C2SMART

CONNECTED CITIES WITH SMART TRANSPORTATION

A USDOT University Transportation Center

New York University

Rutgers University

University of Washington

The University of Texas at El Paso

City College of New York

Lane Changing of Autonomous Vehicles in Mixed Traffic Environments: A Reinforcement Learning Approach

February 2022

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2.Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Lane Changing of Autonomous Vehicles in Mixed Traffic Environments: A Reinforcement Learning Approach		5. Report Date February 2022 6. Performing Organization Code
7. Author(s) Zhong-Ping Jiang, Kaan Ozbay, Sayan Chakraborty, Leilei Cui		8. Performing Organization Report No.
9. Performing Organization Name and Address Connected Cities for Smart Mobility towards Accessible and Resilient Transportation Center (C2SMART), 6 Metrotech Center, 4th Floor, NYU Tandon School of Engineering, Brooklyn, NY, 11201, United States		10. Work Unit No. 11. Contract or Grant No. 69A3551747119
12. Sponsoring Agency Name and Address Office of the Assistant Secretary for Research and Technology University Transportation Centers Program U.S. Department of Transportation Washington, DC 20590		13. Type of Report and Period Covered Final Report, 3/1/2021-2/28/2022 14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

The purpose of this proposal is to develop innovative reinforcement learning control methods for lane changing of connected and autonomous vehicles (CAVs) in mixed traffic. In the proposed framework, before the CAV changes to the target lane, it needs to predict most likely behavior of surrounding vehicles related to the lane change and then determine the optimal path to the other lane. The lane changing maneuver will be completed by solving a linear quadratic regulator (LQR) problem that yields an optimal controller to monitor the CAV's longitudinal and lateral movements during the lane-changing maneuver. A novel aspect of this research is to reduce the trajectory planning and tracking problem down to the minimization of a cost function that depends on the target way-point in the target lane the CAV will reach. In the proposal, the research team will integrate reinforcement learning and adaptive/approximate dynamic programming methods to solve this data-driven LQR control problem under constraints, without assuming the exact knowledge of surrounding vehicles, while avoiding the curses of dimensionality and modeling of conventional dynamic programming. Thanks to the systematic use of systems and control- theoretic methods, the proposed framework aims to yield desirable lane- changing controllers with guaranteed stability for CAVs from small samples of historical and real-time data.

17. Key Words	18. Distribution Statement Public Access		
19. Security Classif (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No of Pages 22	22. Price

Form DOT F 1700.7 (8-69)

Lane Changing of Autonomous Vehicles in Mixed Traffic Environments: A Reinforcement Learning Approach

Zhong-Ping Jiang
New York University
0000-0002-4868-9359

Kaan Ozbay
New York University
0000-0001-7909-6532

Leilei Cui
New York University
0000-0001-8031-7638

Sayan Chakraborty
New York University
0000-0002-8638-4652

C2SMART Center is a USDOT Tier 1 University Transportation Center taking on some of today's most pressing urban mobility challenges. Some of the areas C2SMART focuses on include:

Urban Mobility and Connected Citizens

Urban Analytics for Smart Cities

Resilient, Smart, & Secure Infrastructure

Disruptive Technologies and their impacts on transportation systems. Our aim is to develop innovative solutions to accelerate technology transfer from the research phase to the real world.

Unconventional Big Data Applications from field tests and non-traditional sensing technologies for decision-makers to address a wide range of urban mobility problems with the best information available.

Impactful Engagement overcoming institutional barriers to innovation to hear and meet the needs of city and state stakeholders, including government agencies, policy makers, the private sector, non-profit organizations, and entrepreneurs.

Forward-thinking Training and Development dedicated to training the workforce of tomorrow to deal with new mobility problems in ways that are not covered in existing transportation curricula.

Led by New York University's Tandon School of Engineering, **C2SMART** is a consortium of leading research universities, including Rutgers University, University of Washington, the University of Texas at El Paso, and The City College of NY.

Visit c2smart.engineering.nyu.edu to learn more

C2SMART
CONNECTED CITIES WITH SMART TRANSPORTATION

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

Acknowledgements

This project is supported in part by the USDOT Tier-1 C2SMART Center grant, in part by NYU SUE PhD Fellowship awarded to the student, and in part by the US NSF.

Executive Summary

Autonomous vehicle (AV) and connected vehicle (CV) technologies have been much of the focus of transportation industry lately. In this project, we have focused on the lane changing problem of autonomous vehicles. Inappropriate lane changes due to the inefficiency of human drivers in prediction and estimation of the surrounding environment often lead to accidents. In this project, we have developed methodologies for efficient lane changing of autonomous vehicles (AV) in mixed traffic conditions. The proposed methodology is focused on improving the safety and comfort of the passengers traveling in an AV.

Firstly, we have introduced an optimal data-driven control algorithm to solve the lane changing problem of AVs. We arrive at overcoming the limitations of existing model-based approaches for lane changing. The developed data-driven algorithms can achieve both theoretically provable convergence guarantees and learning-based adaptive responsiveness to environments. In the proposed methodology we make use of the online information of the state and input to solve the algebraic Riccati equation iteratively by using approximate/adaptive dynamic programming (ADP). A linear system model is used to describe the dynamics of the AV.

Secondly, we have developed a lane change decision making algorithm to ensure safe and efficient lane change. Safety is assured during lane changing by maintaining a safe distance from the surrounding vehicles. The safe distance from each surrounding vehicle is chosen as a function of their respective velocities. Thus, ensuring more safety for fast moving vehicles. The purposed lane change decision making algorithm can make the AV abort any initiated lane change maneuver at any time if safety conditions are not met. In such scenarios, the proposed decision-making algorithm makes the AV maneuver back to the original lane.

Thirdly, the linear vehicle model assumes constant longitudinal velocity of the vehicle which is not a practical consideration. Thus, we have considered that the AV can accelerate/decelerate during lane change. To incorporate this, we have proposed a novel data-driven gain-scheduling controller design that learns the controller gains at specified velocity points such that each of the learned controller gain is optimal.

Finally, the optimal data-driven control algorithm and the lane change decision making algorithm have been validated by means of SUMO and MATLAB based computer simulations.

Table of Contents

List of Figures

Background and Contribution

Inappropriate lane changes are responsible for one tenth of all accidents [1], due to human drivers' inaccurate estimation and predication of the surrounding traffic, illegal maneuver, and inefficient driving skill. Automated lane changing is regarded as a solution to reduce these human errors. At present, there are many obstacles to develop automated lane changing technology, including interactions between vehicles, complex routing choice, and interactions between vehicles and the environment. In this work, in order to improve safety and comfort, we aim to develop innovative model-free control methods for lane changing of autonomous vehicles (AVs) in mixed traffic consisting of both autonomous vehicle (AV) and human-driven vehicles (HDVs) based on reinforcement learning and optimal control techniques. The key components of lane changing are: 1) V2V communication, 2) sensing, 3) lane-change decision making and path planning, and 4) vehicle control algorithms [2].

