Transportation Research Part B 145 (2021) 80–117

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Online predictive connected and automated eco-driving on signalized arterials considering traffic control devices and road geometry constraints under uncertain traffic conditions

Shuaidong Zhao, Kuilin Zhang*

1400 Townsend Drive, Department of Civil and Environmental Engineering, Michigan Technological University, Houghton, MI 49931, United States

ARTICLE INFO

Article history:
Received 7 December 2018
Revised 17 November 2020
Accepted 29 December 2020

Keywords:
Connected and automated Eco-driving on signalized arterials
Location-based traffic control devices and road geometry constraints
Model predictive control
Data-driven optimization
Uncertain traffic conditions

ABSTRACT

For energy-efficient Connected and Automated Vehicle (CAV) Eco-driving control on signalized arterials under uncertain traffic conditions, this paper explicitly considers traffic control devices (e.g., road markings, traffic signs, and traffic signals) and road geometry (e.g., road shapes, road boundaries, and road grades) constraints in a data-driven optimization-based Model Predictive Control (MPC) modeling framework. This modeling framework uses real-time vehicle driving and traffic signal data via Vehicle-to-Infrastructure (V2I) and Vehicle-to-Vehicle (V2V) communications. In the MPC-based control model, this paper mathematically formulates location-based traffic control devices and road geometry constraints using the geographic information from High-Definition (HD) maps. The location-based traffic control devices and road geometry constraints have the potential to improve the safety, energy, efficiency, driving comfort, and robustness of connected and automated driving on real roads by considering interrupted flow facility locations and road geometry in the formulation. We predict a set of uncertain driving states for the preceding vehicles through an online learning-based driving dynamics prediction model. We then solve a constrained finite-horizon optimal control problem with the predicted driving states to obtain a set of Eco-driving references for the controlled vehicle. To obtain the optimal acceleration or deceleration commands for the controlled vehicle with the set of Eco-driving references, we formulate a Distributionally Robust Stochastic Optimization (DRSO) model (i.e., a special case of data-driven optimization models under moment bounds) with Distributionally Robust Chance Constraints (DRCC) with location-based traffic control devices and road geometry constraints. We design experiments to demonstrate the proposed model under different traffic conditions using real-world connected vehicle trajectory data and Signal Phasing and Timing (SPaT) data on a coordinated arterial with six actuated intersections on Fuller Road in Ann Arbor, Michigan from the Safety Pilot Model Deployment (SPMD) project.

1. Introduction

As the development of connected and automated vehicle (CAV) technologies, the real-time Vehicle-to-Infrastructure (V2I) and Vehicle-to-Vehicle (V2V) data can be utilized to develop online connected and automated driving control systems on

* Corresponding author.
E-mail addresses: szhao3@mtu.edu (S. Zhao), klzhang@mtu.edu (K. Zhang).

https://doi.org/10.1016/j.trb.2020.12.009
0191-2615/

signalized arterials. The real-time V2I data such as Signal Phasing and Timing (SPaT) and V2V data such as vehicle position, velocity, acceleration, and spacing can be used to improve the safety, energy, efficiency, driving comfort, and robustness of the connected and automated driving. However, to the best of our knowledge, none of the existing optimal control models consider traffic control devices (e.g., road markings, traffic signs, and traffic signals) and road geometry (e.g., road shapes, road boundaries, and road grades) as constraints for connected and automated driving problems. At the same time, the preceding traffic information in front of a CAV is generally not given, especially when traffic oscillations exist, which leads to uncertainty in the connected and automated driving control problem. The traffic oscillations lead to uncomfortable driving, safety issues, time delays, and extra fuel consumption and emissions. The objective of this paper is to develop a data-driven optimization-based Model Predictive Control (MPC) modeling framework, which explicitly considers location-based traffic control devices and road geometry constraints in an online predictive control model, for a more robust, safer, smoother, and energy-efficient connected and automated driving control on signalized arterials under uncertain traffic conditions.

1.1. Literature review

It is still a research question that how to utilize High-Definition (HD) maps, which can serve as the complement to in-vehicle sensors to provide the traffic control device and road geometry information for driving safety, in predictive control models for connected and automated driving problems. Existing automated driving control models solely rely on in-vehicle sensors, which is unsafe and computationally expensive for real-time applications. For example, traffic light detection needs to detect traffic lights through camera captured pictures all the time when an automated vehicle is operating. HD maps such as the maps from HERE and TomTom (Massow et al., 2016) contain high-precision (e.g., at the centimeter level) road information, which is essential for the safety of connected and automated driving. HD maps provide high-precision interrupted flow facility locations, including traffic control devices such as road markings, traffic signs, traffic signal information on signalized arterials (Roess et al., 2004), and road geometry features such as lines, curves, and spirals (Dupuis and Grezlikowski, 2006). The necessity of HD maps for connected and automated driving is obvious because the sensors of automated vehicles are difficult to localize and determine a vehicle's position in reference to its surroundings accurately (Seif and Hu, 2016). The current sensing data from in-vehicle sensors without using HD maps cannot be used to predict future driving states of controlled vehicles, such as energy consumptions.

However, directly using networks from HD maps may result in high computational costs for driving control of connected and automated vehicles because the vehicle dynamics in coordinate-based HD maps are nonlinear (e.g., with both longitudinal and lateral accelerations, Katrakazas et al., 2015). Especially for an optimization-based automated driving control model, the formulation based on a high-density network from HD maps will involve constraint nonlinearity (e.g., nonlinear vehicle dynamics with steering angles, Dixit et al., 2018), which causes a computational burden to real-time applications of a control model (Carvalho et al., 2013). Multiple vehicle manufacturers have released first-generation automated driving functionalities such as Lane Keeping Assist (Level 2 Automation, SAE, 2014). The Lane Keeping Assist functionality is the potential to focus the connected and automated driving control problem on a scope of driving control model with linear vehicle dynamics without considering nonlinear dynamics such as steering angles. Therefore, there is a need to develop an adequate method to use accurate traffic control devices and road geometry information from HD maps in connected and automated driving control with linear vehicle dynamics in real roads.

In addition to the issue of using HD maps, another one for connected and automated driving on signalized arterials is the uncertain traffic conditions. In the literature for automated driving at signalized intersections, there are three categories of research approaches: namely optimal control models based on closed-form controllers (Widodo et al., 2000; Barth et al., 2011; Rakha and Kamalanathsharma, 2011; Hao et al., 2015; He et al., 2015), trajectory optimization models (Ma et al., 2017; Wang et al., 2017; Jiang et al., 2018; Sun et al., 2020; Yang et al., 2020), and MPC models (Asadi and Vahidi, 2009; Kamal et al., 2010). However, these methods focus on the connected vehicle driving at signalized intersections without traffic uncertainties. These methods are suitable for providing speed advisories for energy efficiency to connected vehicles at signalized intersections. Still, they may not be directly used for the online connected and automated driving control on arterials because they do not consider uncertain traffic conditions or assume queue information can be given precisely.

To deal with uncertainty, the stochastic and robust MPC models are able to guarantee stability and constraint satisfaction (Calafiore and Fagiano, 2013). The existing stochastic MPC models assume a known distribution of stochastic parameters (Schwarm and Nikolaou, 1999; Suh et al., 2016). The robust MPC models assume the bounded compactness of stochastic parameters and use the minimax optimization to deal with uncertainty (Scokaert and Mayne, 1998; Wang et al., 2009). Unfortunately, accurate traffic prediction information or the exact distribution of stochastic parameters are rarely available in practice. In this paper, we propose a method to integrate data-driven optimization methods with an MPC model to deal with the uncertainty. Specifically, the data-driven optimization methods proposed in this paper include Distributionally Robust Stochastic Optimization (DRSO) models (with moment bounds) and Distributionally Robust Chance Constraints (DRCC)-based models.

Introduced by Scarf (1958), the idea of DRSO is to define a data-driven uncertainty set (i.e., data-driven ambiguity set, or distributional set) of possible realizations of uncertain parameters and then optimize against worst-case realizations within the set (Calafiore and El Ghaoui 2006; Delage and Ye, 2010), which is used in the objective function to consider a set of Eco-driving references and uncertain traffic conditions. For DRSO models under moment bounds in our paper, a data-driven uncertainty set for the uncertain parameters is constructed using mean and variance information. Then the expected value

of the objective function in the worst case is optimized based on the distributions in that set (Scarf, 1958; Calafiore and El Ghaoui, 2006; Liu et al., 2015; Delage and Ye, 2010; Zymler et al., 2013). In a DRCC-based model, the mean and variance of the uncertain parameters are used to ensure that the probability of satisfying the constraint reaches a certain threshold (Calafiore and El Ghaoui 2006; Jiang and Guan, 2016). In the real-world driving environment, uncertain traffic flow dynamics may cause violations of safety constraints in an MPC model. For driving safety location-based traffic control devices and road geometry constraints, DRCC is used to deal with the uncertain traffic conditions in the MPC model for the proposed connected and automated driving systems.

1.2. Proposed work and contributions

In this paper, we propose a data-driven optimization-based MPC modeling framework explicitly considering location-based traffic control devices and road geometry constraints for connected and automated driving control on signalized arterials under uncertain traffic conditions by using V2V and V2I connected vehicle data. There are six steps in the proposed framework: (1) we propose a method to mathematically formulate the location-based traffic control devices and road geometry constraints in the time-indexed MPC model from the coordinate-based HD maps; (2) we develop an online learning-based driving dynamics prediction method to obtain the predicted driving states of the immediately preceding vehicle (IPV), which directly constrain the controlled vehicle driving dynamics; (3) we predict the driving reference for the controlled vehicle by an Eco-Driving policy for solving a constrained finite-horizon optimal control model; (4) we formulate a DRSO-DRCC based MPC model (using the predicted driving states of the IPV and the controlled vehicle to construct data-driven uncertainty sets) for determining the optimal acceleration or deceleration commands for the controlled vehicle; (5) we propose model reformulations and solution algorithm by applying the strong duality theory and the Semidefinite Relaxation technique to solve the proposed DRSO-DRCC based MPC model; (6) we conduct experiments on a real-world coordinated signalized arterial with six SPaT-enabled Roadside Units (RSUs) to evaluate the proposed model under different traffic conditions (i.e. light, medium, and heavy traffic) and weighting factors using the Safety Pilot Model Deployment (SPMD - http://safetypilot.umtri.umich.edu) dataset.

This paper makes three contributions to the literature as follows:

The first contribution is to propose a method to mathematically formulate the location-based traffic control devices and road geometry constraints from coordinate-based HD maps in a predictive control model for connected and automated driving. We first convert the given coordinate-based path to a distance-based path with traffic control devices and road geometry information along the given path. Then, we apply the traffic control devices and road geometry constraints with location indices to the time-indexed predictive control model. Based on this method, we can use linear vehicle dynamics rather than the nonlinear vehicle dynamics (e.g., the steering angles) in the connected and automated driving control problem. Moreover, the optimal vehicle trajectories along the distance-based path can be mapped back to the given coordinate-based path.

The second contribution is to integrate the location-based traffic control devices and road geometry constraints with the data-driven optimization-based MPC model for the connected and automated driving on signalized arterials. We apply the data-driven optimization-based MPC modeling framework (i.e., for connected and automated driving on freeways in Zhao and Zhang, 2018; Zhao and Zhang, 2020). In the data-driven optimization framework, we apply DRSO to deal with the controlled vehicle driving uncertainty by constructing a data-driven uncertainty set from mean and variance information of the predicted reference driving states of the controlled vehicles. In addition, we use DRCC to address the IPV driving uncertainty by constructing a data-driven uncertainty set by using the predicted driving states of the IPV in the safety constraint. We use an online learning-based driving dynamics prediction method to predict the uncertain driving states of the preceding vehicles and the controlled vehicle using V2V trajectory data and V2I traffic signal data.

The third contribution of this paper is to utilize the real-world V2V connected vehicle trajectory data from Onboard Units (OBUs), and V2I SPaT data from RSUs in the SPMD dataset for the demonstration of the proposed connected and automate driving model. We use the connected vehicle trajectory data from Basic Safety Messages (BSM) to calibrate the connected IDM model for the benchmark human driving behaviors. We extract traffic signal information from the SPaT messages for the actuated control on the real-world SPaT-enabled coordinated corridor (i.e., Fuller Road in Ann Arbor, MI). Then, we measure the performance of the proposed DRSO-DRCC model (i.e., as the automated driving) using the real-world V2V and V2I data.

This paper is structured as follows. Section 2 presents the notation and problem statement. In Section 3, we illustrate the data-driven optimization-based MPC modeling framework for the energy-efficient connected and automated driving control on signalized arterials under uncertain traffic conditions. The model reformulations and solution algorithms are described in Section 4. In Section 5, we demonstrate the proposed connected and automated driving control model using real-world connected vehicle trajectory data and the SPaT data from the SPMD dataset in Southeast Michigan. Concluding remarks are given in Section 6.

2. Problem statement

We begin this section with notation, definitions, and assumptions of the real-time connected and automated driving system on real-world arterials. Then, we discuss the problem of interest in this study. At the end of this section, we also

describe our method to express network geometries (e.g., links/routes, lanes, and interrupted flow facility locations such as stop lines) from HD maps.

2.1. Notation

We list the notation of indices, parameters, state variables, control variables, model matrices, and data-driven inputs as follows.

Indices:

k, control time index, $k = 1, 2, \ldots$
t, prediction horizon time index, $t = k, \ldots, k + PH$
s, distance index, $s = 1, .., S$
w, discretized road geometry index, $w = 1, .., W$

Parameters:

Δt, control time interval
Δs, homogeneous road geometry distance interval
$PH(k)$, time-variant prediction horizon at control time step k
$\theta(s)$, road grade at distance s
ξ, stochastic driving state scenario
α, Value-of-Energy (VOE) weighting factor
T^{obs}, observation time period
T^{p}, driving dynamics prediction time period
ΔT^{p}, driving dynamics prediction update interval
t^{0}, time when a controlled vehicle enters the communication range
$\hat{t}_{f,1}$, estimated arrival time at green
$\hat{t}_{f,2}$, estimated arrival time at red
$\hat{t}_{f}$, estimated arrival time of controlled vehicle
τ^{h}, following time headway
$t_{f,p}$, simulated arrival time of the IPV
$R(k)$, $G(k)$, red and green durations at control time k,
$t_{r,a}(k)$, red signal start time
$t_{r,b}(k)$, red signal end time
γ_1, confidence level of DRCC location-based traffic control devices and road geometry constraint
γ_2, confidence level of DRCC minimum safe time headway constraint
γ_3, confidence level of the first-moment constraint in the data-driven uncertainty set
γ_4, confidence level of the second-moment constraint in the data-driven uncertainty set
s^{min}, minimum safe spacing for stopped vehicles considering vehicle length
τ^{min}, minimum safe time headway
L^{veh}, vehicle length
$p^{int}(s)$, interrupted location at distance index s
$p^{int}(t)$, interrupted location at time t
$\lfloor \bullet \rfloor_d$, rounding operator to d trailing digits
$\boldsymbol{d}\{\bullet, \bullet\}$, Euclidean distance
$\phi(n, k)$, car-following model parameter set
$v_n^{CF}(t)$, velocity calculated by a car-following model
(x_w, y_w, g_w), the three-dimensional point with longitude, latitude, and road grade.

State variables:

$l(t)$, longitudinal location (rear bumper) of the controlled vehicle
$v(t)$, velocity of the controlled vehicle
$\boldsymbol{x}(t)$, driving state vector: longitudinal position and velocity - $x(t) = [l(t), v(t)]^T$

Control variable:

$u(t)$, acceleration of the controlled vehicle

Data-driven inputs:

$\hat{l}_s^{\xi}(t)$, maximum safe driving location
$\hat{v}_s^{\xi}(t)$, safe driving velocity

$\hat{l}_b^{\xi}(t)$, estimated stochastic boundary condition preceding vehicle location
$\hat{v}_b^{\xi}(t)$, estimated stochastic boundary condition preceding vehicle velocity
$\hat{l}_0^{\xi}(t)$, estimated stochastic IPV location
$\hat{v}_0^{\xi}(t)$, estimated stochastic IPV velocity
$\hat{l}_r^{\xi}(t)$, predicted reference driving location
$\hat{v}_r^{\xi}(t)$, predicted reference driving velocity
$\hat{x}_p^{\xi}(t)$, predicted driving states for the IPV - $[\hat{l}_0^{\xi}(t), \hat{v}_0^{\xi}(t)]_{1\times 2}^{T}$
$\hat{x}_r^{\xi}(t)$, predicted driving states for the controlled vehicle - $[\hat{l}_r^{\xi}(t), \hat{v}_r^{\xi}(t)]_{1\times 2}^{T}$
$\hat{x}^{\xi}(t)$, predicted driving states for the IPV and the controlled vehicle - $[\hat{l}_0^{\xi}(t), \hat{v}_0^{\xi}(t), \hat{l}_c^{\xi}(t), \hat{v}_c^{\xi}(t)]_{1\times 4}^{T}$
$\bar{x}_p^{\xi}(t)$, expected driving states of the IPV
$\hat{\sigma}_1$, covariance matrix of the driving states of the IPV for DRCC location-based traffic control devices and road geometry constraint
$\hat{\sigma}_2$, covariance matrix of the driving states of the IPV for DRCC minimum safe time headway constraint
$\hat{\mu}_p$, estimated mean of the predicted driving states of the IPV
$\hat{\Sigma}_p$, estimated covariance matrix of the predicted driving states of the IPV
$\hat{\mu}_r$, estimated mean of the predicted driving states of the controlled vehicle
$\hat{\Sigma}_r$, estimated covariance matrix of the predicted driving states of the controlled vehicle
$\hat{\mu}$, estimated mean of the predicted driving states of the IPV and the controlled vehicle
$\hat{\Sigma}$, estimated covariance matrix of the predicted driving states of the IPV and the controlled vehicle
$Q_p = \begin{bmatrix} 1 & \tau^{min} \\ 0 & 1 \end{bmatrix}$, matrix for the driving state x(t) of the controlled vehicle in the DRSO objective
$Q_p^{\xi} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$, matrix for the predicted driving state $\hat{x}^{\xi}(t)$ of the controlled vehicle in the DRSO objective
$Q_r^{\xi} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, matrix for the predicted driving state $\hat{x}^{\xi}(t)$ in the reference state tracking of the controlled vehicle in the DRSO objective
$R = I$, matrix for the driving comfort of the following controlled vehicle in the DRSO objective
$K_r = [1 \quad 0]$, matrix in the location-based traffic control devices and road geometry DRCC
$G_r = [-1 \quad 0]$, matrix in the location-based traffic control devices and road geometry DRCC
$K_s = [1 \quad \tau^{min}]$, matrix in the minimum time headway safety DRCC
$G_s = [-1 \quad 0]$, matrix for the predicted driving state of the controlled vehicle in the minimum time headway safety DRCC
$K_f = [0 \quad 1]$, matrix in the safety and speed limit constraints
$G_f = v^{limit}$, matrix in the safety and speed limit constraints
$A = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}$, state dynamic matrix for the driving states of the controlled vehicle
$B = \begin{bmatrix} (\Delta t)^2/2 \\ \Delta t \end{bmatrix}$, state dynamic matrix for the acceleration of the controlled vehicle

2.2. Definitions

Before we present the problem statement, we define the following components in this paper.

Definition 1. Network-based path/route. It is a sequence of link/lane and nodes as a planned path from origin to destination (Fig. 1(a)).

Definition 2. Coordinate-based path/route. It is a sequence of waypoints (i.e., longitude, latitude, and road grade) of the center line of lanes along the network-based path (Fig. 1(b)). Noted that the waypoints of the coordinate-based path can capture the curve shape of lanes along the path.

Definition 3. Distance-based path/route. One can convert the coordinate-based path to a distance-based path through a mapping function f: $(x_w, y_w) \rightarrow (s_w), \forall w = 1, \ldots, W$ between the coordinate-based and distance-based positions of waypoints along the path in Fig. 1(c) (See Section 3.1).

Definition 4. HD map: A High-Definition (HD) Map includes detailed road geometry information on coordinates of left, center (i.e., the virtual track for vehicles to drive on a lane), and right lines for each lane. It also contains traffic control device information such as longitudinal markings (e.g., lane markings and edge markings), transverse markings (e.g., stop and yield lines), traffic signs, and traffic signal information (Roess et al., 2004).

Definition 5. Location-based traffic control devices and road geometry constraint. The constraint captures the interrupted flow facility locations from HD maps. For example, it can consider traffic control devices (e.g., road markings, stop

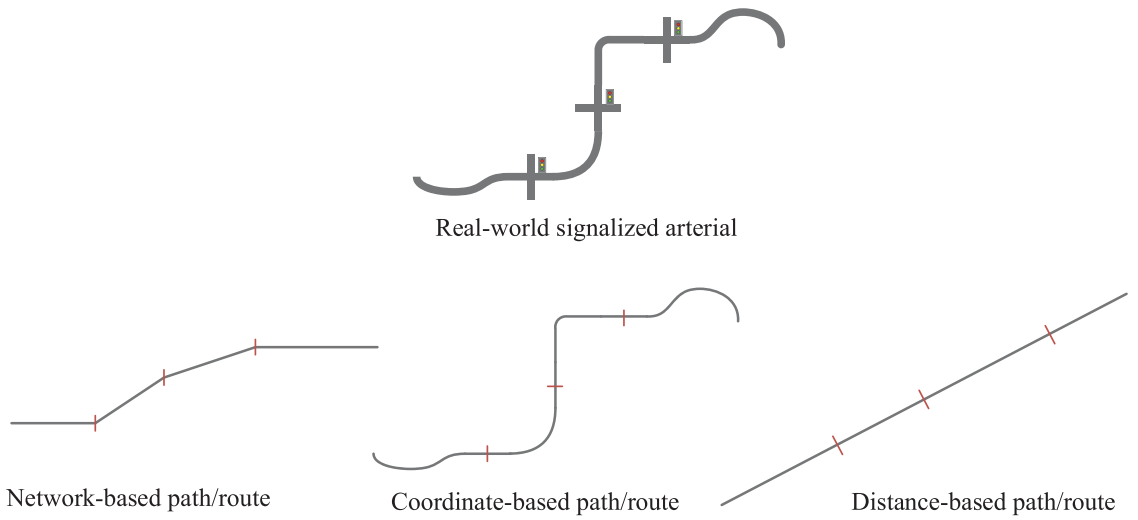

Fig. 1. Three types of representations of a path/route on a signalized arterial.

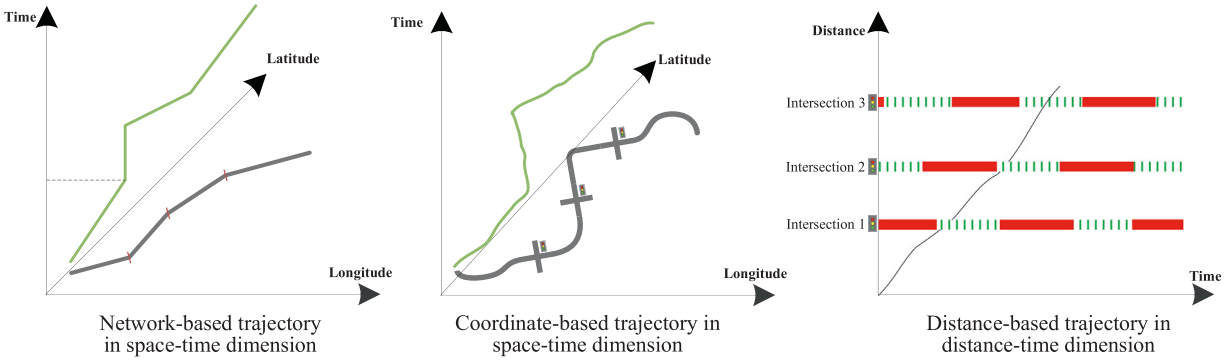

Fig. 2. Three types of representations of a trajectory on a signalized arterial.

lines, and traffic signs) and road geometry (e.g., road shapes and road grades) at a signalized intersection (e.g., the red stop lines in Fig. 1).

Definition 6. Link-based trajectory. It is a time-indexed trajectory that consists of a sequence of links with timestamps (i.e., link enter and exit timestamps) along a given network-based path (Fig. 2(a)).

Definition 7. Coordinate-based trajectory. It is a time-indexed trajectory that consists of a sequence of coordinates (longitude and latitude) with timestamps (e.g., 0.1 seconds) along a given coordinate-based path (Fig. 2(b)).

Definition 8. Distance-based trajectory. It is a time-indexed trajectory that consists of a sequence of distance-based positions with timestamps converted from the coordinate-based trajectory (Fig. 2 (c)).

Definition 9. The controlled vehicle. The controlled vehicle is the CAV to be real-time controlled by following the optimal driving commands by the proposed connected and automated driving system.

Definition 10. The boundary condition preceding vehicle. The boundary condition preceding vehicle is the first (farthest) preceding vehicle within the communication range of the controlled vehicle.

Definition 11. The immediately preceding vehicle (IPV). The IPV is the preceding vehicle just in front of the controlled vehicle.

Definition 12. The space-time region of observations. The space-time region of observations is the region that contains the observed driving states in space and time dimensions of the preceding vehicles and the controlled vehicle.

Definition 13. The space-time region of predictions. The space-time region of predictions is the region that contains the predicted driving states in space and time dimensions of the preceding vehicles and the controlled vehicle.

Definition 14. TheEco-driving policy. The Eco-driving policy is formulated as a constrained finite-horizon optimal control model to obtain an energy-efficient optimal trajectory as the reference for the predictive control model along a given path with road geometry information (e.g., road grade) for each prediction horizon.

