

# Scalable Adaptive Traffic Light Control Over a Traffic Network Including Turns, Transit Delays, and Blocking

Yingqing Chen and Christos G. Cassandras

**Abstract**—We develop adaptive data-driven traffic light controllers for a grid-like traffic network considering straight, left-turn, and right-turn traffic flows. The analysis incorporates transit delays and blocking effects on vehicle movements between neighboring intersections. Using a stochastic hybrid system model with parametric traffic light controllers, we use Infinitesimal Perturbation Analysis (IPA) to derive a data-driven cost gradient estimator with respect to controllable parameters. We then iteratively adjust them through an online gradient-based algorithm to improve performance metrics. By integrating a flexible modeling framework to represent diverse intersection and traffic network configurations with event-driven IPA-based adaptive controllers, we develop a general scalable, adaptive framework for real-time traffic light control in multi-intersection traffic networks.

## I. INTRODUCTION

The Traffic Light Control (TLC) problem aims to dynamically optimize traffic light cycles at one or multiple intersections to improve traffic flow and reduce congestion as measured by metrics such as waiting times, vehicle throughput, or traffic backlog. The single intersection TLC problem has been thoroughly studied using different approaches such as model-based optimization [1], computational intelligence [2], [3], and online optimization methods [4]. In addition to pure vehicle flows, some prior works [5] [6] [1] have incorporated pedestrian flows and their interactions with vehicle traffic in addition to optimizing vehicle movement.

However, the transition from single to coordinated multi-intersection control poses at least four key challenges: (i) Different traffic flows (straight, turning, U-turn) propagate through the network, complicating efforts to model network-wide effects; (ii) *Scalable* solutions that extend one-dimensional approaches to high-dimensional multi-intersection settings are difficult to obtain; (iii) The transit delays experienced by traffic between intersections must be accounted for in order to effectively coordinate an intersection with its downstream counterpart(s); and (iv) When the distance between adjacent intersections is relatively short, traffic is often blocked and such blocking effects must be taken into account. Capturing the dynamics of propagating flows, achieving scalable optimization, modeling delays between-intersections, and handling blocking constraints are critical factors to achieve effective multi-intersection traffic light control.

Y. Chen and C. G. Cassandras are with the Division of Systems Engineering and Center for Information and Systems Engineering, Boston University, Brookline, MA 02446 {yqchen; cgc}@bu.edu. This work was supported in part by Red Hat Research (Red Hat Collaboratory at Boston University), by NSF under grants ECCS-1931600, DMS- 1664644, CNS-2149511, and by ARPA-E under grant DE-AR0001282.

Early works on multi-intersection traffic signal control focused on simple scenarios. For example, [7] presented three self-organizing methods for responsive signal timing, and [8] framed fixed-cycle control as a network synchronization problem, only considering offset costs between adjacent intersections. Later approaches expanded controllable parameters to develop centralized controllers optimizing coordination across multiple signals. For example, [9] proposed a steady-state signal control approach which models the traffic light network as a linear time-varying system, and controls the discharging ratio of internal network links. More recently, [10] developed an adaptive linear quadratic regulator to synchronize a network of intersections. However, it is challenging for these model-based approaches to extend to large networks and accommodate the nonlinearities of flow dynamics and stochastic traffic effects.

To deal with the high computational intensity of such centralized optimization methods, various decomposition approaches have been proposed. In [11], the problem is split into nonlinear non-convex subnetwork optimizations to minimize delays and Model Predictive Control (MPC) is used to coordinate subnetwork controllers. Similarly, [12] considers both deterministic traffic demand from commuters and stochastic demand from infrequent travellers. A two-stage stochastic modeling approach is designed to deal with two kinds of demand sequentially by minimizing the expected total travel time at a traffic equilibrium scenario. To address the computational complexity caused by nonlinearities, [13] presents a linear prediction model and employs an MPC controller using a Quadratic Program (QP) formulation, which quadratically penalizes the number of vehicles in the network and linearly penalizes the control decision.