Over the past few years, many path planning and lane changing methodologies are proposed in the literature. The main aim of these methodologies is to first generate smooth trajectories using safety constraints. The safety constraints which are a part of the lane-change decision making module, are obtained by using the state (position, velocity, acceleration) information of the surrounding vehicles through V2V communication and sensing. Once the safe trajectories are generated, a controller is applied to track the generated trajectories. The authors in [3] proposed a utility function-based lane change and merge technique. The utility function considers the discretionary, anticipatory, and mandatory conditions to judge the desirability of the AV to change lane. Once lane change is deemed desirable, a safe longitudinal and lateral safety corridor is determined to perform the maneuver by selecting an appropriate inter-vehicle traffic gap and time instance. Most of the existing studies on lane changing have assumed that a vehicle in the target lane must decelerate to make space for the CAV due to safety considerations. This might cause traffic disruptions and increase the traffic crash rates. To deal with such an issue, [4] proposes a real-time dynamic cooperative lane-changing model for CAVs with possible accelerations of a preceding vehicle. The lane change decision is based on upper and lower bounds of acceleration and deceleration of the preceding and following vehicles in the target lane respectively. The lane change path is then generated using a cubic polynomial. In [6] an optimization-based lane change methodology is proposed. The objective function is minimized for the longitudinal and lateral jerks, and total distance of lane change. Thus, the objective function is formulated keeping in mind the passenger comfort and traffic efficiency. The safety conditions are considered as constraints of the optimization problem. It is usually assumed that the vehicles neighboring AV are also AVs. This assumption does simplify the lane change methodology to an extent, but are ideal and quite ahead of the time we live in. A more practical scenario is considered in [5] where both human-driven vehicles and CAVs interact for lane change maneuvers. In [7] a dynamic lane-changing model for Autonomous Vehicle (AV) incorporating human driver behavior in mixed traffic is considered. The authors have implemented a model-predictive-control- (MPC-) based joint trajectory control. Field experiments are conducted on a large-scale test track to test and validate the proposed model. Different human driver behaviors were considered in the experimental settings. After comparing with the measured human lane-changing maneuvers, it was found that the AV lane-changing maneuvers from the proposed model are more comfortable and safer.

Once the lane change decision making is done, the next task is to place the AV to the desired position/gap in the desired lane. Many control techniques have been proposed in the past to maneuver the AV to the desired lane while ensuring safety [3]-[13]. In [3], the quadratic programming (QP) has been used to compute the control signals for lateral and longitudinal maneuver, where a double integrator model is used for the AV dynamics. In [4], a linearized vehicle kinematic model with inertia delay τ is used to obtain the safety conditions for lane changing. The lateral trajectory of the AV is obtained using a cubic polynomial, where the linearized vehicle kinematic model was used to obtain the coefficients of the polynomial. In [5] and [7], a model predictive control (MPC) based method is used for the vehicle control which uses the two-wheel kinematic vehicle model, where the two front (or rear) wheels are considered as one wheel. The authors in [6], propose a trajectory-tracking controller based on sliding mode control. The tracking controller is based on the backstepping approach. It can ensure global convergence during the lane-changing process. However, this trajectory-tracking model only considers current tracking errors. It may not handle a high-dynamic motion environment. In [8], a cooperative lane changing methodology is proposed where a decentralized cooperative lane-changing decision-making framework for CAV composed of state prediction, candidate decision generation, and coordination. The state prediction module employs cooperative car-following models to predict the vehicles' future state. In [9], a hierarchical, two-level architecture is employed for the trajectory generation and vehicle control of AV. The high-level planner utilizes a simplified point-mass model and linear collision avoidance constraints, whereas, the low-level controller utilizes a nonlinear vehicle model in order to compute the vehicle control inputs required to execute the planned maneuvers. Both the high-level planner and low-level controller are formulated based on the model predictive control methodology. In [12], the authors formulate a stochastic MPC controller. The MPC controller can predict future states and implement constraints directly into the control algorithm. The proposed algorithm uses a linear parameter varying (LPV) vehicle model.

As evident from the literature, most of the works done to solve the lane change problem in AVs use model-based techniques. The problem with these techniques lies in incorporating environment uncertainties in their methodology that can be introduced by the adversarial situations. Many of the methodologies mentioned above requires solving an optimization problem in real time in order to generate/track safe trajectories for the AV lane change maneuver which requires high computation effort. Because of these observations, we believe that learning-based optimal control is more desirable for practical implementation, which can continually handle the environment uncertainty introduced by the unknown environment-dependent parameters and simultaneously optimize the performance of the AV lane change maneuver by learning from the real-time data.

This report adopts ideas from reinforcement learning [13, 14] and adaptive dynamic programming (ADP) [15] to develop an intelligent and safe lane change maneuver algorithm for AVs in the mixed traffic scenario. By systematic use of control theory, ADP has proven to be a powerful method to learn safe and stable controllers by using real-time data collected along the trajectories of the controlled system. One major advantage of ADP, as opposed to traditional reinforcement learning [13], lies in the fact that the closed-loop stability of the dynamical system is established when the learned control policy is implemented. Meanwhile, the stability/robustness of the CAV controller characterizes the convergence of the AV's lateral position to a desired equilibrium. The main contributions of this report are summarized as follows.

1) The proposed learning-based control algorithm for the AV is implemented in SUMO where the learning is done online during SUMO simulation.

2) We proposed a lane change decision making algorithm to ensure safe and efficient lane change. Safety is assured during lane changing by maintaining a safe distance from the surrounding vehicles. The safe distance from each surrounding vehicle is chosen as a function of their respective velocities. Thus, ensuring more safety for fast moving vehicles.

3) The purposed lane change decision making algorithm can make the AV abort any initiated lane change maneuver at any time if safety conditions are not met. In such scenarios, the proposed decision-making algorithm lets the AV maneuver towards the original lane.

4) We assume that we receive real-time data from a linear system. Thus, to make our methodology more applicable to practical scenarios, we have proposed a gain-scheduled learning-based controller to handle non-linearities.

Dynamic Model and Problem Formulation

Longitudinal Dynamic Model

The vehicle's longitudinal dynamic model is given as follows:

$$\dot{x}_{lo} = A_{lo}x_{lo} + B_{lo}u_{lo}, \qquad (1)$$

where

$$A_{lo} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, B_{lo} = \left[0, \frac{1}{m}\right]^T, x_{lo} = [x_1, x_2]^T$$

with m = mass of the vehicle, and u_{lo} is the force in the acceleration paddle, x_1 = longitudinal position, and x_2 = longitudinal velocity.

Lateral Dynamic Model

The lateral dynamic model is based on the position and orientation error variables as shown in Fig. 1. Let (T_x, T_y) be the coordinates of the target point, ψ be the orientation of the vehicle, e_1 be the error between the distance of the center of gravity of the vehicle and the center line of the target lane, and e_2 be the orientation error of the vehicle with respect to the road.

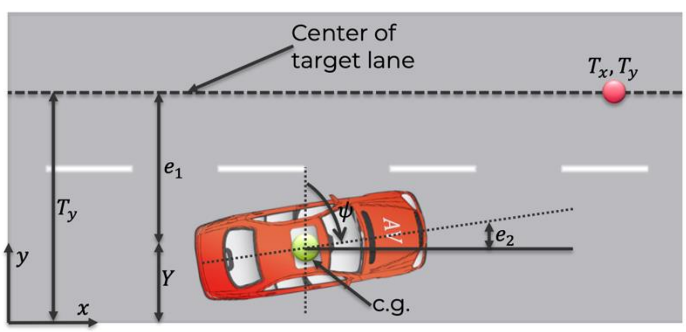

Fig. 1: Defining the errors $e_1(t)$ and $e_2(t)$

Assumption 1: Vehicles travel on a straight road with radius $R = \infty$.