2.3. Assumptions

To properly define our problem for the connected and automated online driving control on signalized arterials, we make the following assumptions:

Assumption 1. The controlled vehicle can obtain preceding vehicle information (e.g., positions, gaps, velocities, and accelerations) and SPaT information (e.g., the current signal status and remaining time before traffic signal change) in real-time (i.e., 0.1-second time interval) via V2V and V2I communications, respectively, within a short-range communication distance (e.g., 300 meters). Communication delays, packet loss, and other issues are not considered for simplicity.

Assumption 2. The network-based path/route is given and can be converted to a corresponding coordinate-based path. We can obtain the corresponding given distance-based path by a mapping function f: $(x_w,y_w) \rightarrow (s_w),\forall w = 1, \ldots, W$ between the coordinates and distance-based positions of the waypoints along the given path. The problem focuses on the driving dynamics of the controlled vehicle on a distance-based trajectory when the controlled vehicle drives on a signalized arterial.

Assumption 3. The predicted driving dynamics of the preceding vehicles are not given by these preceding vehicles or by traffic management centers.

Assumption 4. The distribution of stochastic parameters (i.e., the predictive vehicle position and velocity) are not known in advance.

Assumption 5. The controlled vehicle is equipped with an automated driving system with Level 2 automation (following a given path) to execute the received acceleration or deceleration commands (SAE, 2014).

Assumption 6. The interrupted flow facility locations with traffic control devices (e.g., road markings, traffic signs, and traffic signals) and road geometry (e.g., road shapes, road boundaries, and road grades) information are given from an HD map.

2.4. Problem of interest

The goal of this paper is to seek the optimal driving commands to improve the driving safety, energy efficiency, robustness, and driving comfort for the online connected and automated driving control of a CAV under uncertain traffic conditions on a signalized arterial by using real-time V2I/V2V data. The uncertain traffic conditions pose a challenge of predicting the driving states of the IPV in front of the controlled vehicle on signalized arterials. The problem of interest of this paper is to propose a robust and safe optimization-based control model for connected and automated driving on signalized arterials under uncertain traffic conditions. In this paper, we assume that the controlled vehicle is V2V and V2I enabled, and the preceding vehicles are V2V enabled. Within a communication range such as the Dedicated Short-Range Communications (DSRC) distance, preceding vehicle data such as position, velocity, and acceleration can be transmitted via V2V communications. In addition, Signal Phasing and Timing (SPaT) information such as the current signal state and the remaining time of traffic signal change can be shared by RSU through V2I communications (Fig. 3). The preceding vehicle data and the SPaT information will be collected by the controlled vehicle and sent to a mobile computing platform (which can be hosted in the controlled vehicle). By real-time solving the data-driven optimization-based control problem, the mobile computing platform in the controlled vehicle will determine the driving commands (i.e., the optimal acceleration or deceleration) for the controlled vehicle. The automated driving system in the controlled vehicle will execute the driving commands.

3. Methodology

This section presents mathematical formulations for an MPC modeling framework based on a DRSO-DRCC model for the online connected and automated driving control on signalized arterials under uncertain traffic conditions. The control model updates traffic control devices and road geometry information from HD maps at every control time step. The framework contains three steps for the online connected and automated driving control (Fig. 4). First, inspired by the concept of "forecasting the forecasts of others" in Townsend (1983), we online predict the driving dynamics of the IPV and the controlled vehicles at control time step k. We estimate the uncertain driving states of the boundary condition preceding vehicle as the boundary conditions. Constrained by the boundary conditions, we can predict the uncertain driving states of the IPV by learning the driving behaviors of the preceding vehicles. Restricted by predicted driving states of the IPV, we predict the driving state reference via and an Eco-driving policy for CAVs. Second, we solve the DRSO-DRCC model at control time step k to obtain the optimal acceleration or deceleration commands for the controlled vehicle. The predicted uncertain driving states of the controlled vehicle from the previous step will be used as the tracking reference of the DRSO-DRCC based MPC model. The predicted uncertain driving states of the IPV will be used in the DRCC safety constraints of the model. Third, the optimal acceleration or deceleration commands from the DRSO-DRCC based MPC model are sent to the controlled vehicle to execute at control time step k. Then, the three steps in this procedure will be repeated in the rolling horizon framework.

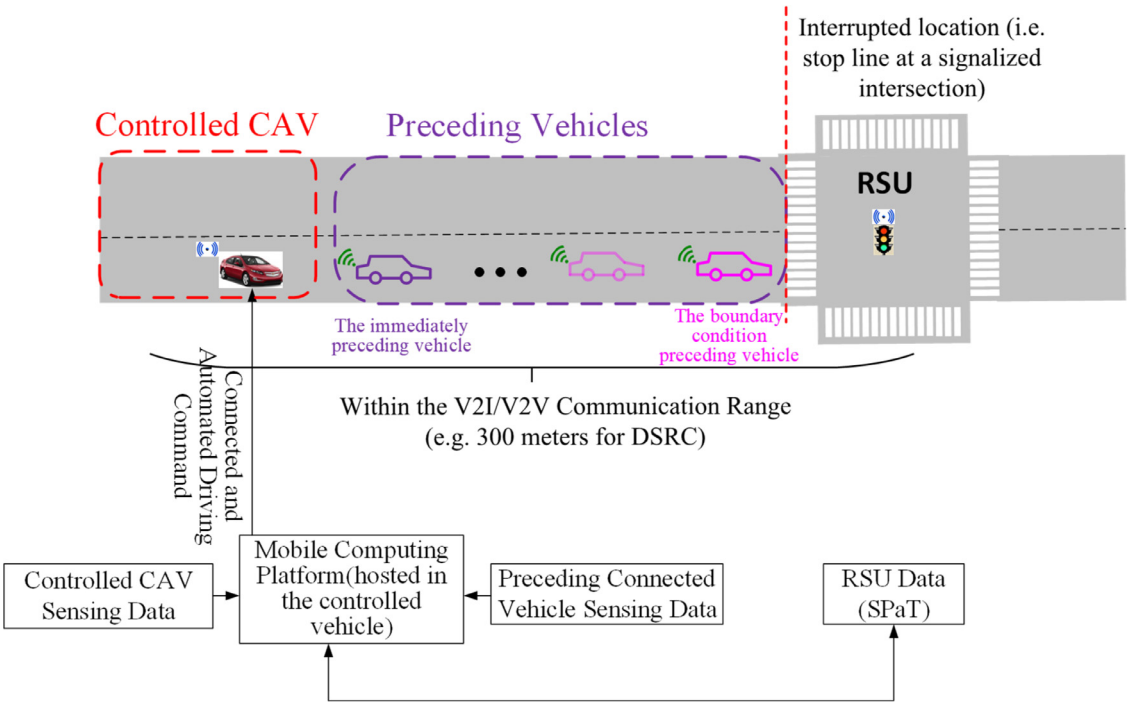

Fig. 3. The illustration of the connected and automated driving control problem on a signalized arterial.

3.1. Location-based traffic control devices and road geometry constraints

This section describes how to integrate location-based traffic control devices and road geometry constraints in the time-indexed MPC formulation. By assuming a given Lane Keeping function in CAVs, linear vehicle dynamics (without considering steering angles) can be used for connected and automated driving in the predictive control model. Within the predictive control model with linear vehicle dynamics, location-based traffic control devices and road geometry constraints can be formulated to consider interrupted flow facility locations such as traffic markings, traffic signs, traffic signals, and other special types of control (Roess et al., 2004).

It is challenging to directly formulate the location-based traffic control devices and road geometry constraints in a time-indexed MPC model due to the difference between the distance index and time index. A vehicle driving control model can be formulated as a time-indexed problem or a distance-indexed problem, depending on whether the indexing is done in the temporal domain or spatial domain (Karlsson et al., 2016). The key beneficial feature of a time-indexed formulation for MPC models is that the time-indexed solutions can be applied for the real-time applications at every predefined time interval as the automated driving commands. In the time-indexed model formulation, the state variables (vehicle location and velocity) are indexed by time steps (e.g., at the 0.1-second interval). However, the distance-indexed term uses the distance a variable subscript in the time-indexed formulation, which makes the problem not tractable in most cases (Hooker et al., 1999).

We propose two steps to mathematically formulate the location-based traffic control devices and road geometry constraints from the coordinate-based HD maps in a time-indexed predictive control model: (1) map the coordinate-based constraints from HD maps to the distance domain; and (2) mathematically formulate the location-based traffic control devices and road geometry constraints by adding binary variables. In Fig. 5, we take road grade selection as an example to show the two steps for integrating location-based traffic control devices and road geometry constraints with a control model.

3.1.1. Map the coordinate-based constraints from HD maps to the distance-based constraints

This section presents a method to map the coordinate-based constraints such as interrupted flow facility locations from HD maps to the distance-based constraints in predictive control models for connected and automated driving.

In Fig. 5, we define Δs as the homogeneous road geometry distance interval. In our problem, the coordinate-based path can be discretized into a set of points from the waypoints of the lanes along the given path. Δs is selected such that two adjacent points have only a small difference in road geometry information (e.g., the difference of the road grades of the two points are small). Along the coordinate-based path, the coordinates of the traffic control devices or the road geometry can be obtained from HD maps.

For the waypoints along the given coordinate-based path with a sequence of three-dimensional points $\{(x_w,y_w,g_w)\forall w = 1, \ldots, W\}$, we can calculate the corresponding distanced-based points from the start position of the path as: $\{(s_w,g_w)\forall w = 1,$

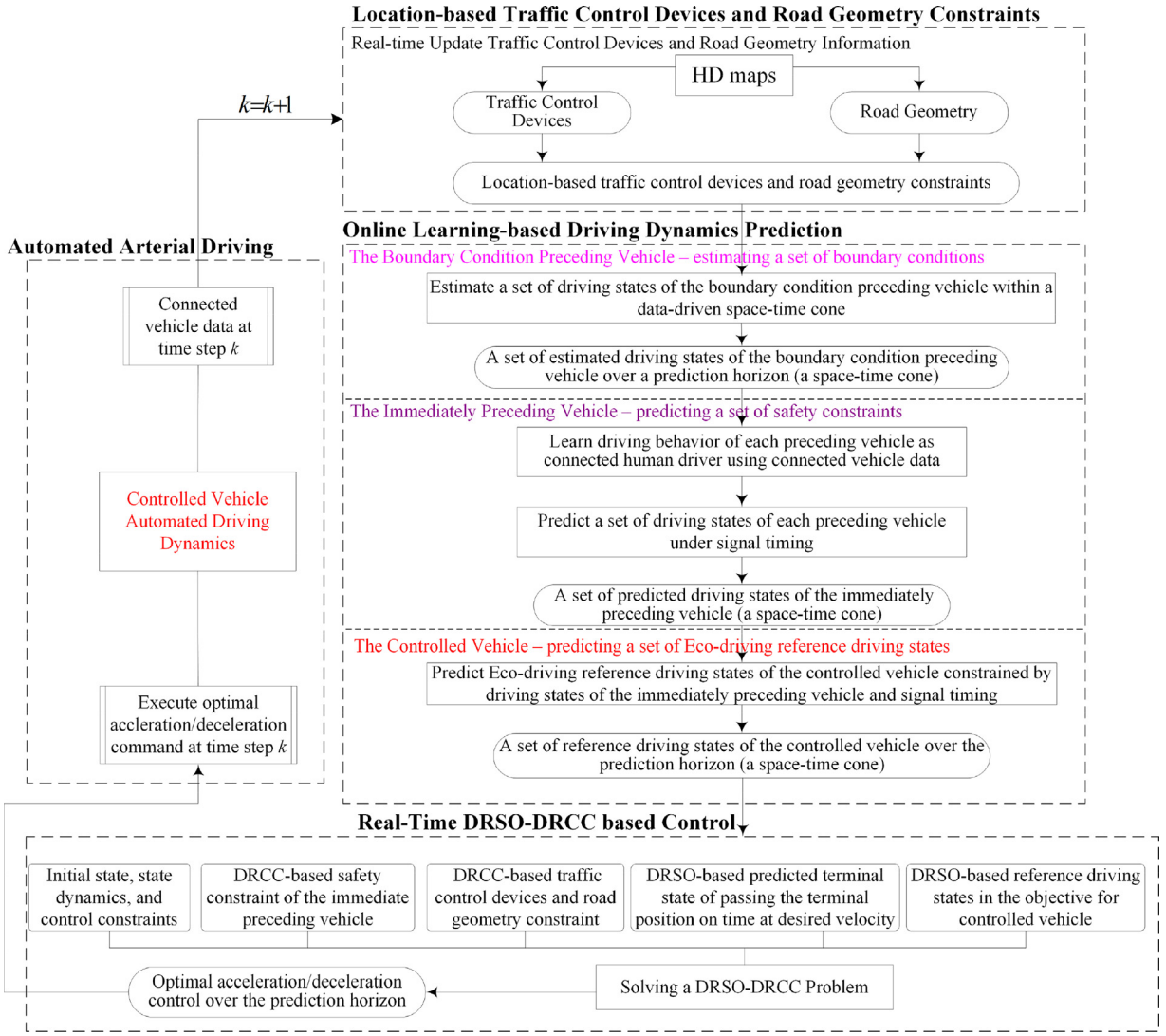

Fig. 4. The data-driven optimization-based MPC framework for connected and automated driving on signalized arterials.

..., W} using the mapping function f: $(x_w,y_w) \rightarrow (s_w)$,$\forall w = 1, ..., W$ in Eq.(1).

$$s_w = \sum_{i=0}^{w-1} \boldsymbol{d}\{(x_i, y_i), (x_{i+1}, y_{i+1})\} \tag{1}$$

where (x_0,y_0) denotes the longitude and latitude of the origin location. Using Eq. (1), we calculate a sequence of distance-based positions as the distance-based path while the traffic control devices and road geometry information are still contained (e.g., the road grade data g is still known as related to distance s in Fig. 5(1)). With the distance-based path, we need to mathematically formulate the location-based traffic control devices, and road geometry constraints from the coordinate-based interrupted flow facility locations and road geometry information along the path.

3.1.2. Mathematically formulation of the location-based traffic control devices and road geometry constraints

Given the mapped distance-based path with traffic control devices and road geometry information, this section shows the step to mathematically formulate the location-based traffic control devices road geometry constraints.

We cannot directly formulate the traffic control devices road geometry constraints with distance indices in our problem because the distance indices are variables in the time-indexed predictive control model. Therefore, we add binary variables to replace the variable indices of vehicle longitudinal locations (see the Travelling salesman problem, Focacci et al., 1999; Ottosson and Thorsteinsson, 2000). We define the distance-indexed road grade data $\theta(l(t)_d)$ in the objective function that

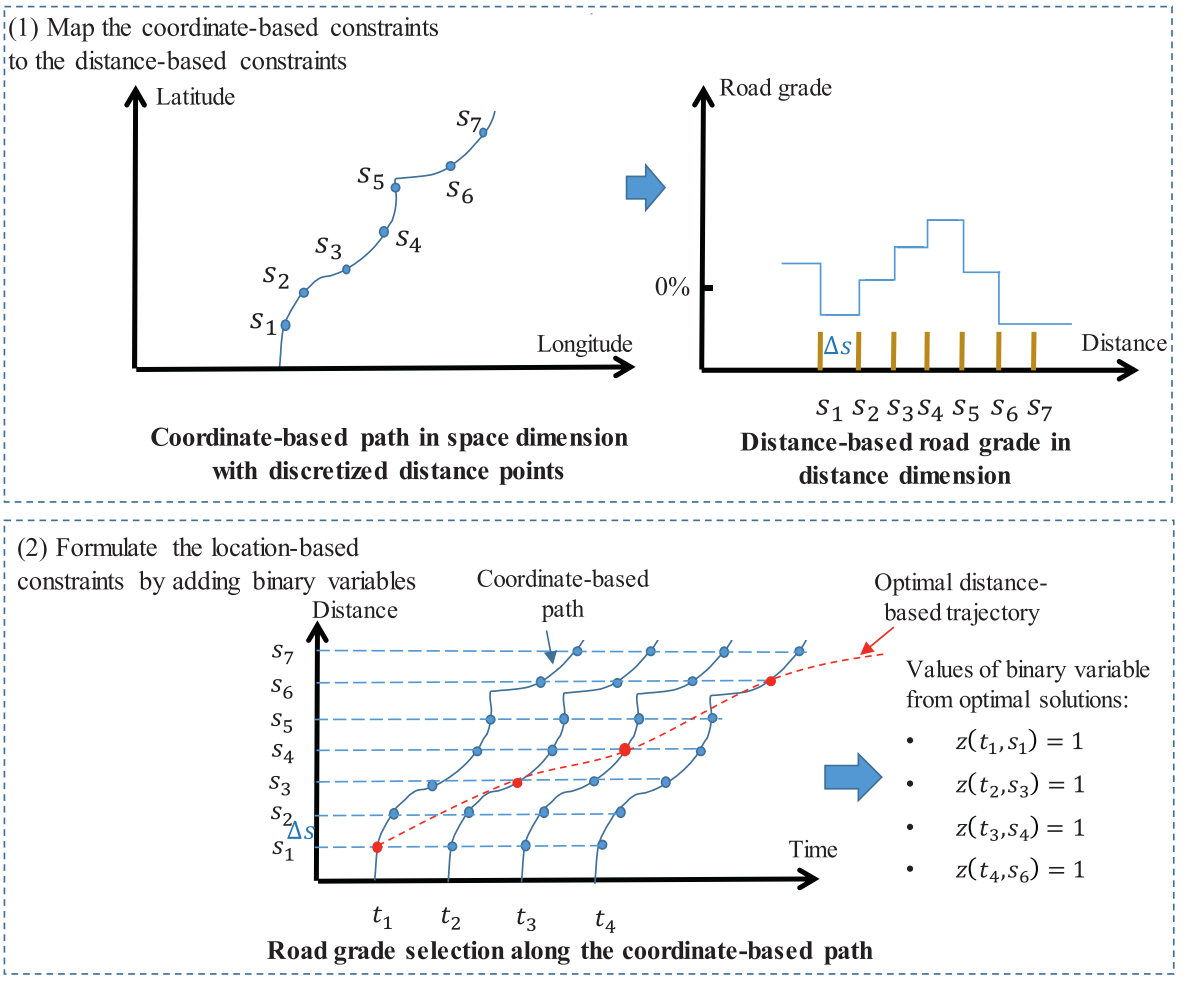

Fig. 5. Two steps to formulate the distance-based road grade selection along a coordinate-based path from HD maps.

minimizes the energy consumption in the Eco-driving reference model Eqs. (4)-(13). The $\lfloor \bullet \rfloor_d$ is the rounding operator to d trailing digits (e.g. $\lfloor 6.581 \rfloor_1 = 6.6$).

We assume that we can obtain the discretized road grade data $\{\theta(s)|s = 1, .., S\}$ (indexed by distance with the homogeneous distance interval Δs) from the given coordinate-based path. Given the road grade data, we define a binary variable $\{z(t, s)|\forall t = k + 1, ..., k + PH(k), \forall s = 1, ..., S\}$ indexed by time and distance. Then, to apply the distance-indexed road grade data in the time-indexed objective function at time step t, we define the following constraints in Eq. (2).

$$\begin{aligned} &\sum_{s=1}^{S} z(t,s) = 1, \forall t = k, \ldots, PH(t) \\ &l(t) \geq \sum_{s=1}^{S} z(t,s) \times (l-1), \forall t = k, \ldots, k + PH(k) \\ &l(t) \leq \sum_{s=1}^{S} z(t,s) \times l, \forall t = k, \ldots, k + PH(k) \end{aligned} \tag{2}$$

The first constraint is to guarantee that only one of the road grade data $\{\theta(s)|s = 1, .., S\}$ is selected at each time step t. The second and third constraints are applied to select the road grade by using the vehicle longitudinal location variable $l(t)$ at time t. With these three constraints, we can replace $\theta(\lfloor l(t) \rfloor)$ with $\sum_{s=1}^{S} \theta(s) \times z(t, s)$ and make sure that only one road grade is selected in the objective function at time step t. In Fig. 5(2), the red dots indicate that $z(t_1,s_1) = 1$, $z(t_2,s_3) = 1$, $z(t_3,s_4) = 1$, and $z(t_4,s_6) = 1$, which implies that only one road grade at a longitudinal location is selected at each time step from the optimal solution of our model. Noted that the sequence of the red dots from the optimal solution consists of the determined vehicle trajectories.

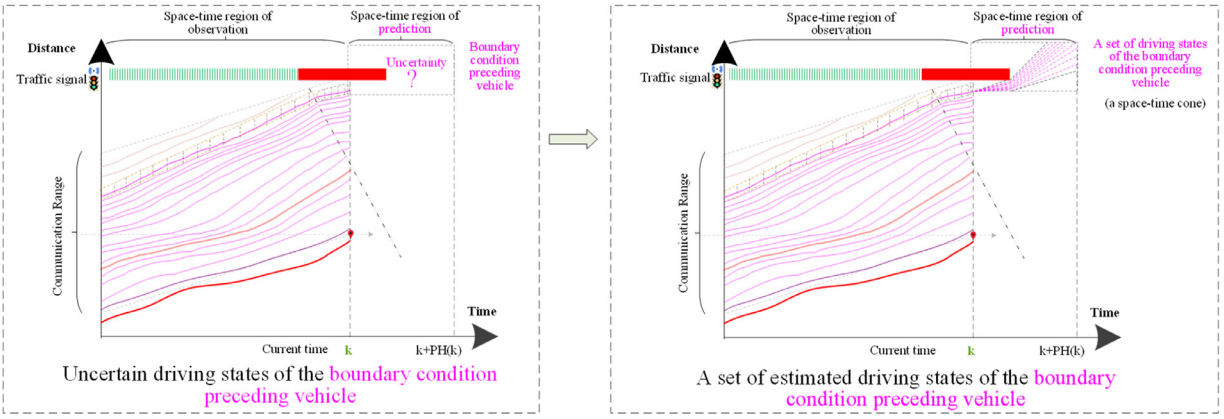

Fig. 6. Estimate a set of driving states (a space-time cone) of the boundary condition preceding vehicle.

3.2. Online learning-based driving dynamics prediction under traffic uncertainties

This section proposes an online method to predict the uncertain driving states of the preceding vehicles and the controlled vehicle over a prediction horizon. Existing trajectory-planning models assume a given feasible vehicle trajectory set as a space-time cone or prism by allowing finite acceleration and deceleration rates (Ma et al., 2017; Zhou et al., 2017). The predictive vehicle driving control models assume that predictive front traffic and traffic signal information can be given by the traffic management center via V2I communications (Zulkefli et al., 2014). However, the predicted information of positions and velocities of the preceding vehicles over a prediction horizon are generally not given in practice. The predictive driving state reference for the controlled vehicle is constrained by the predicted driving states of the preceding vehicles, which may not be given in advance.

We present an online learning-based driving dynamics prediction method, based on the concept of "forecasting the forecasts of others" in Townsend (1983), to predict the driving states of the preceding vehicles. For arterial driving, the driving states of the controlled vehicle are directly affected by its IPV, in addition to traffic signal control. Based on the predicted boundary conditions, we predict the uncertain driving states of the IPV by learning the driving behaviors of other preceding vehicles. By solving a constrained finite-horizon optimal control model (constrained by the IPV), we predict the driving states of the controlled vehicle over a prediction horizon.

3.2.1. Predictive uncertain driving states of the boundary condition preceding vehicle

This section shows the method to estimate a set of uncertain driving states for the boundary condition vehicle within a valid communication range (e.g., a 300-meter DSRC range) of the controlled vehicle. Different from the analytical method to construct a feasible region of possible vehicle trajectories as a space-time cone in Ma et al. (2017) and Zhou et al. (2017), we estimate a set of uncertain driving states of the boundary condition preceding vehicle over a prediction horizon as a set of trajectories, which is also a space-time cone (using historical observations rather by using analytical methods). Noted that one feature of MPC models is that they just implement the first-step optimal control over a prediction horizon [k, ..., $k + PH(k)$] at the current control time step k. This feature makes it possible to estimate the driving dynamics over a prediction horizon by simply using a constant velocity and update during every prediction horizon. Even by assuming a constant velocity of the boundary condition preceding vehicle, the traffic oscillations (i.e., the shocks in Fig. 6) will still be captured by the preceding vehicle prediction model in the next subsection.

At control time step k, we collect velocity data from the historical data during [k, T^{obs},k] (e.g., over the past five or ten seconds) to obtain a set of observed vehicle velocities $\{\hat{v}_b^{\xi_i} \forall i = 1, \ldots, M\}$ in Fig. 6. Based on the sampled velocity observations, we calculate a set of driving states - $\{[\hat{l}_b^{\xi_i}(t), \hat{v}_b^{\xi_i}(t)]^T \forall t = k, \ldots, k + PH(k), \forall i = 1, \ldots, M\}$ of the boundary condition preceding vehicle driving at constant velocities. The set of driving states, which is a space-time cone in space and time dimensions, is considered as a set of boundary conditions for the prediction of the driving states of the other preceding vehicles. The upper bound of the set of the boundary condition preceding vehicle driving states is calculated by speed limits while the lower bound is estimated based on a stopped velocity. From the collected velocity samples at time step k, we can calculate the number of observations $\{N_i i = 1, ..., M\}$ for each of the driving states in the set $\{[\hat{l}_b^{\xi_i}(t), \hat{v}_b^{\xi_i}(t)]^T \forall t = k, \ldots, k + PH(k), \forall i = 1, \ldots, M\}$. The number of observations and the estimated set of driving states of the boundary condition preceding vehicle will be used to construct the data-driven uncertainty set based on moment information.

