ABSTRACT

TAJALLI, MEHRDAD. Dynamic Speed Control in Urban Traffic Networks with Connected and Automated Vehicles. (Under the direction of Dr. Ali Hajbabaie).

Dynamic speed control is a traffic congestion management technique that helps improve the mobility and safety of vehicles on urban roads. Speed control algorithms provide a great potential to control traffic congestion on freeways and arterial streets by changing vehicles' speed to regulate the traffic flow. In particular, controlling the trajectory of connected and automated vehicles (CAVs) on signalized arterial streets would regulate the speed and number of vehicles arriving at signalized intersections during red phases, and therefore, reduce the queue length and number of stops at intersections. However, the high computational complexity of dynamic speed optimization and control algorithms prevents studying the effect of controlling the trajectory of CAVs on large-sized urban street networks. The main objective of this research is to develop efficient and effective speed control algorithms to improve mobility in large-sized and complex transportation networks under highly congested traffic conditions. Hence, we propose developing and testing methodological speed optimization and control algorithms from both macroscopic and microscopic perspectives to improve the mobility of vehicular movements in a mixed traffic environment with presence of both CAVs and human-driven vehicles (HVs).

From macroscopic perspective, this study proposed a dynamic speed harmonization strategy to minimize the temporal and spatial variation of speeds between CAVs in urban street networks. An analytical nonlinear mathematical program is developed to formulate the speed harmonization problem. In addition, the complexity of problem is reduced by developing Distributed Optimization and Coordination Algorithms (DOCA) to provide scalable and real-time solutions for transportation networks of various size. To further improve the mobility of

transportation network, we proposed a methodology for coordinated speed optimization and traffic light control in urban street networks and showed its advantages over optimizing only speed or signal timing plans, separately.

From microscopic perspective, we developed a cooperative signal timing and trajectory optimization at signalized intersections with a mix of CAVs and HVs. A Lagrangian relaxation technique was developed to decompose the optimization problem into several lane-level optimization sub-problems. This setting will allow finding near-optimal solutions with small duality gap for complex intersections with different movements, lane groups, and high demand levels. In addition to the longitudinal trajectory control of CAVs, we controlled their lateral maneuvers to achieve the highest benefit from CAVs in terms of improving the system mobility. A framework is developed for optimal control of CAVs in a fully autonomous environment to improve the mobility in freeway segments with a lane drop. A vehicle-level mixed-integer nonlinear program is proposed to control the longitudinal and lateral maneuvers of CAVs, provide a smooth flow of traffic, and avoid congestion in freeway bottlenecks. To ensure the feasibility of vehicle-level decisions and push them in the direction of system-level optimality, a cooperative distributed algorithm was established so that CAVs coordinate their decisions to find the optimal longitudinal and lateral maneuvers that avoid collisions among all vehicles. The results showed that optimal lane changing of CAVs smoothens the traffic flow and increases freeway capacity in congested traffic conditions.

Dynamic Speed Control in Urban Traffic Networks with Connected and Automated Vehicles

by
Mehrdad Tajalli

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Civil Engineering

Raleigh, North Carolina
2021

APPROVED BY:

_____	_____
Dr. Ali Hajbabaie Committee Chair	Dr. Nagui M. Rouphail
_____	_____
Dr. George List	Dr. Leila Hajibabai
_____	
Dr. John Baugh	

BIOGRAPHY

Mehrdad Tajalli received his B.Sc. degree in civil engineering from the Iran University of Science and Technology, Tehran, Iran, in 2011, and the M.Sc. degree in civil engineering from the Sharif University of Technology, Tehran, in 2014. He pursued his Ph.D. studies in the Department of Civil, Construction, and Environmental Engineering, North Carolina State University. Mehrdad's research interests include Operations Research analysis and their applications on improving the mobility of transportation systems.

ACKNOWLEDGMENTS

First and foremost, I am grateful to my supervisors, Dr. Ali Hajbabaie for his invaluable advice, support, and patience during my Ph.D. study. I would also like to thank my committee members Dr. Rouphail, Dr. List, Dr. Baugh, and Dr. Hajibabai for their invaluable technical support on my research and study. I would like to thank my friends and colleagues SMA Bin Al Islam, Amir Mirheli, Rasool Mohebifard, Mehrzad Mehrabipour, Ramin Niroumand, and Sattar Sattarov for their kind help and support that have made my study and life in the U.S. a wonderful time. Finally, I would like to express my gratitude to my parents and my sister. Without their tremendous understanding and encouragement in the past few years, it would be impossible for me to complete my Ph.D. study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1. Significance of the research

Congestion management through urban traffic control strategies is one the areas that could be improved significantly using new technologies, especially during the era of connectivity and automation. Finding the proper signal timing, metering vehicles on freeways ramps or at the entrance of the network, and providing vehicles with optimal speeds are some well-known dynamic strategies to reduce urban traffic congestion. The speed distribution of vehicles in urban networks can have significant impacts on travel time and mobility of vehicles. Dynamic speed harmonization (DSH) strategy has been widely used to manage traffic flows through highways to prevent oversaturated conditions in case of an incident or the presence of a bottleneck. Speed control problems can be studied from both macroscopic and microscopic perspectives. In macroscopic perspective, the movement of traffic is modeled based on macroscopic traffic flow fundamental diagram and speed control decisions are made for all vehicles is a segment of the road. On the other hand, in microscopic perspective, the movement each individual vehicle is taken into account and speed and trajectory of individual vehicles are controlled, separately. Lyles et al. (2004) showed that DSH is an effective means to reduce travel time and maximize the throughput in work zone areas. Khondaker and Kattan (2015) considered variable speed limit on a freeway facility in a connected environment to estimate the optimal speeds for different segments to minimize travel time, improve safety, and reduce emissions. Furthermore, Grumert et al. (2015) considered DSH on a freeway in a connected environment where speed limits were assigned to individual vehicles based on their distance to the incident location and the desired speed at that location. The results showed a more harmonized traffic flow and reduced emissions compared to a fixed speed limit.

Besides the wide usage of speed limiting strategies for controlling traffic flows on freeways, speed control can be a useful tool for congestion management in urban settings. Vehicle trajectories can be used to assign optimal speeds to each vehicle approaching an intersection to prevent them from stopping in a queue (He et al., 2015; Tajalli and Hajbabaie, 2018a). Zhu and Ukkusuri (2014) addressed dynamic speed limit control problem in a connected vehicle environment and considered a trade-off between the travel time and emissions to find the optimal speed at each link of the transportation network. Based on this study, the total travel time and emissions of transportation network could be reduced by 18% and 20%, respectively compared with no speed limit control cases.

Connected and automated vehicles provide a huge amount of real-time data that could be utilized to make online decisions and improve the mobility of the system. However, traditional network-wide optimization algorithms are insensitive to the dynamic changes in the network and as a result cannot integrate the online data into their static models to provide the highest operational benefits to the transportation systems during the era of connectivity and automation. In addition, the existing control algorithms require central controllers to observe the entire system and decide about the desired control actions to alleviate the congestions. Hence, they would not be applicable for the efficient control of large urban transportation networks due to the presence of huge amount of decision variables and constraints. As a result, centralized algorithms are unable to find optimal solutions in real time. On the other hand, decentralized or distributed solution techniques could be helpful strategies by utilizing several agents and providing an effective coordination between them to share the information online and achieve as high-quality solutions as centralized algorithms. Distributed algorithms let each agent take care of a segment of the network and find the optimal

decisions for the corresponding subproblem. As a result, the complexity of the large-scale problem is significantly reduced and solutions could be achieved in real time.

In this dissertation, we aim to use mathematical programming techniques to find the optimal speeds for connected and automated vehicles in large-sized transportation networks under high traffic volume. However, the presence of a huge number of decision variables and the inherent complexity of nonlinear control algorithms makes it difficult to solve the speed control problem in large or complex networks efficiently in real-time. Hence, a distributed optimization and coordination algorithm (DOCA) is developed to decompose the problem and coordinate solutions of sub-problems in the system to provide near optimal solutions in a reasonable amount of time.

1.2. Problem statement

Controlling the trajectories of connected and automated vehicles in transportation networks not only yields a more efficient utilization of network capacity, but also regulates the movement of vehicles to achieve a "smoother" flow of traffic. The effectiveness of trajectory optimization and speed control of CAVs has been shown in several studies. However, tested scenarios were limited to simple cases due to high complexity of the problem. Therefore, it is essential to develop efficient and accurate optimization and prediction models to demonstrate the effectiveness of controlling the trajectory of CAVs in complex network settings with high traffic volumes. In addition, it is essential to develop appropriate methodologies to implement the dynamic speed optimization problem for various size of transportation networks and achieve near optimal solutions in real-time.

Distributed optimization algorithms are suitable approaches to reduce the complexity of the proposed speed optimization problem and obtain solutions in real-time . However, an effective coordination scheme would be necessary to achieve near optimal solutions that improve the

performance of the entire system. Therefore, it is very important to decompose the network into the most appropriate number of subnetworks and find the most important factors to be communicated between them to achieve the desired performance of the network. In this dissertation, we will show how developing a suitable distributed optimization and coordination algorithm for dynamic speed optimization problem helps improves the operational performance for large-sized transportation networks to the benefit of the system.

Moreover, the impacts of optimal controlling of CAVs need to be considered on the mobility of human-driven vehicles. Therefore, an analytical model is required that considers the interactions between CAVs and HVs in a mixed traffic condition. In addition, developing a lane-changing model for CAVs shows how they could benefit the mobility and safety performance of the system by avoiding the queue buildup in urban and freeway bottlenecks. In this dissertation, we develop novel problem formulations and solution techniques to show the effectiveness of optimal motion planning of CAVs for both urban networks and freeway facilities.

1.3. Research objectives and scope

The primary objective of this dissertation is to investigate the potential impact of controlling the speed and trajectories of connected and automated vehicles (CAVs) to improve the operational performance of transportation systems. In particular, this dissertation proposes the use of optimal speed control methods to develop efficient optimization framework for effective and fluent movement of CAVs in urban street networks and freeway facilities. A distributed optimization and coordination algorithm is also introduced to provide the capability of analyzing the effects of CAVs on large-sized transportation networks without sacrificing on solution run times. This dissertation is aimed at finding answers to the five following fundamental research questions:

1. What is the effect of speed control on the operational performance of connected urban street networks from a macroscopic perspective?
2. How to solve speed optimization problem for a large-sized transportation network in real time?
3. How joint optimization of signal timing parameters and speed of CAVs further improves the performance of large-scale transportation networks?
4. What is the effect of joint signal timing and trajectory optimization on improving the mobility of traffic stream in an isolated intersection under mixed traffic environment with the presence of human driven vehicles (HVs) and various penetration rate of CAVs?
5. How optimal motion and lane-changing maneuver of CAVs can smooth the traffic flow and avoid congestion and stops on freeway bottlenecks?

1.4. Contributions of research

An efficient optimization control algorithm and formulation is rquired to dynamically control the moevement of vehicles in urban street networks. However, most of the existing studies are based on static traffic models as such, are not dynamic and responsive to changes in traffic demand. In addition, despite all the efforts in respective areas, optimal network-level speed optimization problem has not been satisfactorily addressed in the literature. This dissertation develops a novel multi-objective optimization program as well as analytical solution techniques to dynamically control vehicles' speeds to facilitate smooth vehicular traffic on urban street networks.

In addition, the existing speed control techniques for urban street networks are neither scalable nor real-time. As a result, the proposed solution techniques are mostly not applicable for

real-world applications, where transportation networks have various sizes, traffic state is dynamic, and solutions are required in real-time. A distributed optimization and coordination algorithms suitable for dynamic speed control problem is developed in this study to reduce problem complexity; and find near-optimal solutions in real-time and are scalable to transportation networks of various sizes.

In addition, a problem formulation is introduced to show the effectiveness of cooperative signal timing and speed optimization in managing traffic congestion. Most of the existing studies only showed the effectiveness of cooperative signal timing and speed optimization in small networks due to the computational complexities that were associated with the cooperative problem. This study addresses the knowledge gap and enables studying the effects of cooperative signal timing and speed optimization in large scale transportation networks.

Furthermore, we considered the signal timing and trajectory optimization (STTO) problem from a microscopic perspective, to show the effect of controlling connected and automated vehicles (CAVs) on the mobility performance of the network in addition to the movement of human-driven vehicles. We propose a novel solution technique for the NP-hard problem to decompose the intersection-level optimization problem into lane-level sub-problems and easily find solutions with very low optimality gaps with a reasonable runtime.

Finally, we developed cooperative distributed control technique to utilize the maximum functionality of CAVs through optimizing their lane changing maneuvers on multilane freeway facilities with lane drops. We show that the optimal longitudinal and lateral motion planning of CAVs has a great capability to improve the system performance and avoid congestion at the freeway bottlenecks.

1.5. Dissertation organization

The exposition of this doctoral dissertation is as follows.

Chapter 2 provides an in-depth literature review on previous research about controlling the speed of vehicles on freeway facilities, urban arterial streets, isolated intersections, and transportation network. We show how speed harmonization could improve the safety, mobility, and environmental impacts in transportation systems. In addition, the existing limitations on speed control methodologies are highlighted.

Chapter 3 provides a novel problem formulation and solution technique for dynamic speed harmonization in connected urban street network. A linear reformulation is suggested for the nonlinear problem and the results of speed harmonization for various scenarios are presented.

The complexity of speed harmonization problem increases with increasing the size of the network and solutions could not be achieved in real-time for large-scale problems. Therefore, Chapter 4 provides a distributed optimization and coordination algorithm for the dynamic speed optimization of connected and autonomous vehicles on large-sized urban street networks. The efficiency and quality of the solutions from the proposed algorithm are provided in this chapter.

Chapter 5 develops a cooperative signal timing and speed optimization algorithm in a large-sized connected and automated urban street network. The effectiveness of cooperative algorithm is studied and comparisons are provided between a cooperative signal timing and speed optimization and either signal timing optimization or speed optimization in the transportation network.

Chapter 6 provides a novel a mixed integer non-linear program (MINLP) and corresponding solution technique for analyzing the signal timing and trajectory optimization of

CAVs from a microscopic perspective. The effects of optimizing the trajectory of CAVs are evaluated on human-driven vehicles with various market penetration rate of autonomous vehicles on the road.

Chapter 7 presents a cooperative distributed control framework for optimal motion of CAVs in a multilane freeway facility. A novel mathematical programming is proposed to control the longitudinal and lateral movement of CAVs and show its effects on the mobility performance of the system and reducing the traffic congestion.

Chapter 8 summarizes the accomplishments in this research and highlights the findings. In addition, this chapter highlights the gaps and enlightens the possible directions to continue the research in improving the transportation mobility with connected and automated vehicle technology.

CHAPTER 2: LITERATURE REVIEW

Connected vehicle technology has provided highly granular traffic information that can be used to enhance traffic congestion management methods. Controlling the speed of vehicles creates an opportunity to manage traffic congestion on freeways and arterial streets by changing vehicles speed to regulate the traffic congestion and avoid unnecessary accelerations, decelerations, and stops. Speed harmonization (SH) is a strategy that reduces the spatial and temporal variation of speeds between vehicles on the road to improve mobility, safety, and reduce vehicles' fuel consumption (Ma et al., 2016; Tajalli and Hajbabaie, 2018b). The emergence of connected vehicle technology allows the communication of advisory speeds with the human drivers to adjust the speeds according to the road geometry, work zone areas, traffic condition, and weather condition. Furthermore, the real-time optimal speeds could be communicated to connected and automated vehicles (CAVs) to move safely while avoiding congestion on the downstream bottlenecks. Speed harmonization strategies are mostly developed for freeway facilities through Variable Speed Limit (VSL) control methods shown to drivers via variable message signs. Moreover, the Eco-driving models are developed for providing the optimal speeds for vehicles moving on arterial streets when they are approaching a signalized intersection to avoid unnecessary stops. Furthermore, the attention has been recently turned into controlling the trajectory of automated vehicles to enhance the traffic mobility to its highest potential. Following provides an in depth review of the existing research studies on controlling the speed of vehicles on freeway facilities, arterial streets, isolated intersections, and transportation networks.

2.1. Speed harmonization in freeway facilities

The speed harmonization strategy in freeway facilities is mainly implemented through variable speed message signs to inform drivers about the suggested speeds to avoid building congestion at the freeway bottlenecks or improve the safety in freeway work zone areas and under inclement weather conditions (Robinson, 2000). Studies have shown that implementing the speed harmonization strategy in freeway facilities improves the traffic safety (Abdel-Aty et al., 2006; Lee et al., 2006; Sepulcre and Gozalvez, 2012) and helps reduce the emission of environmental pollutants (MacDonald, 2008; Z. Wang et al., 2015). Moreover, from the operational perspective, speed harmonization decreases the travel time (Khondaker and Kattan, 2015) and smoothens the flow of traffic (Ghiasi et al., 2017).

Macroscopic traffic flow fundamental diagram, which indicates the relationship between the traffic flow and density is mostly utilized to capture the homogenous traffic condition on a segment of freeway and find the proper speed limit for vehicles passing that segment (Greenshields et al., 1935; Newell, 1993). Based on the fundamental diagram, the traffic flow increases with the increase of density until reaching to the jam density. Vehicles usually move with the free flow speed up to the jam density point. However, the flow rate decreases and the average speed drops after reaching to the jam density. The main function of VSL is to avoid reaching to the jam density on freeway bottlenecks by controlling the inflow through providing advisory speeds to vehicles upstream of the bottleneck (Ma et al., 2016).

VSL strategies on freeway facilities may be reactive or proactive. Reactive approaches monitor inactive bottlenecks and activate after a certain threshold (Malikopoulos et al. 2016 and Brinckerhoff et al. 2008). The initial implementation of the reactive VSL system in the world comes back to early 1960s in Germany A8 corridor (Schick, 2003). The Washington State

Department of Transportation applied the VSL strategy in the United States for the first time (Brinckerhoff et al., 2008). They used the detectors to measure the flow and speeds to find the proper speed limit based on some pre-defined thresholds. The performance of reactive VSL is mainly a function of traffic volume, where reaching to the maximum capacity yields to lowering the speed limit (Zhicai et al., 2004). For instance, a VSL system in M25 highway in UK activates when traffic volume exceeds 1650 veh/hr/lane and reduces speed limit from 70 mph to 60 mph (Harbord, 1998). In fact, reactive approaches activate the VSL strategy after detecting a problem as such, there might be a lag in managing congestion (Khondaker and Kattan, 2015).

Proactive approaches address this issue by predicting traffic condition in the future and applying appropriate VSL strategies in advance to prevent congestion (Khondaker and Kattan, 2015). Initially, Alessandri et al. (1998) proposed the proactive VSL by predicting the future state of the traffic based on the time-series traffic flow measurement and the Kalman Filter method. However, it turned out that the time-series predictive models are not robust to the unexpected changes of traffic patterns (Malikopoulos et al., 2018). Introducing the model predictive control (MPC) technique into the variable speed limit strategy overcame the drawbacks of time-series predictive models and provided stable flows by focusing on reducing the density upstream of the bottlenecks rather than decreasing the speed variance (Hegyi et al., 2005b).

Later, many studies utilized MPC on developing proactive VSL algorithms (Camacho and Bordons, 2012; Hegyi et al., 2005a; van de Weg et al., 2015). Model predictive control approach, through utilizing a rolling time horizon, allows accounting for changes in the network over time due to the demand changes, incident happening, and lane closure. In addition, it provides feedbacks about the effect of the suggested speed limit on the performance of the system (Carlson et al., 2010; Hegyi et al., 2008; Lu et al., 2010).

The objective of VSL optimization models could vary depending on the focus of VSL and its desired effect on the performance of the system. The main performance measures defined in the objective functions include the travel time (Alessandri et al., 1999), queue length (Lin et al., 2004), throughput (Lin et al., 2004), safety risks (Fang et al., 2015), speed variance and homogeneity (Hegyi et al., 2002), environmental impacts (Lin et al., 2004), and fuel consumptions (Zegeye et al., 2010, 2009). The decision variable of the optimization problem is the optimal speed limit (Alessandri et al., 1999; Hadiuzzaman and Qiu, 2013; Hegyi et al., 2002) and the constraints include the dynamic traffic flow conservation and speed limit boundaries (Hegyi et al., 2005b).

From the safety perspective, Abdel-Aty et al. (2006), Li et al. (2015), and Lee et al. (2006) developed VSL strategies to reduce traffic speed upstream of crash prone areas and increase it downstream of unsafe areas. Ha et al. (2003) mention that VSL strategies increase the average headway while reducing the average speed and its variance. Thus, VSL strategies reduce the number of crashes (Smulders, 1990).

From the environmental perspective, speed harmonization smoothens traffic flow and reduces the number of lane changings in congested conditions as well (Borrough 1997; Ghiasi et al. 2017). Therefore, VSL reduces the emission of environmental pollutants as Aziz and Ukkusuri (2012) show a direct relationship between the CO emissions and vehicle speed. Implementing variable speed limit on the M42 freeway in the UK shows emission reductions between 4% to 10% (MacDonald, 2008). Zegeye et al. (2009) considered speed harmonization from a macroscopic level and showed that harmonizing the speed on a two-lane 12-km long freeway segment reduced fuel consumptions by 14.15%. Zegeye et al. (2010) added the speed, wind direction, and the weather temperature into the macroscopic model for a more accurate analysis and showed that the

total emission on the same freeway segment was reduced by 36% compared to an uncontrolled case.

The emergence of the connected vehicle and autonomous technologies urged a need to reconsider the early VSL algorithms based on offline or homogenious data. CV technology provides the opportunity to collect real-time data from individual vehicles with higher accuracy instead of relying on error prone aggregated detector data. The speed control algorithm based on the CV technology could provide advisory speeds to connected vehicles through information sharing (Talebpour et al., 2013; Yang and Jin, 2014) or control the trajectory of connected and automated vehicles with optimal speed and trajectory decisions (M. Wang et al., 2015a). Talebpour & Mahmassani (2016) studied the possible changes in the traffic flow by introducing these new technologies and showed that connectivity and automation can improve the stability of traffic flow, although the automation is more effective than connectivity in preventing shockwave formation and its propagation. Talebpour et al. (2017) also studied the impact of compliance rate on the effectiveness of speed harmonization techniques for connected vehicles and claimed that a minimum 10% compliance rate would be necessary to attain the desired objective.

Khondaker and Kattan (2015) considered variable speed limit on a freeway facility in a connected vehicle environment to find the optimal speeds for different segments to minimize travel time, improve safety, and reduce emissions. They showed that VSL can reduce total travel time, time to collision, and fuel consumption at most by 20%, 11%, and 16%, respectively. Grumert et al. (2015) showed that harmonizing the speed of vehicles through vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications reduces the emission factors significantly. It has been shown that harmonizing the speed on a basic freeway segment reduced HC and NOx emissions of connected vehicles by 8.0% and 3.8%, respectively.

Note that researchers have not found consistent results on the effects of VSL strategies on the mobility of freeway facilities. For example, Papageorgiou et al. (2008) showed that applying VSL strategies to under-critical occupancies increases travel time. However, VSL reduces travel time when applied to the critical occupancy conditions. In addition, Edara et al. (2016) found an increase of 4% to 8% in travel time and a decrease of 7% to 11% in throughput when variable speed limit is implemented on congested work zone areas.

2.2. Speed harmonization in arterial streets

Signalized intersections are the major bottlenecks in urban street networks. Controlling the speed of vehicles upstream of signalized intersections with an advanced knowledge of signal phasing and timing (SPaT) information improves traffic safety, mobility, and fuel consumption efficiency (Asadi and Vahidi, 2010; He et al., 2015; HomChaudhuri et al., 2015). The national highway traffic safety administration (NHTSA) reported that SPaT broadcasts are capable of reducing red-light violations and energy consumption by 90% and 35%, respectively (Misener et al., 2010). Moreover, vehicles arrival time to the intersection can be planned with the goal of avoiding stops at an intersection and minimizing fuel consumption. The eco-driving models are among the most prevailing strategies to reduce the environmental impacts of vehicles on arterial streets. Eco-driving models provide the optimal speeds to operate personal vehicles as well as public transit with energy efficiency upstream of signalized intersections (Barkenbus, 2010; Zarkadoula et al., 2007).

The objective of eco-driving systems in arterial streets is usually minimizing fuel consumptions and emission of vehicles (Jiang et al., 2017; Kamalanathsharma and Rakha, 2014; Kamalanathsharma et al., 2015; Wan et al., 2016; Xia et al., 2013a). In addition, other objective functions such as minimizing the travel time (Asadi and Vahidi, 2010; Chen et al., 2014; Mahler

and Vahidi, 2012) and minimizing the duration of idle time (Hao et al., 2015; Xia et al., 2013b; Xiang et al., 2015) are considered in previous eco-driving experiments. The reason for minimizing the idle time at the intersection was a strong correlation between higher fuel consumption and stop time at the intersection (Mintsis et al., 2020). However, the idle time minimization problem is much less computationally complex than the fuel consumption minimization (Rakha et al., 2012).

The complexity of eco-driving models highly depends on the modeling of vehicle dynamics in the optimization problems to simulate the vehicle's transition from the current speed to the desired optimal speed. The performance of fuel consumption and emission optimization models are highly impacted by selecting the right dynamic models (Ahn et al., 2002). Constant acceleration (Chen et al., 2015; He et al., 2015; Kamalanathsharma and Rakha, 2016; Mandava et al., 2009; Wang et al., 2016), linear acceleration (Chen et al., 2014; Rakha et al., 2012), and nonlinear acceleration (Barth et al., 2011; Hao et al., 2015; Weber and Winckler, 2013; Xiang et al., 2015) models are among the most commonly used dynamic models in previous studies. The more complex models are able to better capture the dynamic of the vehicular systems with respect to braking and throttling and as a result can be applied to the real-world applications more reliably (Rajamani, 2011).

The advancement of infrastructure-to-vehicle (I2V) communication technology facilitate the propogation of SPaT messages to the vehicles on arterial streets. Therefore, speed planning algorithms are developed to estimate the optimal speed profiles based on remaining green or red signal timings ahead and the corresponding distance between an equipped vehcile and the intersection stop bar (Barth et al., 2011). However, considering the unequipped vehicles, the existing queue at the intersection, and the compliance rate are among the most importatnt factors in a successful implementation of eco-driving algorithms (Mintsis et al., 2020).

The result of initial eco-driving studies without considering the existing queue at the intersection showed that this strategy is only impactful when the demand is light or medium (Barth et al., 2011; Mandava et al., 2009; Xia et al., 2013a). However, considering the queue in the problem formulation yielded better performance of eco-driving system in congested traffic condition (Hao et al., 2015; Jiang et al., 2017; Kamalanathsharma et al., 2015; Xia et al., 2013b).

To account for the queue information around the intersection, Rakha et al. (2012) developed a microscopic fuel consumption model to find the fuel-optimal advisory speed profile for connected vehicles. The objective function of the proposed optimization problem minimizes the fuel consumption for the entire maneuver of vehicles near signalized intersection. For each CAV, the time to intersection is estimated based on the available data (i.e. queued vehicle, lead-vehicle, and signal change information). Then, the optimal speed profile is calculated based on the estimated time to intersection, vehicle acceleration model, roadway characteristics and microscopic fuel consumption models. This study assumed that all vehicles are equipped and they can communicate with the signal controller. Moreover, Xia et al. (2013b) considered the delay of preceding vehicles in the speed optimization algorithm proposed by Barth et al. (2011) to capture the effect of existing queue on the optimal profile of a CAV. Chen et al. (2014) also incorporated constraints about queue discharge information in the speed optimization problem to find more accurate speed profiles for CAVs. He et al. (2015) studied advisory speeds in arterial streets by taking into account queue formation at intersections in addition to minimizing fuel consumption. Thus, the infeasible solutions were prevented and the problem could be solved quicker without converging to sub-optimal solutions. They found 29% reduction in fuel consumption while travel time was increased by 9%. Kamalanathsharma et al. (2015) considered more constraints about the surrounding environment (e.g. safe distance to the following vehicle, maximum acceleration, and deceleration

of a vehicle) to let the advisory speeds be more precise and applicable to real-world conditions. They showed that the fuel consumption was reduced by 25% as a result of advisory speed strategies in arterial streets.

Xia et al. (2012) showed that a 14% reduction in fuel consumption is possible as a result of an advisory speed system experiment at a fixed-time signalized intersection. Wei et al. (2017) showed that optimizing the trajectory of the leading vehicle in a platoon approaching a signalized intersection can effectively manage the traffic congestion and increase the capacity of the intersection. Jiang et al. (2017) proposed an eco-driving system for a simple two-way isolated signalized intersection with fixed signal parameters and showed that optimal advisory speed can reduce the fuel consumption of connected and automated vehicles by 2.02%-58.01%. Wan et al. (2016) showed that advisory speed systems not only reduce fuel consumption significantly, but also affect the behavior of non-equipped vehicles and harmonize their motion as well. Thus, speed advisory systems contribute to more energy efficiency on the roads. In contrast with most of eco-driving algorithms that work based on the pre-time signal timing information, Hao et al. (2015) proposed an eco-driving algorithm for an actuated signalized intersection by providing an upper-bound and lower-bound of the remaining actual green time to approaching vehicles. The simulation results showed significant energy savings and fuel consumption reductions by 12%. Later, the authors applied the proposed strategy in in the field and showed a 2% energy saving, which was less than the expected values from the simulation analysis (Hao et al., 2018).

Most of the studies on optimal advisory speeds assume that vehicles are highly automated and comply with the suggested speeds. However, the lower level automation only provides information to assist human driver (Sharma et al., 2017). Therefore, it is essential to take into account the human factor in complying with the suggested advisory speeds. Rakha et al. (2012)

addressed this problem by providing the ability in the proposed algorithm to recalculate the suggested speed at each time step through rolling horizon strategy. Xiang et al. (2015) also accounted the drivers' behavior in the problem formulation to consider the compliance issue. Liao et al. (2018) utilized a simulation environment to study the effect of compliance rate of CVs on the effectiveness of eco-driving system. The study showed that the higher compliance rate is associated with lower frequency of braking and idling which yields to a smoother movement of vehicles. The study showed that the minimum compliance rate of 30% is required to see the effect of eco-driving on reducing the emission of CO_2, NO_x, HC, and CO. The emission of CO_2 could decrease by 16% to 22% when the compliance rate changes between 30% and 70%.

To account the presence of human drivers around CAVs, Kamal et al. (2015) estimated the state of the un-equipped follower vehicles using Gipps car-following model (Gipps, 1981). They used SPaT data to find optimal acceleration rates of the equipped CAVs. The proposed algorithm dynamically tuned the speed of vehicles to avoid idling in red signals either by speeding up or slowing down. Simulation tests showed that the approach reduced fuel consumption by 4.5% and travel time by 2% compared to the traditional driving system. Moreover, Jiang et al. (2017) developed an eco-driving control algorithm by considering un-equipped vehicles on the road. They showed that at least 40% CAV market penetration rate will be required to see the benefit of the proposed algorithm in terms of improving mobility and reducing the fuel consumption. Kamalanathsharma et al. (2015) also controlled the movement of CAVs in a four-leg single lane intersection by a dynamic programming algorithm and showed that at least 50% market penetration rate of CAVs is required to observe system-wide fuel and delay savings.

The simulation results from Xia et al. (2013a) showed that controlling the speed of equipped vehicles has positive effect on the movement of non-equipped following vehicles in

terms fuel saving and energy reduction. However, this positive effect is more obvious in single-lane configuration when the market penetration rate of CAVs is over 40%-50% (Barth et al., 2011; Xia et al., 2013a, 2013b). On the contrary, human drivers might take over the leading slower CAV and as a result, controlling the equipped vehicles will not improve the mobility of un-equipped vehicle in lower penetration rates (Xia et al., 2013a).

Although previous studies made significant contributions to reducing the fuel consumption and emission of vehicles on arterial streets, few studies considered the safe maneuver of CAVs in real-world applications, where the behavior of human drivers around is hardly predictable. Therefore, more safety focus needs to be added to the dynamic of CAVs through data-driven methodologies. In addition to minimizing the fuel consumption, improving the mobility performance of the system (e.g. travel time) also requires more attention in future studies. Finally, the impact of eco-driving strategies are mostly observed from local perspective within few intersection corridors. Therefore, it will be helpful to study how controlling the speed of CAVs will benefit the large-size transportation networks.

2.3. Speed harmonization in urban networks

Despite all the studies of speed harmonization on freeways and arterial streets, very few published studies have considered its network-wide effect. McLean et al. (1994) showed that lowering the speed limit in urban transportation network is associated with higher safety and less number of accidents. Aside from the safety benefits, Taylor (2000) studied the mobility effects of lowering speed limit on urban network using TrafikcPlan traffic network analysis software (Young et al., 1989) for various traffic congestion conditions. Three speed limit regimes were considered as 40, 50, and 60 km/h on a network with 9 major intersections and 16 minor intersections. The study showed that the difference in overall travel speed for different speed limit are relatively less

than the difference of speed limits. It is shown that the difference between the travel speeds of different speed limit plans is less when the traffic congestion increases.

To capture the network-wide effect of speed harmonization, macroscopic traffic flow models are utilized to represent the flow of traffic. From the static network optimization perspective, Yang et al. (2012) investigated the effects of speed limits on traffic equilibrium. They modified a non-differentiable travel time function to guarantee the uniqueness of flow and travel time in a link at equilibrium condition. As a result of this study, a proper link-based speed limit is found to reduce both the network vehicular emission and total travel time. Wang (2013) has investigated the network-level impact of speed limit on vehicular movements in a macroscopic-level study. They have introduced a bi-level speed limit design problem, in which the upper level determines the optimal speed limits for pre-defined links while the lower level finds static user equilibrium (UE) link flows. This solution approach is static, which limits the application of approach to dynamic and real-world operational conditions. Yang et al. (2013) have proposed a multi-objective bi-level formulation to determine the optimal speed limits in a transportation network. They have formulated a UE traffic assignment in the upper level and introduced a system-wide travel time as a predictive function of speed limit in the lower level along with other objectives on estimated number of accidents and transportation emissions. The problem is solved statically using genetic algorithms, for a study period in a scenario-based fashion and consequently is not applicable to dynamic operational conditions.

Furthermore, Yan et al. (2015) have proposed a bi-level mathematical program to minimize the total travel time through determination of the optimal speed limits in the upper level and solving a UE traffic assignment in the network under the predefined speed limits on some links in the lower level using a stochastic travel time function. The static nature of the approach limits its

application to dynamic operational conditions, just like other previous studies. A study by Zhu and Ukkusuri (2014) has addressed dynamic speed limit control problem in a connected vehicle environment. This study aimed to minimize travel time and environmental emissions and find speed at each network link over the study period using Markov Decision processes. This approach does not guarantee the optimality of the solutions as it is a reinforcement learning process, where the optimal actions may not be found.

The cell transmission model (CTM) developed by Daganzo (1994) is one of the widely used dynamic flow models where queue spill backs and shockwave propagations are captured. CTM is a discrete approximation of the Lighthill, Whitham, and Richards hydrodynamic model (Lighthill & Whitham 1955, Richards 1956), where the relationship between flow and density is assumed to be in a shape of trapezoid (Carlos F. Daganzo, 1994; Ziliaskopoulos, 2000). In the cell transmission model, the time interval is discretized into small values. Consequently, every link in the network is divided into homogenous segments, which are called cells (Li et al., 1999). The length of cells should be at least equal to the distance a vehicle moves with maximum free flow speed in one time step.

In initial efforts for using CTM for speed harmonization, Zhu and Ukkusuri (2016) proposed a cell-based simulation approach to model the proactive driving behavior of connected vehicles in a mixed traffic environment of the transportation network. This study adjusted the exit flow of cells containing the connected vehicles to let them arrive to the intersection during the green traffic light. Although no optimization algorithm is used to control the movement of connected vehicles, testing the result of the simulation methodology on the network of Manhattan downtown in New York City showed that the presence of connected vehicle contributes to the increase of average speed, reduction of emission, and smoothness of traffic flow over the network.

The study showed that 30% and 50% market penetration rate of CVs would be enough to see the significant improvement under heavy and light traffic condition, respectively. Othman et al. (2020) also used CTM to model the flow of traffic in an urban transportation network and find the optimal speed limit at each segment of the road. Moreover, the model predictive control was utilized to handle the dynamic of the problem and reduce the problem complexity. The objective function of this study minimized the total fuel consumption and maximized the distance traveled by all the vehicles in the network. Implementing the proposed method on a network of 16 intersections reduced energy consumption and travel time compared to the cases that speed limit was fixed. The results also showed that the variable speed limit on the network is more effective when the network is under congested traffic condition.

The accuracy of CTM depends on the length of segments. The shorter lengths of segments and shorter time intervals provides more accurate representation of traffic flow in the network. However, it increases the complexity of the models by increasing the number of decision variables. To overcome this drawbacks of CTM, the Variable Length Model (VLM) proposed by De Wit (2011) is introduced to capture the traffic conditions on a long segment by describing the dynamics of the shock wave and its upstream and downstream traffic states. Similar to CTM, VLM is a discrete version of LWR with triangular traffic fundamental diagram. VLM represent the link with only two cells and three variables: the density in the upstream free cell, the density in the downstream congested cell, and the position of the congestion front. De Nunzio et al. (2014) adapted the Variable Length Model to the urban framework considering the traffic lights. They analyzed the macroscopic traffic performance metrics to assess the behavior of the system on the optimal steady-state speed over the links of the network. It is shown that the performance of the system under the optimal speed limit improves in terms of energy consumption and travel time

distance by 29% and 27%, respectively compared to the benchmark, which assumes fixed speed limit for all links. In a later study, Nunzio and Gutman (2017) reduced the computational complexity of previous work by relaxing the differential equations and solving the problem for larger networks and more congested traffic conditions. This study showed the higher effectiveness of variable speed limiting when the network congestion level is higher.

Beside speed control studies based on first order traffic flow models, Papageorgiou et al. (1990) developed a second order macroscopic modeling tool, METANET, that provides more accurate representation of traffic flow despite its higher computational complexity due to the nonlinearity. METANET divides a link into homogenous segments and models the traffic flow based on the average behavior of each segment. Hegyi (2004), Hegyi et al. (2005a), and Hegyi et al. (2005b) utilized MPC to modify and extend the application of METANET for the speed control purposes on freeway facilities. Based on the modified METANET, the new desired speed is find from the minimum of the target speed under the current traffic condition and the actual speed limit of the road. The application of VSL based on METANET on freeway prevented traffic breakdown and maintained a higher outflow. den Berg et al. (2003) extended METANET for the urban network to take into account the effect of traffic light on the blocked traffic flow and its corresponding shockwave. However, this work did not show the effect of speed limiting on the network.

The macroscopic traffic flow models provide general insight about the effect of control strategies from the network-wide perspective. However, the micro-simulation provides more detailed behavior of drivers and more details about the vehicles movements. Since the application of microscopic models is computationally challenging when the network is large, especially when

there is a need to solve an optimization problem, a trade-off between macroscopic and microscopic simulations may lead to better analysis of the network from both perspectives.

Waller et al. (2009) proposed a framework to evaluate the effect of speed harmonization and shoulder use on the urban freeways as strategies to manage traffic dynamically. In fact, dynamic traffic management strategies such as variable speed limit could be employed locally, but has some network-wide consequences. In this study for simulating the traffic flow on mesoscopic level, Vista is utilized which is based on a cell transmission model. Vissim is also used to perform the micro-simulation. Based on the proposed framework, the variable speed limits could be found based on online and offline control strategies. The offline control algorithms depend on the historical demand patterns, while for online purposes real time traffic data will be considered. As a result of this study, VSL and shoulder strategies did not have significant impacts on the freeway throughputs. However, the traffic stream gets homogenized and smoother due to reducing the total number of stops, stop delay, and number of lane changings. Moreover, the average speeds and variation of the speeds decreased. Therefore, total travel time on the network increased slightly.

Abdelghaffar et al. (2020) developed a robust control technique to consider the dynamic of the network in real-time and address the system uncertainty. The sliding mode control was proposed to dynamically identify the bottlenecks in the network and regulate the average space means speed in a link until the congestion is dispersed. The sliding mode control does not have a closed form mathematical formulation and works based on the feedback received from the system. However, the feedbacks utilized in this study are local, which does not ensure the global optimal solution for the entire network. The proposed methodology was applied to the calibrated network of downtown Los Angles consisting of multiple connected freeway and signalized arterials. The

simulation results showed a 12.17% reduction in travel time and a 20.67% reduction in total delay. In addition, the fuel consumption and CO2 emission were reduced by 2.6% and 3.3% respectively.

Panis et al. (2006) also considered the effect of speed limit controlling on the network from microscopic perspective. They used the DRACULA (Dynamic Route Assignment Combining User Learning and microsimulation) micro-simulator to control and study the behavior of individual vehicles after suggesting new speed limit in different links of the network. The Intelligent Speed Adaption (ISA) control as an in-vehicle electronic system was used in different penetration rate of vehicles to regulate the speed of vehicles by capping the maximum speed or informing the driver about the speed limit of the road. The application results of this method on a network of 2 by 2 km in downtown Ghent, Belgium showed that the variation of speed is reduced significantly by increasing the penetration rate of equipped vehicles. However, it did not show significant impact on reducing the emission of vehicles. The reason for that could be the lack of an emission reduction objective function for ISA system, which only help drivers not violating the speed limit. These result are in line with the findings of Liu and Tate (2004), where it was shown that ISA is more effective in less congested traffic conditions of on a street network in the east of Leeds, England, with 120 intersections.

The studies focused on network-wide effects of speed optimization mainly used static traffic flow models to find the optimal speed limits on each link of the network. However, they are not responsive to the change of traffic congestion over time during peak or off-peak hours. The studies that considered the dynamic changes of traffic condition were not able to find near optimal solutions due to the high complexity of the speed optimization problem. In addition, it takes a huge amount of time to solve the problem for large-scale transportation network. As a result, the existing studies are limited to small size case studies. As a result, a real-time algorithm will be required to

handle the computational complexity of the speed optimization problem for large-scale networks. Moreover, controlling the speed of vehicles could be coordinated with signal timing plans at the network intersections to further improve the system mobility.

2.4. Coordinated signal and speed optimization

In addition to controlling the speed of vehicles in the network, the joint signal timing and speed optimization could provide further improvement in the mobility performance of the system. When all the vehicles are connected and automated, controlling the trajectory of CAVs in a signal-free intersection provides the opportunity to achieve the highest capacity while maintaining safety by preventing collisions between vehicles (Dresner and Stone, 2004; Mirheli et al., 2018; Niroumand et al., 2020b, 2020a; B. Xu et al., 2018b; Yu et al., 2019; Zohdy and Rakha, 2014). For instance, Mirheli et al. (2019) showed that the total travel time at a signal-free intersection with 100% CAVs could be reduced up to by 70.5% in comparison with optimized fully actuated signal timing plans. However, the signal free intersection can only operate in a fully autonomous enviorment or at least 90% CAV penetration rate is required to see the mobility improvements compared to the joint signal timing and trajectory optimization cases (Pourmehrab et al., 2020b).

It has also been shown that optimizing the speed and trajectory of vehicles in conjunction with signal control yields a more efficient utilization of intersection capacity. Erdmann (2013) considered the speed and location of vehicles in the vicinity of an isolated intersection and solved the signal timing and trajectory optimization problem in two levels with a dynamic programming procedure. Simulating the proposed algorithm based on Green-Light-Optimal-Speed-Advisory (GLOSA) assistance system (Katsaros et al., 2011) in SUMO (Behrisch et al., 2011) showed that the delay was decreased between 33% and 72% in comparison with fixed time and adaptive signal timings, respectively.

Li et al. (2014) jointly found signal timing and trajectories of CAVs on an intersection of one-way streets, assuming that all vehicles are fully connected and will follow the assigned optimal trajectories. They enumerated all possible signal timing plans and optimized trajectories for each CAV, with the objective of minimizing the average delay. Their proposed approach reduced the average delay by up to 36.9% and increased the throughput by up to 20.2% compared to a fully-actuated signal controller. This approach was not applicable to more complex intersections due to the need to enumerate all possible signal timing plans. Jung et al. (2016) developed a bi-level program to find traffic signal timing parameters and CAV trajectories at a simple four-leg intersection with only one through link at each direction and only CAVs in traffic stream. The total intersection delay was minimized at the upper level by an exhaustive search method based on the estimated arrival time of vehicles to the intersection. Then, the trajectories of CAVs were optimized at the lower level using a Genetic Algorithm to minimize the total fuel consumption. As a result of this study, the travel time and fuel consumption reduced by up to 12% and 10%, respectively.

Yang et al. (2016) found signal timing parameters at an isolated intersection with two one-way streets by solving a bi-level program and enumeration by a tree search algorithm. This study considered three categories of vehicles, including conventional, connected human-driven, and connected self-driving vehicles with different market penetration rates. In the upper-level, the total intersection delay was minimized based on the estimated arrival time of vehicles. The trajectories of CAVs were optimized in the lower-level problem for each signal timing plan to maximize their entry speed to the intersection. This study relied on the assumption that platoon leaders always arrive at the intersection at the unset of green time interval. The algorithm might not be efficient for solving a complex intersections due to its enumerative nature and limiting assumptions. Xu et

al. (2018) developed a bi-level program to minimize the total travel time in the upper level and minimize the fuel consumption of each vehicle at the lower level. The signal timing plans were found by enumeration. Then, the optimal trajectory of vehicles for each plan were found by approximating the state and control variables using an interpolating polynomial function, named pseudospectral control method. Pourmehrab et al. (2019) provided an integrated trajectory optimization and signal control algorithm. In this study trajectory of leading automated vehicles are optimized to minimize the travel time delay. This study mainly focused on the feasibility of the signal timing plans rather than their optimality by adjusting the signal timing parameters to find feasible coordination with optimized trajectory. The signal timing parameters are only controlled to be extended or switched to another phase repeatedly until finding a feasible match between arrival time of vehicles and status of signal timing.

The enumerative methods such as the ones described in previous studies become ineffective when the intersection layout is complex, the number of phases increases, or the demand level is high. Therefore, some studies proposed heuristics or considered simplified assumptions to solve the signal timing and trajectory optimization problem efficiently. For instance, Qi et al. (2020) used particle swarm metaheuristic algorithm and solved the joint signal timing and trajectory optimization problem in under-saturated traffic condition in a mixed environment. Yin et al. (2021) solved a bi-level signal timing and trajectory optimization problem using an ant colony system algorithm. This study also considered the presence of pedestrians at the intersection and minimize the delay of pedestrian in addition of vehicle. Although the metaheuristic algorithms proposed in these studies are helpful in solving nonlinear optimization problem, they might not find the optimal solution efficiently or converge to the local optimal solution, especially for higher traffic volume and more complex intersection layout.

Feng et al. (2018) developed a joint optimization method for signal timing and trajectory of platooning CAVs. They divided the signal phases into several stages and used a dynamic programming method to find the signal timing parameters at each stage. The proposed approach could find optimal solutions analytically with the assumption that the platoon leader arrives at the intersection stop bar exactly at the beginning of the green time interval. However, this assumption may limit the choices of signal timing plans in the dynamic programming framework, especially when platoon leaders are not CAVs. Moreover, the effectiveness of the proposed method for high traffic volume was not clear since the study showed the performance results for simple scenarios, where the traffic volume was very low. The result of this study showed that the joint optimization of signal timing and trajectory of platoon leaders reduced vehicle delay and CO_2 emission by up to 24% and 13.8%, respectively.

Guo et al. (2019) proposed a two-step approach to optimize signal timing and trajectory of CAVs. In the first step, the signal timing parameters were found to minimize the intersection delay. In the second step, the optimal trajectories were designed for the planned signal timings. The signal timing parameters were found by dynamic programming, and the effect of each signal plan on trajectories was evaluated through a shooting heuristic, iteratively. Shooting heuristic was shown to estimate high-quality trajectories for CAVs. However, the feasibility of this approach is limited to the cases that intersection links are long enough to let a vehicle select an appropriate maneuver. In addition, the existence of long queue at the intersection yields finding infeasible solutions by shooting heuristic. The result of this study showed that joint signal timing and trajectory optimization could reduce the travel time and fuel consumption up to by 35.7% and 31.5%, respectively compared to adaptive signal timing control. In addition, they showed that traffic operations improved with higher CAV market penetration rate.

Yu et al. (2018) optimized signal timing parameters such as phase sequence, green start time, duration of each phase, and the cycle length at an isolated intersection. Besides, they optimized arrival time of vehicles and lane changing behaviors in a four-leg intersection with through, right-turn, and left-turn movements. This study utilized a planning horizon procedure to solve an MILP problem. However, it assumed that (1) all vehicles were CAVs and (2) the vehicles that were in the same lane passing through the intersection in the same cycle were in the same platoon. As a result, they only optimized the trajectory (arrival time) of the leading vehicle, and other vehicles followed the leader with a car-following model. In addition, they assumed that all vehicles arrived the intersection with their desired speed and no queue existed at the intersection. These assumptions restricted the application of the proposed methodology to low CAV market penetration rate levels.

Li et al. (2018) optimized trajectory of electric vehicles in combination with signal timing plans for an arterial street with multiple intersections. The objective of this study was to minimize the traffic delay by optimizing signal timings and save the energy of electric vehicles by optimizing their trajectory. To solve this complex optimization problem, a hybrid heuristic technique comprise of genetic algorithm and particle swarm optimization is utilized to generate feasible signal timing plans and consequently evaluate optimal trajectories for that signal. However, the trajectories provided for CAVs ensure vehicle-level optimality rather than system-level optimal performance.

All previous studies are based on mixed integer linear programming that is computationally complex for the problem of signal timing and trajectory optimization. Several recent studies tried to introduce novel technique to reduce the problem complexity and find solutions in real time. For instance, Rodriguez et al. (2020) proposed a multi agent representation of the network, where the traffic light and CAVs are connected to each other within a graph-represented network of the

intersection. In addition, vehicles on the same lane are connected to each other in this graph. Then, a decentralized gradient-based multi-agent control algorithm is developed to find the optimal arrival time for CAVs and the switching time for the signal controller. This method enables scale the problem for larger size networks and higher traffic demands. However, the results in this study are provided for a maximum demand of 500 veh/hr/lane. The result showed that the joint optimization of traffic light and CAV trajectory could save the delay up to 50% in tested scenarios.

Pourmehrab et al. (2020) also attempted to provide a real-time optimization algorithm for the signal timing and trajectory optimization in a mixed traffic flow. To reduce the computational complexity of the problem, the authors divided the signal timing and trajectory optimization problem into two levels by finding the order and duration of each signal timing phases first. Then, the trajectories are optimized based on the found signals. To further reduce the computational complexity of the signal timing optimization, the problem is reformulated as a minimum cost flow problem, which is linear and can be solved centrally. However, the process proposed in this study is not iterative and the effect of lower level problem is not seen on the upper level problem. Therefore, the overall solution might not be global optimal. The result of this study on the maximum demand of 850 veh/hr/lane showed that the proposed signal timing and trajectory optimization algorithm could reduce the travel time by 18%-22%. Similar to this study, Long et al. (2020) also divided the joint problem into upper level signal timing and lower level trajectory optimization problems. However, they considered the iterative process to achieve better performance in the system. The results showed that the average delay and fuel consumption reduced up to by 24.9% and 5.5% respectively compared to the actuated signal timing plan.

In another effort to reduce the computational complexity of signal timing and trajectory optimization problem, Guoa and Ma (2020) improved their previous work by replacing the

dynamic programming strategy with a deep reinforcement learning to find the solutions in real time. In addition, they forced vehicles to build a platoon and just let the platoon leader to coordinate with signal controller and optimize the trajectory. The results based on the maximum demand of 900 veh/hr/lane, showed delay reduction up to 21%. However, the performance of the proposed method might not be consistent with various scenarios since the method should be trained for all possible scenarios to be applicable in general situations. Soleimaniamiri et al. (2020) developed a joint signal timing and trajectory optimization problem to minimize travel delay and energy consumption. To reduce the complexity of the problem and achieve the exact analytical solution, the trajectory shape of CAVs is assumed to be a piece-wise quadratic function. In addition, the nonlinear fuel consumption function is replaced with a simplified quadratic macroscopic approximation. Numerical results in this study confirmed the effectiveness of joint signal timing and trajectory optimization in reducing the travel time and fuel consumption. However, the proposed method is limited to the signal timing control with only two phases.

Li et al. (2020) looked at the signal timing and variable speed limit problem from a macroscopic perspective. In fact, the authors tried to relate the application of VSL on freeway facilities to the signal timing plan of the connecting arterials to reduce the congestion at the entire network. The network is represented from mesoscopic perspective using the queuing theory. The study used the simultaneous perturbation stochastic approximation (SPSA) heuristic approach to minimize the total delay and average travel time of vehicles in the network-wide problem. As a result of this strategy, the local traffic congested is alleviated. Li and Zhou (2017) optimized signal timings and trajectories of CAVs in a mixed environment network with human-driven vehicles. They reduced the complexity of the problem by representing the traffic dynamic and signal timing constraints within a phase-time traffic hyper-network. However, this representation required a

first-order heterogeneous traffic model. Their proposed approach provides invaluable insights in studying large-scale networks; however, its accuracy for operational purposes is not as high as the proposed model in our study that uses car-following models. The performance of the phase-time traffic hyper-network approach is not clear when a second-order traffic flow or microscopic car following model is considered, where the complexity of the problem will increase significantly.

In summary, the signal timing and trajectory optimization problem was mostly solved through enumerative techniques for simple intersection layouts in previous studies. Besides, previous studies relied on simplifying assumptions or heuristic techniques to solve the STTO problem for more complex case studies. As a result, the optimality properties of signal timing plans and vehicle trajectories were not addressed even though significant improvements in the intersection performance were achieved. Therefore, more efficient solution techniques are required to reduce the complexity of the problem and solve such a complex problem in short amount of time.

2.5. Summary of the literature

The speed control problem has been studied from various perspective in freeway facilities, arterial street, isolated intersection, and transportation network. Various problem formulations and solution techniques are provided to find the optimal speed to be communicated with vehicles. On freeway facilities, variable speed limiting is the prevailing strategy to regulate the flow of traffic over segments of the road and avoid the creation of congestion near bottlenecks. Signalized intersections are the major bottlenecks on transportation network, where vehicles have to stop behind traffic light to provide the right of way to the conflicting movements. Therefore, regulating vehicles speed upstream of the intersection helps them arrive the intersection during the green traffic light. Therefore, unnecessary stops will be avoided. Coordinating signal timing plans with

the optimal arrival time of vehicles provides the best mobility performance over the transportation network. The results of all previous studies showed that speed control technique will benefit the transportation system in terms of safety, environmental impacts and fuel consumption, and mobility of the system.

However, the remaining challenge in speed control problem is providing the near optimal solutions for the connected and automated vehicles in real-time. In fact, most of the current studies are limited in finding the solutions for simple case studies with low traffic volume and small size of the network. For more complicated cases, heuristic or meta-heuristic solution techniques are utilized to solve the complex problem. However, the quality of solutions in terms of optimality is not studied in the previous works. In addition, simplifying assumptions are considered in previous studies to reduce the complexity of the problem. Previous studies show the effectiveness of speed control problem (a) in intersections with simple layouts (e.g., one-way streets, or no turning movements), (b) under low traffic volumes, or (c) using simplifying or restrictive assumptions (e.g., using first-order traffic flow model to update trajectory of CAVs, or optimizing the trajectory of a portion of CAVs). Approximation methods and heuristic algorithms are also used to solve the problem in more complex conditions; however, at the expense of sacrificing the quality of the solutions.