The dynamic model is given as:

$$\dot{x}_{la} = A_{la}x_{la} + B_{la}u_{la} \qquad (2)$$

where u_{la}= the front wheel steering angle,

$x_{la} = [e_1(t), \dot{e_1}(t), e_2(t), \dot{e_2}(t)]^T, B_{la} = \left[0, \frac{2C_{\alpha f}}{m}, 0, \frac{2C_{\alpha f} l_f}{I_z}\right]^T$,

$$A_{la} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & -\frac{2C_{\alpha f}+2C_{\alpha r}}{mV_x} & \frac{2C_{\alpha f}+2C_{\alpha r}}{m} & \frac{-2C_{\alpha f}l_f+2C_{\alpha r}l_r}{mV_x} \\ 0 & 0 & 0 & 1 \\ 0 & \frac{-2C_{\alpha f}l_f-2C_{\alpha r}l_r}{I_zV_x} & \frac{2C_{\alpha f}l_f-2C_{\alpha r}l_r}{I_z} & \frac{-2C_{\alpha f}l_f^2+2C_{\alpha r}l_r^2}{I_zV_x} \end{bmatrix}.$$

To keep the development of the model-free learning-based controller simple, in this report, we assume that the longitudinal and lateral dynamics are linear. It can be seen from the A_{la} that the longitudinal velocity V_x appears non-linearly. Thus, Assumption 1 is a valid assumption for a linear lateral dynamical model for lane changing. From Fig. 1, the lateral position $Y(t)$ and yaw angle $\psi(t)$ can be obtained as:

$$Y(t) = T_y - e_1(t),$$

$$\psi(t) = \psi_{des} + e_2(t). \tag{3}$$

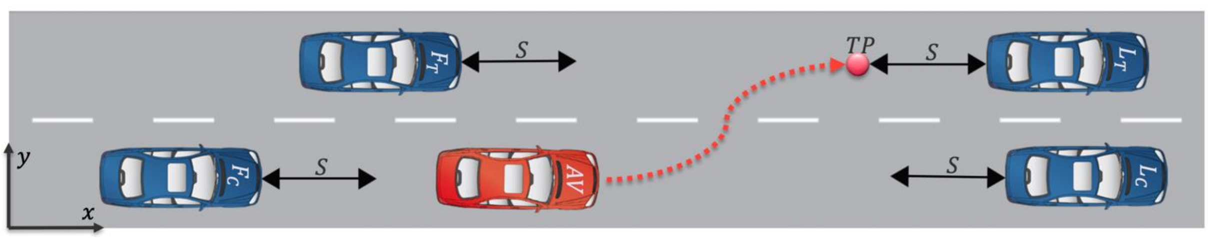

Fig. 2: A typical lane change Scenario

Lane Change Decision Making

In Fig. 2, AV denotes the autonomous vehicle, L_C denotes the lead vehicle in the current lane, F_C denotes the following vehicle in the current lane, L_T denotes the lead vehicle in the target lane, F_T denotes the following vehicle in the target lane, S denotes the safety distance.

In this work, the lane change decision making is introduced for a single lane change maneuver. As shown in Fig. 2, four vehicles are involved in a lane change maneuver. The AV performs a maneuver to change the lane and places itself in the target point (TP). Let$x_{AV}, x_{L_T}, x_{F_T}, x_{L_C}, x_{F_C}$ be the longitudinal positions of the vehicles involved in the lane changing process. Let v_{AV}, and v_{F_T} be the velocities of the AV and FT respectively. Then, the following conditions must hold true for a safe lane change:

(4)

$$x_{\mathrm{AV}} \leq x_{LC} - S_{LC}(t)$$

$$x_{\mathrm{AV}} \geq x_{FC} + S_{FC}(t)$$

$$x_{\mathrm{AV}} \leq x_{LT} - S_{LT}(t)$$

$$x_{\mathrm{AV}} \geq x_{FT} + S_{FT}(t)$$

where the safe distance is $S_i(t) = L + hv_i(t)$, $i \in \{L_T, F_T, L_C, F_C\}$, h=headway time, L=Length of vehicle.

The safe distance $S_i(t)$ is evaluated continuously. If the above inequalities are violated at any time instant during the lane changing, the AV maneuvers back to the original lane. This maneuver is done based on the change of leader vehicle. Thus, when the safe conditions violate, the target point TP is chosen at a safe distance from the leader in the original lane. The complete lane change algorithm is presented in the flowchart below:

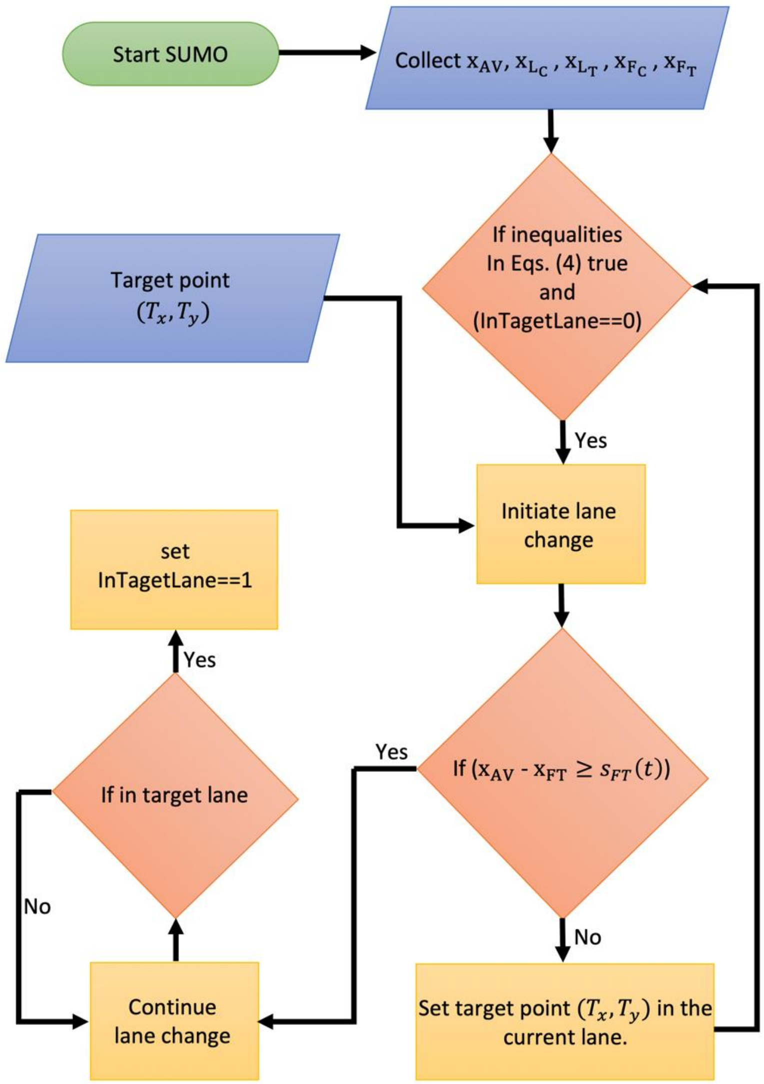

Fig. 3: Lane change decision algorithm.

Problem Definition

Given that the dynamics of the AV is assumed not perfectly known and TP in the target lane is defined, the following problem is addressed in this work:

Problem: Design a lane changing algorithm for the AV that incorporates the following:

(i) an optimal model-free controller for the AV's lateral maneuver such that $e_1 \to 0$, $e_2 \to 0$.

(ii) an optimal model-free controller for the AV's longitudinal maneuver.

(iii) a lane change decision making mechanism based on inequalities in (4).

(iv) an optimal model-free controller for post lane change platooning.

(v) learns and schedules optimal gains for change in longitudinal velocity.

Learning Algorithms

Model-based Learning

Consider a continuous-time linear time-invariant system described as:

$$\dot{x} = Ax + Bu \tag{5}$$

where, $x \in \mathbb{R}^n$ is the system state, $u \in \mathbb{R}^m$ is the control input, $A \in \mathbb{R}^{n\times m}$, and $B \in \mathbb{R}^{n\times m}$ are the unknown state and input matrices, respectively. It is assumed the all the states are available for feedback and the system (5) is stabilizable. We seek to design a linear optimal control law of the form:

$$u = -Kx, \tag{6}$$

such that the following cost function is minimized:

$$\begin{matrix} min \\ u \end{matrix} \quad J = \int_0^\infty (x^T Qx + u^T Ru), \tag{7}$$

where,

$Q = Q^T \geq 0, R = R^T > 0$, and $(A, Q^{1/2})$ is observable.

