obtained with different numbers of training episodes. (a) Motor power consumption indicator \mathcal{E} , (b) RL action of α_{RL} .

TABLE III
PERFORMANCE EVALUATION W.R.T. DIFFERENT TRAINING EPISODES.

Episodes	\mathcal{E}_s	\mathcal{E}_d	\mathcal{E}_m [$N^2 \cdot m^2 \cdot s$]	\mathcal{E}_{M_z} [$N^2 \cdot m^2 \cdot s$]	\mathcal{E}_{v_x} [m^2/s]
0	6.678	92.11	50370	1.40E+06	30.0
200	5.306	85.62	39410	1.41E+06	12.1
500	0.506	18.80	29570	1.02E+06	1.16E-06
700	0.494	16.63	23530	9.56E+05	1.26E-06

We now examine the effect of different fault estimation results on the control performance of the proposed RL-FTC controller. To this end, we compare the evaluation indicators obtained with different values of the time delay t_p , which is the main parameter in the fault factor estimation expression (8). Fig. 7 depicts the behaviors of α_{RL} with respect to different values of t_p . As expected, after the occurrence of the motor fault at 7s, the estimate \hat{k}_{fr} converges quickly to its reference with small values of t_p , and the corresponding responses of the RL action α_{RL} are faster. However, we can see in Table IV that the variations of the performance indicators in function of the time delays are not significant. This means that the proposed RL-FTC controller can robustly perform the driving task despite the delay effects.

D. Test 3: Acceleration on a Split- μ Surface with Motor Fault

This test is used to verify the effectiveness of the proposed control strategy under straight-line acceleration conditions on split- μ surfaces. The road adhesion coefficients on the left side and on

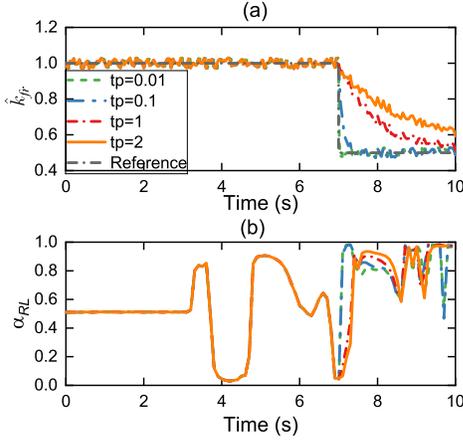


Fig. 7. Test results with different values of time delay. (a) Estimated fault factor k_{rl} , (b) RL action of α_{RL} .

TABLE IV

PERFORMANCE EVALUATION W.R.T. DIFFERENT VALUES OF TIME DELAY.

t_p	\mathcal{E}_s	\mathcal{E}_d	\mathcal{E}_m [N ² · m ² · s]	\mathcal{E}_{M_z} [N ² · m ² · s]	\mathcal{E} [KJ]	\mathcal{E}_{v_x} [m ² /s]
0.01	0.4935	15.14	20670	865200	118.75	1.63E-05
0.1	0.494	16.63	23530	956000	118.77	1.63E-05
1	0.5068	16.2	18420	975600	119.47	1.92E-05
2	0.5695	16.68	20400	871900	119.17	2.08E-05

the right side of the road are respectively set as 0.1 and 0.8. The initial longitudinal velocity is 80 km/h, and it accelerates uniformly to 120 km/h within 20s. For this test, the left front motor is subject to fault $k_{fl} = 0$, *i.e.*, this electric motor is completely damaged, for $t \geq 10$ s.

Fig. 8 shows that from 10s, *i.e.*, 250m, the vehicle stability starts to decrease due to the motor fault. We can see in Fig. 7(a) that the lateral offset of both FTC controllers, *i.e.*, RL-FTC and CO-FTC, are smaller than those without FTC, but the difference is not significant. This is because of the controllers without considering the FTC sacrifice a part of the velocity tracking ability. As shown in Fig. 7(b), the maximum velocity deviations of the RL-FTC controller and the CO-FTC controller are -0.23 km/h and -0.25 km/h, respectively, while the maximum velocity deviations obtained with the CO w/o FTC controller and with the case without lateral control are -1.03 km/h and -1.36 km/h, respectively. Similarly, as shown in Fig. 7(c), these latter sacrifice a part of the speed tracking ability, the difference in the yaw rate between the four control strategies is not significant.

