Safe Reinforcement Learning with a Predictive Safety Filter for Motion Planning and Control: A Drifting Vehicle Example

Bei Zhou[1], Baha Zarrouki[2], Mattia Piccinini[2], Cheng Hu[1], Lei Xie[1]*, Johannes Betz[2]

***Abstract*— Autonomous drifting is a complex and crucial maneuver for safety-critical scenarios like slippery roads and emergency collision avoidance, requiring precise motion planning and control. In this paper, we propose a novel Safe Reinforcement Learning (RL)-based motion planner for autonomous drifting. Our approach integrates an RL agent with model-based drift dynamics to determine desired drift motion states, while incorporating a Predictive Safety Filter (PSF) that adjusts the agent's actions online to prevent unsafe states. This ensures safe and efficient learning, and stable drift operation. We validate the effectiveness of our method through simulations on a Matlab-Carsim platform, demonstrating significant improvements in drift performance, reduced tracking errors, and computational efficiency compared to traditional methods. This strategy promises to extend the capabilities of autonomous vehicles in safety-critical maneuvers. This paper summarizes the methods and results reported in the full paper available at https://arxiv.org/abs/2506.22894.**

I. Motivation and Contributions

A. Limitation of existing literature

- Limited learning and adaptation capability: modeling errors can degrade the drift performance under changing environmental conditions [1], [2], [3], [4].
- Dependence on prior expert knowledge and data in the learning process: drift driving data from professional drivers [5], [6], [7] or initial policies based on prior knowledge [8], [9], [10].
- No examples of PSF within RL-based motion planners, to our current knowledge.
- MPC-based drift planners are computationally expensive [4] and may not be suitable for real-time operation in rapidly-changing scenarios.

B. Our Contributions

1) We propose a Safe RL-based drift motion planner that learns the road friction coefficient and adapts the reference curvature of the path to be tracked, ensuring accurate drift equilibrium calculation for a low-level Model Predictive Drift Controller (MPDC).
2) We design a Predictive Safety Filter (PSF) to enforce safe drifting maneuvers during training and inference while improving the RL agent performance.
3) The proposed Safe RL motion planner requires no prior expert knowledge for training.
4) Simulation results in Matlab-Carsim show that our Safe RL planner outperforms a state-of-the-art MPC planner benchmark in both closed-loop performance and computational efficiency for online operation.

*Corresponding Author

[1]Bei Zhou, Cheng Hu, Lei Xie are with the State Key Laboratory of Industrial Control Technology, Zhejiang University, 310058 Hang zhou, China zhoubei@zju.edu.cn, 22032081@zju.edu.cn, lxie@iipc.zju.edu.cn

[2]Baha Zarrouki, Mattia Piccinini, and Johannes Betz are with the Professorship of Autonomous Vehicle Systems, Technical University of Munich, 85748 Garching, Germany; Munich Institute of Robotics and Machine Intelligence (MIRMI). baha.zarrouki@tum.de, mattia.piccinini@tum.de, johannes.betz@tum.de

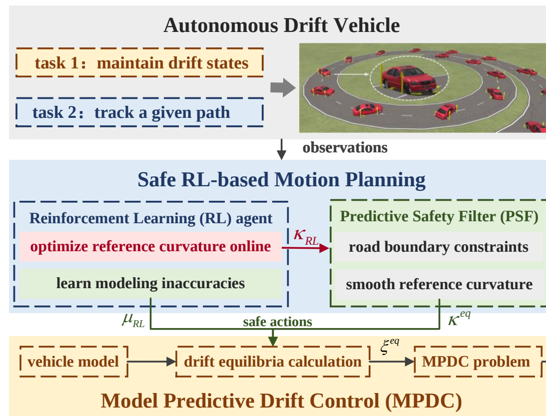


Fig. 1. Overview of the proposed framework for autonomous drifting along a variable-curvature path.