3.2.2. Learning-based predictive uncertain driving states of the IPV

This section presents a learning-based method to predict a set of uncertain driving states of the IPV, given the set of boundary driving conditions from Section 3.2.1. The prediction of the controlled vehicle is greatly affected by the predicted driving states of the IPV. As Fig. 7 shows, given a set of boundary conditions, we online predict the heterogeneous driving dynamics of each preceding vehicle using V2V connected vehicle data. The online learning method dynamically calibrates car-following models for connected human-driven vehicles, and closed-form connected and automated driving models for CAVs. Specifically, we use the connected IDM model in Talebpour and Mahmassani (2016) for preceding vehicles, which utilizes connected vehicle data for driving assistance applications (e.g., safety warning) based on the original IDM model in Kesting et al. (2010) for CAVs. Based on the boundary driving conditions and the online calibrated model parameters of each preceding vehicle, we predict a set of uncertain driving states for all the preceding vehicles over a prediction horizon. Then, we will obtain a set of learning-based predictive uncertain driving states (e.g., the right plot in Fig. 7) of the IPV for a given set of boundary conditions using the observed V2V data.

To learn the heterogeneous driving behaviors of preceding vehicles, we online solve an optimization model to calibrate car-following model parameters. The objective function in Eq. (3) is to obtain the optimal model parameters using observations over a past time horizon.

$$\min_{\varphi(n,k)} \sqrt{\frac{\sum_{t=k-T^{obs}}^{k}\left[v_n(t)-v_n^{CF}(t)\right]^2}{T^{obs}}}, n=1,\ldots,N \tag{3}$$

where $v_n^{CF}(t)$ is the velocity calculated by car-following models of trajectory n at time t (N is the total number of trajectories). We calibrate the model parameters independently for each preceding vehicle.

By car-following models such as connected IDM model, we can obtain the predicted driving states of the IPV for each given boundary condition. The connected IDM model is used to model driving behaviors in a connected environment where drivers are aware of other connected drivers' behaviors:

$$a_{IDM}^{n}(s_n, v_n, \Delta v_n) = \bar{a}_n\left[1-\left(\frac{v_n}{v_0^n}\right)^{\delta_n}-\left(\frac{s^*(v_n,\Delta v_n)}{s_n}\right)^2\right]$$

$$s^*(v_n,\Delta v_n) = s_0^n + T_n v_n + \frac{v_n \Delta v_n}{2\sqrt{\bar{a}_n \bar{b}_n}}$$

where v_0^n is desired speed, δ_n is the free acceleration exponent, s_0^n is the jam density. Even though IDM does not explicitly express the reaction time, it includes T_n as the desired time gap. Unlike the traditional IDM model in Treiber et al. (2000), the connected IDM model considers the average values of maximum acceleration and the desired deceleration of its nearest preceding vehicles (Treiber et al., 2007; Talebpour and Mahmassani, 2016). The predicted driving states of the IPV will be the input for the Eco-driving reference state prediction model in the next subsection.

Noted that one can use any specific model to predict the driving dynamics of the preceding vehicles over each prediction horizon in the proposed DRSO-based MPC framework since the framework uses a set of predictions to address the uncertainty. The uncertainty set will be tighter when a more precise prediction model is employed. As a general form of an MPC framework, our proposed DRSO-based MPC can use vehicle driving prediction methods such as the connected IDM (Talebpour and Mahmassani, 2016) for connected human-driven vehicles and machine learning models for automated vehicles (Lin et al., 2019; Bouton et al., 2019). However, to precisely predict driving trajectories of automated vehicles is still not trivial, which requires further studies in the future. In this paper, without loss of generality, we will focus on the case that preceding vehicles are connected human-driven vehicles (HDV) and select connected IDM as the preceding vehicle driving dynamics prediction model.

3.2.3. An Eco-driving policy to obtain the predictive reference driving states for the controlled vehicle

This section presents a constrained finite-horizon optimal control model as the Eco-driving policy to obtain a set of predictive driving reference for the controlled vehicle, given the set of IPV driving states from Section 3.2.2. With this Eco-driving reference as the reference driving states for the controlled vehicle, the DRSO-DRCC based control model in Section 3.3 is able to improve the energy efficiency by tracking the reference driving states.

As Fig. 9 shows, we can obtain a set of reference driving states of the controlled vehicle through an Eco-driving policy based on a given set of predicted driving states of the boundary condition preceding vehicle. The Eco-driving policy can be the existing velocity trajectory optimization models such as in He et al. (2015) and Ma et al. (2017). In this paper, we propose a constrained finite-horizon optimal control model to obtain the driving reference that explicitly considers the IPV driving and traffic control devices and road geometry constraints in Section 3.1. We also include an energy consumption term in the objective function to improve energy efficiency. Given any predicted driving states of the IPV, we predict the driving state reference of the controlled vehicle over a prediction horizon based on the Eco-driving policy (the middle plot in Fig. 9). In this paper, we formulate a constrained finite-horizon optimal control model as the Eco-driving policy. First, we

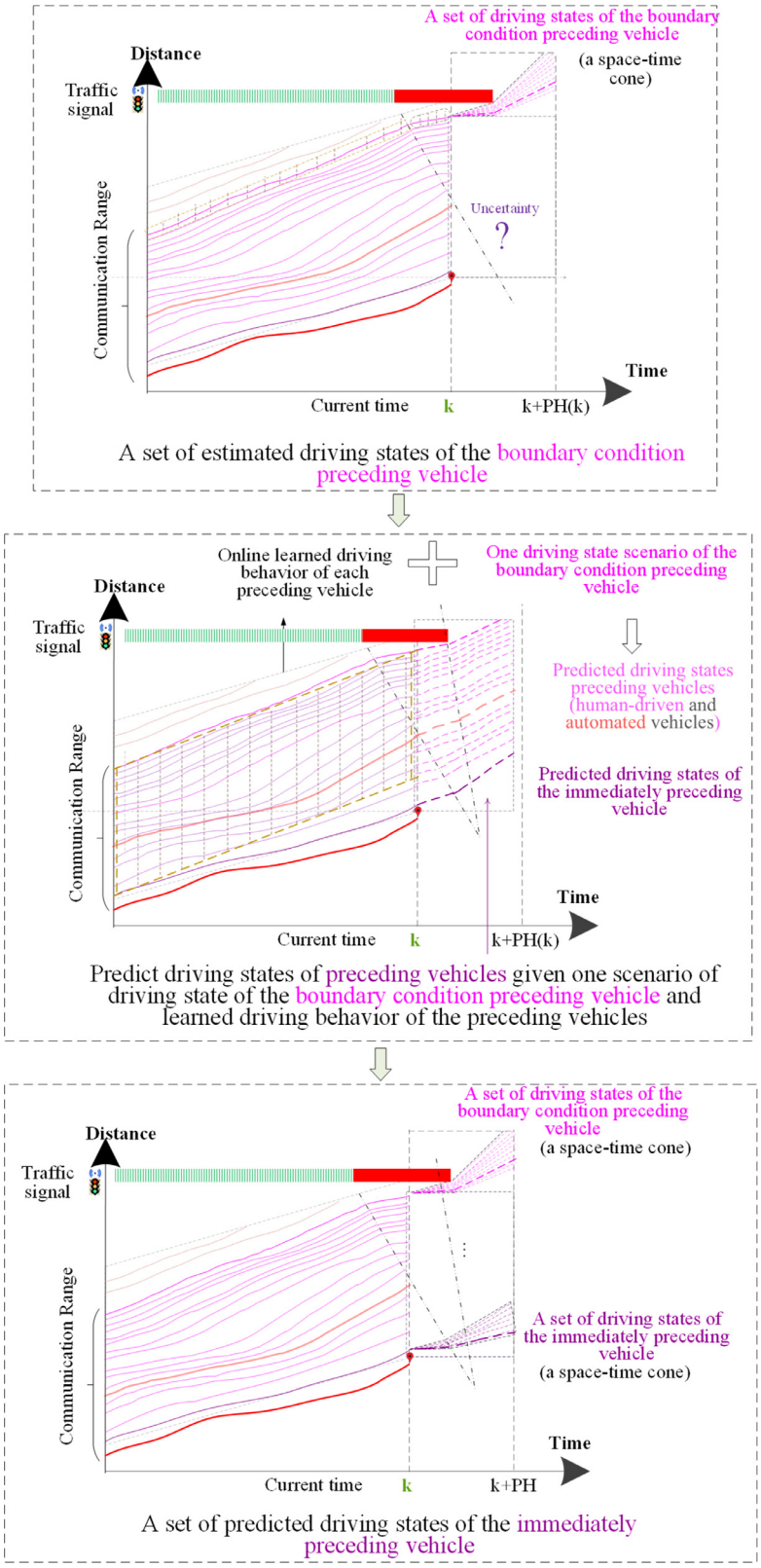

Fig. 7. Predict a set of driving states for the IPV.

define the objective function in Eq. (4).

$$\min_{u} \left\{ \begin{array}{c} \alpha \times \sum_{t=k}^{k+PH(k)} E[v(t), u(t), \theta(l(t))] \\ (1-\alpha) \times \sum_{t=k}^{k+PH(k)} \left\{ \left[\hat{l}_s^{\xi}(t) - l(t) - \tau^{min} \times v(t)\right]^2 + \left[\hat{v}_s^{\xi}(t) - v(t)\right]^2 + [u(t)]^2 \right\} \end{array} \right\} \tag{4}$$

In the objective function, $E[v(t), u(t), \theta(l(t))]$ minimizes the energy consumption by utilizing energy consumption such as the Comprehensive Modal Emissions Model (CMEM) in Barth et al. (2000) and Passenger Car and Heavy-duty Emission Model (PHEM) in Hausberger et al. (2002). The second term regulates the controlled vehicle by target location and velocity constrained by the IPV, interrupted flow facility locations, and current traffic signal timing state. The objective also minimizes the acceleration fluctuation for driving comfort.

By adding binary variable $z(t, s)$ as discussed in Section 3.1.2, the objective function can be reformulated as in Eq. (5).

$$\min_{u} \left\{ \begin{array}{c} \alpha \times \sum_{t=k}^{k+PH(k)} E\left[v(t), u(t), \sum_{s=1}^{S} \theta(s) \times z(t,s)\right] \\ (1-\alpha) \times \sum_{t=k}^{k+PH(k)} \left\{ \left[\hat{l}_s^{\xi}(t) - l(t) - \tau^{min} \times v(t)\right]^2 + \left[\hat{v}_s^{\xi}(t) - v(t)\right]^2 + [u(t)]^2 \right\} \end{array} \right\} \tag{5}$$

s.t.

$$PH(k) \le \hat{t}_f(k) \tag{6}$$

$$\hat{l}_s^{\xi}(t) - l(t) \ge \tau^{\min} \times v(t) + s^{jam}, t = k, \ldots, k + PH(k) \tag{7}$$

$$\mathrm{l(t)} = \mathrm{l^{obs}(k)}, \mathrm{v(t)} = \mathrm{l^{obs}(k)}, \mathrm{t} = \mathrm{k} \tag{8}$$

$$l(t+1) = l(t) + \Delta t \times v(t) + \frac{1}{2}\Delta t^2 \times u(t), \ \ \forall t = k, \ldots, k + PH(k) \tag{9}$$

$$\mathrm{v(t+1)} = \mathrm{v(t)} + \Delta \mathrm{t} \times \mathrm{u(t)}, \ \ \forall \mathrm{t} = \mathrm{k}, \ldots, \mathrm{k} + \mathrm{PH(k)} \tag{10}$$

$$u^{\min} \le u(t) \le u^{\max}, \forall t = k, \ldots, k + PH(k) \tag{11}$$

$$0 \le v(t) \le v^{\lim it}(t), \forall t = k, \ldots, k + PH(k) \tag{12}$$

$$\begin{array}{l} \sum_{s=1}^{S} z(t,s) = 1, \forall t = k, \ldots, PH(k) \\ l(t) \ge \sum_{s=1}^{S} z(t,s) \times (s-1), \forall t = k, \ldots, k + PH(k) \\ l(t) \le \sum_{s=1}^{S} z(t,s) \times s, \forall t = k, \ldots, k + PH(k) \end{array} \tag{13}$$

where $\hat{l}_s^{\xi}(t)$ and $\hat{v}_s^{\xi}(t)$ are the maximum safe driving location and velocity by considering the impacts of preceding vehicle locations and traffic control devices such as the traffic signal. ξ denotes the uncertainty caused by uncertain traffic conditions over the prediction horizon.

The energy term is defined as: $E[v(t), u(t), \theta(l(t))] = [M \times u(t) + M \times g \times \sin\theta + \frac{1}{2}C_d \times A \times \rho \times v(t) + M \times g \times C_r \times \cos\theta] \times v(t)/1000$. where $u(t)$ is vehicle acceleration, $v(t)$ is vehicle velocity, g is gravitational constant, θ is the road grade angle, which is related to the current location $l(t)$ of the vehicle at time t, ρ is air density (kg/m^3), C_d is the constant coefficient of drag (typically 0.7),A is the frontal surface area of the vehicle (m^2), C_r is the constant coefficient of rolling resistance, which is typically 0.01 (Weeks, 2009; Demir et al., 2011).

We define $\hat{l}_s^{\xi}(t) = \min\{p^{int}(\lfloor l(t)_d \rfloor) + s^{min}, \hat{l}_0^{\xi}(t) - L^{veh}\}$ as the maximum safe driving location in this constraint, where$p^{int}(\lfloor l(t) \rfloor_d)$is the distance-indexed interrupted location. $\hat{l}_s^{\xi}(t)$ should be smaller than the safe stop location, which depends on the interrupted location $p^{int}(\lfloor l(t) \rfloor_d)$ (e.g., a stop line at an intersection in case 1 of Fig. 8) and a minimum safe spacing s^{min} for stopped vehicles. $\hat{l}_s^{\xi}(t)$ should also be constrained by the IPV as $\hat{l}_s^{\xi}(t) - L^{veh}$(e.g. case 2 of Fig. 8). We define$\hat{v}_s^{\xi}(t) = \min(v^{limit}(\lfloor l(t) \rfloor_d), \hat{v}_0^{\xi}(t))$as the safe driving velocity. $\hat{l}_0^{\xi}(t)$ and $\hat{v}_0^{\xi}(t)$ are the predicted location and velocity of the IPV for the predicted driving states of the boundary condition preceding vehicle. To apply the location-based terms $p^{int}(\lfloor l(t) \rfloor_d)$ and $v^{limit}(\lfloor l(t)_d \rfloor)$ to the time-indexed MPC formulation, we can add two binary variables $w(t, s)$ and $m(t, s)$ respectively, with corresponding binary variable constraints (see the steps in Section 3.1.2).

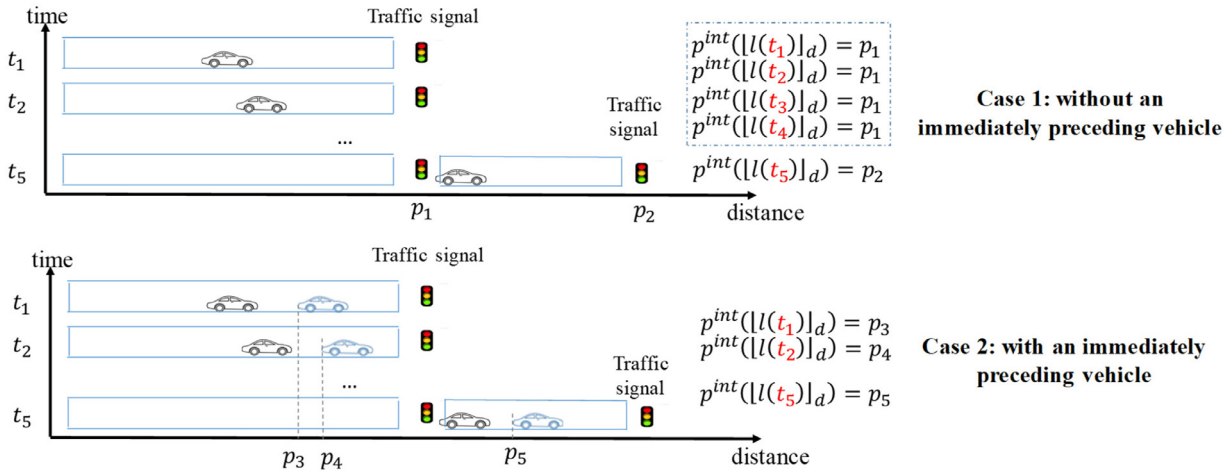

Fig. 8. Two cases of the maximum safe driving location.

Eq. (6) ensure the terminal state constraint by online estimate the arrival time $\hat{t}_f(k)$ at stop lines when the vehicle receives a red signal from SPaT information (see the estimation of $\hat{t}_f(k)$ in Eq. (14)). The location-based traffic control devices and road geometry constraints in Eq. (7) ensures safety by a given time headway τ^{min}. The constraint in Eq. (8) is an initial condition. Eqs. (9)-(10) are longitudinal location and velocity dynamics constraints, respectively. Eqs. (11)-(12) are the acceleration and velocity capability constraints, respectively. Eq. (13) is to ensure the road grade selection after adding the binary variable $z(t, s)$.

Different from existing predictive control models of connected and automated driving, we define a time-variant prediction horizon length $PH(k) \leq \hat{t}_f(k)$ for Eqs. (5)-(13) where $\hat{t}_f(k)$ is the estimated arrival time at the stop line of a signalized intersection. $\hat{t}_f(k)$ is calculated by considering the current traffic signal state and front vehicle trajectories as in Eq. (14).

$$\begin{aligned}
\hat{t}_{f,1}(k) &= \max\left(\begin{matrix} t_{f,p}(k) + \tau^h, \\ t^0 + \frac{L}{v^{limit}} - \left[\frac{0.5\left(v^{limit}+v(t^0)\right)\left(v^{limit}-v(t^0)\right)}{u^{max}\times v^{limit}}\right] + \frac{v^{limit}-v(t^0)}{u^{max}} \end{matrix}\right) \\
\hat{t}_{f,2}(k) &= \left\lfloor \frac{\hat{t}_{f,1}}{R(k)+G(k)} \right\rfloor (R(k)+G(k)) + R(k) \\
\hat{t}_f(k) &= \left\{ \begin{matrix} \hat{t}_{f,1}(k), \forall \hat{t}_{f,1} \in \left\{ t | \hat{t}_{f,1} \leq t_{r,a}, \hat{t}_{f,1} \geq t_{r,b} \right\} \\ \hat{t}_{f,2}(k), \forall \hat{t}_{f,1} \notin \left\{ t | t_{r,a} \leq \hat{t}_{f,1} \leq t_{r,b} \right\} \end{matrix} \right\}
\end{aligned} \tag{14}$$

where $\hat{t}_{f,1}(k)$ is the estimated arrival time during a green signal. The first term in the maximum operator of $\hat{t}_{f,1}(k)$ estimates the arrival time by using IPV estimated arrival time and a given following time headway. The second term in the maximum operator of $\hat{t}_{f,1}(k)$ estimates the arrival time by considering maximum acceleration and maximum velocity to a given velocity limit. $\hat{t}_{f,2}(k)$ is the estimated arrival time considering the current red signal. With Eq. (14), we can estimate a minimum prediction horizon length at control step k.

By solving Eqs. (5)–(14) for each boundary driving condition, the set of optimal predicted driving states - $\hat{l}_r^{\xi}(t)$ and $\hat{v}_r^{\xi}(t)$ will be obtained as the reference states in the tracking term of the DRSO-DRCC based MPC model for driving commands. Based on the optimal value of $\hat{l}_r^{\xi}(t)$ and $\hat{v}_r^{\xi}(t)$, we can calculate the corresponding $\hat{l}_s^{\xi}(t)$ and $\hat{v}_s^{\xi}(t)$ and construct the data-driven uncertainty set for the DRSO-DRCC model.

3.3. A DRSO-DRCC based MPC model for solving the connected and automated arterial driving problem under uncertain traffic conditions

In this section, we formulate the connected and automated driving on signalized arterials as a DRSO-DRCC model in Eqs. (15)-(20). The set of driving states from the Eco-driving policy will be used as reference states for the DRSO-based MPC model. Noted that in the DRSO objective function, the energy term is not considered because it is difficult to solve a DRSO model with the complex nonlinear terms such as the sine term and cubic term in the CMEM energy consumption model (Floudas and Visweswaran, 1995; Ye and Zhang, 2003; Luo et al., 2010). As we will have an Eco-driving reference as the tracking reference in the DRSO-DRCC based MPC objective Eq. (15) for the controlled vehicle, the optimal solution (i.e., optimal acceleration or deceleration) of the DRSO-DRCC model is still able to reduce energy consumption.

The DRSO model allows the incorporation of historical observations and the forecasting of the driving states. At time step k, we predict the driving states of the IPV and controlled vehicle over the prediction horizon by the steps in the previous

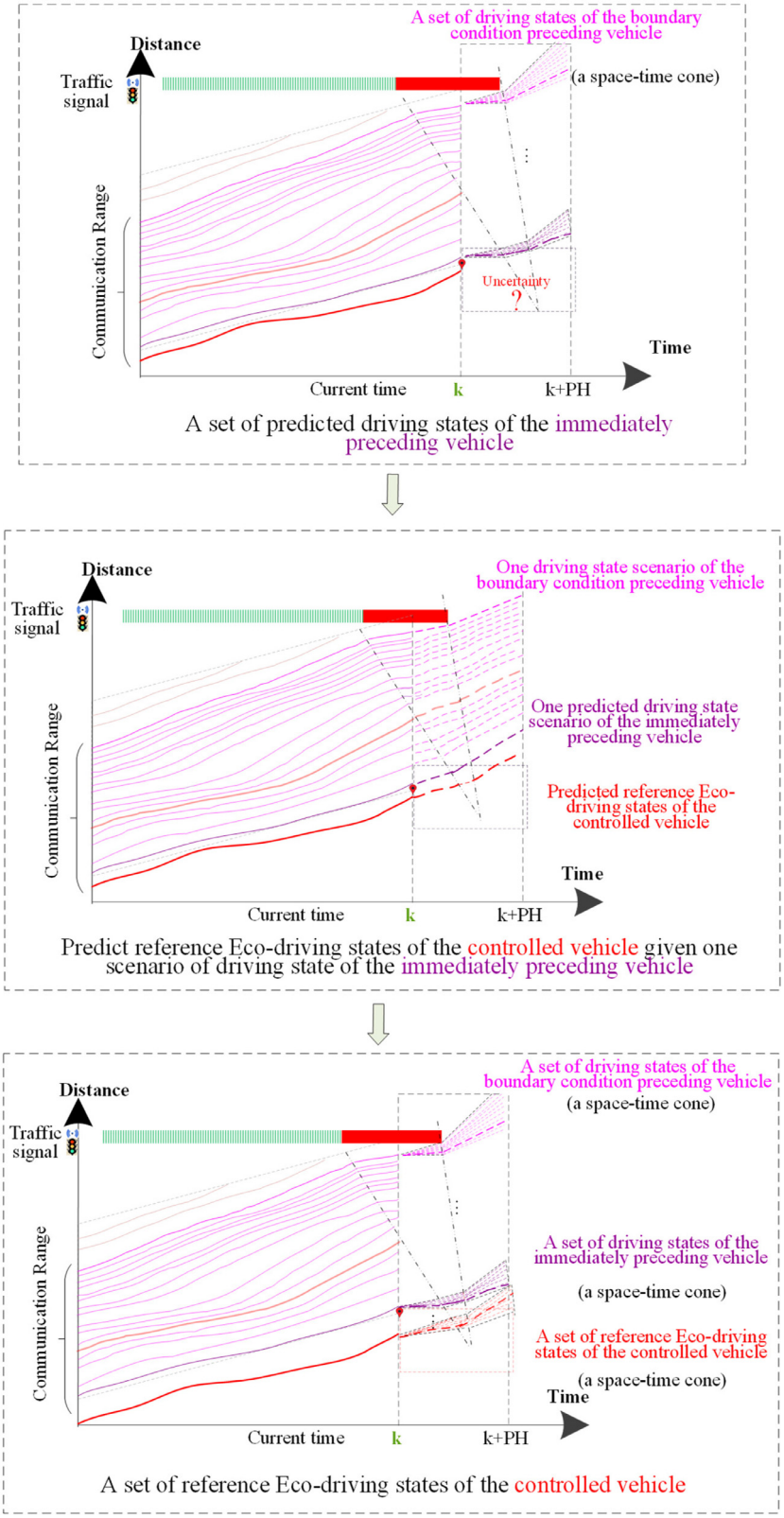

Fig. 9. Predict a set of reference driving states for the controlled vehicle.

subsection. The predicted driving states will be used as the input for the DRSO-DRCC model. Then, we solve the DRSO-DRCC model to determine the optimal acceleration or deceleration commands for the controlled vehicle over the prediction horizon.

$$\min_{u}\left\{\begin{array}{l}\max_{F_p\in\mathbb{D}}\mathbb{E}_{F_p}\left[\sum_{t=k}^{k+PH(k)}\left\{\left[\hat{l}_s^{\xi}(t)-l(t)-\tau^{\min}\times v(t)\right]^2+\left[\hat{v}_s^{\xi}(t)-v(t)\right]^2+[u(t)]^2\right\}\right]\\ \max_{F_p\in\mathbb{D}}\mathbb{E}_{F_r}\left[\sum_{t=k}^{k+PH(k)-1}\left[l(t)-\hat{l}_r^{\xi}(t)\right]^2+\left[v(t)-\hat{v}_r^{\xi}(t)\right]^2\right]\end{array}\right\} \tag{15}$$

s.t.

$$\inf_{F_p\in\mathbb{D}_p}\mathbb{P}_{F_p}\left(\hat{l}_s^{\xi}(t)-l(t)\geq s^{jam}\right)\geq\gamma_1, t\in\{k,\ldots,k+PH(k)|t_{r,a}(k)\leq t\leq t_{r,b}(k)\} \tag{16}$$

$$\inf_{F_p\in\mathbb{D}_p}\mathbb{P}_{F_p}\left(\hat{l}_s^{\xi}(t)-l(t)\geq\tau^{\min}\times v(t)+s^{jam}\right)\geq\gamma_2,\forall t=k,\ldots,k+PH(k) \tag{17}$$

$$l(t)=l^{obs}(k),\ v(t)=l^{obs}(k),\ t=k \tag{18}$$

$$\begin{aligned}&l(t+1)=l(t)+\Delta t\times v(t)+\frac{1}{2}\Delta t^2\times u(t),\ \forall t=k,\ldots,k+PH(k)\\&v(t+1)=v(t)+\Delta t\times u(t),\ \forall t=k,\ldots,k+PH(k)\end{aligned} \tag{19}$$

$$u^{\min}\leq u(t)\leq u^{\max}, 0\leq v(t)\leq v^{\lim it}(t),\forall t=k,\ldots,k+PH(k) \tag{20}$$

where $u(t)$ is the decision variable, and $l(t)$ and $v(t)$ are the state variables. $\hat{l}_r^{\xi}(t)$ and $\hat{v}_r^{\xi}(t)$ are the random parameters, which can be represented as $\hat{\boldsymbol{x}}_r^{\xi}(t)=[\hat{l}_r^{\xi}(t),\hat{v}_r^{\xi}(t)]_{1\times 2}^{T}$. In DRSO models, we can estimate the support of $\hat{\boldsymbol{x}}_r^{\xi}(t)$ from observations such that the random parameters take values in the set of observations $\Xi=\{[\hat{l}_r^{\xi_i}(t)\quad\hat{v}_r^{\xi_i}(t)]^T\forall i=1,\ldots,M\}$ with M samples, i.e., a set of predicted uncertain driving states constrained by a set of boundary driving conditions. $\hat{l}_r^{\xi_i}(t)$ and $\hat{v}_r^{\xi_i}(t)$ are the predicted driving states using historical observations (driving condition i).