As depicted in Fig. 9, the phase trajectory portrait shows a small difference in vehicle stability. However, the RL-FTC controller can save about 53% of motor power consumption compared to the strategy without lateral control, and an improvement compared to both the CO-FTC controller and the CO w/o FTC controller. This is because the current driving conditions are less risky and the RL-FTC and CO-FTC controllers tend to adopt a torque distribution approach that minimizes the power consumption. Note that the CO-FTC controller does not save as much energy as the RL-FTC controller.

As shown in Fig. 10, from 10s, the left front wheel motor is completely damaged and the output torque rapidly turns to 0. After the fault occurrence, the torques of the other wheels quickly

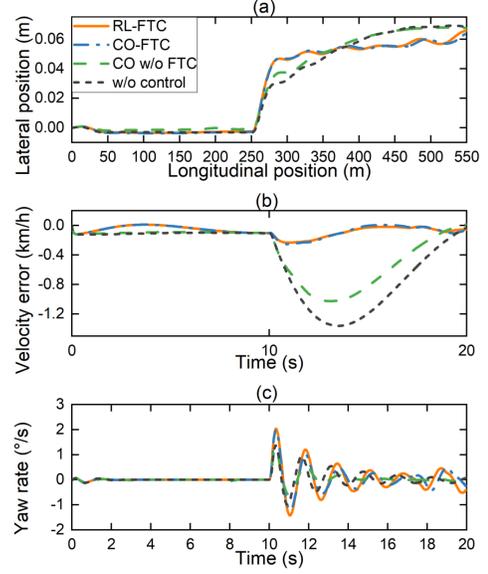


Fig. 8. Performance comparison results obtained with straight-line acceleration conditions. (a) Vehicle trajectory, (b) Velocity error, (c) Yaw rate.

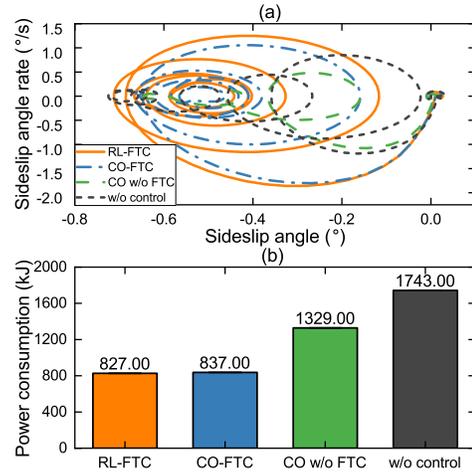


Fig. 9. Performance comparison results obtained with straight-line acceleration conditions. (a) Sideslip angle rate, (b) Motor power consumption \mathcal{E} .

increase to ensure the speed tracking performance. Note from Fig. 10(c) that the left rear wheel torques of both RL-FTC and CO-FTC controllers quickly decrease after a short increase. This is because the left rear wheel is on the same side as the damaged left front wheel, and the left and right side torques need to be maintained in balance to ensure the vehicle stability. Table V shows the evaluation indicators of the four controllers under linear acceleration conditions. For the RL-FTC controller, the stability indicator is slightly reduced. This is because to maintain the speed tracking capability, the controller tends to consider saving the motor power consumption since the current condition is without high risk.