II. Framework Overview

Our framework (Fig. 1) enables autonomous drifting along a variable-curvature path. We design a hierarchical planning and control architecture, where a Safe RL-based motion planner learns the local reference path curvature and road friction coefficient, while a Model Predictive Drift Controller (MPDC) generates low-level drift controls based on the RL planner's output. We introduce a Predictive Safety Filter (PSF) to adjust the RL output curvature and ensure only safe actions reach the MPDC controller. The MPDC computes drift equilibrium points and generates controls to maintain the vehicle in the desired drift states (velocity, sideslip angle, yaw rate) while following the path. Algorithm 1 overviews the training of our safe RL motion planner.

III. Simulation Results

The training process for the proposed RL planners is shown in Fig. 2 and the tracking performance is presented in Fig. 3. The Safe RL planner significantly outperforms state-of-the-art MPC planners, achieving up to 61.1% reduction in mean heading error and 49.6% reduction in mean lateral deviation, while requiring 3.6 times less computational load.

Algorithm 1 Training our Safe RL Motion Planner

Initialize RL_agent
for episode $n = 1$ **to** N **do**
  Initialize the vehicle with $V = V_0$ at the start of the training path
  $\boldsymbol{s}_0 = [e, \Delta\psi, \delta, \kappa_r, e_{la}, h_e, h_{\text{conv}}] \leftarrow$ get initial observations
  for time $t = 0$ **to** T_{sim} **do**
    $\boldsymbol{a}_t = [\kappa_{\text{RL}}, \mu_{\text{RL}}] = \text{RL\_agent}(\boldsymbol{s}_t) \leftarrow$ compute RL actions
    $\kappa_0 = \text{PSF}(\kappa_{\text{RL}}) \leftarrow$ compute safe curvature
    $\boldsymbol{\xi}_t^{eq} = \text{drift\_equil}(\kappa_0, \mu_{\text{RL}}) \leftarrow$ compute drift equilibrium
    $\boldsymbol{\tau}_t = \text{MPDC}(\boldsymbol{\xi}_t^{eq}) \leftarrow$ compute vehicle controls
    $\boldsymbol{s}_{t+1} = \text{vehicle}(\boldsymbol{s}_t, \boldsymbol{\tau}_t) \leftarrow$ get next state
    $r_{t+1} = \text{reward}(\boldsymbol{s}_{t+1}) \leftarrow$ compute reward
    $\text{RL\_agent.update}(\boldsymbol{s}_t, \boldsymbol{a}_t, r_t, \boldsymbol{s}_{t+1}) \leftarrow$ train RL agent
    if $|e| > e_{\text{max}}$ **then** break $\leftarrow$ early termination
  end for
end for

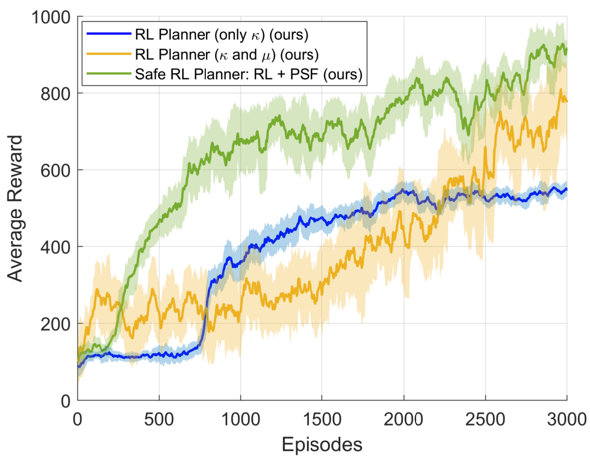
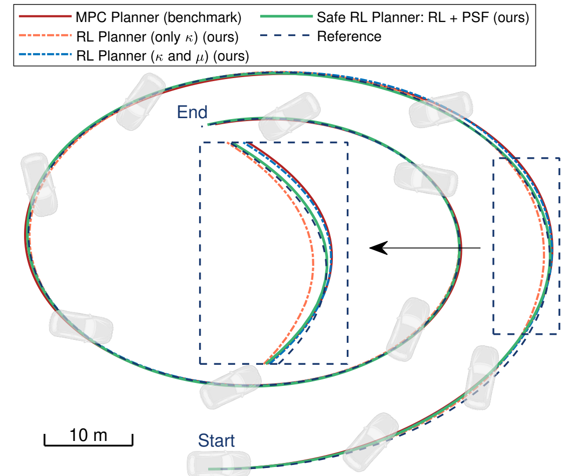


Fig. 2. The average rewards smoothened over 50 episodes.

