

Track based Offline Policy Learning for Overtaking Maneuvers with Autonomous Racecars

Jayanth Bhargav¹, Johannes Betz¹, Hongrui Zheng¹, Rahul Mangharam¹

Abstract—The rising popularity of driver-less cars has led to the research and development in the field of autonomous racing, and overtaking in autonomous racing is a challenging task. Vehicles have to detect and operate at the limits of dynamic handling and decisions in the car have to be made at high speeds and high acceleration. One of the most crucial parts in autonomous racing is path planning and decision making for an overtaking maneuver with a dynamic opponent vehicle. In this paper we present the evaluation of a track based offline policy learning approach for autonomous racing. We define specific track portions and conduct offline experiments to evaluate the probability of an overtaking maneuver based on speed and position of the ego vehicle. Based on these experiments we can define overtaking probability distributions for each of the track portions. Further, we propose a switching MPCC controller setup for incorporating the learnt policies to achieve a higher rate of overtaking maneuvers. By exhaustive simulations, we show that our proposed algorithm is able to increase the number of overtakes at different track portions.

I. INTRODUCTION

A. Autonomous Racing

Autonomous racing has become popular over the recent years and competitions like Roborace or the Indy Autonomous Challenge as well as with small-scale racecars like F1Tenth [10] provide platforms for evaluating autonomous driving software. The overall goal of all these competitions is that researchers and engineers can develop algorithms that operate at the vehicles edge: High speeds, high accelerations, high computation power, adversarial environments. Similar to normal racing series like Formula 1 the development of algorithms for autonomous racing generate trust in the field of autonomous driving and enables the development of advanced autonomous driving algorithms.

The algorithms that were developed in the field of autonomous racing so far are mostly focusing on single vehicle only that try to achieve a human-like lap time. The field of high dynamic overtaking maneuver with dynamic opponents was less displayed so far. In addition, achieving a human-like behavior (e.g. like a Formula 1 race driver) that makes the decision about an overtaking maneuver and executes a secure and reliable maneuver at high speeds is still an unsolved question.

B. Contributions

Based on the state of the art, in this paper we present an offline policy learning for overtaking maneuvers in autonomous racing. This work has two primary contributions:

- 1) We provide a design of experiment (DoE) for an offline driven policy learning approach by track discretization. Based on the specific track we provide a discretization into turns and vary the position and velocity of the ego vehicle for these turns. With this DoE we create policies for the racecar that will enable a better overtaking maneuver. These policies learnt for different track portions are then integrated into path planning and control. We show that this approach is effective in increasing the number of overtakes at different track portions.
- 2) The overtaking policies learnt from the experiments are then incorporated into a MPCC planner. We propose a switched MPCC controller which combines a receding horizon control algorithm and specific driving behaviours like driving on the left side and driving on the right side. These modes restrict the action space for agent/vehicle and helps to enforce practical overtaking strategies like overtake from inside of track or stay on outside of the track at particular turns.

In summary, with the setup defined in this paper we can create more realistic and better overtaking maneuvers for autonomous race vehicles. It also provides the intuition and domain knowledge for tuning learning-based planners and controllers.

II. DESIGN OF EXPERIMENTS

The main idea behind making better overtaking is to learn from the track. Each track has a specific layout and a combination of turns and straights. Each turn has a specific curvature and therefore allows the car to drive only a a maximum velocity. In addition, to drive with the maximum speeds the car needs to follow a specific trajectory that leads through the turns.

We propose now an offline experiment setup which will create a detailed insights on specific track portions and the possibility to overtake in specific turns. With these offline experiments it is possible to create track-based policies that can be used to enhance the MPCC planner by integrating these policies into its cost-function and constraints.

A. Track Portions

In the first step we are defining the track portions that we will examine. We will use here four track portions that the most common kinds of curves/turns found on racetracks:

- *Straightaway*: This is a straight park of the track which has almost no sign of curvature. It is meant to let drivers push the cars to their speed limits. The straightaway

¹University of Pennsylvania, Department of Electrical and Systems Engineering, School of Engineering and Applied Sciences
joebetz@seas.upenn.edu

is the quickest and easiest track portion for overtaking maneuvers.

- *Sweeper*: The sweeper curve is a gradual, long and large turn. These curves are usually very wide and have high banking which makes it easier for race cars to pass by. The cars reach high velocities here and the overtaking can be done on the inside or the outside.
- *Hairpin*: The hairpin has the biggest curvature and is mostly the tightest turn on a track. Here we will see the car slow down to and make very sharp turns. overtaking maneuvers happen here mostly before entering the hairpin.
- *Chicane*: The chicane is a straightaway modified into a slight S shape. The velocities that the car is reaching here are varying from high to low based on the racetrack. Overtaking maneuvers happen here less often because its less space due to the steering maneuvers.