Although many methods have been implemented to reduce computational requirements, especially in large traffic networks, [14] argues that centralized approaches to traffic signal control cannot cope with the increasing complexity of urban traffic networks. Thus, decentralized approaches have emerged. El-Tantawy et al. [15] designed a decentralized framework where agents coordinate through joint policy learning. They later improved agent coordination using modular Q-learning with neighbor information [16]. Modeling duration changing action as a high-dimensional Markov decision process, a deep reinforcement learning framework for traffic light cycle control was developed in [17]. Adding more features, [18] proposed multi-agent reinforcement learning using a Recurrent Neural Network (RNN) and a Graph Neural Network (GNN) to model spatio-temporal influences.

Decentralized techniques aim to improve scalability and

adaptability compared to centralized control. However, both model-based and learning-based approaches mentioned above have downsides in large networks: Centralized optimization faces the “curse of dimensionality” as complexity grows, while decentralized reinforcement learning shifts the computational load to extensive training with large historical datasets, especially for large-scale urban traffic networks. Moreover, many models lack the flexibility to adapt to changing traffic conditions, since they are designed or trained for specific traffic patterns and road layouts. Additional effort is needed to adjust models for new scenarios that may occur. Therefore, key unmet challenges remain centered around the difficulty of balancing optimality, scalability, adaptivity, and flexibility. Open questions persist on coordinating decentralized agents while optimizing system-wide performance under varying demand.

With this motivation, we propose a flexible traffic modeling framework that can incorporate different network topologies. Based on this enhanced network representation, we design scalable adaptive traffic light controllers to optimize signal timings across a multi-intersection structure. Our main contributions are as follows:

(i) We exploit the *scalability* properties of the single-intersection adaptive data-driven TLC approach in [4], [6] based on Infinitesimal Perturbation Analysis (IPA) used to estimate on line performance gradients with respect to the parameters of a TLC controller. In particular, since IPA-based gradient estimators are entirely *event-driven*, these algorithms scale with the (relatively small) number of events in each system intersection, not the (much larger) state space dimensionality [19], [6]. Moreover, IPA is independent of any modeling assumptions regarding the stochastic processes characterizing traffic demand and vehicle behavior, driven only by actual observed traffic data similar to other learning-based approaches.

(ii) We expand the multi-intersection framework in [20] (limited to straight flows) by *incorporating turning flows* into the traffic model. By accounting for vehicle turns at each intersection, the model can better capture the propagation of congestion effects across the intersection network and provides the opportunity to optimize coordinated signal timing plans.

(iii) We incorporate vehicle *transit delays* into the traffic network. Similar to the stochastic hybrid system model in [19], the traffic flow at each intersection depends on the incoming flow from an upstream intersection delayed by the transit time between intersections. In this paper, we provide a simpler way relative to [19] that can capture the flow delay process. With the incorporation of turning flows, this new framework allows us to take advantage of one of the fundamental properties of IPA [21] applied to networks, whereby the effect of a parameter perturbation at one traffic light can only propagate to adjacent traffic lights as a result of a limited and easily tractable number of events; this facilitates the estimation of *network-wide performance* gradients with respect to a local parameter at a specific traffic light.

(iv) We expand the multi-intersection traffic model from

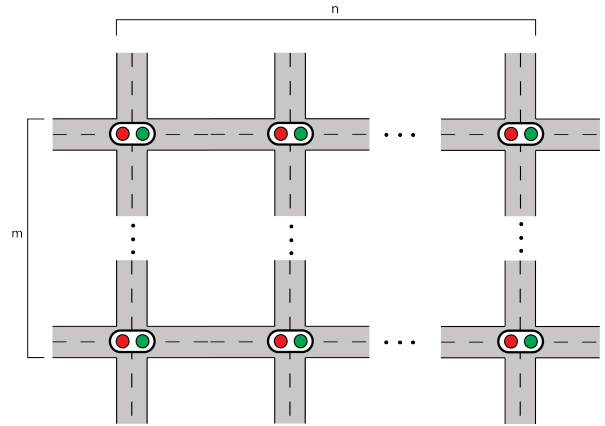


Fig. 1: Grid traffic system

[20] by incorporating finite road capacities and analyzing *blocking effects*. Blocking occurs when downstream congestion causes vehicle queues to propagate backwards, preventing movement at upstream intersections and potentially creating gridlock across the network. The key enhancement is to consider blocking effects not just on the current road segment, but also on upstream segments affected so that the coordinated multi-intersection model can minimize congestion propagation.