The DRSO objective function minimizes the worst-case tracking deviation between the controlled vehicle driving states and the predicted Eco-driving reference. Eq. (16) is the DRCC safety constraint to ensure driving safety when the traffic signal is red. Eq. (17) is the DRCC safety constraint by using a given minimum safe time headway. Noted that the interrupted flow facility locations such as traffic signal are considered by the definition of the maximum safe driving location $\hat{l}_s^{\xi}(t)$ in Section 3.2.3 in both Eq. (16) and Eq. (17). The other constraints are from Eqs. (6)-(12).

3.3.1. The data-driven uncertainty sets for safe driving constraints and driving references

For the safe driving of the DRCC constraints Eqs. (16)-(17), we apply a data-driven uncertainty set to define a set of probability distributions with support in Ξ for $\boldsymbol{x}_p^{\xi}(t)=[l_s^{\xi}(t),v_s^{\xi}(t)]^T$, given that it follows the probability distribution $\boldsymbol{F}_p\in\mathbb{D}_p$. $\mathbb{D}_p$ is defined as the data-driven uncertainty set of the random parameters $\boldsymbol{x}_p^{\xi}(t)$ for the safe driving constraints Eqs. (16)-(17). By solving the Eco-driving policy model in Section 3.2.3 for predictive driving reference, we can obtain a set of maximum safe driving locations and velocities $\{[\hat{l}_s^{\xi_i}(t)\quad\hat{v}_s^{\xi_i}(t)]^T\forall i=1,\ldots,M\}$. With this set, we can calculate the estimated mean $\hat{\boldsymbol{\mu}}_p$ and covariance $\hat{\boldsymbol{\Sigma}}_p$.

$$\mathbb{D}_p\left(S,\hat{\boldsymbol{\mu}}_{\boldsymbol{p}},\hat{\boldsymbol{\Sigma}}_{\boldsymbol{p}},\gamma_3,\gamma_4\right)=\left\{\begin{array}{l}\mathbb{P}\left(\hat{\boldsymbol{x}}_p^{\xi}(t)\in\boldsymbol{S}\right)=1\\\left(\mathbb{E}\left[\hat{\boldsymbol{x}}_p^{\xi}(x)\right]-\hat{\boldsymbol{\mu}}_{\boldsymbol{p}}\right)^T\hat{\boldsymbol{\Sigma}}_{\boldsymbol{p}}^{-1}\left(\mathbb{E}\left[\hat{\boldsymbol{x}}_p^{\xi}(t)\right]-\hat{\boldsymbol{\mu}}_{\boldsymbol{p}}\right)\leq\gamma_3\\\mathbb{E}\left[\left(\hat{\boldsymbol{x}}(t)-\hat{\boldsymbol{\mu}}_{\boldsymbol{p}}\right)\left(\hat{\boldsymbol{x}}_p^{\xi}(t)-\hat{\boldsymbol{\mu}}_{\boldsymbol{p}}\right)^T\right]\leq\gamma_4\hat{\boldsymbol{\Sigma}}_{\boldsymbol{p}}\end{array}\right\} \tag{21}$$

In Eq. (21), we use the constraint parameter $\gamma_3\geq 0$ such that the mean of $\hat{\boldsymbol{x}}_p^{\xi}$ lies in an ellipsoid of size γ_3 centered at the estimate $\hat{\boldsymbol{\mu}}_{\boldsymbol{p}}$. Weuse theconstraint parameter $\gamma_4\geq 1$ such that the second-moment matrix of $\hat{\boldsymbol{x}}_p^{\xi}$ lies in a positive semidefinite cone defined with a matrix inequality. It describes how likely $\hat{\boldsymbol{x}}_p^{\xi}$ is to be close to the estimate $\hat{\boldsymbol{\mu}}_{\boldsymbol{p}}$ with respect to the correlations expressed in $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{p}}$ (Delage and Ye, 2010).

Similarly, by solving Eqs. (5)-(14), we can obtain a set of controlled vehicle driving states $\{[\hat{l}_r^{\xi_i}(t)\quad\hat{v}_r^{\xi_i}(t)]^T\forall i=1,\ldots,M\}$, where M is the number of boundary conditions. We define $\boldsymbol{F}_r\in\mathbb{D}_r$ as the probability distribution and $\mathbb{D}_r$ as the data-driven

uncertainty set of the random parameters $\boldsymbol{x}_r^{\xi}(t) = [l_r^{\xi}(t), v_r^{\xi}(t)]^T$ in Eq. (22). This data-driven uncertainty set will be used as the data-driven input for the reference tracking in the DRSO-DRCC model.

$$\mathbb{D}_r\left(S, \hat{\mu}_r, \hat{\Sigma}_r, \gamma_3, \gamma_4\right) = \left\{ \begin{array}{c} \mathbb{P}\left(\hat{x}_r^{\xi}(t) \in S\right) = 1 \\ \left(\mathbb{E}\left[\hat{x}_r^{\xi}(t)\right] - \hat{\mu}_r\right)^T \hat{\Sigma}_r^{-1} \left(\mathbb{E}\left[\hat{x}_r^{\xi}(t)\right] - \hat{\mu}_r\right) \leq \gamma_3 \\ \mathbb{E}\left[\left(\hat{x}_r^{\xi}(t) - \hat{\mu}_r\right)\left(\hat{x}_r^{\xi}(t) - \hat{\mu}_r\right)^T\right] \leq \gamma_4 \hat{\Sigma}_r \end{array} \right\} \tag{22}$$

4. Model reformulation and solution algorithm

This section presents model reformulations and a solution algorithm to obtain tractable solutions that can be applied for real-time applications. The DRSO-DRCC model formulation in Section 3.3 cannot be directly solved by existing solvers. Moreover, the original minimax problem constrained by a DRCC constraint is computationally expensive (Shapiro and Nemirovski, 2005; Goh and Sim, 2010; Zymler et al., 2013) and is not suitable for real-time applications. In Section 4.1, we present the matrix form of the DRSO-DRCC model as a standard SDP form for simplicity. Based on the standard SDP form, we reformulate the DRSO-DRCC model as a tractable relaxed SDP model in Section 4.2. Specifically, in Section 4.2.1, we replace the DRCC constraints in the DRSO-DRCC model with equivalent deterministic convex constraints. In Sections 4.2.2–4.2.4, we reformulate the DRSO model by applying the strong duality theory and the Semidefinite Relaxation technique. A solution procedure and a solution algorithm are presented in Section 4.3 to solve the reformulated model (i.e., a relaxed SDP problem).

4.1. Matrix form of the DRSO-DRCC model

First, we formulate the proposed DRSO-DRCC model in Section 4.2 in a matrix form. Consider $\boldsymbol{Q}_{\boldsymbol{p}}^{\boldsymbol{\xi}}$ and $\boldsymbol{Q}_{\boldsymbol{p}}$ as the matrices for the constant-time-headway policy and the driving stability in the objective with respect to the leading controlled vehicle. We define $\boldsymbol{Q}_{\boldsymbol{r}}^{\boldsymbol{\xi}}$ as the matrix for the reference tracking under uncertainties for the predicted driving states of the controlled vehicles. We define matrix $\boldsymbol{R}$ as an identity matrix in this paper because we just want to reduce acceleration or deceleration perturbations. Furthermore, we define matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ for the state dynamics constraints. We define $\boldsymbol{K}_{\boldsymbol{r}}$ and $\boldsymbol{G}_{\boldsymbol{r}}$ as matrices in the location-based traffic control devices and road geometry DRCC constraint. We define $\boldsymbol{K}_{\boldsymbol{s}}$ and $\boldsymbol{G}_{\boldsymbol{s}}$ as matrices for the minimum-time-headway safety DRCC constraint. Then, we define $\boldsymbol{K}_{\boldsymbol{f}}$ and $\boldsymbol{G}_{\boldsymbol{f}}$ as matrices in the speed limit constraint.

$$\min_u \left\{ \begin{array}{l} \max_{F_p \in \mathbb{D}} \mathbb{E}_{F_p}\left[\sum_{t=k}^{k+PH(k)} \left[Q_p^{\xi}\hat{x}^{\xi}(t) - Q_p x(t)\right]^T \left[Q_p^{\xi}\hat{x}^{\xi}(t) - Q_p x(t)\right]\right] \\ + \max_{F_r \in \mathbb{D}} \mathbb{E}_{F_r}\left[\sum_{t=k}^{k+PH(k)-1} \left[Q_r^{\xi}\hat{x}^{\xi}(t) - x(t)\right]^T \left[Q_r^{\xi}\hat{x}^{\xi}(t) - x(t)\right]\right] \\ + \sum_{t=k}^{k+PH(k)} [u(t)]^T R[u(t)] \end{array} \right\} \tag{23}$$

s.t.

$$\inf_{F_p \in \mathbb{D}_p} \mathbb{P}_{F_p}\left(k_r x(t) + G_r \hat{x}_p^{\xi}(t) \leq -s^{jam}\right) \geq \gamma_1, t \in \{k, \ldots, k + PH(k) | 0 \leq \mathrm{mod}(t, t_c) \leq t_r\} \tag{24}$$

$$\inf_{F_p \in \mathbb{D}_p} \mathbb{P}_{F_p}\left(K_s x(t) + G_s \hat{x}_p^{\xi}(t) \leq -s^{jam}\right) \geq \gamma_2, \forall t = k, \ldots, k + PH(k) \tag{25}$$

$$x(t) = x^{obs}(k), t = k \tag{26}$$

$$x(t+1) = Ax(t) + Bu(t), \forall t = k, \ldots, k + PH(k) \tag{27}$$

$$u^{\min} \leq u(t) \leq u^{\max}, \forall t = k, \ldots, k + PH(k) \tag{28}$$

$$K_f x(t) - G_f <= 0, \forall t = k, \ldots, k + PH(k) \tag{29}$$

$$\mathbb{D}\left(S, \hat{\mu}, \hat{\Sigma}, \gamma_3, \gamma_4\right) = \left\{ \begin{array}{c} \mathbb{P}\left(\hat{x}^{\xi}(t) \in S\right) = 1 \\ \left(\mathbb{E}\left[\hat{x}^{\xi}(t)\right] - \hat{\mu}\right)^T \hat{\Sigma}^{-1} \left(\mathbb{E}\left[\hat{x}^{\xi}(t)\right] - \hat{\mu}\right) \leq \gamma_3 \\ \mathbb{E}\left[\left(\hat{x}^{\xi}(t) - \hat{\mu}\right)\left(\hat{x}^{\xi}(t) - \hat{\mu}\right)^T\right] \leq \gamma_4 \hat{\Sigma} \end{array} \right\}, \forall t = k, \ldots, k + PH(k) \tag{30}$$

where $\boldsymbol{u}(t)$ denotes acceleration or deceleration as the control variable (noted that we use bold $\boldsymbol{u}$ in the matrix-form formulation even we only control one vehicle in arterial connected and automated driving). $\boldsymbol{x}(t)$ denotes the vector of the driving state variables. $\boldsymbol{F_p}$ and $\boldsymbol{F_r}$ are the probability distributions of the random parameter $\hat{\boldsymbol{x}}^{\boldsymbol{\xi}}(t)$. Eq. (24) is the location-based traffic control devices and road geometry constraint. Eq. (25) is the minimum-time-headway safety constraint. Eq. (26) is the initial condition constraint. Eq. (27) is the state dynamics constraint. Eq. (28) is the acceleration or deceleration capability constraint. Eq. (29) is the speed limit constraint. In the data-driven uncertainty set constraint Eq. (30), we let $\boldsymbol{F} = \{\boldsymbol{F_p}, \boldsymbol{F_r}\}$ and $\mathbb{D} = \{\mathbb{D}_{\boldsymbol{p}}, \mathbb{D}_{\boldsymbol{r}}\}$.

4.2. Model reformulation

This section reformulates the matrix-form DRSO-DRCC model into a relaxed SDP dual problem by applying the strong duality theory and the Semidefinite Relaxation technique. The original formulation with DRCC constraints is difficult to solve since the empirical estimates are random themselves and cannot be directly used as the exact unknown mean and covariance (Calafiore and El Ghaoui, 2006). In addition, the original DRSO formulation of the problem is computationally expensive due to the minimax program (Shapiro and Nemirovski, 2005; Goh and Sim, 2010; Zymler et al., 2013). By applying the strong duality theory, the minimax program will be reformulated as a quadratically constrained minimization problem. However, the complexity of the quadratically constrained minimization problem is unknown and therefore is still computationally expensive (Floudas and Visweswaran, 1995; Ye and Zhang, 2003; Luo et al., 2010).

In this section, the reformulation includes four steps. First, we replace the DRCC constraints with the corresponding equivalent convex constraints in Section 4.2.1. Second, we formulate the dual problem of the inner problem of the DRSO model in Section 4.2.2 by applying the strong duality (Nilim and El Ghaoui, 2005). Third, in Section 4.2.3, we approximate the solution of the quadratically constrained dual problem in a numerically efficient way by the Semidefinite Relaxation technique (Luo et al., 2010). Fourth, in Section 4.2.4, we combine the original outer minimization problem with the relaxed SDP problem of the original inner maximization problem as a relaxed SDP counterpart for the original minimax program.

4.2.1. Equivalent convex constraints of the DRCC constraints

The DRCC constraints in Eqs. (24)-(25) can be explicitly expressed as the explicit counterpart – deterministic convex constraints (Calafiore and El Ghaoui, 2006). We consider that the data-driven uncertainty set $\mathbb{D}$ can be constructed using the mean and covariance information derived from observations.

Proposition 1. For any given $\gamma_p \in (0, 1)$, the DRCC constraint $\inf_{F_p \in \mathbb{D}_p} \mathbb{P}_{F_p}(\boldsymbol{k_r x}(t) + \boldsymbol{G_r}\hat{\boldsymbol{x}}_{\boldsymbol{p}}^{\boldsymbol{\xi}}(t) \le -s^{jam}) \ge \gamma_1$ is equivalent to the convex constraint:

$$\sqrt{\left(\gamma_1/1-\gamma_1\right) G_r \hat{\sigma}_1(t) G_r{}^T} + \begin{bmatrix} K_r & \left(\bar{x}_p^{\xi}(t)\right)^T \end{bmatrix} \begin{bmatrix} x(t) \\ G_r{}^T \end{bmatrix} \le -s^{jam} \tag{31}$$

where $\bar{\boldsymbol{x}}_p^{\xi}(t)$ is the mean and $\hat{\boldsymbol{\sigma}}_1(t)$ is the covariance matrix of $\{\hat{\boldsymbol{x}}_p^{\xi_i}(t) \forall i = 1, \ldots, M\}$.

Proof. To prove that the DRCC constraint is equivalent to its deterministic convex counterpart, we apply the equivalent reformulation from Calafiore and El Ghaoui (2006) based on the one-sided Chebyshev inequality (Marshall and Olkin, 1960; Bertsimas and Popescu, 2005; Calafiore and El Ghaoui, 2006).

First, we express the DRCC constraint in Eq. (24) by:

$$\inf_{F_p \in \mathbb{D}_p} \mathbb{P}_{F_p}\left(K_r x(t) + \left[\hat{x}_p^{\xi}(t)\right]^T G_r{}^T \le -s^{jam}\right) \ge \gamma_1 \tag{32}$$

We define the random vector $d = [\boldsymbol{k_r} \quad [\hat{\boldsymbol{x}}_{\boldsymbol{p}}^{\boldsymbol{\xi}}(t)]^T \boldsymbol{G_r}^T]$ and $\tilde{x} = [\boldsymbol{x}(t) \quad 1]^T$ so that the left-hand side of the DRCC constraint Eq. (32) can be expressed as $\varphi(\boldsymbol{x}) = d\tilde{\boldsymbol{x}}$.

Then the mean of the random vector d is:

$$\bar{d} = \mathbb{E}[d] = \begin{bmatrix} K_r & G_r \bar{x}^{\xi}(t) \end{bmatrix} \tag{33}$$

The covariance of the random vector d is:

$$\Gamma = \mathrm{cov}[d] = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \hat{\sigma}_1(t) \end{bmatrix} \succcurlyeq 0 \tag{34}$$

where mean $\bar{\boldsymbol{x}}^{\boldsymbol{\xi}}(t)$ and covariance $\hat{\boldsymbol{\sigma}}_1(t)$ are calculated from data. In our problem, Γ_{11}, Γ_{12}, and Γ_{21} are constant because $\boldsymbol{K_r}$ is constant.

Furthermore, we have:

$$\bar{\varphi}(x) = \mathbb{E}[\varphi(x)] = \bar{d}\tilde{x} \tag{35}$$

The covariance matrix of the random variable is:

$$\mathrm{var}[\varphi(x)] = \tilde{x}^T \Gamma \tilde{x} = G_r \hat{\sigma}_1(t) G_r{}^T \tag{36}$$

Theorem 3.1 in Calafiore and El Ghaoui (2006) concludes that a convex constraint $\sqrt{\frac{\gamma}{1-\gamma}var[\varphi(\boldsymbol{x})]}+\bar{\varphi}(\boldsymbol{x})\le 0$ is equivalent to $\inf_{d\sim(\bar{d},\Gamma)}\mathbb{P}(d^{\boldsymbol{T}}\tilde{\boldsymbol{x}}\le 0)\ge\gamma$ for any $\gamma\in(0,1)$ by applying the one-sided Chebyshev inequality (See detailed proofs in Marshall and Olkin, 1960; Propositions 6.3 and 6.4 - Bertsimas and Popescu, 2005; Theorem 3.1 - Calafiore and El Ghaoui, 2006). We need to show that their equivalent condition of DRCC constraint applies to our DRCC constraint in Eq. (24).

The equivalent convex constraint of Eq. (24) can be expressed in Eq. (37) based on Theorem 3.1 in Calafiore and El Ghaoui (2006).

$$\sqrt{\left(\gamma_1/1-\gamma_1\right)\tilde{x}^T\Gamma\tilde{x}}+\bar{d}\tilde{x}\le -s^{jam},\gamma_1\in(0,1) \tag{37}$$

The equivalent convex constraint can be expanded in Eq. (38).

$$\sqrt{\left(\gamma_1/1-\gamma_1\right)G_r\hat{\sigma}_1(t){G_r}^T}+\begin{bmatrix}K_r & \left(\bar{x}_p^{\xi}(t)\right)^T\end{bmatrix}\begin{bmatrix}x(t)\\ G_r^T\end{bmatrix}\le -s^{jam} \tag{38}$$

Eq. (38) is the equivalent convex constraint of Eq. (24). This completes the proof.

Similarly, the DRCC constraint Eq. (25) can be equivalent to its convex counterpart Eq. (39).

$$\sqrt{(\gamma_2/1-\gamma_2)G_s\hat{\sigma}_2(t)G_s^T}+\left[K_s\left(\bar{x}_p^{\xi}(t)\right)^T\right]\begin{bmatrix}x(t)\\ G_s^T\end{bmatrix}\le -S^{jam} \tag{39}$$

Then, the DRSO-DRCC model can be reformulated in Eq. (40) as an equivalent DRSO model by applying Proposition 1 into the DRCC constraints.

$$\min_u\max_{F\in\mathbb{D}=\{\mathbb{D}_p,\mathbb{D}_r\}}\sum_{t=k}^{k+PH(k)}\mathbb{E}_F\left\{\begin{array}{c}\left[Q_p^{\xi}\hat{x}^{\xi}(t)-Q_px(t)\right]^T\left[Q_p^{\xi}\hat{x}^{\xi}(t)-Q_px(t)\right]\\ \left[Q_r^{\xi}\hat{x}^{\xi}(t)-x(t)\right]^T\left[Q_r^{\xi}\hat{x}^{\xi}(t)-x(t)\right]\\ +[u(t)]^TRu(t)\end{array}\right\} \tag{40}$$

s.t.

$$\sqrt{(\gamma_1/1-\gamma_1)G_r\hat{\sigma}_1(t)G_r^T}+\left[K_r\quad\left(\bar{x}_p^{\xi}(t)\right)^T\right]\begin{bmatrix}x(t)\\ G_r^T\end{bmatrix}\le -s^{jam},t\in\{k,\ldots,k+PH(k)|0\le \mathrm{mod}(t,t_c)\le t_r\} \tag{41}$$

$$\sqrt{(\gamma_2/1-\gamma_2)G_s\hat{\sigma}_2(t)G_s^T}+\left[K_s\quad\left(\bar{x}_p^{\xi}(t)\right)^T\right]\begin{bmatrix}x(t)\\ G_s^T\end{bmatrix}\le -s^{jam},\forall t=k,\ldots,k+PH(k) \tag{42}$$

$$x(t)=x^{obs}(k),t=k \tag{43}$$

$$x(t+1)=Ax(t)+Bu(t),\forall t=k,\ldots,k+PH(k) \tag{44}$$

$$u^{\min}\le u(t)\le u^{\max},\forall t=k,\ldots,k+PH(k) \tag{45}$$

$$K_fx(t)-G_f<=0,\forall t=k,\ldots,k+PH(k) \tag{46}$$

$$\mathbb{D}\left(S,\hat{\mu},\hat{\Sigma},\gamma_3,\gamma_4\right)=\left\{\begin{array}{c}\mathbb{P}\left(\hat{x}^{\xi}(t)\in S\right)=1\\ \left(\mathbb{E}\left[\hat{x}^{\xi}(t)\right]-\hat{\mu}\right)^T\hat{\Sigma}^{-1}\left(\mathbb{E}\left[\hat{x}^{\xi}(t)\right]-\hat{\mu}\right)\le\gamma_3\\ \mathbb{E}\left[\left(\hat{x}^{\xi}(t)-\hat{\mu}\right)\left(\hat{x}^{\xi}(t)-\hat{\mu}\right)^T\right]\le\gamma_4\hat{\Sigma}\end{array}\right\},\forall t=k,\ldots,k+PH(k) \tag{47}$$

where $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ are the mean vector and covariance of state variables $\hat{l}_i^{\xi}(t)$ and $\hat{v}_i^{\xi}$, $i=0,\ldots,N$, from observations. The set $\boldsymbol{S}$ is any convex set that contains the support of $\boldsymbol{F}$. With the moment constraints in Eq. (47), the equivalent reformulated DRSO model is an SDP problem.

4.2.2. The dual of the inner maximization problem of the reformulated DRSO model

In this section, we show that we can reduce the computational complexity and maintain the solution tractability by solving the dual problem of the inner maximization problem in Eq. (40)–(47) as long as we show that the strong duality holds for the inner problem (**Proposition 2**).

To show the strong duality, we need to explicitly reformulate the moment based uncertainty set to the conic constraints in the inner problem. Specifically, the inner problem is reformulated as $\boldsymbol{F}$-integrable for all $\boldsymbol{F} \in \mathbb{D}$ in Eqs. (49)-(52), which is an infinite-dimensional convex problem. This reformulation allows us to utilize the strong duality theory in the infinite-dimensional convex problems for these moment constraints (Rockafellar, 2015).

For simplicity, we represent the objective function in Eq. (40) as Eq. (48).

$$J\left(u, x, \hat{x}^{\xi}, t\right) = \left\{ \begin{array}{l} \left[Q_p^{\xi}\hat{x}^{\xi}(t) - Q_p x(t)\right]^T \left[Q_p^{\xi}\hat{x}^{\xi}(t) - Q_p x(t)\right] \\ \left[Q_r^{\xi}\hat{x}^{\xi}(t) - x(t)\right]^T \left[Q_r^{\xi}\hat{x}^{\xi}(t) - x(t)\right]^T \\ [u(t)]^T R[u(t)] \end{array} \right\} \tag{48}$$

The inner maximization problem with conic constraints can be described as Eq. (49).

$$\max_{F\in\mathbb{D}} \int_S \sum_{t=k}^{k+PH(k)} J\left(u, x, \hat{x}^{\xi}, t\right) dF\left(\hat{x}^{\xi}(t)\right) \tag{49}$$

s.t.

$$\int_s dF\left(\hat{x}^{\xi}(t)\right) = 1, \forall t = k, \ldots, k + PH(k) \tag{50}$$

$$\int_S \begin{bmatrix} \hat{\Sigma} & \left(\hat{x}^{\xi}(t) - \hat{\mu}\right)^T \\ \hat{x}^{\xi}(t) - \hat{\mu} & \gamma_3 \end{bmatrix} dF\left(\hat{x}^{\xi}(t)\right) \succcurlyeq 0, \forall t = k, \ldots, k + PH(k) \tag{51}$$

$$\int_S \left[\left(\hat{x}^{\xi}(t) - \hat{\mu}\right)\left(\hat{x}^{\xi}(t) - \hat{\mu}\right)^T\right] dF\left(\hat{x}^{\xi}(t)\right) \circ \gamma_4 \hat{\Sigma}, \forall t = k, \ldots, k + PH(k) \tag{52}$$

In Eq. (61), the original first-moment constraint in Eq. (57) is replaced by a semidefinite cone constraint by Schur complement (Boyd and Vandenberghe, 2004). Noted that $\boldsymbol{S}$ is any closed convex set known to contain the support of $\boldsymbol{F}$.