As a showcase in this paper we are using the Silverstone racetrack. The choice of Silverstone track for our experiments is backed by the fact that this track layout has all of the above types of curves which is challenging for an overtaking maneuver. In figure we display 9 different track portions on the Silverstone track that are marked with labels 1 to 8. The track portions are defined by $\mathbf{T} = \{\tau | \tau \in (1, 2, 3, 4, 5, 6, 7, 8)\}$ and are categorized in the four track segment types:

- *Straightaway*: 6
- *Sweeper*: 1, 4, 7
- *Hairpin*: 2, 3, 8
- *Chicane*: 5

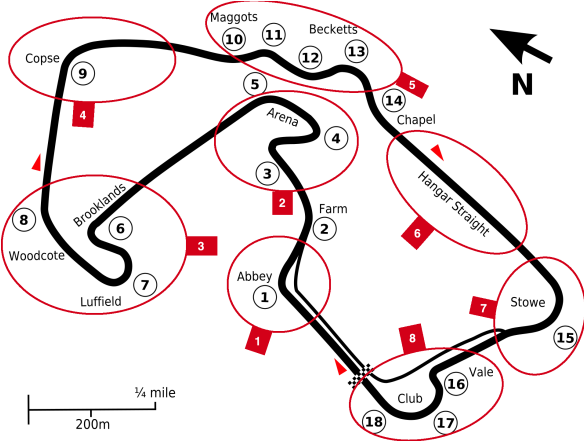


Fig. 1. Silverstone Circuit

B. Sampling based Trajectory Rollouts

To examine these defined track portions we setup an offline simulation that varies different parameters visualized in figure 2.

The simulation setup consists of two agents: One ego and one opponent vehicle. The opponent vehicle follows a pre-computed race line based on [4] and is non-interactive. The ego vehicle is starting behind the opponent vehicle and is an

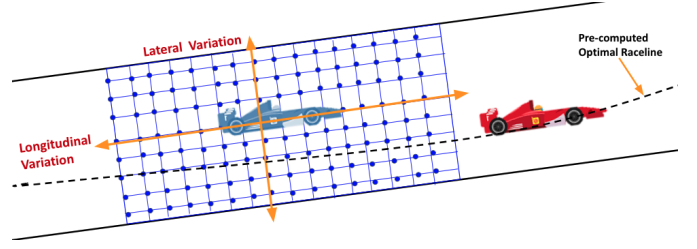


Fig. 2. Ego vehicle (blue car) starting behind the opponent (red car) vehicle. The opponent vehicle follows a pre-computed raceline, the position and starting velocity of the ego vehicle are varied.

intelligent agent that planning its path based on the MPCC algorithm. For every track portion $\tau \in \mathbf{T}$, a uniformly sampled set of positions $\mathcal{P} : \mathcal{X}\mathcal{Y} \subset \mathbf{R}^2$ are chosen as the starting position for the ego. The obstacle vehicle however always tracks the pre-computed race line. The obstacle vehicle speed is varied as $v_{obs} = v_{baseline} * (1 + s)$ where $s \in \{-0.2, 0, +0.2\}$, $v_{baseline}$ being the speed of the obstacle from the pre-computed curvature optimal race line.

The agents are initialised with these positions and set to start the simulation. The ego vehicle synthesises dynamic trajectories based on the MPCC planner with obstacle avoidance. A fully observable model is used for the ego vehicle i.e. the ego vehicle will have the information of the track portion τ which it is driving in and the current state of the obstacle $X_{obs} = [x_{obs}, y_{obs}, \phi_{obs}]$

In this setup, we conduct exhaustive simulations based on the following parameter variations:

- *Lateral offset*: The position of the ego vehicle is varied lateral across the track with an offset from the centerline.
- *Longitudinal Offset*: The position of the ego vehicle is varied longitudinal along the centerline of the track.
- *Opponent Speed Change*: The opponent speed is varied with $\pm 20\%$ from baseline

With this kind of simulations we have now the possibility to examine if a specific position in one of the defined track portions has an advantage for an overtaking maneuver or not.

It is highly expected that the ego vehicle will succeed in an overtaking maneuver when the obstacle speed is 20% lower than its baseline. This establishes the fact that speed advantage always helps in overtaking (e.g. DRS zones in F1). The next set of parameters that influence the overtaking maneuver is the position. In a convoluted track like the Silverstone circuit we can display if starting off at a specific position enables a higher chance of an overtaking maneuver. For each track portion, we define four regions of interest: \mathcal{R}_1 , \mathcal{R}_2 , \mathcal{R}_3 and \mathcal{R}_4 . Starting positions of the ego vehicle are uniformly sampled in all the four regions to generate experimental data.

III. RELATED WORK

Although the state of the art displays tons of algorithms for path and behavioral planning of autonomous vehicles, the explicit algorithm development for autonomous race cars is just done in a small community. As part of the Roborace

competition [1] and [3] presented a planning and control system for real life autonomous racing cars. Both approaches focused on a holistic software architecture that is capable of dynamic overtaking. Nevertheless none of them realized a head to head race with the vehicles. As a part of the same competition, [2] presented a nonlinear model predictive control (NMPC) for racing. The overtaking strategy was implemented as a term in the objective function. The NMPC has the freedom to choose the side for an overtake and was mainly relying on the obstacles velocity to perform the overtaking maneuver.