(v) We propose a flexible traffic modeling framework by abstracting several traffic features and their relationships over network topologies. By focusing on the traffic features and their interactions rather than specific predefined structures, our framework provides flexibility to model heterogeneous road structures, lane configurations, traffic directions, intersection designs, and traffic signal patterns.

The remainder of this paper is organized as follows. In section II, we formulate the TLC problem for a grid-like multi-intersection network and present the stochastic hybrid system modeling framework. Section III details the derivation of the IPA gradient estimators for a network-wide cost function with respect to a controllable parameter vector. The IPA estimators are then incorporated into a gradient-based optimization algorithm. In Section IV, we conduct multiple simulation experiments under different settings and demonstrate the adaptivity and scalability properties of this approach. Finally, we conclude and discuss future work in Section V.

## II. PROBLEM FORMULATION

We consider a traffic system with  $m$  parallel arterials, each of which contains  $n$  consecutive intersections. A total of  $N$  traffic lights are included in this grid-like network, where  $N = m * n$ , as shown in Fig.1. We assume that the YELLOW cycle is combined with the RED one. The traffic flow towards each intersection  $n$  ( $n = 1, \dots, N$ ) can be identified by the orientation of its origin (East, South, West, North) and direction (straight, left, right), and we define each combination of these three attributes as a *movement*. For example, movement [n-E-I] denotes the direction for

traffic that starts from East and turns left at intersection  $n$ . We categorize movements into two types: controllable and uncontrollable. Controllable traffic flows can be directed by traffic lights, while uncontrolled traffic is free to move without signals, such as some right-turning vehicle flows.

We can further group movements by their location since the traffic flow for certain movements is confined to specific lanes on the road. Note that several different movements can share a single lane rather than having dedicated lanes for each movement. We model each lane segment between two adjacent intersections as a *queue* where vehicles stop when the light for the associated movement is RED. We define the set of all queues at intersection  $n$  as  $Q_n$ . As in actual traffic systems, a traffic light gives right-of-way by showing GREEN to a particular set of controllable movements. After its GREEN cycle time, the light pattern changes to allow the vehicles from the next group of movements to pass through. Therefore, we define a set of movements that are enabled at the same time as a *phase*, and a traffic light cycle for each intersection  $n = 1, \dots, N$  contains a predefined sequence of  $k$  phases  $P_n = (p_{n,1}, p_{n,2}, \dots, p_{n,k})$ . When a phase is enabled, all movements inside that phase are facing a GREEN light with all associated queues such that their movements are consistent in terms of traffic rules. Therefore, we denote the corresponding set of enabled queues within phase  $p \in P_n$  as  $Q_p$ , where  $Q_p \subset Q_n$ . Note that a single queue could appear in different phases. This indicates that when the GREEN light switches from one phase to the next consecutive phase, some queues may remain enabled if they are included in both phases. Additionally, note that each queue with controllable movements should be enabled at least once during a full signal cycle. This ensures that all movements get a chance to proceed during the cycle.

We now define the corresponding queue content state variable  $x_q(t)$  for each queue  $q$  in  $Q_n$ , where  $n = 1, \dots, N$ , and  $x_q(t) \in \mathbb{R}_0^+$ . The potential downstream queues for each queue  $q$  are determined based on the movements accepted in  $q$  and are denoted as  $Q_q^D$ . Similarly, the set of upstream queues for queue  $q$  is denoted as  $Q_q^U$ . With these set definitions, we capture all the flow connection relationships between queues.

As an example, Fig.2 shows a portion of a traffic network structure where each road segment contains 2 queues - one for left-turn movements and one for straight and right-turn movements. The two intersections have a similar structure, so we will focus on modeling intersection 1. The queue set for intersection 1 is  $Q_1 = [q_1, q_2, q_3, q_4, q_5, q_6, q_7, q_8]$ . We designate all right-turn movements as uncontrollable, aligning with many real-world traffic regulations. For the remaining movements, we define 4 phases: (a) Phase  $p_1$  includes straight movement from West [W-s] on queue  $q_2$  and straight movement from East [E-s] on queue  $q_6$  ( $q_2, q_6 \in Q_{p_1}$ ); (b) Phase  $p_2$  includes left turns from West [W-l] on  $q_1$  and from East [E-l] on  $q_5$  ( $q_1, q_5 \in Q_{p_2}$ ); (c) Phase  $p_3$  includes straight flows from South [S-s] on  $q_4$  and from North [N-s] on  $q_8$  ( $q_4, q_8 \in Q_{p_3}$ ); and (d) Phase  $p_4$  includes left turns from South [S-l] on  $q_3$  and from North [N-l]