E. Test 4: Double Lane Change with Motor Fault

This test aims to analyze the control performance under extreme driving conditions. To simulate an emergency lane change behavior under a low adhesion road surface, we set $\mu = 0.3$ and $v_{xd} = 72$

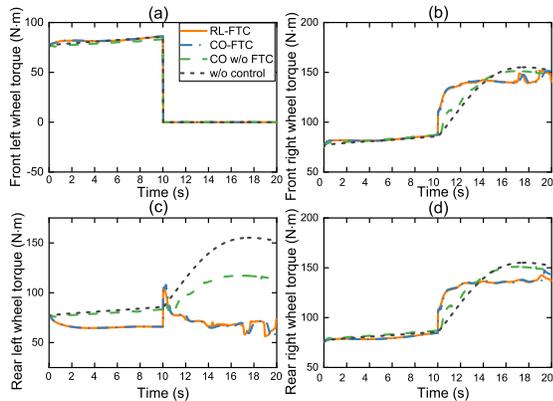


Fig. 10. Results of four-wheel torques in straight-line conditions. (a) Front left wheel torque, (b) Front right wheel torque, (c) Rear left wheel torque, (d) Rear right wheel torque.

TABLE V
PERFORMANCE INDICATORS IN STRAIGHT-LINE CONDITIONS.

Controller	\mathcal{E}_s	\mathcal{E}_d	\mathcal{E}_m [N ² · m ² · s]	\mathcal{E}_{M_z} [N ² · m ² · s]	\mathcal{E}_{v_x} [m ² /s]
RL-FTC	0.0093	6.548	4575	2.35E+05	0.014
CO-FTC	0.0089	6.579	4724	2.30E+05	0.016
CO w/o FTC	0.0031	6.702	3102	2.36E+05	0.355
w/o control	0.0053	6.719	3134	0.00E+00	0.658

km/h. The electric vehicle performs a DLC task with the reference trajectory shown in Fig. 11(a). A fault of the rear left motor occurs at 7s with $k_{rl} = 0.5$.

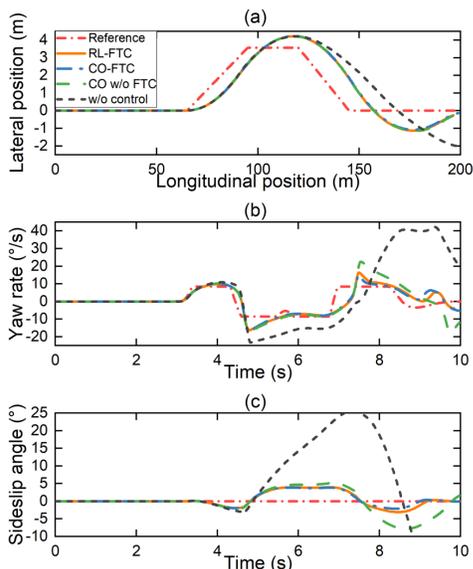


Fig. 11. Performance comparison results obtained with DLC conditions. (a) Vehicle movement trajectory, (b) Yaw rate, (c) Sideslip angle rate.

We can see from Fig. 11(a) that RL-FTC, CO-FTC, and CO w/o FTC controllers can complete the considered DLC task, which is not the case of the strategy without lateral control. In Figs. 11(b) and (c), the performance of RL-FTC and CO-FTC controllers is better than that of CO w/o FTC controller. In particular, after 7s the CO w/o FTC controller shows a significant variation, which

indicates the effectiveness of the FTC in maintaining the vehicle stability with faulty motors. There is no significant difference in terms of stability performance between both RL-FTC and CO-FTC controllers, which guarantee a stable vehicle driving. This is also explicitly confirmed with the phase trajectory portrait in Fig. 12(a). Both RL-FTC and CO-FTC controllers have similar trajectories in the phase plane with lower boundaries than CO w/o FTC and w/o control controllers. Note that without lateral control, the vehicle state exceeds the stability boundary and cannot return to a stable point. As shown in Fig. 12(b), the RL-FTC controller leads to the lowest motor power consumption, which confirms that the RL-FTC controller can reduce the power consumption while guaranteeing the vehicle stability even with extreme driving conditions. With the quantitative results in Table VI, we can also see the advantages of the RL-FTC controller over other compared ones in terms of power consumption, driver burden and vehicle stability.