Beside these classical control driven approaches, new machine learning algorithms were used to learn how to race. [14] displayed an end-to-end framework for a computer game (Formula One (F1) Codemasters) that is using Convolution Neural Network (CNN) integrated with Long Short-Term Memory (LSTM) to learn how to drive fast and reliable around a racetrack. In [12] a reinforcement learning algorithm was presented that learns competitive visual control policies through self-play in imagination. Although this setup is not focusing on real vehicle dynamic behavior, it provided interesting strategies for multi vehicle interaction.

When it comes to multi vehicle racing, only a few researchers have considered the whole pipeline of path and behavioral planning in combination with the interactions the vehicle can make. [8] presented a non-cooperative game theory approach where autonomous racing is formulated as racing decisions as a non-cooperative nonzero-sum game. Liniger et al. displayed that different games can be modelled that achieve successfully different racing behaviors and generate interesting racing situations e.g. blocking and overtaking. Notomista et al. [9] considered a two-player racing game where the ego vehicle is based on a Sensitivity-Enhanced Nash equilibrium seeking (SENNA) method, which uses an iterated best response algorithm in order to optimize for a trajectory in a two-car racing game. Focusing on collision avoidance, the SENNA method exploits interactions between the ego and the opponent vehicle. Unfortunately this approach was only displayed in an unknown simulation environment and therefore its performance is unknown. Wang et. al [13] proposed a nonlinear receding horizon game-theoretic planner for autonomous cars in competitive scenarios with other cars. By implementing the trajectory as a piecewise polynomial and incorporating bicycle kinematics into the trajectory the authors were able to show that the ego vehicle exhibits rich game strategies such as blocking, faking, and opportunistic overtaking maneuvers.

The state of the art displays that the autonomous racing community is focusing on integrating effective learning techniques and strategies into dynamic path and behavioral planning. Additionally, authors have displayed individual algorithms and methods that are trying to make the car faster, more reliable and more interactive [11] [5] [6]. Majority of the literature is centered around extensions to existing path and behavioral planning approaches. Improvements in planning/control for overtaking maneuvers have not yet been explored rigorously. Techniques to optimize path planning and

control algorithms to ensure a better overtaking maneuver have to be addressed.

IV. PLANNING & CONTROL SETUP

The continuous time system dynamics described in Section VII-A is used to develop a constrained optimal controller to steer the vehicle in the track. The optimal planner plans the path for a horizon of N steps ahead, steers the vehicle with the first step, and again repeats the process for the specified amount of time. This is a modified form of the Model Predictive Controller (MPC).

A. Model Predictive Contouring Control

The Model Predictive Contouring Control (MPCC) problem defined in [7] is re-formulated into a finite-continuous time optimal control problem as follows:

$$\begin{aligned} \min \quad & \int_0^T [\epsilon_c^{lin}(t) \quad \epsilon_l^{lin}(t)] \begin{bmatrix} Q_c & 0 \\ 0 & Q_l \end{bmatrix} \begin{bmatrix} \epsilon_c^{lin}(t) \\ \epsilon_l^{lin}(t) \end{bmatrix} \\ & - Q_\theta \dot{\theta}(t) + u^T(t) R u(t) dt \\ \text{s.t.} \quad & \dot{x} = f(x, u, \Phi) \\ & b_{lower} \preceq x(t) \preceq b_{upper} \\ & l_{lower} \preceq u(t) \preceq l_{upper} \\ & h(x, \Phi) \leq 0 \end{aligned}$$

given the system dynamics f and the arclength parametrization of the contour (the track) Φ . A single-track bicycle model is used. Section VII-A contains the details of the vehicle dynamics. Here $x(t)$ denotes the system state, $u(t)$ the inputs to the system, b the box constraints on the state, l the box constraints on the input and h captures the track boundary constraints. The state of the system is augmented with the advancing parameter θ

$$x = \begin{bmatrix} x_{model} \\ \theta \end{bmatrix}$$

and the virtual input $\dot{\theta}$ is appended to the inputs from the original system dynamics.

$$u = \begin{bmatrix} u_{model} \\ \dot{\theta} \end{bmatrix}$$

The track boundary constraint is realized as a convex disk constraint.

$$h(x, \Phi) = (x - x_t^{lin}(\theta))^2 + (y - y_t^{lin}(\theta))^2 - r_\Phi(\hat{\theta})^2$$

Here $r_\Phi(\hat{\theta})$ is the half-width of the track at the last predicted arc length.

The linearized contouring error ϵ_c^{lin} and lag error ϵ_l^{lin} are computed as shown in Figure 3. To make the problem real-time feasible they are approximated by linearizing both them and the track around the previous solution θ as:

$$\begin{aligned} \Phi(\theta) &= \begin{bmatrix} x_t(\theta) \\ y_t(\theta) \end{bmatrix} \approx \Phi(\hat{\theta}) + \partial_\theta \Phi(\hat{\theta})(\theta - \hat{\theta}) \\ \Rightarrow \Phi^{lin}(\theta) &= \begin{bmatrix} x_t(\hat{\theta}) + \cos(\phi(\hat{\theta}))(\theta - \hat{\theta}) \\ y_t(\hat{\theta}) + \sin(\phi(\hat{\theta}))(\theta - \hat{\theta}) \end{bmatrix} \end{aligned}$$