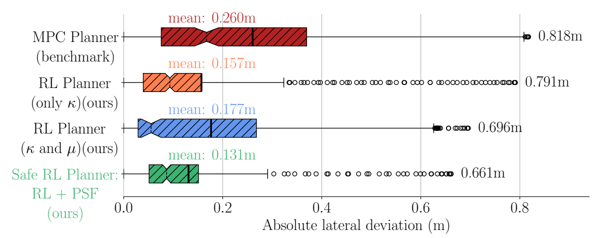
To assess generalization, we test a trained RL agent on three new unseen tracks with varying curvatures, and compare it to the MPC benchmark planner (Table I). The Safe RL planner consistently outperforms the MPC, demonstrating its ability to handle unpredictable environmental changes and its potential for real-world applications.

TABLE I
GENERALIZATION IN UNSEEN TEST TRACKS: SAFE RL VERSUS MPC.

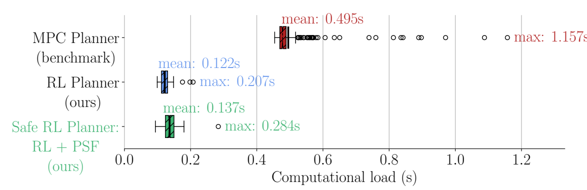
Track	**Curvature** (m^{-1}) Start	End	**MPC Planner** RMSE e (m)	**Safe RL Planner** RMSE e (m)
Training	1/40	1/20	0.36	0.13
Test Track 1	1/45	1/20	0.49	0.35
Test Track 2	1/40	1/25	0.37	0.26
Test Track 3	1/45	1/20	0.53	0.39



(a) Path tracking performance.



(b) Box plot for lateral path deviations.



(c) Computational times.

Fig. 3. Comparisons of path tracking performance among the proposed safe RL planner and benchmark planners.

REFERENCES

[1] G. Chen, X. Zhao, Z. Gao, and M. Hua, “Dynamic drifting control for general path tracking of autonomous vehicles,” *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 3, pp. 2527–2537, Mar. 2023.

[2] T. P. Weber and J. C. Gerdes, “Modeling and control for dynamic drifting trajectories,” *IEEE Transactions on Intelligent Vehicles*, vol. 9, no. 2, pp. 3731–3741, Feb. 2024.

[3] B. Lenzo, T. Goel, and J. Christian Gerdes, “Autonomous drifting using torque vectoring: Innovating active safety,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 11, pp. 17 931–17 939, Nov. 2024.

[4] C. Hu, L. Xie, Z. Zhang, and H. Xiong, “A novel model predictive controller for the drifting vehicle to track a circular trajectory,” *Vehicle System Dynamics*, vol. 0, no. 0, pp. 1–30, 2024.

[5] M. Acosta and S. Kanarachos, “Teaching a vehicle to autonomously drift: A data-based approach using neural networks,” *Knowledge-Based Systems*, vol. 153, pp. 12–28, Aug. 2018.

[6] P. Cai, X. Mei, L. Tai, Y. Sun, and M. Liu, “High-speed autonomous drifting with deep reinforcement learning,” *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1247–1254, Apr. 2020.

[7] X. Hou, J. Zhang, C. He, Y. Ji, J. Zhang, and J. Han, “Autonomous driving at the handling limit using residual reinforcement learning,” *Advanced Engineering Informatics*, vol. 54, p. 101754, Oct. 2022.

[8] M. Cutler and J. P. How, “Autonomous drifting using simulation-aided reinforcement learning,” in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, May 2016, pp. 5442–5448.

[9] S. Zhao, J. Zhang, X. He, C. He, X. Hou, H. Huang, and J. Han, “A harmonized approach: Beyond-the-limit control for autonomous vehicles balancing performance and safety in unpredictable environments,” *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–14, 2024.

[10] S. Zhao, J. Zhang, N. Masoud, Y. Jiang, H. Huang, and T. Liu, “High-speed cornering control and real-vehicle deployment for autonomous electric vehicles,” Nov. 2024.