Proposition 2. Suppose that $\gamma_3 \geq 0$, $\gamma_4 \geq 1$ and $J(\boldsymbol{u}, \boldsymbol{x}, \hat{\boldsymbol{x}}^{\xi}, t)$ is $\boldsymbol{F}$-integrable where $\boldsymbol{F} \in \mathbb{D}$. Given any $\boldsymbol{u}(t)$, we denote $\Psi(\boldsymbol{u}(t), \gamma_3, \gamma_4)$ as the optimal solution of the inner maximization problem Eqs. (49)–(52). Then, the duality gap between the primal optimal solution $\Psi(\boldsymbol{u}(t), \gamma_3, \gamma_4)$. And the optimal solution of the dual problem is zero.

Proof. Strong duality does not hold in general while it usually holds for convex problems Calafiore and El Ghaoui, 2014). To show that the strong duality holds for the inner problem, we first need to demonstrate that the objective function of the inner problem is convex. We also need to prove that the regularity conditions and the Slater's constraint qualifications (i.e., the conditions that guarantee strong duality in convex problems) are satisfied (Shapiro, 2001; Liu et al., 2015; Wiesemann et al., 2014). Then, we can show that solving the dual problem is equal to solving the primal inner maximization problem in Eqs. (49)-(52).

First, we need to show the quadratic terms in the objective function are convex quadratic. Without loss of generality, we denote that any quadratic term in the objective function of the conic constrained inner maximization problem as $f(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{Z}\boldsymbol{x}$, where $\boldsymbol{Z}$ is a given positive semidefinite matrix. Noted that in the objective function Eq. (40), the matrices $\boldsymbol{I}$ and $\boldsymbol{R}$ are positive semidefinite. By the definition of convex functions in Boyd and Vandenberghe (2004), for any $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}$, we need to show that $f(\lambda\boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2) \leq \lambda f(\boldsymbol{x}_1) + (1-\lambda) f(\boldsymbol{x}_2)$. For the quadratic function $f(\boldsymbol{x})$, we have:

$$\begin{aligned} &[\lambda\boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2]^T Z[\lambda\boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2] \leq \lambda\boldsymbol{x}_1{}^T\boldsymbol{Z}\boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2{}^T\boldsymbol{Z}\boldsymbol{x}_2\lambda^2\boldsymbol{x}_1{}^T\boldsymbol{Z}\boldsymbol{x}_1 \\ &+ (1-\lambda)^2\boldsymbol{x}_2{}^T\boldsymbol{Z}\boldsymbol{x}_2 + \lambda(1-\lambda)\boldsymbol{x}_1{}^T\boldsymbol{Z}\boldsymbol{x}_2 + \lambda(1-\lambda)\boldsymbol{x}_2{}^T\boldsymbol{Z}\boldsymbol{x}_1 \leq \lambda\boldsymbol{x}_1{}^T\boldsymbol{Z}\boldsymbol{x}_1 \\ &+ (1-\lambda)\boldsymbol{x}_2{}^T\boldsymbol{Z}\boldsymbol{x}_2\lambda(1-\lambda)\left[\boldsymbol{x}_1{}^T\boldsymbol{Z}\boldsymbol{x}_1 + \boldsymbol{x}_2{}^T\boldsymbol{Z}\boldsymbol{x}_2\right] \geq \lambda(1-\lambda)\left[\boldsymbol{x}_1{}^T\boldsymbol{Z}\boldsymbol{x}_2 + \boldsymbol{x}_2{}^T\boldsymbol{Z}\boldsymbol{x}_1\right]\boldsymbol{x}_1{}^T\boldsymbol{Z}\boldsymbol{x}_2 \\ &+ \boldsymbol{x}_2{}^T\boldsymbol{Z}\boldsymbol{x}_1 \leq \boldsymbol{x}_1{}^T\boldsymbol{Z}\boldsymbol{x}_1 + \boldsymbol{x}_2{}^T\boldsymbol{Z}\boldsymbol{x}_2(\boldsymbol{x}_1 - \boldsymbol{x}_2)^T\boldsymbol{Z}(\boldsymbol{x}_1 - \boldsymbol{x}_2) \geq 0 \end{aligned} \tag{53}$$

In Eq. (53), $(\boldsymbol{x}_1 - \boldsymbol{x}_2)^T\boldsymbol{Z}(\boldsymbol{x}_1 - \boldsymbol{x}_2) \geq 0$ is true for any positive semidefinite matrix $\boldsymbol{Z}$, which shows the convexity of $f(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{Z}\boldsymbol{x}$. This applies to the quadratic terms with semidefinite matrices in the objective function Eq. (40). Thus, the objective function of the conic constrained inner maximization problem is quadratic convex.

Then, we need to show that the Slater's constraint qualification conditions hold. Based on the definition of convex problems in Boyd and Vandenberghe (2004), strong duality holds if the primal problem is convex and the Slater's condition holds as long as the inner maximization problem (i.e., the primal problem) has a nonempty and bounded interior solution (i.e., the regularity condition for finite-dimensional mathematical programs holds). According to our definition of the data-driven uncertainty set $\mathbb{D}$ in Eqs. (19)-(20), the probability distribution $\boldsymbol{F}$ is bounded, and then the regularity condition is applicable as long as the inner maximization problem has finite nonempty solutions (Liu et al., 2015; Wiesemann et al., 2014). Such

regularity conditions regarding nonempty and bounded interior solutions for the proof of strong duality of convex programs are discussed in Proposition 3.4 of Shapiro (2001) and Bertsimas and Popescu (2005). With the regularity condition, we can show that the Slater's condition holds if there exists a strictly feasible solution that satisfying the linear equality constraint and convex inequality constraint.

To prove Slater's condition holds for the conic constrained inner maximization problem, we need to show that the equality constraint in Eq. (50) is linear and the inequality constraints in Eqs. (51)-(52) are convex. Eq. (50) is linear because it is just the integral of the first constraint in Eq. (47). For the first-moment inequality constraint in Eq. (47) – the equivalent of Eq. (51), we define $\boldsymbol{z} = \mathbb{E}[\hat{\boldsymbol{x}}^{\xi}(t)] - \hat{\boldsymbol{\mu}}$ such that the inequality constraint can be expressed as $z^T \hat{\Sigma}^{-1} z - \gamma_3 \leq 0$, which is quadratic convex based on the procedure in Eq. (53). Similarly, the second-moment inequality constraint in Eq. (47), the equivalent of Eq. (52), is quadratic convex.

The optimal duality gap is zero if there is a feasible solution that satisfies the linear equality constraint and the convex inequality constraint of the conic constrained inner maximization problem. This completes the proof.

To formulate the dual problem of the conic constrained inner maximization problem, we begin with the Lagrangian of the conic constrained inner optimization problem as Eq. (54) by relaxing moment based conic constraints.

$$\begin{aligned}\mathcal{L}\left(\hat{x}^{\xi}, s_d, Q_d, P_d, p_d\right) &= J\left(u, x, \hat{x}^{\xi}, t\right) + \gamma_4 \hat{\Sigma} - \left[\hat{x}^{\xi}(t) - \hat{\mu}\right]\left[\hat{x}^{\xi}(t) - \hat{\mu}\right]^T, Q_d(t) \\ &+ \left\langle \begin{bmatrix} \hat{\Sigma} & \hat{x}^{\xi}(t) - \hat{\mu} \\ \left(\hat{x}^{\xi}(t) - \hat{\mu}\right)^T & \gamma_3 \end{bmatrix}, \begin{bmatrix} P_d(t) & p_d(t) \\ (p_d(t))^T & s_d(t) \end{bmatrix} \right\rangle \end{aligned} \tag{54}$$

where $\langle \bullet \ , \ \bullet \rangle$ denotes the inner product. $\boldsymbol{Q_d}(t) \in \mathbb{R}^{m \times m}$ is the dual variable for the second-moment constraint in Eq. (47) and $\boldsymbol{P_d}(t) \in \mathbb{R}^{m \times m}$, $\boldsymbol{p_d}(t) \in \mathbb{R}^m, s_d(t) \in \mathbb{R}$ are the dual variables for the first-moment constraint in Eq. (47). Noted that m is the length of the random vector $\hat{\boldsymbol{x}}^{\xi}$.

To remove the inner products in Eq. (54), the Lagrangian of the inner optimization problem can be further expanded as Eq. (55).

$$\begin{aligned}\mathcal{L}\left(\hat{\boldsymbol{x}}^{\xi}, s_d, \boldsymbol{Q_d}, \boldsymbol{P_d}, \boldsymbol{p_d}\right) &= J\left(\boldsymbol{u}, \boldsymbol{x}, \hat{\boldsymbol{x}}^{\xi}, t\right) + \gamma_4 \hat{\boldsymbol{\Sigma}} \bullet \boldsymbol{Q_d}(t) - \boldsymbol{tr}\left(\left[\hat{\boldsymbol{x}}^{\xi}(t) - \hat{\boldsymbol{\mu}}\right]^{\boldsymbol{T}} \boldsymbol{Q_d}(t)\left[\hat{\boldsymbol{x}}^{\xi}(t) - \hat{\boldsymbol{\mu}}\right]\right) \\ &+ \boldsymbol{tr}\left(\begin{bmatrix} \hat{\boldsymbol{\Sigma}} & \left(\hat{\boldsymbol{x}}^{\xi}(t) - \hat{\boldsymbol{\mu}}\right)^T \\ \hat{\boldsymbol{x}}^{\xi}(t) - \hat{\boldsymbol{\mu}} & \gamma_3 \end{bmatrix}\begin{bmatrix} \boldsymbol{P_d}(t) & \boldsymbol{p_d}(t) \\ (\boldsymbol{p_d}(t))^T & s_d(t) \end{bmatrix}\right) \\ &= \left\{J\left(\boldsymbol{u}, \boldsymbol{x}, \hat{\boldsymbol{x}}^{\xi}, t\right) - \left[\hat{\boldsymbol{x}}^{\xi}(t)\right]^T \boldsymbol{Q_d}(t)\hat{\boldsymbol{x}}^{\xi}(t) - 2\left[\hat{\boldsymbol{x}}^{\xi}(t)\right]^T \boldsymbol{Q_d}(t)\hat{\boldsymbol{\mu}} + 2\left[\hat{\boldsymbol{x}}^{\xi}(t)\right]^T \boldsymbol{p_d}(t)\right\} \\ &- 2\hat{\boldsymbol{\mu}}^T \boldsymbol{p_d}(t) + \gamma_3 s_d(t) + \gamma_4 \hat{\boldsymbol{\Sigma}} \bullet \boldsymbol{Q_d}(t) + \hat{\boldsymbol{\mu}}^{\boldsymbol{T}} \boldsymbol{Q_d}(t)\hat{\boldsymbol{\mu}} + \hat{\boldsymbol{\Sigma}} \bullet \boldsymbol{P_d}(t) \end{aligned} \tag{55}$$

where $(\bullet)$ denotes the Frobenius inner product.

For illustration, we decompose the Lagrangian into two parts such that $\mathcal{L}(\hat{\boldsymbol{x}}^{\xi}, s_d, \boldsymbol{Q_d}, \boldsymbol{P_d}, \boldsymbol{p_d}) = \mathcal{L}_1(\hat{\boldsymbol{x}}^{\xi}, \boldsymbol{Q_d}, \boldsymbol{p_d}) + \mathcal{L}_2(s_d, \boldsymbol{Q_d}, \boldsymbol{P_d}, \boldsymbol{p_d})$.

$$\begin{aligned}\mathcal{L}_1\left(\hat{\boldsymbol{x}}^{\xi}, \boldsymbol{Q_d}, \boldsymbol{p_d}\right) &= \left\{J\left(\boldsymbol{u}, \boldsymbol{x}, \hat{\boldsymbol{x}}^{\xi}, t\right) - \left[\hat{\boldsymbol{x}}^{\xi}(t)\right]^T \boldsymbol{Q_d}(t)\hat{\boldsymbol{x}}^{\xi}(t) - 2\left[\hat{\boldsymbol{x}}^{\xi}(t)\right]^T \boldsymbol{Q_d}(t)\hat{\boldsymbol{\mu}} + 2\left[\hat{\boldsymbol{x}}^{\xi}(t)\right]^T \boldsymbol{p_d}(t)\right\} \\ \mathcal{L}_2(s_d, \boldsymbol{Q_d}, \boldsymbol{P_d}, \boldsymbol{p_d}) &= -2\hat{\boldsymbol{\mu}}^T \boldsymbol{p_d}(t) + \gamma_3 s_d(t) + \gamma_4 \hat{\boldsymbol{\Sigma}} \bullet \boldsymbol{Q_d}(t) + \hat{\boldsymbol{\mu}}^{\boldsymbol{T}} \boldsymbol{Q_d}(t)\hat{\boldsymbol{\mu}} + \hat{\boldsymbol{\Sigma}} \bullet \boldsymbol{P_d}(t) \end{aligned} \tag{56}$$

Based on the definition of the Lagrangian dual problem in Bazaraa et al. (2013), the dual problem can be formulated as minimizing the supreme (i.e., the least upper bound) of the Lagrangian function Eq. (55). This dual problem can be further expressed by incorporating the function $\mathcal{L}_2(s_d, \boldsymbol{Q_d}, \boldsymbol{P_d}, \boldsymbol{p_d})$ that consists of dual variables and minimizing the supremum of the function $\mathcal{L}_1(\hat{\boldsymbol{x}}^{\xi}, \boldsymbol{Q_d}, \boldsymbol{p_d})$ in Eq. (57).

$$\min_{Q_d, P_d, p_d} \mathcal{L}_2(s_d, Q_d, P_d, p_d) + \left[\sup_{\hat{x}^{\xi}} \mathcal{L}_1\left(\hat{x}^{\xi}, Q_d, p_d\right)\right] \tag{57}$$

s.t.

$$\begin{bmatrix} P_d(t) & p_d(t) \\ (p_d(t))^T & s_d(t) \end{bmatrix} \succcurlyeq 0, Q_d(t) \succcurlyeq 0, \forall t = k, \ldots, k + PH(k) \tag{58}$$

where the objective function in Eq. (57) is to minimize the Lagrangian function, in which $\sup_{\hat{\boldsymbol{x}}^{\xi}} \mathcal{L}_1(\hat{\boldsymbol{x}}^{\xi}, \boldsymbol{Q_d}, \boldsymbol{p_d})$ considers the worst-case distribution of the stochastic parameter $\hat{\boldsymbol{x}}^{\xi}$. Eq. (58) contains dual variable constraints.

By adding a dual variable r_d, we can rewrite the dual problem as Eqs. (59)-(61) (see Delage and Ye, 2010).

$$\min_{Q_d, P_d, p_d, r_d, s_d} \sum\nolimits_{t=k}^{k+PH} \mathcal{L}_2(s_d, Q_d, P_d, p_d) + r_d(t) \tag{59}$$

s.t.

$$r_d(t) \geq J\left(u, x, \hat{x}^{\xi}, t\right) - \left[\hat{x}^{\xi}(t)\right]^T Q_d(t)\hat{x}^{\xi}(t) - 2\left[\hat{x}^{\xi}(t)\right]^T Q_d(t)\hat{\mu} + 2\left[\hat{x}^{\xi}(t)\right]^T p_d(t), \forall t = k, \ldots, k + PH(k) \tag{60}$$

$$\begin{bmatrix} P_d(t) & p_d(t) \\ (p_d(t))^T & s_d(t) \end{bmatrix} \succcurlyeq 0, Q_d \succcurlyeq 0, r_d(t) \geq 0, \forall t = k, \ldots, k + PH(k) \tag{61}$$

The dual problem in Eqs. (59)-(61) is a quadratically constrained program. Based on Proposition 2, we can obtain tractable solutions by solving the dual problem Eqs. (59)-(61), which is equivalent to solving the primal conic constrained inner maximization problem in the DRSO model. However, the complexity of the dual problem is unknown and is not suitable for the real-time application of the proposed connected and automated control model due to the quadratic constraint in Eq. (60) (Floudas and Visweswaran, 1995; Ye and Zhang, 2003; Luo et al., 2010).

4.2.3. The relaxed SDP of the dual problem

To obtain a computationally efficient approximation of the quadratically constrained optimization problems in Eqs. (59)-(61), this section applies the Semidefinite Relaxation technique in Luo et al. (2010) to deal with the quadratic constraint in Eq. (60). By the Semidefinite Relaxation technique, we introduce two new variables - $\boldsymbol{X}(t) = \boldsymbol{x}(t)[\boldsymbol{x}(t)]^T$ and $\boldsymbol{U}(t) = \boldsymbol{u}(t)[\boldsymbol{u}(t)]^T$. Then, the quadratic constraint in Eq. (70) can be reformulated as Eq. (62).

$$\begin{aligned} & r_d + \left[\hat{x}^{\xi}(t)\right]^T \left(Q_d(t) - \left(Q_p^{\xi}\right)^T Q_p^{\xi} - \left(Q_r^{\xi}\right)^T Q_r^{\xi}\right) \hat{x}^{\xi}(t) \\ & + \left[\hat{x}^{\xi}(t)\right]^T \left(-2p_d(t) - 2Q_d(t)\hat{\mu} + 2\left(\left(Q_p^{\xi}\right)^T Q_p^{\xi} + \left(Q_r^{\xi}\right)^T Q_r^{\xi}\right) x(t)\right) \\ & - tr\left(\left(Q_p{}^T Q_p + I\right) X(t)\right) - tr(RU(t)) \geq 0, rank(X(t)) = 1, rank(U(t)) = 1, X(t) \succcurlyeq 0, U(t) \\ & \succcurlyeq 0, \forall t = k, \ldots, k + PH(k) \end{aligned} \tag{62}$$

In Eq. (62), we show that $[\boldsymbol{u}(t)]^T \boldsymbol{R} \boldsymbol{u}(t) = \boldsymbol{tr}([\boldsymbol{u}(t)]^T \boldsymbol{R} \boldsymbol{u}(t)) = \boldsymbol{tr}(\boldsymbol{R}\boldsymbol{U}(t))$. Noted that the constraint $rank(\boldsymbol{X}(t)) = 1$, $\boldsymbol{X}(t) \succcurlyeq 0$ is equivalent to restricting that $\boldsymbol{X}(t) = \boldsymbol{x}(t)[\boldsymbol{x}(t)]^T$, which makes the problem still difficult to solve Luo et al., 2010). To reduce the computational expense while obtaining a tractable approximation of the optimal solution of the quadratically constrained dual problem Eqs. (59)-(61), we can drop the rank constraints on $\boldsymbol{X}(t)$ and $\boldsymbol{U}(t)$ to obtain a numerically efficient approximation (Zhang, 2000; Kim and Kojima, 2003; Luo et al., 2010).

Then, without the rank constraints, the quadratic form of the constraint in Eq. (62) can be written as Eq. (63).

$$\begin{aligned} & \begin{bmatrix} \left(\hat{x}^{\xi}(t)\right)^T & 1 \end{bmatrix} \begin{bmatrix} Q_d(t) - \left(Q_p^{\xi}\right)^T Q_p^{\xi} - \left(Q_r^{\xi}\right)^T Q_r^{\xi} & -p_d(t) - Q_d(t)\hat{\mu} + \left(\left(Q_p^{\xi}\right)^T Q_p^{\xi} + \left(Q_r^{\xi}\right)^T Q_r^{\xi}\right) x(t) \\ -p_d(t)^T - \hat{\mu}^T Q_d(t)^T + x(t)^T \left(\left(Q_p^{\xi}\right)^T Q_p^{\xi} + \left(Q_r^{\xi}\right)^T Q_r^{\xi}\right)^T & r_d(t) - tr\left(\left(Q_p{}^T Q_p + I\right) X(t)\right) - tr(RU(t)) \end{bmatrix} \begin{bmatrix} \hat{x}^{\xi}(t) \\ 1 \end{bmatrix} \\ & \succcurlyeq 0, X(t) \succcurlyeq 0, U(t) \succcurlyeq 0, \forall t = k, \ldots, k + PH(k) \end{aligned} \tag{63}$$

The stochastic term $\begin{bmatrix} \left(\hat{x}^{\xi}(t)\right)^T & 1 \end{bmatrix}$ in Eq. (63) makes the quadratic constraint difficult to solve. By the S-Lemma (Yakubovic, 1971; Pólik and Terlaky, 2007; Lemma 2.2 in Zymler et al., 2013), we remove the stochastic term and replace the constraint in Eq. (63) by Eq. (64).

$$\begin{aligned} & \begin{bmatrix} Q_d(t) - \left(Q_p^{\xi}\right)^T Q_p^{\xi} - \left(Q_r^{\xi}\right)^T Q_r^{\xi} & -p_d(t) - Q_d(t)\hat{\mu} + \left(\left(Q_p^{\xi}\right)^T Q_p^{\xi} + \left(Q_r^{\xi}\right)^T Q_r^{\xi}\right) x(t) \\ -p_d(t)^T - \hat{\mu}^T Q_d(t)^T + x(t)^T \left(\left(Q_p^{\xi}\right)^T Q_p^{\xi} + \left(Q_r^{\xi}\right)^T Q_r^{\xi}\right)^T & r_d(t) - tr\left(\left(Q_p{}^T Q_p + I\right) X(t)\right) - tr(RU(t)) \end{bmatrix} \\ & \succcurlyeq 0, X(t) \succcurlyeq 0, U(t) \succcurlyeq 0, \forall t = k, \ldots, k + PH(k) \end{aligned} \tag{64}$$

Finally, the dual problem in Eqs. (57)-(58) is reformulated as a relaxed dual problem in Eqs. (65)-(67).

$$\min_{Q_d, P_d, p_d, r_d, s_d} \sum_{t=k}^{k+PH(k)} -2\hat{\mu}^T p_d(t) + \gamma_1 s_d(t) + \gamma_2 \hat{\Sigma} \bullet Q_d(t) + \hat{\mu}^T Q_d(t)\hat{\mu} + \hat{\Sigma} \bullet P_d(t) + r_d(t) \tag{65}$$

s.t.

$$\begin{aligned} & \begin{bmatrix} Q_d(t) - \left(Q_p^{\xi}\right)^T Q_p^{\xi} - \left(Q_r^{\xi}\right)^T Q_r^{\xi} & \frac{1}{2} q_d(t) + \left(\left(Q_p^{\xi}\right)^T Q_p^{\xi} + \left(Q_r^{\xi}\right)^T Q_r^{\xi}\right) x(t) \\ \frac{1}{2}(q_d(t))^T + x(t)^T \left(\left(Q_p^{\xi}\right)^T Q_p^{\xi} + \left(Q_r^{\xi}\right)^T Q_r^{\xi}\right)^T & r_d - tr\left(\left(Q_p{}^T Q_p + I\right) X(t)\right) - tr(RU(t)) \end{bmatrix} \\ & \succcurlyeq 0, q_d(t) = -p_d(t) - Q_d(t)\hat{\mu}, X(t) \succcurlyeq 0, U(t) \succcurlyeq 0, \forall t = k, \ldots, k + PH(k) \end{aligned} \tag{66}$$

$$\begin{bmatrix} P_d(t) & p_d(t) \\ (p_d(t))^T & s_d(t) \end{bmatrix} \succcurlyeq 0, Q_d(t) \succcurlyeq 0, r_d(t) \geq 0, \forall t = k, \ldots, k + PH(k) \tag{67}$$

where the relaxed dual problem in Eqs. (65)-(67) is an SDP program. Solving this relaxed dual problem will obtain an approximation of the optimal solution of the primal inner maximization problem of the DRSO model in Eqs. (41)-(47). Noted that the Semidefinite Relaxation works well empirically in many applications that can generate reliable solutions efficiently (Ma et al., 2002; Ma et al., 2004; Sidiropoulos et al., 2006; Li and Jiang, 2007; Luo et al., 2010). The relaxed dual problem will be integrated with the outer minimization problem of the DRSO model in Section 4.2.4.

4.2.4. The reformulated DRSO-DRCC model – a relaxed SDP

This section reformulates the proposed DRSO-DRCC model in Section 4.1 as a relaxed SDP problem based on the discussions of Sections 4.2.1, 4.2.2, and 4.2.3. We formulate an equivalent convex constraint for the DRCC constraint in Section 4.2.1 using the mean and variance information of the predictive uncertain driving states of the IPV. For the conic constrained inner maximization problem of the DRSO model, we formulate its dual problem as a quadratically constrained program by applying the strong duality theory in Section 4.2.2. Then, we reformulate a relaxed SDP in Section 4.2.3 to obtain the approximation of the optimal solution of the quadratically constrained dual problem by using the Semidefinite Relaxation technique.