If A, B are completely known, the solution to the above-mentioned problem is well known and the optimal gain matrix $K^* \in \mathbb{R}^{m\times n}$ can be found as follows:

$$A^T P + PA + Q - PBR^{-1}B^T P = 0 \tag{8}$$

$$K^* = R^{-1}B^T P \tag{9}$$

where (8) is called the algebraic Riccati equation. Since the Riccati equation is non-liner in P, it is generally difficult to solve. In the literature, many efficient iterative approaches have been proposed to solve (8). One such approach is given in [17], which is reproduced for completeness:

Theorem 1*:* If K_0 is any stabilizing control gain, and P_k is the symmetric positive definite solution of the Lyapunov equation:

$$(A - BK_k)^T P_k + P_k(A - BK_k) + Q + K_k^T R K_k = 0, \tag{10}$$

$$K_{k+1} = R^{-1} B^T P_k, \tag{11}$$

Then, the following conditions hold:

- $(A - BK_k)$ is Hurwitz
- $P^* \leq P_{k+1} \leq P_k,$
- $\lim_{k\to\infty} K_k = K^*, \lim_{k\to\infty} P_k = P^*$

Note that (10) is linear in P. Thus, one can iteratively solve (10) and update K_k to numerically approximate the solution. But this assumes the complete knowledge of the system matrices A, B.

Model-free Learning

Here, we present an online model-free learning-based controller design strategy that does not assume any knowledge of the system matrices A, B. Consider the modified system equation as follows:

$$\dot{x} = A_k x + B(K_k x + u), \tag{12}$$

where, $A_k = A - BK_k$. Then, using (10), (11), and (12), we have:

$$\begin{aligned} &(t + \delta t)^T P_k x(t + \delta t) - x(t)^T P_k x(t) \\ &= \int_t^{t+\delta t} \left[x^T \left(A_k^T P_k + P_k A_k\right) x + 2(u + K_k x)^T B^T P_k x\right] d\tau \\ &= \int_t^{t+\delta t} x^T Q_k x \, d\tau + 2 \int_t^{t+\delta t} (u + K_k x)^T R K_{k+1} x \, d\tau \end{aligned} \tag{13}$$

where, $Q_k = Q + K_k^T R K_k$**.** It must be noted that (13) is independent of the systems matrices A, B.

Define the following:

$$x^T Q_k x = (x^T \otimes x^T)\mathrm{vec}(Q_k) \tag{14}$$

$$(u + K_k x)^T R K_{k+1} x = [(x^T \otimes x^T)(I_n \otimes K_k R) + (x^T \otimes u^T)(I_n \otimes R)]\mathrm{vec}(K_{k+1}) \tag{15}$$

For any positive integer l, define: $\Delta_{xx} \in \mathbb{R}^{l\times n^2}$, $I_{xx} \in \mathbb{R}^{l\times n^2}$, $I_{xu} \in \mathbb{R}^{l\times nm}$ as follows for $0 \le t_1 < t_2 < \cdots < t_l$:

$$\Delta_{xx} = \left[x^T \otimes x \,|_{t_1}^{t_1+\delta t}, x^T \otimes x \,|_{t_2}^{t_2+\delta t}, \cdots, x^T \otimes x \,|_{t_l}^{t_l+\delta t}\right]^T \tag{16}$$

$$I_{xx} = \left[\int_{t_1}^{t_1+\delta t} x \otimes x d\tau, \int_{t_2}^{t_2+\delta t} x \otimes x d\tau, \cdots, \int_{t_l}^{t_l+\delta t} x \otimes x d\tau\right]^T, \tag{17}$$

$$I_{xu} = \left[\int_{t_1}^{t_1+\delta t} x \otimes u d\tau, \int_{t_2}^{t_2+\delta t} x \otimes u d\tau, \cdots, \int_{t_l}^{t_l+\delta t} x \otimes u d\tau\right]^T. \tag{18}$$

Using (14)-(18), (13) can be written as:

$$\Gamma_k \begin{bmatrix} \mathrm{vec}(P_k) \\ \mathrm{vec}(K_{k+1}) \end{bmatrix} = \Psi_k, \tag{19}$$

where,

$$\Gamma_k = \left[\Delta_{xx}, -2I_{xx}\left(I_n \otimes K_k^T R\right) - 2I_{xu}(I_n \otimes R)\right], \tag{20}$$

$$\Psi_k = -I_{xx}\mathrm{vec}(Q_k). \tag{21}$$

Thus, given an initial stabilizing control input, the trajectories of the system can be recorded online in (16)-(18), which can then be recorded in the data matrices (20), (21). The learning-based control algorithm is presented in Fig. (4).

Assumption 3*:* There exists a sufficiently large integer $l > 0$, such that:

$$\mathrm{rank}([I_{xx}, I_{xu}]) = \frac{n(n+1)}{2} + mn \tag{22}$$

Theorem 2 [14,15]*:* Under the assumption (22), there is a unique pair of matrices P_k, K_{K+1}, with $P_k = P_k^T$, $\forall k \in \mathbb{Z}_+$, such that:

$$\Gamma_k \begin{bmatrix} \mathrm{vec}(P_k) \\ \mathrm{vec}(K_{k+1}) \end{bmatrix} = \Psi_k \tag{23}$$

Theorem 3 [14,15]*:* Given an initial stabilizing gain K_0 if (22) holds, the sequence $\{P_i\}_0^\infty$ and $\{K_i\}_0^\infty$ obtained by solving (23) converge to the optimal values P^* and K^*, respectively.

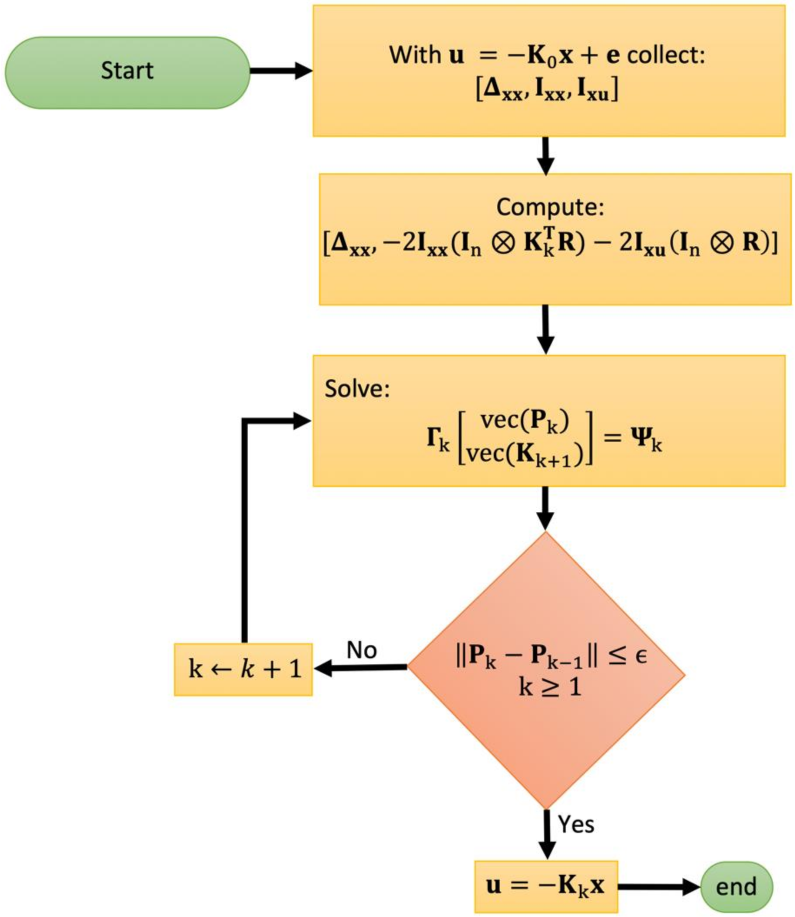

Fig. 4: Learning-based control algorithm.