this allows us to compute the errors

$$\begin{aligned}
x_t^{lin}(\theta) &= x(t) + \epsilon_l^{lin} \cos(\phi(\hat{\theta})) + \epsilon_c^{lin} \sin(\phi(\hat{\theta})) y_t^{lin}(\theta) \\
&= y(t) + \epsilon_l^{lin} \sin(\phi(\hat{\theta})) - \epsilon_c^{lin} \cos(\phi(\hat{\theta})) \Leftrightarrow \\
\begin{cases} \epsilon_l^{lin} \\ \epsilon_c^{lin} \end{cases} &= \begin{cases} \cos(\phi(\hat{\theta})) (x_t^{lin}(\theta) - x(t)) + \sin(\phi(\hat{\theta})) (y_t^{lin}(\theta) - y(t)) \\ \sin(\phi(\hat{\theta})) (x_t^{lin}(\theta) - x(t)) - \cos(\phi(\hat{\theta})) (y_t^{lin}(\theta) - y(t)) \end{cases}
\end{aligned}$$

These approximations hold good if $\epsilon_l \approx 0$ and $\hat{\theta} - \theta \approx 0$. In practice the first can be incentivized by increasing Q_l and the second by warmstarting the problem correctly.

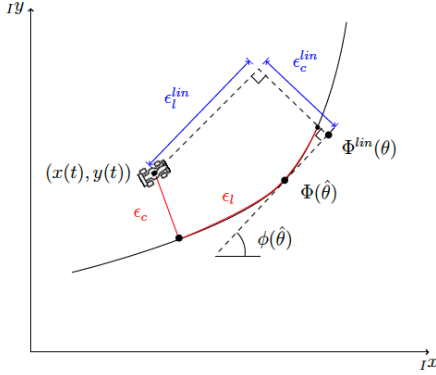


Fig. 3. Linearized contouring and lag errors

The MPCC is optimizing to move the position of a virtual point $\theta(t)$ along the track to achieve as much progress as possible while steering the model of the vehicle to keep contouring and lag errors as small as possible.

B. Switched Model Predictive Contouring Control

The MPCC controller defined above ensures that the vehicle stays inside the track, drives fast and avoids obstacles. The track and obstacle constraints are realised as non-linear convex disk constraints. However, such a setup does not give enough control over achieving different driving behaviours like driving on the inside of a chicane (left side of the track), driving on the outside of a sweeper (right side of the track). To achieve more control over the path planning of the ego vehicle, a switched MPCC setup is proposed and displayed in figure 4.

In this, the agent switches between different modes defined by different solver formulations. The overtaking possibilities for the MPCC in this paper are:

- **Normal Mode:** In this, the search space for the N-horizon MPCC is the complete racetrack
- **Drive Right:** In this, the search space of the N-horizon MPCC is restricted only right half of the track.
- **Drive Left:** In this, the search space of the N-horizon MPCC is restricted only left half of the track.

The "Drive Right" and "Drive Left" modes are achieved by imposing additional constraints on the contouring error of

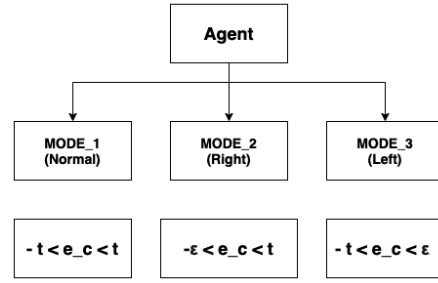


Fig. 4. Switched MPCC Controller

the vehicle. The constraints of the optimisation problem are tweaked with slack variables (like ϵ) to ensure that the planner does not get stuck into an in-feasibility loop leading to a crash. Note that many more of these modes can be defined and added to the MPCC. These modes can then integrate practical driving behaviours like ADAS mode (making tighter bounds on the problem and not letting the vehicle drive at its limits) and Push-to-Pass (giving a speed boost for a short window of time).

The MPCC control problem with a horizon of $N = 35$ is modeled as a quadratic problem (QP) with linear and non linear inequalities solved by efficient interior point solvers in FORCES [15]. At each step of the simulation, the solver generates the optimal control input u^* for a horizon of N . The first value $u^*(0)$ is used to steer the vehicle and the procedure is repeated.

C. Track Parametrization

The center-line of the track is given in way-points (X-and Y-position). To implement MPCC an arc-length parametrization Φ is required. This is realized by interpolating the way-points using cubic splines with a cyclic boundary condition, and creating a dense lookup table with the track location and the linearization parameters. Note that in the optimization it is not practical to pass the full parametrization since it contains functions that the used solver has difficulties dealing with such as floor and modulo. Instead the linearization parameters are precomputed offline and passed at every stage.

V. OFFLINE POLICY LEARNING FOR OVERTAKING

Algorithm 1 elucidates the offline experiment based policy learning developed in this paper.

\mathcal{X} , \mathcal{Y} are the set of x and y coordinate offsets (expressed as percentage of track width) and $\mathcal{S} = \{-0.2, 0, +0.2\}$ is the speed offset expressed as a percentage change from the baseline obstacle speed.

The obstacle update model is g , which is a pre-computed curvature optimal race line of the Silverstone track. One run of the algorithm populates the policy map Π with the track regions for each of the 8 curves, having highest probability of overtakes. In total we are running a number of 576 experiments based on the 16 lateral, 12 longitudinal and 3 velocity variations for each of the 8 track portions.