Therefore, a general solution technique needs to be investigated to be able to provide near-optimal speeds for vehicles in the network regardless of the problem complexity and traffic congestion condition. In addition, the presence of human drivers beside CAVs should be considered to plan for the transition period before fully automated condition.

CHAPTER 3: DYNAMIC SPEED HARMONIZATION IN CONNECTED URBAN STREET NETWORKS

This chapter proposes a mathematical formulation for dynamic speed harmonization in urban street networks. The underlying assumption is that all vehicles are connected and will receive advisory speed information via infrastructure to vehicle communications. The dynamic speed harmonization (DSH) problem is formulated into a Non-Linear Program and then converted to an equivalent Linear Program utilizing the characteristics of fundamental traffic flow diagram to reduce its complexity. Moreover, to account for stochasticity in traffic demand and reduce the computational complexity of the problem, a Model Predictive Control (MPC) algorithm is developed that predicts the future traffic state in the network and optimizes the speeds in advance to be implemented in the network.

3.1. Research gaps and contributions

The review of the literature reveals that most of previous studies on speed harmonization strategies in urban street networks are based on static traffic models, as such, are not dynamic and responsive to changes in traffic demand. We develop a novel multi-objective optimization program as well as analytical solution techniques to dynamically harmonize speeds in urban street networks to facilitate smooth vehicular traffic on urban street networks. The nonlinear DSH problem is converted to an equivalent linear program using fundamental traffic flow relationship. This linearization reduces the complexity of the problem significantly, so that it can be solved in a reasonable amount of time. Furthermore, to account for stochastic transportation demand and further reducing the complexity of the solution technique, we will develop a model predictive control approach to solve the problem.

3.2. Problem formulation

The proposed formulation is based on the cell transmission model (Carlos F. Daganzo, 1994). The cell transmission model is a discrete approximation of the hydrodynamic models developed by Lighthill and Whitham (1955) and Richards (1956), where the time interval is discretized into small values. Consequently, every link in the network is divided into homogenous cells. Homogeneity is defined as considering the same driving characteristics for all vehicles in a segment of the road. The length of cells is equal to the distance traveled in free flow conditions by a vehicle in one time-step. Let C, C_R, C_S, and C_I respectively denote the set of all cells, source cells, sink cells, and intersection cells. We define T as the set of all discrete time steps, Γ_i as the set of all cell successor to cell $i \in C$, and Γ_i^{-1} as the set of all cells predecessor to cell $i \in C$. The decision variables of the problem are the number of vehicles x_i^t in cell $i \in C$ at time step $t \in T$ and the number of vehicles y_{ij}^t flowing from cell $i \in C$ to its subsequent downstream cell $j \in \Gamma_i$ at time step $t \in T$. We also define v_i^t as the space mean speed in cell $i \in C$ at time step $t \in T$. Table 3-1 summarizes the notations used in this study.

We maximize the total number of vehicles x_i^t in sink cells $i \in C_S$ over all time steps, which maximizes network throughput and favors reaching to sink cells earlier, see the first term in (3.1). We also include a second term to address speed harmonization in the network by minimizing the difference of average speed both spatially and temporally. Minimizing speed difference between two subsequent cells addresses spatial and minimizing speed difference over two subsequent time steps in a cell addresses temporal speed differences, see the second term in (3.1). We also define γ (vehicle/mph) as a weight factor to define the desired emphasis on each terms of the objective function.

$$\text{Max}\left[\sum_{t\in T}\sum_{i\in C_S} x_i^t - \gamma\sum_{t\in T}\sum_{i\in C\setminus\{C_s\}}\sum_{j\in\{i,\Gamma_i\}}\left|v_i^t - v_j^{t+1}\right|\right] \tag{3.1}$$

Cell transmission model assumes that the fundamental flow-density diagram follows a trapezoid shape. Figure 3-1 shows flow and density relationship as defined by this relationship (Carlos F. Daganzo, 1994). In this figure, k_j is the jam density and q_{max} is the maximum flow. As the fundamental traffic flow relationship suggests, the average space mean speed for cell $i \in \mathrm{C}$ at time step $t \in T$ could be found based on equations (3.2).

$$v_i^t = \frac{q_i^t}{k_i^t} = \frac{\sum_{j\in\Gamma_i} y_{ij}^t/\Delta t}{x_i^t/\Delta x} = \frac{\sum_{j\in\Gamma_i} y_{ij}^t}{x_i^t}\frac{\Delta x}{\Delta t} \tag{3.2}$$

Considering Δt and Δx as the duration of a time step and the length of a cell, respectively, the ratio of $\frac{\Delta x}{\Delta t}$ would be equal to the free flow speed v_f in the cell transmission model. Thus, the space mean speed in cell $i \in C$ at time step $t \in T$ is found by equations (3.3) (Zhu and Ukkusuri, 2014). Average space mean speed is equal to the ratio of outflow ($\sum_{j\in\Gamma_i} y_{ij}^t$) to the occupancy ($x_i^t$) at cell $i \in C$ at time step $t \in T$, which is multiplied by the free flow speed (v_f) when the occupancy is greater than zero. Otherwise, it is equal to the free flow speed.

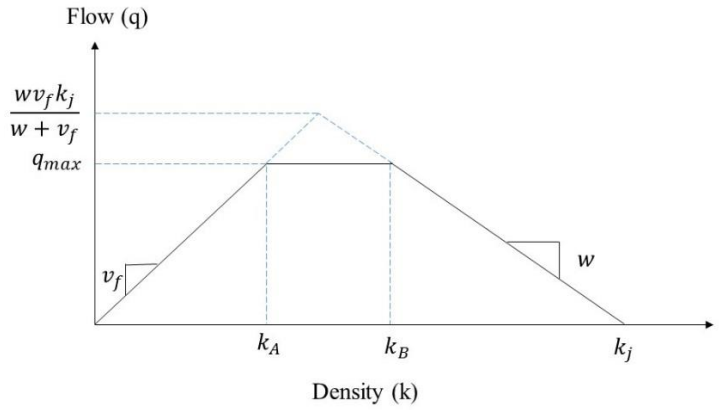


Figure 3-1: Fundamental diagram of traffic flow in CTM

Table 3-1: Summary of notations

Sets	
C	set of all cells
C_R	set of all source cells
C_S	set of all sink cells
C_I	set of all intersection cells
Γ_i	set of all successors of cell $i \in C$
Γ_i^{-1}	set of all predecessors of cell $i \in C$
T	Set of all time intervals
Variables	
x_i^t	number of vehicles in cell $i \in C$ at time step $t \in T$
y_{ij}^t	number of vehicles moving from cell $i \in C$ to cell $j \in \Gamma_i$ at time step $t \in T$
v_i^t	space mean speed in cell $i \in C$ at time step $t \in T$
z_{ii}^t	auxiliary variable
u_{ii}^t	auxiliary variable
Parameters	
N_i	jam density of cell $i \in C$
Q_i^t	outflow capacity of cell $i \in C$ at time step $t \in T$
D_r^t	demand from source cell $r \in C_R$ at time step $t \in T$
f_i^t	green time indicator for signals at cell $i \in C_I$ at time step $t \in T$
v_f	free flow speed
w	shock wave speed
β_i	turning percentage at intersection cell $i \in C_I$
k_j	jam density
q_{max}	maximum flow
γ	weight factor
α	weight factor
θ	the angle between speed vectors
Indices	
i	index for cells
j	index for cells
u	index for cells
p	index for cells
k	index for cells
t	index for time interval
r	index for source cells
s	index for sink cells

$$v_i^t = \begin{cases} \dfrac{\sum_{j\in\Gamma_i} y_{ij}^t}{x_i^t} v_f & x_i^t > 0 \\ v_f & x_i^t = 0 \end{cases} \qquad \forall\, i \in C, t \in T \qquad (3.3)$$

Constraints (3.4) ensure conservation of flow for ordinary cells ($i \in C \setminus \{C_S, C_R\}$), source cells ($i \in C_R$), and sink cells ($i \in C_S$), where δ_{ij} represents Kronecker delta ($\delta_{ij} = 1$ when $i = j$; otherwise $\delta_{ij} = 0$).

$$(\delta_{il} + \delta_{is}) \sum_{u \in \Gamma_i^{-1}} y_{ui}^t - (\delta_{ir} + \delta_{il}) \sum_{j \in \Gamma_i} y_{ij}^t + D_i^t(\delta_{ir}) = (\delta_{il} + \delta_{ir} + \delta_{is})(x_i^{t+1} - x_i^t) \quad \forall i \in C, \forall l \in C \setminus \{C_S, C_R\}, \forall r \in C_R, \forall s \in C_S, t \in T \tag{3.4}$$

The flows in a cell transmission model are limited by the occupancies of the cells, the saturation flow rates, and the traffic light signal status. Constraints (3.5) to (3.9) limit the flow between two cells over all time steps. Constraints (3.5) ensure that the outflow from cell $i \in C \setminus C_s$ to cell $j \in \Gamma_i$ is not greater than the occupancy of the cell $i \in C \setminus C_s$ at time step $t \in T$. Constraints (3.6) and (3.7) respectively limit the flow between cells $i \in C \setminus C_s$ and $j \in \Gamma_i$ to their saturation flow rates Q_i^t and Q_j^t at each time step $t \in T$. Constraints (3.8) ensure no vehicle passes an approach when its corresponding signal is red, by multiplying the intersections cell's saturation flow rate by a binary variable f_i^t that takes on a value of zero when the signal indication is red. It should be noted that constraints (3.8) interrupt the flow when the signal is red and vehicles are accumulated in the intersections cells (and other cells upstream of the signal) and stopping shockwave starts traveling backward. In addition, these constraints prevent conflicting movements to proceed at intersection cells at the same time. Constraints (3.9) ensure that the inflow to cell $j \in C \setminus C_R$ from cell $i \in \Gamma_j^{-1}$ is less than or equal to the available capacity of cell $j \in C \setminus C_R$.

Constraints (3.10) ensure that turning percentages at intersection cells are equal to the predefined turning proportions (β_j^t). The number of vehicles travelling to cell $j \in C_I$ from cell $i \in \Gamma_j^{-1}$ should be equal to the product of corresponding turning proportion β_j^t and the total number of

vehicles $\sum_{k\in\Gamma_i} y_{ik}^t$ leaving cell $i \in \Gamma_j^{-1}$ to all immediately downstream cells. Finally, Constraints (3.11) and (3.12) ensure non-negativity of the decision variables.

$$\sum_{j\in\Gamma_i} y_{ij}^t \le x_i^t \qquad \forall\, i \in C\backslash C_S, t \in T \qquad (3.5)$$

$$\sum_{j\in\Gamma_i} y_{ij}^t \le Q_i^t \qquad \forall\, i \in C\backslash C_S, t \in T \qquad (3.6)$$

$$\sum_{i\in\Gamma_j^{-1}} y_{ij}^t \le Q_j^t \qquad \forall\, j \in C\backslash C_R, t \in T \qquad (3.7)$$

$$\sum_{j\in\Gamma_i} y_{ij}^t \le f_i^t Q_i^t \qquad \forall\, i \in C_I, t \in T \qquad (3.8)$$

$$\sum_{i\in\Gamma_j^{-1}} y_{ij}^t \le \frac{\mathrm{w}}{\mathrm{v}_f}\left(N_j - x_j^t\right) \qquad \forall\, j \in C\backslash C_R, t \in T \qquad (3.9)$$

$$y_{ij}^t = \beta_j^t \sum_{k\in\Gamma_i} y_{ik}^t \qquad \forall\, j \in C_I, i \in \Gamma_j^{-1}, t \in T \qquad (3.10)$$

$$x_i^t \ge 0 \qquad \forall\, i \in C, t \in T \qquad (3.11)$$

$$y_{ij}^t \ge 0 \qquad \forall\, i \in C\backslash C_S, j \in \Gamma_i, t \in T \qquad (3.12)$$

3.3. Methodology

All constraints of the developed dynamic speed harmonization formulation are linear with no integer variable. However, the objective function of the problem is nonlinear and nonconvex due to presence of fractional terms. We utilize fundamental diagram of traffic flow characteristics to address this issue. The space mean speed of vehicles in cell $i \in C$ at time step $t \in T$ is equal to the slope of a vector that connects the origin to point $\left(x_i^t, \sum_p y_{ip}^t\right)$ on the fundamental diagram, see

Figure 3-2. Similarly, the space mean speed of vehicles in the subsequent cell $j \in \Gamma_i$ at time step $t+1 \in T$ is equal to the slope of a vector that connects the origin to point $(x_j^{t+1}, \sum_k y_{jk}^{t+1})$, where $k \in \Gamma_j$. Minimizing speed difference between the two cells is equivalent to minimizing the angle between the two speed vectors, as shown by θ in Figure 3-2.

Furthermore, minimizing the angle between space mean speed vectors v_i^t and v_j^{t+1} is equivalent to minimizing the difference between the distances of points $\left(x_i^t, \sum_p y_{ip}^t\right)$ and $\left(x_j^{t+1}, \sum_k y_{jk}^{t+1}\right)$ to the vector representing the free flow speed, see Figure 3-2. Note that in the cell transmission model, the average speed of a cell is equal to free flow speed when all vehicles in the cell exit it in a single time step. In other words, the outflow $\sum_p y_{ip}^t$ is equal to the occupancy x_i^t of the cell as shown in (3.13).

$$\frac{\sum_{p \in \Gamma_i} y_{ip}^t}{x_i^t} = 1 \qquad \forall\, i \in C, t \in T \qquad (3.13)$$

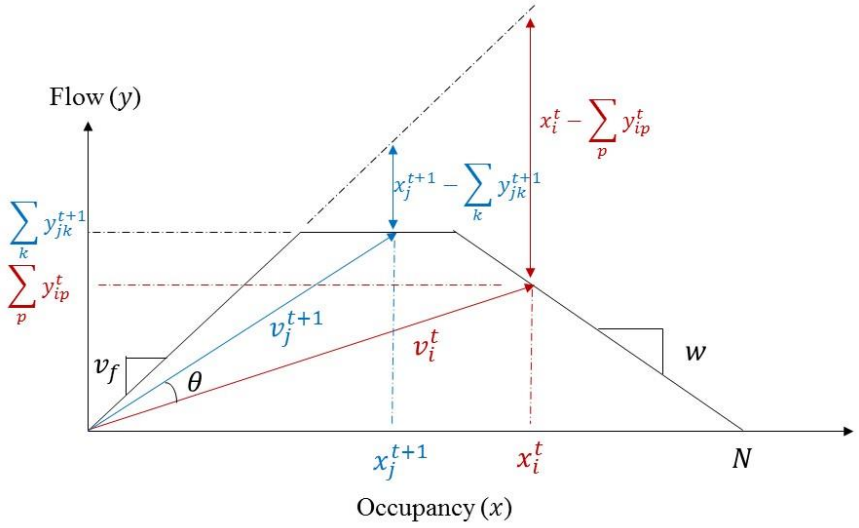


Figure 3-2: Linearizing speed differences using traffic flow fundamental diagram

The distance of point $\left(x_i^t, \sum_p y_{ip}^t\right)$ from the free flow speed vector is equal to $x_i^t - \sum_p y_{ip}^t$. Therefore, the objective function (3.1) can be equivalently rewritten as (3.14). Note that both terms

of the objective functions have similar units. As such, rather than using term γ with vehicle/mph unit, we use unit-less term α.

$$\alpha \sum_{t\in T}\sum_{i\in C_S} x_i^t - (1-\alpha)\sum_{t\in T}\sum_{i\in C\setminus C_s}\sum_{j\in\{\mathrm{i},\,\Gamma_i\}}\left|\left(x_i^t - \sum_p y_{ip}^t\right) - \left(x_j^{t+1} - \sum_k y_{jk}^{t+1}\right)\right| \tag{3.14}$$

By this transformation, the fractional terms of the objective function are removed and the complexity of the problem reduced; however, the absolute value function still makes the mathematical program non-linear. We define non-negative variables z_{ij}^t and u_{ij}^{t+1} to represent the differences between the values in the absolute function. Using these two variables, the objective function is replaced with the new objective function (3.15) and constraints (3.16) - (3.18). As such, the new objective function and all other constraints are linear and the problem can be solved efficiently.

$$\alpha \sum_{t\in T}\sum_{i\in C_S} x_i^t - (1-\alpha)\sum_{t\in T}\sum_{i\in C\setminus C_s}\sum_{j\in\{\mathrm{i},\,\Gamma_i\}}\left(z_{ij}^t + u_{ij}^{t+1}\right) \tag{3.15}$$

$$z_{ij}^t - u_{ij}^{t+1} = \left(x_i^t - \sum_{p\in\,\Gamma_i} y_{ip}^t\right) - \left(x_j^{t+1} - \sum_{k\in\,\Gamma_j} y_{jk}^{t+1}\right) \quad \forall\, i \in \mathrm{C}\setminus\mathrm{C_S}, j \in \{i,\,\Gamma_i\}, t \in T \tag{3.16}$$

$$z_{ij}^t \geq 0 \qquad \forall\, i \in \mathrm{C}\setminus\mathrm{C_S}, j \in \{i,\,\Gamma_i\}, t \in T \tag{3.17}$$

$$u_{ij}^t \geq 0 \qquad \forall\, i \in \mathrm{C}\setminus\mathrm{C_S}, j \in \{i,\,\Gamma_i\}, t \in T \tag{3.18}$$

To account for unforeseen changes in traffic demand and capacity and further improve the efficiency of the optimization problem, a model predictive control approach (i.e., rolling horizon) is used. This will allow harmonizing speed in large-scale transportation networks. With a rolling horizon strategy, the problem is broken into fewer number of time steps. The main concept is to

solve the linear problem over a prediction period including N_p time steps. Then the solution for the first time step is implemented as the time horizon is rolled forward for one time step. The process is continued to cover the entire analysis period.

3.4. Case Study

The case study network is a portion of downtown Springfield, Illinois. The network has 20 intersections and a mix of one-way and two-way streets with different number of lanes and turning configurations at signalized intersections. Figure 3-3 shows the cell representation of the network with 316 cells.

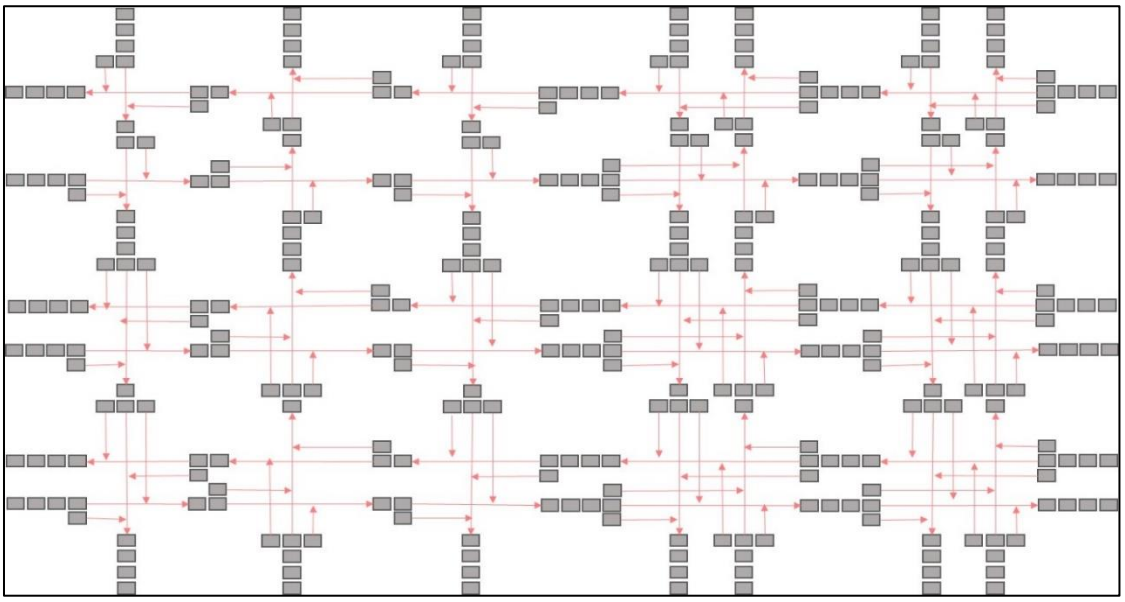


Figure 3-3: Springfield network

The signal timing parameters are obtained according to Hajbabaie and Benekohal (2011, 2012, and 2015), maximizing the total number of the completed trips in the network. The cycle times and the number of phases could vary and the left turn movements are protected. It should be noted that the signal timings are optimized before implementing DSH strategy and there is no

cooperation between DSH and signal timing optimization. The maximum speed limit is 30 mph. Table 3-2 presents the general characteristics of the network.

Table 3-2: Characteristics of Springfield network

Link data	
Number of lanes per link	1, 2, or 3
Maximum free-flow speed (mph)	30
Link saturation flow (vphpl)	1800
Optimization period (time steps)	250
Duration of each time step (seconds)	6
Shock wave speed (mph)	30
Number of cells	316
Cell jam density (veh/cell/ln)	12
Maximum cell saturation flow (veh/cell/ln)	3

We used an asymmetric demand pattern covering both under- and oversaturated conditions. Demand profiles for east-west and north-south streets are shown in Figure 3-4. The total study period is 250 time steps (25 minutes) and the prediction horizon is 50 time steps (5 minutes). The time steps are six seconds and decisions for advisory speeds are implemented for the next time step.

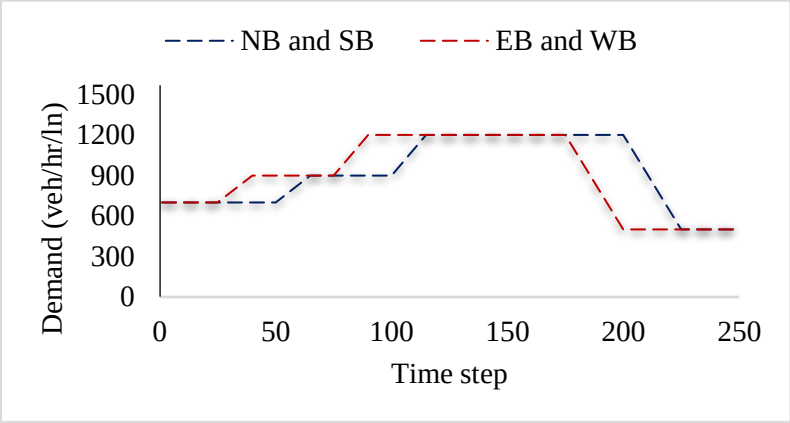


Figure 3-4: Demand profile

3.5. Results

In this section, the effects of dynamic speed harmonization on network performance are studied. Figure 3-5 shows the network-level average speed over time for three different scenarios: (1) no speed harmonization, (2) speed harmonization without rolling horizon, and (3) speed harmonization with rolling horizon.

For the symmetric undersaturated demand conditions (500 veh/hr/ln on all entry points), the DSH strategy did not significantly change network-wide average speed as the network was uncongested. Having said that, DSH increased the average speed in intersection cells while reducing it in ordinary cells in all time steps. Harmonizing the speed reduced average speed on cells upstream of intersection cells to prevent unnecessary stops at intersection cells. By reducing the average speed, speed variations in ordinary cells are decreased as well. As shown in Figure 3-5, when the problem is solved with the rolling time horizon approach, the same results are achieved.

When the demand increases to 900 veh/hr/ln in all entry points (saturated demand pattern), the network is at the capacity level. Like the previous demand pattern, the DSH strategy increased speed in intersections cells; however, the difference was less than the observed difference in the uncongested demand pattern. In addition, average speed and its variation were both less at ordinary cells as a result of speed harmonization. Note that the solutions of the approach with rolling horizon was similar to those obtained without it.

In the symmetric oversaturated demand pattern (1200 veh/hr/ln), speed difference in network links is less, as the network is highly congested and vehicles travel at low speeds. As a result, DSH would not be as effective as the cases with lower congestion levels. Having said that, speed harmonization still reduced speed variation and increased average speed at intersection cells.

As shown in Figure 3-5, when the network gets more congested, the average speed with speed harmonization would be closer to the average speed without harmonizing the speed.

The case with asymmetric demand pattern supports the findings of other cases. Speed harmonization increased average speed at intersection cells while reducing speed variation. In addition, the average speed and speed variation at ordinary cells were less than the case without speed harmonization. It should be noted that the network is at the saturated demand level at the beginning of the analysis period. In undersaturated demand pattern, the network is less congested and average speed increases in both not-harmonized and harmonized cases. However, the average speed in harmonized case increases smother than the not-harmonized case. Increasing the demand in saturated and over-saturated demand patterns leads to an increase in the network congestion and therefore, the average speed decreases over time. Again, the rate of changes in the average speed is smoother in harmonized cases.

Figure 3-6 shows average speed (over time in each cell) distribution over the network for all demand patterns. DSH increased the average speed in intersection cells, while reducing it at most of their upstream cells. By utilizing this strategy, the number of stops at intersections was reduced (see Table 3-3). Moreover, increasing the demand yielded larger speed reductions when the speeds were not harmonized, indicating the capability of the DSH technique to maintain higher speeds in higher demand levels.

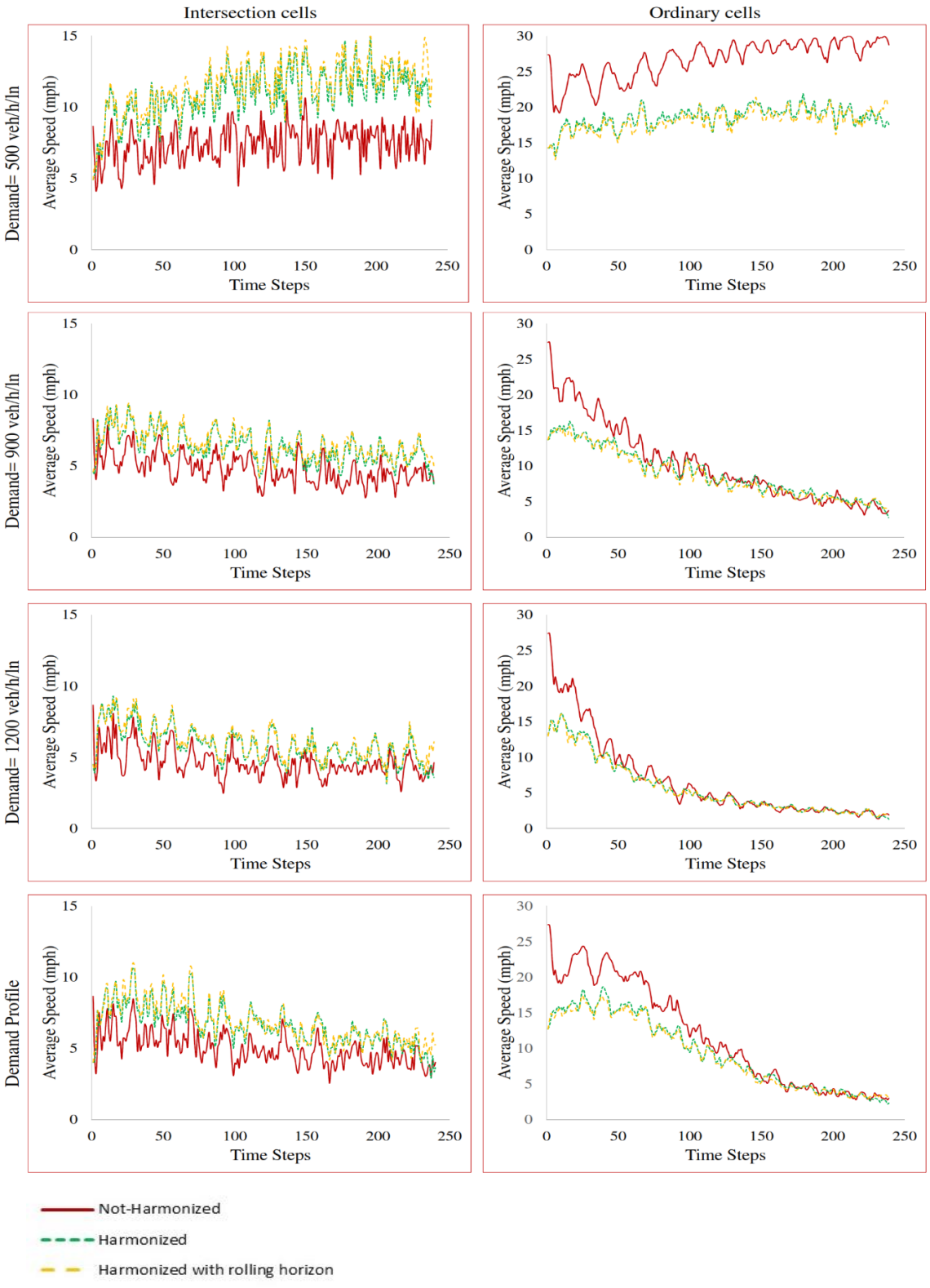


Figure 3-5: The comparison of speed variations on the network with different demand patterns

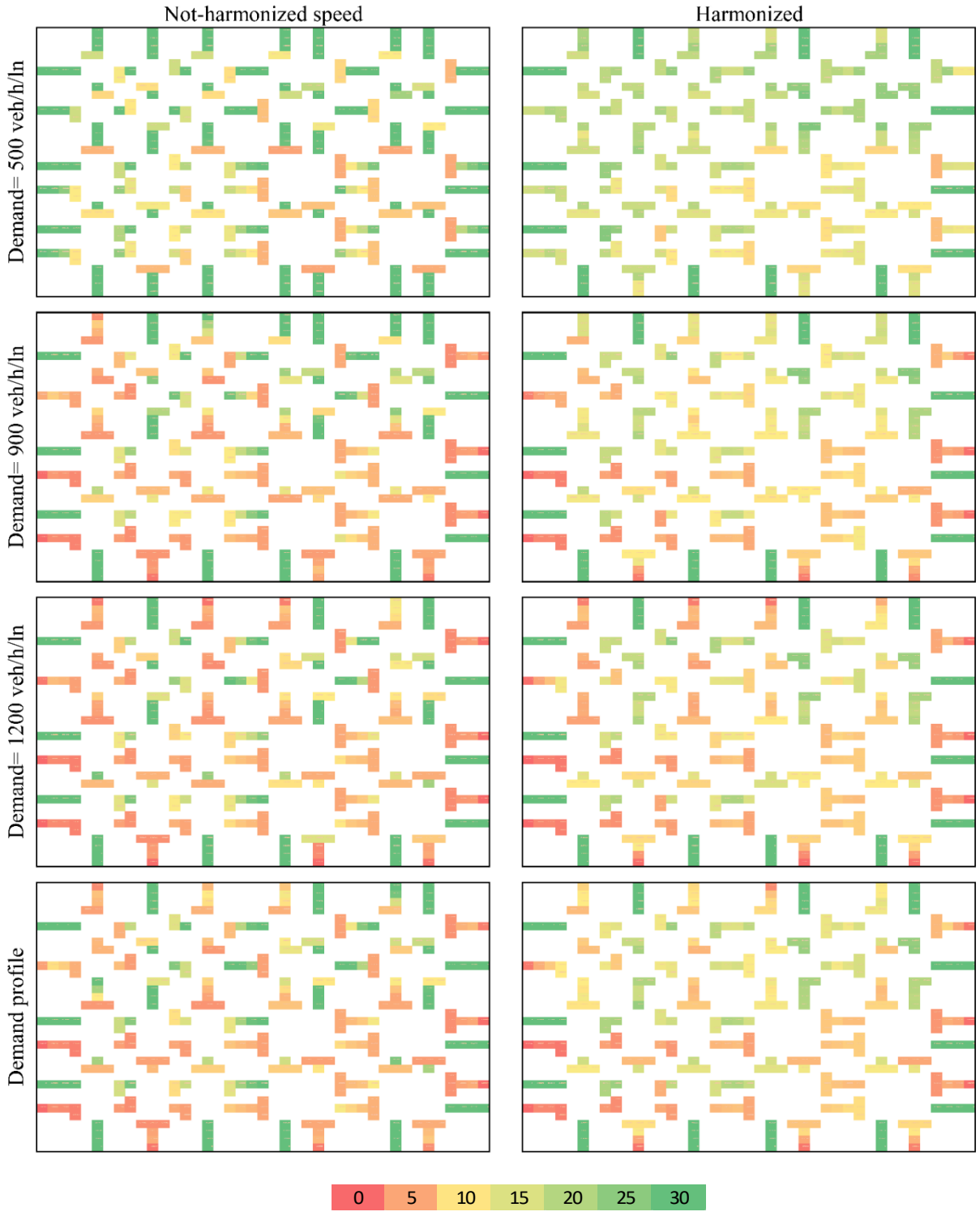


Figure 3-6: Distribution of the average speed over the network

As discussed in the methodology section, there is a trade-off between the two terms of the objective function: maximizing network throughput and minimizing speed variation. Figure 3-7 shows the results of a sensitivity analysis on the value of α, the weight between the two terms of the objective function. Increasing the value of α, decreases the travel time while average speed and the variance of the speed increases. This is expected as a higher α value gives more priority to maximizing network throughput, which consequently results in lower total travel time. In addition, we observed a negative relationship between the travel time and the speed variance. As the variance of the speed decreases the total travel time in the network increases, indicating that highly harmonized speeds reduce the efficiency of network performance. Hence, an appropriate value of α can provide the desired balance. It should be noted that the results provided in Figure 3-7 are for interrupted traffic flow where vehicles have to stop at the intersections when traffic light is red. The pattern could be different in un-interrupted traffic flow, where vehicles do not necessarily need to stop.

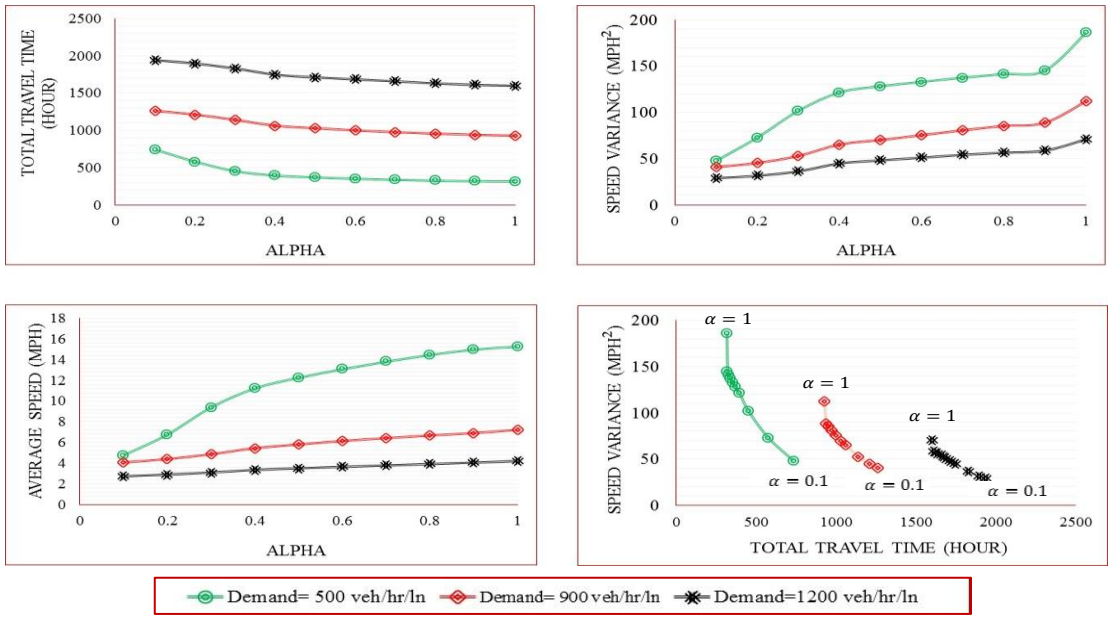


Figure 3-7: The relationship between travel time, average speed, and speed variance in the network with alpha

Table 3-3 illustrates the mobility performance results in Springfield network for different demand patterns when α is 0.95. In symmetric undersaturated demand pattern, while speed harmonization kept the travel time, average speed, and network throughput, similar to not harmonizing the speeds, it reduced speed variance by more than 29% and the number of stops by more than 8%, see Table 3-3. Note that we do not expect huge changes in network performance as a result of speed harmonization in undersaturated demand pattern. As shown in Figure 3-5, the average speed at intersection cells are much higher than those in non-harmonized case, indicating that DSH reduced the number of unnecessary stops.

Speed harmonization shows significant improvements in network performance in the symmetric saturated demand pattern. DSH reduced the total travel time by 5.4%. In addition, the average speed and the network throughput were increased by 5.9% and 4.0%, respectively. The speed variance decreased by 20.2% and the number of stops by 16.8%.

For the symmetric oversaturated demand pattern (1200 veh/hr/ln in all entry points), the travel time, speed variance, and number of stops are reduced by 1.5%, 19.8%, and 18.5%, respectively. On the other hand, the average speed and the network throughput increased by 0.1% and 1.6% respectively. Similar trends were observed for the asymmetric demand pattern.

Table 3-3 shows that the DSH yielded lower travel times and more completed trips at the saturation demand level than it did at the under- and oversaturated demand levels. Vehicles move efficiently in the network in under-saturated conditions (i.e., 500 veh/hr/ln). Therefore, DSH does not lead to substantial improvement in network performance, while it still reduces speed variance. The flow is often in stop and go condition in the oversaturated demand level (i.e., 1200 veh/hr/ln) with vehicles traveling uniformly at low speeds. As a result, DSH leads to some improvement in network performance and reductions in speed variance. On the other hand, there is more variations

in speeds and more room for improvement in network performance around the saturation demand level (i.e., 900 veh/hr/ln) as such DSH yields the highest reduction in total travel time and highest increase in the number of completed trips.

Table 3-3: Mobility performance of the network

Demand (veh/hr/ln)	Mobility performance	Harmonized Speed	Not-Harmonized Speed	Difference (%)
500	Travel time (hour)	53.0	53.0	0.0
	Average speed (mph)	15.4	15.6	-1.2
	Variance (mph^2)	140.1	198.4	-29.4
	Number of stops	90954.1	99150.4	-8.3
	Number of completed trips	6197.2	6202.2	-0.1
900	Travel time (hour)	155.1	163.9	-5.4
	Average speed (mph)	7.0	6.7	5.9
	Variance (mph^2)	90.7	113.7	-20.2
	Number of stops	125560.2	150897.7	-16.8
	Number of completed trips	8032.5	7727.0	4.0
1200	Travel time (hour)	267.8	271.9	-1.5
	Average speed (mph)	4.1	4.1	0.1
	Variance (mph^2)	59.33	74.0	-19.8
	Number of stops	127750.2	156695.9	-18.5
	Number of completed trips	7980.2	7857.5	1.6
Time-Variant Demand Profile	Travel time (hour)	168.9	172.5	-2.1
	Average speed (mph)	6.4	6.3	0.5
	Variance (mph^2)	87.9	111.5	-21.2
	Number of stops	122948.7	146733.2	-16.2
	Number of completed trips	7850.8	7693.3	2.0

Figure 3-8 compares the solutions of DSH with and without the rolling horizon feature. The rolling horizon approach provided solutions that were very close to the true optimal solutions of the problem (obtained by the central approach). At the same time, the rolling horizon approach is computationally less expensive, and can account for unforeseen changes in traffic demand and capacity. Furthermore, the rolling horizon approach allows us to analyze driver compliance rates with the speeds suggested by the algorithm. As a result, it can be used to harmonize speeds in urban street networks.

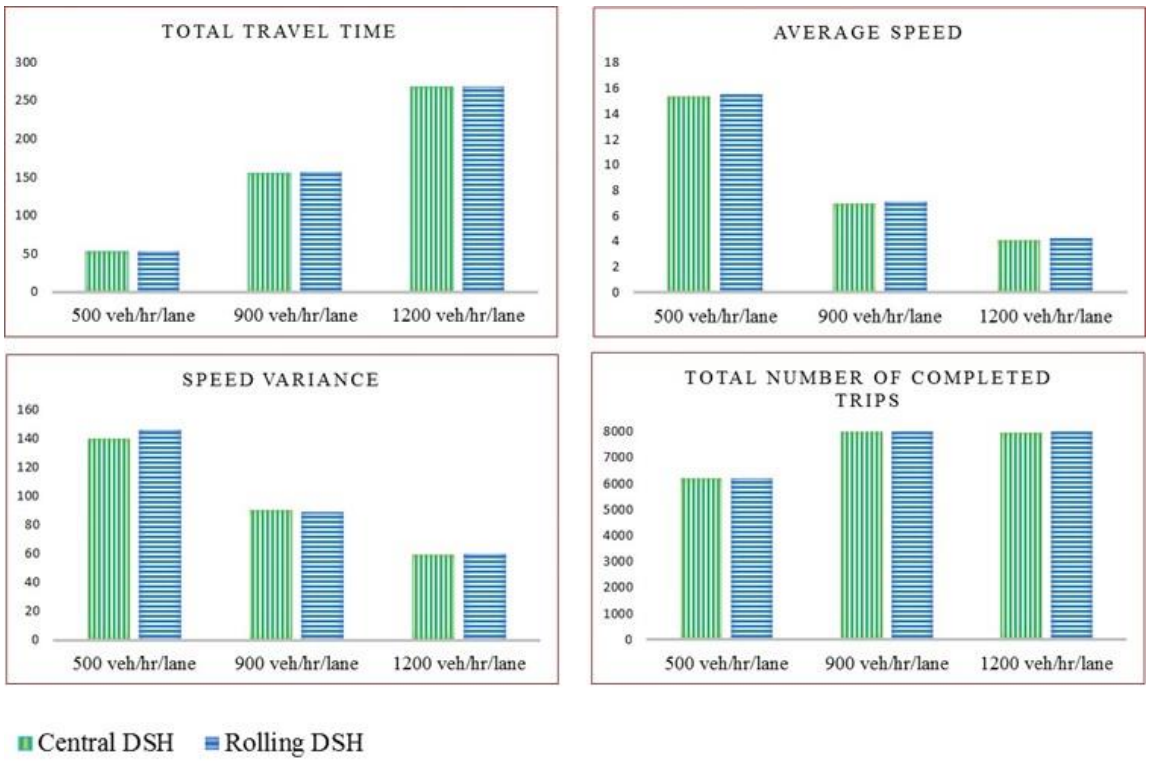


Figure 3-8: Comparison of the network performances between Central DSH and Rolling DSH

3.6. Conclusion

This study formulates dynamic speed harmonization in urban street networks based on the cell transmission model. We developed a non-linear multi-objective mathematical program to represent the problem and converted it to a linear program utilizing characteristics of the fundamental traffic flow relationship. As a result of this transformation, the complexity of the problem was significantly reduced. To account for unforeseen changes in traffic demand and to further reduce problem complexity, a model predictive control approach was developed. The approach collected data at a certain time, dynamically harmonized speeds over an extended prediction period, and implemented the optimal speed only for the next time step and continued throughout the entire study period.

We tested the proposed formulation and solution technique in a realistic case study network of 20 intersections with a combination of one-way and two-way streets with various number of lanes. The results indicated that the solution technique effectively optimized dynamic advisory speeds for connected vehicles throughout the network. Although the proposed approach is based on a first order model, it improves the network performance significantly since it regulates the flow by changing speeds and prevents spillovers and gridlocks. As a result of the optimization, speed in the network was harmonized and network performance was improved. The algorithm reduced the travel time by 2.1%, speed variance by 21.2% and number of stops by 16.2%. At the same time, DSH increased the average speed by up to 0.5% and increased the number of completed trips by 2% in our case study network under an asymmetric demand pattern. Furthermore, the proposed DSH formulation decreased average speed on ordinary cells and increased it at intersection cells. The spatial distribution of the speed over the network indicated smoother flow as a result of the DSH strategy.

The proposed approach can efficiently determine optimal advisory speeds in the case study network and in urban street networks of larger spatial and temporal scales. Having said that, the solutions will not be real-time as the size of the network grows. Developing real-time methods for DSH in urban street networks is provided in next chapter. In addition, it should be noted that the results are achieved based on a deterministic and fixed capacity assumption. However, capacity has a stochastic nature. Addressing stochastic capacity in speed harmonization problem is an interesting topic that should be addressed by further research.

CHAPTER 4: DISTRIBUTED OPTIMIZATION AND COORDINATION ALGORITHMS FOR DYNAMIC SPEED OPTIMIZATION

As discussed in previous chapter, dynamic speed harmonization provides great potential to smoothen the flow of traffic and reduce travel time in urban street networks. The existing methods, while providing great insights, are neither scalable nor real-time. This chapter develops Distributed Optimization and Coordination Algorithms (DOCA) to solve the dynamic speed optimization (DSO) problem for CAVs in urban street networks. DOCA decomposes the network-level nonlinear DSO problem into several subnetwork-level (a.k.a. computation node level) nonlinear problems. As a result of this decomposition, the computational complexity of the problem is reduced significantly, and solutions can be found in real-time. DOCA creates effective coordination between each pair of neighboring subnetworks by sharing information on the number of vehicles traveling from one subnetwork to another, the available capacity, and the average speed of the receiving subnetwork. Besides, the signal status of the downstream subnetwork is shared between the controllers. Information from all neighboring subnetworks will be implemented in the nonlinear DSO program of each subnetwork. The coordination pushes the solution toward optimality and ensures finding near-optimal solutions. Finally, DOCA is integrated into a Model Predictive Control (MPC) approach to reduce problem complexity further and account for adaptive transportation demand and capacity.

4.1. Strategies to reduce the computational complexity of optimization problems

The complexity of traffic control problems is often reduced in the literature by decomposing the network to several subnetworks and optimizing them individually. Tettamanti & Varga (2010) solved the signal timing problem using Lagrangian Relaxation (LR) utilizing an

MPC solution technique. The quadratic objective function of this problem aimed at minimizing queue length at intersections. They found the dual of this problem and solved it iteratively such that each intersection found its corresponding Lagrangian multipliers. Determining the optimal Lagrangian multipliers led to finding the near-optimal signal timing parameters. The MPC approach transfered the link inflow, outflow, and occupancy and the demand to the next time step. Papageorgiou and Mayr (1982) optimized a nonlinear ramp metering problem by solving its equivalent Hamiltonian function obtained from the Lagrangian decomposition technique. They used a hierarchical approach to decompose the problem by dividing the roadway into different segments and including a central algorithm to compute the Lagrangian multipliers at each segment and time step. However, the computation time of the hierarchal approach was higher than the central approach due to the slow convergence rate of the Lagrangian approach.

Most of the reviewed studies were based on decomposition solution techniques, and they needed central computations to coordinate the decisions of subproblems. However, the connected vehicle environment provides an opportunity for controllers to interact with each other directly without the need to have a central unit. Araghi et al. (2015) studied the signal timing optimization problem in a network that was decomposed into several intersections. They used the Q-learning concept to solve the problem and considered cycle timing parameters as the actions and average delay as the corresponding reward. Adjacent intersection controllers shared information about the number of incoming vehicles in an effort to coordinate their decisions. Their results showed that the solutions reduced the average delay in the network by 35.75% compared to a set of predefined signal timing parameters. Islam and Hajbabaie (Al Islam and Hajbabaie, 2017) proposed distributed optimization and coordination algorithms for network-level signal timing optimization in a connected vehicle environment. Their approach was scalable and could find solutions in real-

time that reduced network travel time between 17% and 48% under various demand patterns. Mehrabiour and Hajbabaie (2017) developed a CTM-based distributed-coordinated methodology for signal timing optimization. Their approach was real-time and scalable with gaps of no more than 2% from a theoretical upper bound on the objective function value. Moreover, Mohebifard and Hajbabaie (2018) developed a real-time traffic metering methodology that increased the network throughput by 2.8-5.9% compared to a no-metering strategy. Their approach was distributed and had a maximum optimality gap of 2.2%.

4.2. Research gap and contributions

The existing speed harmonization methods for urban street networks either find static solutions or find optimal solutions but are not scalable nor real-time. As a result, the solutions of neither approaches are suitable for real-world applications, where transportation networks have various sizes, traffic state is dynamic, and solutions are required in real-time. Decomposition approaches are used to reduce problem complexity; however, they do not scale well and cannot find solutions in real-time. The main contribution of this study is to develop distributed optimization and coordination algorithms suitable for DSO problem that can find near-optimal solutions in real-time and are scalable to transportation networks of various sizes.

4.3. Solution technique

Even though the proposed problem formulation in previous chapter is linear, it is not scalable to large-sized urban street networks or long study periods due to the excessive number of decision variables it has. Therefore, we developed scalable Distributed Optimization and Coordination Algorithms for Dynamic Speed Optimization (DOCA-DSO) capable of finding real-time and near-optimal solutions to the problem. Distributed optimization decomposes the network-

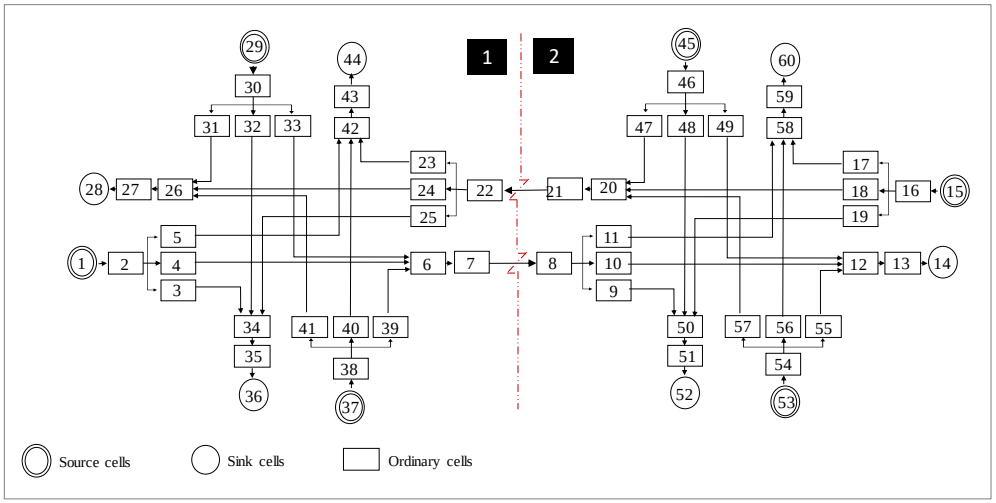
level LP into several subnetwork-level LPs and solves them simultaneously. As such, the computational complexity of the problem is significantly reduced. Distributed coordination among each two subnetwork-level LPs pushes the solutions toward optimality and ensures attaining near-optimal solutions. We will show how far the DOCA-DSO solutions are from the optimal solutions. To further reduce the problem complexity, create consensus between subproblems, and address probabilistic changes in traffic demand and capacity, the solution algorithm is incorporated into an MPC approach.

4.3.1. Distributed optimization

Distributed optimization decomposes the network into several subnetworks, and rather than solving a single network-level LP, solves several subnetwork level LPs. Previous studies showed that decomposing the network to individual intersections (i.e., each subnetwork is an intersection) reduces the problem complexity sufficiently (Al Islam and Hajbabaie, 2017; Mehrabipour and Hajbabaie, 2017). As a result, we decomposed the network into individual intersections. As such, a network of "n" intersections was converted to "n" subnetworks each including a single intersection. Note that a processor exists in traffic signal controller of each intersection. The presence of this computational power will further support intersection-level decomposition, as the signal controller processor can be utilized for speed optimization in each computation node.

The network decomposition is achieved by relaxing each constraint that connects an intersection to its neighbors. In other words, some links between consecutive intersections are eliminated. Figure 4-1-a shows a two-intersection network in the cell representation. By cutting the links that connect cell 7 to 8 and cell 21 to 22, the network will be divided to intersections 1

and 2. Cutting these links in the network shown in Figure 4-1-a is equivalent to relaxing the following constraints:



a) before decomposition

b) after decomposition

Figure 4-1: Two-intersection network

$$x_i^{t+1} = x_i^t + \sum_{u \in \Gamma_i^{-1}} y_{ui}^t - \sum_{j \in \Gamma_i} y_{ij}^t \qquad \forall\, i \in \{7,8,21,22\}, t \in T \quad (4.1)$$

$$\sum_{j\in\Gamma_i} y_{ij}^t \le Q_i \qquad \forall\, i \in \{7,21\}, t \in T \qquad (4.2)$$

$$\sum_{i\in\Gamma_j^{-1}} y_{ij}^t \le Q_j \qquad \forall\, j \in \{8,22\}, t \in T \qquad (4.3)$$

$$\sum_{i\in\Gamma_j^{-1}} y_{ij}^t \le \frac{W}{v_f}\left(N_j - x_j^t\right) \qquad \forall\, j \in \{8,22\}, t \in T \qquad (4.4)$$

$$z_{ij}^t - u_{ij}^{t+1} = \left(x_i^t - \sum_{p\in\,\Gamma_i} y_{ip}^t\right) - \left(x_j^{t+1} - \sum_{k\in\,\Gamma_j} y_{jk}^{t+1}\right) \qquad \forall\, i \in \{6,7,20,21\},\ j \in \{7,8,21,22\}, t \in \mathrm{T} \qquad (4.5)$$

Each intersection requires a source cell at each entry and a sink cell at each exit point; however, source and sink cells do not exist at locations where links are cut. We defined dummy source and sink cells on the links that are cut to make each intersection a stand-alone system. For instance, dummy source cell 22' and dummy sink cell 8'' were added to intersection 1, and dummy source cell 8' and dummy sink cell 22'' were added to intersection 2 as shown in Figure 4-1-b. Let S denote the set of all intersections (i.e., sub-problems) of the network. We defined C_{SD}^n to represent the set of all dummy sink cells and C_{RD}^n to represent the set of all dummy source cells at intersection $n \in S$. Similarly, we added component n to all previously defined sets. Note that the numbering of cells at each intersection is identical to the original numbering in the network. As a result of the decomposition, the original problem is converted to the following LP for each intersection $n \in S$:

$$\mathrm{Max}\left[\alpha \sum_{t\in T} \sum_{i\in\{C_S^n \cup C_{SD}^n\}} x_i^t - (1-\alpha) \sum_{t\in T} \sum_{i\in C\setminus\{C_S^n \cup C_{SD}^n\}} \sum_{j\in\{\mathrm{i},\,\Gamma_i\}} \left(z_{ij}^t + u_{ij}^{t+1}\right)\right] \qquad (4.6)$$

s.t.

$$x_r^{t+1} = D_r^t + x_r^t - \sum_{j\in\Gamma_i} y_{rj}^t \qquad \forall\, r \in \{C_R^n \cup C_{RD}^n\}, t \in T \qquad (4.7)$$

$$x_s^{t+1} = x_s^t + \sum_{i\in\Gamma_s^{-1}} y_{is}^t \qquad \forall\, s \in \{C_S^n \cup C_{SD}^n\}, t \in\ T \qquad (4.8)$$

$$x_i^{t+1} = x_i^t + \sum_{u\in\Gamma_i^{-1}} y_{ui}^t - \sum_{j\in\Gamma_i} y_{ij}^t \qquad \forall\, i \in C^n\backslash\{C_S^n \cup C_{SD}^n \cup C_R^n \cup C_{RD}^n\}, t \in T \qquad (4.9)$$

$$\sum_{j\in\Gamma_i} y_{ij}^t \le x_i^t \qquad \forall\, i \in C^n\backslash\{C_S^n \cup C_{SD}^n\}, t \in T \qquad (4.10)$$

$$\sum_{j\in\Gamma_i} y_{ij}^t \le Q_i \qquad \forall\, i \in C^n\backslash\{C_S^n \cup C_{SD}^n\}, t \in T \qquad (4.11)$$

$$\sum_{i\in\Gamma_j^{-1}} y_{ij}^t \le Q_j \qquad \forall\, j \in C^n\backslash\{C_R^n \cup C_{RD}^n\}, t \in T \qquad (4.12)$$

$$\sum_{i\in\Gamma_j^{-1}} y_{ij}^t \le \frac{\mathrm{W}}{v_f}\left(N_j - x_j^t\right) \qquad \forall\, j \in C^n\backslash\{C_R^n \cup C_{RD}^n\}, t \in T \qquad (4.13)$$

$$\sum_{j\in\Gamma_i} y_{ij}^t \le f_i Q_i \qquad \forall\, i \in C_I^n, t \in T \qquad (4.14)$$

$$y_{ij}^t = \beta_j \sum_{k\in\Gamma_i} y_{ik}^t \qquad \forall\, j \in C_I^n, i \in \Gamma_j^{-1}, t \in T \qquad (4.15)$$

$$z_{ij}^t - u_{ij}^{t+1} = \left(x_i^t - \sum_{p\in\ \Gamma_i} y_{ip}^t\right) - \left(x_j^{t+1} - \sum_{k\in\ \Gamma_j} y_{jk}^{t+1}\right) \qquad \forall\, i \in C^n\backslash\{C_S^n \cup C_{SD}^n\}, j \in\ \{i,\ \Gamma_i\}, t \in T \qquad (4.16)$$

$$z_{ij}^t \ge 0, \qquad \forall\, i \in C^n\backslash\{C_S^n \cup C_{SD}^n\}, j \in\ \{i,\ \Gamma_i\}, t \in T \qquad (4.17)$$

$$u_{ij}^{t} \geq 0, \qquad \forall\, i \in C^{n} \backslash \{C_{S}^{n} \cup C_{SD}^{n}\}, j \in \{i,\ \Gamma_{i}\}, t \in T \qquad (4.18)$$

$$x_{i}^{t} \geq 0, \qquad \forall\, i \in C^{n}, t \in T \qquad (4.19)$$

$$y_{ij}^{t} \geq 0, \qquad \forall\, i \in C^{n} \backslash \{C_{S}^{n} \cup C_{SD}^{n}\}, j \in \Gamma_{i}, t \in T \qquad (4.20)$$

Each subnetwork level LP can be solved efficiently using commercial software packages. In this study, we utilized CPLEX (CPLEX, 2009) to solve them. It is apparent that the demand in dummy source cells along with other information are needed in each LP. This information is provided by the distributed coordination as explained below.