We combine the relaxed dual problem in Eqs. (65)-(67) with the outer minimization problem in Eqs. (41)-(47) as a minimization problem in Eqs. (78)-(84).

$$\min_{u,Q_d,P_d,p_d,r_d,s_d} \sum_{t=k}^{k+PH(k)} \begin{bmatrix} \gamma_4 \hat{\Sigma} \bullet Q_d(t) - \hat{\mu}\hat{\mu}^T \bullet Q_d(t) + r_d(t) + \hat{\Sigma} \bullet P_d(t) \\ -2\hat{\mu}^T p_d(t) + \gamma_3 s_d(t) \end{bmatrix} \tag{68}$$

s.t.

$$\begin{bmatrix} Q_d(t) - \left(Q_p^{\xi}\right)^T Q_p^{\xi} - \left(Q_r^{\xi}\right)^T Q_r^{\xi} & \frac{1}{2} q_d(t) + \left(\left(Q_p^{\xi}\right)^T Q_p^{\xi} + \left(Q_r^{\xi}\right)^T Q_r^{\xi}\right) x(t) \\ \frac{1}{2}(q_d(t))^T + x(t)^T \left(\left(Q_p^{\xi}\right)^T Q_p^{\xi} + \left(Q_r^{\xi}\right)^T Q_r^{\xi}\right)^T & r_d - tr\left(\left(Q_p{}^T Q_p + I\right) X(t)\right) - tr(RU(t)) \end{bmatrix} \succcurlyeq 0, q_d(t) = -p_d(t) - Q_d(t)\hat{\mu}, X(t) \succcurlyeq 0, U(t) \succcurlyeq 0, \forall t = k, \ldots, k + PH(k) \tag{69}$$

$$\begin{bmatrix} P_d(t) & p_d(t) \\ (p_d(t))^T & s_d(t) \end{bmatrix} \succcurlyeq 0, Q_d(t) \succcurlyeq 0, r_d(t) \geq 0, \forall t = k, \ldots, k + PH(k) \tag{70}$$

$$\sqrt{\left(\gamma_1 / 1 - \gamma_1\right) G_r \hat{\sigma}_1(t) G_r{}^T} + \begin{bmatrix} K_r & \left(\bar{x}_p^{\xi}(t)\right)^T \end{bmatrix} \begin{bmatrix} x(t) \\ G_r{}^T \end{bmatrix} \leq -s^{jam}, t \in \{k, \ldots, k + PH(k) | t_{r,a}(k) \leq t \leq t_{r,b}(k)\} \tag{71}$$

$$\sqrt{(\gamma_2 / 1 - \gamma_2) G_s \hat{\sigma}_2(t) G_s^T} + \begin{bmatrix} K_s & \left(\bar{x}_p^{\xi}(t)\right)^T \end{bmatrix} \begin{bmatrix} x(t) \\ G_s^T \end{bmatrix} \leq -s^{jam}, \forall t = k, \ldots, k + PH(k) \tag{72}$$

$$x(t) = x^{obs}(k), t = k \tag{73}$$

$$x(t+1) = Ax(t) + Bu(t), \forall t = k, \ldots, k + PH(k) \tag{74}$$

$$u^{\min} \leq u(t) \leq u^{\max}, \forall t = k, \ldots, k + PH(k) \tag{75}$$

$$K_f x(t) - G_f <= 0, \forall t = k, \ldots, k + PH(k) \tag{76}$$

$$u^{\min} \leq u(t) \leq u^{\max}, \forall t = k, \ldots, k + PH(k) \tag{77}$$

$$K_f x(t) - G_f <= 0, \forall t = k, \ldots, k + PH(k) \tag{78}$$

where $\boldsymbol{Q_d}(t) \in \mathbb{R}^{m \times m}$ is the dual variable for the second-moment constraint in Eqs. (51)-(52) and $\boldsymbol{P_d}(t) \in \mathbb{R}^{m \times m}$, $\boldsymbol{p_d}(t) \in \mathbb{R}^m, s_d(t) \in \mathbb{R}$ are the dual variables for the first-moment constraint in Eqs. (51)-(52). Noted that m is the length of the random vector $\hat{\boldsymbol{x}}^{\xi}$.

This minimization problem is a relaxed SDP problem. In the objective function, $r_d(t)$, is the dual variable for the Lagrangian function of the inner maximization problem of Eqs. (23)-(30). All the other terms in the objective function are from the dual variables. The constraint Eq. (69) is the relaxed constraint after applying the dual variable $r_d(t)$. Eq. (70) is the constraint by applying dual variables $\boldsymbol{Q_d}(t)$ - the dual variable for the second-moment constraint Eq. (52), and $\boldsymbol{P_d}(t)$,$\boldsymbol{p_d}(t)$,$s_d(t)$ - the dual variables for the first-moment constraint Eq. (30). Eqs. (71)-(72) are the equivalent convex constraint of the DRCC constraints. Eqs. (73)-(78) are the original constraints of the outer minimization problem.

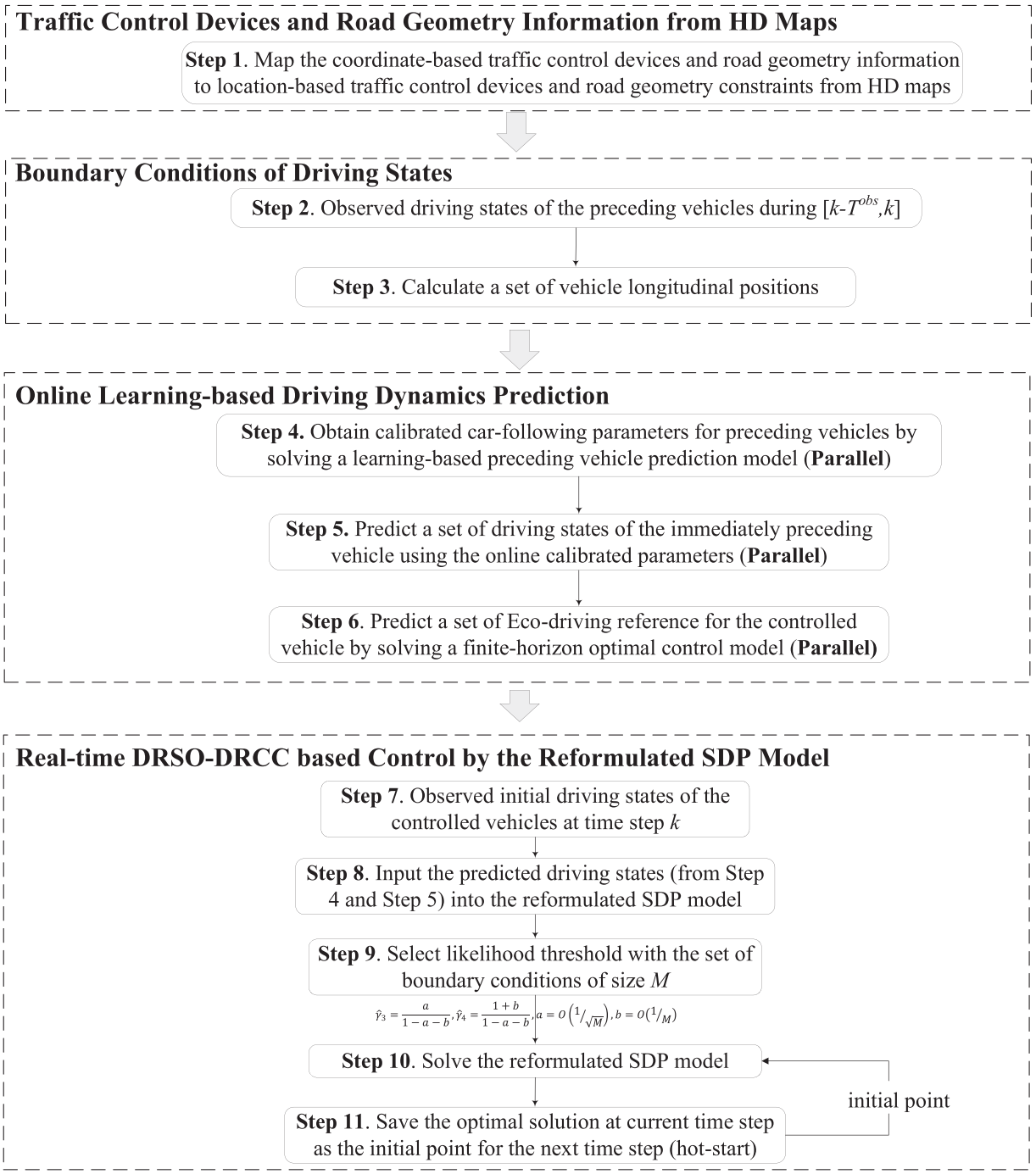

Fig. 10. The solution algorithm based on parallel computing and hot start.

4.3. Solution algorithm

In this section, we propose an efficient algorithm for solving the reformulated SDP problem to implement a real-time control system of the DRSO-DRCC model. The computational time of solving the proposed model can be reduced in two aspects: (1) applying parallel computing to the online learning-based driving dynamics prediction (Steps 4-6 in Fig. 10); (2) using a hot-start method for the real-time DSRO-DRCC based control (Step 7-11 in Fig. 10).

The solution procedure is summarized as a step-by-step algorithm in Fig. 10. Step 1 real-time updates traffic control devices and road geometry information from HD maps for location-based traffic control devices and road geometry constraints.

Step 2 and Step 3 predict the uncertain driving states of the boundary condition preceding vehicle by randomly sampling from historical observations. Step 4 solves the learning-based prediction model to obtain the calibrated car-following parameters for the preceding vehicles in a parallel computing framework. We iteratively check whether the online calibrated car-following parameters are stable at Step 4. For example, we assume that the online calibrated parameters are stable after T^p. If $k \geq T^p$, we will solve the learning-based prediction model to calibrate car-following models at a time interval ΔT^p instead of solving it at every time step. We set the driving dynamic prediction update interval (e.g., $\Delta T^p = 2$ seconds) greater than the control update interval (e.g., $\Delta t = 0.1$ seconds) to reduce the computation time. Furthermore, we break Steps 4-6 into independent parts so that each vehicle under each boundary condition can be processed concurrently on a computer with multi-core processors. Step 5 predicts the uncertain driving states of the preceding vehicles using the online calibrated parameters from Step 4. Step 6 solves the proposed finite-horizon optimal control model to predict the uncertain driving states of the controlled vehicle due to the Eco-driving reference constrained by the predicted driving states of the preceding vehicles from Step 5. Step 7 collects historical observations of the controlled vehicle due to the initial conditions for the MPC model. Step 8 uses the predicted driving states of the IPV in the safety constraint of the proposed model. Step 9 calculates the parameters for the data-driven uncertainty set of the proposed DRSO-DRCC model. Step 10 solves the reformulated SDP to obtain the optimal acceleration or deceleration commands for the controlled CAV with an initial point. Step 11 saves the optimal solution at the current time step as the initial point for Step 10 at the next control time step.

For solving SDP problems, cold-start methods initialize from an arbitrarily chosen vector, which leads to uncertain complexity and long computation time (Sturm, 2002). The warm-start or hot-start techniques can be applied to reduce the computational expense in solving SDP problems (Biswas and Ye, 2006; Skajaa et al., 2013; Sun et al., 2017), especially in an MPC framework (Zeilinger et al., 2014). Different from the existing hot-start or warm-start methods that are trying to solve a simplified problem as an initial guess, the hot-start in this paper does not solve any extra problem but takes advantage of the previous optimal solution in an iterative way. We utilize the previous optimal solution as an initial starting point because the problem does not change, and we just add a new data point to the constraints and the data-driven uncertainty set for the reformulated SDP model. In addition, the optimal solution of the reformulated SDP model will not change dramatically from previous optimal solutions due to the driving comfort objective and the acceleration or declaration capability constraint during a very short time interval (i.e., 0.1 seconds). Noted that our hot-start method does not sacrifice the solution quality since the optimality criterion is the same as the cold-start method.

In the hot-start method, we first input the previous primal and dual optimal solution at time step k-1 as the initial starting point for the SDP model at time step k. Noted that if it is the first time step, we just use an arbitrarily chosen start point. Then, we solve the reformulated SDP with the initial starting point from the last time step k-1 and the selected γ_1 and γ_2 using an SDP solver. We save the optimal solution at the current time step k as the initial point for the next time step k+1. In step 11, we keep the optimal solution at step k as the initial point of Step 10 for step $k + 1$. The reformulated SDP can be solved using an SDP solver such as MOSEK(https://www.mosek.com/).

5. Numerical analysis

In this section, we use the real-world data from the SPMD dataset to evaluate and validate the proposed DRSO-DRCC model for connected and automated driving on signalized arterials. In Section 5.1, we describe the dataset and the steps to extract signal timing information from the SPaT data and calibrate the IDM model from the connected vehicle trajectory data in the SPMD dataset. Then, we design five experiments to evaluate the implemented connected and automated driving control model in Section 5.2 - 1) evaluate DRSO-DRCC model performance under different traffic conditions; 2) measure the model performance of the proposed DRSO-DRCC model using different energy weighting factors; 3) computation time of the proposed DRSO-DRCC model under different prediction horizon lengths; and 4) duality gap of the relaxed SDP and the computation time comparison with the original SDP.

5.1. Overview of the SPMD data

To illustrate the proposed method, we use SPaT data (i.e., signal phasing and timing information) and Data Acquisition Systems (DAS) data (i.e., vehicle trajectories with radar sensing information) from the SPMD dataset. It is a high-frequency (0.1 seconds) connected vehicle dataset collected over two separate months, in which around 2,800 connected vehicles were deployed. We use the network of a signalized arterial with six signalized intersections on Fuller Road in Ann Arbor, Michigan (Fig. 11). In the SPMD project, the six SPaT-enabled RSUs are installed at the six intersections for coordinated actuated signal control (Bezzina and Sayer, 2015). We focus on the eastbound path on this signalized arterial with a 2.1-mile driving distance. Among the SPMD dataset, we only select SPaT data of the six RSUs and the DAS data of 50 vehicle trajectories driving through this arterial with 125,410 observations. The SPMD dataset is published on the ITS Public Data Hub and can be downloaded from the U.S. Department of Transportation (USDOT) website provided by the USDOT Intelligent Transportation Systems (ITS) Joint Program Office (2018).

5.1.1. Set the coordinated actuated control parameters using the SPaT data

To set the coordinated actuated control for the six intersections in the experiments, we simulate the SPaT information using the calculated parameters from the broadcast SPaT messages (in the SPMD dataset) by the RSUs at the six signalized

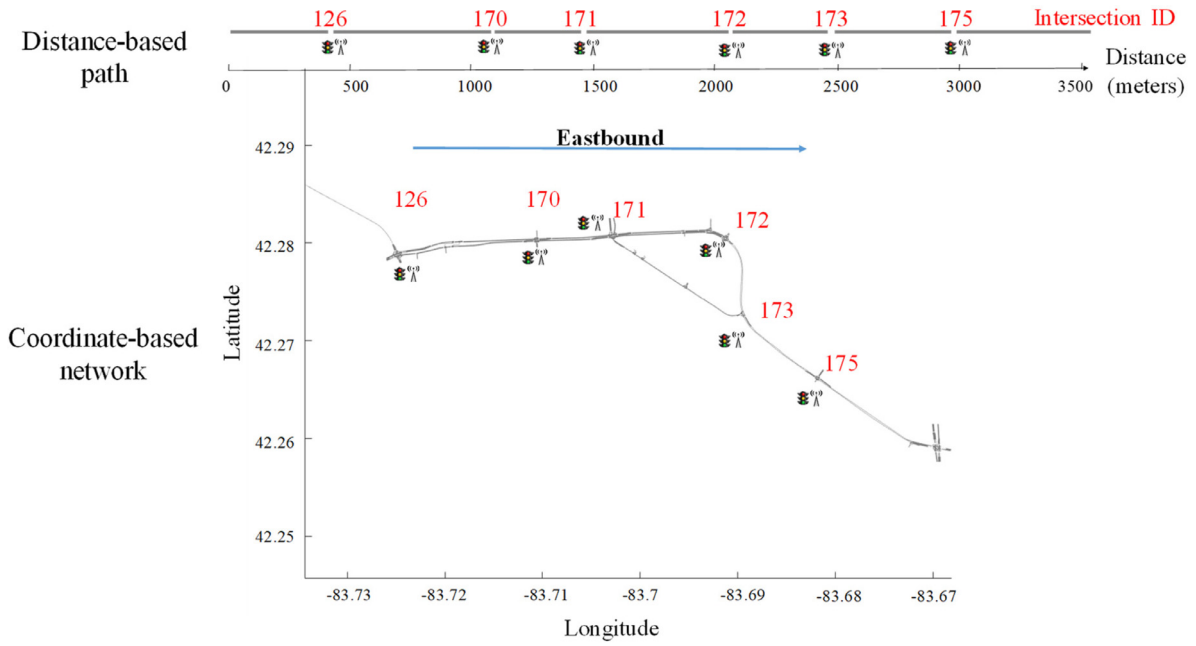

Fig. 11. The signalized arterial network with six RSUs on Fuller Road in Ann Arbor, Michigan.

Table 1
The actuated signal control settings of the six signalized intersections in the experiments (12:00 AM – 12:05 AM on April 2, 2013).

Intersection ID	Cycle lengths (s)	Offsets (s)	Maximum green times (s)	Minimum green times (s)	Yellow intervals (s)
126	55	0	35	4	4
170	55	49	37	4	4
171	55	28	40	4	4
172	55	38	35	4	4
173	55	28	35	4	4
175	55	32	25	4	4

intersections (Fig. 11) on Fuller Road in Ann Arbor, Michigan. The raw SPaT and SPaTMovement files in the SPMD dataset contain intersection ID, signal phase, signal status, minimum time remaining before a traffic signal change, maximum time remaining before a traffic signal change, and timestamps. The data in the current signal state column are represented by 8-bit hexadecimal values and indicate control signal states on different lanes. However, there some missing records and invalid messages in the raw SPaT and SPaTMovement files. Among the very few valid and continuous raw SPaT messages at the six intersections, we selected a segment of continuous SPaT records over five minutes during 12:00 AM – 12:05 AM on April 2, 2013.

From the raw SPaT and SPaTMovement files in the SPMD dataset over the five minutes, we calculate actuated signal control parameters such as the cycle lengths, offsets, minimum green times, maximum green times, and yellow intervals for the coordinated actuated control for the six signalized intersections in the experiments (Table 1).

To simulate the traffic signal control of our experiments, we implement a broadcast module and the VISSIM Vehicle Actuated Programming (VAP) signal control interface using VISSIM-COM interface in C++. We configure the VAP signal control using the calculated actuated signal control parameters and implement the traffic-responsive logic in VISSIM through VISSIM-COM interface (PTV, 2008). Then, we start the VISSIM simulation in our C++ COM program and activate the signal control based on the configured VAP module. In the simulation, we broadcast the simulated SPaT information to the controlled vehicle.

5.1.2. Calibrate IDM model for simulating traffic conditions using the DAS data

The DAS data in the SPMD dataset includes vehicle trajectories with in-vehicle radar sensing information of 50 vehicles. Each record of DAS data contains anonymous ID, longitude, latitude, velocity, front vehicle distance, front vehicle speed, and timestamp. We use the map-matching method in (Zhao and Zhang, 2017, 2019) to match the DAS vehicle trajectories (with radar sensing data) to the transportation network of Fuller Road (Fig. 11). From the map-matched result, we only select vehicle trajectories that were driving through this signalized arterial.

Among the selected vehicle trajectories in the DAS dataset, we select 50 driving trajectories of vehicles equipped with GPS and radar devices over one week (between October 11, 2013, and October 18, 2013). There are a total of 125,410 observations

Table 2
The calibrated IDM model parameters using the DAS data from the SPMD project.

Parameter	Description	Calibrated value	Unit
a	Maximum acceleration	2.45	m/s^2
b	Desired deceleration	-3.88	m/s^2
V_{des}	Desired velocity	18.13	m/s
s_0	Minimum net stopped distance from the preceding vehicle	2.04	m
T_{hw}	Desired safety time headway	0.95	s

of the 50 vehicle trajectories. With these vehicle trajectories (with radar sensing data), we use Eq. (2) in Section 3.2.2 to calibrate IDM model parameters for vehicle driving behaviors on Fuller Road in Ann Arbor, Michigan. In our experiments, all the preceding vehicles are simulated to drive along the signalized arterial (i.e., Fuller Road) using the IDM model with the calibrated parameters (Table 2).

5.2. Experimental analysis

In this section, we evaluate the performance of the proposed DRSO-DRCC model for connected and automated driving using five experiments (i.e., different traffic conditions, and different VOEs). In the experiments, we simulate different driving conditions by generating a different number of vehicles in the network (Fig. 11) where their driving behaviors are controlled by the calibrated IDM model (Table 2). For the traffic signal control of the six intersections on the network, we code the coordinated actuated control using VAP module in VISSIM (PTV, 2008) with the SPaT data (Table 1).

Simulation experiments have been employed in the literature of stability analysis of car-following models for both traditional and connected environments (Sun et al., 2018). We implement the solution algorithm (Section 4.3) in C++ and send real-time driving commands to control the vehicles in VISSIM through COM interface programming (control the network-based path choices) and VISSIM External Driver Model (control the microscopic driving behaviors). In the two experiments, we only control one CAV using the proposed DRSO-DRCC model for connected and automated arterial driving. We also use the calibrated IDM model to control the CAV as the benchmark and use the CMEM energy consumption model (see Barth and Boriboonsomsin, 2009) in the Eco-driving policy (Section 3.2.3) to compare with the performance of the proposed DRSO-DRCC model. We also investigate the computation time of the proposed model under different prediction horizon lengths. To compare the relaxed SDP with the original SDP problem, we use another experiment to show the duality gap and computation time comparison between the relaxed SDP and the original SDP.

5.2.1. Model performance of the connected and automated arterial driving under different traffic conditions

To demonstrate the proposed model, we compare the results of the proposed DRSO-DRCC based MPC model as **automated driving** of the controlled CAV under different traffic conditions. We use the calibrated IDM model to simulate the controlled vehicle driving behaviors of HDVs as **human driving** under different traffic conditions. We design three scenarios for the different traffic conditions: (1) light traffic, (2) medium traffic, and (3) heavy traffic. We use three different OD matrices for all the 14 activity locations (i.e., two activity locations for each intersection and two activity locations for Fuller Road) as the input demand for the simulation. By using the three OD matrices, we add 85, 360, and 500 vehicles into the Fuller Road network (with 2.1-mile driving distance) to simulate different traffic conditions for the three scenarios, respectively.

In this experiment, all the preceding vehicles, in addition to the IPV, are controlled by the calibrated connected IDM model using the DAS data from the SPMD dataset. We apply the IDM model in the learning-based prediction of the IPV to predict its driving states. We use the constrained finite-horizon optimal control model in Section 3.2.3 as the Eco-driving policy to predict the reference driving states for the controlled vehicle. We set the minimum safe time headway to guarantee the following safety of the controlled vehicle based on the experiments in González-Villaseñor et al. (2007). We use 5.6 meters as the minimum safe spacing according to the minimum spacing in the experiments in Nowakowski et al. (2016) and Darbha et al. (2018). The prediction horizon is calculated as a time-variant parameter, as in Eq. (14). Without loss of generality, we use 0.8 as the confidence levels γ_1 and γ_2 for the two DRCC constraints, and use 0.85 as the confidence levels γ_3 and γ_4 for the first and second moment constraints in the data-drive uncertainty set of the proposed DRSO-DRCC model. The VOEs for the three different traffic conditions are set to be 0.25.

In Fig. 12, the trajectory profile under light traffic shows that the controlled vehicle has a smooth acceleration process when the vehicle is driving through the six intersections. The trajectory profile under medium and heavy traffic shows that the idling time of the controlled vehicle is reduced at the intersections. Under the three traffic conditions, the trajectory profiles of the controlled vehicle show that the proposed DRSO-DRCC model intentionally slows down the controlled vehicle before it is driving through the intersections under red signals, which mitigate decelerations and reduce time delays of the red signal. On the contrary, the IDM model stops the controlled vehicle under red signals with high deceleration perturbations.

To demonstrate that the vehicle trajectory determined by the proposed DRSO-DRCC model can be reversed back to the given coordinate-based network, we plot the controlled vehicle trajectory in the space-time network in Fig. 13. The location records when the controlled vehicle passes the intersections can be projected to the arterial network in the space

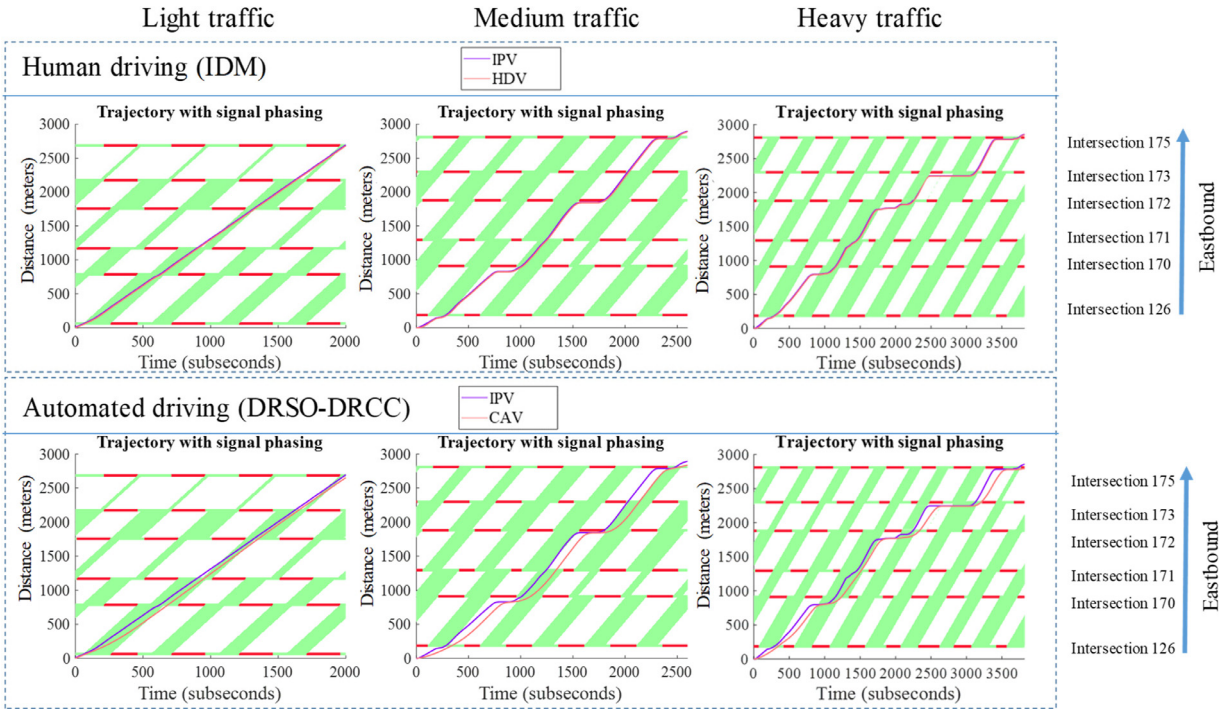

Fig. 12. Vehicle trajectories with traffic signal timing information.