Algorithm 1: Offline Policy Learning

```

Function MPCC Planner( $X_{obs}$ ):
    Solve MPCC Problem defined in Section IV
    return  $u^*$ ;
Function Check Overtake( $X_{obs}, X_{ego}$ ):
    Project  $X_{obs}, X_{ego}$  as  $s_1, s_2$  on track
    if  $s_1 > s_2$  then
        overtake_status = 0;
    else
        overtake_status = 1;
    end
    return overtake_status;
Function Main():
    initialize:  $\Pi = \{\}$ 
    for  $\tau \in \mathbf{T}$  do
        initialize:  $p = \{\}$ , overtakes =  $\{\}$ , total =  $\{\}$ 
        for  $x, y, s \in \mathcal{X} \mathcal{X} \mathcal{Y} \mathcal{X} \mathcal{S}$  do
            initialize:  $X_{ego}, X_{obs}$ 
            for  $t = 0$  to  $T_{sim}$  do
                 $u^* = \text{MPCC Planner}(X_{obs})$ 
                Steer the ego:  $X_{ego}^+ = f(X_{ego}, u^*)$ 
                Update the obstacle position:
                 $X_{obs}^+ = g(X_{obs})$ 
                 $X_{obs}, X_{ego} = X_{obs}^+, X_{ego}^+$ 
                Identify track portion  $i \in \{0, 1, 2, 3\}$ 
                if Check Overtake( $X_{obs}, X_{ego}$ ) then
                    overtakes[ $\mathcal{R}_i$ ] ++;
                    total[ $\mathcal{R}_i$ ] ++;
                end
            end
             $p[\mathcal{R}_i] = \text{overtakes}[\mathcal{R}_i] / \text{total}[\mathcal{R}_i]$ 
            Compute:  $p[\mathcal{R}_1], p[\mathcal{R}_2], p[\mathcal{R}_3], p[\mathcal{R}_4]$ 
             $\Pi[\tau] = \text{argmax}(p)$ 
        end
    end

```

A. Offline Experiment Results

The offline policy learning algorithm is run on the described experimental setup. Based on the outcome of the algorithms different results are shown. Firstly we display the results for one specific track portions.

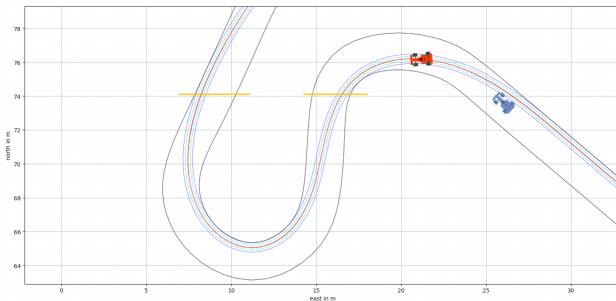


Fig. 5. Hairpin Curve - Overtake analysis

We choose the hairpin (track portion label 3) which has a predefined overtaking window (marked in yellow). The

obstacle vehicle (red) is tracking the optimal race line and the ego vehicle (blue) starts behind the obstacle (5)

The ego vehicle is said to overtake the obstacle if it is ahead of the obstacle before the overtaking corridor ends. The overtaking corridor defines the portion of the track in which a successful overtake must happen. If the ego vehicle passes the obstacle vehicle outside of the corridor, it will not count as a successful overtake. Similar overtaking corridors are defined for all 8 track portions. As a results we get for each experiment the feedback if the overtaking maneuver was successful (figure 6).

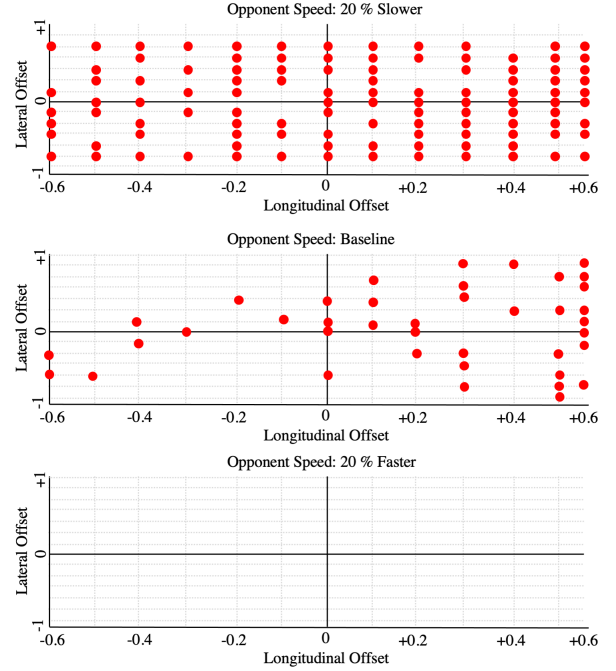


Fig. 6. Display of successful overtaking maneuvers over different positions across different speed offsets for the hairpin (track portion 3)

The Algorithm 1 is run for the track portion and the overtaking probability distributions are obtained. With this, the policy for this curve is also obtained. Table I below shows the probability distribution of overtakes at this curve.