4.3.2. Distributed coordination

Each intersection-level LP can be solved very efficiently; however, their optimal solutions will not lead to the network-level optimal solution. Effective coordination between LPs of each two neighboring intersections is essential to push solutions towards optimality. We created effective coordination by sharing information between each two adjacent intersections (LPs) on (a) the outflow of cells immediately upstream of the broken links, (b) the available capacity of the receiving cells immediately downstream of the broken links, (c) the average speed in the receiving cells immediately downstream of the broken links, and (d) signal timing status at intersection cells downstream to the broken links. For each intersection at a computation node, the information coming from neighboring intersections was incorporated as input parameters in its optimization model. We assume that connected vehicle technology and the Internet of Things will facilitate generating and sharing this information among neighboring intersections in the future.

It is necessary to share signal timing parameters of downstream intersections with the computation nodes. The reason is that if the signal is red at a downstream intersection, there is no benefit in sending too many vehicles towards it. This is especially important for through

movements as they often have higher traffic volume levels. As a result, the objective function at each intersection needs to be updated with the signal timing information received from downstream intersections. For this purpose, we defined F_i^n as the set of intersection cells that represent the through movement at the downstream of dummy sink cell $i \in C_{SD}^n$. Note that these cells always belong to receiving intersections. Moreover, we just considered the signal timing for through movements because they have the highest impact at intersections. For example, in Figure 4-1-a, F_8^1 is equal to $\{10\}$. Let $\hat{f}_i^{t,n}$ take on a value of one when the signal for through movement is green at the intersection cell $i \in F_i^n$, and zero otherwise. To ensure that the number of completed trips at the intersection located at the computation node is maximized only when the receiving intersections have a green signal, we changed the first term of the objective function (4.6) as follows:

$$\alpha\left(\sum_{t\in T}\sum_{i\in C_S^n} x_i^t + \sum_{t\in T}\sum_{i\in C_{SD}^n}\sum_{j\in F_i} \hat{f}_j^{t,n} x_i^t\right) \qquad \forall\, n \in S \quad (4.21)$$

The objective function of the problem minimizes the difference of occupancy and outflow at each two subsequent cells at each two subsequent time steps. Note that this minimization is equivalent to minimizing speed difference between those cells. However, this difference is not minimized for the cells that have a broken link between them. As a result, the speed difference between them may become significant. To avoid this problem, we added the following term to the objective function of each intersection at the computation node to ensure that the speed difference in cells with broken links is minimized similar to other cells:

$$\sum_{t\in T}\sum_{j\in C_{SD}^n}\sum_{i\in \Gamma_j^{-1}} \left|\left(x_i^t - \sum_p y_{ip}^t\right) - \left(\hat{x}_j^t - \sum_p \hat{y}_{jp}^t\right)\right| \qquad \forall\, n \in S \quad (4.22)$$

As a result of adding these two terms, the objective function of the problem at each computation node is as follows:

$$Z_n = \text{Max}\left(\alpha\left(\sum_{t\in T}\sum_{i\in C_S^n} x_i^t + \sum_{t\in T}\sum_{i\in C_{SD}^n}\sum_{j\in F_i}\hat{f}_j^{t,n}x_i^t\right) - (1-\alpha)\left(\sum_{t\in T}\sum_{i\in C^n\setminus\{C_S^n\cup C_{SD}^n\cup C_{RD}^n\}}\sum_{j\in\{i,\Gamma_i\}}\left|\left(x_i^t-\sum_p y_{ip}^t\right)-\left(x_j^{t+1}-\sum_k y_{jk}^{t+1}\right)\right| + \sum_{t\in T}\sum_{j\in C_{SD}^n}\sum_{i\in\Gamma_j^{-1}}\left|\left(x_i^t-\sum_p y_{ip}^t\right)-\left(\hat{x}_j^t-\sum_p \hat{y}_{jp}^t\right)\right|\right)\right) \quad \forall\, n\in S \qquad (4.23)$$

Let $\widehat{D}_i^{t,n}$ represent the entry demand in dummy source cell $i \in C_{RD}^n$ in intersection $n \in S$ at time step $t \in$ T. This demand should be equal to the flow $\hat{y}_{ui}^t$ where $u \in \Gamma_i^{-1}$ on the broken link, as shown below:

$$\widehat{D}_i^{t,n} = \hat{y}_{ui}^t \qquad \forall\, i \in C_{RD}^n, t \in T, n \in S \qquad (4.24)$$

We define $\widehat{N}_i^{t,n}$ as the available capacity of dummy sink cell $i \in C_{SD}^n$ at intersection $n \in S$. The capacity should be equal to the capacity of the cell downstream of the broken links, as shown in equation (4.25).

$$\widehat{N}_i^{t,n} = \frac{w}{v_f}(N_i^t - \hat{x}_i^t) \qquad \forall\, i \in C_{SD}^n, t \in T, n \in S \qquad (4.25)$$

Constraints (4.7) - (4.20) remain the same in the speed harmonization problem at subproblem $n \in S$. Besides, constraint (4.26) is added to avoid the infeasibility of the solution. This constraint ensures that the outflow of the segment is less than the available capacity of the receiving cell $j \in C_{SD}^n$ at the downstream computation node. The information about the available capacity can be shared between the computation nodes.

$$\sum_{i \in \Gamma_j^{-1}} y_{ij}^t \leq \widehat{N}_j^{t,n} \qquad \forall\, j \in C_{SD}^n, t \in T, n \in S \quad (4.26)$$

The information that needs to be shared between adjacent intersections is estimated during the prediction period by either the optimization program at each computation node or by implementing the solutions found by each computation node in a CTM simulation. The CTM simulation is very fast and can be performed for large transportation networks in a fraction of a second. In this paper, we implemented the solutions in a Modified CTM (MCTM). Changing the average speed in a network results in a change in the flow. In other words, it changes the number of vehicles that exit the subnetwork. We limited the outflow of a cell to the value that will be achieved with the updated speed by adding the information about the optimal outflow of a cell to the cell transmission simulation through equations (4.27) to (4.29). The value of $\sum_{j \in \Gamma_i} \hat{y}_{ij}^t$ indicates the optimal outflow of cell $i \in C$ at time $t \in T$ to all the subsequent cells according to speed changes.

$$y_{ij}^t = \min\left\{x_i^t\,, Q_i, Q_j, \frac{w}{v_f}\left(N_j - x_j^t\right), \sum_{j \in \Gamma_i} \hat{y}_{ij}^t\right\} \qquad \forall i \in C \backslash \{C_s, C_I, C_d\}, j \in \Gamma_i, t \in T \quad (4.27)$$

$$y_{ij}^t = \beta_j \min\left\{x_i^t\,, Q_i, \frac{Q_j}{\beta_j}, \frac{\frac{w}{v_f}\left(N_j - x_j^t\right)}{\beta_j}, \frac{Q_k}{\beta_k}, \frac{\frac{w}{v_f}\left(N_k - x_k^t\right)}{\beta_k}, \sum_{j \in \Gamma_i} \hat{y}_{ij}^t\right\} \qquad \forall i \in C_d,\ j \in \Gamma_i,\ k \in \Gamma_i,\ j \neq k, t \in T \quad (4.28)$$

$$y_{ij}^t = \min\left\{x_i^t\,, f_i, Q_j, \frac{w}{v_f}\left(N_j - x_j^t\right), \sum_{j \in \Gamma_i} \hat{y}_{ij}^t\right\} \qquad \forall i \in C_I, j \in \Gamma_i, t \in T \quad (4.29)$$

After finding link flows, cell occupancies are found from the flow conservation constraints. Therefore, inflows to dummy source cells $i \in C_{RD}$ as well as available cell capacity in dummy sink cells $i \in C_{SD}$ are found in the MCTM and transferred to the computation nodes.

4.3.3. DOCA-DSO framework

To account for unforeseen changes in traffic demand and capacity, create consensus among subproblems, and further reduce problem complexity, we integrated the DSO-DOCA in a model predictive control approach. The main idea is to solve the linear problem over a prediction period of τ time steps. Then, the solution for the first time step is implemented, while the time horizon is rolled forward for one time step, and the process is continued through the entire study period. Note that in optimizing the decision variables of each subproblem, information from other subproblems is required. This information is estimated at the network level by the MCTM, based on the locally optimized solutions from all subproblems. The shared information about the last time steps of the prediction period is not as accurate as the information about the first time periods. As the MPC moves forward, it optimizes and simulates the information of the last time steps several times before they are ready for implementation. This iterative process that happens as the time step moves forward, increases the accuracy of the estimations and creates consensus between different subproblems without the need to run subproblems sequentially.

Figure 4-2 shows the general procedure of DOCA-DSO approach. After initializing the network, the DSO problem will be solved for each subnetwork separately over a prediction period consisting several time steps. Each subnetwork will find the optimal average speeds and send them to the MCTM simulator. Considering the optimal average speed, feasible flows in the entire network will be found and the required information to be shared between subnetworks will be estimated over the prediction period. This information will be sent back to each subnetwork to let

them solve their own mathematical problems in a distributed way. Finally, the algorithm stops when the optimal speeds are found for the entire study period. Note that at the very first iteration, the average speed is assumed to be equal to free flow speed.

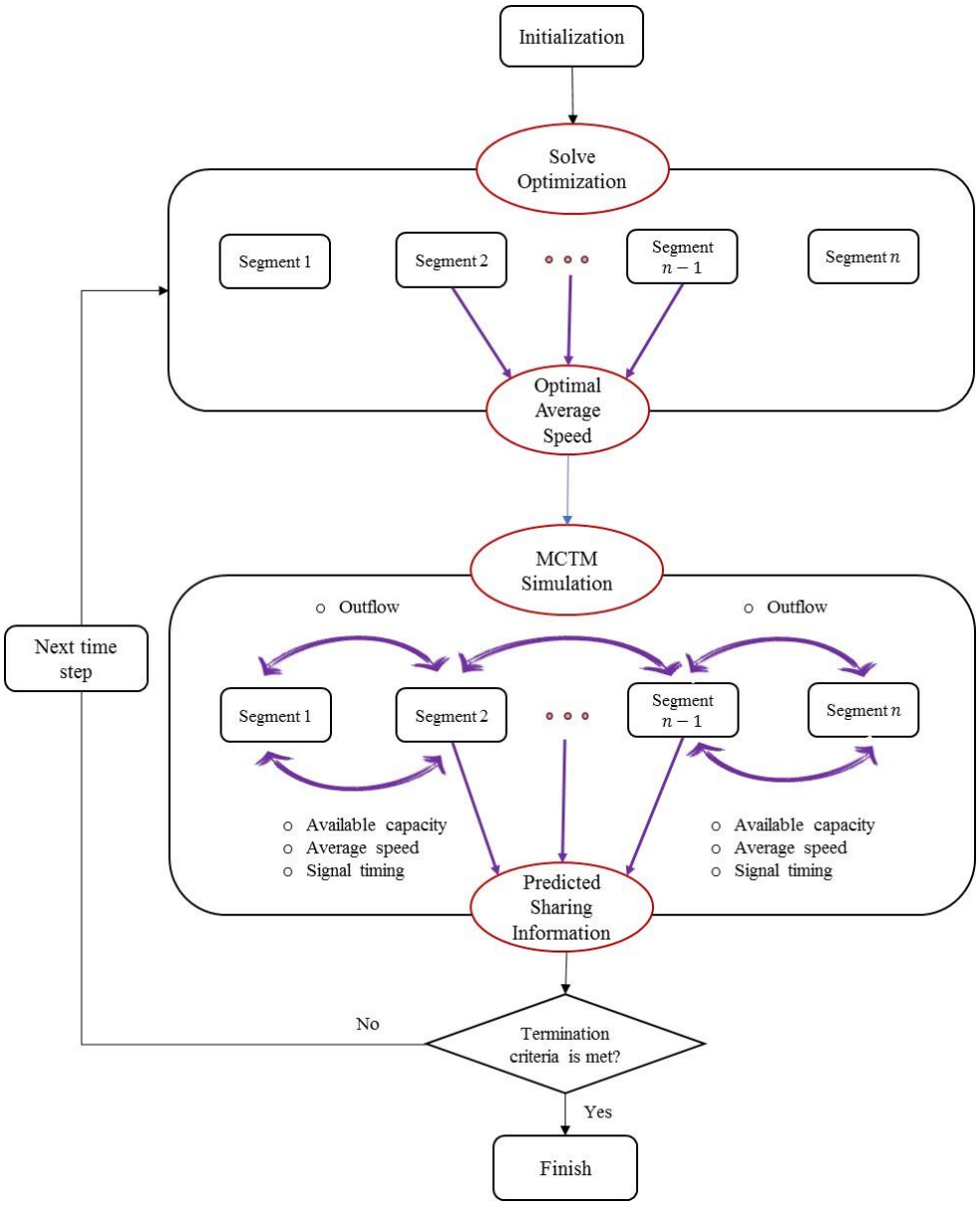


Figure 4-2: Information flow in Distributed Optimization and Coordination

Figure 4-3 shows the framework of DOCA-DSO in more details. The network is initialized based on the number of vehicles that exist in each cell. The number of vehicles in each cell is determined by aggregating vehicle location data (from connected vehicles) over CTM cells on the

network. Then the MCTM simulation will be executed for τ time steps to generate the data that needs to be shared between computation nodes. The data will be passed to the corresponding computation nodes, and the distributed optimization will be performed for each node. The average speeds are calculated for all cells and the speed for the first time steps will be saved for implementation. Then, the network state will be updated, and the time horizon will be rolled forward for one time step. The procedure repeats until the end of the study time period.

The proposed MPC approach collects the cell occupancies and finds the optimal speeds over a prediction period constituting τ time steps (in the numerical results the prediction period is 300 seconds or 50 time steps). After sharing the information among various subnetworks through the MCTM model, the optimal speeds for the next time step are implemented (transferred to the CAVs) and system state is updated. The whole process is continued until the end of the study period.

The proposed speed optimization algorithm is designed for transportation networks with 100% connected and autonomous vehicle (CAV) market penetration rate. The assumption is that the speed that is found by the algorithm is transmitted to the CAVs that are traveling inside the network and they comply with it. Note that this study does not impose a speed limit on the facility as such, the free flow speed does not change. Besides, the approach does not lead to a change in the capacity and jam density. As a result, the study will yield a new point on the existing fundamental flow density curve, and does not create a new curve.

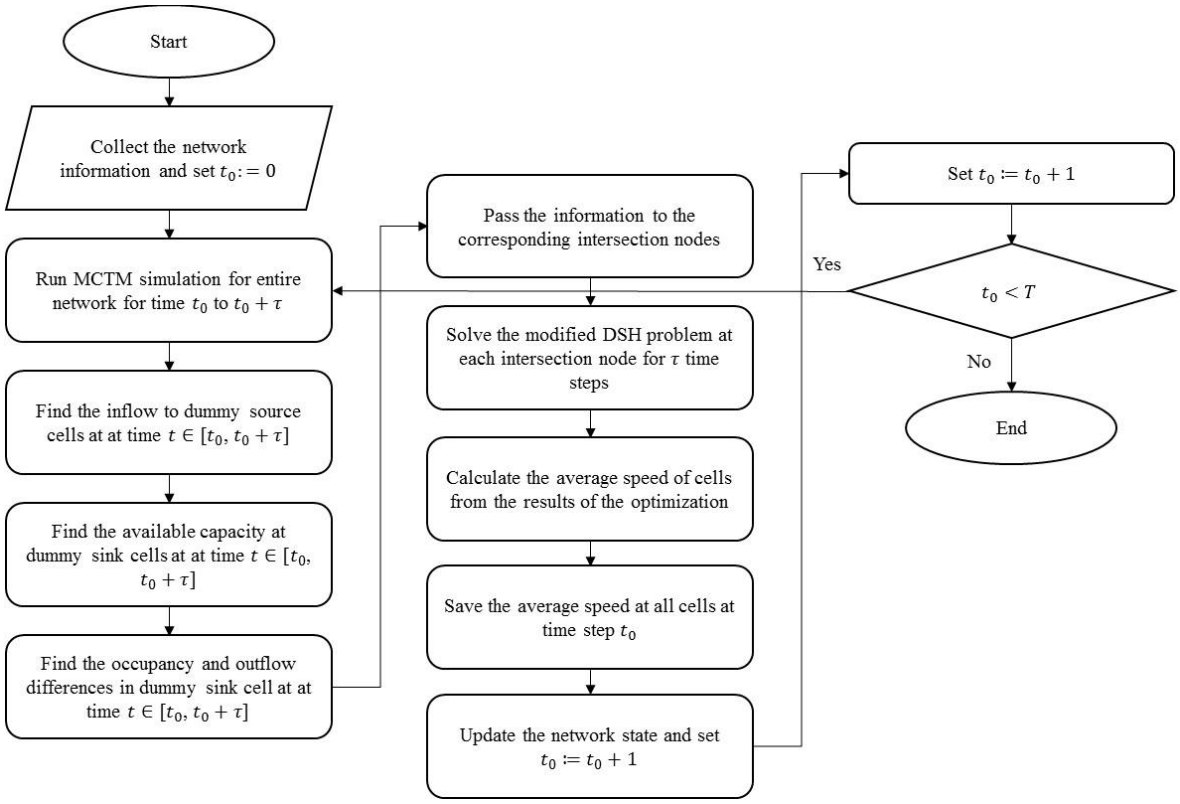


Figure 4-3: DOCA-DSO Framework

4.4. Benchmark solution

In previous chapter, we developed a linear formulation for DSH in urban street networks. The approach finds the optimal solution to the problem in medium size urban street network; however, its runtime increases with the size of the network. As such, we used this approach as the benchmark and compared the solutions of DOCA-DSO to the true optimal solutions to the problem under identical conditions.

4.5. Case study

Three example networks with eight (4×2), twenty (4×5), and forty (8×5) intersections were tested to show the performance of DOCA-DSO approach in comparison with the benchmark algorithm. The main case study network with twenty intersections is a portion of downtown Springfield, Illinois. This network contains a mix of one-way and two-way streets with a different

number of lanes and turning configurations at signalized intersections. Figure 4-4 shows the cell representation of the network with 342 cells, which is divided into 20 computation nodes.

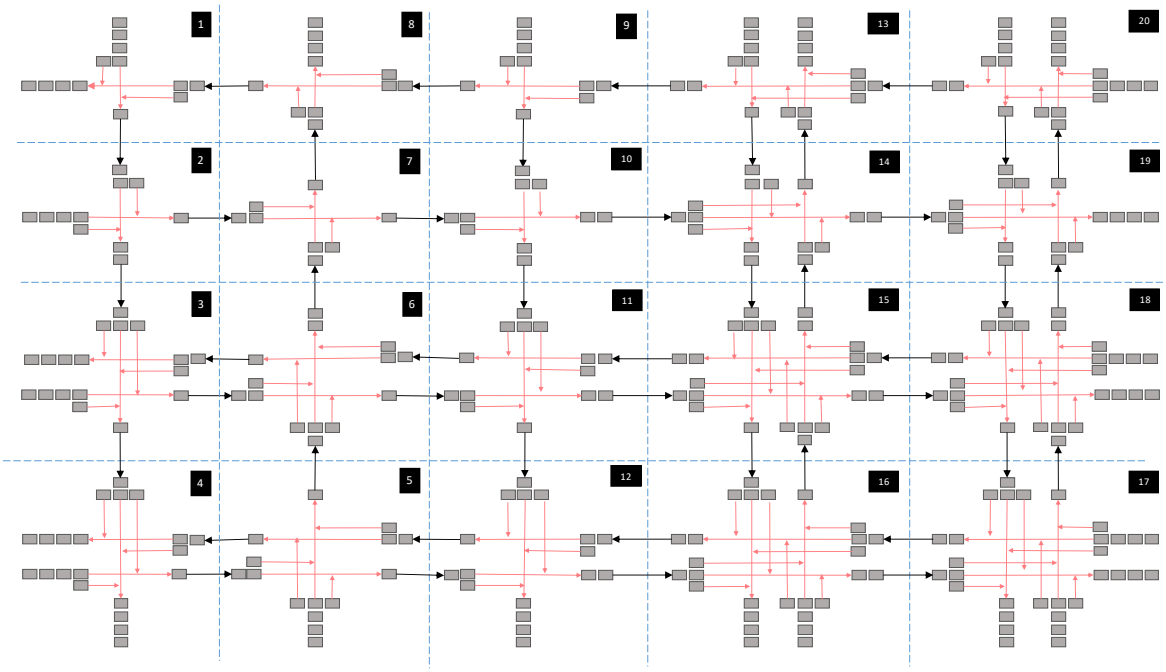


Figure 4-4: Springfield network decomposed into intersection-based subproblems

The signal timing parameters are optimized by Intelligent Dynamic Signal Timing Optimization Program (IDSTOP) developed by Hajbabaie and Benekohal (2012, 2013, 2015). The cycle times and the number of phases could vary, and the left turn movements are protected. It should be noted that signal timings are optimized before implementing the DSO strategy and there is no cooperation between the DSO strategy and signal timing optimization. The maximum speed limit is assumed to be 30 mph. Table 4-1 presents the general characteristics of the Springfield network. The total study period is 500 time steps (50 minutes), where each time step is six seconds and decisions for advisory speeds are implemented in the next time step. It is assumed that the ratio $\frac{w}{v_f}$ in the case studies is equal to 1.

Table 4-1: Characteristics of Springfield network

Link data	
Number of lanes per link	1, 2, or 3
Maximum free-flow speed (mph)	30
Link saturation flow (veh/hr/lane)	1800
Optimization period (time steps)	500
Prediction period (time steps)	50
Duration of each time step (seconds)	6
Shock wave speed (mph)	30
Number of cells	342
Cell jam density (veh/cell/ln)	12
Maximum cell saturation flow (veh/cell/ln)	3

We used four demand patterns in the case study networks as follows:

1- Symmetric undersaturated demand pattern: 500 veh/hr/ln on all entry points,

2- Symmetric saturated demand pattern: 900 veh/hr/ln on all entry points,

3- Symmetric oversaturated demand pattern: 1200 veh/hr/ln on all entrty points, and

4- Asymmetric demand pattern covering both under and oversaturated conditions. Demand profiles for east-west and north-south streets are shown in Figure 4-5.

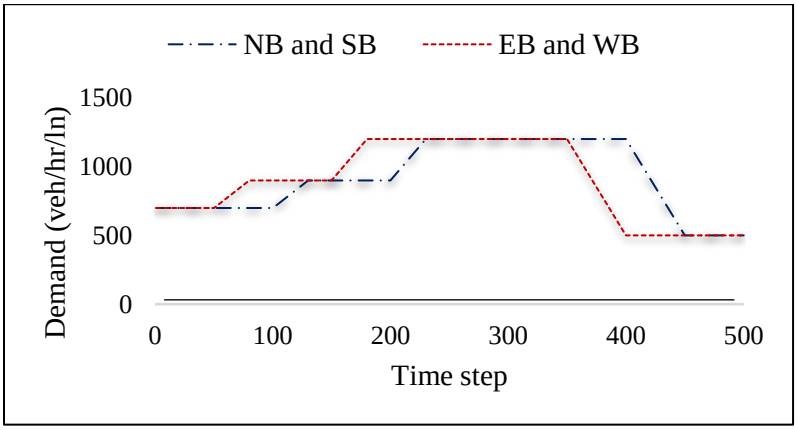
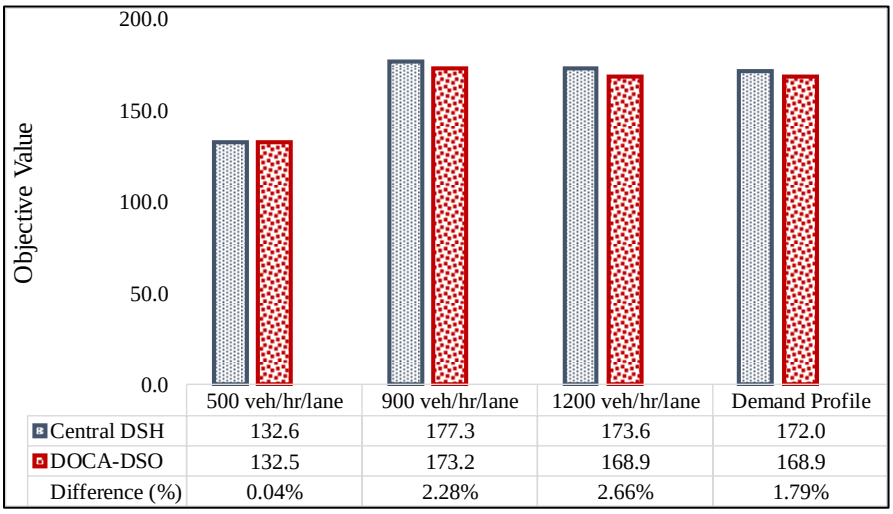


Figure 4-5: Demand profile

4.6. Results

Figure 4-6 compares the objective function values obtained by DOCA-DSO and the benchmark solutions for the network with twenty intersections. Based on the results, DOCA-DSO finds near-optimal solutions with a maximum optimality gap of 2.66%. For all the cases, the DOCA-DSO approach underestimates the optimal solution. In other words, MCTM simulation provides a feasible solution for the entire network by coordinating the decisions from each subnetwork. Therefore, the solution from DOCA-DSO is always a lower-bound for the maximization problem. The trends show that when the network is in the undersaturated condition, the gap is at the minimum value and increasing the demand increases the optimality gap. The maximum observed gap is less than 3% which is acceptable considering that demand is too high: we do not expect demands even close to 1200 veh/hr/lane in urban street networks.

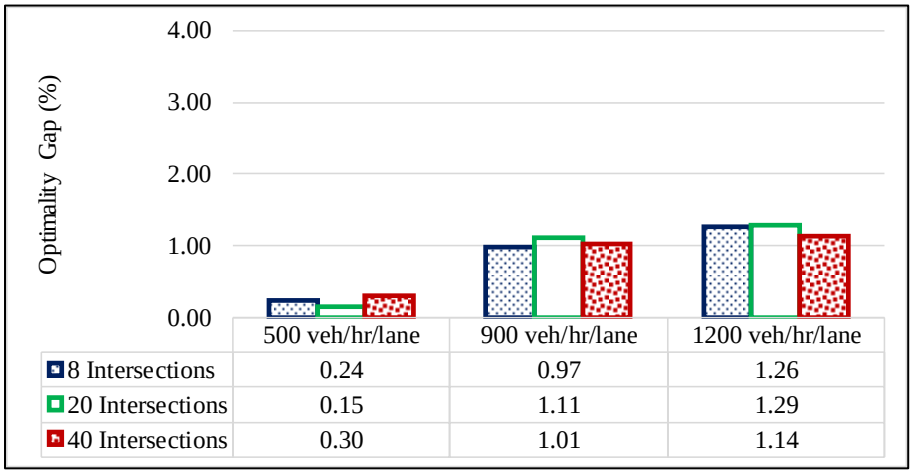


	500 veh/hr/lane	900 veh/hr/lane	1200 veh/hr/lane	Demand Profile
Central DSH	132.6	177.3	173.6	172.0
DOCA-DSO	132.5	173.2	168.9	168.9
Difference (%)	0.04%	2.28%	2.66%	1.79%

Figure 4-6: Objective values for different demand patterns in the network with 20 intersections ($\times 10^5$)

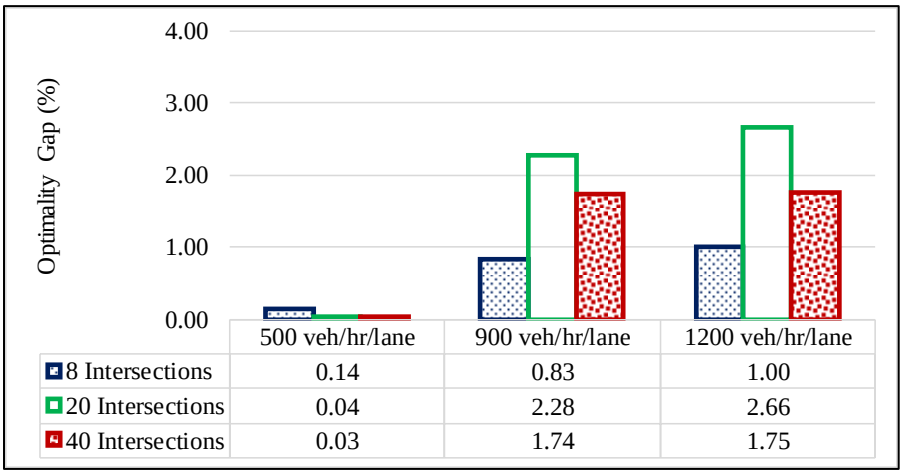
Moreover, the gaps between DOCA-DSO and the central solution were investigated for different network sizes and study period durations. Figure 4-7 shows the optimality gaps for the networks with eight, twenty, and forty intersections with a study period of 250 time steps.

Increasing the demand enlarged the optimality gap. The same trend was observed when the network size went from 8 intersections to 20; however, increasing the network size to 40 intersections did not lead to a considerable increase in the optimality gap. Figure 4-8 shows the optimality gaps for different networks sizes when study period has 500 time steps. The same trends were observed. besides, increasing the study period slightly increased the optimality gap, especially in the network with 20 intersections.



	500 veh/hr/lane	900 veh/hr/lane	1200 veh/hr/lane
8 Intersections	0.24	0.97	1.26
20 Intersections	0.15	1.11	1.29
40 Intersections	0.30	1.01	1.14

Figure 4-7: The optimality gap for networks with different sizes with study period of 250 time steps



	500 veh/hr/lane	900 veh/hr/lane	1200 veh/hr/lane
8 Intersections	0.14	0.83	1.00
20 Intersections	0.04	2.28	2.66
40 Intersections	0.03	1.74	1.75

Figure 4-8: The optimality gap for the networks with different sizes with study period of 500 time steps

Figure 4-9 shows average speed profile over time at intersection and ordinary cells in the case study network for three strategies: (a) DOCA-DSO, (b) central DSO (benchmark solution), and (c) no-speed-optimization. Average speed profiles obtained by DOCA-DSO are very similar to those obtained by the benchmark solution. In addition, the DOCA-DSO and the benchmark solutions yielded higher average speeds at intersection cells for all demand patterns. In the ordinary cells, the speed optimization reduced the average speed in the undersaturated condition for all time steps. However, in the saturated and oversaturated cases, the average speed was higher in the no-speed-optimization at the beginning of study period; however, it got lower than the optimized cases over time when the network congestion increased. In other words, the controlled speeds on the links at the beginning of the study period helped keep higher average speed over the entire study period.

In addition, the speed optimization reduced the average speed on links upstream to intersections to allow vehicles to arrive the intersection when the signal is green. Therefore, speed optimization increased the average speed at the intersections cells. This change of speed reduced the number of stops and facilitated eco-driving in the network. Moreover, the speed optimization cases provided less variation in average speed over time and created a smoother flow of traffic. When the network was loaded with higher demands, the results of optimized and not-optimized speeds got closer to each other because the network was more congested and there was less possibility for speed variation. Therefore, there was not enough room to improve the network performance by controlling the speed of vehicles.

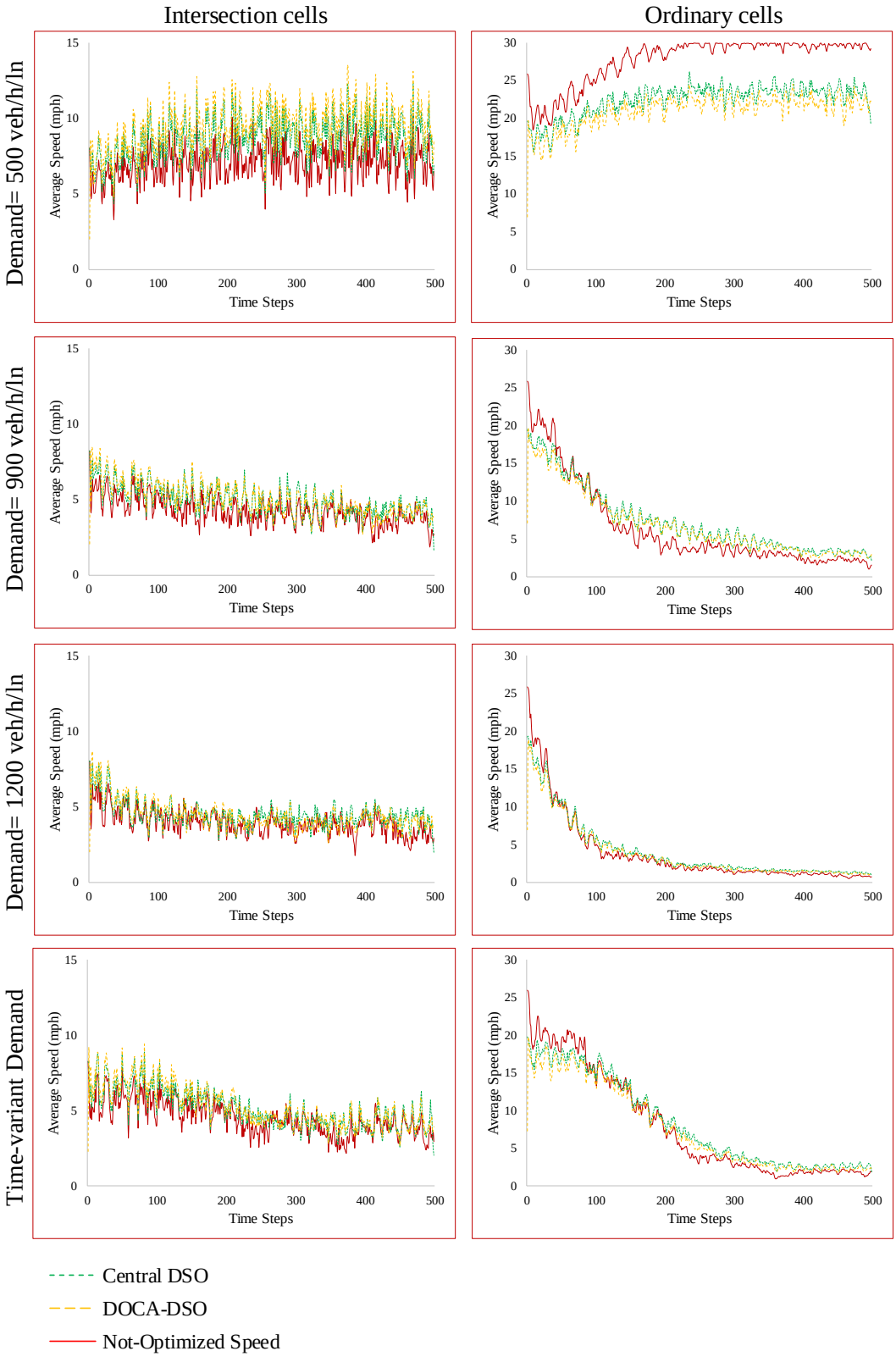


Figure 4-9: The comparison of speed variations on the network with different demand patterns

The runtimes are found by solving the algorithm in the Java environment on a PC with a Core i7 CPU and 24 GB of memory. Figure 4-10 shows the average runtime for each subproblem. The runtimes range between 0.07 to 0.53 seconds for all subproblems across the entire study period under all tested demand patterns. This range indicates that the algorithm is realtime as the six-second duration of each time step provides enough time for communication and exchange of information between sub-problems. Moreover, Figure 4-10 shows that increasing the demand increases the runtime only slightly.

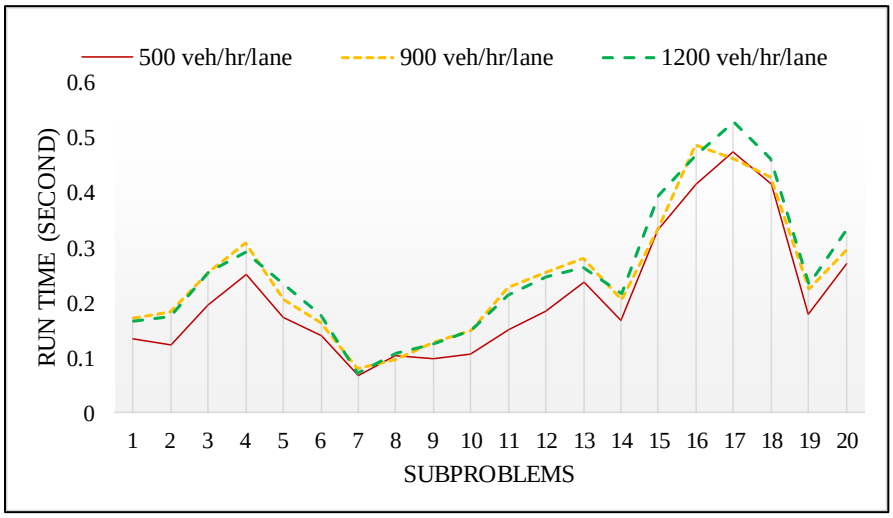


Figure 4-10: Average runtimes of subproblems

Table 4-2 illustrates the performance measures in the Springfield network for different demand patterns when α is equal to 0.95. Besides, the results for the benchmark solution and DOCA-DSO are compared to a no-speed-optimization strategy. In symmetric undersaturated demand pattern (500 veh/hr/lane), central and distributed DSO solutions yielded similar travel time, average speed, speed variation, and the number of completed trips. The travel time in DOCA-DSO was 0.4% higher than the no-speed-optimization case, and the network average speed was 0.2% less. Moreover, the total number of completed trips were the same as the no-speed-optimization case. These results were as expected since speed optimization cannot improve

network performance in low demand conditions. Vehicles mostly travel at the free flow speed in undersaturated conditions and since speed optimization does not assigned speeds faster than the free flow speed, it does not improve traffic operations. Having said that, speed optimization reduces speed variation by 13.4%.

In the symmetric saturated demand pattern (900 veh/hr/lane), the differences between the benchmark solution and DOCA-DSO were higher than the undersaturated demand pattern. Both speed optimization cases performed better in comparison to the no-speed-optimization case. DOCA-DSO reduced travel time and speed variation by 14.8% and 9.7%, respectively. Moreover, it increased the network average speed and throughput by 28.0% and 8.4%, respectively.

In the symmetric oversaturated demand pattern (1200 veh/hr/ln in all entry points), the differences between the benchmark solution and DOCA-DSO were slightly higher than the saturated demand pattern. DOCA-DSO reduced the travel time and speed variation by 6.0% and 11.1%, respectively in comparison to the no-speed-optimization case. In addition, both the network average speed and network throughput were increased by 15.4% and 9.5% respectively. Similar trends were observed for the asymmetric demand pattern.

The same pattern can be seen in the benchmark (i.e. centralized) solutions. As expected, the benchmark solutions yielded a more efficient network performance. Based on the results of both central and distributed approaches, speed optimization provided lower travel times and more completed trips at the saturation and oversaturation demand level than it did at the undersaturated demand levels. Vehicles moved efficiently in the network in undersaturated conditions (i.e., 500 veh/hr/ln). Therefore, the DSO did not lead to improvement in network performance measures, while it still reduced speed variations. On the other hand, there is more variations in speeds and more room for improvement in network performance in the saturated and oversaturated demand

levels (i.e., 900 and 1200 veh/hr/ln). As such, DSO yielded the highest reduction in total travel time and highest increase in the number of completed trips in these cases.

Table 4-2: Mobility performance of the network

Demand (veh/hr/lane)	Mobility performance	DOCA-DSO		Central DSO		Not-optimized speed
		Value	Difference (%)	Value	Difference (%)	
500	Travel time (hour)	691.8	0.4	689.3	0.0	689.3
	Throughput	12197	0.0	12197	0.0	12197
	Speed variations (mph × 10^5)	25.9	-13.4	26.4	-11.6	29.9
	Network average speed (mph)	16.0	-0.2	16.1	0.1	16.1
	Intersection average speed (mph)	9.3	33.4	8.5	23.1	6.9
	link average speed (mph)	20.7	-23.7	22.1	-18.5	27.2
900	Travel time (hour)	2943.0	-14.8	2851.0	-17.5	3455.5
	Throughput	15977	8.4	16286	10.5	14740
	Speed variations (mph × 10^5)	23.5	-9.7	24.0	-8.0	26.1
	Network average speed (mph)	5.2	28.0	5.4	33.9	4.0
	Intersection average speed (mph)	4.9	16.6	4.9	15.6	4.2
	link average speed (mph)	5.3	31.6	5.6	40.4	4.0
1200	Travel time (hour)	5969.7	-6.0	5769.9	-9.1	6348.1
	Throughput	15509	9.5	16062	13.4	14170
	Speed variations (mph × 10^5)	22.2	-11.1	23.0	-7.9	25.0
	Network average speed (mph)	2.5	15.4	2.6	23.4	2.1
	Intersection average speed (mph)	4.4	11.5	4.4	12.8	3.9
	link average speed (mph)	2.2	16.6	2.4	25.7	1.9
Demand profile	Travel time (hour)	3297.4	-9.3	3176.9	-12.7	3637.5
	Throughput	15662	8.5	16089	11.4	14436
	Speed variations (mph × 10^5)	23.2	-12.1	23.7	-10.3	26.4
	Network average speed (mph)	4.5	18.7	4.8	25.8	3.8
	Intersection average speed (mph)	5.0	14.4	4.9	11.4	4.4
	link average speed (mph)	4.4	20.1	4.8	30.5	3.6

Table 4-3 shows the total runtime for demand of 900 veh/hour/lane for all tested networks and study periods. When the study period increased from 250 to 500 time steps, the runtime for DOCA-DSO increased proportionally and was almost doubled. In addition, increasing the size of the network for a fix study period did not increase DOCA-DSO's runtime because the subproblems were solved in parallel. In contrast, increasing the size of the network or the study period yielded a significant increase in the runtimes of the Central DSO.

Table 4-3: Total runtimes (second) for different sizes and study periods when demand is 900 veh/hour/lane

Network	Time steps	DOCA-DSO	Central DSO	Difference (%)
8 Intersections	250	123.1	120.4	2.3
	500	252.0	643.3	-60.8
20 Intersections	250	125.4	386.5	-67.6
	500	272.0	1601.9	-83.0
40 Intersections	250	132.1	1679.7	-92.1
	500	279.5	12722.9	-97.8

We have also analyzed the average runtime when the subproblem consists one, two, and three intersections. Table 4-4 shows that increasing the subproblem size and prediction period length increase the runtime for each subproblem. The single-intersection subnetwork can provide the real-time results for all predication periods because the average runtime is less than 6 seconds (implementation period).

Table 4-4: Comparing the average runtimes (seconds) for one, two, and three intersection-based subproblems

Prediction period (seconds)	Subproblem size		
	1 Intersection	2 Intersections	3 Intersections
50	0.51	1.64	2.07
100	1.31	3.87	7.08
150	2.23	8.22	13.57

4.7. Conclusion

This chapter developed a scalable and real-time algorithm to find near-optimal solutions to dynamic speed optimization in urban street networks. DOCA-DSO has two main components: distributed optimization and distributed coordination. The distributed optimization component decomposes the single network-level mathematical program to several intersection level mathematical programs and consequently reduces the problem complexity significantly. The distributed coordination component creates effective coordination between intersection level mathematical programs by exchanging information on a) the outflow of cells at boundaries of a computation node, (b) the available capacity of the receiving cells, (c) the average speed in the receiving cells, and (d) signal timing status at intersection cells. The problem formulation was modified for each computation node to accommodate incorporation of the information.

We tested the developed methodology in networks with eight, twenty, and forty intersections under four different demand patterns. The results indicated that DOCA-DSO provided solutions with at most 2.7% optimality gap. Furthermore, DOCA-DSO and the true optimal solutions yielded similar spatial and temporal speed distribution over the network. The developed methodology reduced the travel time by up to 14.8% and speed variation by 9.7% – 13.4% over different demand patterns compared to a no-speed-harmonization strategy. Moreover, the average speed and the total number of completed trips were increased by up to 28.0% and 9.5% respectively in a network with 20 intersections and 500 study periods. The runtime analysis showed that controllers could solve the problem in real-time as it took at most 0.53 seconds for a subproblem to solve the corresponding problem.

The results of this study indicate significant improvement in network performance as a result of speed optimization. However, the capacity drop cannot be captured by the CTM. As such,

the results need to be interpreted cautiously and we suggest further study for incorporating capacity drop in future speed harmonization/optimization studies. It would be interesting to integrated signal timing optimization and speed optimization in the future studies to arrive at more significant improvements in network performance. This study assumed full market penetration rate of CAVs. Research is needed to show how the algorithm performs under various market penetration rates and which rate is necessary to observe the benefits of speed harmonization. Moreover, this study assumed that vehicle positions are accurate, and the communications are instantaneous. Further research on the effects of error in vehicle positions and communication latency is needed.

CHAPTER 5: NETWORK-LEVEL COORDINATED SPEED OPTIMIZATION AND TRAFFIC LIGHT CONTROL FOR CONNECTED AND AUTOMATED VEHICLES

As shown in previous chapters, advisory speed systems can adjust automated vehicles' speeds and consequently their arrival time to signalized intersections to reduce the number of stops and unnecessary acceleration/deceleration. Preventing stop-and-go conditions reduces travel delay, fuel consumption, and yields a more efficient network performance (Al-Khalili, 1985). Traffic operations can be further improved by coordinated signal timing and speed optimization. In other words, signal timing parameters and vehicle speeds can be optimized jointly to plan the arrival of vehicles to signalized intersections more accurately to utilize green durations more efficiently.

In this chapter, we develop a mathematical program for Coordinated Signal timing and Speed Optimization (CSSO) in urban-street networks assuming that all vehicles are connected and automated, and intersection controllers can communicate to each other. It is assumed that connected and automated vehicles receive the assigned speeds through communication with intersection controllers and accept them. The mathematical program is based on the Cell Transmission Model (CTM) (Carlos F. Daganzo, 1994; Daganzo, 1995) network loading concept. We develop a Distributed Optimization and Coordination Algorithm (DOCA) to find near-optimal solutions to the coordinated signal timing and speed optimization problem in real-time in networks of various sizes. To reduce the computational complexity, the proposed methodology decomposes the network-level CSSO problem into several intersection-level sub-problems by relaxing the constraints that represent interrelationship between intersections. Effective coordination among sub-problems pushes their solutions towards global optimality (Al Islam and Hajbabaie, 2017; Camponogara and De Oliveira, 2009; Camponogara and Scherer, 2011; Mehrabipour and

Hajbabaie, 2017; Mohebifard and Hajbabaie, 2018a; Shen and Kong, 2009). The coordination is achieved by exchanging information among sub-problems, re-enforcing the re-introduced relaxed constraints and incorporating the information in them. The complexity of the problem is further reduced by using Model Predictive Control (MPC) and solving the problem over a planning horizon rather than the entire study period. Moreover, the proposed objective function in this study takes into account the trade-off between maximizing the intersection throughput and minimizing the spatial and temporal speed variations over the entire network.

5.1. Research gaps and contributions

The existing studies show the effectiveness of cooperative signal timing and speed optimization in managing traffic congestion in either an isolated intersection or at most four intersections due to the computational complexities that were associated with the cooperative problem. The existing algorithms are complex due to their microscopic nature, where the trajectory of each connected and automated vehicle is controlled. The existing approaches are enumerative or centralized heuristic/metaheuristic techniques that limit their scalability, real-time application, and optimality. This study addresses the knowledge gap and enables studying the effects of cooperative signal timing and speed optimization in large transportation networks. We formulate the problem using macroscopic network loading concept and develop a scalable solution technique that can find near-optimal solutions in real-time. The solution technique distributes the complex network-level signal timing and speed optimization problem into several intersection-level sub-problems and implements them in a model predictive controller. As such, it significantly reduces computational complexity and finds solutions in real-time.

5.2. Problem formulation

The problem formulation utilizes the CTM network loading concept that is introduced by Daganzo (Carlos F Daganzo, 1994; Daganzo, 1995) and used in other traffic control studies (Mehrabipour, 2018; Mehrabipour et al., 2019; Mohebifard et al., 2019; Mohebifard and Hajbabaie, 2019). CTM divides a network link into homogeneous segments and discretizes the study period to short time steps. The cell length is the distance that a vehicle can travel with the free flow speed during a time step. Let C, C_O and C_S respectively denote the sets of all cells, source cells, and sink cells in the network. In addition, we define C_S^N and C_S^I as the sets of network sink cells and internal intersection sink cells, respectively. The sets of predecessors $P(i)$ and successors $S(i)$ cells are defined for each cell $i \in C$. Moreover, we defined T as the set of discrete time steps.

Table 5-1 summarizes the notations used in this study.