Gain Scheduling

The design of controllers by the gain scheduling technique is as follows: linear time-invariant approximations of the system at a good number of operating points of the system are obtained; linear time-invariant controllers are designed for each linearized representation of the system at the selected operating points, so that the stability and certain performance objectives are achieved; these controllers are then linked together in order to obtain a single controller for the entire range of the system operation [18]. Consider the following linear parameter varying (LPV) system:

$$\dot{x} = A(\alpha)x(t) + B(\alpha)u(t),$$

$$\alpha = \alpha(t) \in [\alpha_0, \alpha_n] =: I \subset R \tag{24}$$

where, the system matrices $A(\alpha) \in \mathbb{R}^{n\times n}, B(\alpha) \in \mathbb{R}^{n\times m}, C(\alpha) \in \mathbb{R}^{q\times n}$ are functions of time varying parameter $\alpha = \alpha(t) \in [\alpha_0, \alpha_n] =: I \subset \mathbb{R}$.

Assumption 3: The system in (25) is controllable for all $\alpha \in I$.

In gain scheduling technique, we try to design a feedback control law of the form:

$$u(t) = -K(\alpha)x(t), \tag{25}$$

where, $K(\alpha) \in \mathbb{R}^{m\times n}$ is the state feedback gain matrix. To design the state feedback control law in (25), we first select finite number of fixed $\alpha \in I$. Next, for each fixed α the state feedback gain matrix $K(\alpha)$ is computed so that the stability and certain design goals are achieved for the closed-loop system. Let $K(\alpha_l)$ and $K(\alpha_{l+1})$, respectively, denote the gain matrices computed at two adjacent points α_l, and α_{l+1}. At each $\alpha \in [\alpha_l, \alpha_{l+1}]$, the gain $K(\alpha)$ in (25) is chosen to be the linear interpolation between $K(\alpha_l)$ and $K(\alpha_{l+1})$, given as [18]:

$$(\alpha) = K(\alpha_l) + \frac{K(\alpha_{l+1}) - K(\alpha_l)}{\alpha_{l+1} - \alpha_l}(\alpha - \alpha_l). \tag{26}$$

Note that $\alpha(t) = V_x(t) = x_2(t)$, and $\dot{\alpha}(t) = \dot{V}_x(t) = \mathrm{u_{lo}}/\mathrm{m}$ which is the longitudinal acceleration. In this work, we obtain each $K(\alpha_i)$ by means of the learning-based control technique discussed above. Thus, for all fixed α_i, we obtain an optimal controller by Theorems 2 and 3. Thus, stability is guaranteed for each of the fixed α_i's. To guarantee the stability of the overall system, we need that $V_x(t)$ is slowly varying [18]. Since $\dot{V}_x(t) = \mathrm{u_{lo}}/\mathrm{m}$, we can design $\mathrm{u_{lo}}$ such that $V_x(t)$ is slowly varying. Thus, to guarantee overall system stability, one needs to design a feedback law K_{lo} for the longitudinal motion such that that vehicle acceleration has a small magnitude. The algorithm of the learning-based gain scheduled controller is given in Fig. 5 and is explained next.

The algorithm starts by initializing the initial stabilizing controllers $K_0^1, K_0^2, \cdots, K_0^n$ for the finite velocity points $\mathrm{V_x^1}, \mathrm{V_x^2}, \cdots, \mathrm{V_x^n}$ where we wish to freeze our system and learn the optimal controller gains $K^1, K^2, \cdots, K^n$. Next, we collect the position data $\mathrm{x_{AV}}$, $\mathrm{x_{L_C}}$, $\mathrm{x_{L_T}}$, $\mathrm{x_{F_C}}$, $\mathrm{x_{F_T}}$ and feed it to the lane change decision module (see Fig. 3). The lane change decision module decides whether to do a lane change or remain in the desired lane based on the safety conditions explained above. In any situation, we collect the actual velocity data $\mathrm{V_x^{AV}}$ of the AV. The gain scheduling technique suggests that we need to freeze our system at all the velocity points $\mathrm{V_x^1}, \mathrm{V_x^2}, \cdots, \mathrm{V_x^n}$, but in a practical scenario this is not possible as when the AV is on the road it would be very challenging to maintain a constant $\mathrm{V_x^{AV}}$. Thus, we define a tolerance value $\epsilon_1$such that when $\mathrm{V_x^{AV}}$ is close to one of the $\mathrm{V_x^i}$'s and $\left|\mathrm{V_x^{AV}} - \mathrm{V_x^i}\right| \le \epsilon_1$ we assume $\mathrm{V_x^{AV}} = \mathrm{V_x^i}$ and start collecting data for learning K^i for $\mathrm{V_x^i}$ and store in the database. It must be noted that the longitudinal velocity of the car might vary and the condition $\left|V_x^{AV} - V_x^i\right| \leq \epsilon_1$ might not be always satisfied when we start collecting data for V_x^i. In such scenario, we again start from the very beginning step of collecting the position data x_{AV}, x_{L_C} , x_{L_T} , x_{F_C} , x_{F_T} and repeat all the steps as explained above until we have collected enough data. Once we collect enough data, say m samples for a particular velocity V_x^i, we pass the data to the learning module (see Fig. 4). The learning module then returns the learned controller gain K^i and it is stored in the database. The flag learned V_x^i is used to avoid repeated learning for the same V_x^i. Once, a gain K^i is learned for a V_x^i, we change K_0^1 with K^i and use the new controller for the AV maneuvers. Once K^i and K^{i+1} for two given adjacent points are learned, we define the interval $[V_x^i, V_x^{i+1})$ and use the use the interpolated gain given in (26) to obtain the control signal whenever V_x^i lies the interval $[V_x^i, V_x^{i+1})$.