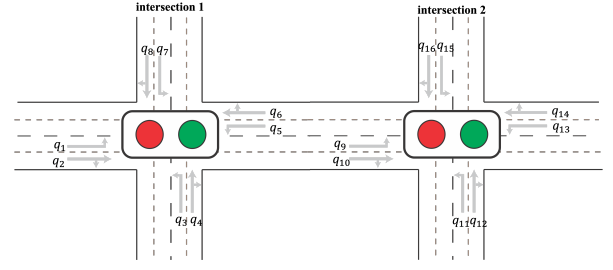


Fig. 2: Traffic Structure Demonstration

on  $q_7$  ( $q_3, q_7 \in Q_{p_4}$ ). To connect adjacent intersections, we can specify downstream queue sets under 2-intersection scope, e.g.  $Q_{q_2}^D = \{q_9, q_{10}\}$ , and upstream queue sets, e.g.  $Q_{q_9}^U = \{q_2, q_4, q_7\}$ .

With these definitions, we can represent several key features of transportation networks: (a) Different acceptable traffic directions (e.g. straight, left-turn or right-turn) are defined by specifying allowable movements; (b) Road structures with different numbers of lanes are represented by defining multiple queues for road segments; (c) Lane occupancy rules regarding which movements can use which lanes are encoded by assigning sets of allowed movements to each queue; (d) Intersection structures like 4-way or 3-way intersections are captured by properly defining the upstream and downstream queue connectivity for each queue based on the intersection geometry; (e) Traffic light patterns are specified by defining the corresponding sequence of phases. In summary, this network representation is flexible enough to model heterogeneous road structures, lane configurations, traffic directions, intersection designs and traffic signal patterns through careful specification of movements, queues, phases and their relationships. This allows the analytical modeling of traffic flows through diverse transportation network topologies.

**Blocking.** Blocking arises when the road length between two neighboring intersections is not long enough or the green cycle for the in-between queues is too short so that the probability that one or more queues are filled with vehicles is strictly positive. We denote the capacity of queue  $q$  as  $c_q$  so that when  $x_q(t)$  reaches  $c_q$  from below, blocking occurs and no more vehicles can join this queue. Note that such blocking may not only affect the arrival process at queue  $q$ , but also halt the departure processes of all upstream queues  $q^u \in Q_q^U$ , since the vehicles in them are not able to proceed.

**State Dynamics.** As in prior work, e.g., [4], we model the input and output to each queue as two stochastic flow processes  $\{\alpha_q(t)\}$  and  $\{\beta_q(t)\}$ , where  $\alpha_q(t)$  and  $\beta_q(t)$  are the *stochastic* instantaneous arrival and departure rate at queue  $q$  respectively. We can then write the queue flow dynamics as

$$\dot{x}_q(t) = \alpha_q(t) - \beta_q(t), \quad q \in Q_n, n = 1, \dots, N \quad (1)$$

Unlike a single intersection, where  $\{\alpha_q(t)\}$  is an exogenous arrival process independent of the state of the system, in the multi-intersection case, the arrival flow process is determined by the departure process of its upstream queues, except

when there are no upstream queues, i.e.,  $q$  is such that  $Q_q^U = \emptyset$ . The arrival rate  $\alpha_q(t)$  for  $q \in \{q | Q_q^U \neq \emptyset\}$  will be defined in the sequel so as to properly capture its dependence on  $\beta_{q^u}(t), q^u \in Q_q^U$  and include the effect of flow transit delays. On the other hand,  $\{\beta_q(t)\}$  depends on the corresponding traffic light control, denoted by  $v_q(t)$ , where  $v_q(t) = 1$  corresponds to a GREEN phase for queue  $q$ , and  $v_q(t) = 0$  to a RED phase, respectively. We set  $v_q(t)$  to be a right-continuous function of  $t$ . The departure process is also influenced by whether or not the downstream queue is blocked, since a blocked downstream queue prevents vehicles in the corresponding upstream queue from moving, hence  $\beta_q(t) = 0$ . Therefore,  $\beta_q(t)$  can be expressed as follows:

$$\beta_q(t) = \begin{cases} h_q(t), & \text{if } x_q(t) > 0, v_q(t) = 1, \\ & \text{and } x_{q_d}(t) < c_{q_d} \quad \forall q_d \in Q_q^D \\ \alpha_q(t), & \text{if } x_q(t) = 0, v_q(t) = 1, \\ & \text{and } x_{q_d}(t) < c_{q_d} \quad \forall q_d \in Q_q^D \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

where  $q \in Q_n, n = 1, \dots, N$  and  $h_q(t)$  is the unconstrained departure rate for which appropriate models can be used (see Section IV).

Next, we ensure that the control of a queue is consistent with phase control. Given a phase sequence  $P_n = (p_{n,1}, p_{n,2}, \dots, p_{n,k})$ , we require that when the GREEN cycle ends for movements in one phase, then the GREEN signals for the next phase movements in the sequence immediately turn on, while the remaining movements face a RED light. We denote the control for phase  $p$  as  $u_p(t)$  for  $p \in P_n, n = 1, \dots, N$ , where  $u_p(t) = 1$  indicates a GREEN signal for all movements in  $p$ , otherwise  $u_p(t) = 0$ . We define  $u_p(t)$  to be right-continuous in order to accurately represent the control policy defined in the sequel. Moreover, for any intersection  $n$ , we require  $\sum_{p \in P_n} u_p(t) = 1$ . Therefore, for any  $q \in Q_p$ , we can always set the control of queue  $q \in Q_p$  once the phase  $p$  controls are decided:

$$v_q(t) = \begin{cases} 1, & \text{if } u_p(t) = 1 \text{ for some } p \in P_n, q \in Q_p \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

Note that it is possible for a particular queue to be enabled during several consecutive phase cycle.

Finally, we define clock state variables  $z_p(t)$  to measure the time since the current enabled phase last switches to  $p$ . The dynamics of  $z_p(t)$  are:

$$\dot{z}_p(t) = \begin{cases} 1, & \text{if } u_p(t) = 1 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

We define  $z_p(t)$  to be left-continuous, so that when the enabled phase switches from  $p$  to the next phase,  $z_p(t) > 0$ ,  $z_p(t^+) = 0$ , and  $u_p(t) = 0$  (since  $u_p(t)$  is right-continuous).

**Traffic Light Control.** An effective controller design needs to address two issues: (i) maintaining a proper balance between allocating a GREEN light to competing queues and (ii) preventing the undesired phenomenon where vehicles wait at a RED light while competing queues are empty during

their GREEN phase. Such “waiting-for-nothing” instances waste the resources of vehicles that wait unnecessarily and can be eliminated through a proper controller design as detailed next. Towards these two goals, we design a quasi-dynamic controller such that only partial state information is needed. Such control is based on the ability of current sensors to detect events of interest in the state dynamics above, such as a queue content becoming empty. On the other hand, it may not be possible to detect the exact number of vehicles in a queue (e.g., using cameras); therefore, we assume that this number can be estimated so as to classify a queue content  $x_q(t)$  as being empty and either below or above some *queue content threshold*, as well as the time such transitions occur. We define an adjustable queue content threshold, denoted by  $s_p$ , for each phase  $p$ . When phase  $p$  is enabled, the detector uses the same threshold value  $s_p$  for all queues  $q \in Q_n$ . We also assign a guaranteed minimum GREEN light cycle time  $\theta_p^{\min}$  and a maximum cycle time  $\theta_p^{\max}$  for each phase  $p$ . This is to ensure that traffic light switches are not overly frequent nor can they be excessively long. Thus, a controller for phase  $p$  depends on a controllable parameter vector for each phase:

$$\theta_p = [\theta_p^{\min}, \theta_p^{\max}, s_p] \quad (5)$$

for  $p \in P_n, n = 1, \dots, N$ , where  $\theta_p^{\min} \geq 0, \theta_p^{\max} \geq \theta_p^{\min}, s_p \geq 0$ . We also define the complete parameter set  $\Theta = \{\theta : \theta \in \theta_p, p \in P_n, n = 1, \dots, N\}$ .