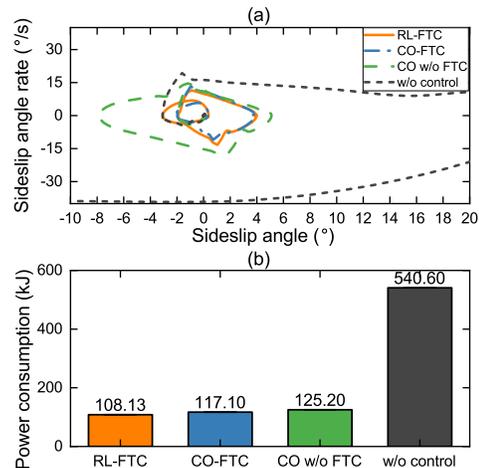


Fig. 12. Performance comparison results obtained with DLC conditions. (a) Phase trajectory portrait. (b) Motor power consumption \mathcal{E} .

TABLE VI
PERFORMANCE INDICATORS IN DLC CONDITIONS.

Controller	\mathcal{E}_s	\mathcal{E}_d	\mathcal{E}_m [N ² · m ² · s]	\mathcal{E}_{M_z} [N ² · m ² · s]	\mathcal{E}_{v_x} [m ² /s]
RL-FTC	0.4485	17.34	16300	747600	9.3E-07
CO-FTC	0.4391	19.01	24270	869300	1.2E-06
CO w/o FTC	0.9724	13.99	18510	458700	2.7E-06
w/o control	10.010	88.28	356100	0	2.2E+01

V. CONCLUDING REMARKS

A new fault-tolerant model predictive control method considering both the vehicle stability and the motor power consumption has been proposed. To achieve the best tradeoff between vehicle stability and power consumption, the weighting factor of the optimization-based torque distribution problem is adjusted online based on the REDQ deep RL algorithm. Moreover, several tests and analysis are performed with different values of the weighting factor to highlight that fuzzy logic cannot allow for an optimal control strategy for the whole vehicle operating range. Different evaluation indicators are given to quantify the performance of various control strategies under multiple driving scenarios. The co-simulation results obtained with a CarSim electric vehicle model show that the proposed FTC scheme can reduce the power consumption and

the driver workload while guaranteeing the vehicle stability despite the presence of motor faults. For future works, we consider FTC control design for different types of motor faults and different steering characteristics of drivers. Moreover, vehicle experiments should be conducted to verify the real-time performance of the proposed control results.