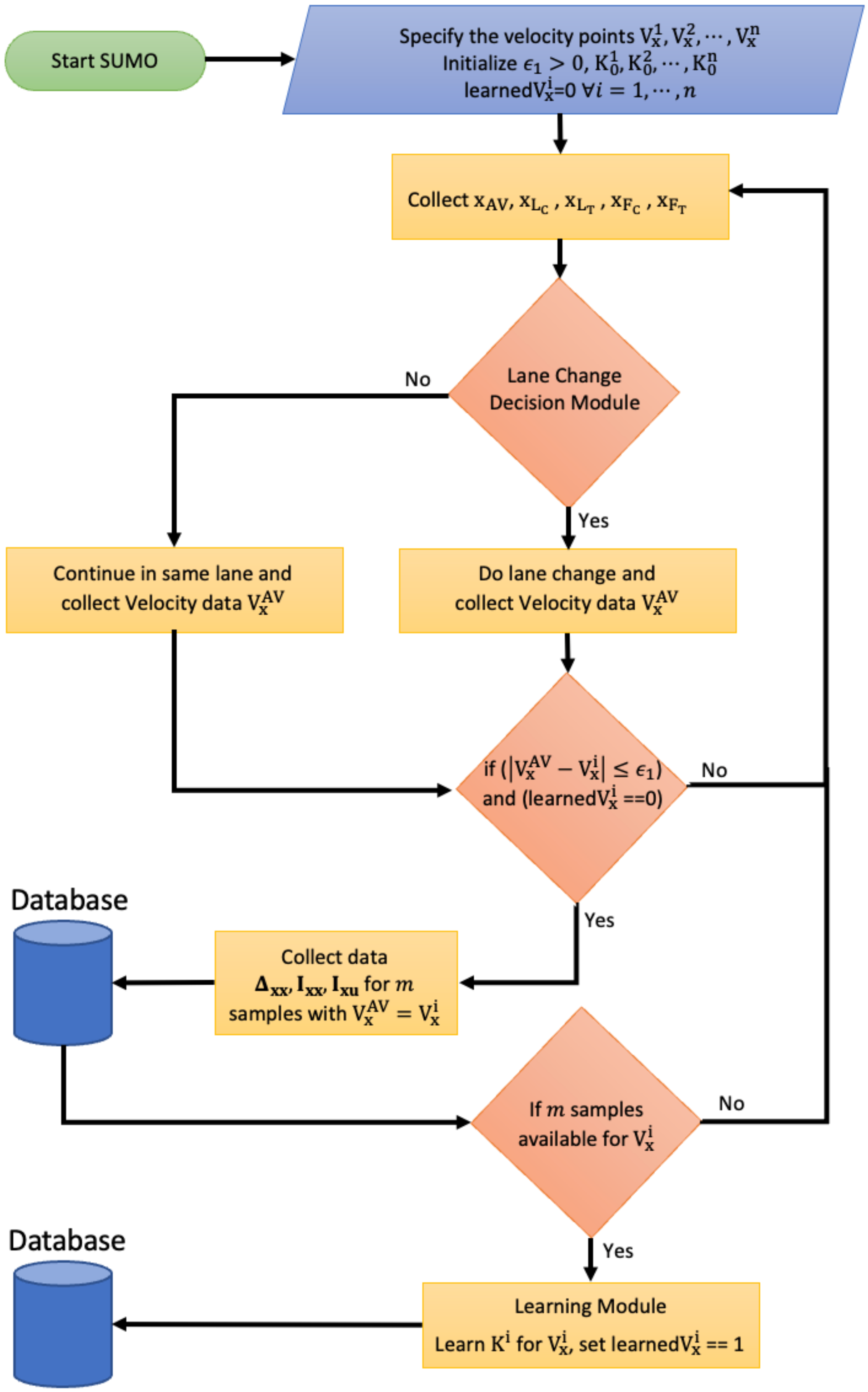

Fig. 5: Learning-based gain scheduling algorithm.

Results and Discussions

In this section, simulation results are conducted to show the efficacy of the data-driven controller and the lane change decision making algorithm. We have implemented the proposed gain-scheduled learning-based controller and the lane change decision algorithm in SUMO to test the effectiveness of the proposed methodology. We have obtained all the simulation data from SUMO environment and plotted them using MATLAB. This section explains the SUMO simulation data via different plots. We use the following weight matrices for the lateral control:

$$Q = \mathrm{diag}([1, 2, 2000, 3000]), \qquad R = 1.$$

For the purpose of learning with an initial control gain K_0, we apply the control input $u = K_0 x + e$, where e is noise which is obtained using the summations of sinusoidal signals with randomly distributed frequencies. Note that the noise e is deterministic [14-15]. The choice of the Q and R matrices is done considering the passenger and driver comfort and low fuel use. The last two diagonal entries in Q**,** i.e., 2000 and 3000 will penalize any large error in the yaw angle $(e_2(t), \dot{e}_2(t))$ of the AV this will ensure passenger and driver comfort. The first two diagonal entries in Q**,** i.e., $1,$ and 2 will penalize the error in lateral position $(e_2(t), \dot{e}_2(t))$ of AV. If the first two diagonal entries are increased, then the error in lateral position of AV will need to be penalized more but this will require a more aggressive controller which might increase fuel consumption. We have chosen $R = 1$ using trial and error. We have found that with $R = 1$, the control input to the AV, i.e., the steering angle of the AV, can be computed such that the driver comfort is assured.

As explained above, we use the proposed technique of gain scheduling learning-based controller to learn the controller gains. It was observed that the AV longitudinal velocity varies roughly between 20m/s to 22m/s. Thus, we choose the velocity intervals for gain scheduling as $[20\mathrm{m/s}, 21\mathrm{m/s})$, and $[21\mathrm{m/s}, 22\mathrm{m/s}]$. Thus, we need to learn optimal controllers for three velocities, i.e., $\mathrm{V}_\mathrm{x}^1 = 20\mathrm{m/s}, \mathrm{V}_\mathrm{x}^2 = 21\mathrm{m/s},$ and $\mathrm{V}_\mathrm{x}^3 = 22\mathrm{m/s}$. We use the algorithm presented in Fig. 5 to perform gain scheduling-based learning. We use the following initial stabilizing control gains:

$$K_0^1 = [0.535, 0.023, 88.546, 92.4412]$$

$$K_0^2 = [0.535, 0.0258, 89.214\ , 92.445]$$

$$K_0^3 = [0.535, 0.028, 89.883\ , 92.448]$$

The tolerance ϵ_1 is set as 0.2. To demonstrate the learning process and application of the learned gains we perform the lane changing two times. The Fig. 6 shows the velocity profiles of the AV

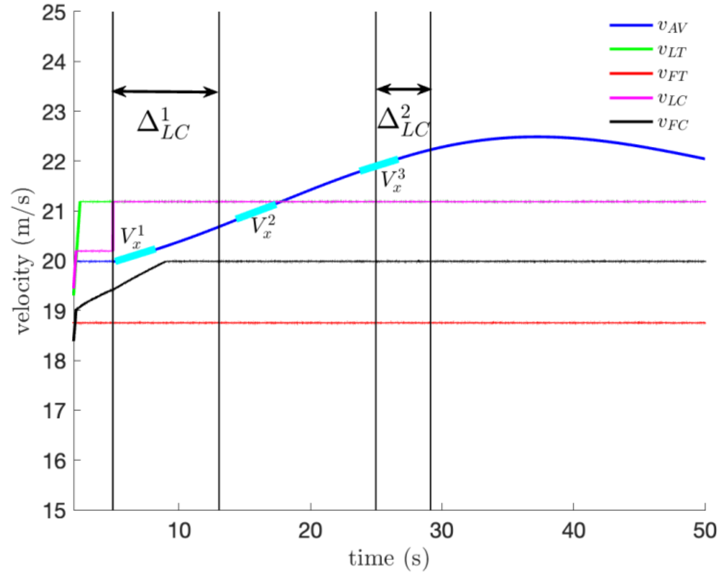

Fig. 6: Velocities of the vehicles.

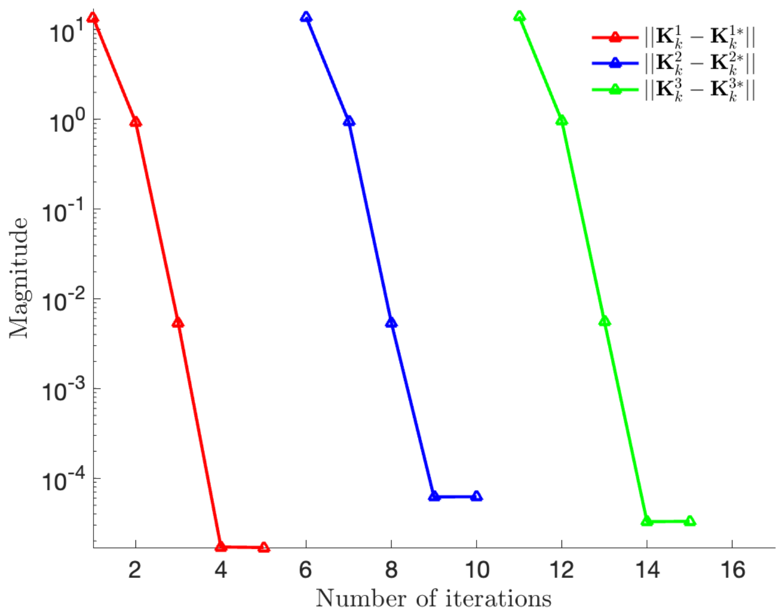

Fig. 7: The convergence of the of the optimal gains

and the surrounding vehicles that are obtained from the SUMO environment, where Δ_{LC}^{1} and Δ_{LC}^{2} are the lane pre-learning and post-learning lane changing times. The light blue strips in Fig. 6 indicate the intervals where $\left|\mathrm{V}_{\mathrm{x}}^{\mathrm{AV}}-\mathrm{V}_{\mathrm{x}}^{\mathrm{i}}\right| \leq \epsilon_1$ was satisfied for each $\mathrm{V}_{\mathrm{x}}^{\mathrm{i}}$. Here, each of these intervals comprises of 300 data points. Thus, with a sampling rate of 0.01s, we collect data for 3s for learning. Observe that the data for $\mathrm{V}_{\mathrm{x}}^{1}$ was collected during the first lane change duration Δ_{LC}^{1}, the data for $\mathrm{V}_{\mathrm{x}}^{2}$ was collected when the vehicle was already in the desired lane, and the data for $\mathrm{V}_{\mathrm{x}}^{3}$ was collected when the vehicle was maneuvering to perform to second lane change. Thus, we collect data to cover different scenarios during lane changing. After the updated optimal controller gains were obtained for every $\mathrm{V}_{\mathrm{x}}^{\mathrm{i}}$, the initial gains are replaced with the updated gains and the AV maneuver is performed using the updated gains and using the interpolation formula presented in (26). It is evident for Fig. 6 that $\Delta_{\mathrm{LC}}^{2} \approx 8s < \Delta_{\mathrm{LC}}^{1} \approx 4s$. Thus, a significant improvement is seen in the lane changing time with the update optimal gain scheduled controller.