The decision variables are (a) the status of traffic signal g_i^t (one if green, zero otherwise) on intersection cell $i \in C_I$ at time step $t \in T$, (b) space mean speed v_i^t on cell $i \in C$ at time step $t \in T$, and (c) the number of vehicles y_{ij}^t flowing from cell $i \in C$ to successor cell $j \in S(i)$ at time step $t \in T$. The state variable of the system is the number of vehicles x_i^t in cell $i \in C$ at time step $t \in T$, which is equivalent to cell occupancy assuming that each cell is one length unit long. The space mean speed is defined as the ratio of the outgoing flow y_{ij}^t from a cell to its occupancy x_i^t as shown in equation (5.1) (Tajalli and Hajbabaie, 2018c). Note that the space mean speed is equal to the free flow speed when a cell is empty.

$$v_i^t = \begin{cases} \frac{\sum_{j \in \Gamma_i} y_{ij}^t}{x_i^t} v_f & x_i^t > 0 \\ v_f & x_i^t = 0 \end{cases} \qquad \forall\, i \in C, t \in T \quad (5.1)$$

Table 5-1: Definition of sets, decision variables, and parameters

Sets:	
T	set of all time steps
C	set of all network cells
C_O	set of all source cells
C_S	set of all sink cells
C_S^N	set of all network sink cells
C_S^I	set of internal intersection sink cells
C_I	set of all intersection cells
C_{OD}	set of all dummy source cells
C_{SD}	set of all dummy sink cells
C_R	set of all right turn movement cells
C_T	set of all through movement cells
C_L	set of all left turn movement cells
C_F	set of all conflicting movement pairs at an intersection
C_{RT}	set of all concurrent adjacent right and through movements
$P(i)$	set of all cells predecessors
$S(i)$	sets of all cells successors
State Variable:	
x_i^t	number of vehicles in cell $i \in C$ at time step $t \in T$
Decision Variables:	
g_i^t	a binary variable for signal status at cell $i \in C_I$ at $t \in T$
y_{ij}^t	number of vehicles flowing from cell $i \in C$ to downstream cell $j \in S(i)$ at time step $t \in T$
v_i^t	space mean speed in cell $i \in C$ at time step $t \in T$
Other Variables:	
f_i^t	variable saturation flow rate in cell $i \in C_I$ at $t \in T$
z_{ij}^t	auxiliary variable
u_{ij}^t	auxiliary variable
Parameters:	
d_i^t	entry demand on source cell $i \in C_O$ at time step $t \in T$
F_i	saturation flow rate of cell $i \in C$
N_j	capacity of cell $j \in C$
r_i^t	turning proportion at intersection cell $i \in C_I$ at $t \in T$
f'	the star-up lost time reduction factor
$G_{\max}$	maximum green time
$G_{\min}$	minimum green time
$\hat{x}_j^t$	predicted occupancy of cell $j \in C$ at $t \in T$ from CTM
v_f	free flow speed

The objective function of the CSSO problem has two terms that aim at maximizing the cumulative throughput and minimizing speed variations. Past research has shown the effectiveness of throughput maximization in congestion management especially when the demand level is high. The cumulative throughput maximization accounts for both network and intersection throughputs:

a larger weight (M) is assigned to the network sink cells $i \in C_S^N$ to prioritize the number of completed trips, while a smaller weight (m) is assigned to the number of vehicles exiting each intersection through the internal sink cells $i \in C_S^I$. In addition, the difference of space mean speeds between two adjacent cells at two consecutive time steps as well as the space mean speed differences in one cell at two consecutive time steps are minimized in the second term of the objective function. This term prioritizes the smooth movement of vehicles in the network. There is a trade-off between these two terms, as smooth vehicle speed does not necessarily yield a higher throughput value. Therefore, we defined γ (vehicle/mph) as a weight factor to define the desired emphasis on each term of the objective function and convert the units of the two terms, see objective function (5.2).

$$\text{Max} \left[\sum_{t \in T} \sum_{i \in C_S^N} M x_i^t + \sum_{t \in T} \sum_{i \in C_S^I} m x_i^t - \gamma \sum_{t \in T} \sum_{i \in C \setminus \{C_s\}} \sum_{j \in \{\text{i},\ S(i)\}} \left| v_i^t - v_j^{t+1} \right| \right] \tag{5.2}$$

The defined objective function is nonlinear due to the inherent nonlinearity of the space mean speed. In Chapter 3:, we showed that the speed harmonization term in the objective function is equivalent to the difference of occupancy and flow between two subsequent cells and two subsequent time steps, see objective function (5.3). This term is linear as such, significantly reduces problem complexity. Since all terms of the objective function have a unit of vehicles, we use a unit-less factor α to assign priority to each term.

$$\begin{aligned} \text{Max} \Bigg[\alpha \Bigg(\sum_{t \in T} \sum_{i \in C_S^N} M x_i^t + \sum_{t \in T} \sum_{i \in C_S^I} m x_i^t \Bigg) - (1 \\ - \alpha) \sum_{t \in T} \sum_{i \in C \setminus C_s} \sum_{j \in \{\text{i}, S(i)\}} \left| \left(x_i^t - \sum_p y_{ip}^t \right) - \left(x_j^{t+1} - \sum_k y_{jk}^{t+1} \right) \right| \Bigg] \end{aligned} \tag{5.3}$$

Equations (5.4) to (5.6) represent the state transition of the system. Let D_i^t denote the entry demand level on source cell $i \in C_O$ at time step $t \in T$. Constraints (5.4) to (5.6) ensure the flow conservation in source cells $i \in C_O$, sink cells $i \in C_S$, and ordinary cells $i \in C \setminus \{C_S, C_O\}$, respectively. The number of vehicles in a cell in the next time step is equal to the number of vehicles that are in that cell in the current time step, minus those who are leaving, plus those who are entering during the current time step.

$$x_o^{t+1} = D_o^t + x_o^t - \sum_{j \in S_o} y_{oj}^t \qquad \forall\, o \in C_O, t \in T \qquad (5.4)$$

$$x_s^{t+1} = x_s^t + \sum_{i \in P_s} y_{is}^t \qquad \forall\, s \in C_S, t \in T \qquad (5.5)$$

$$x_i^{t+1} = x_i^t + \sum_{u \in P_i} y_{ui}^t - \sum_{j \in S(i)} y_{ij}^t \qquad \forall\, i \in C \setminus \{C_S \cup C_O\}, t \in T \qquad (5.6)$$

The number of vehicles $\sum_{j \in S(i)} y_{ij}^t$ moving between cell $i \in C \setminus C_s$ and all successor cells $j \in S(i)$ at time step $t \in T$ must be less than or equal to the number of vehicles x_i^t that exists in cell $i \in C$ at time $t \in T$, as follows:

$$\sum_{j \in S(i)} y_{ij}^t \leq x_i^t \qquad \forall\, i \in C, t \in T \qquad (5.7)$$

Let us define F_i as the saturation flow rate of cell $i \in C$. Constraints (5.8) and (5.9) limit the number of vehicles flowing from a cell to its successor to the saturation flow rates of the sending and receiving cells, respectively.

$$\sum_{j \in S(i)} y_{ij}^t \leq F_i \qquad \forall\, i \in C \setminus C_S, t \in T \qquad (5.8)$$

$$\sum_{i \in P(j)} y_{ij}^t \leq F_j \qquad \forall\, j \in C \backslash C_O, t \in T \quad (5.9)$$

Let N_j denote the maximum number of vehicles that cell $j \in C$ can accommodate. Constraints (5.10) ensure that the number of vehicles flowing between two cells is less than the available capacity of the receiving cell.

$$\sum_{i \in P(j)} y_{ij}^t \leq N_j - x_j^t \qquad \forall\, j \in C \backslash C_O, t \in T \quad (5.10)$$

Constraints (5.11) ensure that turning percentages are equal to the pre-defined turning proportions (r_j^t). Let C_I be the set of intersection cells from which, right turning, through, and left turning movements are completed. The number of vehicles travelling to each intersection cell $j \in C_I$ from cell $i \in P(j)$ should be equal to the product of corresponding turning proportion r_j^t and the total number of vehicles $\sum_{k \in S(i)} y_{ik}^t$ leaving cell $i \in P(j)$ to all immediately downstream intersection cells at time step $t \in T$.

$$y_{ij}^t = r_j \sum_{k \in S(i)} y_{ik}^t \qquad \forall j \in C_I, i \in P(j), t \in T \quad (5.11)$$

Signal controllers at each intersection find the optimal signal timing parameters (i.e. green time duration and phase sequences) in cooperation with vehicles' average speed. Optimizing the sequence of phases yields a more efficient network performance. To optimize signal timing parameters, let the binary variable g_i^t take on the value of one when the signal is green, otherwise zero. Constraints (5.12) adjusts the saturation flow rate (f_i^t) of an intersection cell based on its signal status, and constraints (5.13) ensure that the flow of vehicles leaving intersections cells cannot exceed the adjusted saturation flow rate (f_i^t).

$$f_i^t = g_i^t F_i \qquad \forall\, i \in C_I, t \in T \quad (5.12)$$

$$\sum_{j \in S(i)} y_{ij}^t \leq f_i^t \qquad \forall\, i \in C_I, t \in T \quad (5.13)$$

Constraints (5.14) reduce the saturation flow rate of the intersection when a green signal is initiated. Parameter f' is defined as the saturation flow rate reduction factor due to start-up lost time. A proper joint optimization of signal timing parameters and speeds should regulate the arrival of vehicles at intersections such that the signal has already turned green and the queue is cleared. Therefore, speed optimization reduces the impacts of start-up lost time on traffic operations.

$$\sum_{j \in S(i)} y_{ij}^{t+1} \leq F_i - F_i f'(g_i^{t+1} - g_i^t) \qquad \forall\, i \in C_I,\ t \in T \quad (5.14)$$

We define C_F as the set of all conflicting movements at an intersection. Constraints (5.15) ensure that only one of the two non-conflicting movements $(i, j) \in C_F$ receives the green time at time step $t \in T$. It should be noted that all turning movements are assumed to be protected.

$$g_i^t + g_j^t \leq 1 \qquad \forall\, (i,j) \in C_F, t \in T \quad (5.15)$$

Constraints (5.16) limit the green duration at intersection cell $i \in C_I$ to a maximum green duration G_{max}^i. Constraints (5.17) ensure a minimum green duration G_{min}^i for each signal.

$$\sum_{j=t}^{t+G_{max}^i+1} g_j^t \leq G_{max}^i \qquad \forall\, i \in C_I, t < T - G_{max}^i \quad (5.16)$$

$$\sum_{j=t+1}^{t+G_{min}^i} g_j^t \geq (g_i^{t+1} - g_i^t) G_{min}^i \qquad \forall\, i \in C_I, t \leq T - G_{min}^i \quad (5.17)$$

We define C_{RT} as the set of concurrent adjacent right and through movements. Constraints (5.18) ensure that the adjacent right turn and through movements have the same signal timing, either red or green.

$$g_i^t = g_j^t, \qquad \forall\, (i,j) \in C_{RT}, t \in T \quad (5.18)$$

5.3. Methodology

5.3.1. Distributed optimization and coordination

The presented formulation has mixed-integer decision variables and will not scale well with the size of the network when a centralized algorithm is utilized to solve it. We utilized a distributed optimization and coordination algorithm, described in Chapter 4:, that can handle the computational complexity of the problem and find near-optimal solutions in real-time. Figure 5-1 shows a conceptual representation of the proposed methodology. The distributed optimization decomposes the network-level problem into several stand-alone intersection-level sub-problems and solves them in parallel. The decomposition is achieved by identifying the constraints that represent interrelationships between sub-problems and relaxing them. Distributed coordination pushes solutions towards global optimality by exchanging information among sub-problems that share a relaxed constraint, and re-enforcing the relaxed constraints in sub-problems by incorporating the information that is received from other sub-problems in them. This method is incorporated in MPC to account the dynamic nature of the problem and further reduce the computational complexity.

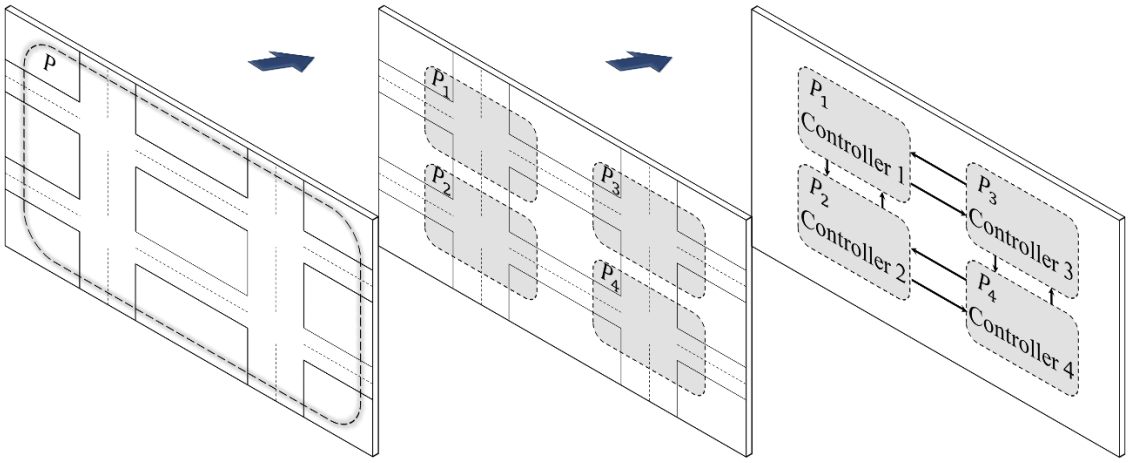


Figure 5-1: The intersection-level decomposition.

5.3.2. Model predictive control

In addition, a model predictive control (MPC) approach is implemented to increase the computational efficiency of the algorithm. First, the state variables are predicted for near future. Then, the required information between sub-networks is shared through the distributed coordination. After receiving the information, each controller solves the corresponding sub-problem and finds the optimal signal timing and speeds for each sub-network. This process is repeated until the study period is finished. Note that the length of prediction horizon is selected long enough to ensure the feasibility and stability of the system with MPC. The overall framework of the distributed optimization and coordination algorithm for coordinated signal timing and speed optimization using model predictive control is illustrated in Figure 5-2.

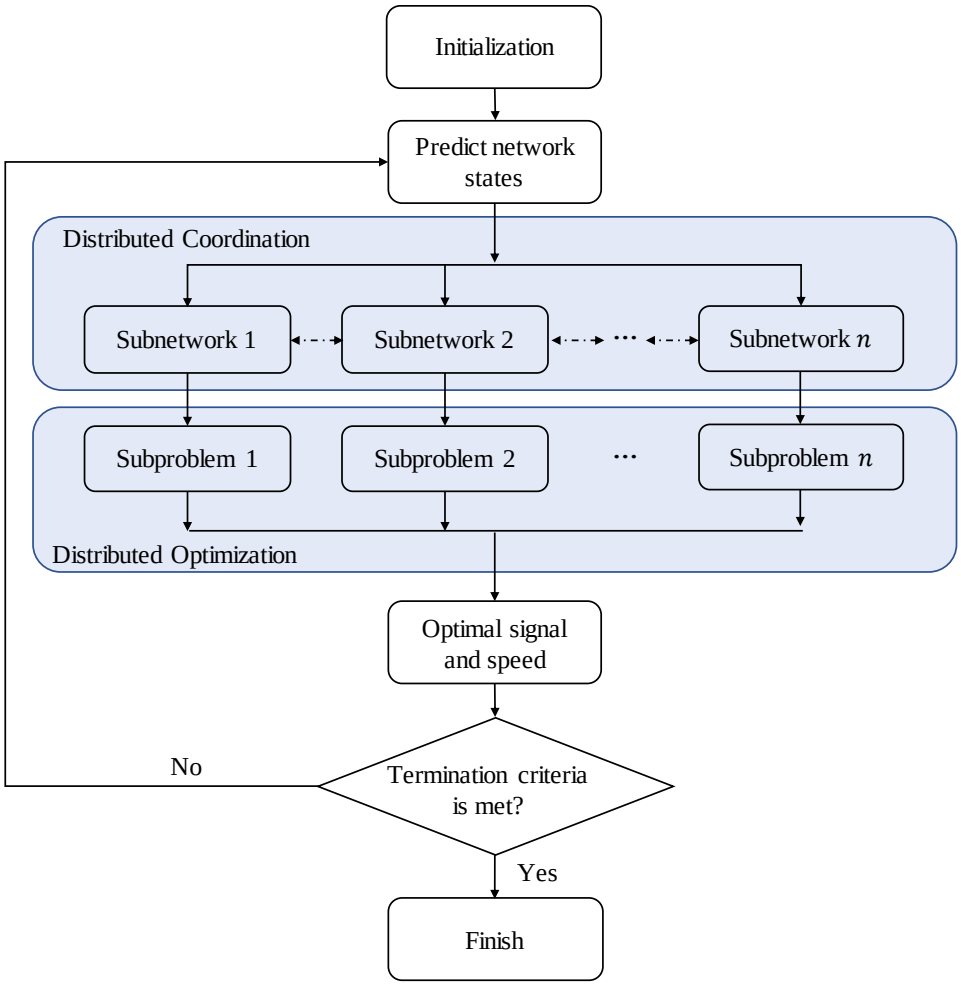


Figure 5-2: DOCA-CSSO general framework with MPC

5.3.3. Accounting for the trade-offs between traffic operations and speed variations

CSSO is a multi-objective program with a trade-off between maximizing the cumulative intersection throughputs and minimizing speed variance. Thus, there is a set of Pareto optimal solutions for CSSO problem. Hwang and Masud (Hwang and Masud, 2012) classified the solution techniques for accounting these trade-offs into three categories including Priori, Interactive, and Posteriori methods. In Priori method, the optimal solution is found in a way that satisfies the preference of decision makers. The Interactive method allows decision makers to search iteratively for the most preferred solution by receiving feedbacks. However, decision makers might not see

the entire pareto. The Posterior method finds a set of pareto optimal solutions and let decision makers select among them. We utilized the Posteriori method because it provides the opportunity to observe the tradeoffs between different terms of the objective function. Based on the weighting method in (Mavrotas, 2009), we multiply the intersection throughput maximization term by α and speed variation minimization term by $(1-\alpha)$ in objective function (5.3). Therefore, an α of zero corresponds to minimizing speed variation and an α of one corresponds to intersection throughput maximization.

5.3.4. Benchmark

The benchmark solutions are found using the Benders decomposition technique (Geoffrion, 1972). Benders decomposition technique finds an upper and a lower bound to the objective function of the problem through an iterative process. It is proven that the gap will be reduced to zero (i.e., the exact solution is found) after a finite number of iterations. Benders technique decomposes the coordinated signal timing and speed optimization problem to master and primal sub-problems. The signal timing decision variables are found by the master problem and the average speeds are optimized by the primal problem based on the fixed values of signal timing variables. The algorithm iterates between the two sub-problems until the convergence criteria is met. For more details see Mohebifard and Hajbabaie (Mohebifard and Hajbabaie, 2018a).

5.3.5. Implementation of the algorithm in Vissim

The proposed algorithm is implemented in Vissim (PTV Group, 2015) to allow a more accurate evaluation of the results in a more realistic simulated environment. We used COM interface to provide required communications between vehicles and signal controllers. Figure 5-3

shows the general implementation procedure. In general, the network in Vissim is divided into several segments to match the cells in the CTM. Vehicle location data is passed from Vissim to DOCA through the COM interface as the initial state of the system. The optimization problem is solved, and optimized signal timing and speed variables are sent back to Vissim.

DOCA finds the average speed of a cell in CTM. This speed is assigned to the vehicles travelling in the corresponding segment of the network as their desired speed. Note that the optimal speed from the CTM should be calibrated for Vissim to ensure that both models have similar outflows in the road segment. In other words, the CTM is a macroscopic first-order model and does not consider the interaction between vehicles. Therefore, vehicles take more time to achieve the speed in Vissim than the CTM (instantaneous). As a result, we quantified this difference and accounted for it when transmitting speeds to Vissim.

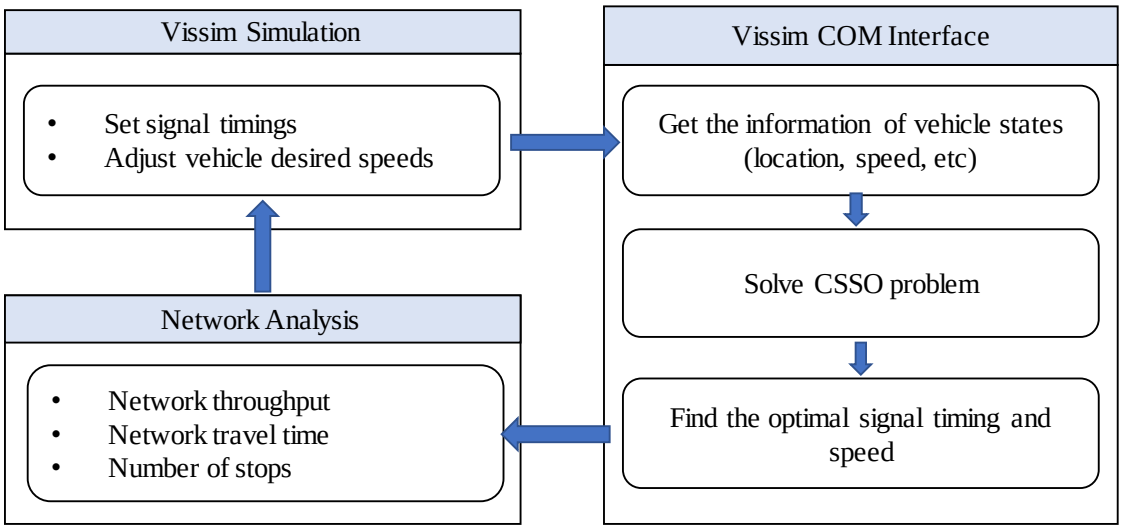


Figure 5-3: Vissim COM interface to apply the optimized signal timings and speeds

5.4. Case study

The case study network is a portion of downtown Springfield, Illinois. The network has 20 intersections and a mix of one-way and two-way streets with different number of lanes and turning

configurations at signalized intersections. Figure 5-4 shows the network, which is divided into 20 sub-networks, each corresponding to a sub-problem. Table 5-2 presents the general characteristics of the Springfield network in the CTM representation and the signal timing parameters. The total study period is 500 time steps (50 minutes), where each time step is six seconds. Four demand patterns, as described in previous chapter, were used in the case study network.

Four demand patterns were used in the case study network:

1. Symmetric undersaturated demand pattern: 500 veh/hr/ln on all entry points,
2. Symmetric saturated demand pattern: 900 veh/hr/ln on all entry points,
3. Symmetric oversaturated demand pattern: 1200 veh/hr/ln on all entrty points, and
4. Asymmetric demand pattern covering both under and oversaturated conditions. Demand profiles for east-west and north-south streets of the network are shown in Figure 5-5.

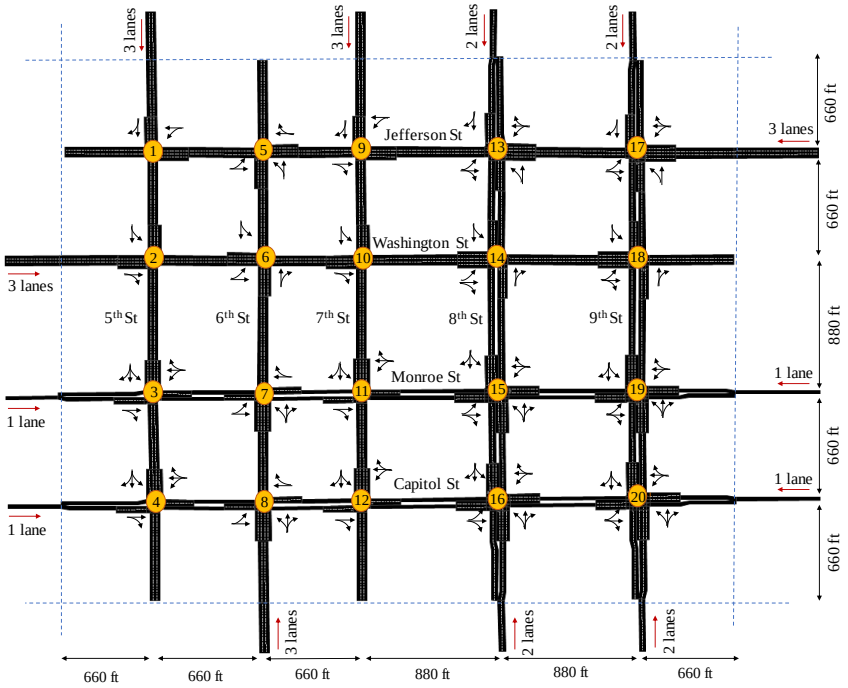


Figure 5-4: Springfield network.

Table 5-2: Characteristics of Springfield network in CTM

Link data	
Number of lanes per link	1, 2, or 3
Maximum free-flow speed (mph)	30
Link saturation flow (veh/hour/lane)	1800
Optimization period (time steps)	500
Prediction period (time steps)	15
Duration of each time step (seconds)	6
Number of cells	342
Cell jam density (veh/cell/lane)	12
Cell saturation flow (veh/cell/ lane)	3
Signal timing parameters	
Maximum green for through (seconds)	60
Minimum Green time for through (seconds)	18
Maximum green for left turn (seconds)	24
Minimum green for left turn (seconds)	6

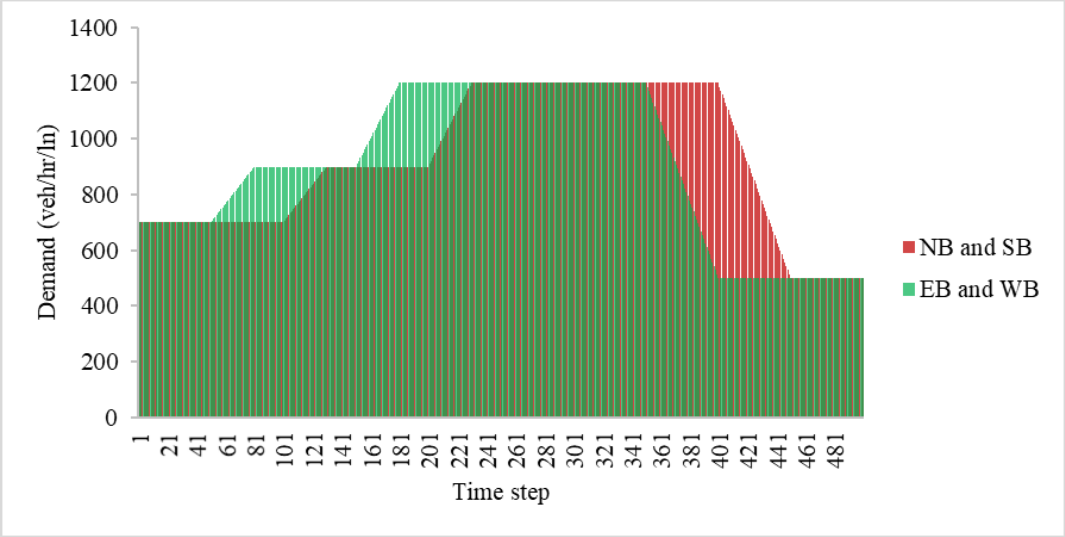


Figure 5-5: Time-variant Demand profile for the network

Note that demand patterns 1, 2, and 3 are considered in the CTM to show the analytical optimality gap of DOCA. Demand pattern 4 is used in Vissim to represent a more realistic

evaluation of the proposed algorithm in a more realistic simulated environment. Three scenarios are considered to evaluate the effectiveness of CSSO in congestion management:

1. Independent speed optimization with pre-defined signals (signals are optimized using genetic algorithm (Hajbabaie et al., 2011)),
2. Independent signal timing optimization, and
3. Coordinated signal timing and speed optimization

Under independent speed optimization, the signal timing parameters are fixed and input to an optimization program who only optimizes speed across different network links. In independent signal timing optimization, the desired speed of vehicles is not changed and only signal timing parameters are optimized in the network.

5.5. Results

5.5.1. Mathematical programs results

Figure 5-6 (a)-(c) shows DOCA's solutions and the best upper and lower bounds found by the Benders decomposition in the undersaturated, saturated, and oversaturated conditions, respectively. Since, DOCA provides a feasible solution, it can always be considered as a lower bound for a maximization problem. Therefore, the optimality gaps are calculated from the difference between the upper bound found by the Benders decomposition technique and the DOCA-CSSO solution. Note that the optimality gaps are found when the study period was set to 200 time steps since the runtime of the Benders decomposition technique does not allow increasing the study period further. The optimality gaps were 4.9%, 5.4%, and 5.2% for the undersaturated, saturated, and oversaturated demand levels, respectively.

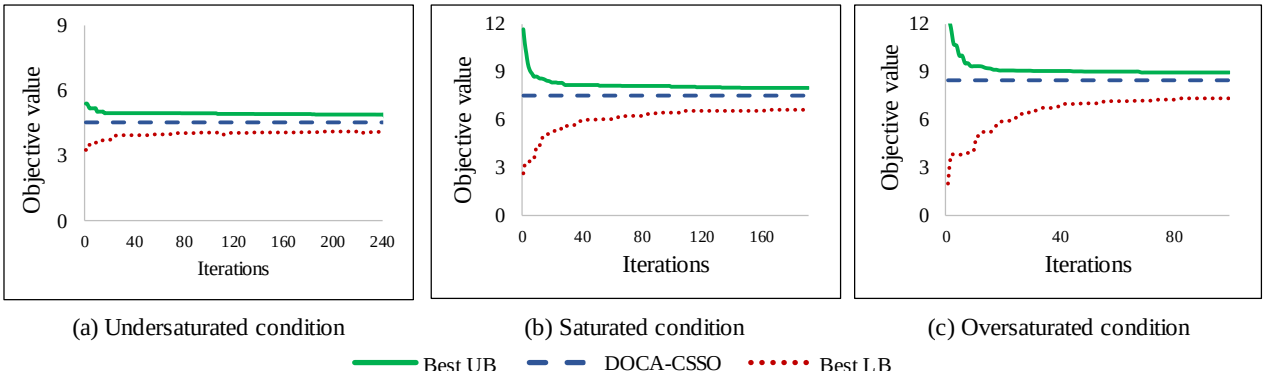


Figure 5-6: DOCA and the benchmark solutions objective values ($\times 10^5$) for three demand patterns.

We used the Posteriori technique (Mavrotas, 2009) to select the most preferred weight (alpha) in the objective function. For this purpose, the weights of objective function terms are changed incrementally. Figure 5-7 shows the trade-off between the travel time and speed variations for three demand patterns based on different alpha values with an increment of 0.05. For all demand patterns, the travel time decreases by increasing alpha. This is expected since a higher value of alpha prioritizes network throughput more. The speed variance also strictly increases by increasing the value of alpha, which indicates that highly harmonized speeds can reduce the network performance. We selected the value of 0.75 for alpha in this study.

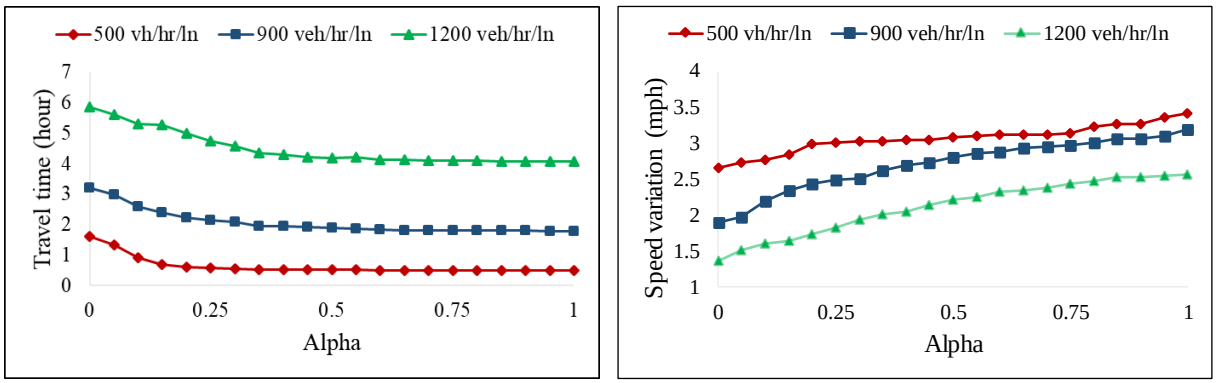


Figure 5-7: Trade-off between travel time and speed variations.

Table 5-3 shows network performance measures for CSSO, independent signal timing optimization, and independent speed optimization for the same three demand patterns. In the

undersaturated demand conditions, CSSO significantly improved network performance compared to speed optimization. The travel time decreased by 26.1% and the network throughput and average speed increased by 1.5% and 36.8%, respectively. In comparison to signal timing optimization, CSSO only improved the travel time and average speed by 2.1% and 2.2%, respectively. Both signal timing optimization and CSSO yield similar improvements in network throughputs in undersaturated flow conditions when a macroscopic first-order traffic flow model is utilized.

In the saturated demand conditions, CSSO showed significant improvements in comparison with the two other scenarios. The coordinated approach reduced the travel time by 36.7% and increased the network throughput and average speed by 17.1% and 81.2%, respectively, compared to speed optimization. CSSO reduced the travel time by 2.5% and increased the network throughput and average speed by 2.0% and 4.8%, respectively, compared to signal timing optimization. These trends show that the cooperation between signal timing and speed optimization offers great potential for reducing traffic congestion in saturated demand conditions, even when a first-order traffic flow model is in use.

The results of oversaturated demand conditions show trends similar to those that were observed in the saturated demand conditions. The coordinated approach reduced the travel time by 29.9% and respectively increased the network throughput and average speed by 28.6% and 82.2% compared to the speed optimization. In comparison to the signal timing optimization, CSSO reduced the travel time by 5.6% and increased the network throughput and average speed by 5.0% and 10.6%, respectively. These trends indicate that CSSO offers great potential for congestion management in oversaturated demand conditions.

Note that the signal timing optimization improved the network performance significantly in comparison with the speed optimization in all demand patterns. CSSO improved the network performance further in more congested conditions. The trends suggest that CSSO yields higher improvement in traffic operations with more congested demand levels in comparison with the signal timing optimization.

Table 5-3: The network mobility performances for three scenarios and three demand patterns based on CTM

Demand (veh/hour/lane)	Mobility performance	(3) CSSO			(2) Signal optimization	(1) Speed optimization
		Value	% Diff to (2)	% Diff to (1)		
500	Travel time (hour)	511.2	-2.1	-26.1	522	691.8
	Throughput	12385	0.0	1.5	12389	12197
	Average speed (mph)	21.9	2.2	36.8	21.5	16.0
900	Travel time (hour)	1863	-2.5	-36.7	1911.4	2943.0
	Throughput	18717	2.0	17.1	18342	15977
	Average speed (mph)	9.4	4.8	81.2	8.9	5.2
1200	Travel time (hour)	4187.3	-5.6	-29.9	4436.9	5969.7
	Throughput	19950	5.0	28.6	19005	15509
	Average speed (mph)	4.5	10.6	82.2	4.1	2.5

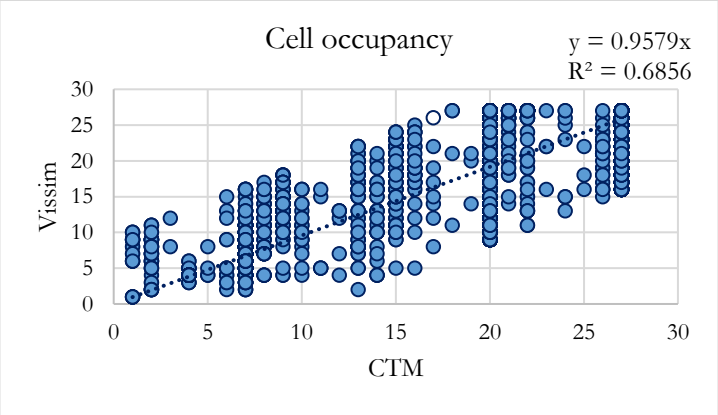
*Diff= Difference

5.5.2. Vissim results

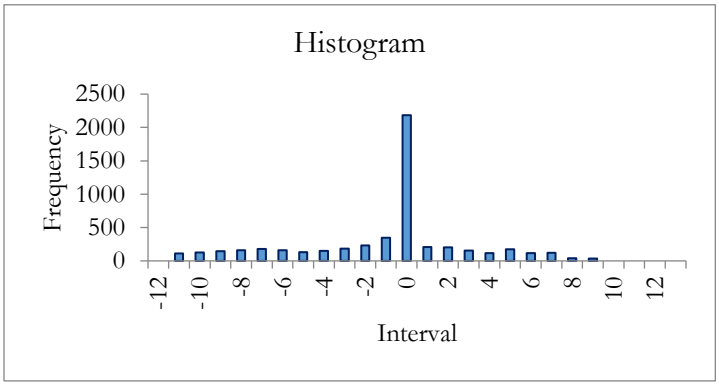
In this section, we present the results that are obtained by incorporating the proposed methodology in Vissim. Vissim provides a more realistic representation of traffic operations on urban-street networks and accounts for the interactions between vehicles using car following and lane changing models. For this purpose, the time variant demand profile is used.

We simulated the movement of vehicles in an arterial streets using both Vissim and CTM. We calibrated the parameters of cell transmission model including cell jam density and cell saturation flow rate to make it accurate enough to represent the flow of traffic in Vissim. Figure

5-8-a compares the occupancy of segments of the roads (i.e. cells) based on the actual traffic in Vissim and the estimated values from CTM. We can see that there is a direct relationship between the actual and estimated occupancy. In addition, Figure 5-8-b shows the frequency of the difference between the actual and estimated values. We see that for most of the cases, CTM provides estimation with high accuracy.



a) Occupancy of segment in Vissim vs. CTM



b) The frequency of differnces between Vissim and CTM observations

Figure 5-8: Comparing the flow movement in Vissim vs. CTM

To select the right prediction period in model predictive control, we performed a sensitivity analysis. Figure 5-9 shows the changes of objective value for demand pattern 4 for different prediction horizons. Increasing the prediction horizon up to 90 seconds improves the performance of the system by reducing the objective function. However, after 90 seconds the performance of

the system remains stable and no further improvement is achieved. Therefore, 90 seconds would be enough to look forward without increasing the complexity of the optimization problem significantly.

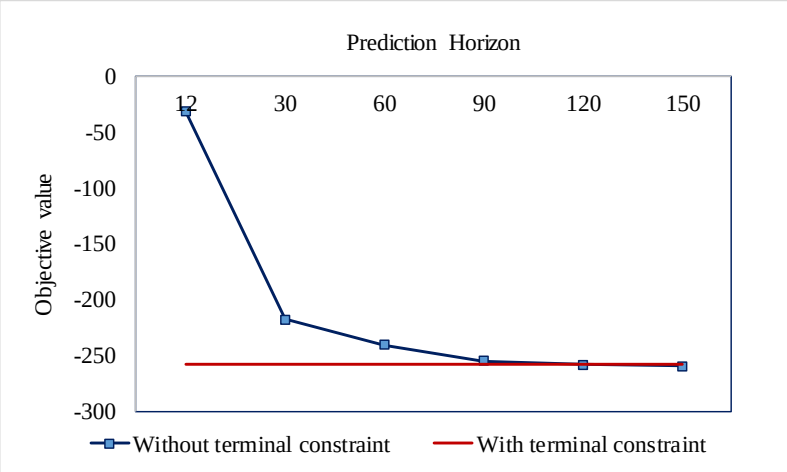


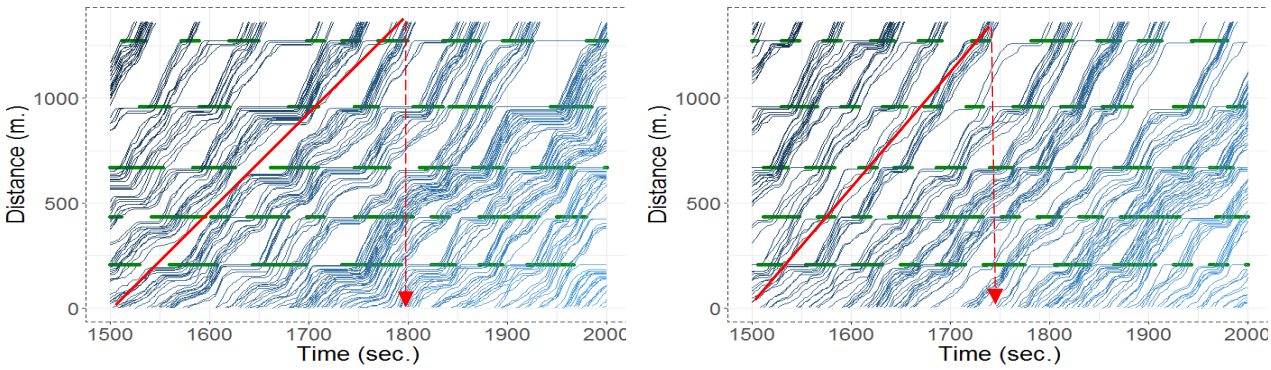
Figure 5-9: The changes of objective value for different prediction horizon in MPC

Table 5-4 shows the network performance for CSSO, signal timing optimization, and speed optimization strategies in Vissim. The coordinated approach respectively reduced the travel time, average delay, average number of stops, and average delay at stops by 32.5%, 38%, 35.3% and 42.1% compared to the case that only vehicles speeds are optimized. Moreover, the network throughput and average speed were increased by 41.4% and 104.2%, respectively. CSSO respectively reduced the travel time, average delay, average number of stops, and average delay at stops by 1.9%, 5.3%, 28.5%, and 5.4% compared to the case that only signal timing parameters are optimized. In addition, CSSO increased the network throughput and average speed by 1.7% and 3.4%, respectively. These results are consistent with the findings from the mathematical analysis indicating that the coordinated approach has a positive effect on traffic.

Table 5-4: The network mobility performances for three scenarios based on Vissim

Mobility performance	(3) CSSO			(2) Signal optimization	(1) Speed optimization
	Value	% Diff to (2)	% Diff to (1)		
Travel time (hour)	3533.8	-1.9	-32.5	3600.5	5236.7
Throughput (vehicle)	15093	1.7	41.4	14839	10671
Average Delay (sec.)	464.4	-5.3	-38.0	490.5	749.0
Average speed (mph)	3.9	3.4	104.2	3.8	1.9
Average number of stops	12.0	-28.5	-35.3	16.9	18.6
Average delay at stop (sec.)	385.5	-5.4	-42.1	407.5	666.2

Figure 5-10 shows vehicle trajectories when only the signal timing parameters were optimized (part a) and when both signal timing parameters and speeds were optimized (part b). Vehicles pass through five intersections on eastbound Washington St. with three lanes. The CSSO smoothened the movement of vehicles and reduced the number of stops. Moreover, it is shown that vehicles that entered the network at the same time left the network earlier when signal and speeds are optimized together.



(a) Signal timing optimization (b) Coordinated signal and speed optimization

Figure 5-10: Vehicles trajectories

The algorithm was solved for demand pattern 3 (the highest demand) on a PC with a Core i9 CPU and 64 GB of memory. Figure 5-11 shows the runtime distribution for solving the

optimization problem at each intersection in the network. The maximum runtime was 2.45 seconds. Since the implementation period is six seconds the algorithm works in real time even with considering a conservative time for communication between sub-problems.

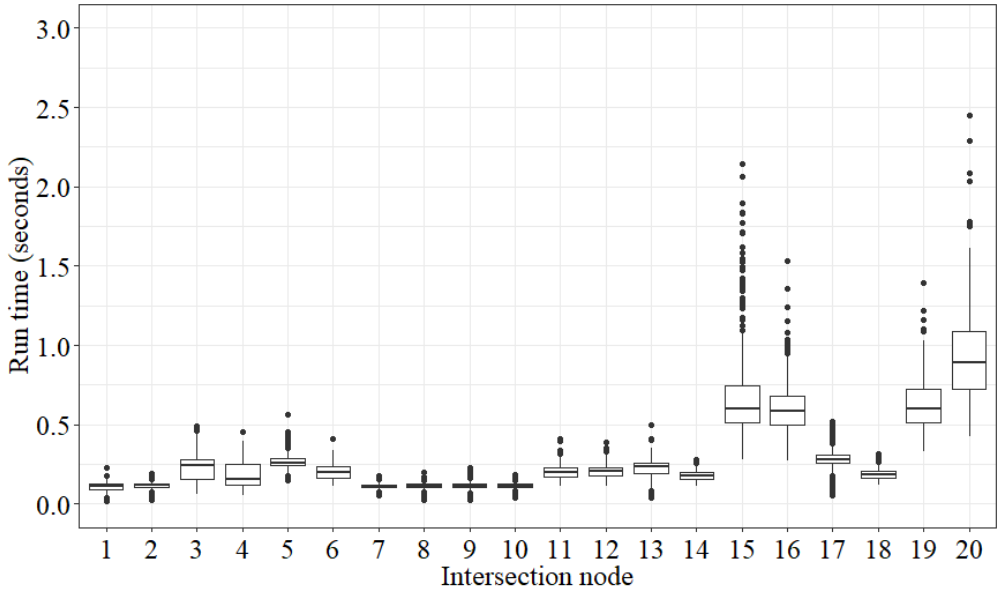


Figure 5-11: DOCA runtimes at each intersection node

5.6. Conclusion

In this chapter, we formulated coordinated signal timing and speed optimization problem in urban street networks as an MILP based on the CTM network loading concept. Since the problem is a mixed integer linear program, it does not scale well with the size of the network and cannot be solved in real-time. Therefore, a distributed optimization and coordination algorithm is developed to improve the scalability and provide real-time solutions. Distributed optimization decomposes the network-level problem to several sub-network level sub-problems by relaxing the constraints that represent the interrelationship between sub-problems. The decomposition significantly reduces computational complexity; however, may affect solution quality adversely. To avoid this issue, an effective coordination scheme is designed that re-enforces the re-introduced relaxed constraints by exchanging information among adjacent sub-problems and implementing

them in the constraints. The required information is (a) the outflow of cells at boundaries of a computation node, (b) the available capacity of the receiving cells, and (c) the average speed in the receiving cells. The problem formulation was modified for each computation node to accommodate the incorporation of the information.

We tested the proposed algorithm in a network with twenty intersections. Compared to the Benders decomposition algorithm (benchmark), DOCA found solutions with at most 5.4% optimality gap. Moreover, it was shown that the CSSO is more effective when the network is congested. In comparison to the signal timing optimization, DOCA-CSSO reduced the travel time by 0.5%, 1.1%, and 2.7% in the undersaturated, saturated, and oversaturated demand conditions, respectively. Results that were obtained by implementing the proposed algorithm in Vissim were positive too. Compared to the speed optimization, CSSO reduced the travel time, average delay, average number of stops, and average delay at stops by 32.5%, 38%, 35.3%, and 42.1%, respectively. In addition, the network throughput and average speed increased by 41.4% and 104.2%, respectively. In comparison with the signal optimization, CSSO reduced the travel time, average delay, average number of stops, and average delay at stops by 1.9%, 5.3%, 28.5%, and 5.4%, respectively. Moreover, the network throughput and average speed increased by 1.7% and 3.4%, respectively.

This study assumed that all vehicles are connected and automated and the information about their position and speeds are available. In addition, it was assumed that all vehicles follow the assigned speeds. Further research is needed to investigate how CSSO performs under various market penetration rates of connected and automated vehicles or their compliance with the assigned speed. Moreover, this study assumed that vehicle positions are accurate, and the

communications are instantaneous. Further research on the effects of error in vehicle positions and communication latency is needed.

CHAPTER 6: TRAFFIC SIGNAL TIMING AND TRAJECTORY OPTIMIZATION IN A MIXED AUTONOMY TRAFFIC STREAM

As discussed in previous chapters, signalized intersections are major bottlenecks in urban street networks. Establishing effective communications between signal controllers and connected vehicles can lead to finding signal timing plans that benefit system users with reduced delay and improved mobility. In addition, optimizing the trajectory of connected automated vehicles (CAVs) through cooperation with signal controllers and the information received from CAVs and human-driven vehicles (HVs) can smoothen the traffic flow and prevent excessive stops at signalized intersection. However, in previous chapter, we assumed all vehicles are connected and automated vehicles.

Past research shows that optimizing the trajectory of connected automated vehicles (CAVs) and the timing of traffic lights at signalized intersections offers a great potential to improve traffic operations (Guo et al., 2019; Li and Zhou, 2017; Niroumand et al., 2020a; Tajalli et al., 2020; K. Yang et al., 2016; Yu et al., 2018). Joint CAV trajectory and signal timing optimization helps plan the arrival time of vehicles to the intersection more accurately to utilize green durations more efficiently. Therefore, the number of stops, fuel consumption, and travel delay at intersections will be reduced significantly. However, the cooperation between a signal controller and approaching vehicles requires an extensive amount of communication and computational power (Mohebifard et al., 2019). It is not likely that a signal controller can handle all the required computations. In fact, previous studies show the effectiveness of signal timing and trajectory optimization (a) in intersections with simple layouts (e.g., one-way streets (Li et al., 2014; K. Yang et al., 2016), or no turning movements (Jung et al., 2016)), (b) under low traffic volumes (Jung et al., 2016; Li et al., 2014; B. Xu et al., 2018a; K. Yang et al., 2016), or (c) using simplifying or restrictive

assumptions (e.g., using first-order traffic flow model to update trajectory of CAVs (Li et al., 2015; Li and Zhou, 2017) or optimizing the trajectory of a portion of CAVs (Yu et al., 2018)). Approximation methods and heuristic algorithms are also used to solve the problem in more complex conditions; however, at the expense of sacrificing the quality of the solutions.

This chapter presents a methodology for joint CAV trajectory and signal timing optimization at signalized intersections that provides a balance between computational efficiency and solution quality. As Figure 6-1shows, the methodology is designed for a mixed traffic stream of CAVs and HVs, where the movement of CAVs is controlled centrally and communicated with them through a vehicle-to-infrastructure communication system at the signalized intersection. The methodology requires the initial location and speed of all vehicles (CAVs and HVs) in the vicinity of the intersection and predicts the location of HVs over a planning horizon using car following concepts. The movement of HVs are not optimized. We assume that either all vehicles are connected (just to collect vehicle location and speed), or the intersection is equipped with detectors (e.g., radar units or video cameras) that can provide the location and speed of vehicles. Note that the proposed algorithm works if HVs do not send information to the signal controller as long as detectors are available to collect the required data. This study formulates the joint optimization as a mixed-integer nonlinear program, whose objective reduces the total travel time and speed variance at the intersection. The decision variables are the acceleration rates of CAVs and signal timing parameters. The interaction between HVs and CAVs is incorporated into the optimization problem using a linear car-following model developed by Helly (Helly, 1959). We modified the linear model to become responsive to traffic lights by incorporating signal timing parameters. The signal timing parameters are optimized through a cycle- and phase-free plan that satisfies the minimum and maximum green constraints in addition to yellow time interval.

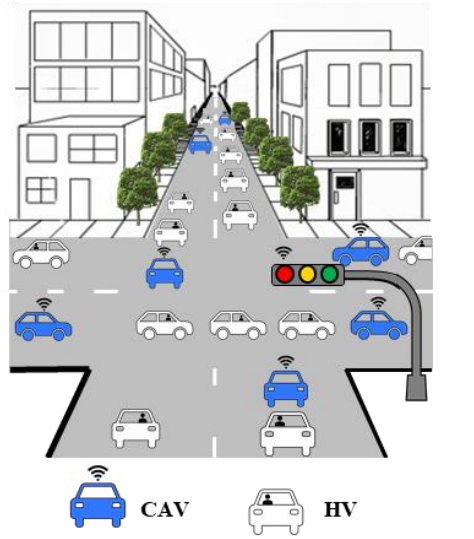


Figure 6-1: Connected signal controller interacting with CAVs and HVs

6.1. Research gaps and contributions

The signal timing and trajectory optimization problem is complex and there is a need to develop algorithms providing a balance between computational complexity and solution quality. Previous studies showed that there is a trade-off between the computational efficiency and solution quality, where more efficient algorithms are associated with more simplifying assumptions. The contributions of this study are three-fold: 1) this study introduces a solution technique that provides a balance between computational efficiency and solution quality: it reduces the computational complexity of STTO without significantly sacrificing the quality of the solution. We achieved this balance by developing an efficient solution technique using Lagrangian relaxation method to decompose a centralized STTO problem into several lane-level optimization sub-problems. Therefore, intersections with various layouts and different demand levels can be analyzed; 2) we reduced the optimality gap of the proposed Lagrangian relaxation approach so that the solutions that it finds are close to the optimal solutions. We did so by proposing a problem reformulation using maximal clique sets to tighten the convex hull of mixed-integer feasible region. As a result, we improve the convergence property of Lagrangian relaxation technique for STTO problem and

ensure finding near optimal solutions efficiently with reduced duality gap; 3) the proposed algorithm does not require the following simplifying assumptions: a) fleet of all automated vehicles, b) forcing vehicles to move in a platoon, c) arriving at the intersection with the desired speed, or d) no initial queue existing at the intersection during red phases. Therefore, the effect of joint signal timing and trajectory optimization of CAVs can be studied in a mixed autonomy environment with various market penetration rates of CAVs.

6.2. Problem formulation

This study optimizes the trajectory of all CAVs and signal timing parameters cooperatively. The human-driven vehicles are assumed to be connected to the infrastructures and move following a linear car-following model. We define L as the set of all lanes at an intersection and I as the set of all vehicles at the intersection neighborhood. Moreover, I^l, I_A^l, and I_H^l represent the sets of all vehicles, CAVs, and HVs on lane $l \in L$, respectively. We define C_l as the set of all lanes conflicting with lane $l \in L$. Figure 6-2 shows the defined sets in an isolated intersection.

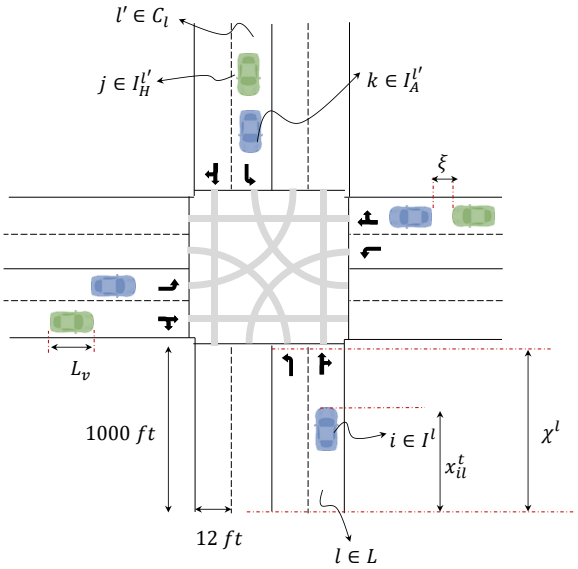


Figure 6-2: Defined sets and parameters in STTO problem

We define T_s and T_t as the set of signal timing and trajectory optimization time steps, respectively. Note that vehicle trajectory update time steps ΔT_t are different than the signal timing parameter update time steps ΔT_s. The trajectory update time step is much shorter than the signal timing time step to take into account the uncertainty in driver behaviors and capture any differences between the predicted and actual vehicle trajectory. The signal timing parameters do not need to be updated as often as vehicle trajectories. Equations (6.1) show the relation between these two time steps. Figure 6-3 shows an example assuming vehicle trajectories are updated every 0.5 seconds and the signal timing parameters are updated every two seconds. For instance, at time 1.5 sec, the time index $t \in T_{trj}$ for trajectory takes the value of 3. Based on equation (6.1), the time index for signal $p \in T_s$ is equal to $0 = \left\lfloor 3 \times \frac{0.5}{2} \right\rfloor$. The $\lfloor\ \rfloor$ operator rounds down the corresponding argument inside it.

$$p = \left\lfloor t \times \frac{\Delta T_t}{\Delta T_s} \right\rfloor \qquad t \in T_t, p \in T_s \quad (6.1)$$

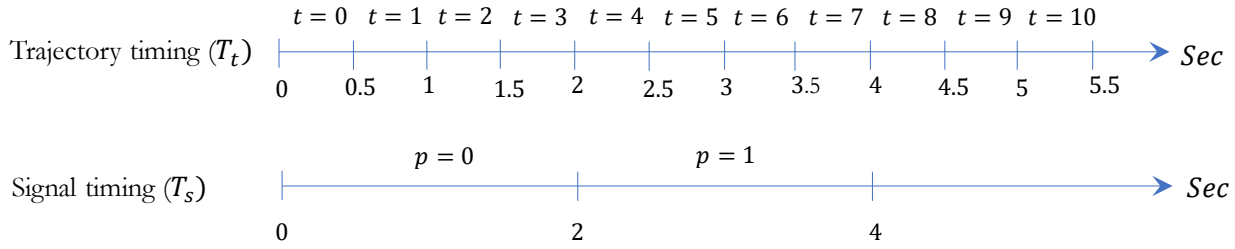


Figure 6-3: The relationship between signal timing and trajectory update intervals

The status of the signal timing variables including green time g_l^p and yellow time y_l^p on lane $l \in L$ and time-step $p \in T_s$ will be optimized based on information on location x_{il}^t and speed v_{il}^t of vehicle $i \in I^l$ approaching the intersection on lane $l \in L$ at time-step $\in T_t$. The acceleration rate a_{il}^t of vehicle $i \in I_A^l$ at time $t \in T_t$ is also the control variable in motion planning of CAVs.

Table 6-1 provides a detailed definition of variables, sets, and parameters used in the problem formulation.