REFERENCES

- [1] Y. Mao, S. Zuo, and J. Cao, "Effects of rotor position error on longitudinal vibration of electric wheel system in in-wheel PMSM driven vehicle," *IEEE/ASME Trans. Mechatron.*, vol. 23, no. 3, pp. 14–25, 2018.
- [2] J.-S. Hu, Y. Wang, H. Fujimoto, and Y. Hori, "Robust yaw stability control for in-wheel motor electric vehicles," *IEEE/ASME Trans. Mechatron.*, vol. 22, no. 3, pp. 1360–1370, 2017.
- [3] A. Karki, S. Phuyal, D. Tuladhar, S. Basnet, and B. Shrestha, "Status of pure electric vehicle power train technology and future prospects," *Appl. Syst. Innov.*, vol. 3, no. 3, p. 35, 2020.
- [4] M. Dalboni, D. Tavernini, U. Montanaro, A. Soldati, C. Concaro, M. Dhaens, and A. Sorniotti, "Nonlinear model predictive control for integrated energy-efficient torque-vectoring and anti-roll moment distribution," *IEEE/ASME Trans. Mechatron.*, vol. 26, no. 3, pp. 1212–1224, 2021.
- [5] W. Zhao, A. Wang, S. Zou, and H. Zhang, "Individual auxiliary and fault-tolerant control of steer-by-wire system considering different drivers steering characteristics," *IEEE/ASME Trans. Mechatron.*, vol. 26, no. 3, pp. 1558–1569, 2020.
- [6] D. Zhang, G. Liu, H. Zhou, and W. Zhao, "Adaptive sliding mode fault-tolerant coordination control for four-wheel independently driven electric vehicles," *IEEE Trans. Indus. Electron.*, vol. 65, no. 11, pp. 9090–9100, 2018.
- [7] R. Wang and J. Wang, "Fault-tolerant control with active fault diagnosis for four-wheel independently driven electric ground vehicles," *IEEE Trans. Veh. Technol.*, vol. 60, no. 9, pp. 4276–4287, 2011.
- [8] B. Zhang, S. Lu, W. Wu, C. Li, and J. Lu, "Robust fault-tolerant control for four-wheel individually actuated electric vehicle considering driver steering characteristics," *J. Franklin Inst.*, vol. 358, no. 11, pp. 5883–5908, 2021.
- [9] R. Wang, H. Zhang, and J. Wang, "Linear parameter-varying controller design for four-wheel independently actuated electric ground vehicles with active steering systems," *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 4, pp. 1281–1296, 2013.
- [10] S. Aouaouda, M. Chadli, M. Boukhnifer, and H. Karimi, "Robust fault tracking controller design for vehicle dynamics: A descriptor approach," *Mechatronics*, vol. 30, pp. 316–326, 2015.
- [11] J. Guo, J. Wang, Y. Luo, and K. Li, "Robust lateral control of autonomous four-wheel independent drive electric vehicles considering the roll effects and actuator faults," *Mech. Syst. Signal Process.*, vol. 143, p. 106773, 2020.
- [12] D. Aravindh, R. Sakthivel, B. Kaviarasan, S. Anthoni, and F. Alzahrani, "Design of observer-based non-fragile load frequency control for power systems with electric vehicles," *ISA Trans.*, vol. 91, pp. 21–31, 2019.
- [13] H. Tang, Y. Chen, and A. Zhou, "Actuator fault-tolerant control for four-wheel-drive-by-wire electric vehicle," *IEEE Trans. Transp. Electrification*, vol. 8, no. 2, pp. 2361–2373, 2022.
- [14] B. Guo and Y. Chen, "Robust adaptive fault-tolerant control of four-wheel independently actuated electric vehicles," *IEEE Trans. Indus. Inf.*, vol. 16, no. 5, pp. 2882–2894, 2019.
- [15] Y. Wang, S. Yu, J. Yuan, and H. Chen, "Fault-tolerant control of electric ground vehicles using a triple-step nonlinear approach," *IEEE/ASME Trans. Mechatron.*, vol. 23, no. 4, pp. 1775–1786, 2018.
- [16] A. Parra, D. Tavernini, P. Gruber, A. Sorniotti, A. Zubizarreta, and J. Pérez, "On nonlinear model predictive control for energy-efficient torque-vectoring," *IEEE Trans. Veh. Technol.*, vol. 70, no. 1, pp. 173–188, 2020.
- [17] H. Peng, W. Wang, C. Xiang, L. Li, and X. Wang, "Torque coordinated control of four in-wheel motor independent-drive vehicles with consideration of the safety and economy," *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, pp. 9604–9618, 2019.
- [18] H. Deng, Y. Zhao, S. Feng, Q. Wang, C. Zhang, and F. Lin, "Torque vectoring algorithm based on mechanical elastic electric wheels with consideration of the stability and economy," *Energy*, vol. 29, p. 63, 2021.
- [19] Y. Zhu, Z. Wang, C. Chen, and D. Dong, "Rule-based reinforcement learning for efficient robot navigation with space reduction," *IEEE/ASME Trans. Mechatron.*, vol. 27, no. 2, pp. 846–857, 2022.