Figure 7 shows the convergence of the optimal gains. The ϵ in Fig. 4 is set as 10^{-4}. One can set the tolerance ϵ lower but this would increase the computation time that might affect the real time application of the learned controller. It is clearly seen from Fig. 7 that the gains converge to the optimal gains with just 5 iterations. Thus, it can be said that for the proposed algorithm, 300 samples or in other words 3s data is enough for the learning algorithm (Fig. 4) to converge.

Figure 8 shows the safe distance of the AV with the surrounding vehicle. The safe distance is defined as $S_i(t) = L + hv_i(t)$, $i \in \{LT, FT, LC, FC\}$, h= 0.5s is the headway time, L= 5m is the length of vehicle. It can be observed at $t = 0s$, the AV was not at a safe distance from FT, thus the AV does not start a lane change maneuver. At $t \approx 5s$, the safety conditions for lane changing (4) satisfy for all the surrounding vehicles and the AV starts the first lane change maneuver that is completed in approximately $8s$. For the second lane change maneuver, the vehicles are already at safe distance, thus the AV can safely start the lane change maneuver.

Figure 9 shows the states of the lateral system. It can be seen that the states converge to zero with the application of the controllers obtained using the proposed methodology. It was mentioned above that the gain scheduled controller can guarantee overall system stability if the feedback law K_{lo} for the longitudinal motion can be obtained such that that vehicle acceleration has a small magnitude. Here, we have obtained the K_{lo}^{*}= [1, 52.63] with $Q_{lo} = \mathrm{diag}([1, 1])$, $R_{lo} = 1$ using historical data. The choice of Q_{lo} and R_{lo} must be such that the acceleration has a lower magnitude. Figure 10 shows the longitudinal acceleration profile of the AV. It can be seen that the acceleration magnitude is low.

Remark: If K_{lo}^{*} is very conservative such that the AV response in tracking the x-coordinate of the target point (T_x) is sluggish, one can change K_{lo}^{*} to a more aggressive gain for target tracking. But this must only be done when the error states of the lateral motion are negligible. In other words, when the AV has already reached T_y, and there is no dependence on the lateral dynamics, one can switch to an aggressive K_{lo}^* for better tracking of T_x.

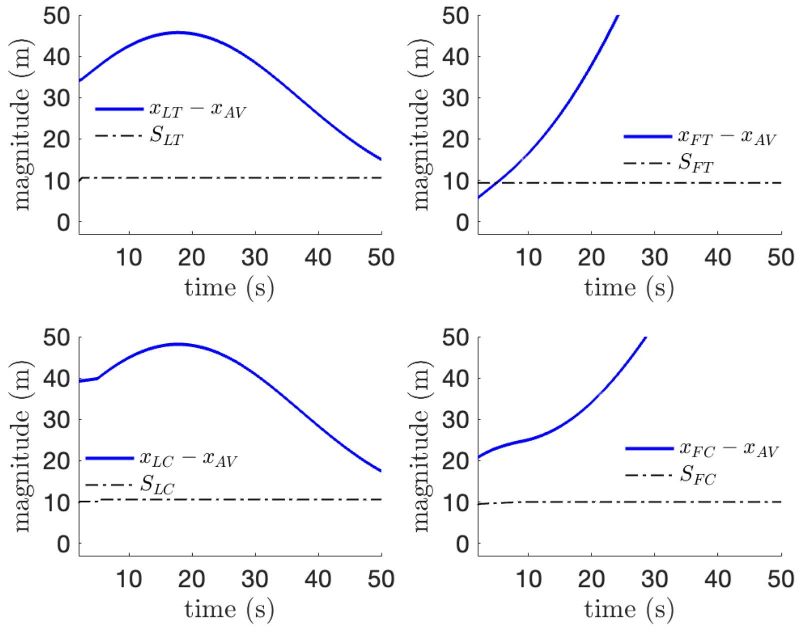

Fig. 8: Safe distances of AV from surrounding vehicles.

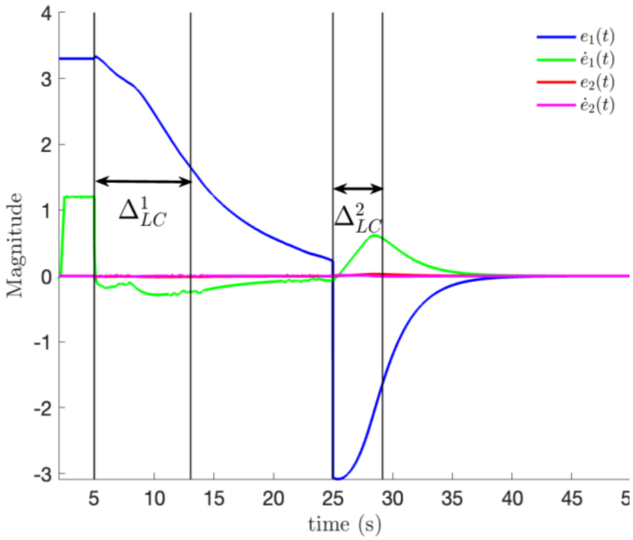

Fig. 9: System states.

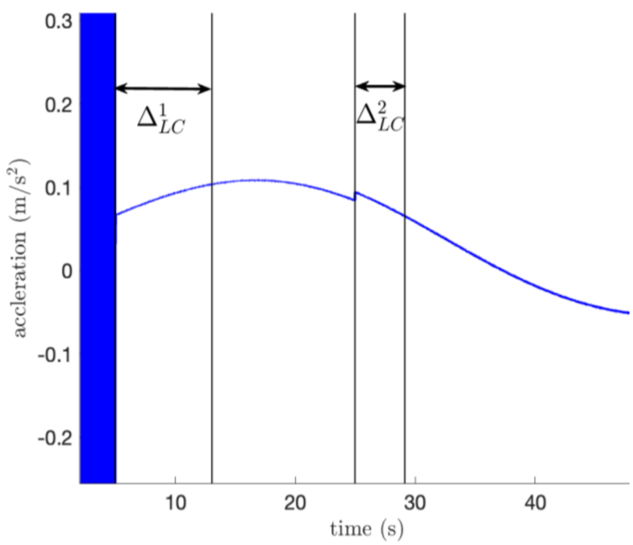

Fig. 10: Acceleration of AV.

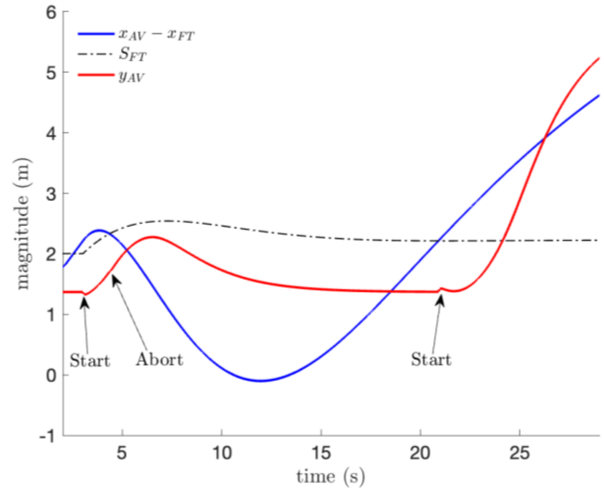

Fig. 11: Lane abortion of AV.