Table 6-1: Definition of sets, decision variables, and parameters

Sets:	
T_t	set of all time steps for trajectory updates
T_s	set of all time steps for signal timing updates
L	set of all lane groups at the intersection
I^l	set of all vehicles on lane $l \in L$
I_A^l	set of all CAVs on lane $l \in L$
I_H^l	set of all HVs on lane $l \in L$
C_l	set of all lanes conflicting with lane $l \in L$
Decision Variables:	
a_{il}^t	the acceleration of vehicle $i \in I_l$ on lane $l \in L$ at time step $t \in T_t$
g_l^p	a binary variable for green signal status on lane $l \in L$ at time step $p \in T_s$
y_l^p	a binary variable for yellow signal status on lane $l \in L$ at time step $p \in T_s$
State Variables:	
x_{il}^t	the distance of vehicle $i \in I_l$ at time step $t \in T_t$ from the beginning of lane $l \in L$
v_{il}^t	the speed of vehicle $i \in I_l$ on lane $l \in L$ at time step $t \in T_t$
Parameters:	
v'	maximum speed
v	minimum speed
σ'	maximum acceleration
σ	minimum acceleration
χ^l	the location of intersection stop line based on the coordination of lane $l \in L$
w_l	weight factor
L_v	the length of vehicle
ξ	desired safety distance between vehicles
τ_H	the reaction time of HVs
τ_A	the reaction time of CAVs
β	weight factor in the objective function
ω_l'	maximum green time for lane $l \in L$
ω_l	minimum green time for lane $l \in L$
γ	yellow time interval
M	a very large value

6.2.1. Objective function

The objective function of STTO is shown in (6.2) with two terms. The first term maximizes the distance of each vehicle from the beginning of lane $l \in L$, where vehicle $i \in I^l$ is located on

(see Figure 6-2). Maximizing the summation of distance for all vehicles at the intersection is equivalent to minimizing the total travel time at the intersection. Weight factor w_{il} for vehicle $i \in I_l$ on lane $l \in L$ is considered to prevent building a long queue at each lane of the intersection. This weight is equal to the delay that each vehicle is experiencing after entering the intersection vicinity. Therefore, a higher priority will be assigned to vehicles on a lane with higher experienced delays. The second term of the objective function smoothens the motion of CAVs by minimizing the difference of speeds of each CAVs between two consecutive time steps. In Chapter 3:, we showed that minimizing the difference in speeds yields to a smaller number of stops at intersections. Factor $\beta \in (0,1)$ provides a proper balance between the two terms of the objective function and is an input.

$$f = \underset{\boldsymbol{a}}{\text{Max}} \left(\beta \sum_{t \in T_t} \sum_{l \in L} \sum_{i \in I^l} w_{il}(x_{il}^t) - (1-\beta) \sum_{t \in T_t} \sum_{l \in L} \sum_{i \in I_A^l} \left| v_{il}^{t+1} - v_{il}^t \right| \right) \quad (6.2)$$

6.2.2. Constraints

The speed and position of all vehicle are updated based on the basic equations of motion, as shown in constraints (6.3) and (6.4).

$$v_{il}^{t+1} = v_{il}^t + a_{il}^t \Delta T_t \qquad \forall\, i \in I^l, l \in L, t \in T_t \quad (6.3)$$

$$x_{il}^{t+1} = x_{il}^t + v_{il}^t \Delta T_t + \frac{1}{2} a_{il}^t \Delta T_t^2 \qquad \forall\, i \in I^l, l \in L, t \in T_t \quad (6.4)$$

6.2.2.1. Car following constraints

The estimated location and speed of HVs are needed for optimizing the signal timing parameters and trajectory of CAVs. Hence, the future trajectory of HVs are predicted based on a car-following behavior. This study considered the linear car-following model developed by Helly

(1959). This car-following model is used in modeling adaptive and cooperative adaptive cruise control systems (Schakel et al., 2010; Wang et al., 2014a). Furthermore, Panwai and Dia (2005) showed an appropriate fit of Halley's car following model to real-world traffic data. The followers respond to both relative speed and distance from their preceding vehicle in this car following model, see equations (6.5) for calculation of the acceleration rate. Parameters α_1 and α_2 are fixed, have positive values, and should be within the ranges of $[0.17, 1.3]$ and $[\frac{1}{4}\alpha_1, \frac{1}{2}\alpha_1]$, respectively (Askari et al., 2016; Brackstone and McDonald, 1999). The first term in equations (6.5) considers the relative speed between the leading and following vehicles. A positive speed difference (i.e., leader travels at a higher speed) results in a positive acceleration rate for the follower, while a negative speed difference forces the follower to decelerate. The second term takes into account the relative distance between successive vehicles.

$$a_{il}^t = \alpha_1\left(v_{i-1,l}^t - v_{il}^t\right) + \alpha_2\left(\left(x_{i-1,l}^t - x_{il}^t - L_v\right) - \xi - \tau_H v_{il}^t\right) \qquad \forall\, i \in I_H^l, l \in L, t \in T_t \qquad (6.5)$$

We enhance the car-following model to account for traffic signals so that a vehicle decelerates when approaching a red signal. The traffic light is considered as a virtual vehicle with either zero (for a red signal) or max speed (for a green signal). Equations (6.6) show how a vehicle's acceleration rate is updated when reaching a traffic light when vehicle's location is before the intersection stop bar. Note that when the signal is red with $g_l^p = 0$, the virtual stopped vehicle is located at the intersection stop bar χ^l and there would be a desired safety distance ξ between the vehicle and the intersection stop bar. However, when the traffic light is green, the location of preceding virtual vehicle changes to a far distance from the following vehicle and the desired safety distance to the intersection stop bar will be reduced to zero. Moreover, it should be noted that the connection between vehicles and traffic light is only needed when vehicles are before the

intersection stop bar with $x_{il}^t \le \chi^l$ and equations (6.6) would become inactive after passing the intersection. Multiplying signal timing variable g with a big coefficient (M) in the second term of equations (6.6) results in having the virtual car in a very long distance from the following vehicle when signal is green.

$$\text{If } x_{il}^t \le \chi^l: \qquad \forall\, i \in I_H^l, l \in L; t \in T_t,\ p \in T_s$$
$$a'^t_{il} = \alpha_1\left(g_l^p v' - v_{il}^t\right) + \alpha_2\left(\left(\chi^l + M g_l^p - x_{il}^t\right) - (1 - g_l^p)\xi\right) \tag{6.6}$$

An HV may be too close to the intersection to stop safely before the stop-bar when the signal turns red. Therefore, we considered yellow signal timing parameter y_l^p between each green and red times. The duration of yellow time should be long enough to let a vehicle with maximum speed either completely stop before the intersection stop bar or pass the intersection before the unset of the red signal.

The linear car-following model presented by equations (6.5) and (6.6) does not limit the acceleration rate and speed of vehicles between minimum and maximum values. As a result, we formulate the car following model based on a max-min function, as shown in equations (6.7). This formulation is a complete form of the car-following model that describes situations such as moving in free-flow condition, following other vehicles in stationary and non-stationary conditions, and approaching slow or standing vehicles and red signals.

$$a_{il}^t = \max\left\{\sigma,\ \frac{v - v_{il}^t}{\Delta T_t},\ \min\left\{\begin{matrix}\sigma', \\ \dfrac{v' - v_{il}^t}{\Delta T_t}, \\ \alpha_1\left(v_{i-1,l}^t - v_{il}^t\right) + \alpha_2\left(\left(x_{i-1,l}^t - x_{il}^t - L_v\right) - \xi - \tau_H v_{il}^t\right), \\ \alpha_1\left(g_l^p v' - v_{il}^t\right) + \alpha_2\left(\left(\chi^l + M g_l^p - x_{il}^t\right) - (1 - g_l^p)\xi\right) + M\gamma_{il}^t\end{matrix}\right\}\right\}$$

$$\forall\, i \in I_H^l, l \in L, t \in T_t, p \in T_s \quad (6.7)$$

Binary variable γ_{il}^t is introduced in equations (6.7) to break the connection between the traffic light and vehicles after they pass the intersection stop bar ($x_{il}^t > \chi^l$). Variable γ_{il}^t becomes zero when the vehicle is downstream of the intersection stop bar and becomes one when the vehicle has passed the intersection stop bar. Equations (6.8) define the value of γ_{il}^t for vehicle $i \in I^l$ on lane $l \in L$ at time step $t \in T_t$. The product of γ with a big value (M) yield to the deactivation of the last term of min function in equations (6.7) when a vehicle passed the intersection stop bar.

$$\gamma_{il}^t = \begin{cases} 0 & \text{if } x_{il}^t \le \chi^l \\ 1 & \text{otherwise} \end{cases} \qquad \forall\, i \in I^l, l \in L, t \in T_t \quad (6.8)$$

6.2.2.2. CAV motion constraints

CAVs can receive the optimal trajectory without being in a car following mode. As mentioned earlier, objective function (6.2) smoothens the trajectory of CAVs to prevent frequent stops at the intersection. Constraints (6.9) ensure a safe distance between a CAV and its preceding vehicle. The distance between two consecutive vehicles is a function of desired safety distance ξ, average vehicle length L_v, and distance $\tau_A v_{il}^t$ that can be passed within the reaction time of the following vehicle. Parameter τ_A represent the reaction time for CAVs.

$$x_{i-1,l}^t - x_{il}^t \ge \xi + L_v + \tau_A v_{il}^t \qquad \forall\, i \in I_A^l, l \in L, t \in T_t \quad (6.9)$$

Constraints (6.10) are used to prevent a CAV from entering the intersection area when the signal is not green. When a CAV has not arrived at the intersection and the signal is red, the distance between the vehicle and intersection stop bar should be greater than or equal to a safe distance, as shown on the right-hand side of constraints (6.10). The safe distance is a function of the maximum distance $\Delta T_t v_{il}^t$ that a vehicle can travel in one time step with its current speed at

time $t \in T_t$. This constraint needs to become inactive when either the signal is green, or the vehicle has passed the stop bar. Therefore, variables g and γ are multiplied to a big value (M) to ensure that constraints (6.10) are always satisfied when either the signal is green, or the vehicle has passed the stop bar.

$$\chi^l - x_{il}^t \geq \Delta T_t v_{il}^t - M g_l^p - M\gamma_{il}^t \qquad \forall\, i \in I_A^l, l \in L, t \in T_t, p \in T_s \quad (6.10)$$

Constraints (6.11) and (6.12) limit the acceleration and speed of CAVs between allowed ranges, respectively.

$$\sigma \leq a_{il}^t \leq \sigma' \qquad \forall\, i \in I_A^l, l \in L, t \in T_t \quad (6.11)$$

$$v \leq v_{il}^t \leq v' \qquad \forall\, i \in I_A^l, l \in L, t \in T_t \quad (6.12)$$

6.2.2.3. Signal timing constraints

The signal timings are assumed to be cycle- and phase-free; however, restricted to non-conflicting movements. Figure 6-4 shows all allowed movements for northbound and southbound approaches. This study assumes that through and right-turn movements can go on the same lane.

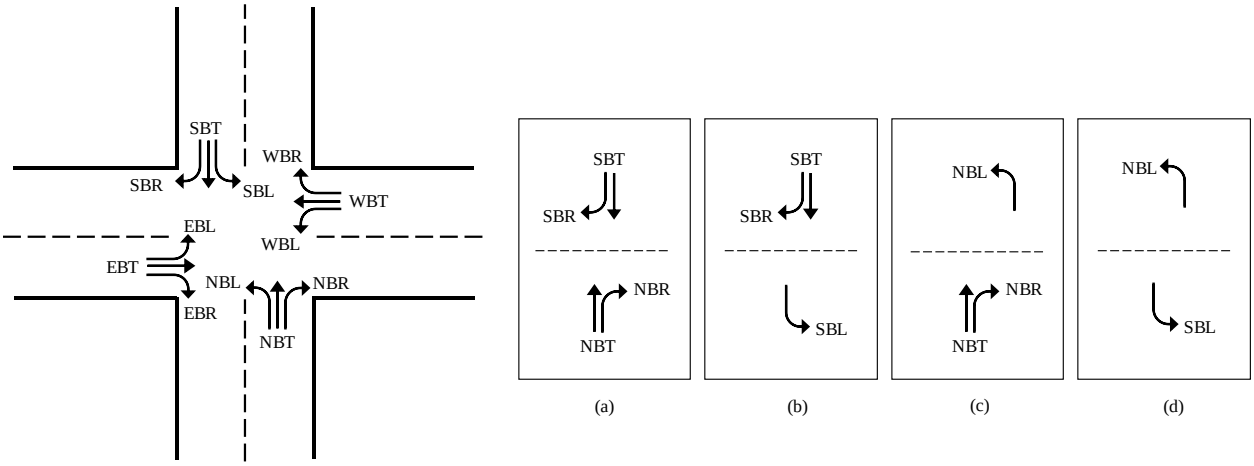


Figure 6-4: The allowed movements for northbound and southbound approaches

Several constraints are considered to prevent collisions between vehicles on conflicting movements. The signal timing decision variables g_l^p and y_l^p respectively take on the value of one when the traffic light serving lane $l \in L$ at time step $p \in T_s$ is green or yellow. Otherwise, both g_l^p and y_l^p take on the value of zero simultaneously, which means the signal is red.

Constraints (6.13) ensure that no pair of conflicting movements can receive a non-red traffic light at time step $p \in T_s$. In other words, the traffic signals serving lanes with conflicting movements can be red at the same time, but only one of them can be either green or yellow at a time. In addition, constraints (6.13) prevent the signal status of a lane to be green and yellow at the same time.

$$g_l^p + g_{l'}^p + y_l^p + y_{l'}^p \leq 1 \qquad \forall\, l \in L, l' \in C_l, p \in T_s \quad (6.13)$$

Constraints (6.14) ensure that the green time assigned to a lane group is less than or equal to maximum green time ω_l'. Constraints (6.15) ensure that the green time duration for lane group $l \in L$ is greater than or equal to minimum green time ω_l.

$$\sum_{j=p}^{p+\omega_l'} g_l^j \leq \omega_l' \qquad \forall\, l \in L, p \in T_s \quad (6.14)$$

$$\sum_{j=p+1}^{p+\omega_l} g_l^j \geq \left(g_l^{p+1} - g_l^p\right)\omega_l \qquad \forall\, l \in L, p \in T_s \quad (6.15)$$

Constraints (6.16) define the duration of yellow time, and constraints (6.17) make sure that the signal switches from green to yellow at the end of the green time interval. Parameter $\mathcal{y}$ is the yellow time interval. Constraints (6.18) ensures the integrality of signal timing variables.

$$\sum_{j=p}^{p+Y} y_l^j \leq \mathcal{Y} \qquad \forall\, l \in L, p \in T_s \quad (6.16)$$

$$\sum_{j=p+1}^{p+Y} y_l^j \geq \left(g_l^p - g_l^{p+1}\right)\mathcal{Y} \qquad \forall\, l \in L, p \in T_s \quad (6.17)$$

$$g_l^p, y_l^p \in \{0,1\} \qquad \forall\, l \in L, p \in T_s \quad (6.18)$$

6.3. Methodology

The proposed formulation is a mixed-integer nonlinear program to cooperatively find optimal signal timing parameters and trajectory of CAVs in an intersection. However, this optimization program is intractable and cannot be solved efficiently due to the presence of nonlinear constraints and binary variables. We first linearize the nonlinear form of the objective function (6.2), car-following model (6.7), and conditional constraints (6.8) to reduce the computational complexity of the proposed formulation. We use the Lagrangian relaxation technique to decompose the intersection-level problem into several lane-level sub-problems with reduced computational complexity and the possibility of allocating one controller to each of them. The controllers will reach a consensus on signal timing parameters and CAV trajectories through sharing Lagrangian multipliers to ensure that a near-optimal solution is found when conflict avoidance constraints (6.13) are satisfied.

6.3.1. Linearization

The second term of the objective function (6.2) contains an absolute value function, which is convex but nonlinear. We introduce two auxiliary non-negative variables z_i^t and u_i^t for each vehicle $i \in I_A^l$ at time step $t \in T_t$ to linearize the absolute value function. Constraints (6.20) and

(6.21) are added to the original problem to enforce that the difference of z_i^t and u_i^t equals the terms in the absolute value function. The linear form of objective function (6.2) is shown in (6.19), where the sum of auxiliary variables is minimized.

$$\underset{\boldsymbol{a}}{\text{Max}}\left(\beta\sum_{t\in T_t}\sum_{l\in L}\sum_{i\in I^l}w_{il}(x_{il}^t)-(1-\beta)\sum_{t\in T_t}\sum_{l\in L}\sum_{i\in I_A^l}(z_{il}^t+u_{il}^t)\right) \tag{6.19}$$

$$z_{il}^t-u_{il}^t=v_{il}^{t+1}-v_{il}^t \qquad \forall\, i\in I^l, l\in L, t\in T_t \tag{6.20}$$

$$z_{il}^t\geq 0, u_{il}^t\geq 0 \qquad \forall\, i\in I^l, l\in L, t\in T_t \tag{6.21}$$

Constraints (6.7) are nonlinear due to the existence of a max-min function. We provide a linear form of these constraints by converting the equality constraints to several inequalities and adding a penalty term to the objective function. Auxiliary variable $b_{il}^t\in R^n$ is introduced to represent the min part of the max-min function (6.7). Note that b_{il}^t has the same unit as acceleration (ft/s^2) and can take both positive and negative values. As constraints (6.22)-(6.25) show, the min part of constraints (6.7) are represented by less than or equal inequalities.

$$b_{il}^t\leq \sigma' \qquad \forall i\in I_H^l, l\in L, t\in T_t \tag{6.22}$$

$$b_{il}^t\leq \frac{v'-v_{il}^t}{\Delta T_t} \qquad \forall i\in I_H^l, l\in L, t\in T_t \tag{6.23}$$

$$b_{il}^t\leq \alpha_1\left(v_{i-1,l}^t-v_{il}^t\right)+\alpha_2\left(\left(x_{i-1,l}^t-x_{il}^t-L_v\right)-\xi-\tau_H v_{il}^t\right) \qquad \forall i\in I_H^l, l\in L, t\in T_t \tag{6.24}$$

$$b_{il}^t\leq \alpha_1\left(g_l^p v'-v_{il}^t\right)+\alpha_2\left(\left(\chi^l+Mg_l^p-x_{il}^t\right)-\left(1-g_l^p\right)\xi\right)+M\gamma_{il}^t \qquad \forall i\in I_H^l, l\in L, t\in T_t, p\in T_s \tag{6.25}$$

Constraints (6.26)-(6.27) relax the max part of function (6.7), which is represented by greater than or equal inequalities. Constraints (6.28) ensure that a_{il}^t is greater than or equal to b_{il}^t from the relaxed min function.

$$a_{il}^t \geq \sigma \qquad \forall i \in I_H^l, l \in L, t \in T_t \qquad (6.26)$$

$$a_{il}^t \geq \frac{v - v_{il}^t}{\Delta T_t} \qquad \forall i \in I_H^l, l \in L, t \in T_t \qquad (6.27)$$

$$a_{il}^t \geq b_{il}^t \qquad \forall i \in I_H^l, l \in L, t \in T_t \qquad (6.28)$$

The proposed linearization (6.22)-(6.28) is loose and cannot equivalently represent the max-min form of (6.7). To fix this issue, we penalize the non-negative difference $a_{il}^t - b_{il}^t \geq 0$ with a big penalty coefficient M in the objective function. The new objective function is shown by (6.29) that pushes b_{il}^t towards a_{il}^t.

$$\max_{\boldsymbol{a}} \left(\beta \sum_{t \in T_t} \sum_{l \in L} \sum_{i \in I^l} w_{il}(x_{il}^t) - (1-\beta) \sum_{t \in T_t} \sum_{l \in L} \sum_{i \in I_A^l} \left| v_{il}^{t+1} - v_{il}^t \right| - M \sum_{t \in T_t} \sum_{l \in L} \sum_{i \in I_H^l} (a_{il}^t - b_{il}^t) \right) \qquad (6.29)$$

Constraints (6.8) are nonlinear due to the if-else condition. The Big-M method allows linearizing the constraints as shown in (6.30) and (6.31). If γ_{il}^t takes the value of 0, constraints (6.30) show that the location of vehicle $i \in I^l$ is before the location of the intersection stop bar on lane $l \in L$ at time $t \in T_t$. However, when γ_{il}^t takes the value of one, constraints (6.31) ensure that the vehicle's location is after the intersection stop bar.

$$x_{il}^t - \chi^l \leq M\gamma_{il}^t \qquad \forall\, i \in I^l, l \in L, t \in T_t \qquad (6.30)$$

$$x_{il}^t - \chi^l \geq M(\gamma_{il}^t - 1) \qquad \forall\, i \in I^l, l \in L, t \in T_t \qquad (6.31)$$

6.3.2. Lagrangian relaxation

The linearization techniques described in the previous section converts an MINLP to a mixed-integer linear problem (MILP). Although this conversion reduces the complexity of the problem, the existence of integer signal timing and other variables makes the problem still intractable and difficult to solve. We develop a Lagrangian relaxation technique that decomposes the problem into several lane-level optimization sub-problems, where the optimal signal timing and vehicles trajectories on each lane of the intersection are found separately in parallel. However, Lagrangian relaxation might not hold a strong duality for MILPs (Fisher, 1981). Hence, we reformulate the problem with a superior solution space structure to overcome this issue. Then, a simple optimization problem is introduced to ensure having a high-quality feasible solution for signal timing plans.

6.3.2.1. Problem reformulation

Pairwise constraints (6.13) are the only common constraints between the lanes that prevent conflicting movements to receive a non-red signal at the same time. Relaxing these constraints and adding them to the objective function with Lagrangian multipliers leads to a lane-level decomposition. The Lagrangian relaxation provides an upper-bound to the optimal solution of non-convex signal timing and trajectory optimization problem based on the weak duality theory. The feasible convex polyhedron provided by the pairwise constraints has non-integer extreme points due to the weak structure of pairwise constraints (Murray and Church, 1996). In other words, there are too many constraints associated with pairwise constraints (6.13), hence the continuous linear relaxation contains many fractions in most of the cases (Murray and Church, 1996). Therefore, there is no guarantee that the Lagrangian relaxation converges to the desired integer solutions with non-zero duality gap (Fisher, 1981). Finding hyperplanes that define the convex hull of integer

solutions helps overcome finding infeasible solutions after the relaxation in addition to satisfying the pairwise constraints (Mirheli et al., 2020).

Let Z_{IP}, Z_{LR}, and Z_{LP} respectively represent the optimal solution of mixed-integer problem, Lagrangian relaxation problem when pairwise constraints (6.13) are relaxed, and the linear relaxation of STTO problem. The optimal value of Z_{LR} provides an upper-bound to the original problem Z_{IP}, which is as good as the upper-bound provided by Z_{LP} when maximizing the objective function (Fisher, 1981). Therefore, the optimal solution of Z_{LR} is within a range between Z_{IP} and Z_{LP}. In other words, $Z_{IP} \leq Z_{LR} \leq Z_{LP}$. One way to reduce the duality gaps between Z_{IP} and Z_{LR} is tightening the convex hull of Z_{LP} and make the corner points of the convex hull of the feasible area closer to the integer values. As a result, the value of Z_{LP} decreases, which yields to lower duality gaps between Z_{IP} and Z_{LP}.

Pairwise conflicting constraints (6.13) can be represented by an undirected conflict graph $G = (N, E)$ that contains edge $\{i, j\} \in E$ if and only if two nodes within a pair of binary nodes $i, j \in N$ cannot be selected at the same time (i.e., two conflicting movements). In other words, E is the edge between two binary nodes when at most one of them can take the value one in the solution of MILP. Figure 6-5 shows the conflict graph for an intersection with four approaches and eight lanes containing through and left-turn movements. The nodes of the graph are signal heads associated with lane $l \in L$, and each edge represents a pairwise conflict constraint represented by (6.13).

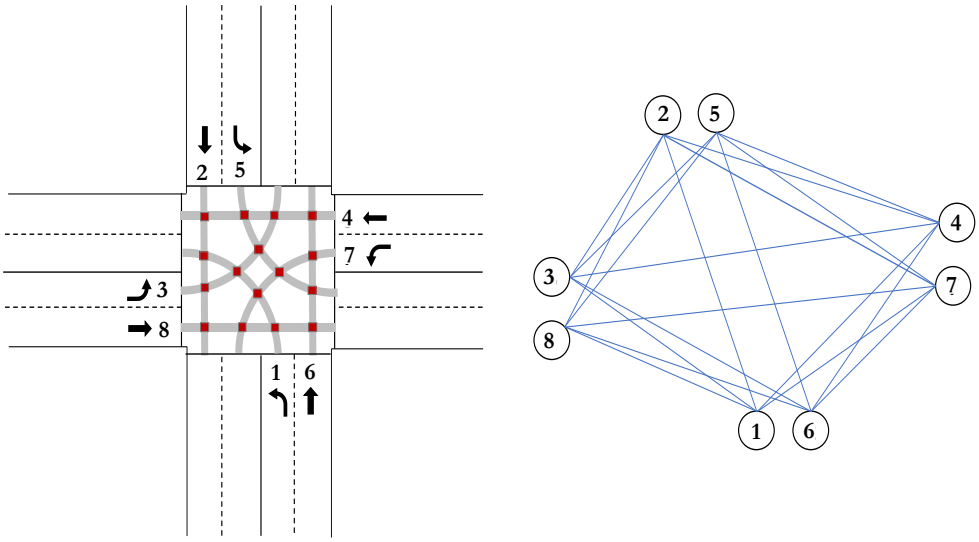


Figure 6-5: Conflict graph for a four-leg intersection with through and left-turn movements

The graph shown in Figure 6-5 represents all conflicts at intersections with exclusive left-turn movements, assuming the right turn movements are operated with the through movements. Pairwise conflicting constraints (6.13) are a special case of a more powerful set of constraints called cliques (Murray and Church, 1997). A clique is a set of several mutually conflicting movements that at most one of them can receive a non-red signal at a time in this study. We define clique $k \in K$ as a set of mutually connected nodes in the conflict graph G, where K is the set of all possible cliques. Respectively, set C_k represents all members of clique k, where each member is mutually connected in graph G. In order to have set C_k in clique k, it is required that all potential nodes $i, j \in C_k$ be connected to each other through the edges of graph G. For instance, $k = \{8,6,7\}$ is a clique for the intersection shown in Figure 6-5 since at most one of the movements among $8, 6,$ and 7 can get a non-red signal. Multiple pairwise constraints (6.13) between different movements can be represented in one clique. Hence, it is possible to define a complete set of clique constraints that impose all edge restrictions with a smaller number of constraints. Respectively, we define the maximal clique as a clique that cannot be enlarged by adding any additional node.

Let define variable s_l^p as the sum of green g_l^p and yellow y_l^p signal timing variables on lane $l \in L$ and time $p \in T_s$. This is binary because at most one of the green and yellow signal timing status take on the value of one, see constraints (6.32)-(6.34).

$$s_l^p = g_l^p + y_l^p \qquad \forall\, l \in L, p \in T_s \qquad (6.32)$$

$$\sum_{l \in C_k} s_l^p \leq 1 \qquad \forall\, l \in L, p \in T_s \qquad (6.33)$$

$$s_l^p \in \{0, 1\} \qquad \forall\, l \in L, p \in T_s \qquad (6.34)$$

Table 6-2 shows all cliques of an intersection with eight lanes. Index $p \in T_s$ in variable s_l^p is removed for the sake of simplicity. The cliques have a size of 2, 3, and 4. The size of each clique is defined as the number of binary variables covered by that clique. Based on Table 6-2, the number of cliques decreases as the size of cliques increases. As shown in Table 6-2, the intersection graph with exclusive left-turn movements cannot have cliques of a size of larger than four. Therefore, any clique of size four is a maximal clique. In addition, it can be confirmed that the maximal cliques in the intersection conflict graph maintain all the necessary restrictions to prevent two conflicting movements to receive a non-red signal at the same time. Besides, the maximal cliques with size four are facet-defining for the convex hull of all feasible area from pairwise constraints (6.13) (Padberg, 1973). As a result, the feasible convex hull is tightened and the solutions of linear relaxation (Z_{LP}) get closer to the integer problem (Z_{IP}) as much as possible. In addition, the number of relaxed constraints is reduced significantly. Therefore, the duality gap reduces in Lagrangian relaxation problem (Z_{LR}).

Replacing the pairwise constraints with the maximal clique constraints enforces at most one of the conflicting approaches in the clique to take the non-red signal. Therefore, the signal and

trajectory optimization problem with the maximal clique constraints provides one of the tightest formulations in terms of LP relaxation since all of the node coefficients in the constraints are 0 or 1 (Murray and Church, 1996).

Table 6-2: All Cliques represented by the conflict graph of intersections with exclusive left-turn movement

Cliques with size 2	Cliques with size 3	Cliques with size 4
$s_1 + s_2 \leq 1$	$s_1 + s_2 + s_3 \leq 1$	$s_1 + s_2 + s_3 + s_4 \leq 1$
$s_1 + s_3 \leq 1$	$s_1 + s_2 + s_4 \leq 1$	$s_1 + s_2 + s_7 + s_8 \leq 1$
$s_1 + s_4 \leq 1$	$s_1 + s_2 + s_7 \leq 1$	$s_3 + s_4 + s_5 + s_6 \leq 1$
$s_1 + s_7 \leq 1$	$s_1 + s_2 + s_8 \leq 1$	$s_5 + s_6 + s_7 + s_8 \leq 1$
$s_1 + s_8 \leq 1$	$s_1 + s_3 + s_4 \leq 1$	
$s_2 + s_3 \leq 1$	$s_1 + s_7 + s_8 \leq 1$	
$s_2 + s_4 \leq 1$	$s_2 + s_3 + s_4 \leq 1$	
$s_2 + s_7 \leq 1$	$s_2 + s_7 + s_8 \leq 1$	
$s_2 + s_8 \leq 1$	$s_3 + s_4 + s_5 \leq 1$	
$s_3 + s_4 \leq 1$	$s_3 + s_4 + s_6 \leq 1$	
$s_3 + s_5 \leq 1$	$s_3 + s_5 + s_6 \leq 1$	
$s_3 + s_6 \leq 1$	$s_5 + s_6 + s_7 \leq 1$	
$s_4 + s_5 \leq 1$	$s_5 + s_6 + s_8 \leq 1$	
$s_4 + s_6 \leq 1$	$s_5 + s_7 + s_8 \leq 1$	
$s_5 + s_6 \leq 1$		
$s_5 + s_7 \leq 1$		
$s_5 + s_8 \leq 1$		
$s_6 + s_7 \leq 1$		
$s_6 + s_8 \leq 1$		
$s_7 + s_8 \leq 1$		

6.3.2.2. Primal and dual formulation

As described in the previous section, constraints (6.13) are replaced with constraints (6.32)-(6.34). As a result, the primal problem of signal timing and trajectory optimization will be represented as follows.

IP:

$$f' = \underset{\boldsymbol{a}}{\text{Max}}\left(\beta \sum_{t\in T_t}\sum_{l\in L}\sum_{i\in I^l} w_l(x_{il}^t) - (1-\beta)\sum_{t\in T_t}\sum_{l\in L}\sum_{i\in I_A^l}(z_i^t + u_i^t) - M\sum_{t\in T_t}\sum_{l\in L}\sum_{i\in I_H^l}(a_{il}^t - b_{il}^t)\right) \quad (6.35)$$

s. t.

(6.3)-(6.4), (6.9)-(6.12), (6.14)-(6.18),(6.20)-(6.28),(6.30)-(6.31), (6.32)-(6.34)

$$\sum_{l \in C_k} s_l^t \leq 1 \qquad \forall\, k \in K, p \in T_s \qquad (6.36)$$

$$s_l^p = g_l^p + y_l^p \qquad \forall\, l \in L, p \in T_s \qquad (6.37)$$

$$s_l^p \in \{0,1\} \qquad \forall\, l \in L, p \in T_s \qquad (6.38)$$

Relaxing complicating constraints (6.36) decomposes the primal problem into lane-level sub-problems. Therefore, the Lagrangian problem is obtained by dualizing clique constraints (6.36) as follows:

LR:

$$L(\boldsymbol{\mu}) = \underset{\boldsymbol{a}}{\text{Max}} \Bigg(\beta \sum_{t \in T_t} \sum_{l \in L} \sum_{i \in I^l} w_l (x_{il}^t) - (1-\beta) \sum_{t \in T_t} \sum_{l \in L} \sum_{i \in I_A^l} (z_i^t + u_i^t) - M \sum_{t \in T_t} \sum_{l \in L} \sum_{i \in I_H^l} (a_{i,l}^t - b_{i,l}^t) + \sum_{p \in T_s} \sum_{k \in K} \sum_{l \in C_k} \mu_{lk}^p (1 - s_l^p) \Bigg) \qquad (6.39)$$

s. t.

(6.3)-(6.4), (6.9)-(6.12), (6.14)- (6.18),(6.20)-(6.28),(6.30)-(6.31), (6.37)-(6.38)

where $\mu_{lk}^p \in R^+$ is the Lagrangian multiplier for lane $l \in C_k$ that belongs to clique $k \in K$ at time step $p \in T_s$. Vector $\boldsymbol{\mu}$ is defined as the vector of all Lagrangian multipliers. Since the objective function and the remaining constraints of problem LR are separable over lanes, the sub-problem for each lane $l \in L$ can be solved separately in parallel when the dual multiplier $\boldsymbol{\mu}$ is available. The values of the dual function $L(\boldsymbol{\mu})$ at dual feasible point $\boldsymbol{\mu}$ are always upper-bounds

to the optimal value f'. Hence, the sharpest upper-bound can be found from the optimal value of the dual problem (6.40), which is defined as $\boldsymbol{\mu}^*$.

$$L(\boldsymbol{\mu}^*) = \min_{\boldsymbol{\mu} \in R} L(\boldsymbol{\mu}) \tag{6.40}$$

By duality theory, dual problem (6.40) is always convex. In other words, solving Lagrangian dual problem (6.40) is equivalent to minimizing a convex piecewise linear function. A function $f: R^n \to R$ is a piecewise linear convex function if f is attained as the maximum of the finite number of affine functions $f_i: R^n \to R$. We can use this feature to find the optimal Lagrangian multipliers $\boldsymbol{\mu}$ through the dual cutting plane method.

6.3.2.3. Update Lagrangian multipliers

The subgradient method is a common approach to solve the Lagrangian dual problem and update the Lagrangian multipliers. However, it suffers from a slow convergence (Kappes et al., 2012; Zhao and Luh, 2002). The subgradient method utilizes the information of only the last iteration to update the Lagrangian multipliers. On the other hand, using the dual cutting plane method helps store the information of all previously found Lagrangian multipliers ($\boldsymbol{\mu^n}$), the optimal Lagrangian relaxation function $L(\boldsymbol{\mu^n})$, and the subgradient $r(\boldsymbol{\mu^n})$ up to iteration n and use them to find the new Lagrangian multiplier $\boldsymbol{\mu^{n+1}}$ at next iterations (Nowak, 2005). The subgradient r_k^p corresponding to each relaxed constraint is found from (6.41).

$$r_k^p = 1 - \sum_{l \in C_k} s_l^t \qquad \forall\, k \in K, p \in T_s \tag{6.41}$$

Based on the definition of subgradient, inequality (6.42) is satisfied for all $\boldsymbol{\mu}$.

$$L(\boldsymbol{\mu}) \geq L(\boldsymbol{\mu^n}) + r(\boldsymbol{\mu^n})^T(\boldsymbol{\mu} - \boldsymbol{\mu^n}) \qquad \forall\, \boldsymbol{\mu} \in R^n \tag{6.42}$$

The cutting plane optimization problem that solves dual problem (6.40) overtime is presented in (6.43) and (6.44). Solving this optimization problem over iterations provide a sequence of Lagrangian multipliers $\{\boldsymbol{\mu^n}\}_{n=1,\dots,N}$. Iteration N is considered as the last iteration in the cutting plane optimization.

CP:

$$\min_{\mu \in R} v \tag{6.43}$$

s. t.

$$v \geq L(\boldsymbol{\mu^n}) + r(\boldsymbol{\mu^n})^T(\boldsymbol{\mu} - \boldsymbol{\mu^n}) \qquad n = \{1, \dots, N\} \tag{6.44}$$

Although the cutting plane method theoretically works, there are some drawbacks that prevent its successful application in practice. Objective function (6.43) is unbounded in the first few iterations, and there would not exist enough constraints to find the optimal Lagrangian multipliers. Moreover, the convergence rate of the cutting plane method is slow (Kiwiel, 1990). To fix the mentioned issues, the stabilized version of the cutting plane is introduced as the proximal bundle method (Kiwiel, 1990). Similar to the cutting plane method, a polyhedral model of dual function (6.46) is considered in the proximal bundle approach. In addition, a quadratic penalty term is added to the objective function (6.45) to stabilize the optimal Lagrangian multiplier around the center point $\boldsymbol{\mu_{cp}}$. The center point is considered as the best Lagrangian multiplier that has been found so far that significantly improves the solutions. Parameter $h \in R^+$ controls the weight of the quadratic term.

BM:

$$\min_{\mu \in R} v + \frac{1}{2h} \left\| \boldsymbol{\mu} - \boldsymbol{\mu_{cp}} \right\|^2 \tag{6.45}$$

s. t.

$$v \geq L(\boldsymbol{\mu^n}) + r(\boldsymbol{\mu^n})^T (\boldsymbol{\mu} - \boldsymbol{\mu^n}) \qquad n = \{1, \dots, N\} \tag{6.46}$$

6.3.2.4. Making infeasible solutions feasible

Solving the Lagrangian relaxation problem iteratively converges to $\boldsymbol{\mu}^*$, $\mathbf{g}^*$, and $\mathbf{y}^*$ as the optimal Lagrangian multipliers, green signal status, and yellow signal status, respectively. In addition, it yields the optimal trajectory of all CAVs. Although reformulating the problem reduces the duality gap to a very small value, the optimal solution of Lagrangian relaxation still may be an infeasible signal timing parameter, that does not satisfy relaxed constraints (6.36). We introduce a simple optimization problem that finds a good feasible signal timing solution in such conditions. Reformulating constraints (6.13) with (6.36) pushes the solution of signal timing variables from Lagrangian relaxation problem to close proximity of the feasible integer values in the convex polyhedron. As a result, there would be no need to change the structure of the Lagrangian relaxation solution. We need to project the final infeasible signal timing solutions to the closest feasible integer point. The complementary optimization problem *FP* is introduced in (6.47)-(6.53) to ensure the feasibility of the solutions from the relaxed problem.

The decision variables in FP are all signal timing variables. In addition, $\mathbf{g}^*$ and $\mathbf{y}^*$ found from the Lagrangian relaxation are inputs. The objective function of the proposed optimization problem minimizes the difference between signal timing variables and the corresponding solutions from the Lagrangian relaxation. In addition, all signal timing constraints, including the relaxed

pairwise constraints, are considered in FP. Constraints (6.49) are also added to the problem, assuming that the optimal signal timing parameters always give non-red signal status to at least two movements at the same time. The optimization problem *FP* is always feasible if the prediction horizon of signal timing T_s is greater than the minimum green time for all approaches of the intersection.

FP:

$$\min_{g,y\in\{0,1\}} \sum_{l\in L}\sum_{p\in T_s} \left\|g_l^p - \boldsymbol{g}_l^{*p}\right\|^2 + \left\|y_l^p - \boldsymbol{y}_l^{*p}\right\|^2 \tag{6.47}$$

s. t.

$$g_l^p + g_{l'}^p + y_l^p + y_{l'}^p \le 1 \qquad \forall\, l \in L,\ l' \in C_l, p \in T_s \tag{6.48}$$

$$\sum_{l\in L} g_l^p + y_l^p \ge 2 \qquad \forall\, l \in L, p \in T_s \tag{6.49}$$

$$\sum_{j=p}^{p+\omega_l'} g_l^j \le \omega_l' \qquad \forall\, l \in L, p \in T_s \tag{6.50}$$

$$\sum_{j=p+1}^{p+\omega_l} g_l^j \ge \left(g_l^{p+1} - g_l^p\right)\omega_l \qquad \forall\, l \in L, p \in T_s \tag{6.51}$$

$$\sum_{j=p}^{p+Y} y_l^j \le Y \qquad \forall\, l \in L, p \in T_s \tag{6.52}$$

$$\sum_{j=p+1}^{p+Y} y_l^j \ge \left(g_l^p - g_l^{p+1}\right)Y \qquad \forall\, l \in L, p \in T_s \tag{6.53}$$

6.3.3. Receding horizon control

The Lagrangian relaxation for finding the optimal signal and trajectory optimization problem is embedded in a receding horizon control (RHC) to take into account the dynamic nature of the problem. Figure 6-6 shows the general framework. First, we initialize the problem with initial Lagrangian multipliers at time step 0. Then, the corresponding signal controller for each lane group of the intersection solves sub-problem *LR*, and then they share solutions to update Lagrangian multipliers form solving the proximal bundle optimization problem *BM*. The convergence criteria are evaluated with calculating the difference between the upper- and lower-bounds of original problem *IP*. Then the feasibility of the signal timing parameters is checked. If the solutions are infeasible, the optimization problem *FP* is solved, which updates the optimal CAV and HV trajectories based on the feasible signal plan. If the solutions from Lagrangian relaxation are feasible, they are the optimal solutions to the original problem and there is no need to solve *FP*. RHC implements the trajectory and signal timing plans at the first time step into Vissim micro-simulator. Then, the Lagrangian multipliers will be updated from the last updated solutions of Lagrangian relaxation, and the planning horizon rolls one time step forward until the study period is finished.

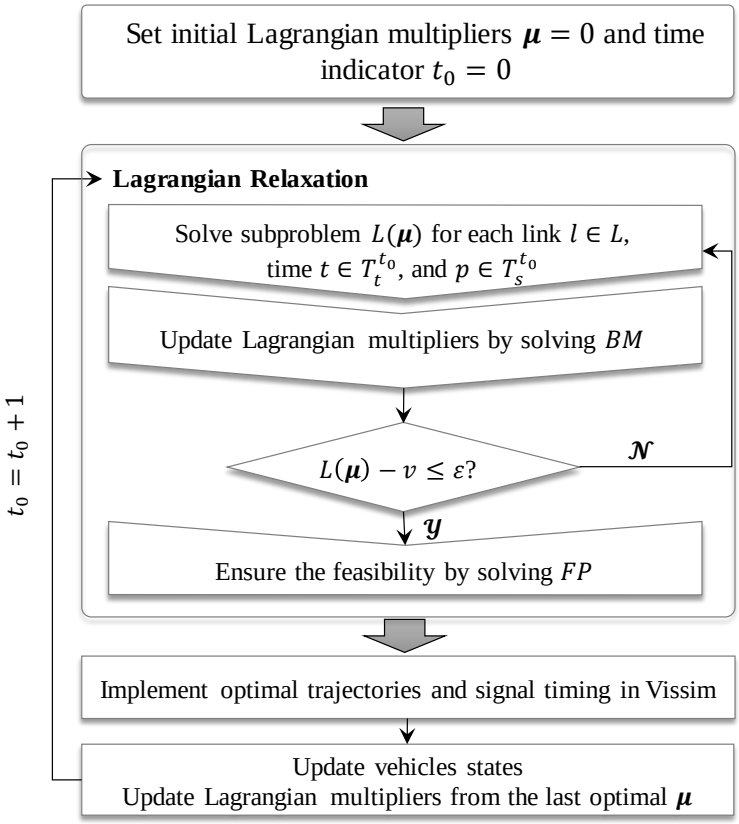


Figure 6-6: Lagrangian relaxation procedure is embedded in a receding horizon framework

6.4. Case study

We applied the proposed solution technique to an isolated four-leg intersection with exclusive left-turn lanes, as shown in Figure 6-5. It is assumed that vehicles are on the desired lane before arriving the intersection neighborhood. The detection range is 1000 ft before and after the intersection. The signal status is updated every two seconds, while vehicle acceleration, speed, and position are updated every 0.5 seconds. The prediction horizon of RHC is 20 seconds. More details are provided in Table 6-3. The STTO problem is solved for a 15-minute study period.

Table 6-3: Case study parameters

Parameters	Value
Desired speed (ft/s)	44
Maximum acceleration (ft/s^2)	6.5
Maximum deceleration (ft/s^2)	11.5
Human driver reaction time (s)	1
Self-driving vehicle reaction time (s)	0.1
Detection range (ft)	1000
Average length of the vehicles (ft)	13
Safety distance between following vehicles (ft)	12
Car-following parameter α_1	0.95
Car-following parameter α_2	0.25
Trajectory updating interval (s)	0.5
Signal updating interval (s)	2
Minimum green time for through movement (s)	12
Minimum green time for left turns (s)	6
Maximum green time for through movement (s)	60
Maximum green time for left turns (s)	20
Yellow time interval (s)	4
Planning time horizon (s)	20
Study period (min)	15

Table 6-4 summarizes different scenarios tested in this study. For each scenario, six different market penetration rates of CAVs (i.e., 0%, 20%, 40%, 60%, 80% and 100%) are considered to evaluate the effects of cooperative signal timing and trajectory optimization planning. Vissim microsimulation (PTV Group, 2015) is used to test the proposed algorithm. The proposed algorithm is coded in Java running on a desktop computer with inter core i-9-9900 CPU and 64 GB memory. CPLEX (CPLEX, 2009) is used to solve the MILP optimization problem.

Table 6-4: Demand patterns in STTO case study

	East and Westbound		North and Southbound	
Scenario	Through demand $(veh/h/ln)$	Left-turn demand $(veh/h/ln)$	Through demand $(veh/h/ln)$	Left-turn demand $(veh/h/ln)$
1	500	100	500	100
2	600	120	600	120
3	750	150	750	150
4	900	180	900	180
5	900	180	750	150

6.5. Results

The Helly car-following model provides a linear description of driving behavior, which is suitable for the purpose of optimization algorithm. Analyzing the performance of Helly's car-following model shows that it provides very well behavior in terms of stability. We have shown the following behavior of seven consecutive vehicles within a platoon in Figure 6-7, where the first vehicle is forced to stop behind the red-light at the intersection and then start over moving forward. Five plots are provided below to show the location, speed, acceleration, relative location, and relative speed among these seven vehicles. We can see that the propagation of relative location and speeds are pretty much stable.

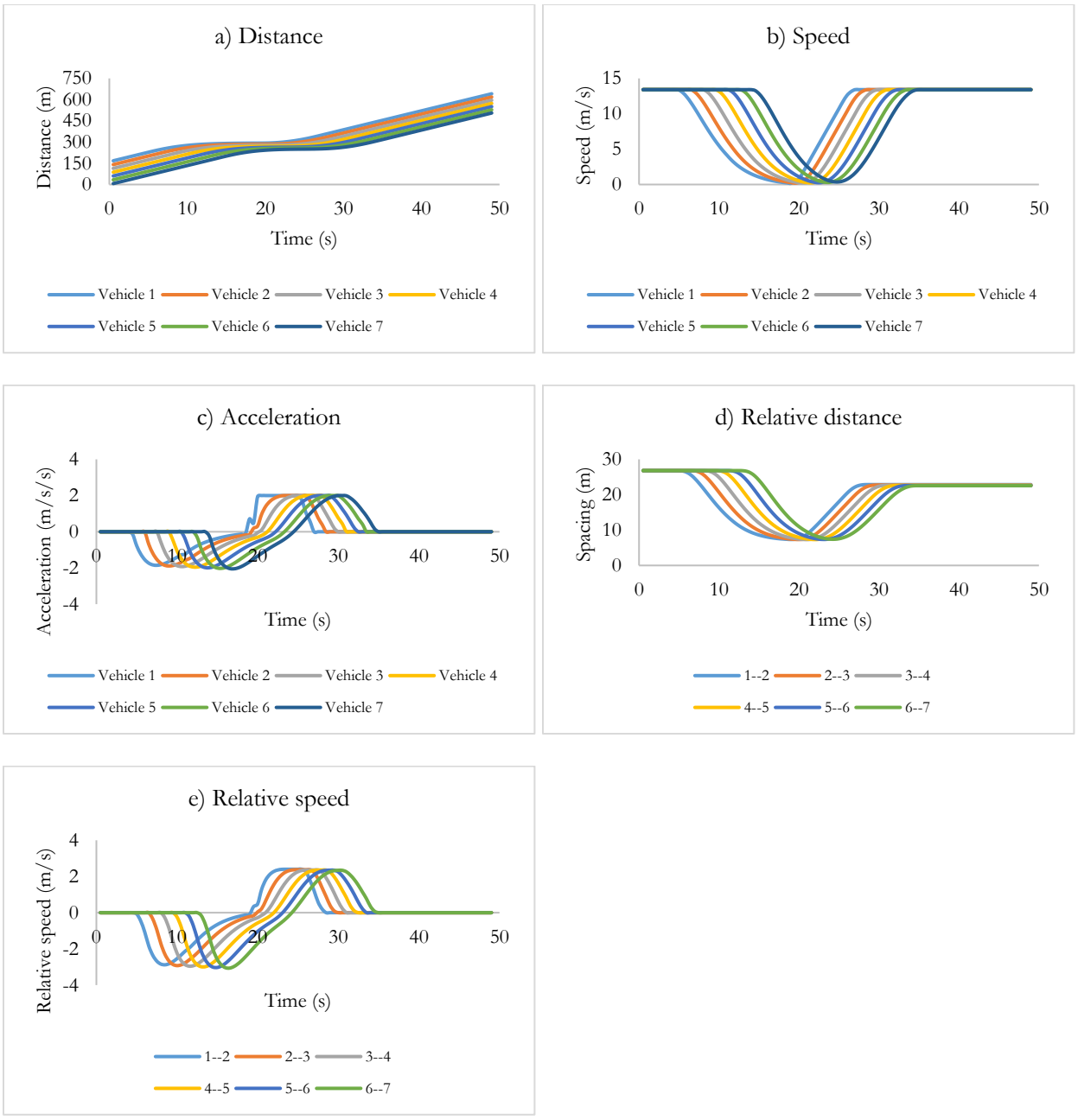


Figure 6-7: The stability performance of Helly's car-following model

Figure 6-8 shows the duality gaps in scenario 4, the highest demand level tested in this study, as a result of solving STTO problem with the proposed Lagrangian relaxation technique when pairwise conflicting constraints (6.13) are replaced with tighter cliques. Since the problem is solved through receding horizon control, the Lagrangian relaxation problem is solved every two seconds to find the optimal signal timing plans. Hence, the duality gaps are reported overtime for

all relaxed problems that are solved dynamically. Figure 6-8 indicates that the duality gap is mostly zero, meaning that the proposed solution technique mostly holds the strong duality. Moreover, it is shown that the duality gaps are always less than 0.1% in all cases tested with different CAV market penetration rates in scenario 4. Note that results for other scenarios also confirm the same patterns.

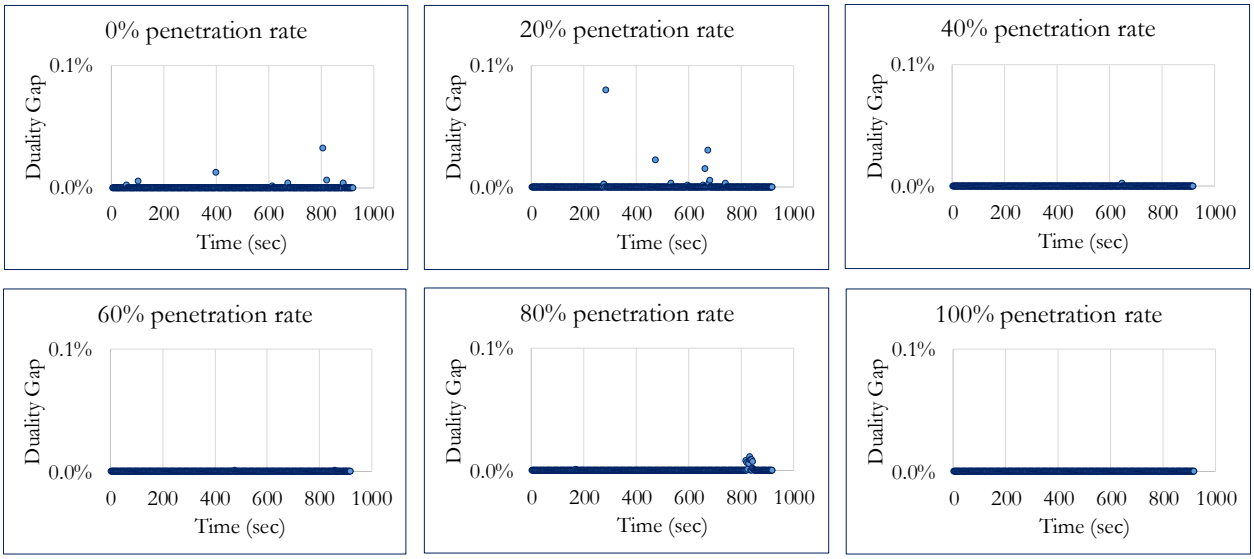


Figure 6-8: Duality gaps for Lagrangian relaxation after reformulation

Table 6-5 shows the average computational runtime of the proposed model based on scanrios 1 to 4 and different market penetration rate of CAVs. The average runtime is shown for two optimization programs where (1) signal and trajectories are optimized simultaneously, and (2) only trajectories are optimized with fixed signal timing parameters. In both cases, the optimization runtime decreases when CAV market penetration rate increases since the number of variables associated with car-following model of HVs decreases. It is also shown that increasing the traffic volume is associated with higher computational runtime. The minimum and maximum average runtime for signal and trajectory optimization with enhanced Lagrangian relaxation technique was 0.8 s and 5.9 s, respectively. It should be noted that the proposed method is an inetrative approach and might not find the optimal solution in real time since the signal timing parameters are

optimized each 2 s. However, real time solutions are achievable by setting the optimality gap to higher values or updating signal timing parameters in higher intervals. Table 6-5 also shows that the minimum and maximum runtime for trajectory optimization with fixed signal parameters are 0.04 s and 0.22 s, respectively.

Table 6-5: The Average runtime for different scenarios and market penetration rate of CAVs (second)

CAV Penetration rate	(1) Signal and trajectory optimization				(2) Trajectory optimization with fixed signal			
	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 1	Scenario 2	Scenario 3	Scenario 4
0%	3.2	3.9	5.2	5.9	0.10	0.12	0.19	0.22
20%	2.3	3.3	4.8	5.4	0.09	0.12	0.17	0.21
40%	1.8	2.7	3.5	4.1	0.08	0.10	0.14	0.17
60%	1.6	1.8	3.1	3.4	0.07	0.08	0.12	0.19
80%	1.3	1.4	2.8	2.2	0.05	0.07	0.12	0.13
100%	0.8	1.0	1.5	1.8	0.04	0.06	0.08	0.10

Table 6-6 compares the average travel time for different CAV market penetration rates of the STTO strategy with three state-of-practice signal control approaches: (a) fixed-time (Hajbabaie, 2012; Hajbabaie and Benekohal, 2015, 2013), (b) actuated, and (c) adaptive signal control. The fixed-time signal control can be seen as a baseline, where signal timing parameters are optimized based on a prediction on demand levels in different times of day. Fixed-time signal control is not responsive to unforeseen demand variations. The actuated control utilizes vehicle detectors and can react to changes in traffic demand as they are observed. However, does not predict near future conditions. The adaptive signal control can predict traffic conditions in near future and can proactively change signal timing parameters. We used PTV Vistro (America, 2014) to find the optimal fixed-time and actuated signal timing plans. Moreover, the results of an adaptive signal control methodology based on the Cell Transmission Model (Al Islam et al., 2020; Al Islam and Hajbabaie, 2017) is provided for further comparisons.