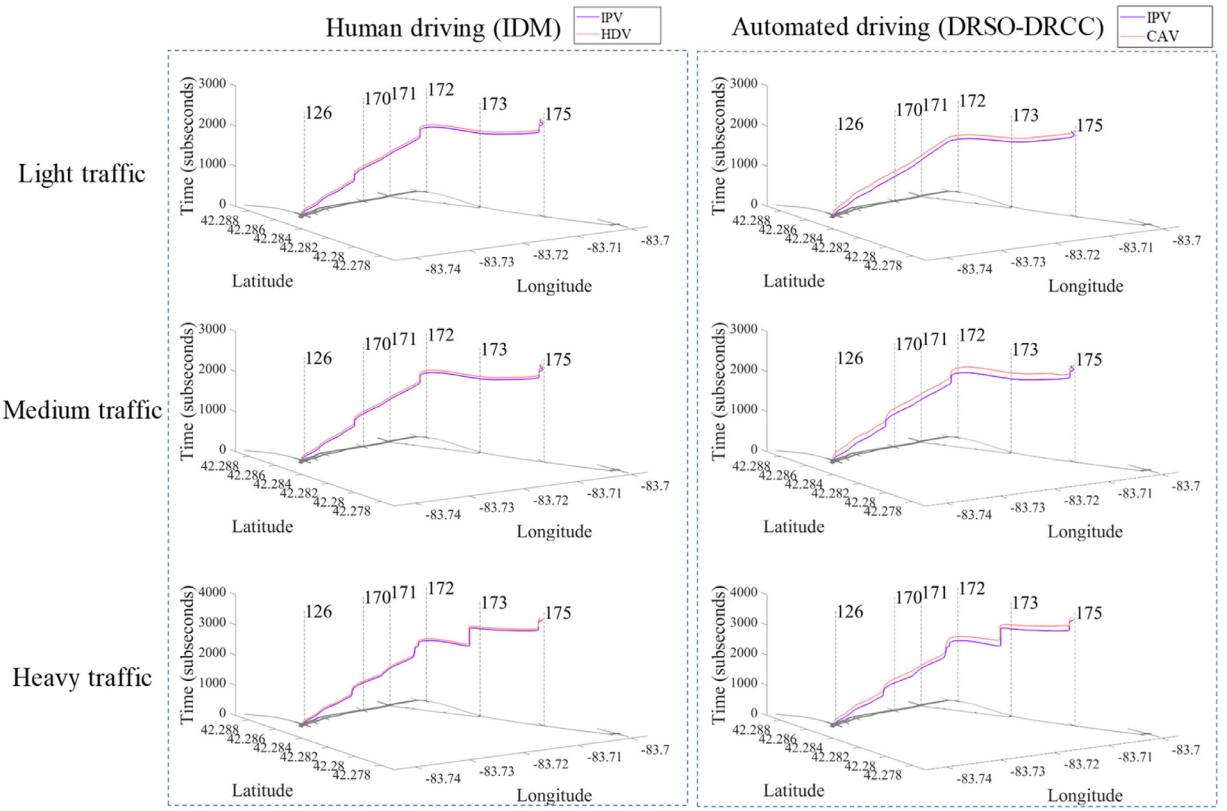

Fig. 13. The space-time trajectories of the controlled vehicle by the IDM and DRSO-DRCC models.

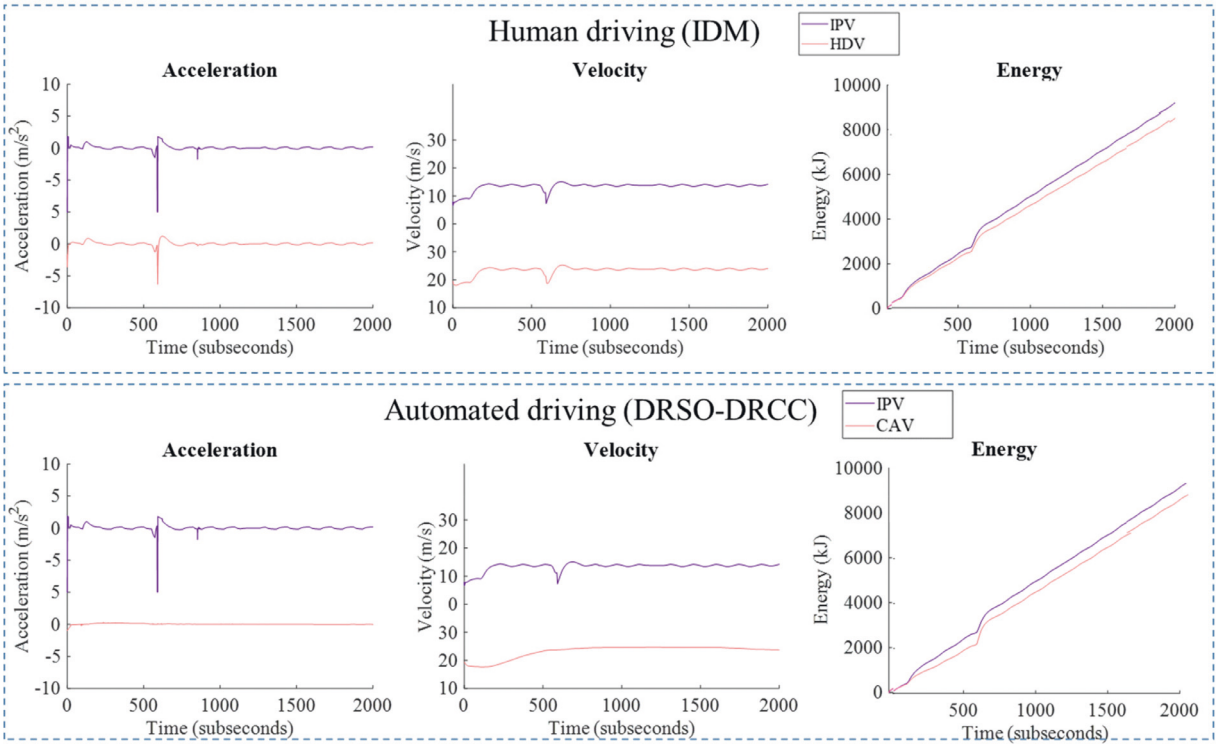

Fig. 14. Comparison of acceleration, velocity, and energy consumption profiles between the IDM and the DRSO-DRCC model under the light traffic condition.

dimension. Using the method described in Section 3.1, the distance-based trajectories in Fig. 12 can be converted to the coordinate-based trajectories in Fig. 13. Noted that the arterial network is shown in the space dimension (i.e., the longitude and latitude), in which the space-time trajectories of the controlled vehicle can be exactly projected. In addition, the controlled vehicle trajectories of the DRSO-DRCC model are smoother when approaching the signalized intersections compared with the IDM model, especially in the medium and heavy traffic conditions in Fig. 13.

In Fig. 14, we plot the acceleration, velocity, trajectory, and energy consumption profiles of the controlled vehicle and its IPV in the time domain under the light traffic condition. It is shown that the acceleration and velocity perturbations of the controlled vehicle of the DRSO-DRCC are smoother than its benchmark (i.e., the IDM-controlled vehicle).

For the energy efficiency in Fig. 15, the DRSO-DRCC model generates lower energy consumptions than the calibrated IDM model for the same driving distance. The smooth acceleration and velocity profiles and the lower energy consumption are caused by the Eco-driving reference that considers driving comfort and energy saving in the DRSO-DRCC model. The results show that the DRSO-DRCC model saves 7.1% energy for the controlled vehicle driving through this signalized arterial under the light traffic condition.

For the medium traffic conditions, the velocity profile of the controlled vehicle shows that the controlled vehicle has three stops at three intersections (i.e., intersections 126, 171, and 175 in Fig. 15) in the IDM controlled case. However, the DRSO-DRCC controlled vehicle just has one fully stop at intersection 171 around simulation time 1500 subseconds. In addition, the acceleration perturbations of the controlled vehicle by the DRSO-DRCC model are smaller than the controlled vehicle by the connected IDM model. The energy consumption using the DRSO-DRCC model saves 18.1% energy of the controlled vehicle compared with the energy consumption using the connected IDM model.

Under the heavy traffic condition, the acceleration profile of the controlled vehicle by the DRSO-DRCC model is much smoother than the controlled vehicle by the IDM model (Fig. 16). The velocity profiles show that the proposed DRSO-DRCC model reduces the idling time at intersections (especially around 700 subseconds) compared to the controlled vehicle by the IDM model. The energy consumption profile shows that the controlled vehicle by the DRSO-DRCC model has much lower energy consumption when driving through the signalized arterial than the case by the IDM model. The DRSO-DRCC model saves 23.4% energy consumption for the controlled vehicle compared to the IDM model under the heavy traffic condition.

The results of the three traffic conditions consistently show that the controlled vehicle's acceleration can be optimized by the proposed DRSO-DRCC model to reduce acceleration perturbations, velocity variations, and time delays along the signalized arterial.

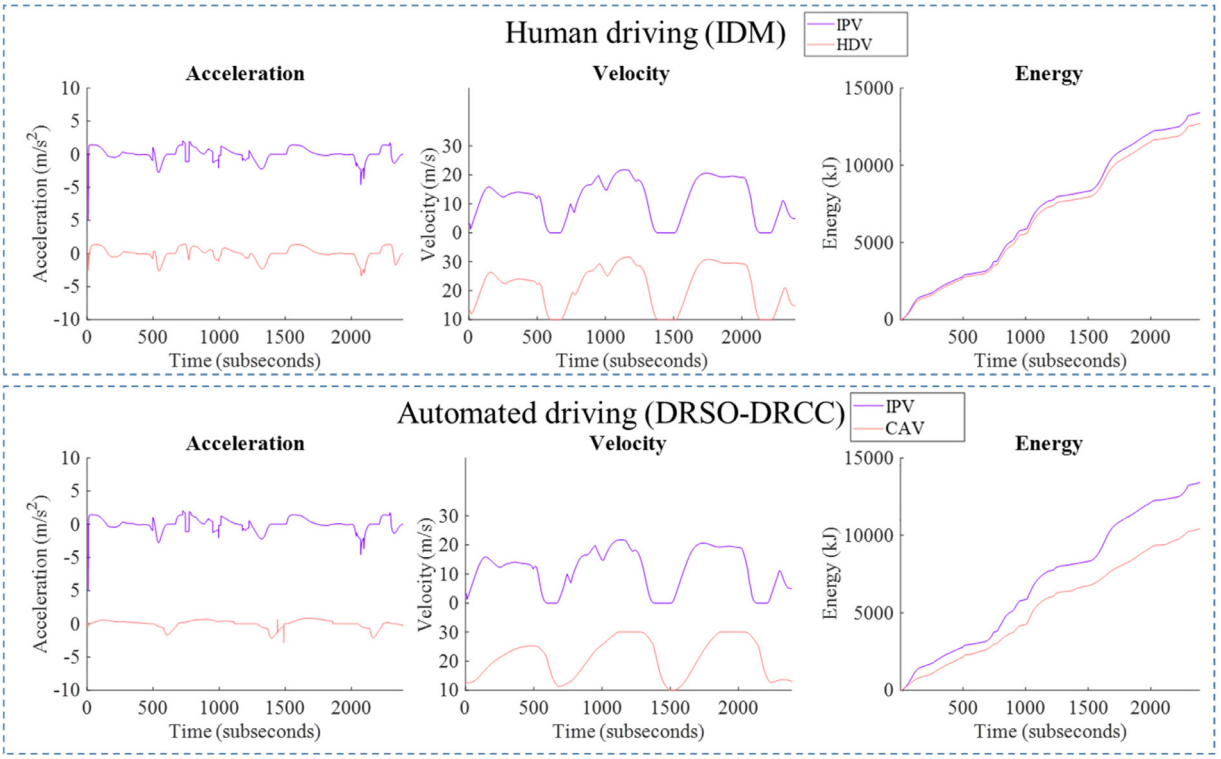

Fig. 15. Comparison of acceleration, velocity, and energy consumption profiles between the IDM and the DRSO-DRCC model under the medium traffic condition.

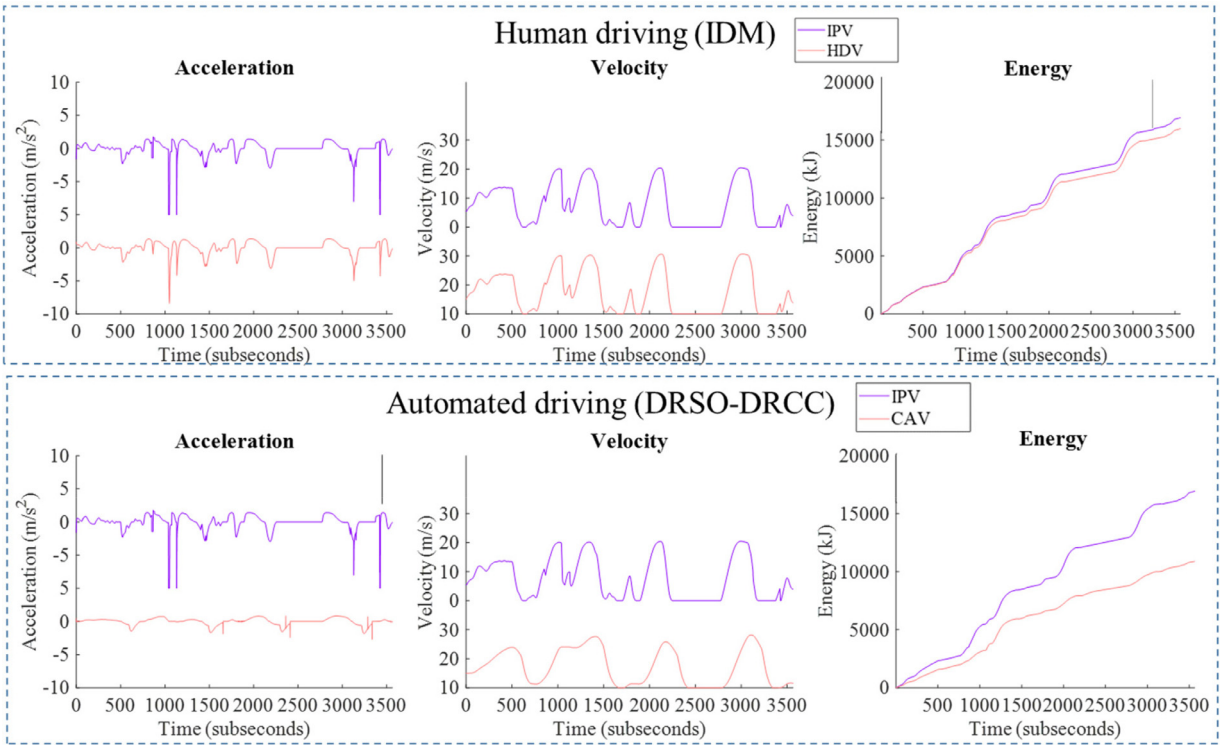

Fig. 16. Comparison of acceleration, velocity, and energy consumption profiles between the IDM and the DRSO-DRCC model under the heavy traffic condition.

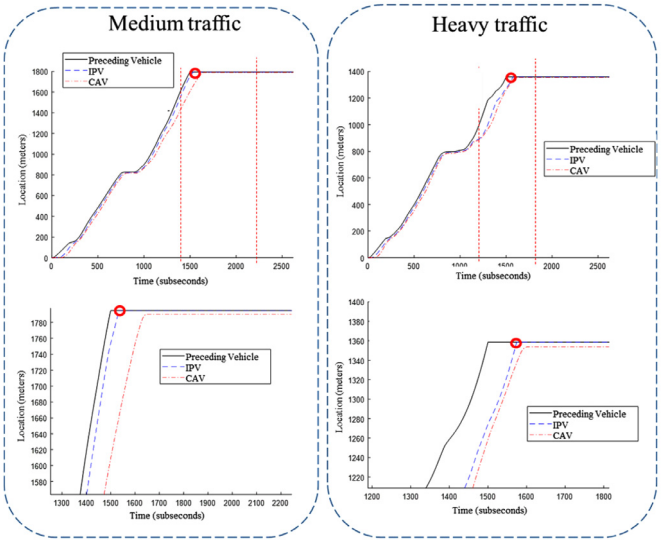

Fig. 17. Vehicle trajectories with one incident under medium and heavy traffic conditions.

Table 3
The acceleration ($\boldsymbol{m/s^2}$) variance, velocity ($\boldsymbol{m/s}$) variance, and energy savings under different VOEs.

Scenario	Acceleration (variance)	Velocity (variance)	Energy Saving (%)
Human Driving (IDM)	0.829	9.766	0
Automated Driving (DRSO-DRCC) VOE=0	0.331	5.919	4.493
Automated Driving (DRSO-DRCC) VOE=0.25	0.068	4.795	7.120
Automated Driving (DRSO-DRCC) VOE=0.5	0.067	4.35	7.346
Automated Driving (DRSO-DRCC) VOE=0.75	0.034	4.966	7.698

5.2.2. Model performance of connected and automated arterial driving under incidents

To evaluate the model performance for connected and automated arterial driving under some uncertain conditions, we design an experiment of connected and automated driving under traffic incidents. Similar to the vehicle collision experiment in Wang et al. (2015), we include an incident during the simulation. We apply an erroneous acceleration command to the IPV such that it hits its preceding vehicle. The incident happens at 1500 subseconds during the control time. In the upper two plots of Fig. 17, we show the trajectories of two preceding vehicles and the controlled vehicle. Noted that there are ten other front vehicles, including the boundary preceding vehicle in front of the IPV in the simulation, within a 300-meter communication range. We also select a period when the incident happens in the lower two plots in the figure. In this experiment, the controlled vehicle avoids collisions under both medium and heavy traffic. During the control period, our proposed approach maintains a minimum safe time headway τ^{min} and a minimum safe spacing s^{min} for the controlled vehicle to guarantee safety.

5.2.3. Model performance of the connected and automated arterial driving with different VOEs

In this section, we select the light traffic condition (where vehicles can travel through the six intersections without stop) to evaluate the performance of the proposed DRSO-DRCC based MPC using the value of energy (VOE) factors of 0, 0.25, 0.5, and 0.75. The performance of the controlled vehicle will be measured in terms of vehicle driving states (acceleration, velocity, and trajectory) and energy efficiency benefits. Noted that, in Ma et al. (2017), they use a minimum 0.1 energy weighting factor. However, in our DRSO-DRCC model, we also consider safety, mobility, and driving comfort in addition to energy efficiency. We use a minimum 0.25 weighting factor for $(1 - \alpha)$ for vehicle driving considerations (i.e., safety, mobility, and driving comfort).

In Fig. 18, we plot the acceleration, velocity, and energy consumption profiles of the controlled vehicles determined by the DRSO-DRCC (with different VOEs) and the connected IDM. The acceleration and velocity profiles from the DRSO-DRCC model using different VOEs are all smoother than the controlled vehicle by the connected IDM model. When we increase the VOE for the DRSO-DRCC model, the acceleration and velocity profiles are slightly different, which may be caused by the light traffic volume under coordinated actuated control on the signalized arterial. However, the energy consumption profiles still show a reduced value when the VOE is larger.

For more detailed comparisons between the DRSO-DRCC model and the connected IDM model in the experiment, we show the acceleration variance, velocity variance, and energy saving of the IDM model and the DRSO model using different VOEs in Table 3. When the controlled vehicle is commanded by the IDM model, the perturbations of acceleration and

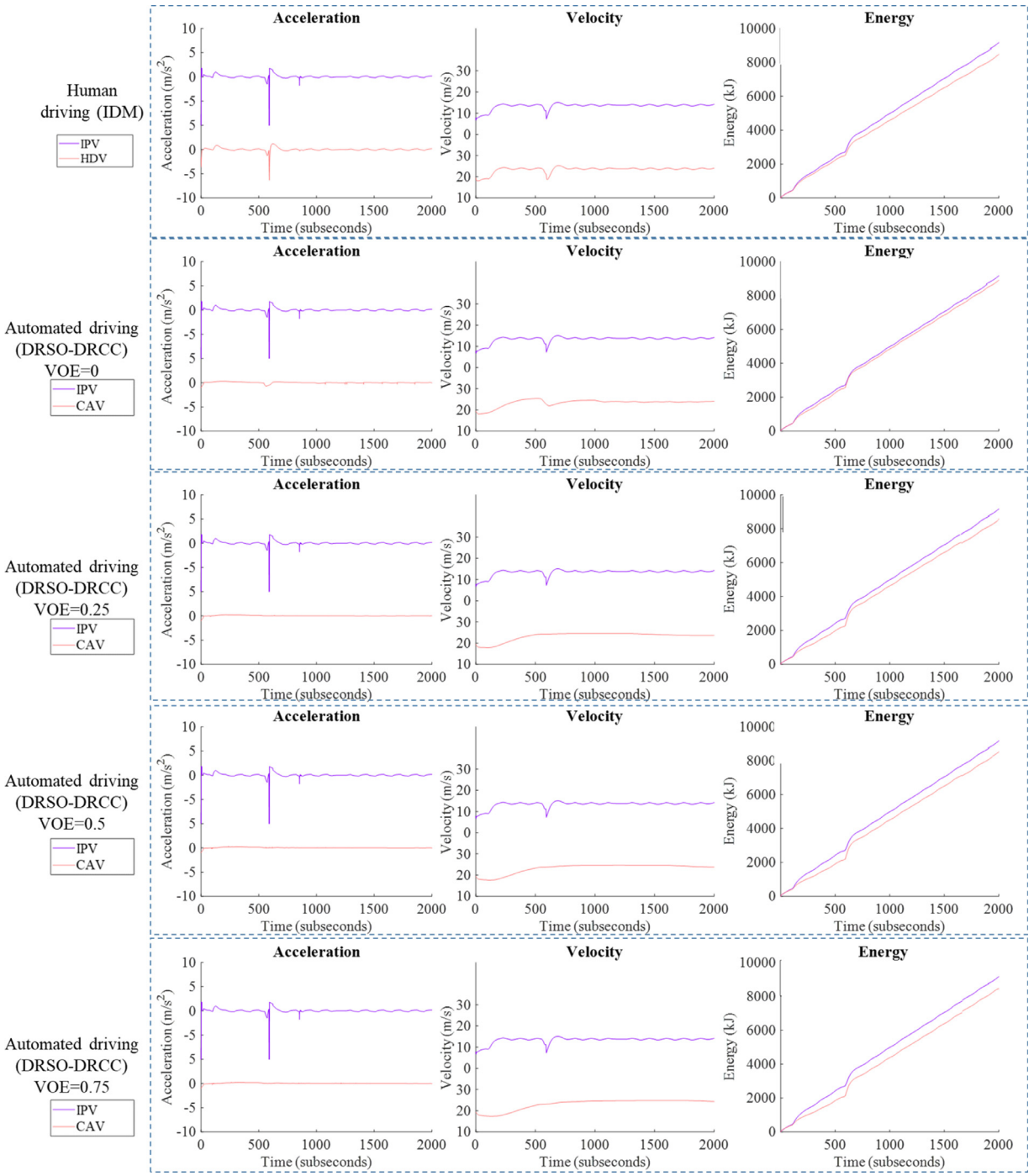

Fig. 18. Comparison of acceleration, velocity, and energy consumption profiles between the IDM and the DRSO-DRCC model with different VOEs under the light traffic condition.

velocity (i.e., the variance) of the controlled vehicle are the largest among all the scenarios. The acceleration and velocity perturbations of the controlled vehicle reduce as we set larger VOEs for the DRSO-DRCC model. When we set the VOE to be no less than 0.25, the DRSO-DRCC model saves more than 7.1% energy consumptions for the controlled vehicle without sacrificing travel time. The results in Table 3 demonstrate the effectiveness of the proposed DRSO-DRCC model to reduce energy consumptions and improve vehicle trajectory smoothness and driving comfort even with zero VOE.

With the data-driven optimization-based online control framework, the proposed connected and automated driving control model determines acceleration or deceleration rates for the controlled vehicle by dynamically update the learning-based

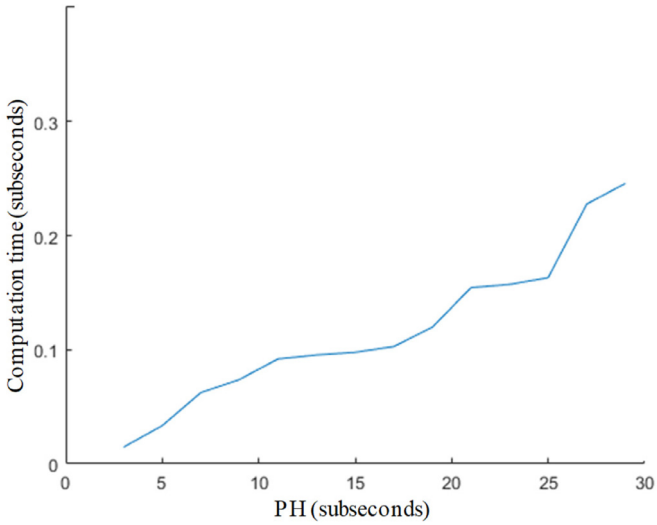

Fig. 19. Computation time of the proposed model.