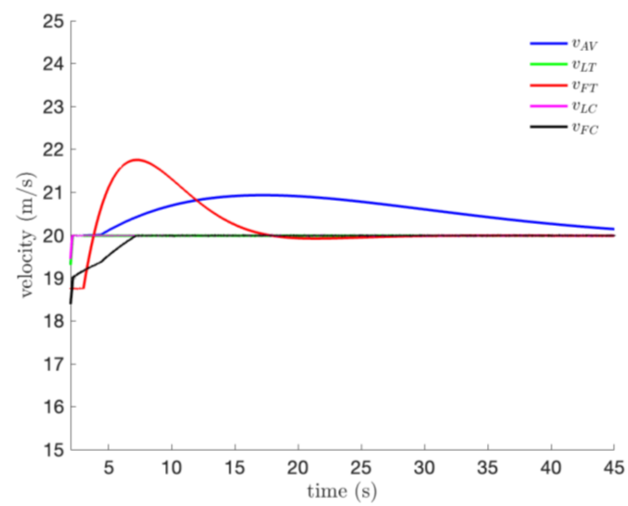

Fig. 12: Velocities during lane abortion.

Figure 11 shows the effectiveness of the lane change algorithm discussed in Fig. 3, and the velocities of the vehicles are shown in Fig. 12. The lane change starts at 3s. From Fig. 12, it can be seen that the FT starts accelerating more than the AV. At around 4s, FT comes close the AV and it starts aborting the lane change and maneuvers back to the current lane. Again, at 21s when the safety conditions satisfied, the AV starts maneuvering to the target lane. It must be noted that the graphs in Fig. 11 are normalized to bring them in the same scale for the sake of clarity in understanding the AV maneuver during lane abortion and lane changing.

Conclusion and Contributions

In this work, we have introduced an optimal data-driven control algorithm to solve the lane changing problem of AVs, where we make use of the online information of the state and input to solve the algebraic Riccati equation iteratively by using approximate/adaptive dynamic programming (ADP). In this work, we have assumed that the state and control input are received from a linear system. In order to make the proposed methodology applicable to non-linear and/or parameter varying systems, we have proposed a gain scheduling-based data driven control technique to learn optimal gains. Also, we have developed a lane change decision making algorithm to ensure safe and efficient lane change. Safety is assured during lane changing by maintaining a safe distance from the surrounding vehicles. The proposed lane changing algorithm can make the AV perform a lane abortion if safety conditions are violated during lane change. The optimal data-driven gain scheduled control algorithm and the lane change decision making algorithm has been validated by means of SUMO and MATLAB based computer simulations.

As compared to existing methodologies in the literature, our proposed method is completely data driven. We do not use or assume any information of the system parameters. We only assume the knowledge of the state vector and the control input, and derive a model-free optimal controller with guaranteed stability. It must be noted that many methodologies in the literature of lane changing does not guarantee optimal control of their AV. Many of the techniques that are proposed in the literature requires to solve an optimization problem at every time step whereas our proposed methodology only requires to learn at specific time intervals with a smaller number of data points when the longitudinal velocity changes. Also, due to the fast convergence of the proposed methodology, it is suitable for real-time applications. Although, we assume that we receive data from a linear model, the gain scheduling-based data-driven controller design adds to the versatility of the methodology that makes the proposed methodology applicable to non-linear systems and/or parameter varying systems as well. Also, with the obtained optimal gain parameters, the lane changing time is seen to considerably improve.

References

[1] J. D. Chovan, L. Tijerina, G. Alexander, and D. L. Hendricks, “Examination of lane change crashes and potential IVHS countermeasures,” Nat. Highway Traffic Safety Admin., U.S. Dept. Transp., Washington, DC, USA, Tech. Rep., 1994.

[2] Z. Wang, X. Shi, and X. Li. “Review of lane-changing maneuvers of connected and automated vehicles: models, algorithms and traffic impact analyses,” Journal of the Indian Institute of Science, vol. 99, no. 4, pp. 589-599, 2019.

[3] J. Nilsson, J. Silvlin, M. Brannstrom, E. Coelingh, and J. Fredriksson, “If, when, and how to perform lane change maneuvers on highways,” IEEE Intelligent Transportation Systems Magazine, vol. 8, no. 4, pp. 68-78, 2016.

[4] Z. Wang, X. Zhao, Z. Chen, and X. Li, “A dynamic cooperative lane-changing model for connected and autonomous vehicles with possible accelerations of a preceding vehicle,” Expert Systems with Applications, vol. 173, pp. 114675, 2021.

[5] Z. Wang, X. Zhao, Z. Xu, X. Li, and X. Qu, “Modeling and field experiments on autonomous vehicle lane changing with surrounding human‐driven vehicles,” Computer-Aided Civil and Infrastructure Engineering, vol. 36, no. 7, pp. 877-889, 2021.

[6] Y. Luo, Y. Xiang, K. Cao, and K. Li, “A dynamic automated lane change maneuver based on vehicle-to-vehicle communication,” Transportation Research Part C: Emerging Technologies, vol. 62, pp. 87-102, 2016.

[7] Z. Wang, X. Zhao, Z. Xu, X. Li, Z. Qu, “Modeling and field experiments on lane changing of an autonomous vehicle in mixed traffic”, Computer-aided Civil and Infrastructure Engineering, 2020.

[8] J. Nie, J. Zhang, W. Ding, X. Wan, X. Chen, and B. Ran, “Decentralized cooperative lane-changing decision-making for connected autonomous vehicles”, IEEE access, vol. 4, pp. 9413-9420, 2016.

[9] J. Nilsson, Y. Gao, A. Carvalho, and F. Borrelli, “Manoeuvre generation and control for automated highway driving”, IFAC Proceedings Volumes, vol. 47, no. 3, pp. 6301-6306, 2014.

[10] B. Li, Y. Zhang, Y. Ge, Z. Shao, and P. Li, “Optimal control-based online motion planning for cooperative lane changes of connected and automated vehicles”, 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 3689-3694, 2017.

[11] A. Carvalho, Y. Gao, S. Lefevre, and F. Borrelli, “Stochastic predictive control of autonomous vehicles in uncertain environments”, 12th International Symposium on Advanced Vehicle Control, pp. 712-719, 2014.

[12] J. Suh, H. Chae, K. Yi, “Stochastic model-predictive control for lane change decision of automated driving vehicles”, vol. 67, no. 6, pp. 4771-4782, 2018.

[13] G. Xu, L. Liu, Y. O. Li, and Z. Song, “Dynamic modeling of driver control strategy of lane-change behavior and trajectory planning for collision prediction”, vol. 13, no. 3, pp 1138-1155, 2012.

[14] Y. Jiang and Z. P. Jiang, “Robust Adaptive Dynamic Programming”. Hoboken, NJ, USA: Wiley, 2017.Cambridge, MA, USA: MIT Press, 2018.

[15] Y. Jiang, Z. P. Zhong-Ping, “Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics”, vol. 48, no. 10, pp. 2699-2704, 2012.

[16] W. Gao, Z. P. Jiang, and K. Ozbay, “Data-driven adaptive optimal control of connected vehicles,” *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 5, pp. 1122–1133, May 2017.

[17] D. Kleinman, “On an iterative technique for Riccati equation computations”, IEEE Transactions on Automatic Control, vol. 13, no. 1, pp. 114-115, 1968.

[18] S. M. Shahruz, S. Behtash, “Design of controllers for linear parameter-varying systems by the gain scheduling technique”, vol. 168, no. 1, pp. 195-217, 1992.