TABLE I
OVERTAKING PROBABILITIES FOR HAIRPIN

Region	Probability
\mathcal{R}_1	0.41
\mathcal{R}_2	0.38
\mathcal{R}_3	0.25
\mathcal{R}_4	0.21

To use this probability later for our switching MPCC algorithm we define these four regions in latent space on the track as displayed in figure 7.

The distribution of overtakes with respect to change in obstacle speed is shown in Figure 6. It can be observed that the number of overtakes is highest if the obstacle speed is 20% lesser than its baseline. However, when there is no speed

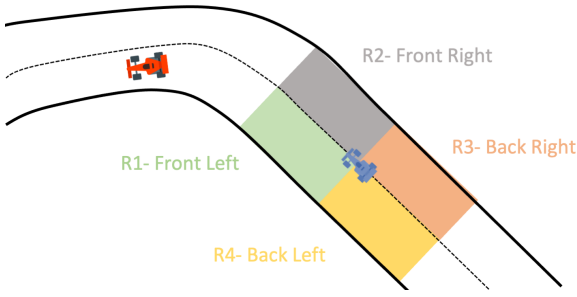


Fig. 7. Hairpin curve split in four track regions of interest

advantage (baseline scenario), we can observe that the region \mathcal{R}_1 has higher overtakes than any other region. Therefore, the policy for this track portion would ensure that the ego vehicle starts at \mathcal{R}_1 before entering the hairpin which gives it a positional advantage and increases its overtaking probability. This process is repeated for all track portions and the policies are learnt from the offline experiments.

Besides the case study on the hairpin Table II contains the overtaking probabilities for all the eight derived track portions. The policy map is developed using the results from the experiments.

TABLE II
OVERTAKING PROBABILITIES FOR ALL TRACK PORTIONS

Track Portion (τ)	Type	$p(\mathcal{R}_1)$	$p(\mathcal{R}_2)$	$p(\mathcal{R}_3)$	$p(\mathcal{R}_4)$
1	Sweeper	0.99	0.93	0.52	0.59
2	Hairpin	0.63	0.52	0.31	0.33
3	Hairpin	0.41	0.38	0.25	0.21
4	Sweeper	0.65	0.67	0.57	0.59
5	Chicane	0.25	0.21	0.14	0.21
6	Straight	0.99	1.0	0.95	0.99
7	Sweeper	0.47	0.52	0.33	0.31
8	Hairpin	0.40	0.36	0.37	0.38

From the statistical data obtained from offline experiments, we have the following observations and outcomes:

- At the straightaway, it does not really matter where we are on the track when trying to overtake. Track region 3 (position: back right) is the one region with the lowest overtaking probability. This is due to the cause that the on the straight the optimal raceline leads from left to right across the straight. Starting in the back right would lead to an braking and left overtaking maneuver which is causing problems for turn 7 afterwards.
- We can see that in each hairpin we have the highest overtaking probability in region 1 (position: front left). This is due to the fact that being on the inside in the hairpin that car is able to achieve a better trajectory inside the hairpin.

- The sweeper curve generally has a high overtaking probability due to the high speeds of the car. We only get an advantage here if we are close enough to the car and therefore we need to be in region $p(\mathcal{R}_1)$ or $p(\mathcal{R}_2)$. In both track portions 4 and 7 we see a higher overtaking maneuver from starting in the front right track regions.
- The chicane has generally a low overtaking probability due to the fact that it is a complex region to handle for the car and with little space for an overtaking maneuver. We see a highest overtaking probability here on the front left region and in contrary the lowest overtaking probability in the back right region. Both front right and back left region show the same overtaking probability and therefore being far behind the car lowers here again the overtaking possibility of the ego car.

B. Online switching MPCC Results

Firstly we define the algorithm for the switching MPCC:

Algorithm 2: Policy based Switched MPCC

```

Function Switched MPCC ( $X_{obs}, mode$ ):
    if  $mode = 'normal'$  then
        |  $u^* = \text{Solve MPCC Problem in Section IV}$ 
    else
        | Modify MPCC to integrate policy (left,right)
        |  $u^* = \text{Solve Modified MPCC Problem}$ 
    end
    return  $u^*$ ;
initialize  $X_{ego}, X_{obs}$   $mode = 'normal'$ 
for  $t = 0$  to  $T_{sim}$  do
     $u^* = \text{Switched MPCC} (X_{obs}, mode)$ 
    Steer the ego:  $X_{ego}^+ = f(X_{ego}, u^*)$ 
    Update the obstacle position:  $X_{obs}^+ = g(X_{obs})$ 
     $X_{obs}, X_{ego} = X_{obs}^+, X_{ego}^+$ 
    Identify track portion  $\tau$  where ego is present
    Policy lookup:  $mode = \Pi[\tau]$ 
end

```

With the results of the track region probability we now have the possibility to use the policies obtained from algorithm 1 in our switching MPCC algorithm. We use this information to make the car move to the regions of high overtaking probability as displayed before

We repeat the experiments at all track portions with the learnt policies and compute the number of successful overtakes. Table III summarizes the number of overtakes before and after the policy is integrated into the ego vehicle's MPCC planner. For every track portion, there are 576 trajectories owing to $\mathcal{X} = \{-0.8, -0.7, -0.6, -0.5, -0.4, -0.3, -0.2, -0.1, 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8\}$, $\mathcal{Y} = \{-0.6, -0.5, -0.4, -0.3, -0.2, -0.1, 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6\}$ and $\mathcal{S} = \{-0.2, 0, 0.2\}$.