Table 6-6: Comparing average travel time between actuated signal timing and STTO

Scenario	Fixed-time signal control	Actuated signal control	Adaptive signal control	STTO with different penetration rates					
				0%	20%	40%	60%	80%	100%
1	87.7	84.3	75.6	71.2	66.1	65.7	64.3	63.7	62.2
2	104.3	97.1	85.9	74.1	72.4	67.9	64.5	63.7	62.9
3	141.3	130.6	117.3	105.6	95.1	85.2	76.2	73.0	70.9
4	185.1	171.1	137.8	117.6	109.1	92.4	80.8	73.6	66.4
5	166.2	155.0	130.3	112.6	101.9	90.8	81.3	73.3	70.1

The results show that STTO yields significantly shorter average travel time compared to the fixed time, actuated, and adaptive signal control in all scenarios with different CAV penetration rates. The results also show that the average travel time decreases as the CAV penetration rates increase, which is as expected. However, this reduction is less significant in low traffic volumes. For instance, the average travel time is almost the same when the CAV penetration rates are 60%, 80%, and 100% in scenarios 1 and 2. On the other hand, increasing the CAV penetration rate in scenarios with higher demand levels (e.g., scenario 4) still improves the performance of the intersection by reducing the average travel time of all vehicles. Moreover, we compared the performance of STTO with fixed-time signal timing plan when the market penetration rate of CAVs changes. Scenario 4 was considered for this analysis since it has the highest demand. When signal timing is fixed, only the trajectory of CAVs is optimized to utilize the green time efficiently and reduce the number of stops at the intersection. Figure 6-9 shows that STTO always outperforms the fixed-time signal timing plan in terms if average travel time. In addition, the rate of decrease in travel time is higher with STTO compared to the fixed-time plan when the market penetration rate of CAVs increases.

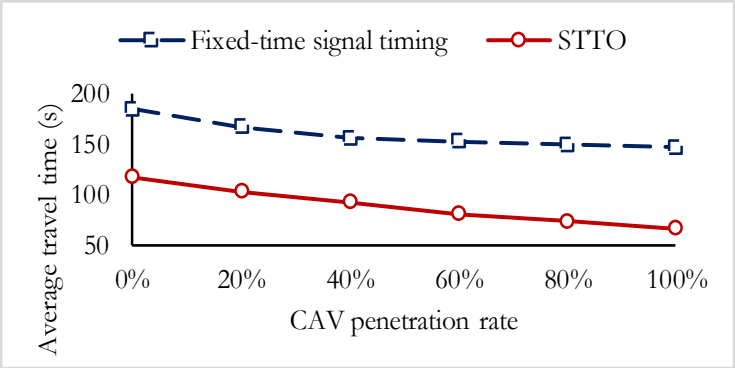


Figure 6-9: Comparing the performance of STTO with fixed-time signal for various market penetration rate of CAVs

Having a minimum green time constraint prevents a further reduction of average travel time in high CAV penetration rates when traffic volume is low. CAVs can be processed very efficiently and do not require a long green duration. Therefore, a long minimum green duration will reduce intersection efficiency. We reduced the minimum green times from 12 seconds to 4 seconds for through movement, and from six seconds to two seconds for left turn movements. Figure 6-10 compares the average travel time for different CAV penetration rates before and after reducing the minimum green time for scenarios 1-5. As the figure shows, increasing the market penetration rate of CAVs further improves the intersection performance when the minimum green duration is shortened and as such, its corresponding constraint is not binding. In addition, the effect of reduced minimum green time is higher on improving the average travel time of scenarios with lower demand since vehicles require lower green time to proceed.

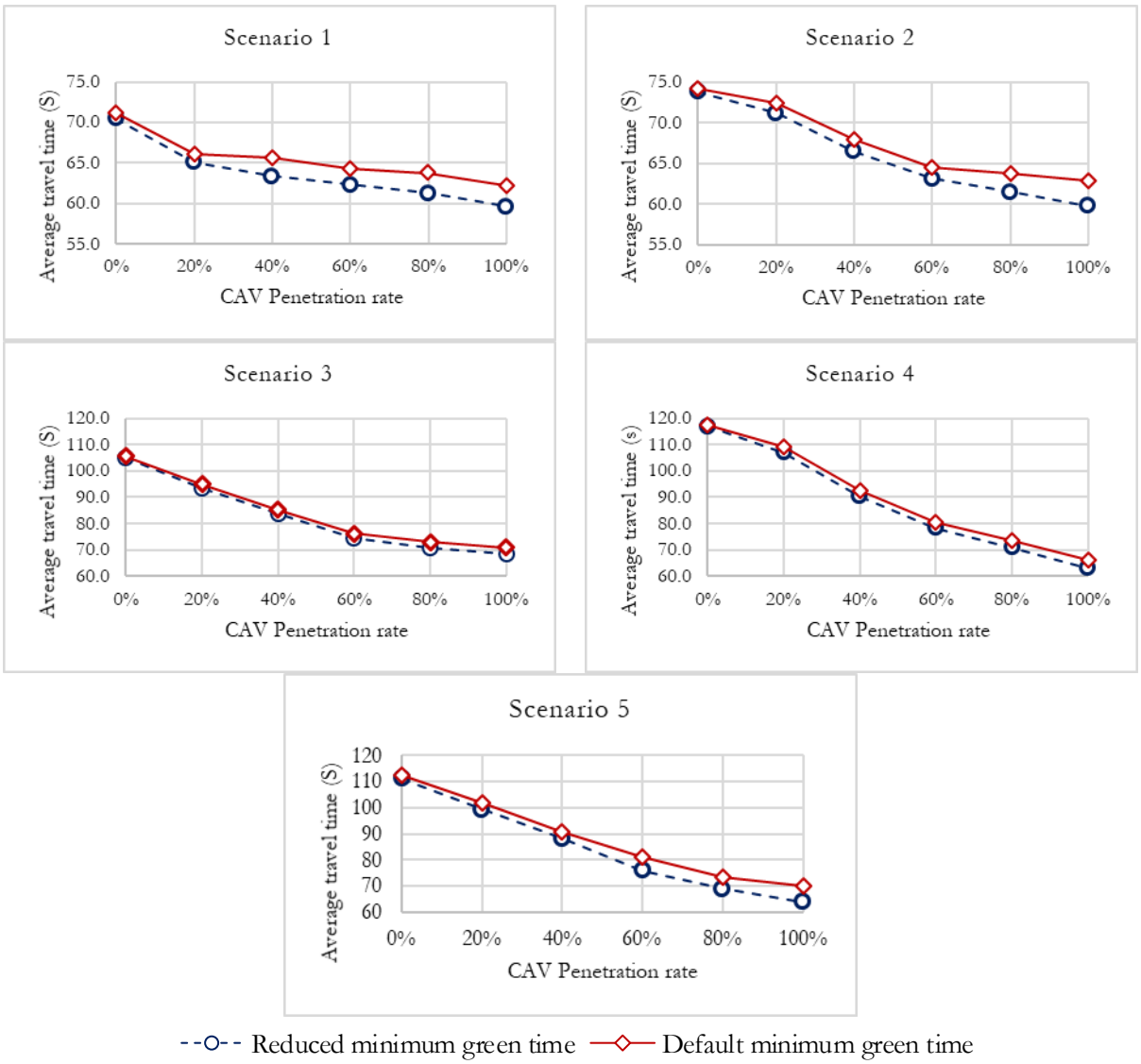


Figure 6-10: Comparing the average travel time before and after reducing the minimum green time in four scenarios

Figure 6-11 shows the average queue length in all movements in different approaches of the case study intersection with different CAV penetration rates in scenario 4, which has the highest tested demand level. The average queue length decreases with an increase in CAV market penetration rate. The average queue length is about 450 ft when there is no CAV on the road to control the movement of human-driven vehicles. The average queue length reduces to less than 50 ft when all vehicles are CAVs.

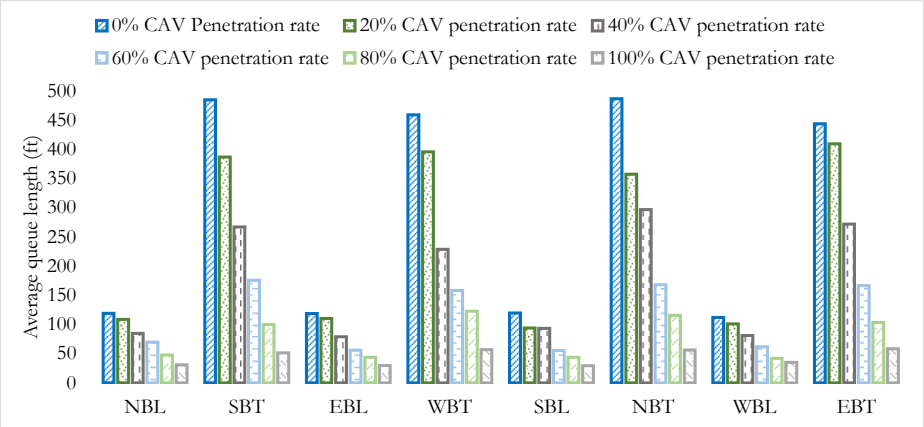


Figure 6-11: Average queue length for different CAV penetration rates in scenario 4

Figure 6-12 shows the signal timing parameters for all movements in different approaches of the intersection with 0% and 100% CAV penetration rates in scenario 4 with the highest tested demand level. The green period for all movements of the intersection is longer when the CAV penetration rate is zero. This is due to the high start-up-lost time of human-driven vehicles. On the other hand, the green period is shorter and more frequent when all vehicles are CAVs. This is due to a low start-up lost time of CAVs when they are stopped before the intersection. Moreover, CAVs adjust their speed to pass the intersection with the maximum speed; hence, they require shorter green durations to pass the intersection.

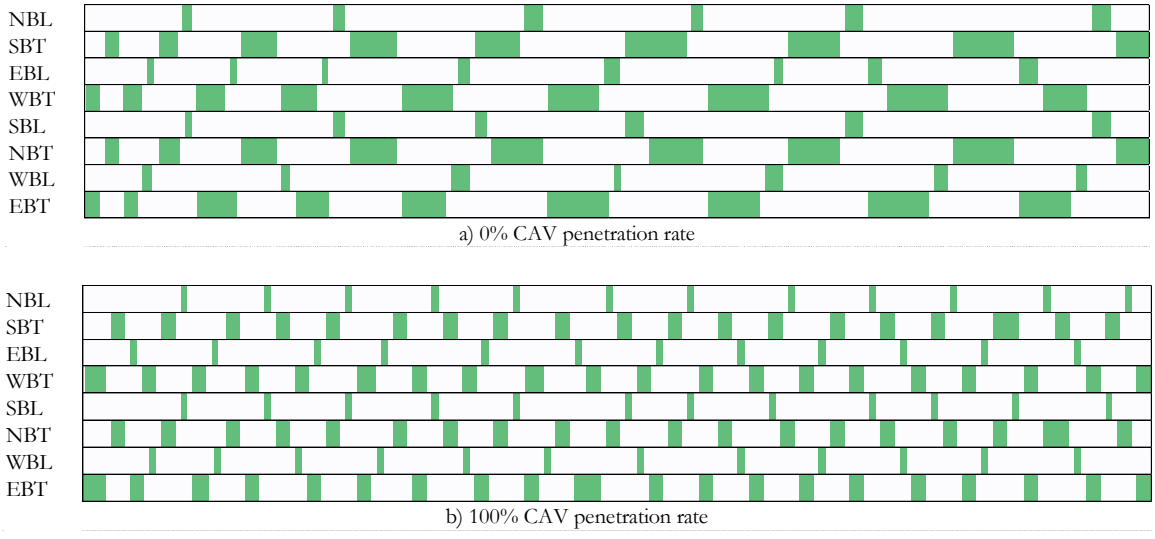


Figure 6-12: Comparing signal timing plans under 0% and 100% CAV penetration rate

Figure 6-13 shows the trajectory of CAVs and HVs on eastbound through (EBT) for different market penetration rates in Scenario 4 with the highest tested demand. Increasing the penetration rate of CAVs is associated with a smoother trajectory for all vehicles to travel through the intersection. In addition, the back of the queue gets closer to the intersection as the market penetration rate of CAVs increase. Moreover, Figure 6-14 compares the trajectory of vehicles on two conflicting movements (i.e. EBT and SBT) for scenarios with 40% and 100% CAV penetration rates to show that the operation of one approach is not associated with the poor performance of the conflicting approaches. Figure 6-15 also provide the cumulative distribution function for each approach of the intersection. These distributions confirm the results of Figure 6-14 and show that the delay experience in different approaches of the intersection are within the same range. Note that the delay experienced in left-turning movement is higher than through movement, but they are consistent in different scenarios.

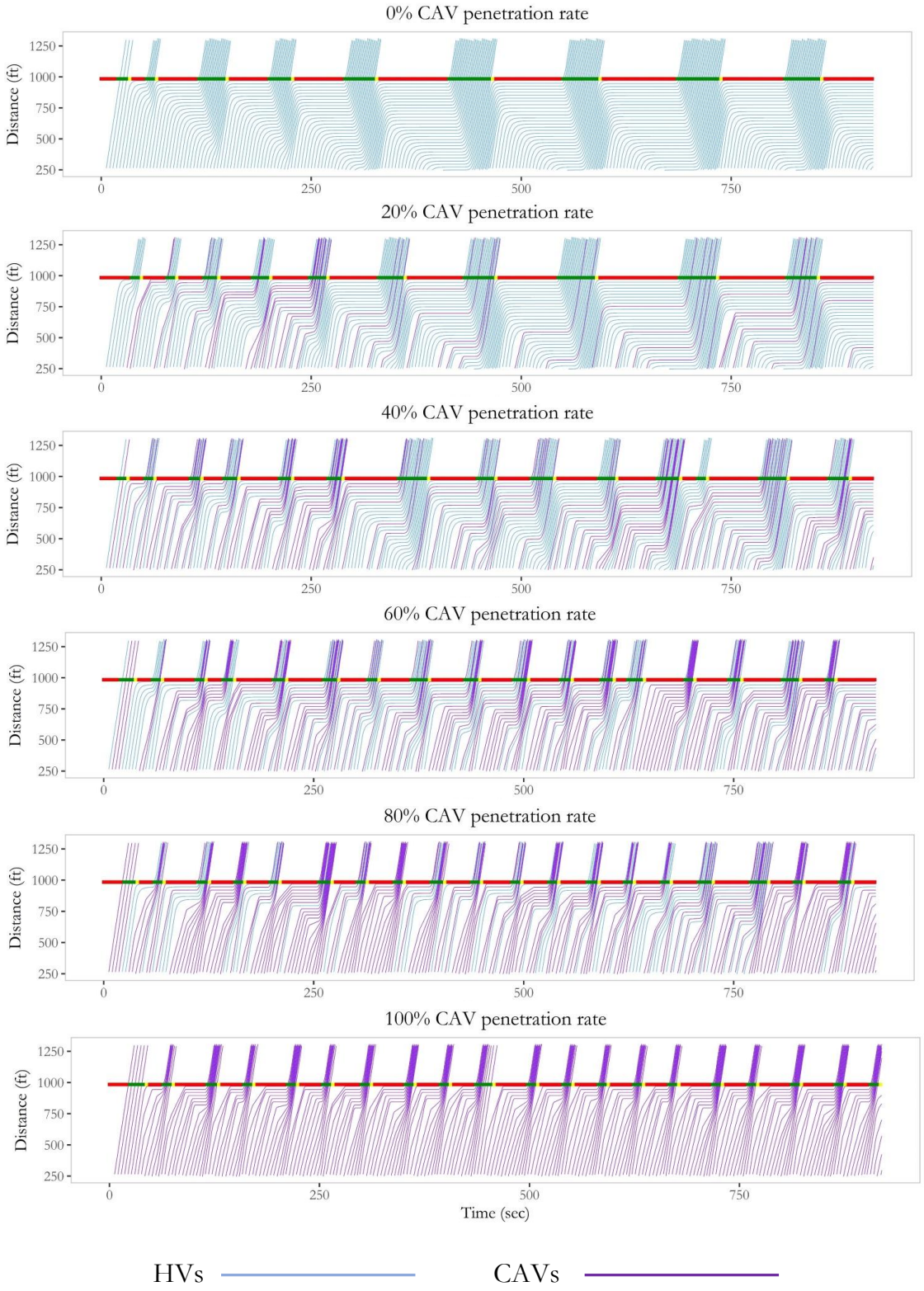


Figure 6-13: The trajectory of CAVs and HVs in Lane 1 under different market penetration rates of CAVs

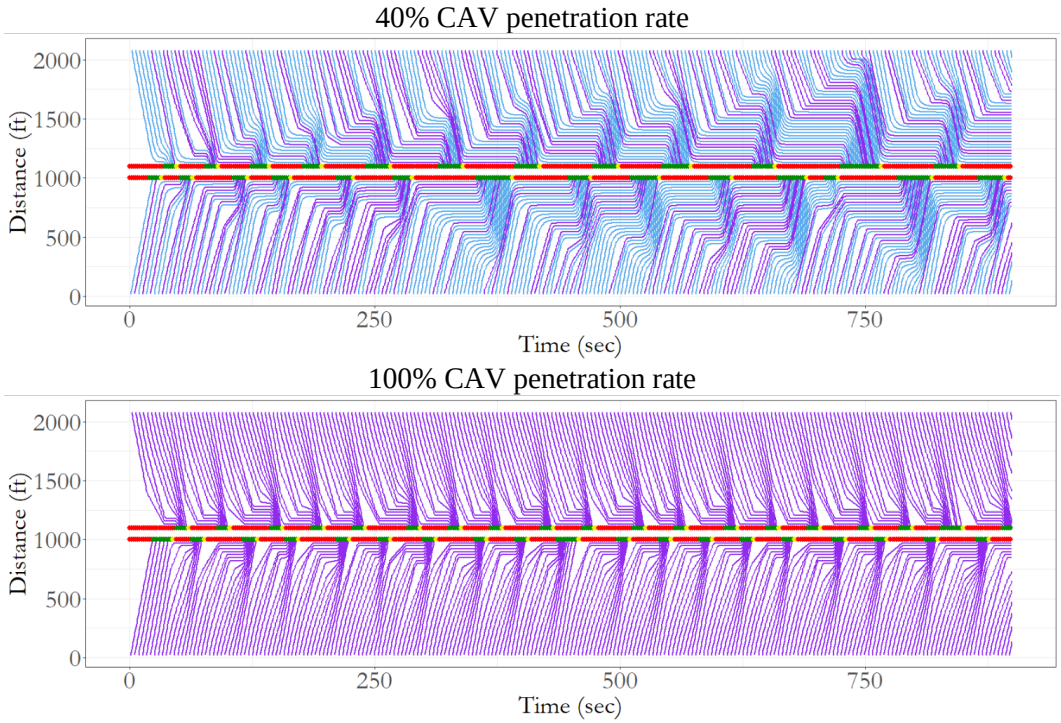


Figure 6-14: Comparing the trajectory of conflicting approaches of the intersection

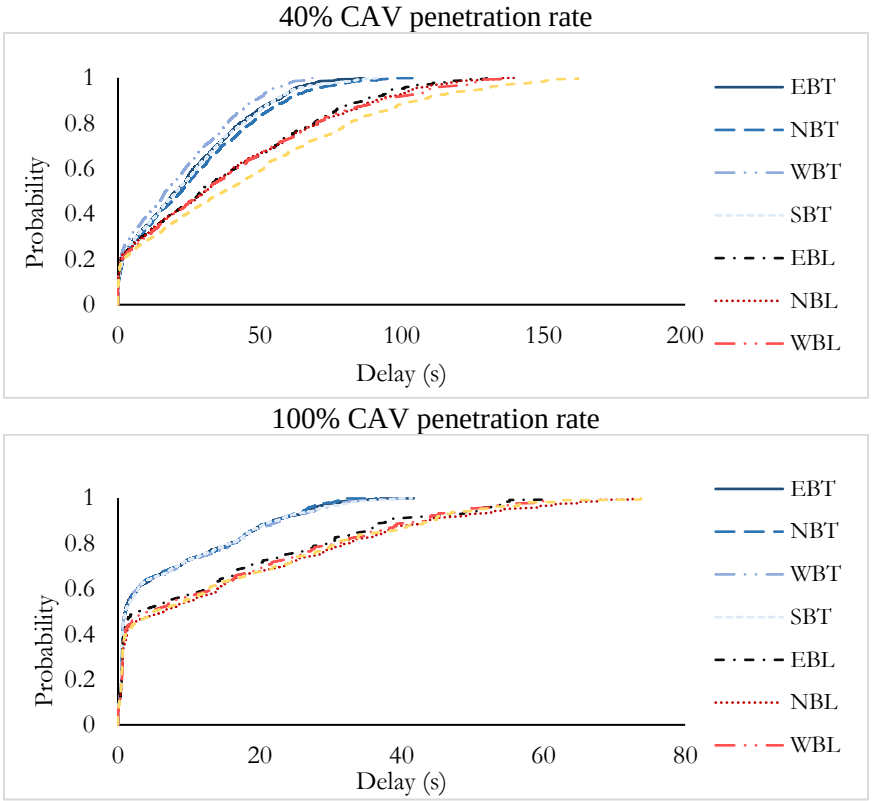


Figure 6-15: Comparing the cumulative delay distribution of different approaches of the intersection

Since none of current state-of-practice strategies jointly optimize the signal timing plans with the trajectory of CAVs, we have compared the result of STTO with a recently developed state-of-art strategy in Guo et al. (2019). This study optimizes the signal timing plan and trajectory of CAVs in a mixed environment using dynamic programming and shooting heuristic. The findings show that the proposed algorithm in our paper could find lower average travel time and fuel consumption with lower computational times. The fuel consumption is calculated based on VT-Micro model (Ahn et al., 2013) with similar parameters to Ma et al. (2017). The following parameters are set exactly the same as the studied case in Guo et al. Guo et al. (2019): Length of intersection link: 1312 ft, Saturation rate f_s: 0.6, Maximum speed for through movement: 98 ft/s, Maximum speed for left turn movements: 79 ft/s, Planning horizon: 122 s, Car following parameter: maximum acceleration= 4.72 ft/S^2, maximum deceleration= 5.48 ft/S^2, and Step size to optimize signal timing plans: 8 seconds.

Table 6-7 shows that the average travel time and fuel consumption is reduced by increasing the market penetration rate of CAVs. In addition, our proposed algorithm performs better than the best results in Guo et al. (2019) for different market penetration rates of CAVs. It should be noted that Guo et al. (2019) assumed that there are four phases for signal timing plans. However, our study includes eight phases, which represent a more general condition. Moreover, phases in Guo et al. (2019) are considered in fixed order in a cycle even though some phases could be skipped. On the other hand, in our study, no order for phases are considered. Therefore, these differences in parameters could potentially affect the result of this comparison. All these differences created a larger feasible area for our problem, which led to the significant difference in the solution quality presented in Table 6-7.

Table 6-7: Comparing with Guo et al. (2019)

CAV Penetration rate	Average travel time (s)		Average fuel (lit/veh)	
	This study	Guo et al.	This study	Guo et al.
0%	72.81	114.10	0.150	0.180
20%	68.25	96.46	0.147	0.149
40%	65.29	75.06	0.145	0.146
60%	55.47	75.06	0.140	0.146
80%	49.49	75.06	0.123	0.146
100%	43.88	75.06	0.100	0.146

Table 6-8 shows average travel time, average fuel consumption, and total runtime obtained by STTO for 5, 10, 15, and 30 minutes of study period. The performance measures are provided for scenario 4, which has the highest demand, and two CAV market penetration rates of 40% and 100%. The trends in Table 6-8 shows that increasing the study period is associated with higher average travel time and average fuel consumptions due to presence of more vehicles in the network. In addition, the runtime increases by increasing the study period due to higher number of variables and memory usage.

Table 6-8: STTO performance with different study periods

Study period (min)	40 % CAV penetration rate			100 % CAV penetration rate		
	Average travel time (s)	Average fuel (lit/veh)	Total Runtime (min)	Average travel time (s)	Average fuel (lit/veh)	Total Runtime (min)
5	68.4	0.134	11.15	60.8	0.115	5.85
10	80.2	0.144	24.97	65.3	0.124	12.33
15	83.5	0.147	38.82	66.7	0.127	22.06
30	85.1	0.151	71.95	67.4	0.129	59.47

The saturation flow rate of traffic flow with the presence of CAVs highly depends on selecting the parameters in defining the behavior of CAVs. The parameter τ_A in (6.9) affects the reaction time of CAVs to keep a safe distance with the leading vehicle. To show how the saturation flow rate at the intersection will change based on different rate of CAV penetration rates, we performed a sensitivity analysis on the reaction time parameter. Four values are selected for the CAV reaction time as 0.1s, 0.4s, 0.7s, and 1s. The reaction time of human drivers is fixed as 1s. Figure 6-16 shows that lower reaction time is associated with lower saturation headway at the intersection. In addition, Figure 6-17 shows that the intersection throughput is higher when the reaction time decreases. The figures also show that the rate of mobility improvement with various penetration rate of CAVs is dependent on the reaction time. In other words, the presence of CAVs with lower reaction time in the network provides better mobility.

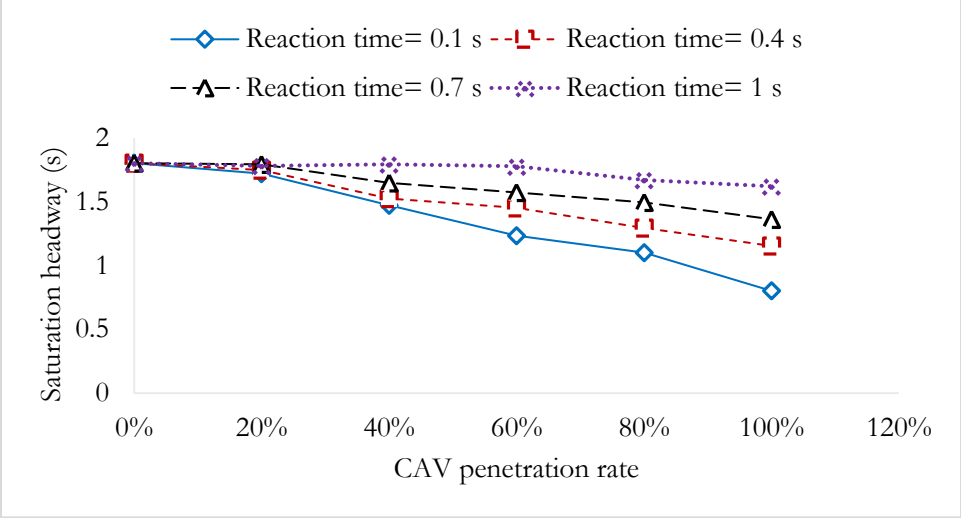


Figure 6-16: Comparing the saturation headway for different reaction time of CAVs

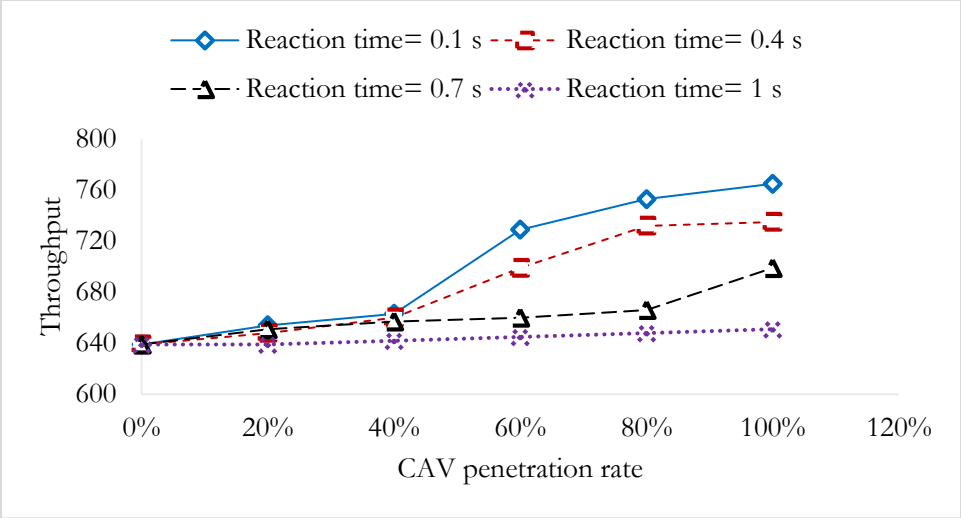


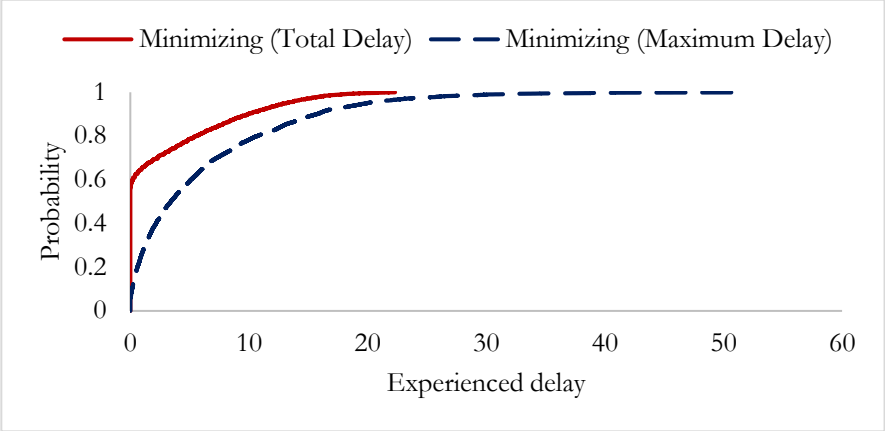
Figure 6-17: Comparing the saturation headway for different reaction time of CAVs

We tried a new objective function for signal timing and trajectory optimization problem by minimizing the maximum delay at the intersection. This objective function avoids experiencing a very high delay in one approach of the intersection. Table 6-9 shows the results of running the optimization algorithm with new objective function in comparison with the original objective function (i.e. minimizing total delay) for the case with 100% CAV penetration rate. The results show that minimizing the total delay outperforms minimizing the maximum delay in terms of network-wide performances. However, the maximum delay that a vehicle experienced is much lower when the maximum delay is minimized.

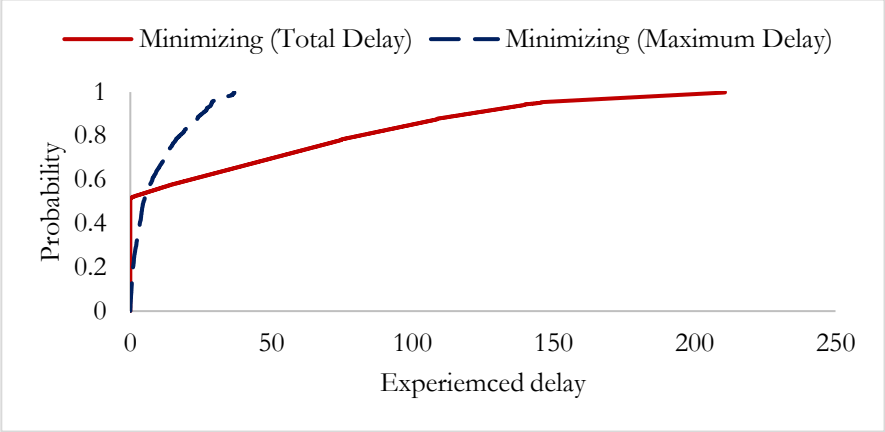
Table 6-9: STTO performance with two different objective functions

Performance measure	Minimizing (Total Delay)	Minimizing (Maximum Delay)	Difference
Average delay (s)	8.11	9.35	15.3 %
Maximum delay (s)	210.8	51.9	-75.3%
Average speed (mph)	18.41	15.24	-17.2 %
Throughput	252	241	4.4 %

Figure 6-18-a and Figure 6-18-b compare the performance of two different objective functions by showing the cumulative experienced delay distributions for through movement and left-turn movements, respectively. The through movements are processed quickly when the total delay is minimized in the objective function due to higher traffic volume and therefore, receiving higher rates in the objective. On the other hand, the delay of left-turning movements is higher when the total delay is minimized in the objective function due to lower traffic volume. However, minimizing the maximum delay considers the trade-off between the through movements and left-turning movements to avoid experiencing high delay for left-turning vehicles.



a) Cumulative delay distribution for through movements



b) Cumulative delay distribution for left-turning movements

Figure 6-18: Comparing Cumulative delay distributions

Figure 6-19 also compares the trajectory of vehicles based on two different objective functions. Minimizing the total delay pushes vehicles to stay closer to the intersection stop bar even though the traffic light is red. On the other hand, minimizing the maximum delay makes the trajectories smoother and avoids vehicles from stopping at the intersection. In other words, vehicles adjust their speed to arrive during the green traffic light.

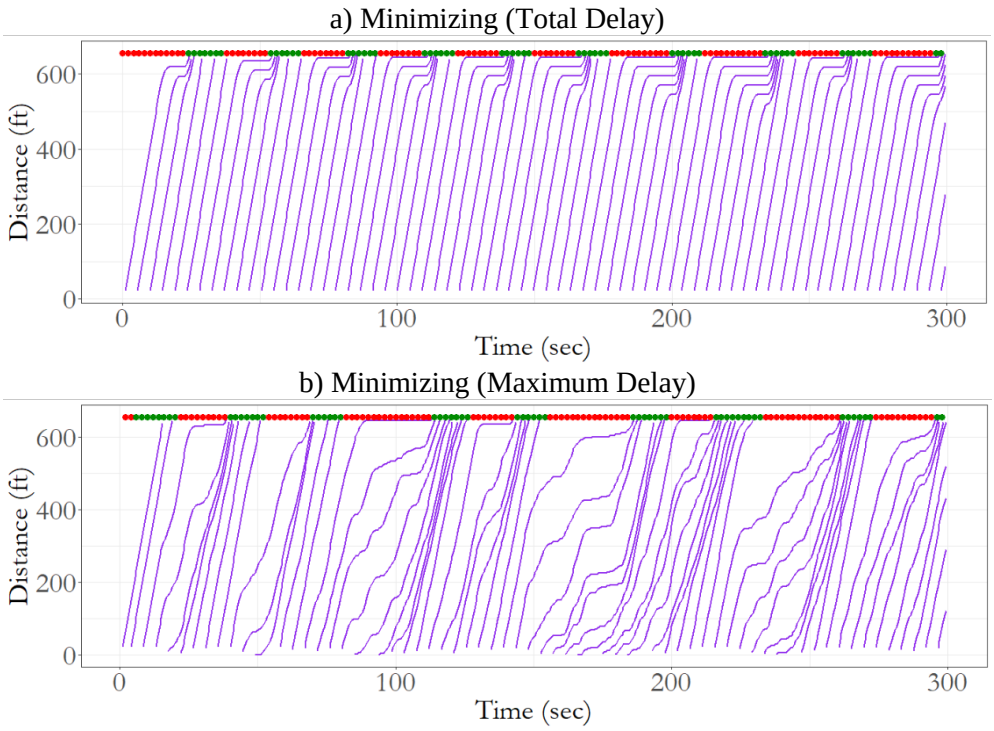


Figure 6-19: Comparing trajectory of CAVs with two different objective functions

6.6. Conclusion

This chapter developed a mixed integer nonlinear signal timing and trajectory optimization problem in a mixed traffic environment of connected automated vehicles and human-driven vehicles. The signal timing parameters and trajectory of CAVs are optimized based on the data from CAVs and HVs. The trajectories of HVs are incorporated into the optimization program using Helly's car-following model. Furthermore, the linear car-following model is extended to take into account the signal timing variables, including the green and yellow times. The proposed problem

is complex due to nonlinearity and the presence of binary variables. Hence, we reduced the complexity of the problem by linearizing the nonlinear constraints and decomposing the interactable optimization problem into lane-level sub-problems using Lagrangian relaxation technique. Thus, the signal timing parameters and trajectory of vehicles in each lane of the intersection can be controlled with a single controller. After converting the MINLP problem into MILP we reformulated the STTO problem with a tight convex hull of the feasible area to reduce the duality gap. In addition, a complementary optimization problem is introduced to find high-quality feasible signal timing parameters when the relaxed constraints are not satisfied after the convergence of Lagrangian relaxation problem. The proposed solution technique is embedded in a receding horizon control technique to capture the dynamic nature of the problem.

The result of this study shows that the Lagrangian relaxation for STTO can find solutions within am optimality gap of at most 0.1%. It is also shown that STTO outperforms the optimized fully-actuated signal timing plans by decreasing the average travel time by 19%-61% for different CAV market penetration rates tested in this study. Furthermore, increasing the penetration rate of CAVs reduces the average travel time for all vehicles at the intersection. This reduction is more significant in higher traffic volumes. The results also show that vehicles can pass the intersection very quickly when the market penetration rate of CAVs increases. As a result, there would be no need to assign long green time duration to each approach.

The proposed methodology is applicable to intersection layouts that operate left-turn movements exclusively. It will be worthwhile in the future to generalize the methodology and find optimal signal timing parameters and vehicles trajectory in all intersection types. In addition, this study assumed that vehicles do not change the lane within the detected range of intersection and did not use CAV lane change to further control the flow of traffic. Developing an algorithm for

predicting the lane changing behavior of HVs and optimizing the lane changing decision of CAVs could further improve traffic operations. This study utilized a linear car following model and it is worthwhile to explore using more complex car fallowing models. Studying the effects of the signal timing and trajectory control in a transportation network, where intersections communicate and coordinate their decisions with each other offers great potential for further improvement in traffic operations and safety and needs to be studied.

CHAPTER 7: DISTRIBUTED COOPERATIVE TRAJECTORY AND LANE CHANGING OPTIMIZATION OF CONNECTED AUTOMATED VEHICLES

As shown in previous chapter, optimizing the longitudinal movement of connected automated vehicles (CAVs) improves the mobility performance of the transportation system in a mixed traffic condition with the presence of human-driven vehicles. In addition to the longitudinal trajectory planning, several studies have shown that lane-changing maneuvers have significant impacts on the safety and operational performance of traffic flow characteristics due to the interfering effects that they have on surrounding vehicles (Daganzo et al., 1999). Vehicles change lanes to achieve desired speeds, to avoid unsafe conditions, or to move into turning/exit lanes. Finding a proper lane-changing maneuver for self-driving vehicles not only improves the mobility performance of automated vehicles, but also controls the movement of human-driven vehicles in a multi-lane arterial street. In other words, the optimal lateral motion planning for self-driving vehicles provides a uniform distribution on the road to keep human-driven vehicles from inefficient movement and prevent them from unnecessary stops at signalized intersections.

Lane drops may create bottlenecks on freeway facilities when the demand flow rate is high. The bottleneck is often activated at a flow rate significantly less than the capacity of the remaining lanes after the lane drop due to the excessive number of lane changing maneuvers that are needed to place vehicles in the appropriate lanes. Controlling the maneuvers of CAVs upstream of the bottleneck will be necessary to avoid traffic congestion. Considerable efforts have been done on the longitudinal control of autonomous vehicles in freeway facilities through adaptive cruise control (Shladover et al., 2012; Vander Werf et al., 2002; Wang et al., 2014b) or cooperative adaptive cruise control (Monteil et al., 2013; Ploeg et al., 2011; H. Yang et al., 2016). However, the problem of optimal lane changing decision making for CVAs remains unresolved, especially

during congested traffic conditions. In fact, coupling the discrete lane changing decision with the continuous longitudinal and lateral vehicle dynamics increases the complexity of trajectory optimization problem (Katzourakis et al., 2012; Sarvi et al., 2004; Toledo et al., 2007). Moreover, the cooperation among CAVs requires an extensive amount of communicational and computational power, which requires developing a scalable solution technique that ensures the safety of vehicles and finds near-optimal solution in a short amount of time regardless of traffic volume or the geometric characteristics of the freeway facilities (Islam et al., 2020; Mehrabipour et al., 2019; Mohebifard and Hajbabaie, 2018b; Tajalli and Hajbabaie, 2021).

This chapter presents a methodology to cooperatively control longitudinal and lateral motions and lane changing decisions of CAVs in a highway facility. We introduce a novel vehicle-level mixed integer nonlinear program (MINLP) that integrates the discrete lane changing decision of CAVs with their granular longitudinal and lateral trajectory planning. CAVs can make both mandatory and discretionary lane changes while keeping a safe distance with all surrounding vehicles. We apply the methodology to a freeway segment with a lane drop with different number of lanes and show how CAVs can cooperate with each other through distributed programming and find near-optimal solutions in a short amount of time.

A distributed cooperative strategy is developed to reduce the complexity of cooperative lane changing decision making, while achieving an agreement among all vehicles in the system to push the solutions in the direction of system-level optimality. As a result of the cooperation, CAVs negotiate with each other to determine collision-free trajectories that reduce the system-level travel time while increase the CAVs' driving comfort. To ensure the feasibility of vehicle-level solutions and push them towards system-level optimality, we develop a coordination scheme between CAVs with conflicting trajectories. The proposed coordination scheme will share traffic state (i.e., vehicle

location and speeds) over a prediction period among CAVs, and implement such information in their respective MINLPs.

7.1. Background

About 4%-10% of total traffic related crashes in the US are associated with lane-changing maneuvers (Ammoun et al., 2007), where the chance of driving errors increases due to the significant increase in drivers' workload and stress (Mattes, 2003; Pande and Abdel-Aty, 2006). Moreover, lane-changing maneuvers increase the probability of traffic breakdowns and decrease the bottleneck discharge flow rate due to spreading the queue laterally, which yields to capacity drop in congested traffic condition (Cassidy and Rudjanakanoknad, 2005). In addition, lane changing maneuvers trigger the formation and propagation of stop-and-go oscillations (Ahn and Cassidy, 2007; Zheng et al., 2011) and traffic flow instabilities (Gazis et al., 1962; Zheng et al., 2011). On the other hand, it is shown that controlling the lane changing of automated vehicles can reduce traffic congestion (Mohebifard and Hajbabaie, 2020; Roncoli et al., 2015) and mitigate traffic oscillation (Schakel and Van Arem, 2014).

Studies addressing the lane changing maneuver of CAVs can be grouped into two main categories. Non-cooperative control strategies make decisions for only a single CAV and do not consider the responses of surrounding vehicles to the control actions. Therefore, there is no negotiation nor consensus in the decision-making process (Wang et al., 2014b). On the other hand, cooperative control strategies let CAVs coordinate their decisions and take into account the expected response of other vehicles when making lane changing decisions (Wang et al., 2014a).

7.1.1. Non-cooperative lane changing decision making for CAVs

Non-cooperative lane changing models mainly focus on the safety and mobility of a single CAV and ignore the potential effects that a CAV's maneuver could have on the traffic flow. The decision variables of non-cooperative control approaches generally concern the dynamics of a subject CAV. For instance, Luo et al. (2016) proposed a closed-form constrained optimization problem to accomplish an automated lane change and eliminate potential collisions during the maneuver based on the perceived information from surrounding vehicles. Xi et al. (2019) developed a mixed integer quadratic optimization program with logical obstacle avoidance constraints to control the motion of a CAV during the lane changing maneuver. The proposed problem formulation ensures keeping a safe distance with the immediate surrounding vehicles on both current and target lanes. Wang et al. (2020) considered the dynamics of surrounding human-driven vehicles during the lane changing maneuver of a CAV. A CAV makes a lane change when a safe distance with the leading and following human drivers in the target lane is ensured. This study assumes that all surrounding vehicles maintain their current speed and lanes during the lane changing process of the subject CAV. Although the exiting non-cooperative lane changing algorithms find a safe lane changing maneuver for the subject CAV, they cannot ensure making decisions that improve the mobility of the entire system. In addition, non-cooperative algorithms are limited in finding feasible maneuvers for mandatory lane changes during congested traffic condition since a CAV might not receive the required safe gap in the target lane for making a safe lane change.

7.1.2. Cooperative lane changing decision making for CAVs

The main advantage of cooperative strategies is that vehicles are assumed to be selfless and take actions that maximize the collective or group utility rather than their own objective.

Several types of cooperative lane changing decision-making models have been proposed in the literature, which could be categorized as incentive-, game theoretic-, and learning-based algorithms.

The minimizing overall braking induced by lane changes (MOBIL) developed by Kesting et al. (2007) has been widely used for cooperative lane changing of CAVs. MOBIL provides a simplified decision-making process defining an incentive to make decisions for a CAV regarding keeping or changing the lanes based on some predefined criterion. Nie et al. (2016), Zheng et al. (2019), Treiber & Kesting (2009), and Khan et al. (2014) used the incentive-based MOBIL for cooperative lane changing decision making to study whether the lane changing of a CAV affects the upstream and downstream vehicles of the original and target lanes in a positive or negative way. Then, the vehicle performs the lane changing maneuver if there is enough incentive compared to a predefined threshold. However, the lane-changing execution process in MOBIL is considered as an instantaneous maneuver, which is not realistic. Ni et al. (2020) and Nilsson et al. (2016) resolved this issue by adding another level to the lane changing decision-making problem to provide realistic longitudinal and lateral accelerations after making the lane changing decision in the upper level. The incentive-based model could provide reasonable performance in uncongested traffic condition, where a CAV's lane change only affects the performance of direct surrounding vehicles. However, this approach might not provide the best mobility performance in congested conditions since a lane changing maneuver could affect the movement of several vehicles upstream.

In contrast to incentive-based approaches, game theoretical approaches could adopt realistic utility functions to achieve a more efficient system-level mobility performance through vehicle-to-vehicle communications and information sharing. Lin et al. (2019) proposed a

cooperative lane changing strategy through a game theoretic framework assuming that CAVs share their corresponding utilities dynamically until they reach a win-win decision that yields lower travel times for all players. This study utilized a discrete cellular automaton model to represent the movement of vehicles and the application of the proposed strategy was successful on a case study between two vehicles that desired to make a lane change. However, establishing the game between a greater number of vehicles, especially during congested traffic conditions, could be cumbersome and might not yield a win-win decision for all vehicles in real time. Similarly, Ding et al. (2018) proposed a synchronous scheme for passive-proactive lane changing based on a game theory model that considers the interaction between three vehicles. The resulting performance on the proposed technique yields a smooth merging of vehicles from different lanes.. Wang et al. (2015) developed a game-based optimization framework, where a controlled CAV predicts the state of neighboring vehicles and plans for both lane changing and trajectory within a planning horizon. However, the proposed algorithm could work with maximum of four vehicles in the tested cases since the dimension of the optimization problem for the subject CAV increases significantly with the number of surrounding vehicles. Overall, the game theoretical approaches are able to provide solutions that benefit the entire system, but their performance highly depends on the number of vehicles competing to make a lane change. Therefore, these methods require to be enhanced to be scalable and applicable in real-time control algorithms.

Learning-based algorithms are among the strategies that can ensure the optimality of cooperative lane changing algorithms for the entire system. Several studies used Q-learning (Ngai and Yung, 2011; Wang et al., 2018; X. Xu et al., 2018), or Markov decision processes (Kuge et al., 2000; Ulbrich and Maurer, 2013) to learn optimal lane changing decisions in addition to safe longitudinal and lateral maneuvers of CAVs. In contrast to game theoretical approaches, the

decisions regarding the movement of vehicles could be made in real-time after learning various scenarios. However, learning the large dimension of state and actions could limit the application of learning-based techniques in real-world. To resolve this problem, deep reinforcement learning approaches are used to make a lane changing decision for CAVs considering the continuous state and action spaces (Alizadeh et al., 2019; Mukadam et al., 2017; Wang et al., 2019). However, complex driving behaviors, especially in congested traffic conditions, and extensive number of scenarios might limit the application of learning-based techniques or prevent finding the solutions that ensure the safety and optimality for the entire system in unexpected situations.

7.2. Research gaps and expected contributions

The review of literature shows that various strategies are employed to perform a safe and efficient lane changing maneuver for CAVs. Cooperative decision-making techniques help improve the mobility of traffic flow and reduce the negative impacts of lane changing maneuvers on the movement of surrounding vehicles. However, the existing incentive-based and game theoretical methods can consider the cooperation between only a few CAVs. Therefore, the system-level optimal performance will not be achieved during the congested traffic condition. Although learning-based strategies can theoretically consider all different situations and come up with the global optimal decisions, the extensive number of scenarios in addition to complex driving situations prevent their application in real-world condition, especially during unexpected conditions. As a result, this study presents a cooperative distributed algorithm that address the existing gaps in the literature. The contribution of this study follows: 1) We develop a novel problem formulation to couple the discrete lane changing decisions with the polynomial longitudinal and lateral equations of motion. 2) We establish cooperation among all vehicles on the road to achieve the system-level operational and safe performance. All connected and

automated vehicles interact with each other by sharing their current and predicted states to achieve a consensus to the optimal lane changing and trajectory planning that benefit the entire system. 3) We propose a vehicle-level distributed algorithm to reduce the complexity of cooperative problem and find the optimal solution for each CAV in short time. Therefore, the proposed strategy will be scalable to various number of vehicles in the transportation network. 4) In contrast with existing studies, the impact of cooperative lane changing is considered in congested traffic condition with very high traffic demand flow rates, where the benefit of all vehicles on the road in terms of mobility and safety are considered.

7.3. Model formulation

We present a vehicle-level program to determine the optimal non-conflicting trajectories for a CAV based on predictive information from other vehicles in traffic stream. Due to discrete nature of lane changing decisions and without loss of generality, CAVs update their trajectories every Δt^t seconds (i.e. 0.2 seconds in this study) and decide about making a lane change every Δt^l seconds (i.e. 1 second in this study). We define T as the set of time steps and I as the set of all vehicles in the network. Figure 7-1 depicts these sets on a freeway facility with three lanes, where the right-most lane is dropped. Note that the detection range of each vehicle $i \in I$ (shown in colored disks) is considered as $\mathcal{R}$ (i.e. 350 ft in this study, but it is significantly down-sized compared to the vehicle and the link dimensions due to representation purposes). Let o denote the origin of the coordinate system, which is defined at the beginning of the controlled segment along the roadway and perpendicular to that. State variables x_i^t and y_i^t represent the relative longitudinal and lateral locations of vehicle $i \in I$ at time $t \in T$ from the origin, respectively. Similarly, variables $v_{x_i}^t$ and $v_{y_i}^t$ denote the longitudinal and lateral speed of vehicle $i \in I$ at time $t \in T$, respectively. Therefore, the state of each vehicle i at time t is defined by its longitudinal and lateral

location and speed (i.e. $s_i^t = \left\{x_i^t, y_i^t, v_{x_i}^t, v_{y_i}^t\right\}, \forall i \in I, t \in T$). Table 7-1 provides a detailed definition of variables, sets, and parameters used in the problem formulation.

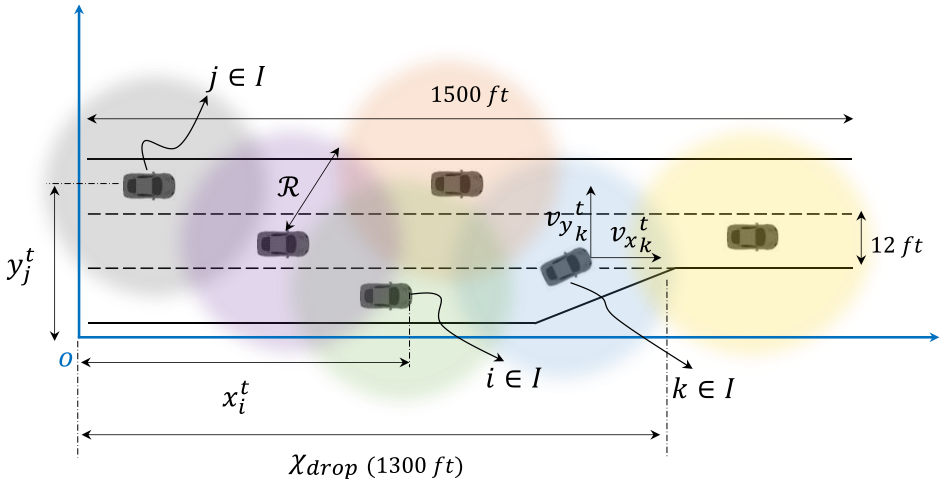


Figure 7-1: Freeway layout with dropped lane, sets of CAVs, and state variables

Utilizing a distributed model structure and capturing each CAV's predicted states at time t, the optimal collision-free trajectories will be found by assigning longitudinal and lateral accelerations $a_{x_i}^t$ and $a_{y_i}^t$ (i.e. the decision variables of the program) to vehicle $i \in I$ at time $t \in T$. The lateral acceleration depends on the two binary decision variables α_i^t and β_i^t to decide if CAV $i \in I$ changes its lane to the relative right or left lane at time $t \in T$, respectively. In other words, either α_i^t or β_i^t takes on the value of 1 if a vehicle makes a right or left lane change, respectively. Otherwise, they take on zero values and the CAV keeps moving on its current lane. A CAV shares its current state in addition to its predicted state with other CAVs through a vehicle to vehicle communication system (e.g., dedicated short-range communication) at every time step.