- [20] H. Wei, N. Zhang, J. Liang, Q. Ai, W. Zhao, T. Huang, and Y. Zhang, "Deep reinforcement learning based direct torque control strategy for distributed drive electric vehicles considering active safety and energy saving performance," *Energy*, vol. 238, p. 121725, 2022.
- [21] L. Zhang, R. Zhang, T. Wu, R. Weng, M. Han, and Y. Zhao, "Safe reinforcement learning with stability guarantee for motion planning of autonomous vehicles," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 12, pp. 5435–5444, 2021.
- [22] D. Zhang and Z. Gao, "Reinforcement learning-based fault-tolerant control with application to flux cored wire system," *Meas. Control*, vol. 51, no. 7–8, pp. 349–359, 2018.
- [23] —, "Fault tolerant control using reinforcement learning and particle swarm optimization," *IEEE Access*, vol. 8, pp. 168 802–168 811, 2020.
- [24] L. Xiong, Z. Yu, Y. Wang, C. Yang, and Y. Meng, "Vehicle dynamics control of four in-wheel motor drive electric vehicle using gain scheduling based on tyre cornering stiffness estimation," *Veh. Syst. Dyn.*, vol. 50, no. 6, pp. 831–846, 2012.
- [25] C. Hu, R. Wang, F. Yan, and N. Chen, "Should the desired heading in path following of autonomous vehicles be the tangent direction of the desired path?" *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 6, pp. 3084–3094, 2015.
- [26] R. Rajamani, *Vehicle Dynamics and Control*. Springer US, 2012.
- [27] H. Zhang, J. Liang, H. Jiang, Y. Cai, and X. Xu, "Stability research of distributed drive electric vehicle by adaptive direct yaw moment control," *IEEE Access*, vol. 7, pp. 225–237, 2019.
- [28] H. Pacejka and E. Bakker, "The magic formula tyre model," *Veh. Syst. Dyn.*, vol. 21, pp. 1–18, 1992.
- [29] R. Wang and J. Wang, "Actuator-redundancy-based fault diagnosis for four-wheel independently actuated electric vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 1, pp. 239–249, 2013.
- [30] G. Zhang, H. Zhang, X. Huang, J. Wang, H. Yu, and R. Graaf, "Active fault-tolerant control for electric vehicles with independently driven rear in-wheel motors against certain actuator faults," *IEEE Trans. Control Syst. Technol.*, vol. 24, no. 5, pp. 1557–1572, 2015.
- [31] B. Houska, H. Ferreau, and M. Diehl, "An auto-generated real-time iteration algorithm for nonlinear MPC in the microsecond range," *Automatica*, vol. 47, no. 10, pp. 2279–2285, 2011.
- [32] L. Gurobi Optimization, "Gurobi optimizer reference manual," 2018. [Online]. Available: <https://www.gurobi.com>
- [33] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," in *Int. Conf. Machine Learning*, 2018, pp. 1861–1870.
- [34] X. Chen, C. Wang, Z. Zhou, and K. Ross, "Randomized ensembled double Q-learning: Learning fast without a model," *arXiv preprint arXiv:2101.05982*, 2021.
- [35] X. Wu, B. Zhou, G. Wen, L. Long, and Q. Cui, "Intervention criterion and control research for active front steering with consideration of road adhesion," *Veh. Syst. Dyn.*, vol. 56, no. 4, pp. 553–578, 2018.
- [36] L. Zhai, T. Sun, and J. Wang, "Electronic stability control based on motor driving and braking torque distribution for a four in-wheel motor drive electric vehicle," *IEEE Trans. Veh. Technol.*, vol. 65, no. 6, pp. 4726–4739, 2016.
- [37] M. Diehl, H. Bock, H. Diedam, and P.-B. Wieber, "Fast direct multiple shooting algorithms for optimal robot control," in *Fast Motions in Biomechanics and Robotics*. Springer, 2006, pp. 65–93.

Huifan Deng is a Ph.D. student at Nanjing University of Aeronautics and Astronautics. His research interests include vehicle dynamics control, and human-machine shared control.

Youqun Zhao is currently a Professor at Nanjing University of Aeronautics and Astronautics, China. His research interests include vehicle dynamics and vehicle control.

Anh-Tu Nguyen is an Associate Professor at the INSA Hauts-de-France, Valenciennes, France. His research interests include robust control and estimation, human-machine shared control with a strong emphasis on mechatronic applications. Dr. Nguyen is an Associate Editor for several international journals.

Chao Huang is a Research Assistant Professor at the Hong Kong Polytechnic University. Her research interests include human-machine collaboration, fault-tolerant control, mobile robots.