The results display that the offline policy learning approach is successful and leads to an increased number of overtaking maneuvers when implemented in a path planning approach.

TABLE III
NUMBER OF OVERTAKES WITH AND WITHOUT POLICY

Track Portion (τ)	Track Portion Type	Number of Overtakes Policy OFF	Number of Overtakes Policy ON
1	Sweeper	436	452
2	Hairpin	256	337
3	Hairpin	308	426
4	Sweeper	342	357
5	Chicane	117	302
6	Straight	565	566
7	Sweeper	237	283
8	Hairpin	218	394

For all eight track portions we see more overtaking maneuvers. We can observe that overtaking at the straightaway is usually easy (even without policy) and therefore we don not see many more overtaking maneuvers with switching policy. The same is for the sweeper curve. Since sweeper curves are usually wide track portions that allow high speeds and not complicated steering maneuver, both with and without switching policy achieve high maneuvers. Although we see that the switching policy leads to more overtaking maneuvers.

We see the biggest impact in both the hairpins and the chicane. This is mainly due to the fact, that overtaking at these hairpins is usually complicated and needs a good strategy beforehand. We see that we can nearly triple the amount of overtaking maneuvers in the chicane which shows, that having the right starting position for an overtaking maneuver is indispensable. The best results in this experiment is achieved in track portion 8 (hairpin) where a total of 186 more overtakes are done with the MPCC switching.

In summary we see that the gains for the overtaking maneuver are less for straight sections such as straightaway and sweeper curve and better for curved sections such as the hairpin and the chicane.

VI. CONCLUSION AND FUTURE WORK

In this paper, a track based policy learning from offline experiments is proposed to learn effective overtaking strategies based on position advantage at different track portions. A switched Model Predictive Contouring Control (MPCC) scheme was proposed to integrate driving behaviours/policies into the motion planning and control of the vehicle. Extensive simulations on real world racetrack layout with a naive, non-interactive obstacle shows that the offline policy learning algorithm is able to provide areas for high overtaking maneuvers. By integrating these areas in a switching MPCC method we could show that the policy based Switching MPCC approach has more overtakes than the regular MPCC planner. Some track portions like straight have not shown considerable increase in the number of overtakes, as the overtaking probabilities are already highest (around 0.99) in a straightway. However, the policy based algorithm was found to be highly effective for convoluted track portions like chicanes, where a positional advantage plays a major role in successful overtaking maneuvers.

The offline policy learning shows promising results for a naive opponent following a pre-computed raceline with no interactions. Future directions of research include extending the policy learning algorithm to race against a variety of opponent types. We consider: non-reactive, reactive and aggressive opponents which are defensive and sophisticated to overtake. Other dimensions of research include the evaluating the policy based algorithm across different tracks and study the effect of changing track parameters on track-policy based overtaking.

REFERENCES

- [1] Johannes Betz et al. “A Software Architecture for the Dynamic Path Planning of an Autonomous Racecar at the Limits of Handling”. In: *2019 IEEE International Conference on Connected Vehicles and Expo (ICCVE)*. IEEE, Nov. 2019. DOI: 10.1109/iccve45908.2019.8965238.
- [2] Alexander Buyval et al. “Deriving overtaking strategy from nonlinear model predictive control for a race car”. In: *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, Sept. 2017. DOI: 10.1109/iros.2017.8206086.
- [3] Danio Caporale et al. “A Planning and Control System for Self-Driving Racing Vehicles”. In: *2018 IEEE 4th International Forum on Research and Technology for Society and Industry (RTSI)*. IEEE, Sept. 2018. DOI: 10.1109/rtsi.2018.8548444.
- [4] Alexander Heilmeyer et al. “Minimum curvature trajectory planning and control for an autonomous race car”. In: *Vehicle System Dynamics* 58.10 (June 2019), pp. 1497–1527. DOI: 10.1080/00423114.2019.1631455. URL: <https://doi.org/10.1080/00423114.2019.1631455>.
- [5] Juraj Kabzan et al. “Learning-Based Model Predictive Control for Autonomous Racing”. In: *IEEE Robotics and Automation Letters* 4.4 (Oct. 2019), pp. 3363–3370. DOI: 10.1109/lra.2019.2926677.
- [6] Nitin R. Kapania and J. Christian Gerdes. “Learning at the Racetrack: Data-Driven Methods to Improve Racing Performance Over Multiple Laps”. In: *IEEE Transactions on Vehicular Technology* 69.8 (Aug. 2020), pp. 8232–8242. DOI: 10.1109/tvt.2020.2998065.
- [7] Alexander Liniger, Alexander Domahidi, and Manfred Morari. “Optimization-based autonomous racing of 1: 43 scale RC cars”. In: *Optimal Control Applications and Methods* 36.5 (2015), pp. 628–647.
- [8] Alexander Liniger and John Lygeros. “A Noncooperative Game Approach to Autonomous Racing”. In: *IEEE Transactions on Control Systems Technology* 28.3 (May 2020), pp. 884–897. DOI: 10.1109/tcst.2019.2895282.
- [9] Gennaro Notomista et al. “Enhancing Game-Theoretic Autonomous Car Racing Using Control Barrier Functions”. In: *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, May 2020. DOI: 10.1109/icra40945.2020.9196757.