Table 7-1: Definition of sets, decision variables, and parameters

Sets:	
T	set of all time steps for trajectory updates
I	set of all CAVs
Decision Variables:	
${a_x}_i^t$	the longitudinal acceleration of vehicle $i \in I$ at time $t \in T$
${a_y}_i^t$	the lateral acceleration of vehicle $i \in I$ at time $t \in T$
α_i^t	a binary variable to decide on making a right lane change for CAV $i \in I$ at time $t \in T$
β_i^t	a binary variable to decide on making a left lane change for CAV $i \in I$ at time $t \in T$
State Variables:	
x_{il}^t	the longitudinal location of vehicle $i \in I$ at time $t \in T$
y_i^t	the lateral location of vehicle $i \in I$ at time $t \in T$
${v_x}_i^t$	the longitudinal speed of vehicle $i \in I$ at time $t \in T$
${v_y}_i^t$	the lateral speed of vehicle $i \in I$ at time $t \in T$
s_i^t	the overall state of vehicle $i \in I$ at time $t \in T$
l_i^t	the current lane of vehicle $i \in I$ at time $t \in T$
${l_g}_i^t$	the lane of vehicle $i \in I$ at time $t \in T$ after T_l seconds
Parameters:	
v	Maximum longitudinal speed
v'	maximum lateral speed
$\overline{\sigma_x}$	maximum longitudinal acceleration
$\underline{\sigma_x}$	minimum longitudinal acceleration
$\overline{\sigma_y}$	maximum lateral acceleration
$\underline{\sigma_y}$	minimum lateral acceleration
L_v	the length of vehicle
ξ	desired safety distance between vehicles
τ	the reaction time of CAVs
Δt^t	time interval to update the trajectory decisions
Δt^l	time interval to update the lane changing decisions
$\mathcal{R}$	the detection range around a CAV
χ_{drop}	the location of lane drop in longitudinal coordinate
l_{drop}	the lane number of the dropped lane
w	lane width
γ	weight factor
δ	weight factor
M	a very large value
Auxiliary Variables:	
ρ_i^t	auxiliary variable
γ_{ij}^t	auxiliary variable
λ_{ij}^t	auxiliary variable
${v'_x}_i^t$	auxiliary variable
${v''_x}_i^t$	auxiliary variable
${v'_y}_i^t$	auxiliary variable
${v''_y}_i^t$	auxiliary variable

While maintaining the safe distance between CAVs, we aim at maximizing the longitudinal distance of each CAV from the origin of the controlled freeway segment, which is equivalent to minimizing their travel time, as shown by the first term of the objective function (7.1), $\sum_{t\in T} x_i^t$. Besides, the longitudinal and lateral speed variation of the CAV are minimized to reduce the likelihood of disruptive and frequent changes in vehicles' speeds over time, see the second term of the objective function (7.1), $\sum_{t\in T}\left|{v_{x_i}}^{t+1} - {v_{x_i}}^{t}\right| + \sum_{t\in T}\left|{v_{y_i}}^{t+1} - {v_{y_i}}^{t}\right|$. The third term in the objective function is introduced to penalize frequent lane changing of the CAV for an infinitesimal marginal benefit, $\sum_{t\in T}(\alpha_i^t + \beta_i^t)$. Parameters $\gamma > 0$ and $\delta > 0$ assign the desired weights to the second and third terms of the objective function f, respectively and make all terms to have the same units.

$$f_i = \underset{a_{x_i}, a_{y_i}}{\text{Max}} \left(\sum_{t\in T} x_i^t - \gamma \left[\sum_{t\in T} \left|{v_{x_i}}^{t+1} - {v_{x_i}}^{t}\right| + \sum_{t\in T} \left|{v_{y_i}}^{t+1} - {v_{y_i}}^{t}\right| \right] - \delta \left[\sum_{t\in T} (\alpha_i^t + \beta_i^t) \right] \right)$$

$$\forall\, i \in I \qquad (7.1)$$

Constraints (7.2)-(7.5) use polynomial equations of motion with degree of two to update the longitudinal and lateral location and speed of each CAV at time $t+1$ based on the acceleration decisions ${a_{x_i}}^t$ and ${a_{y_i}}^t$ and state s_i^t at time t.

$$x_i^{t+1} = x_i^{t+1} + {v_{x_i}}^t \Delta t^t + \frac{1}{2} {a_{x_i}}^t {\Delta t^t}^2 \qquad \forall\, i \in I, t \in T \qquad (7.2)$$

$$y_i^{t+1} = y_i^{t+1} + {v_{y_i}}^t \Delta t^t + \frac{1}{2} {a_{y_i}}^t {\Delta t^t}^2 \qquad \forall\, i \in I, t \in T \qquad (7.3)$$

$${v_{x_i}}^{t+1} = {v_{x_i}}^t + {a_{x_i}}^t \Delta t^t \qquad \forall\, i \in I, t \in T \qquad (7.4)$$

$${v_{y_i}}^{t+1} = {v_{y_i}}^t + {a_{y_i}}^t \Delta t^t \qquad \forall\, i \in I, t \in T \qquad (7.5)$$

Constraints (7.6)-(7.10) ensure finding feasible state and decision variables. Constraints (7.6) and (7.7) keep the longitudinal and lateral speeds from taking values higher than the allowable maximum speeds v and v', respectively. Without loss of generality, the maximum longitudinal speed is considered the same as speed limit. Note that the lateral speed can take both positive and negative values depending on moving left or right. Constraints (7.8) limit the second norm of speed to the speed limit. Similarly, Constraints (7.9) and (7.10) limit the longitudinal and lateral accelerations within the maximum and minimum allowable ranges.

$$0 \leq v_{x_i}^t \leq v \qquad \forall\, i \in I,\, t \in T \quad (7.6)$$

$$-v' \leq v_{y_i}^t \leq v' \qquad \forall\, i \in I,\, t \in T \quad (7.7)$$

$$\sqrt{{v_{x_i}^t}^2 + {v_{y_i}^t}^2} \leq v \qquad \forall\, i \in I,\, t \in T \quad (7.8)$$

$$\underline{\sigma_x} \leq a_{x_i}^t \leq \overline{\sigma_x} \qquad \forall\, i \in I,\, t \in T \quad (7.9)$$

$$\underline{\sigma_y} \leq a_{y_i}^t \leq \overline{\sigma_y} \qquad \forall\, i \in I,\, t \in T \quad (7.10)$$

Constraints (7.11) ensure that a CAV changes its lane either to the left or right lanes for consecutive T_l seconds. It is assumed that a CAV arrives to the centerline of the target lane T_l seconds after starting the lane changing maneuver from the centerline of its current lane. Based on empirical observations, it is shown that a lane changing maneuver usually takes between 3 and 5 seconds to be completed (Thiemann et al., 2008; Toledo and Zohar, 2007). Therefore, without loss of generality, we assumed 4 seconds for completing a lane change maneuver.

$$\sum_{t'=t-T_l}^{t} (\alpha_i^{t'} + \beta_i^{t'}) \leq 1 \qquad \forall\, i \in I, t \in T \quad (7.11)$$

Constraints (7.12) and (7.13) provide the boundary conditions before and after a lane changing maneuver. The lateral speed and acceleration of a CAV are set to zero before starting and after finishing a lane changing maneuver, where the value of $\sum_{t'=t-T_l}^{t}(\alpha_i^{t'} + \beta_i^{t'})$ is equal to zero. In other words, a CAV can take non-zero lateral speed and acceleration only during the lane changing maneuver. To satisfy these conditional constraints, we use big M method.

$$-M \sum_{t'=t-T_l}^{t} (\alpha_i^{t'} + \beta_i^{t'}) \leq {v_y}_i^t \leq M \sum_{t'=t-T_l}^{t} (\alpha_i^{t'} + \beta_i^{t'}) \qquad \forall\, i \in I, t \in T \quad (7.12)$$

$$-M \sum_{t'=t-T_l}^{t} (\alpha_i^{t'} + \beta_i^{t'}) \leq {a_y}_i^t \leq M \sum_{t'=t-T_l}^{t} (\alpha_i^{t'} + \beta_i^{t'}) \qquad \forall\, i \in I, t \in T \quad (7.13)$$

In addition, to preserving the boundary conditions, Constraints (7.14) are defined to ensure that the lateral position of a CAV will be at the centerline of the adjacent target lane T_l seconds after beginning of the lane changing maneuver. Parameter w is the lane width. We assume similar widths for all lanes, but it can be variant over the location or lanes. Similar to Constraints (7.12) and (7.13), the big M method is used to linearize the conditional Constraints (7.14). To illustrate, assume a CAV starts a lane change to the right at time t. Thus, variable α_i^t takes the value of 1 and β_i^t takes zero. Based on Constraints (7.14), the lateral position of the CAV after T_l seconds will be equal to $y_i^{t+T_l} = y_i^t - w$, which is the centerline of the adjacent right lane.

$$y_i^t - \alpha_i^t w + \beta_i^t w - (1 - \alpha_i^t - \beta_i^t)M \leq y_i^{t+T_l} \leq y_i^t - \alpha_i^t w + \beta_i^t w + (1 - \alpha_i^t - \beta_i^t)M$$

$$\forall\, i \in I, t \in T \quad (7.14)$$

Constraints (7.15) and (7.16) update the lane that a CAV is located on at time $t + 1$ based on the lane at time t and lane changing decision variables α_i^t and β_i^t. Constraints (7.15)

immediately update the CAV lane after taking the decision to start the lane change. On the other hand, Constraints (7.16) update the CAV's lane with a lag of T_l seconds.

$$l_i^{t+1} = l_i^t + \alpha_i^t - \beta_i^t \qquad \forall\, i \in I, t \in T \qquad (7.15)$$

$$l_{g_i}^{\ t+1} = l_{g_i}^{\ t} + \alpha_i^{t-T_l} - \beta_i^{t-T_l} \qquad \forall\, i \in I, t \in T \qquad (7.16)$$

The information about the immediate and lagged lane updates could be used to satisfy the safety Constraints (7.17)-(7.20). The basic idea is that vehicle $i \in I$ should keep a safe distance $\zeta + L_V$ from vehicle $j \in I \backslash i$ when two vehicles are currently on the same lane or they will be merged into the same lane after the completion of the lane changing maneuver of either vehicles. Parameter ζ is the desired distance from the rear-end of vehicle i to the front-end of vehicle j. Parameter L_v is the vehicle length. Moreover, the distance travelled over the reaction time $v_{x_i}^{\ t} \tau$ will be added to the desired distance since the vehicle's headway depends on its corresponding speed. Parameter τ represents the reaction time of CAVs. The safety constraints in the optimization problem of vehicle i should be satisfied only when vehicle i is located behind vehicle j in longitudinal coordinate. In other words, x_i^t should be less than or equal to $\hat{x}_j^t$ to satisfy the safety constraints for vehicle i. It should be noted that the information about the longitudinal and lateral position of surrounding vehicles in addition to their corresponding immediate and lagged lanes are considered as an input to the optimization problem of vehicle i through vehicle-to-vehicle communications.

If $x_i^t \leq \hat{x}_j^t$:

$$\left|x_i^t - \hat{x}_j^t\right| + \left|y_i^t - \hat{y}_j^t\right| + M\left|l_i^t - \hat{l}_j^t\right| \geq \zeta + L_v + v_{x_i}^{\ t} \tau \qquad \forall\, i \in I, j \in I \backslash i, t \in T \quad (7.17)$$

$$\left|x_i^t - \hat{x}_j^t\right| + \left|y_i^t - \hat{y}_j^t\right| + M\left|l_{g_i}^{\ t} - \hat{l}_j^t\right| \geq \zeta + L_v + v_{x_i}^{\ t} \tau \qquad \forall\, i \in I, j \in I \backslash i, t \in T \quad (7.18)$$

$$\left|x_i^t - \hat{x}_j^t\right| + \left|y_i^t - \hat{y}_j^t\right| + M\left|l_i^t - \widehat{l_{g_j}}^t\right| \geq \zeta + L_v + v_{x_i}^t \curlyvee \qquad \forall\, i \in I, j \in I \backslash i, t \in T \quad (7.19)$$

$$\left|x_i^t - \hat{x}_j^t\right| + \left|y_i^t - \hat{y}_j^t\right| + M\left|{l_{g_i}}^t - \widehat{l_{g_j}}^t\right| \geq \zeta + L_v + v_{x_i}^t \curlyvee \qquad \forall\, i \in I, j \in I \backslash i, t \in T \quad (7.20)$$

Although Constraints (7.17)-(7.20) avoid collision between a vehicle and other vehicles, there is a need to sense the stationary obstacles and avoid collision with them. We assume that, the stationary obstacles such as blocked lane due to the lane drop is sensed through static high-definition (HD) map information. The location of lane drop χ_{drop} is fixed in the formulation and a vehicle needs to make a proper lane change ahead of the dropped location to avoid the collision. Therefore, Constraints (7.21) make sure that vehicle i changes the lane before reaching χ_{drop}. Without loss of generality, we assume that the right-most lane is dropped (i.e. $l_{drop} = 1$). Therefore, vehicle i needs to change the lane to other left-hand side lanes when $x_i^t \leq \chi_{drop}$. The auxiliary variable ρ_i^t is defined in (7.22) to satisfy the conditional constraints.

$$l_i^t \geq l_{drop} + 1 - M\rho_i^t \qquad \forall\, i \in I, t \in T \quad (7.21)$$

$$\rho_i^t = \begin{cases} 1 & \text{if } x_i^t \leq \chi_{drop} \\ 0 & \text{otherwise} \end{cases} \qquad \forall\, i \in I, t \in T \quad (7.22)$$

7.4. Solution technique

The proposed optimization program (7.1)-(7.22) finds a feasible trajectory for a single CAV based on the predicted trajectory of other vehicles in the system. However, there is no guarantee that the solution of a single CAV is feasible for other vehicles in the system, especially during congested traffic condition with significant number of interactions between vehicles. In addition, a selfish CAV-level optimal solution might not provide the system-optimal solution for all vehicles in the network. The exchange of the state information between CAVs, i.e. capturing

the current and predicted trajectories of neighboring vehicles in Constraints (7.17)-(7.20), can help build a coordination between CAV-level problems to achieve a system-level optimal performance.

Note that constraints (7.17)-(7.20) are non-convex since the absolute value functions divide the feasible region into separate regions without any intersects to prevent crashes between neighboring vehicles. Hence, the absolute value functions are linearized using binary variables to account for the conditional safety. We have two sets of binary variables γ_{ij}^t and λ_{ij}^t to define the relative longitudinal and lateral location of vehicles i and j with respect to each other and used big M method to linearize the absolute value functions. Equations (7.23) and (7.24) define the logical rules for these two conditional variables.

$$\gamma_{ij}^t = \begin{cases} 1 & \text{if } x_i^t \le x_j^t \\ 0 & otherwise \end{cases} \qquad \forall\, i \in I, j \in I \backslash i, t \in T \quad (7.23)$$

$$\lambda_{ij}^t = \begin{cases} 1 & if\ y_i^t \le y_j^t \\ 0 & otherwise \end{cases} \qquad \forall\, i \in I, j \in I \backslash i, t \in T \quad (7.24)$$

For each sets of Constraints (7.17)-(7.20), four new linearized constraints are defined depending on the relative location of two vehicles. Constraints (7.25)-(7.28) illustrate the linear forms for Constraints (7.17). Similarly, new constraints are defined to linearize other safety Constraints (7.18)-(7.20).

$$\left(x_i^t - \hat{x}_j^t\right) + \left(y_i^t - \hat{y}_j^t\right) + M\left(l_i^t - \hat{l}_j^t\right) \ge \zeta + L_v + v_{x_i}{}^t r - M\gamma_{ij}^t - M\lambda_{ij}^t \qquad \forall\, i \in I, j \in I \backslash i, t \in T \quad (7.25)$$

$$\left(-x_i^t - \hat{x}_j^t\right) + \left(y_i^t - \hat{y}_j^t\right) + M\left(l_i^t - \hat{l}_j^t\right) \ge \zeta + L_v + v_{x_i}{}^t r - M\gamma_{ji}^t - M\lambda_{ij}^t \qquad \forall\, i \in I, j \in I \backslash i, t \in T \quad (7.26)$$

$$\left(x_i^t - \hat{x}_j^t\right) + \left(-y_i^t + \hat{y}_j^t\right) + M\left(-l_i^t + \hat{l}_j^t\right) \ge \zeta + L_v + v_{x_i}{}^t r - M\gamma_{ij}^t - M\lambda_{ji}^t \qquad \forall\, i \in I, j \in I \backslash i, t \in T \quad (7.27)$$

$$\left(-x_i^t + \hat{x}_j^t\right) + \left(-y_i^t + \hat{y}_j^t\right) + M\left(-l_i^t + \hat{l}_j^t\right) \ge \zeta + L_v + v_{x_i}{}^t r - M\gamma_{ji}^t - M\lambda_{ji}^t \qquad \forall\, i \in I, j \in I \backslash i, t \in T \quad (7.28)$$

In addition, objective function (7.1) contains two convex but nonlinear terms with absolute value functions. We linearized these functions by introducing non-negative auxiliary variables ${v'}_{x_i}^{\ t}$ and ${v''}_{x_i}^{\ t}$ for the longitudinal and ${v'}_{y_i}^{\ t}$ and ${v''}_{y_i}^{\ t}$ for the lateral speeds in the second term of the objective function. The equivalent objective function (7.29) is introduced by minimizing the sum of the auxiliary variables subsequent to the addition of constraints (7.30)-(7.32) in each vehicle-level subproblem.

$$f'_i = \underset{a_{x_i}, a_{y_i}}{\text{Max}} \left(\sum_{t \in T} x_i^t - \gamma \left[\sum_{t \in T} \left({v'}_{x_i}^{\ t} + {v''}_{x_i}^{\ t} \right) + \sum_{t \in T} \left({v'}_{y_i}^{\ t} + {v''}_{y_i}^{\ t} \right) \right] - \delta \left[\sum_{t \in T} \left(\alpha_i^t + \beta_i^t \right) \right] \right) \quad \forall\, i \in I \tag{7.29}$$

$$v_{x_i}^{\ t+1} - v_{x_i}^{\ t} = {v'}_{x_i}^{\ t} - {v''}_{x_i}^{\ t} \qquad \forall\, i \in I, t \in T \tag{7.30}$$

$$v_{y_i}^{\ t+1} - v_{y_i}^{\ t} = {v'}_{y_i}^{\ t} - {v''}_{y_i}^{\ t} \qquad \forall\, i \in I, t \in T \tag{7.31}$$

$${v'}_{x_i}^{\ t}, {v''}_{x_i}^{\ t}, {v'}_{y_i}^{\ t}, {v''}_{y_i}^{\ t} \geq 0 \qquad \forall\, i \in I, t \in T \tag{7.32}$$

We also linearized the conditional constraints (7.22)-(7.24) using the big M method. Constraints (7.22) are replaced with (6.30) and (6.31). Similarly, Constraints (7.23) and (7.24) are replaced with (7.35)-(7.36) and (7.37)-(7.38), respectively.

$$x_i^t - \chi_{drop} \leq M(1 - \rho_i^t) \qquad \forall\, i \in I, t \in T \tag{7.33}$$

$$x_i^t - \chi_{drop} \geq M\rho_i^t \qquad \forall\, i \in I, t \in T \tag{7.34}$$

$$x_i^t - x_j^t \leq M(1 - \gamma_{ij}^t) \qquad \forall\, i \in I, j \in I \backslash i, t \in T \tag{7.35}$$

$$x_i^t - x_j^t \geq M\gamma_{ij}^t \qquad \forall\, i \in I, j \in I \backslash i, t \in T \tag{7.36}$$

$$y_i^t - y_j^t \leq M(1 - \lambda_{ij}^t) \qquad \forall\, i \in I, j \in I \backslash i, t \in T \tag{7.37}$$

$$y_i^t - y_j^t \geq M\lambda_{ij}^t \qquad \forall\, i \in I, j \in I \backslash i, t \in T \tag{7.38}$$

The most recent CAVs' decisions ${a_x}_j^t$, ${a_y}_j^t$, α_i^t and β_i^t will be used to update the trajectories of subject CAV $i \in I$ at time $t \in T$ to find the feasible state s_i^t. The proposed distributed cooperative algorithm is embedded into a model predictive control (MPC) strategy that solves the problem over a prediction horizon τ. Therefore, the dynamic nature of the system is taken into account and the future state of each CAV could be predicted. The prediction horizon τ is selected long enough to capture all possible longitudinal and lateral maneuvers of a CAV to allow complete at least one lane changing maneuver. We implement the MPC through three steps: (i) initialization, (ii) cooperation, and (iii) implementation and update, as follow.

Step 1. We first initialize the problem at time step $t = 0$ by defining an initial set of vehicles I^0 and vehicle states s_i^0 for each CAV $i \in I^0$.

Step 2. We then follow an iterative procedure by sharing all predicted trajectories of CAVs in an agreement framework to reach a consensus on feasible maneuvers that benefit the entire system over K iterations or until convergence. All variables at each iteration of the cooperative procedure are indexed by $\kappa \in \{1, 2, \dots, \mathrm{K}\}$. The cooperation procedure is described as follows.

1. Set $\kappa = 1$ and $s_i^{t,\kappa} \leftarrow s_i^t$.

2. Collect $s_i^{t,\kappa}$ for all CAVs $i \in I^t$ in prediction horizon $t \rightarrow t + \tau$.

3. Optimize CAV-level trajectory.

Each CAV $i \in I^t$ optimizes its own trajectory by solving problem $\mathcal{A}$: (7.2)-(7.21) and (7.29)-(7.38) acquiring predicted trajectories of other CAVs as inputs. In particular, at time instant t, vehicle i receives the predicted information about the future states of other vehicles from t to $t + \tau$ through V2V communications. Based on this information, the optimal lane changing

decisions $\alpha_i^{t'}$ and $\beta_i^{t'}$, and consequently optimal trajectory plan ${a_x}_j^{t'}$ and ${a_y}_j^{t'}$ for $t' \in \{t, t+1, \dots, t+\tau\}$ are made to determine a unique path based on the desired objective value of vehicle i, as shown in (7.29).

4. Find system-level optimal feasible trajectories.

In this step, we evaluate if the recently-found trajectories by CAV $i \in I^t$ based on input states $s_j^{t,\kappa}$ of other vehicles $j \in I^t \backslash i$ is improving the system-level performance. For this purpose, all predicted trajectories of CAVs (i.e. $s_i^{t,\kappa'}$ for $i \in I^t$ and iterations $\kappa' \in \{1, \dots, \kappa\}$) are gathered in the optimization problem $\mathcal{B}$: (7.39)-(7.42). The objective function (7.39) takes into account the benefit of the entire system by maximizing the distance of all CAVs from the origin and minimizing all CAVs' speed differences and avoiding unnecessary lane changing. In fact, this optimization problem looks over different predictions provided by all CAVs up to the iteration κ and finds a feasible trajectory for each CAV that benefits the entire system rather than just vehicle-level performances. The longitudinal and lateral locations $x_i^{t,\kappa'}$ and $y_i^{t,\kappa'}$, the longitudinal and lateral speeds ${v_x}_i^{t,\kappa'}$ and ${v_y}_i^{t,\kappa'}$, lane numbers $l_i^{t,\kappa'}$ and ${l_g}_i^{t,\kappa'}$, and lane changing decision variables $\alpha_i^{t,\kappa'}$ and $\beta_i^{t,\kappa'}$ for all vehicles $i \in I^t$ and all iterations $\kappa' \in \{1, \dots, \kappa\}$ are considered into this optimization. The main decision variable, ${\theta_i}^{\kappa'}$, for each vehicle $i \in I^t$, selects the best combination of trajectories among all vehicles to the benefit of the system. Variable ${\theta_i}^{\kappa'}$ takes the value of 1 when the state $s_i^{t,\kappa'}$ at iteration $\kappa' \in \{1, \dots, \kappa\}$ is selected as the optimal maneuver for vehicle $i \in I^t$. Otherwise it takes zero.

Problem $\mathcal{B}$:

$$f = \underset{a_x, a_y}{\text{Max}} \left(\sum_{t \in T} \sum_{i \in I} (x_{il}^t) - \gamma \left[\sum_{t \in T} \sum_{i \in I} \left({v'_{x_i}}^t + {v''_{x_i}}^t \right) + \sum_{t \in T} \sum_{i \in I} \left({v'_{y_i}}^{t+1} + {v''_{y_i}}^t \right) \right] - \delta \left[\sum_{t \in T} \sum_{i \in I} (\alpha_i^t + \beta_i^t) \right] \right) \quad (7.39)$$

s.t.

(7.17)-(7.20) $\forall\, i \in I, j \in I \backslash i, t \in T$ (7.40)

$$s_i^t \leq s_i^{t,\kappa'} + M\theta_i^{\kappa'} \qquad \forall\, i \in I, t \in T, \kappa' \in \{1, \dots, \kappa\} \quad (7.41)$$

$$s_i^t \geq s_i^{t,\kappa'} - M\theta_i^{\kappa'} \qquad \forall\, i \in I, t \in T, \kappa' \in \{1, \dots, \kappa\} \quad (7.42)$$

Step 3. Set $s_i^{t+1} \leftarrow s_i^{t,\kappa+1}$ and implement the trajectory at time step t for all CAVs only for the first time step (e.g., starting from $t = 0$). The remaining predicted trajectories (i.e., $t + 1 \rightarrow t + \tau$) will be considered as predicted trajectories for near future and shared with corresponding CAVs. Therefore, at the next time step (e.g., $t = 1$), CAV i obtains all recently updated trajectories of other CAVs as inputs and follows the proposed procedure from Step 2. Moreover, update the set I^{t+1} roll one time step, and set $t \leftarrow t + 1$.

Steps 2 and 3 continue until CAVs' trajectories are implemented at $\Gamma - 1 + \tau$ that is defined as the end of the study period. Figure 7-2 shows the general framework for the proposed approach.

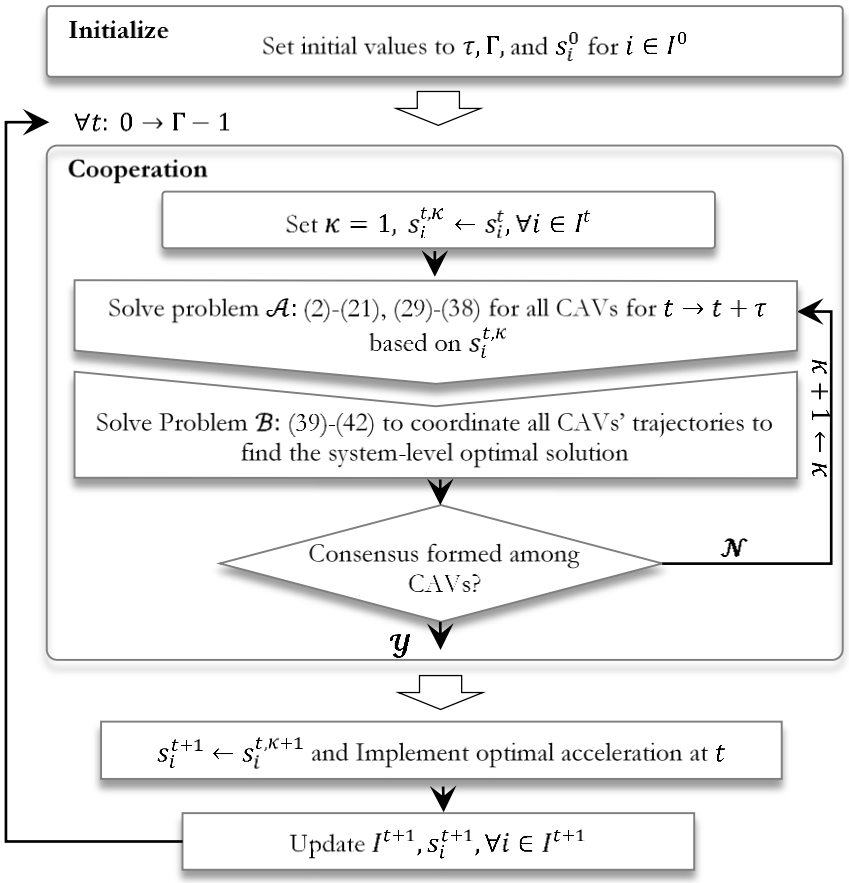


Figure 7-2: Model predictive control framework for distributed cooperative algorithm

7.5. Case Study

We applied the proposed cooperative distributed optimization algorithm in a freeway segment with 1500 ft length, where the right-most lane is dropped at the location 1300 ft. Twenty seven scenarios are considered to evaluate the performance of the system based on variations in demand, number of lanes, and speed limit. We considered three different numbers of lanes (two, three, and four lanes), three different speed limits of 60, 65, and 70 mph and three demand flow rates of 900, 1500, and 2400 $veh/h/ln$. Vehicle to vehicle communications are considered within 350 ft distance range from each vehicle. To accurately update the location of CAVs, sufficiently short time steps of 0.2 sec are selected in the case studies. Besides, the length

L_v of each CAV is assumed to be 13 ft and the safety distance ζ is considered to be 7 ft. Table 7-2 shows the parameters used for case studies, where the maximum longitudinal and lateral acceleration rates are considered 13 ft/s^2 and 6.5 ft/s^2, respectively. The maximum longitudinal and lateral deceleration rates are considered 11.5 ft/s^2 and 6.5 ft/s^2, respectively. The study period is considered as 900 s with 10 s planning horizon.

Table 7-2: Parameters for implementing case studies

Parameters	Value
Maximum longitudinal speed (mph)	$60-65-70$
Maximum longitudinal acceleration (ft/s^2)	13
Maximum longitudinal deceleration (ft/s^2)	11.5
Maximum lateral speed (ft/s)	6.5
Maximum lateral acceleration (ft/s^2)	6.5
Maximum lateral deceleration (ft/s^2)	6.5
Maximum lane changing duration (s)	4
CAV reaction time (s)	0.6
Detection range (ft)	330
Dropped-lane length (ft)	300
Average length of the vehicles (ft)	13
Safety distance between following vehicles (ft)	7
Trajectory updating interval (s)	0.2
Planning time horizon	10
Study period (s)	300

We used the all-known CAV parameters in Widemann 99 car-following model in Vissim (PTV Group, 2015) as the benchmark to compare the performance of the proposed automated lane changing algorithm with. The algorithm is implemented through multithreading in JAVA and run on a high performance computer with 96 virtual cores, 2.9 GHz CPU, and 512 GB memory. We used Vissim Component Object Model (COM) interface to exchange CAVs' states between Vissim and our JAVA-coded algorithm.

7.6. Results

First, we tested our proposed algorithm in a simple scenario to show how a CAV can find its optimal path when slow moving vehicles are in front of it. As Figure 7-3 shows, we have a

three-lane road with the length of 250 meters. Vehicles 1 and 2 move with 20 mph speed on the right and left edge lanes. The speed of vehicle 3 is 15 mph on the middle lane. The target is vehicle 4 with 30 mph speed. Tracking the movement of vehicle 4 over time shows two lane changing maneuvers to overtake slow leading vehicles. The first lane changing starts at time 3s and finishes at time 6s from the middle lane to the left lane. The second lane changing maneuver takes place at time 10 s and finishes at time 13 s from the left lane to the middle lane. Finally, the vehicle with higher speed leaves the network earlier than all vehicles at time 18s.

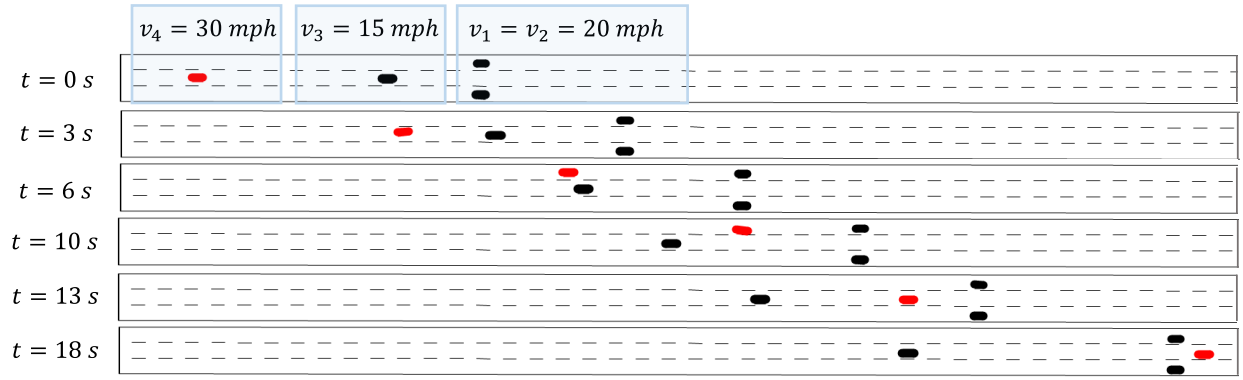


Figure 7-3: Optimal lane changing for a simple scenario on a freeway facility

Figure 7-4 shows the summation of objective values (7.39) for all CAVs on a three-lane freeway with three demand patterns. In the maximization problem, increasing the number of iterations κ^t is associated with increasing the objective value (7.39) and therefore, feasible solutions converge towards the system optimal solution. We observe considerable increase in the objective value when the number of iterations κ^t increases from 1 to 5. However, the increase become insignificant by adding more iterations while higher iterations have a direct relationship with increasing the computational time. Hence, we have selected $\kappa^t = 5$ in this study since no significant improvement is observed in the objective value afterwards, see Figure 7-4. In addition, the iteration histograms on the right-hand side of Figure 7-4 show that the solution of the problem for higher traffic demands requires more number of iterations to converge to the optimal solution.

This is due to negotiation among more number of CAVs and as a result, it will take more time to agree on the system-level optimal solution that benefit all CAVs.

Figure 7-5 shows how the trajectory of a single CAV is found through the cooperative distributed procedure. Figure 7-5-a shows that the trajectory of a CAV is adjusted over several iterations to converge to the optimal solution after negotiating with other CAVs in the system. Figure 7-5-b and Figure 7-5-c also show the optimal longitudinal and lateral speeds and accelerations, respectively for making the lane changing maneuver after the convergence.

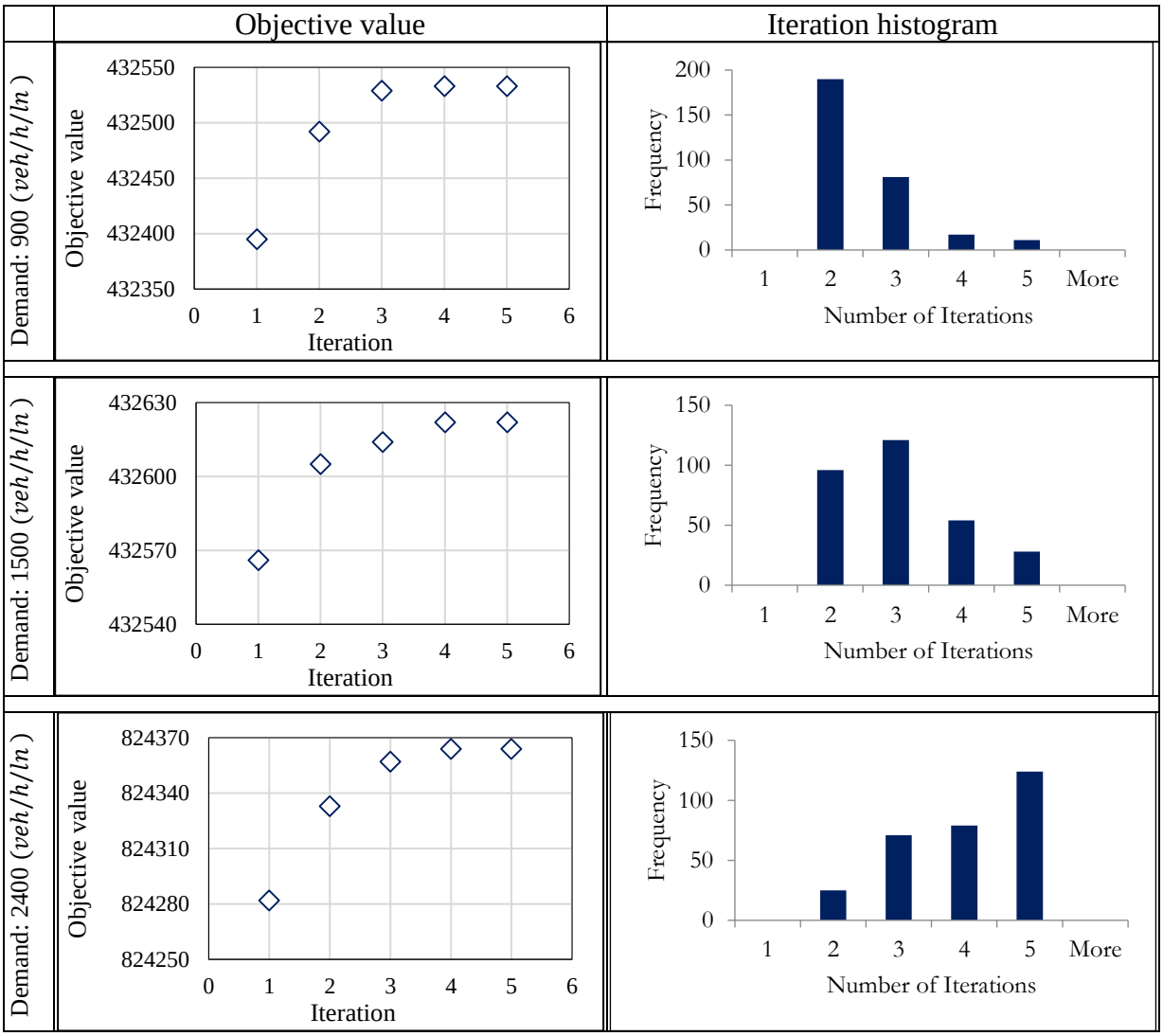
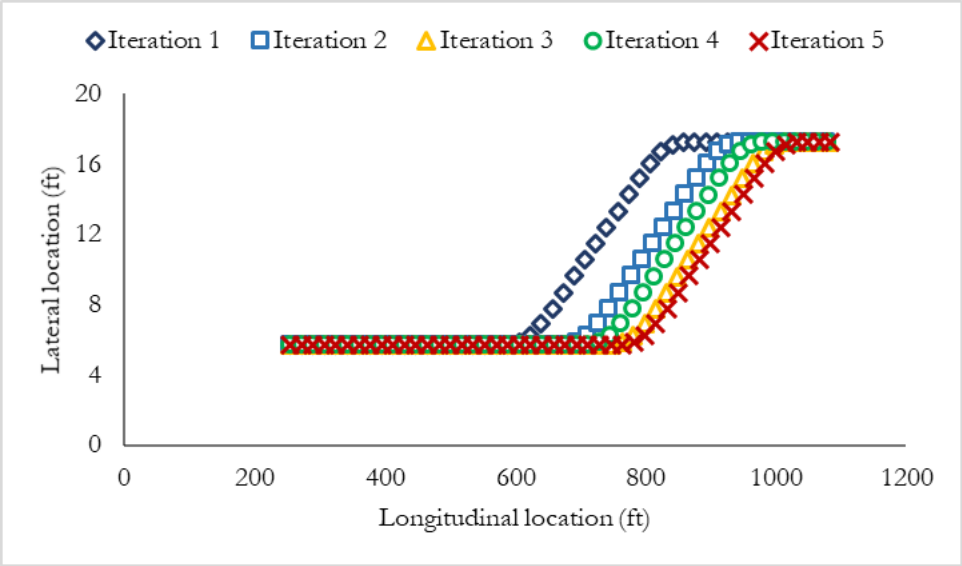
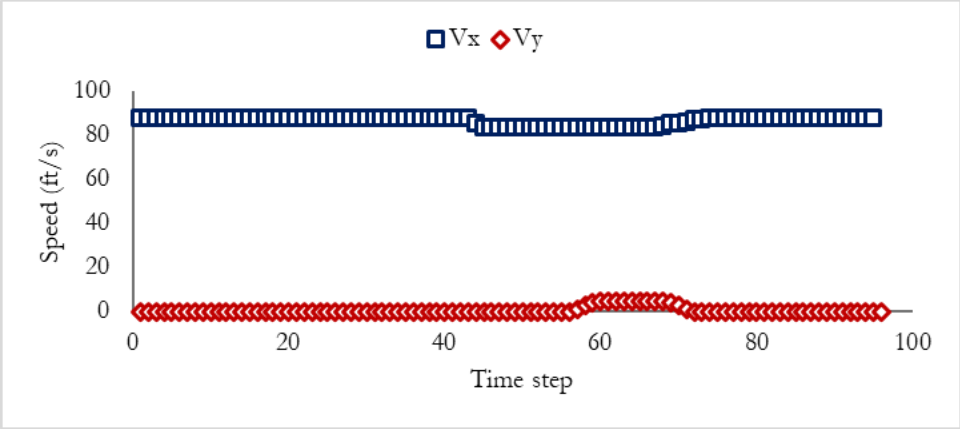


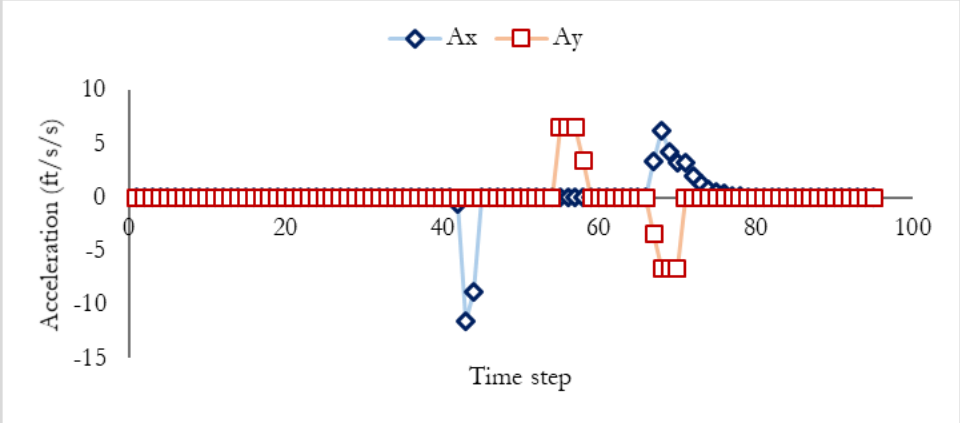
Figure 7-4: The change of objective values over iterations, and their corresponding iteration histograms



a) Adjustment of trajectory over iterations



b) Longitudinal and lateral speed



c) Longitudinal and lateral acceleration

Figure 7-5: Optimal trajectory, speed, and acceleration of a CAV after convergence

Table 7-3 shows the comparison between the mobility performance of the optimal automated lane changing algorithm and the all-knowing CAV behavior in Vissim simulation. Average travel time, total throughput, and total number of stops are considered as measures of the system mobility performance. The results show that automated lane changing of CAVs is more effective with higher demand, smaller number of lanes, and slower speed limit. In other words, the most improvement is achieved during more congested traffic conditions, when vehicles have less freedom to move with their desired speed and have less opportunity to make a lane change. Table 7-3 also shows an interesting observation that operating CAVs with automated lane changing algorithm will achieve zero number of stops. This is due to adjusting the speed of vehicles upstream of the lane drop and smoothening the movement of CAVs with the second term of the objective function (7.29).

Figure 7-6 shows the average delay per lane for different scenarios depending on demand, speed limit, and number of lanes. The patterns show that increasing the demand is associated with higher delay. On the other hand increasing the speed limit reduces the average delay as expected. Increasing the number of lanes from 2 to 3 and from 3 to 4 decreases the average delay. However, the pattern of delay reduction is different for three demand levels. For demand 900 veh/h/ln, the delay is decreasing linearly when the number of lanes increases from 2 to 3 and from 3 to 4. For demand 1500 veh/h/ln the reduction rate is steeper when the number of lanes increases from 2 to 3 compared to increasing the lane numbers from 3 to 4. For demand 2400 veh/h/ln, the trend is different from other two demand patterns, where the rate of delay reduction from 2 to 3 lane is steeper than from 3 to 4 lane. As a result, the effect of automated lane changing decision on reducing the delay is dependent of number of lanes in addition to demand level and speed limit.

Table 7-3: The mobility performance of automated lane changing algorithm in comparison with Vissim simulation

Demand Flow Rate (veh/h/ln)	Number of lanes	Speed limit (mph)	Average travel time (s)			Throughput (vehicle)			Number of stops		
			Automated lane changing	Vissim	Difference (%)	Automated lane changing	Vissim	Difference (%)	Automated lane changing	Vissim	Difference (%)
900	2	60	18.3	19.1	-4.0	465	427	8.9	0	14	-100.0
		65	16.9	17.7	-4.3	465	429	8.4	0	9	-100.0
		70	15.7	17.2	-8.7	466	427	9.1	0	11	-100.0
	3	60	18.3	18.3	0.0	696	646	7.7	0	0	0.0
		65	16.9	16.9	0.0	697	646	7.9	0	1	-100.0
		70	15.7	15.8	-0.4	697	648	7.6	0	2	-100.0
	4	60	18.2	18.3	-0.6	899	854	5.3	0	3	-100.0
		65	16.8	17.1	-1.9	901	855	5.4	0	6	-100.0
		70	15.6	16.5	-5.2	902	855	5.5	0	8	-100.0
1500	2	60	18.7	116.9	-84.0	757	505	49.9	0	2553	-100.0
		65	17.3	110.2	-84.3	758	518	46.3	0	2457	-100.0
		70	16.0	117.7	-86.4	758	506	49.8	0	2657	-100.0
	3	60	18.4	19.0	-3.0	1119	1082	3.4	0	10	-100.0
		65	17.0	17.5	-2.4	1120	1083	3.4	0	8	-100.0
		70	15.9	16.7	-4.8	1122	1084	3.5	0	10	-100.0
	4	60	18.2	19.0	-4.0	1476	1414	4.4	0	12	-100.0
		65	16.9	19.1	-11.7	1477	1416	4.3	0	38	-100.0
		70	15.7	17.5	-10.5	1478	1418	4.2	0	29	-100.0
2400	2	60	26.1	144.6	-82.0	1161	496	134.1	0	3278	-100.0
		65	24.7	141.5	-82.6	1176	502	134.3	0	3347	-100.0
		70	21.7	141.2	-84.6	1179	506	133.0	0	3563	-100.0
	3	60	18.9	85.1	-77.8	1782	1164	53.1	0	3080	-100.0
		65	17.4	81.1	-78.5	1786	1184	50.8	0	3043	-100.0
		70	16.4	80.4	-79.6	1788	1196	49.5	0	3140	-100.0
	4	60	18.7	68.8	-72.8	2393	1791	33.6	0	2734	-100.0
		65	17.4	68.6	-74.7	2397	1770	35.4	0	2790	-100.0
		70	16.3	66.9	-75.6	2399	1797	33.5	0	2885	-100.0

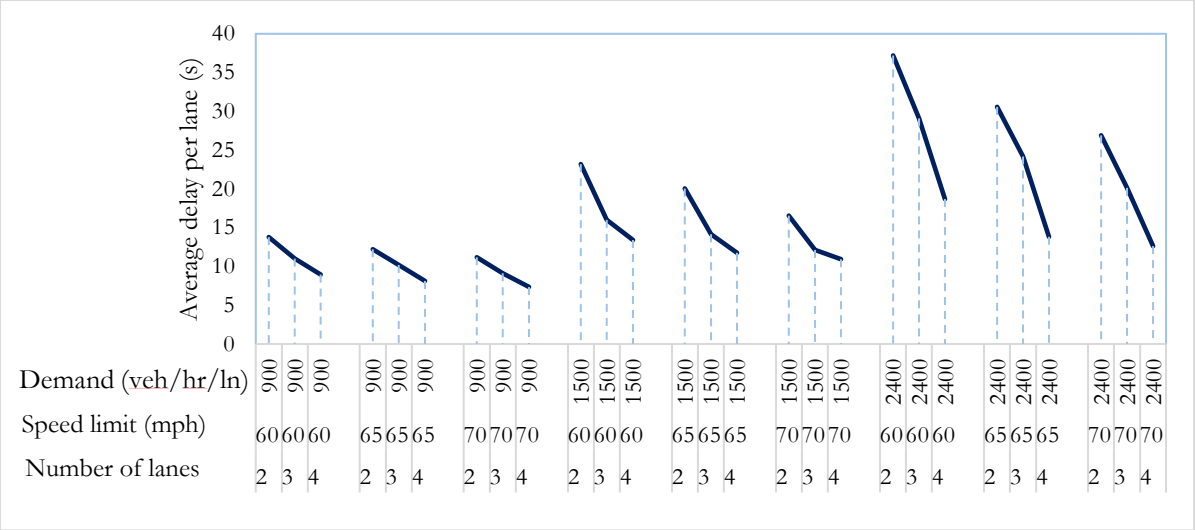


Figure 7-6: Average delay per lane for various scenarios

Figure 7-7 shows the number of lane changing maneuvers that took place in different test scenarios based on demands and number of lanes. The lane changing maneuvers are divided into mandatory and discretionary maneuvers. The mandatory lane changing refers to the maneuvers that a CAV has to change the lane from the dropped lane to the adjacent left lane to avoid stationary obstacles. The discretionary lane changing refers to the lane changing maneuvers that a CAV does not necessarily have to make it to avoid obstacles, but changing the lane benefit the mobility performance of the system and reduces the experienced delay or increases the driving comfort. As Figure 7-7 shows, the number of mandatory lane changing maneuvers increases by increasing the demand from 900 veh/h/ln to 2400 veh/h/ln. However, the number of discretionary lane changing depends on the number of lanes. In other words, more number of lanes provide more freedom for CAVs to select the desired lane that improves their mobility performance by making a lane change.

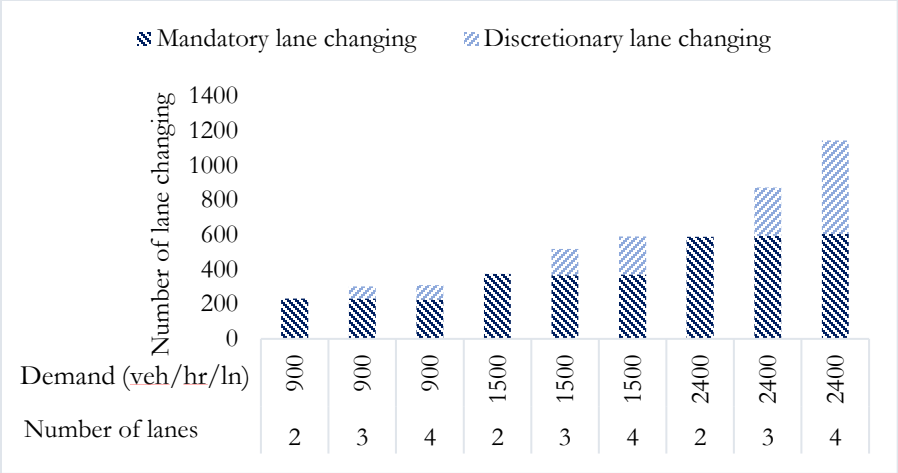


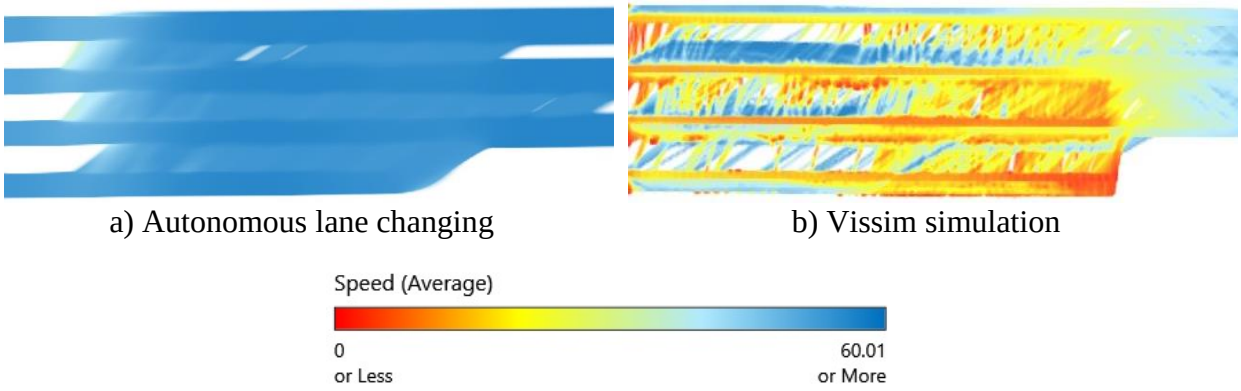
Figure 7-7: Number of mandatory and discretionary lane changing based on demand levels and number of lanes

Table 7-4 shows the result of automated lane changing algorithm in terms of safety. We used Time to Collision (TTC) as a surrogate safety measure for the rear-end collision that could occur in freeway facilities. The higher value of TTC indicates less dangerous situation. As the results show, no near collision situation occurred when the automated lane changing algorithm is used to perform the lane changing of CAVs. On the other hand, the simulated car-following model shows a relatively high number of near-crash conditions, where higher demands and lower number of lanes contribute to higher near-crash situations with smaller TTC value.

Table 7-4: Comparing time to collision between automated lane changing and Vissim simulation

Demand (veh/h/ln)	Number of lanes	Autonomous lane changing		Vissim simulation	
		Ave. TTC (sec)	# near-crash conditions	Ave. TTC (sec)	# near-crash conditions
900	2	>1.5	0	1.4	4
	3	>1.5	0	>1.5	0
	4	>1.5	0	>1.5	0
1500	2	>1.5	0	1.27	2036
	3	>1.5	0	1.32	51
	4	>1.5	0	1.33	22
2400	2	>1.5	0	1.26	3039
	3	>1.5	0	1.27	2298
	4	>1.5	0	1.31	2164

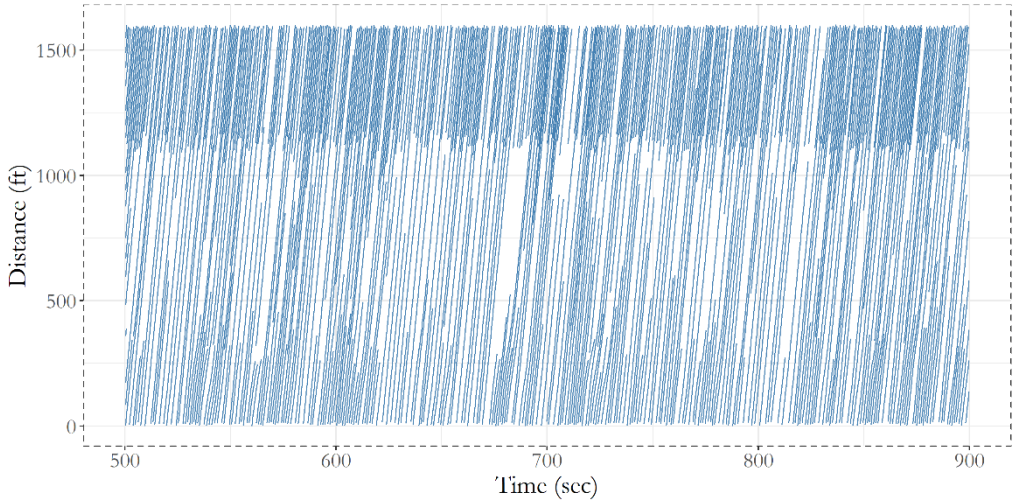
Figure 7-8 shows the spatial average speed distribution on a freeway with four lanes, where the demand is 2400 veh/h/ln and speed limit is 60 mph. As Figure 7-8-a shows, CAVs mostly move with their desired speed in freeway and slow down slightly during the lane changing maneuver. On the other hand, vehicles on the dropped lane with car-following movement have to stop near the bottleneck to achieve the right of way by other vehicles on the road, see Figure 7-8-b.



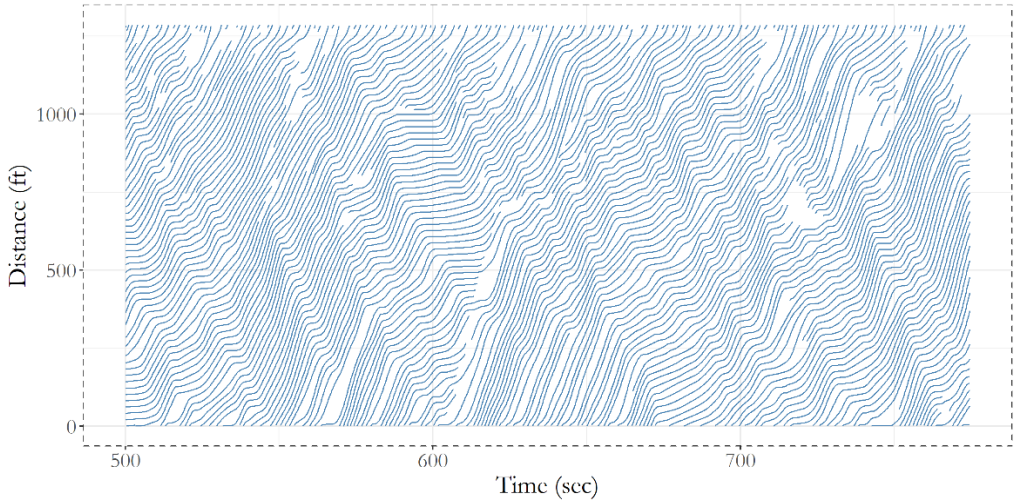
a) Autonomous lane changing b) Vissim simulation

Figure 7-8: Average speed distribution over the freeway facility

Figure 7-9 shows the trajectory of vehicles on the adjacent lane to the dropped lane. Figure 7-9-a shows, CAVs mostly move with their free flow speed before and after the merge of new vehicles from the dropped lane. This performance is achieved due to the cooperation between vehicles to find the right time to merge from the dropped lane to the next lane available. However, Figure 7-9-b shows that vehicles without automated lane changing have to slow down significantly to let the vehicles on the dropped lane make a lane change and move forward.