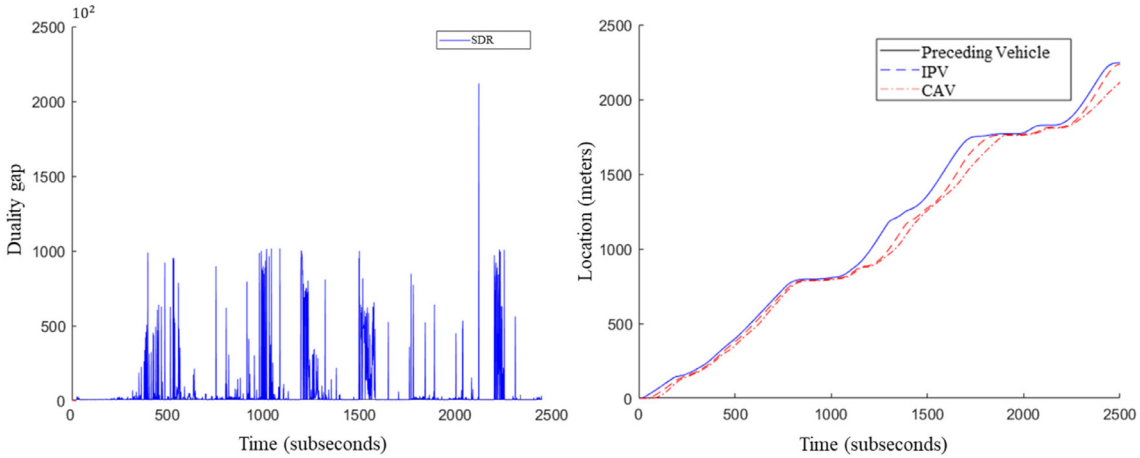

Fig. 20. Duality gap of the SDR problem and the vehicle trajectories.

predictions of its preceding vehicles. With the predictions, the proposed method online generates an Eco-driving reference as the tracking driving reference for the DRSO-DRCC model. Under this data-driven optimization-based online control framework, the predicted information will be updated online such that the controlled vehicle is able to reduce time delays and energy consumptions on signalized arterials.

5.2.4. The computation time of different prediction horizon lengths

In this section, we study the computation time of the proposed model. The two proposed solution methods for solving the proposed model are implemented in C++ and run on a computer with Intel Xeon E5-2699 v4 2.2 GHz CPU (22 cores) and 192 GB RAM. For the experiment, we use the connected IDM model to predict the driving states of the IPV and the Eco-driving policy to predict the driving states of the controlled vehicles.

We use 0.8-second minimum safe time headway in this experiment. We select the prediction horizon length from 0.2 to 3.0 seconds. We conduct 50 simulation runs for each of the prediction horizon lengths to get the average computation time in Fig. 19. The computation time is lower than 0.1 seconds when the prediction horizon is not greater than 1.5 seconds in our experiments. If the prediction horizon is smaller than 0.8 seconds, then the computation time will be shorter than 0.05 seconds. The computation time will increase if the duration of the prediction horizon increases, as shown in Fig. 19. But we can have at least 10 time steps in our prediction horizon in our results for real-time applications.

The computation time may be further reduced by using computers with high-speed CPUs such as Intel Core i7- 8086k CPU (4.0 GHz clock rate) and 5.0 GHz max turbo clock rate) and Intel Xeon E3-1280 v6 CPU (3.9 GHz clock rate and 4.2 GHz max turbo clock rate).

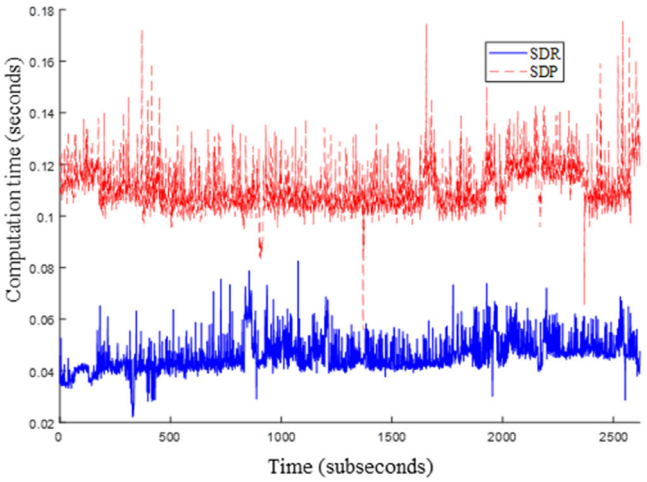

Fig. 21. Computation times of SDR and SDP.

5.2.5. Duality gap of the relaxed SDP and the computation time compared with the original SDP

In this section, we run the original SDP problem and the relaxed SDP (i.e., SDR) problem under the heavy traffic scenario, respectively, to investigate the duality gap. We use a prediction horizon length of 8time steps in this experiment.

In Fig. 20, results show that the duality gap of the solution of our SDR problem is stable (i.e., from 0.01 to 10) under most of the control period. The duality gap increases only when there is a velocity fluctuation from the IPVs (e.g., the traffic shock at 1100 subseconds). the original objective function has the quadratic terms of velocity in $^{m}/_{s}$ and the acceleration in $^{m}/_{s^2}$. In general, the duality gap of the SDR is under $1 \times 10^2$for the heavy traffic scenario over the control period.

We run the SDR and SDP on a computer with Intel Xeon E5-2699 v4 2.2 GHz CPU (22 cores) and 192 GB RAM. From the computation times of the SDR and the SDP in Fig. 21, the computation time of the SDR is always shorter than the SDP. With a prediction horizon length of 8 steps, the computation time of SDR fluctuates around 0.05 seconds. We are noted that the CPU frequency of the computer we are using is relatively low (2.2 GHz). We think the computation time can be further reduced even with a longer prediction horizon length using a more powerful CPU such as Intel Core i7- 8086k CPU (4.0 GHz clock rate) and 5.0 GHz max turbo clock rate).

6. Conclusion

In this paper, we integrate a data-driven optimization-based MPC-based control model with traffic control devices (e.g., road markings, traffic signs, and traffic signals) and road geometry (e.g., road shapes, road boundaries, and road grades) constraints for an online connected and automated driving problem on signalized arterials under traffic uncertainties. In the MPC-based control models, we mathematically formulate the location-based traffic control devices and road geometry constraints from the high-definition network data in HD maps. The location-based traffic control devices and road geometry constraints explicitly consider interrupted locations and road geometry information. With the location-based traffic control devices and road geometry constraints from HD maps, the robustness and safety of connected and automated driving are improved because the control model will not only rely on in-vehicle sensors but also utilize the HD map data to capture the interrupted flow facility locations.

Based on the integrated data-driven optimization-based MPC model with location-based traffic control devices and road geometry constraints, the proposed connected and automated driving control framework includes four parts for the online control system. First, we predict a set of uncertain driving states of the first (farthest) preceding vehicle (i.e., as the boundary driving conditions) within the V2V communication range of the controlled vehicle. Second, we develop an online learning-based prediction method to individually calibrate the car-following model (e.g., the connected IDM) parameters. Using the online calibrated parameters, we predict a set of uncertain driving states of the IPV based on the concept of forecasting the forecasts of others. Third, based on the set of predicted driving states of the IPV, we predict a set of driving state reference of the controlled vehicle by an Eco-driving policy. Fourth, given the predicted driving states of the IPV and the predicted driving state reference (i.e., the Eco-driving reference) of the controlled vehicle, we can obtain the optimal acceleration or deceleration commands for the controlled vehicle by solving the proposed DRSO-DRCC model at each control step. To solve the proposed DRSO-DRCC model, we reformulate a relaxed SDP problem by applying the strong duality theory and the Semidefinite Relaxation technique. Then, we present a solution algorithm for solving the relaxed problem of the dual problem as an SDP using parallel computing and hot-start techniques.

To demonstrate the DRSO-DRCC model, we conduct two numerical experiments to evaluate the safety, smoothness, and energy efficiency using real V2V connected vehicle trajectory (DAS) data from OBUs and V2I SPaT data from RSUs in the

SPMD dataset on Fuller Road in Ann Arbor, Michigan. We use the DAS vehicle trajectory data with radar sensing information to calibrate the connected IDM model for simulating the uncontrolled vehicles as the benchmark human driving in the experiments. For the signal phasing and timing in the experiments, we use the SPaT data of the six signalized intersections on Fuller Road for the coordinated actuated control. The first experiment shows that the proposed DRSO-DRCC model generates smoother trajectories with shorter idling time at intersections, and saves 7.1% to 23.6% energy consumption compared to the connected IDM model under three different traffic conditions (i.e., light, medium, and heavy traffic). The second experiment indicates that the proposed model can avoid collisions of the controlled vehicle when collisions happen in its preceding vehicles. The third experiment shows that a higher VOE factor will lead to larger energy savings (compared with the connected IDM model) for the controlled vehicle driving through the signalized arterial. The fourth experiment evaluates the computational complexity of the proposed method. The fifth experiments analyze the duality gap of the relaxed SDP model. In the future, the proposed method can be extended in two ways: (1) by applying customized implementations to reduce the computation time of the proposed model; (2) by incorporating a communication delay inside the control loop using a delay compensation method.

Acknowledgments

This research was supported by a grant from the National Science Foundation award https://www.nsf.gov/awardsearch/showAward?AWD_ID=1846795, CMMI- 1846795: CAREER: Tackling Congestion in Smart Cities via Data-Driven Optimization-Based Control of Connected and Automated Vehicles. The authors are grateful to the anonymous referees and associate editors for helpful comments that have considerably improved the manuscript. The authors would also like to thank Qinjie Lyu for her assistance in preparing the figures and tables in the Numerical Analysis section. We also acknowledge Dr. James R. Sayer at the University of Michigan Transportation Research Institute (UMTRI) to share the SMPD data with us. Naturally, the authors are solely responsible for the content of the paper.

References

Asadi, B., Vahidi, A., 2009. Predictive use of traffic signal state for fuel saving. IFAC Proceedings Volumes 42 (15), 484–489.
Barth, M., An, F., Younglove, T., Scora, G., Levine, C., Ross, M., Wenzel, T., 2000. Comprehensive Modal Emission Model (CMEM), version 2.0 user's guide. University of California, Riverside.
Barth, M., Boriboonsomsin, K., 2009. Energy and emissions impacts of a freeway-based dynamic Eco-driving system. Transport. Res. Part D: Transport Environ. 14 (6), 400–410.
Barth, M., Mandava, S., Boriboonsomsin, K., Xia, H., 2011. Dynamic Eco-driving for arterial corridors. In Integrated and Sustainable Transportation System (FISTS), 2011 IEEE Forum, pp. 182-188.
Boyd, S., Vandenberghe, L., 2004. Convex optimization. Cambridge University Press.
Bouton, M., Nakhaei, A., Fujimura, K. and Kochenderfer, M.J., 2019, June. Safe reinforcement learning with scene decomposition for navigating complex urban environments. In 2019 IEEE Intelligent Vehicles Symposium (IV) (pp. 1469-1476). IEEE.
Bazaraa, M.S., Sherali, H.D., Shetty, C.M., 2013. Nonlinear programming: theory and algorithms. John Wiley & Sons.
Bertsimas, D., Popescu, I., 2005. Optimal inequalities in probability theory: A convex optimization approach. SIAM J. Optim. 15 (3), 780–804.
Bezzina, D., Sayer, J., 2015. Safety pilot model deployment: Test conductor team report. (Report No. DOT HS 812 171). National Highway Traffic Safety Administration, Washington, DC.
Biswas, P., Ye, Y., 2006. A distributed method for solving semidefinite programs arising from ad hoc wireless sensor network localization. In: Multiscale optimization methods and applications. Springer, Boston, MA, pp. 69–84.
Calafiore, G.C., El Ghaoui, L., 2006. On distributionally robust chance-constrained linear programs. J. Optim. Theo. Appl. 130 (1), 1–22.
Calafiore, G.C., Fagiano, L., 2013. Robust model predictive control via scenario optimization. IEEE Trans. Autom. Contr. 58 (1), 219–224.
Carvalho, A., Gao, Y., Gray, A., Tseng, H.E., Borrelli, F., 2013. Predictive control of an autonomous ground vehicle using an iterative linearization approach. In Intelligent Transportation Systems-(ITSC), 2013 16th International IEEE Conference on IEEE. pp. 2335-2340.
Calafiore, G.C., El Ghaoui, L., 2014. Optimization models. Cambridge University Press.
Dupuis, M., Grezlikowski, H., 2006. OpenDRIVE®-an open standard for the description of roads in driving simulations. Proc. Driv. Simul. Conf. 25–36.
Delage, E., Ye, Y., 2010. Distributionally robust optimization under moment uncertainty with application to data-driven problems. Oper. Res. 58 (3), 595–612.
Demir, E., Bektaş, T., Laporte, G., 2011. A comparative analysis of several vehicle emission models for road freight transportation. Transp. Res. Part D Transp. Environ. 16 (5), 347–357.
Dixit, S., Fallah, S., Montanaro, U., Dianati, M., Stevens, A., Mccullough, F., Mouzakitis, A., 2018. Trajectory planning and tracking for autonomous overtaking: State-of-the-art and future prospects. Ann. Rev. Contr..
Focacci, F., Lodi, A., Milano, M., Vigo, D., 1999. Solving TSP through the integration of OR and CP techniques. Electron. Note. Discrete Math. 1, 13–25.
Floudas, C.A., Visweswaran, V., 1995. Quadratic optimization. In Handbook of global optimization. Springer, Boston, MA., pp. 217–269.
Goh, J., Sim, M., 2010. Distributionally robust optimization and its tractable approximations. Oper. Res. 58 (4-NaN-1), 902–917.
Hausberger, S., Engler, D., Ivanisin, M., Rexeis, M., 2002. Update of the emission functions for heavy duty vehicles in the handbook emission factors for road traffic. Inst. Intern. Combust. Eng. Thermodyn. 2002–2003.
He, X., Liu, H.X., Liu, X., 2015. Optimal vehicle speed trajectory on a signalized arterial with consideration of queue. Transport. Res. Part C: Emerg. Technol. 61, 106–120.
Hao, P., Wu, G., Boriboonsomsin, K., Barth, M.J., 2015. Developing a framework of eco-approach and departure application for actuated signal control. In Intelligent Vehicles Symposium (IV), 2015 IEEE, pp. 796-801.
Hooker, J.N., Ottosson, G., Thorsteinsson, E.S., Kim, H.J., 1999. On integrating constraint propagation and linear programming for combinatorial optimization. AAAI/IAAI 136–141.
Jiang, R., Guan, Y., 2016. Data-driven chance constrained stochastic program. Math. Prog. 158 (1–2), 291–327.
Jiang, H., An, S., Wang, J., Cui, J., 2018. Eco-approach and departure system for left-turn vehicles at a fixed-time signalized intersection. Sustainability 10 (1), 273.
Kamal, M.A.S., Mukai, M., Murata, J., Kawabe, T., 2010. On board eco-driving system for varying road-traffic environments using Model Predictive Control. CCA 1636–1641.
Katrakazas, C., Quddus, M., Chen, W.H., Deka, L., 2015. Real-time motion planning methods for autonomous on-road driving: State-of-the-art and future research directions. Transport. Res. Part C: Emerg. Technol. 60, 416–442.

Kim, S., Kojima, M., 2003. Exact solutions of some nonconvex quadratic optimization problems via SDP and SOCP relaxations. Comput. Optim. Appl. 26 (2), 143–154.
Kesting, A., Treiber, M., Helbing, D., 2010. Enhanced intelligent driver model to access the impact of driving strategies on traffic capacity. Philos. Trans. R. Soc. Lond. A Math. Phys. Eng. Sci. 368 (1928), 4585–4605.
Karlsson, J., Murgovski, N., Sjöberg, J., 2016. Temporal vs. spatial formulation of autonomous overtaking algorithms. In Intelligent Transportation Systems (ITSC), 2016 IEEE 19th International Conference on IEEE. pp. 1029-1034.
Li, X., Jiang, H., 2007. Solving large-margin hidden Markov model estimation via semidefinite programming. IEEE Trans. Audio Speech Lang. Process. 15 (8), 2383–2392.
Liu, H., Han, K., Gayah, V.V., Friesz, T.L., Yao, T., 2015. Data-driven linear decision rule approach for distributionally robust optimization of on-line signal control. Transport. Res. Part C: Emerg. Technol. 59, 260–277.
Lin, Y., McPhee, J., Azad, N.L., 2019. Longitudinal dynamic versus kinematic models for car-following control using deep reinforcement learning. In 2019 IEEE Intelligent Transportation Systems Conference (ITSC) (pp. 1504-1510). IEEE
Luo, Z.Q., Ma, W.K., So, A.M.C., Ye, Y., Zhang, S., 2010. Semidefinite relaxation of quadratic optimization problems. IEEE Sign. Process. Mag. 27 (3), 20–34.
Ma, W.K., Davidson, T.N., Wong, K.M., Luo, Z.Q., Ching, P.C., 2002. Quasi-ML multiuser detection using SDR with application to sync. CDMA. IEEE Trans. Signal Process 50 (4), 912–922.
Ma, W.K., Ching, P.C., Vo, B.N., 2004. Crosstalk resilient interference cancellation in microphone arrays using capon beamforming. IEEE Trans. Speech Audio Process. 12 (5), 468–477.
Ma, J., Li, X., Zhou, F., Hu, J., Park, B.B., 2017. Parsimonious shooting heuristic for trajectory design of connected automated traffic part II: computational issues and optimization. Transport. Res. Part B: Methodol. 95, 421–441.
Marshall, A., Olkin, I., 1960. Multivariate Chebyshev Inequalities. Ann. Math. Stat. 31 (4), 1001–1014.
Massow, K., Kwella, B., Pfeifer, N., Häusler, F., Pontow, J., Radusch, I., Hipp, J., Dölitzscher, F., Haueis, M., 2016. Deriving HD maps for highly automated driving from vehicular probe data. In: Intelligent Transportation Systems (ITSC), 2016 IEEE 19th International Conference on IEEE, pp. 1745–1752.
Ottosson, G., Thorsteinsson, E.S., 2000. Linear relaxation and reduced-cost based propagation of continuous variable subscripts. Proc. Int. WS. on Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems–CPAIOR'00, Paderborn Center for Parallel Computing, Technical Report tr-001-2000, pp. 129-138.
Pólik, I., Terlaky, T., 2007. A survey of the S-lemma. SIAM Rev. 49 (3), 371–418.
PTV, T.M.L., 2008. VAP 2.16. Interface Use Manual. Karisruhe, Alemania.
Roess, R.P., Prassas, E.S., McShane, W.R., 2004. Traffic engineering. Pearson/Prentice Hall.
Rockafellar, R.T., 2015. Convex analysis. Princeton University Press.
Rakha, H., Kamalanathsharma, R.K., 2011. Eco-driving at signalized intersections using V2I communication. In Intelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference, pp. 341-346.</bib'>
Society of Automotive Engineers (SAE), 2014. Automated driving: levels of driving automation are defined in new SAE international standard J3016. SAE International.
Scokaert, P.O., Mayne, D.Q., 1998. Min-max feedback model predictive control for constrained linear systems. IEEE Transactions on Automatic Control 43 (8), 1136–1142.
Scarf, H., 1958. A min-max solution of an inventory problem. In: Arrow, K.J., Karlin, S., Scarf, H. (Eds.), Studies in the Mathematical Theory of Inventory and Production. Stanford University Press, Stanford, CA, pp. 201–209.
Seif, H.G., Hu, X., 2016. Autonomous driving in the iCity—HD maps as a key challenge of the automotive industry. Engineering 2 (2), 159–162.
Sturm, J.F., 2002. Implementation of interior point methods for mixed semidefinite and second order cone optimization problems. Optim. Method Softw. 17 (6), 1105–1154.
Schwarm, A.T., Nikolaou, M., 1999. Chance-constrained model predictive control. AIChE J. 45 (8), 1743–1752.
Shapiro, A., 2001. On duality theory of conic linear problems. In: Semi-infinite programming. Springer, Boston, MA, pp. 135–165.
Shapiro, A., Nemirovski, A., 2005. On complexity of stochastic programming problems. In: Continuous optimization. Springer, Boston, MA., pp. 111–146.
Sidiropoulos, N.D., Davidson, T.N., Luo, Z.Q., 2006. Transmit beamforming for physical-layer multicasting. IEEE Trans. Signal Processing 54 (6-1), 2239–2251.
Skajaa, A., Andersen, E.D., Ye, Y., 2013. Warmstarting the homogeneous and self-dual interior point method for linear and conic quadratic problems. Math. Program. Comput. 5 (1), 1–25.
Sun, D., Toh, K.C., Yuan, Y., Zhao, X.Y., 2017. SDPNAL+: A Matlab software for semidefinite programming with bound constraints (version 1.0) arXiv preprint arXiv:1710.10604.
Sun, J., Zheng, Z., Sun, J., 2018. Stability analysis methods and their applicability to car-following models in conventional and connected environments. Transp. Res. Part B Method. 109, 212–237.
Sun, C., Guanetti, J., Borrelli, F., Moura, S., 2020. Optimal Eco-Driving Control of Connected and Autonomous Vehicles Through Signalized Intersections. IEEE Internet of Things J..
Suh, J., Yi, K., Jung, J., Lee, K., Chong, H., Ko, B., 2016. Design and evaluation of a model predictive vehicle control algorithm for automated driving using a vehicle traffic simulator. Control Engineering Practice 51, 92–107.
Treiber, M., Hennecke, A., Helbing, D., 2000. Congested traffic states in empirical observations and microscopic simulations. Phys. Rev. E 62 (2), 1805.
Treiber, M., Kesting, A., Helbing, D., 2007. Influence of reaction times and anticipation on stability of vehicular traffic flow. Transp. Res. Rec. 1999 (1), 23–29.
Townsend, R.M., 1983. Forecasting the forecasts of others. J. Pol. Econ. 91 (4), 546–588.
Talebpour, A., Mahmassani, H.S., 2016. Influence of connected and autonomous vehicles on traffic flow stability and throughput. Transport. Res. Part C: Emerg. Technol. 71, 143–163.
Wang, C., Ong, C.J., Sim, M., 2009. Convergence properties of constrained linear system under MPC control law using affine disturbance feedback. Automatica 45 (7), 1715–1720.
Wang, J.Q., Li, S.E., Zheng, Y., Lu, X.Y., 2015. Longitudinal collision mitigation via coordinated braking of multiple vehicles using model predictive control. Integrat. Comp.-Aid. Eng. 22 (2), 171–185.
Wang, Z., Wu, G., Hao, P., Barth, M., 2017. Cluster-wise cooperative eco-approach and departure application along signalized arterials. Proc. IEEE 20th Int. Conf. Intell. Transp. Syst..
Widodo, S., Hasegawa, T.,Tsugawa, S., 2000. Vehicle fuel consumption and emission estimation in environment-adaptive driving with or without inter-vehicle communications. In Intelligent Vehicles Symposium, 2000. IV 2000. Proceedings of the IEEE, pp. 382-386.
Wiesemann, W., Kuhn, D., Sim, M., 2014. Distributionally robust convex optimization. Oper. Res. 62 (6), 1358–1376.
Weeks, A.J., 2009. Modeling the Emissions of Heavy-Duty Diesel Vehicles on Interstate 89/189 and US Route 7 in the Burlington Area (No. UVM TRC Report# 09-006).
Yakubovic, V.A., 1971. In: S-procedure in nonlinear control theory, 1. Vestnik Leningrad University, pp. 62–77.
Yang, H., Almutairi, F., Rakha, H.A., 2020. Eco-Driving at Signalized Intersections: A Multiple Signal Optimization Approach. IEEE Intell. Transp. Syst. Trans. Mag. arXiv preprint arXiv:2001.01117.
Ye, Y., Zhang, S., 2003. New results on quadratic minimization. SIAM J. Optim. 14 (1), 245–267.
Zymler, S., Kuhn, D., Rustem, B., 2013. Distributionally robust joint chance constraints with second-order moment information. Math. Program. 137 (1–2), 167–198.
Zeilinger, M.N., Raimondo, D.M., Domahidi, A., Morari, M., Jones, C.N., 2014. On real-time robust model predictive control. Automatica 50 (3), 683–694.
Zhang, S., 2000. Quadratic maximization and semidefinite relaxation. Math. Program. 87 (3), 453–465.
Zhou, F., Li, X., Ma, J., 2017. Parsimonious shooting heuristic for trajectory design of connected automated traffic part I: theoretical analysis with generalized time geography. Transport. Res. Part B: Methodol. 95, 394–420.

Zhao, S., Zhang, K., 2017. A data-driven optimization model to observe individual dynamic choices of activity-travel-path using connected vehicles as mobile sensors. The 96th Annual Meeting of Transportation Research Board 4702, 2017.
Zhao, S., Zhang, K., 2018. A data-driven Model Predictive Control framework for robust Cooperative Adaptive Cruise Control using mobile sensing data. The 97th Annual Meeting of Transportation Research Board 2018.
Zhao, S., Zhang, K., 2019. A distributionally robust optimization approach to reconstructing missing locations and paths using high-frequency trajectory data. Transp. Res. Part C: Emerg. Technol. 102, 316–335.
Zhao, S., Zhang, K., 2020. A distributionally robust stochastic optimization-based model predictive control with distributionally robust chance constraints for cooperative adaptive cruise control under uncertain traffic conditions. Transp. Res. Part B Method. 138, 144–178.
U.S. Department of Transportation Intelligent Transportation Systems Joint Program Office, 2018. Safety Pilot Model Deployment Data (SPMD) dataset. https://data.transportation.gov/Automobiles/Safety-Pilot-Model-Deployment-Data/a7qq-9vfe.