- [10] Matthew O’Kelly et al. “FITENTH: An Open-source Evaluation Environment for Continuous Control and Reinforcement Learning”. In: *NeurIPS 2019 Competition and Demonstration Track*. PMLR. 2020, pp. 77–89.
- [11] Ugo Rosolia and Francesco Borrelli. “Learning How to Autonomously Race a Car: A Predictive Control Approach”. In: *IEEE Transactions on Control Systems Technology* 28.6 (Nov. 2020), pp. 2713–2719. DOI: 10.1109/tcst.2019.2948135.
- [12] Wilko Schwarting et al. “Deep Latent Competition: Learning to Race Using Visual Control Policies in Latent Space”. In: *Conference on Robot Learning (CoRL) 2020*. 2020.
- [13] Mingyu Wang et al. “Game-Theoretic Planning for Self-Driving Cars in Multivehicle Competitive Scenarios”. In: *IEEE Transactions on Robotics* (2021).
- [14] Trent Weiss and Madhur Behl. “DeepRacing: A Framework for Autonomous Racing”. In: *2020 Design, Automation & Test in Europe Conference & Exhibition (DATE)*. IEEE, Mar. 2020. DOI: 10.23919/date48585.2020.9116486.
- [15] A. Zanelli et al. “FORCES NLP: an efficient implementation of interior-point methods for multistage nonlinear nonconvex programs”. In: *International Journal of Control* 93.1 (May 2017), pp. 13–29.

VII. APPENDIX

A. Vehicle Model

The simulation is built around the dynamic model used in [7] and is implemented with adapted parameters. It is a single track bicycle model with pajecka tire forces. The single-track

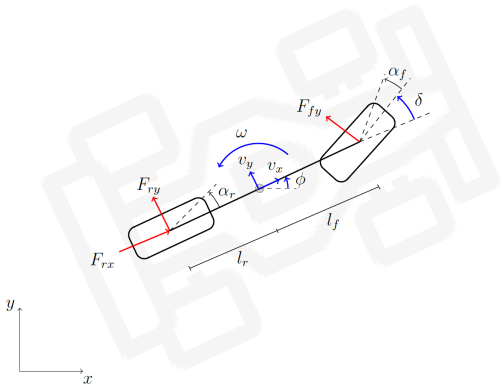


Fig. 8. Single track model with pajecka tire forces

bicycle model is used for modelling the race car (ego-vehicle). The model is based on in-plane motions and pitch, roll as well as load-changes are neglected. Due to the rare-wheel drive nature and absence of active braking, longitudinal forces on the front wheel are also neglected.

The dynamics of this model read:

$$\dot{X} = v_x \cos(\phi) - v_y \sin(\phi)$$

$$\dot{Y} = v_x \sin(\phi) + v_y \cos(\phi)$$

$$\dot{\phi} = \omega$$

$$\dot{v}_x = \frac{1}{m} (F_{rx} - F_{fy} \sin(\delta) + m v_y \omega)$$

$$\dot{v}_y = \frac{1}{m} (F_{ry} + F_{fy} \cos(\delta) - m v_x \omega)$$

$$\dot{\omega} = \frac{1}{I_z} (F_{fy} l_f \cos(\delta) - F_{ry} l_r)$$

$$\dot{d} = u_d$$

$$\dot{\delta} = u_\delta$$

$$\dot{\theta} = u_\theta$$

where

$$\alpha_f = -\arctan\left(\frac{\omega l_f + v_y}{v_x}\right) + \delta$$

$$\alpha_r = \arctan\left(\frac{\omega l_r - v_y}{v_x}\right)$$

$$F_{fy} = D_f \sin(C_f \arctan(B_f \alpha_f))$$

$$F_{ry} = D_r \sin(C_r \arctan(B_r \alpha_r))$$

$$F_{rx} = (C_{m1} - C_{m2} v_x) d - C_{roll} - C_d v_x^2$$

$X \in \mathcal{X} : X = [x, y, \phi, v_x, v_y, \omega, d, \delta, \theta]$ is the state of the car. x, y are position of the centre of gravity, ϕ is the yaw angle of the vehicle with reference to world frame, v_x, v_y are longitudinal and lateral velocities and ω is the yaw rate.

The parameter θ is known as the advancing parameter (which is augmented to form the complete state vector), d is the integrated motor torque and δ is the steering angle.

$u \in \mathcal{U} : u = [u_d, u_\delta, u_\theta]$ is the input to the model. These inputs are the derivatives of the commandable inputs to the vehicle chosen so as to penalise their smoothness.

$\mathcal{X} \subset \mathbf{R}^9$ and $\mathcal{U} \subset \mathbf{R}^3$ are the set of admissible states and control inputs for the vehicle defined by the optimal control problem in Section IV