(a) Automated lane changing



(b) Vissim simulation

Figure 7-9: The trajectory of vehicles on the lane adjacent to the dropped lane

Figure 7-10 shows the average runtime that takes a CAV to find its near-optimal trajectory in a planning horizon. The trends show that increasing the demand is associated with higher runtime because a CAV needs to communicate and negotiate with more CAVs within the detection range. As a results, the number of constraints in the optimization problem of a CAV increases and the problem gets more complex. The runtime results show that CAVs are able to achieve their near optimal solution in real time when demand is not very high, i.e. 900 veh/h/ln or 1500 veh/h/ln. However, the runtime for scenarios with demand 2400 veh/h/ln is more than 0.2 seconds (i.e. the

implementation period). Considering less number of vehicles in a vicinity of a CAV would reduce the problem complexity and help achieve a real-time solution for high-demand scenarios at the cost of finding suboptimal solutions.

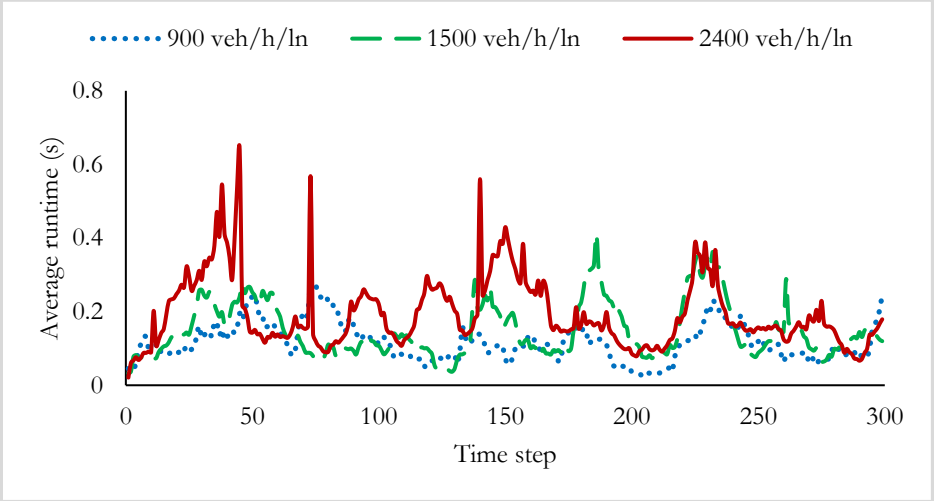


Figure 7-10: Average computational runtime of a CAV-level optimization based on different demand levels

7.7. Conclusion

This study developed a cooperative distributed algorithm for controlling the motion of CAVs integrated with the automated lane changing decisions. The problem is formulated as vehicle-level MINLP to couple the discrete lane changing decision with continuous dynamic of vehicles on a freeway facilities with lane drops. Furthermore, a cooperation scheme is proposed to let CAVs negotiate with each other through vehicle-to-vehicle communications and find collision-free solutions that consider the benefit of the system. An MPC is utilized to capture the dynamics of the problem and provide predictions on vehicle states and future trajectories. Therefore, CAVs solve their trajectory planning problem and update their solutions until they reach agreement on the same solution.

The numerical experiments quantify the effects of the proposed methodology on traffic safety and performance measures in freeway facilities with various number of lanes and speed

limits under various traffic demand scenarios. The results show that automated lane changing improves the mobility performance of freeway facilities by reducing the average travel time by 0%-86.4% and increasing the throughput by 0%-134.3% depending on traffic demand and lane configuration. The most significant improvements are achieved during the over-congested traffic condition and lower number of lanes, where vehicles have less freedom to avoid stop-and-go condition with making a lane change. The automated lane changing algorithm could successfully operate CAVs smoothly with no collision and avoid unnecessary stops upstream of the bottleneck.

As a result of vehicle-level formulation, the complexity of the problem and the computational time are reduced, where the problem could be solved in real-time for under-congested and congested traffic condition. However, we need to further improve the methodology in the future to be able to solve the automated lane changing problem in real-time for over-congested traffic condition. In addition, the effect of automated lane changing of CAVs need to be further investigated when interacting with human-driven vehicles.

CHAPTER 8: CONCLUSIONS

8.1. Summary

Connected and automated vehicle technology, the Internet of Things, and other advance communication technologies create possibilities to facilitate the movement of vehicles through transportation networks and reduce their travel time. Controlling the speed and trajectory of vehicles in transportation networks not only yields a more efficient network capacity utilization, but also regulates the movement of vehicles to achieve a "smoother" flow of traffic. Advisory speed systems can adjust automated vehicles' speeds and consequently their arrival time to signalized intersections to reduce the number of stops and unnecessary acceleration/deceleration. Preventing stop-and-go conditions reduces travel delay, fuel consumption, and yields a more efficient network performance. This dissertation aimed at evaluating the effectiveness of controlling connected and automated vehicles on urban transportation networks.

In this doctoral dissertation, we developed a novel mathematical nonlinear formulation for dynamic speed harmonization in urban street networks to improve traffic operations, especially at signalized intersections. A linear form of complex optimization problem could be achieved by utilizing the characteristics of fundamental flow-density relationship. Developing a predictive control model suitable for dynamic speed optimization problem enabled us to account for stochastic changes in traffic demand and further improve the efficiency of the developed solution algorithm. The results of dynamic speed harmonization on urban street network showed that the proposed mathematical programming could significantly reduce travel time (up to 5.4%), speed variance (19.8%-29.4%), and number of stops (8.3-18.5%), while increasing the average speed (up to 5.9%) and number of completed trips (up to 4%) in tested case studies.

Furthermore, we developed a state-of-the-art algorithm that provides the ability to reduce the complexity of difficult optimization problems and solve them in real time. Distributed optimization and coordination algorithm (DOCA) showed promising results in terms of finding near optimal solutions in real time. DOCA helped analyze large-sized networks in real time by decomposing the network into several subnetworks and provide effective coordination between subnetworks by exchanging the required information among them to push the solutions from subproblems toward global optimality. The result of testing the proposed dynamic speed harmonization problem with DOCA on networks with eight, twenty, and forty intersections yielded a maximum optimality gap of 2.7%. The runtime analysis also showed that controllers could solve the problem in real-time as it took at most 0.53 seconds for a subproblem to solve the corresponding optimization problem.

This research also showed that traffic operations can be further improved by coordinated signal timing and speed optimization. In other words, jointly optimizing signal timing parameters and vehicle speeds can plan the arrival of vehicles to signalized intersections more accurately to utilize green durations more efficiently. The results showed that cooperative signal timing and speed optimization could reduce the average delay 38% and 5.3% in comparison with only speed optimization and only signal timing optimization, respectively. In addition, the number of stops could be reduced by 35.3% and 28.5% in comparison with only speed optimization and only signal timing optimization, respectively.

Since signalized intersections are major bottlenecks in urban street networks, we tried to establish a more accurate strategy that not only benefits self-driving vehicles, but also improves the mobility of human-driven vehicles at signalized intersections. Therefore, we developed a mixed integer nonlinear signal timing and trajectory optimization problem in a mixed traffic

environment of connected automated vehicles (CAVs) and human-driven vehicles. We developed a novel control algorithm that let signal timing parameters and trajectory of vehicles in each lane of the intersection be controlled by a single controller. This algorithm helped reduce the complexity of an NP-hard problem significantly and solve the problem in a reasonable time. The results of the proposed algorithm showed that increasing the penetration rate of CAVs significantly improves the mobility performance of the network. This reduction is more significant in higher traffic volumes.

Finally, we developed an optimal motion planning mathematical program that enables utilizing the potential of CAVs in improving the transportation network to its highest level. A cooperative distributed algorithm is developed for controlling the motion of CAVs integrated with the automated lane changing decisions. A cooperation scheme is proposed to let CAVs negotiate with each other through vehicle-to-vehicle communications and find collision-free solutions that consider the benefit of the system. The numerical experiments showed that the automated lane changing algorithm could successfully operate CAVs smoothly with no collision and avoid unnecessary stops upstream of the bottleneck. In addition, the automated lane changing improves the mobility performance of freeway facilities by reducing the average travel time by 0%-86.4% and increasing the throughput by 0%-134.3% depending on traffic demand and lane configuration.

8.2. Limitations

In this research, we developed algorithms for the optimal control of connected and automated vehicles assuming a perfect connection between vehicles and infrastructures without any latency or interruption. In addition, we assumed that CAVs sense the environment around them perfectly and are aware of traffic condition around them. The information about the current location and speed of CAVs are considered to be accurate, while there is a high chance of error in

measuring these values. Moreover, CAVs are assumed to comply with the advisory speeds all the time. However, these assumptions might not get satisfied in real world applications. It is also assumed that the driving behavior of all CAVs in the system remains the same by considering similar parameters, models, and objective functions for all vehicles. In other words, the driving behavior of CAVs assumed to be homogeneous. More realistic models could be achieved by considering the heterogeneity in the behavior of CAVs.

We developed the network-wide speed harmonization problem based on the cell transmission model, which is a first-order traffic flow model. A more accurate problem formulation could be designed by considering the second-order traffic flow models. In addition, the CTM-based model is assumed to be deterministic with fixed capacity. However, capacity has a stochastic nature. Addressing stochastic capacity in speed harmonization problem is an interesting topic that should be addressed by further research. In addition, the capacity drop cannot be captured by the CTM. As such, the results need to be interpreted cautiously and we suggest further study for incorporating capacity drop in future speed harmonization/optimization studies. We also developed speed harmonization problem for an interrupted traffic flow, where the signalized intersection is the bottleneck. The results provided for our case studies might not be applicable to uninterrupted traffic flow without the presence of bottlenecks.

8.3. Future research suggestions

Following provides a list of suggestions for continuing this research study in the future:

1- This research mainly focused on finding system optimal solutions for the entire system. Studying other objective functions such as user-equilibrium traffic condition provides better insight about the impacts of speed control problems on the driving behavior of CAVs from user perspective.

2- Game theoretic models could better consider the interaction of CAVs when they have different objective functions rather than maximizing the system-level objectives. It is suggested to consider the game theoretic models for analyzing various objective function in solving the speed and trajectory problem.

3- This study mainly developed model-based optimization algorithms to control the speed of CAVs. In addition, the behavior of human drivers is modeled with car-following models. However, the closed form mathematical models might not accurately capture the realistic behavior of both CAVs and human drivers. As a result, it is suggested to utilize data-driven solution techniques with machine learning algorithm to predict the behavior of drivers and state of the system more accurately. Future studies can integrate the machine learning algorithms and mathematical programming to achieve solutions with higher quality that are applicable to real world situations.

4- We showed the effectiveness of Distributed Optimization and Coordination Algorithm (DOCA) in reducing the complexity of the speed control problem. In addition, we showed that DOCA provides near optimal solutions in real time. It is suggested to implement DOCA into other transportation problems such as vehicle routing or ride sharing problems and study its effectiveness on providing high quality solutions in real time for other applications.

5- It is suggested to study the impact of controlling CAVs on the behavior of other drivers on the road. In addition, a sensitivity analysis could be performed on the aggressiveness level of CAVs and its impact on the traffic flow behavior.

6- It is also suggested to consider heterogeneous behaviors for CAVs and consider different driving parameters for different types of CAV to achieve more realistic results.

REFERENCES

Abdel-Aty, M., Dilmore, J., Dhindsa, A., 2006. Evaluation of variable speed limits for real-time freeway safety improvement. Accid. Anal. Prev. 38, 335–345.

Abdelghaffar, H.M., Elouni, M., Bichiou, Y., Rakha, H.A., 2020. Development of a connected vehicle dynamic freeway variable speed controller. IEEE Access 8, 99219–99226.

Ahn, K., Rakha, H., Trani, A., Van Aerde, M., 2002. Estimating vehicle fuel consumption and emissions based on instantaneous speed and acceleration levels. J. Transp. Eng. 128, 182–190.

Ahn, K., Rakha, H.A., Park, S., 2013. Ecodrive application: Algorithmic development and preliminary testing. Transp. Res. Rec. 2341, 1–11.

Ahn, S., Cassidy, M.J., 2007. Freeway traffic oscillations and vehicle lane-change maneuvers. Transp. Traffic Theory 1, 691–710.

Al-Khalili, A.J., 1985. Urban traffic control—A general approach. IEEE Trans. Syst. Man. Cybern. 260–271.

Al Islam, S.M.A. Bin, Aziz, H.M.A., Hajbabaie, A., 2020. Stochastic gradient-based optimal signal control with energy consumption bounds. IEEE Trans. Intell. Transp. Syst.

Al Islam, S.M.A. Bin, Hajbabaie, A., 2017. Distributed coordinated signal timing optimization in connected transportation networks. Transp. Res. Part C Emerg. Technol. 80, 272–285.

Alessandri, A., Di Febbraro, A., Ferrara, A., Punta, E., 1998. Optimal control of freeways via speed signalling and ramp metering. Control Eng. Pract. 6, 771–780.

Alessandri, A., Di Febbraro, A., Ferrara, A., Punta, E., 1999. Nonlinear optimization for freeway control using variable-speed signaling. IEEE Trans. Veh. Technol. 48, 2042–2052.

Alizadeh, A., Moghadam, M., Bicer, Y., Ure, N.K., Yavas, U., Kurtulus, C., 2019. Automated

lane change decision making using deep reinforcement learning in dynamic and uncertain highway environment. In: 2019 IEEE Intelligent Transportation Systems Conference (ITSC). pp. 1399–1404.

America, P., 2014. PTV VISTRO User Manual.

Ammoun, S., Nashashibi, F., Laurgeau, C., 2007. An analysis of the lane changing manoeuvre on roads: the contribution of inter-vehicle cooperation via communication. In: 2007 IEEE Intelligent Vehicles Symposium. pp. 1095–1100.

Araghi, Sahar and Khosravi, Abbas and Creighton, D., 2015. Distributed Q-learning controller for a multi-intersection traffic network. In: International Conference on Neural Information Processing. Springer, pp. 337--344.

Asadi, B., Vahidi, A., 2010. Predictive cruise control: Utilizing upcoming traffic signal information for improving fuel economy and reducing trip time. IEEE Trans. Control Syst. Technol. 19, 707–714.

Askari, A., Farias, D.A., Kurzhanskiy, A.A., Varaiya, P., 2016. Measuring impact of adaptive and cooperative adaptive cruise control on throughput of signalized intersections. arXiv Prepr. arXiv1611.08973.

Aziz, H.M.A., Ukkusuri, S. V., 2012. Integration of environmental objectives in a system optimal dynamic traffic assignment model. Comput. Civ. Infrastruct. Eng. 27, 494–511.

Barkenbus, J.N., 2010. Eco-driving: An overlooked climate change initiative. Energy Policy 38, 762–769.

Barth, M., Mandava, S., Boriboonsomsin, K., Xia, H., 2011. Dynamic ECO-driving for arterial corridors. In: 2011 IEEE Forum on Integrated and Sustainable Transportation Systems. pp. 182–188.

Behrisch, M., Bieker, L., Erdmann, J., Krajzewicz, D., 2011. Sumo--simulation of urban mobility. In: The Third International Conference on Advances in System Simulation (SIMUL 2011), Barcelona, Spain.

Borrough, P., 1997. Variable speed limits reduce crashes significantly in the UK. Urban Transp. Monit.

Brackstone, M., McDonald, M., 1999. Car-following: a historical review. Transp. Res. Part F Traffic Psychol. Behav. 2, 181–196.

Brinckerhoff, P., Farradyne, T., Burgess, J.C., 2008. Active traffic management concept of operations.

Camacho, E.F., Bordons, C., 2012. Model predictive control in the process industry. Springer Science & Business Media.

Camponogara, E., De Oliveira, L.B., 2009. Distributed optimization for model predictive control of linear-dynamic networks. IEEE Trans. Syst. Man, Cybern. A Syst. Humans 39, 1331–1338.

Camponogara, E., Scherer, H.F., 2011. Distributed optimization for model predictive control of linear dynamic networks with control-input and output constraints. IEEE Trans. Autom. Sci. Eng. 8, 233–242.

Carlson, R.C., Papamichail, I., Papageorgiou, M., Messmer, A., 2010. Optimal mainstream traffic flow control of large-scale motorway networks. Transp. Res. Part C Emerg. Technol. 18, 193–212.

Cassidy, M.J., Rudjanakanoknad, J., 2005. Increasing the capacity of an isolated merge by metering its on-ramp. Transp. Res. Part B Methodol. 39, 896–913.

Chen, W., Liu, Y., Yang, X., Bai, Y., Gao, Y., Li, P., 2015. Platoon-based speed control

algorithm for ecodriving at signalized intersection. Transp. Res. Rec. 2489, 29–38.

Chen, Z., Zhang, Y., Lv, J., Zou, Y., 2014. Model for optimization of ecodriving at signalized intersections. Transp. Res. Rec. 2427, 54–62.

CPLEX, I.B.M.I., 2009. V12. 1: User's Manual for CPLEX. Int. Bus. Mach. Corp. 46, 157.

Daganzo, Carlos F., 1994. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. Transp. Res. Part B Methodol. 28, 269–287.

Daganzo, Carlos F, 1994. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. Transp. Res. Part B 28, 269–287.

Daganzo, C.F., 1995. The cell transmission model, part II: Network traffic. Transp. Res. Part B Methodol. 29, 79–93.

Daganzo, C.F., Cassidy, M.J., Bertini, R.L., 1999. Possible explanations of phase transitions in highway traffic. Transp. Res. Part A Policy Pract. 33, 365–379.

De Nunzio, G., De Wit, C.C., Moulin, P., 2014. Urban traffic Eco-Driving: A macroscopic steady-state analysis. In: 2014 European Control Conference (ECC). pp. 2581–2587.

De Wit, C.C., 2011. Best-effort highway traffic congestion control via variable speed limits. In: 2011 50th IEEE Conference on Decision and Control and European Control Conference. pp. 5959–5964.

den Berg, M., Hegyi, A., De Schutter, B., Hellendoorn, J., 2003. A macroscopic traffic flow model for integrated control of freeway and urban traffic networks. In: 42nd IEEE International Conference on Decision and Control (IEEE Cat. No. 03CH37475). pp. 2774–2779.

Ding, G., Aghli, S., Heckman, C., Chen, L., 2018. Game-theoretic cooperative lane changing

using data-driven models. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 3640–3647.

Dresner, K., Stone, P., 2004. Multiagent traffic management: a reservation-based intersection control mechanism. Proc. Third Int. Jt. Conf. Auton. Agents Multiagent Syst. 2004. AAMAS 2004. 530–537.

Edara, P., Sun, C., Hou, Y., 2017. Evaluation of Variable Advisory Speed Limits in Congested Work Zones. J. Transp. Saf. Secur. 9, 123–145.

Erdmann, J., 2013. Combining adaptive junction control with simultaneous green-light-optimal-speed-advisory. In: Wireless Vehicular Communications (WiVeC), 2013 IEEE 5th International Symposium On. pp. 1–5.

Fang, J., Luo, Y., Hadiuzzaman, M., Liu, G., Qiu, T.Z., 2015. Safety oriented variable speed limit control method with enhanced driver response modeling.

Feng, Y., Yu, C., Liu, H.X., 2018. Spatiotemporal intersection control in a connected and automated vehicle environment. Transp. Res. Part C Emerg. Technol. 89, 364–383.

Fisher, M.L., 1981. The Lagrangian relaxation method for solving integer programming problems. Manage. Sci. 27, 1–18.

Gazis, D.C., Herman, R., Weiss, G.H., 1962. Density oscillations between lanes of a multilane highway. Oper. Res. 10, 658–667.

Geoffrion, A.M., 1972. Generalized Benders decomposition. J. Optim. Theory Appl. 10, 237–260.

Ghiasi, A., Ma, J., Zhou, F., Li, X., 2017. Speed Harmonization Algorithm Using Connected Autonomous Vehicles. In: Transportation Research Board.

Gipps, P.G., 1981. A behavioural car-following model for computer simulation. Transp. Res.

Part B Methodol. 15, 105–111.

Greenshields, B.D., Bibbins, J.R., Channing, W.S., Miller, H.H., 1935. A study of traffic capacity. In: Highway Research Board Proceedings.

Grumert, E., Ma, X., Tapani, A., 2015. Analysis of a cooperative variable speed limit system using microscopic traffic simulation. Transp. Res. Part C Emerg. Technol. 52, 173–186.

Guo, Y., Ma, J., Xiong, C., Li, X., Zhou, F., Hao, W., 2019. Joint optimization of vehicle trajectories and intersection controllers with connected automated vehicles: Combined dynamic programming and shooting heuristic approach. Transp. Res. Part C Emerg. Technol. 98, 54–72.

Guoa, Y., Ma, J., 2020. DRL-TP3: A Learning and Control Framework for Signalized Intersections with Mixed Connected Automated Traffic.

Ha, T.-J., Kang, J.-G., Park, J.-J., 2003. The effects of automated speed enforcement systems on traffic-flow characteristics and accidents in Korea. Inst. Transp. Eng. ITE J. 73, 28.

Hadiuzzaman, M., Qiu, T.Z., 2013. Cell transmission model based variable speed limit control for freeways. Can. J. Civ. Eng. 40, 46–56.

Hajbabaie, A., 2012. Intelligent dynamic signal timing optimization program. University of Illinois at Urbana-Champaign.

Hajbabaie, A., Benekohal, R.F., 2013. Traffic Signal Timing Optimization: Choosing the Objective Function. Transp. Res. Rec. J. Transp. Res. Board 2355, 10–19.

Hajbabaie, A., Benekohal, R.F., 2015. A program for simultaneous network signal timing optimization and traffic assignment. IEEE Trans. Intell. Transp. Syst. 16, 2573–2586.

Hajbabaie, A., Medina, J.C., Benekohal, R.F., 2011. Traffic signal coordination and queue management in oversaturated intersection. USDOT Reg. V Reg. Univ. Transp. Cent. Final

Rep. 108.

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.

Hao, P., Wu, G., Boriboonsomsin, K., Barth, M.J., 2018. Eco-approach and departure (EAD) application for actuated signals in real-world traffic. IEEE Trans. Intell. Transp. Syst.

Harbord, B., 1998. M25 controlled motorway-results of the first two years. In: Road Transport Information and Control, 1998. 9th International Conference on (Conf. Publ. No. 454). pp. 149–154.

He, X., Liu, H.X., Liu, X., 2015. Optimal vehicle speed trajectory on a signalized arterial with consideration of queue. Transp. Res. Part C Emerg. Technol. 61, 106–120.

Hegyi, A., 2004. Model predictive control for integrating traffic control measures. Netherlands TRAIL Research School.

Hegyi, A., De Schutter, B., Hellendoorn, H., 2005a. Model predictive control for optimal coordination of ramp metering and variable speed limits. Transp. Res. Part C Emerg. Technol. 13, 185–209.

Hegyi, A., De Schutter, B., Hellendoorn, H., Van Den Boom, T., 2002. Optimal coordination of ramp metering and variable speed control-an MPC approach. In: Proceedings of the 2002 American Control Conference (IEEE Cat. No. CH37301). pp. 3600–3605.

Hegyi, A., De Schutter, B., Hellendoorn, J., 2005b. Optimal coordination of variable speed limits to suppress shock waves. IEEE Trans. Intell. Transp. Syst. 6, 102–112.

Hegyi, A., Hoogendoorn, S.P., Schreuder, M., Stoelhorst, H., Viti, F., 2008. SPECIALIST: A dynamic speed limit control algorithm based on shock wave theory. In: 2008 11th

International IEEE Conference on Intelligent Transportation Systems. pp. 827–832.

Helly, W., 1959. Simulation of bottlenecks in single-lane traffic flow.

HomChaudhuri, B., Vahidi, A., Pisu, P., 2015. A fuel economic model predictive control strategy for a group of connected vehicles in urban roads. In: 2015 American Control Conference (ACC). pp. 2741–2746.

Hwang, C.-L., Masud, A.S.M., 2012. Multiple objective decision making—methods and applications: a state-of-the-art survey. Springer Science & Business Media.

Islam, S.M.A.B. Al, Hajbabaie, A., Aziz, H.M.A., 2020. A real-time network-level traffic signal control methodology with partial connected vehicle information. Transp. Res. Part C Emerg. Technol. 121, 102830.

Jiang, H., Hu, J., An, S., Wang, M., Park, B.B., 2017. Eco approaching at an isolated signalized intersection under partially connected and automated vehicles environment. Transp. Res. Part C Emerg. Technol. 79, 290–307.

Jung, H., Choi, S., Park, B.B., Lee, H., Son, S.H., 2016. Bi-level optimization for eco-traffic signal system. In: Connected Vehicles and Expo (ICCVE), 2016 International Conference On. pp. 29–35.

Kamal, M.A.S., Taguchi, S., Yoshimura, T., 2015. Intersection vehicle cooperative eco-driving in the context of partially connected vehicle environment. In: Intelligent Transportation Systems (ITSC), 2015 IEEE 18th International Conference On. pp. 1261–1266.

Kamalanathsharma, R., Rakha, H., 2014. Agent-Based Simulation of Ecospeed-Controlled Vehicles at Signalized Intersections. Transp. Res. Rec. J. Transp. Res. Board 1–12.

Kamalanathsharma, R.K., Rakha, H.A., 2016. Leveraging connected vehicle technology and telematics to enhance vehicle fuel efficiency in the vicinity of signalized intersections. J.

Intell. Transp. Syst. 20, 33–44.

Kamalanathsharma, R.K., Rakha, H.A., Yang, H., 2015. Networkwide Impacts of Vehicle Ecospeed Control in the Vicinity of Traffic Signalized Intersections. Transp. Res. Rec. J. Transp. Res. Board 91–99.

Kappes, J.H., Savchynskyy, B., Schnörr, C., 2012. A bundle approach to efficient MAP-inference by Lagrangian relaxation. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition. pp. 1688–1695.

Katsaros, K., Kernchen, R., Dianati, M., Rieck, D., 2011. Performance study of a Green Light Optimized Speed Advisory (GLOSA) application using an integrated cooperative ITS simulation platform. In: IWCMC 2011-7th International Wireless Communications and Mobile Computing Conference. pp. 918–923.

Katzourakis, D., de Winter, J.C.F., de Groot, S., Happee, R., 2012. Driving simulator parameterization using double-lane change steering metrics as recorded on five modern cars. Simul. Model. Pract. Theory 26, 96–112.

Kesting, A., Treiber, M., Helbing, D., 2007. General lane-changing model MOBIL for car-following models. Transp. Res. Rec. 1999, 86–94.

Khan, U., Basaras, P., Schmidt-Thieme, L., Nanopoulos, A., Katsaros, D., 2014. Analyzing cooperative lane change models for connected vehicles. In: 2014 International Conference on Connected Vehicles and Expo (ICCVE). pp. 565–570.

Khondaker, B., Kattan, L., 2015. Variable speed limit: A microscopic analysis in a connected vehicle environment. Transp. Res. Part C Emerg. Technol. 58, 146–159.

Kiwiel, K.C., 1990. Proximity control in bundle methods for convex nondifferentiable minimization. Math. Program. 46, 105–122.

Kuge, N., Yamamura, T., Shimoyama, O., Liu, A., 2000. A driver behavior recognition method based on a driver model framework. SAE Trans. 469–476.

Lee, C., Hellinga, B., Saccomanno, F., 2006. Evaluation of variable speed limits to improve traffic safety. Transp. Res. Part C Emerg. Technol. 14, 213–228.

Li, M., Wu, X., He, X., Yu, G., Wang, Y., 2018. An eco-driving system for electric vehicles with signal control under V2X environment. Transp. Res. Part C Emerg. Technol. 93, 335–350.

Li, P., Mirchandani, P., Zhou, X., 2015. Solving simultaneous route guidance and traffic signal optimization problem using space-phase-time hypernetwork. Transp. Res. Part B Methodol. 81, 103–130.

Li, P.T., Zhou, X., 2017. Recasting and optimizing intersection automation as a connected-and-automated-vehicle (CAV) scheduling problem: A sequential branch-and-bound search approach in phase-time-traffic hypernetwork. Transp. Res. Part B Methodol. 105, 479–506.

Li, S., Wang, T., Ren, H., Kong, X., Wang, X., 2020. Coordination Optimization of VSL Strategy on Urban Expressway and Main Road Intersection Signal. IEEE Access 8, 223976–223987.

Li, Y., Ziliaskopoulos, A., Waller, S., 1999. Linear programming formulations for system optimum dynamic traffic assignment with arrival time-based and departure time-based demands. Transp. Res. Rec. J. Transp. Res. Board 1667, 52–59.

Li, Z., Elefteriadou, L., Ranka, S., 2014. Signal control optimization for automated vehicles at isolated signalized intersections. Transp. Res. Part C Emerg. Technol. 49, 1–18.

Li, Z., Liu, P., Xu, C., Wang, W., 2016. Optimal mainline variable speed limit control to improve safety on large-scale freeway segments 31, 366–380.

Liao, R., Chen, X., Yu, L., Sun, X., 2018. Analysis of emission effects related to drivers'

compliance rates for cooperative vehicle-infrastructure system at signalized intersections. Int. J. Environ. Res. Public Health 15, 122.

Lighthill, M.J., Whitham, G.B., 1955. On kinematic waves. II. A theory of traffic flow on long crowded roads. In: Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences. pp. 317–345.

Lin, D., Li, L., Jabari, S.E., 2019. Pay to change lanes: A cooperative lane-changing strategy for connected/automated driving. Transp. Res. Part C Emerg. Technol. 105, 550–564.

Lin, P.-W., Kang, K.-P., Chang, G.-L., 2004. Exploring the effectiveness of variable speed limit controls on highway work-zone operations. In: Intelligent Transportation Systems. pp. 155–168.

Liu, R., Tate, J., 2004. Network effects of intelligent speed adaptation systems. Transportation (Amst). 31, 297–325.

Long, K., Ma, C., Jiang, Z., Wang, Y., Yang, X., 2020. Integrated optimization of traffic signals and vehicle trajectories at intersection with the consideration of safety during signal change. IEEE Access 8, 170732–170741.

Lu, X.-Y., Varaiya, P., Horowitz, R., Su, D., Shladover, S.E., 2010. A new approach for combined freeway variable speed limits and coordinated ramp metering. In: 13th International IEEE Conference on Intelligent Transportation Systems. pp. 491–498.

Luo, Y., Xiang, Y., Cao, K., Li, K., 2016. A dynamic automated lane change maneuver based on vehicle-to-vehicle communication. Transp. Res. Part C Emerg. Technol. 62, 87–102.

Lyles, R.W., Taylor, W.C., Lavansiri, D., Grossklaus, J., 2004. A field test and evaluation of variable speed limits in work zones. Transp. Res. Board 83rd Annu. Meet. 1–21.

Ma, J., Li, X., Shladover, S., Rakha, H.A., Lu, X.-Y., Jagannathan, R., Dailey, D.J., 2016.

Freeway speed harmonization. IEEE Trans. Intell. Veh. 1, 78–89.

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. Transp. Res. Part B Methodol. 95, 421–441.

MacDonald, M., 2008. ATM monitoring and evaluation, 4-lane variable mandatory speed limits 12 month report (primary and secondary indicators). Publ. by Dep. Transp.

Mahler, G., Vahidi, A., 2012. Reducing idling at red lights based on probabilistic prediction of traffic signal timings. In: 2012 American Control Conference (ACC). pp. 6557–6562.

Malikopoulos, A.A., Hong, S., Lee, J., Park, B.B., 2016. Development and evaluation of speed harmonization using optimal control theory. arXiv Prepr. arXiv1611.04647.

Malikopoulos, A.A., Hong, S., Park, B.B., Lee, J., Ryu, S., 2018. Optimal control for speed harmonization of automated vehicles. IEEE Trans. Intell. Transp. Syst. 20, 2405–2417.

Mandava, S., Boriboonsomsin, K., Barth, M., 2009. Arterial velocity planning based on traffic signal information under light traffic conditions. In: Intelligent Transportation Systems, 2009. ITSC'09. 12th International IEEE Conference On. pp. 1–6.

Mattes, S., 2003. The lane-change-task as a tool for driver distraction evaluation. Qual. Work Prod. Enterp. Futur. 57, 60.

Mavrotas, G., 2009. Effective implementation of the epsilon-constraint method in Multi-Objective Mathematical Programming problems. Appl. Math. Comput. 213, 455–465.

McLean, A.J., Anderson, R.W.G., Farmer, M.J.B., Lee, B.H., Brooks, C.G., 1994. Vehicle travel speeds and the incidence of fatal pedestrian collisions (volume 1).

Mehrabipour, M., 2018. Real-time network-level signal timing optimization. Washington State University.

Mehrabipour, M., Hajbabaie, A., 2017. A cell based distributed-coordinated approach for network level signal timing optimization. Comput. Aided Civ. Infrastruct. Eng. an Int. J. 32, 599--616.

Mehrabipour, M., Hajibabai, L., Hajbabaie, A., 2019. A decomposition scheme for parallelization of system optimal dynamic traffic assignment on urban networks with multiple origins and destinations. Comput. Civ. Infrastruct. Eng. 34, 915–931.

Mintsis, E., Vlahogianni, E.I., Mitsakis, E., 2020. Dynamic eco-driving near signalized intersections: Systematic review and future research directions. J. Transp. Eng. Part A Syst. 146, 4020018.

Mirheli, A., Hajibabai, L., Hajbabaie, A., 2018. Development of a signal-head-free intersection control logic in a fully connected and autonomous vehicle environment. Transp. Res. Part C Emerg. Technol. 92, 412–425.

Mirheli, A., Tajalli, M., Hajibabai, L., Hajbabaie, A., 2019. A consensus-based distributed trajectory control in a signal-free intersection. Transp. Res. part C Emerg. Technol. 100, 161–176.

Mirheli, A., Tajalli, M., Mohebifard, R., Hajibabai, L., Hajbabaie, A., 2020. Utilization management of highway operations equipment. Transp. Res. Rec. 0361198120927400.

Misener, J., Shladover, S., Dickey, S., 2010. Investigating the potential benefits of broadcasted signal phase and timing (SPAT) data under IntelliDriveSM. In: ITS America Annual Meeting, Houston, Texas.

Mohebifard, R., Al Islam, S.M.A. Bin, Hajbabaie, A., 2019. Cooperative traffic signal and perimeter control in semi-connected urban-street networks. Transp. Res. Part C Emerg. Technol. 104, 408–427.

Mohebifard, R., Hajbabaie, A., 2018a. Distributed optimization and coordination algorithms for dynamic traffic metering in urban street networks. IEEE Trans. Intell. Transp. Syst. 20, 1–12.

Mohebifard, R., Hajbabaie, A., 2018b. Dynamic traffic metering in urban street networks: Formulation and solution algorithm. Transp. Res. Part C Emerg. Technol. 93, 161--178.

Mohebifard, R., Hajbabaie, A., 2019. Optimal network-level traffic signal control: A benders decomposition-based solution algorithm. Transp. Res. Part B Methodol. 121, 252–274.

Mohebifard, R., Hajbabaie, A., 2020. Effects of automated vehicles on traffic operations at roundabouts. In: 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC). pp. 1–6.

Monteil, J., Billot, R., Sau, J., Armetta, F., Hassas, S., Faouzi, N.-E. El, 2013. Cooperative highway traffic: Multiagent modeling and robustness assessment of local perturbations. Transp. Res. Rec. 2391, 1–10.

Mukadam, M., Cosgun, A., Nakhaei, A., Fujimura, K., 2017. Tactical decision making for lane changing with deep reinforcement learning.

Murray, A.T., Church, R.L., 1996. Analyzing cliques for imposing adjacency restrictions in forest models. For. Sci. 42, 166–175.

Murray, A.T., Church, R.L., 1997. Facets for node packing. Eur. J. Oper. Res. 101, 598–608.

Newell, G.F., 1993. A simplified theory of kinematic waves in highway traffic, part I: General theory. Transp. Res. Part B Methodol. 27, 281–287.

Ngai, D.C.K., Yung, N.H.C., 2011. A multiple-goal reinforcement learning method for complex vehicle overtaking maneuvers. IEEE Trans. Intell. Transp. Syst. 12, 509–522.

Ni, J., Han, J., Dong, F., 2020. Multivehicle cooperative lane change control strategy for

intelligent connected vehicle. J. Adv. Transp. 2020.

Nie, J., Zhang, J., Ding, W., Wan, X., Chen, X., Ran, B., 2016. Decentralized cooperative lane-changing decision-making for connected autonomous vehicles. IEEE Access 4, 9413–9420.

Nilsson, J., Brännström, M., Coelingh, E., Fredriksson, J., 2016. Lane change maneuvers for automated vehicles. IEEE Trans. Intell. Transp. Syst. 18, 1087–1096.

Niroumand, R., Tajalli, M., Hajbabaie, A., Hajibabai, L., 2020a. Joint optimization of vehicle-group trajectory and signal timing: introducing the white phase for mixed-autonomy traffic stream. Transp. Res. part C Emerg. Technol. 116, 102659.

Niroumand, R., Tajalli, M., Hajibabai, L., Hajbabaie, A., 2020b. The effects of the "white phase" on intersection performance with mixed-autonomy traffic stream. In: The 23rd IEEE International Conference on Intelligent Transportation Systems. IEEE.

Nowak, I., 2005. Relaxation and decomposition methods for mixed integer nonlinear programming 152.

Nunzio, G. De, Gutman, P., 2017. An application of shock wave theory to urban traffic control via dynamic speed advisory. 6th Symp. Eur. Assoc. Res. Transp. Eur. Assoc. Res. Transp. 1–7.

Othman, B., De Nunzio, G., Di Domenico, D., Canudas-de-Wit, C., 2020. Variable speed limits control in an urban road network to reduce environmental impact of traffic. In: 2020 American Control Conference (ACC). pp. 1179–1184.

Padberg, M.W., 1973. On the facial structure of set packing polyhedra. Math. Program. 5, 199–215.

Pande, A., Abdel-Aty, M., 2006. Assessment of freeway traffic parameters leading to lane-change related collisions. Accid. Anal. Prev. 38, 936–948.

Panis, L.I., Broekx, S., Liu, R., 2006. Modelling instantaneous traffic emission and the influence of traffic speed limits. Sci. Total Environ. 371, 270–285.

Panwai, S., Dia, H., 2005. Comparative evaluation of microscopic car-following behavior. IEEE Trans. Intell. Transp. Syst. 6, 314–325.

Papageorgiou, M., Blosseville, J.-M., Hadj-Salem, H., 1990. Modelling and real-time control of traffic flow on the southern part of Boulevard Peripherique in Paris: Part I: Modelling. Transp. Res. Part A Gen. 24, 345–359.

Papageorgiou, M., Kosmatopoulos, E., Papamichail, I., 2008. Effects of variable speed limits on motorway traffic flow. Transp. Res. Rec. J. Transp. Res. Board 37–48.

Papageorgiou, M., Mayr, R., 1982. Optimal decomposition methods applied to motorway traffic control. Int. J. Control 35, 269–280.

Ploeg, J., Scheepers, B.T.M., Van Nunen, E., de Wouw, N., Nijmeijer, H., 2011. Design and experimental evaluation of cooperative adaptive cruise control. In: 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC). pp. 260–265.

Pourmehrab, M., Elefteriadou, L., Ranka, S., 2020a. Real-time intersection optimization for signal phasing, timing, and automated vehicles' trajectories. arXiv Prepr. arXiv2007.03763.

Pourmehrab, M., Elefteriadou, L., Ranka, S., Martin-Gasulla, M., 2019. Optimizing signalized intersections performance under conventional and automated vehicles traffic. IEEE Trans. Intell. Transp. Syst.

Pourmehrab, M., Emami, P., Martin-Gasulla, M., Wilson, J., Elefteriadou, L., Ranka, S., 2020b. Signalized intersection performance with automated and conventional vehicles: A comparative study. J. Transp. Eng. Part A Syst. 146, 4020089.

PTV Group, 2015. PTV Vissim 7 User Manual. PTV AG.

Qi, H., Dai, R., Tang, Q., Hu, X., 2020. Coordinated intersection signal design for mixed traffic flow of human-driven and connected and autonomous vehicles. IEEE Access 8, 26067–26084.

Rajamani, R., 2011. Vehicle dynamics and control. Springer Science & Business Media.

Rakha, H.A., Kamalanathsharma, R.K., Ahn, K., others, 2012. AERIS: Eco-vehicle speed control at signalized intersections using I2V communication.

Richards, P.I., 1956. Shock waves on the highway. Oper. Res. 4, 42–51.

Robinson, M.D., 2000. Examples of variable speed limit applications. Transp. Res. Board.

Rodriguez, M., Zhao, X., Song, H., Mavrommati, A., Valenti, R.G., Rajhans, A., Mosterman, P.J., Diaz-Mercado, Y., Fathy, H., 2020. A Gradient-Based Approach for Coordinating Smart Vehicles and Traffic Lights at Intersections. IEEE Control Syst. Lett.

Roncoli, C., Papageorgiou, M., Papamichail, I., 2015. Traffic flow optimisation in presence of vehicle automation and communication systems--Part II: Optimal control for multi-lane motorways. Transp. Res. Part C Emerg. Technol. 57, 260–275.

Sarvi, M., Kuwahara, M., Ceder, A., 2004. Freeway ramp merging phenomena in congested traffic using simulation combined with a driving simulator. Comput. Civ. Infrastruct. Eng. 19, 351–363.

Schakel, W.J., Van Arem, B., 2014. Improving traffic flow efficiency by in-car advice on lane, speed, and headway. IEEE Trans. Intell. Transp. Syst. 15, 1597–1606.

Schakel, W.J., Van Arem, B., Netten, B.D., 2010. Effects of cooperative adaptive cruise control on traffic flow stability. In: 13th International IEEE Conference on Intelligent Transportation Systems. pp. 759–764.

Schick, P., 2003. Influence of traffic control systems on freeway capacity and stability of traffic

flow. Univ. Stuttgart, Stuttgart, Ger. Tech. Rep 20.

Sepulcre, M., Gozalvez, J., 2012. Experimental evaluation of cooperative active safety applications based on V2V communications. In: Proceedings of the Ninth ACM International Workshop on Vehicular Inter-Networking, Systems, and Applications. pp. 13–20.

Sharma, A., Ali, Y., Saifuzzaman, M., Zheng, Z., Haque, M.M., 2017. Human factors in modelling mixed traffic of traditional, connected, and automated vehicles. In: International Conference on Applied Human Factors and Ergonomics. pp. 262–273.

Shen, G., Kong, X., 2009. Study on road network traffic coordination control technique with bus priority. IEEE Trans. Syst. Man, Cybern. Part C (Applications Rev. 39, 343–351.

Shladover, S.E., Su, D., Lu, X.-Y., 2012. Impacts of cooperative adaptive cruise control on freeway traffic flow. Transp. Res. Rec. 2324, 63–70.

Smulders, S., 1990. Control of freeway traffic flow by variable speed signs. Transp. Res. Part B Methodol. 24, 111–132.

Soleimaniamiri, S., Ghiasi, A., Li, X., Huang, Z., 2020. An analytical optimization approach to the joint trajectory and signal optimization problem for connected automated vehicles. Transp. Res. Part C Emerg. Technol. 120, 102759.

Tajalli, M., Hajbabaie, A., 2018a. Distributed optimization and coordination algorithms for dynamic speed optimization of connected and autonomous vehicles in urban street networks. Transp. Res. Part C Emerg. Technol. 95, 497--515.

Tajalli, M., Hajbabaie, A., 2018b. Collision mitigation at signalized intersection using connected vehicles data and technologies. Transp. Res. Board 18.

Tajalli, M., Hajbabaie, A., 2018c. Dynamic speed harmonization in connected urban street

networks. Comput. Civ. Infrastruct. Eng. 33, 510–523.

Tajalli, M., Hajbabaie, A., 2021. Traffic Signal Timing and Trajectory Optimization in a Mixed Autonomy Traffic Stream. IEEE Trans. Intell. Transp. Syst. In press.

Tajalli, M., Mehrabipour, M., Hajbabaie, A., 2020. Network-level coordinated speed optimization and traffic light control for connected and automated vehicles. IEEE Trans. Intell. Transp. Syst. 1–12.

Talebpour, A., Mahmassani, H.S., 2016. Influence of connected and autonomous vehicles on traffic flow stability and throughput 71, 143–163.

Talebpour, A., Mahmassani, H.S., Hamdar, S.H., 2013. Speed harmonization: evaluation of effectiveness under congested conditions. Transp. Res. Rec. 2391, 69–79.

Taylor, M.A.P., 2000. Network modelling of the traffic, environmental and energy effects of lower urban speed limits. Road Transp. Res. 9, 48.

Tettamanti, T., Varga, I., 2010. Distributed traffic control system based on model predictive control 1, 3–9.

Thiemann, C., Treiber, M., Kesting, A., 2008. Estimating acceleration and lane-changing dynamics from next generation simulation trajectory data. Transp. Res. Rec. 2088, 90–101.

Toledo, T., Koutsopoulos, H.N., Ben-Akiva, M., 2007. Integrated driving behavior modeling. Transp. Res. Part C Emerg. Technol. 15, 96–112.

Toledo, T., Zohar, D., 2007. Modeling duration of lane changes. Transp. Res. Rec. 1999, 71–78.

Treiber, M., Kesting, A., 2009. Modeling lane-changing decisions with MOBIL. In: Traffic and Granular Flow’07. Springer, pp. 211–221.

Ulbrich, S., Maurer, M., 2013. Probabilistic online POMDP decision making for lane changes in fully automated driving. In: 16th International IEEE Conference on Intelligent

Transportation Systems (ITSC 2013). pp. 2063–2067.

van de Weg, G.S., Hegyi, A., Hoogendoorn, S.P., De Schutter, B., 2015. Efficient model predictive control for variable speed limits by optimizing parameterized control schemes. In: 2015 IEEE 18th International Conference on Intelligent Transportation Systems. pp. 1137–1142.

Vander Werf, J., Shladover, S.E., Miller, M.A., Kourjanskaia, N., 2002. Effects of adaptive cruise control systems on highway traffic flow capacity. Transp. Res. Rec. 1800, 78–84.

Waller, S.T., Ng, M., Ferguson, E., Nezamuddin, N., Sun, D., 2009. Speed harmonization and peak-period shoulder use to manage urban freeway congestion.

Wan, N., Vahidi, A., Luckow, A., 2016. Optimal speed advisory for connected vehicles in arterial roads and the impact on mixed traffic. Transp. Res. Part C Emerg. Technol. 69, 548--563.

Wang, G., Hu, J., Li, Z., Li, L., 2019. Cooperative lane changing via deep reinforcement learning. arXiv Prepr. arXiv1906.08662.

Wang, J.F., Li, C.C., Lv, J.R., Yan, X.D., 2016. Speed guidance model during the green phase based on a connected vehicle. Simulation 92, 899–905.

Wang, M., Daamen, W., Hoogendoorn, S.P., van Arem, B., 2014a. Rolling horizon control framework for driver assistance systems. Part II: Cooperative sensing and cooperative control. Transp. Res. part C Emerg. Technol. 40, 290–311.

Wang, M., Daamen, W., Hoogendoorn, S.P., van Arem, B., 2014b. Rolling horizon control framework for driver assistance systems. Part I: Mathematical formulation and non-cooperative systems. Transp. Res. part C Emerg. Technol. 40, 271–289.

Wang, M., Daamen, W., Hoogendoorn, S.P., Van Arem, B., 2015a. Connected variable speed

limits control and vehicle acceleration control to resolve moving jams. In: 94th Annual Meeting Transportation Research Board, Washington, USA, 11-15 January 2015; Authors Version.

Wang, M., Hoogendoorn, S.P., Daamen, W., van Arem, B., Happee, R., 2015b. Game theoretic approach for predictive lane-changing and car-following control. Transp. Res. Part C Emerg. Technol. 58, 73–92.

Wang, P., Chan, C.-Y., de La Fortelle, A., 2018. A reinforcement learning based approach for automated lane change maneuvers. In: 2018 IEEE Intelligent Vehicles Symposium (IV). pp. 1379–1384.

Wang, S., 2013. Efficiency and equity of speed limits in transportation networks. Transp. Res. Part C Emerg. Technol. 32, 61–75.

Wang, Z., Chen, X.M., Ouyang, Y., Li, M., 2015. Emission mitigation via longitudinal control of intelligent vehicles in a congested platoon. Comput. Civ. Infrastruct. Eng. 30, 490–506.

Wang, Z., Zhao, X., Xu, Z., Li, X., Qu, X., 2020. Modeling and field experiments on autonomous vehicle lane changing with surrounding human-driven vehicles. Comput. Civ. Infrastruct. Eng.

Weber, A., Winckler, A., 2013. Advanced traffic signal control algorithms, appendix A: Exploratory advanced research project: BMW final report.

Wei, Y., Avc\i, C., Liu, J., Belezamo, B., Ayd\in, N., Li, P.T., Zhou, X., 2017. Dynamic programming-based multi-vehicle longitudinal trajectory optimization with simplified car following models. Transp. Res. part B Methodol. 106, 102–129.

Xi, C., Shi, T., Li, Y., Wang, J., 2019. An efficient motion planning strategy for automated lane change based on mixed-integer optimization and neural networks. In: 2020 IEEE 23rd

International Conference on Intelligent Transportation Systems (ITSC). pp. 1--6.

Xia, H., Boriboonsomsin, K., Barth, M., 2013a. Dynamic eco-driving for signalized arterial corridors and its indirect network-wide energy/emissions benefits. J. Intell. Transp. Syst. 17, 31–41.

Xia, H., Boriboonsomsin, K., Schweizer, F., Winckler, A., Zhou, K., Zhang, W.-B., Barth, M., 2012. Field operational testing of eco-approach technology at a fixed-time signalized intersection. In: 2012 15th International IEEE Conference on Intelligent Transportation Systems. pp. 188–193.

Xia, H., Wu, G., Boriboonsomsin, K., Barth, M.J., 2013b. Development and evaluation of an enhanced eco-approach traffic signal application for connected vehicles. In: 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013). pp. 296–301.

Xiang, X., Zhou, K., Zhang, W.-B., Qin, W., Mao, Q., 2015. A closed-loop speed advisory model with driver's behavior adaptability for eco-driving. IEEE Trans. Intell. Transp. Syst. 16, 3313–3324.

Xu, B., Ban, X.J., Bian, Y., Li, W., Wang, J., Li, S.E., Member, S., Li, K., 2018a. Cooperative method of traffic signal optimization and speed control of connected vehicles at isolated intersections. IEEE Trans. Intell. Transp. Syst. PP, 1–14.

Xu, B., Li, S.E., Bian, Y., Li, S., Ban, X.J., Wang, J., Li, K., 2018b. Distributed conflict-free cooperation for multiple connected vehicles at unsignalized intersections. Transp. Res. Part C Emerg. Technol. 93, 322–334.

Xu, X., Zuo, L., Li, X., Qian, L., Ren, J., Sun, Z., 2018. A reinforcement learning approach to autonomous decision making of intelligent vehicles on highways. IEEE Trans. Syst. Man,

Cybern. Syst.

Yan, C.-Y., Jiang, R., Gao, Z.-Y., Shao, H., 2015. Effect of speed limits in degradable transport networks. Transp. Res. Part C Emerg. Technol. 56, 94–119.

Yang, H., Jin, W.-L., 2014. A control theoretic formulation of green driving strategies based on inter-vehicle communications. Transp. Res. Part C Emerg. Technol. 41, 48–60.

Yang, H., Rakha, H., Ala, M.V., 2016. Eco-cooperative adaptive cruise control at signalized intersections considering queue effects. IEEE Trans. Intell. Transp. Syst. 18, 1575–1585.

Yang, H., Wang, X., Yin, Y., 2012. The impact of speed limits on traffic equilibrium and system performance in networks. Transp. Res. part B Methodol. 46, 1295–1307.

Yang, K., Guler, S.I., Menendez, M., 2016. Isolated intersection control for various levels of vehicle technology: Conventional, connected, and automated vehicles. Transp. Res. Part C Emerg. Technol. 72, 109–129.

Yang, Y., Lu, H., Yin, Y., Yang, H., 2013. Optimization of variable speed limits for efficient, safe, and sustainable mobility. Transp. Res. Rec. J. Transp. Res. Board 37–45.

Yin, B., Menendez, M., Yang, K., 2021. Joint optimization of intersection control and trajectory planning accounting for pedestrians in a connected and automated vehicle environment. Sustainability 13, 1135.

Young, W., Taylor, M.A.P., Gipps, P.G., 1989. Microcomputers in traffic engineering series: No. 1.

Yu, C., Feng, Y., Liu, H.X., Ma, W., Yang, X., 2018. Integrated optimization of traffic signals and vehicle trajectories at isolated urban intersections. Transp. Res. Part B Methodol. 112, 89–112.

Yu, C., Feng, Y., Liu, H.X., Ma, W., Yang, X., 2019. Corridor level cooperative trajectory

optimization with connected and automated vehicles. Transp. Res. Part C Emerg. Technol. 105, 405–421.

Zarkadoula, M., Zoidis, G., Tritopoulou, E., 2007. Training urban bus drivers to promote smart driving: A note on a Greek eco-driving pilot program. Transp. Res. Part D Transp. Environ. 12, 449–451.

Zegeye, S.K., De Schutter, B., Hellendoorn, H., Breunesse, E., 2009. Model-based traffic control for balanced reduction of fuel consumption, emissions, and travel time. IFAC Proc. Vol. 42, 149–154.

Zegeye, S.K., De Schutter, B., Hellendoorn, J., Breunesse, E.A., 2010. Variable speed limits for area-wide reduction of emissions. In: Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference On. pp. 507–512.

Zhao, X., Luh, P.B., 2002. New bundle methods for solving Lagrangian relaxation dual problems. J. Optim. Theory Appl. 113, 373–397.

Zheng, Y., Ran, B., Qu, X., Zhang, J., Lin, Y., 2019. Cooperative lane changing strategies to improve traffic operation and safety nearby freeway off-ramps in a connected and automated vehicles environment. IEEE Trans. Intell. Transp. Syst.

Zheng, Z., Ahn, S., Chen, D., Laval, J., 2011. Freeway traffic oscillations: microscopic analysis of formations and propagations using wavelet transform. Procedia-Social Behav. Sci. 17, 702–716.

Zhicai, J., Xiaoxiong, Z., Hongwei, Y., 2004. Simulation research and implemented effect analysis of variable speed limits on freeway. In: Proceedings. The 7th International IEEE Conference on Intelligent Transportation Systems (IEEE Cat. No. 04TH8749). pp. 894–898.

Zhu, F., Ukkusuri, S. V., 2014. Accounting for dynamic speed limit control in a stochastic traffic

environment: A reinforcement learning approach. Transp. Res. Part C Emerg. Technol. 41, 30–47.

Zhu, F., Ukkusuri, S. V, 2016. Modeling the Proactive Driving Behavior of Connected Vehicles: A Cell-Based Simulation Approach. Comput. Civ. Infrastruct. Eng.

Ziliaskopoulos, A.K., 2000. A Linear Programming Model for the Single Destination System Optimum Dynamic Traffic Assignment Problem. Transp. Sci. 34, 37–49.

Zohdy, I.A., Rakha, H.A., 2014. Intersection Management via Vehicle Connectivity: The Intersection Cooperative Adaptive Cruise Control System Concept. J. Intell. Transp. Syst. 00, 1–16.