In order to partition the queue content state space of intersection  $n$  which is currently under phase  $p \in P_n$  ( $u_p(t) = 1$ ), we start by defining two auxiliary variables indicating the (instantaneous) largest queue content of any queue in phase  $p$  and that of any queue not in phase  $p$  as  $x_p^{\max}(t) = \max_{q \in Q_p} x_q(t)$  and  $x_{\bar{p}}^{\max}(t) = \max_{q \in Q_n \setminus Q_p} x_q(t)$ . Based on the value of these two variables, we partition the queue content state space into the following regions (as shown in Fig.3):

$$\begin{aligned} X_{n,0} &= \{(x_p^{\max}, x_{\bar{p}}^{\max}) : x_p^{\max}(t) = 0, x_{\bar{p}}^{\max}(t) = 0\} \\ X_{n,1} &= \{(x_p^{\max}, x_{\bar{p}}^{\max}) : x_p^{\max}(t) > 0, x_{\bar{p}}^{\max}(t) = 0\} \\ X_{n,2} &= \{(x_p^{\max}, x_{\bar{p}}^{\max}) : x_p^{\max}(t) = 0, x_{\bar{p}}^{\max}(t) > 0\} \\ X_{n,3} &= \{(x_p^{\max}, x_{\bar{p}}^{\max}) : 0 < x_p^{\max}(t) < s_p, 0 < x_{\bar{p}}^{\max}(t) < s_p\} \\ X_{n,4} &= \{(x_p^{\max}, x_{\bar{p}}^{\max}) : 0 < x_p^{\max}(t) < s_p, x_{\bar{p}}^{\max}(t) \geq s_p\} \\ X_{n,5} &= \{(x_p^{\max}, x_{\bar{p}}^{\max}) : x_p^{\max}(t) \geq s_p, 0 < x_{\bar{p}}^{\max}(t) < s_p\} \\ X_{n,6} &= \{(x_p^{\max}, x_{\bar{p}}^{\max}) : x_p^{\max}(t) \geq s_p, x_{\bar{p}}^{\max}(t) \geq s_p\} \end{aligned}$$

Therefore, if intersection  $n$  is currently in phase  $p \in P_n$ , the quasi-dynamic traffic light controller follows the following rules:

1.  $(x_p^{\max}, x_{\bar{p}}^{\max}) \in \{X_{n,1}\}$ : In this case, there is no vehicle flow for any queue in  $\{\bar{q} | \bar{q} \in Q_n, \bar{q} \notin Q_p\}$  competing with queue  $q$  while queues enabled in the current phase are not empty. In this case, the light for the current phase should remain GREEN:

$$u_p(t) = 1 \quad (6)$$

2.  $(x_p^{\max}, x_{\bar{p}}^{\max}) \in \{X_{n,2}\}$ : In this case, there is no vehicle flow for queues enabled in the current phase  $p$ , while

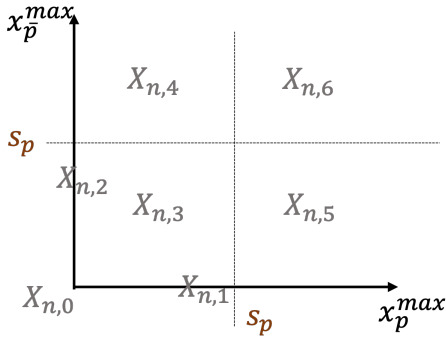


Fig. 3: State space representation

competing queues are not empty. In this case, the current phase should switch to RED immediately:

$$u_p(t) = 0 \quad (7)$$

3.  $(x_p^{max}, x_{\bar{p}}^{max}) \in \{X_{n,0}, X_{n,3}, X_{n,5}, X_{n,6}\}$ : In this case, vehicle traffic for queues in the current phase  $p$  is either higher or balanced compared to competing queues. Thus, the current phase should be prioritized and remain GREEN until it reaches its maximum cycle time:

$$u_p(t) = \begin{cases} 1, & \text{if } z_p(t) \in (0, \theta_p^{max}) \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

4.  $(x_p^{max}, x_{\bar{p}}^{max}) \in \{X_{n,4}\}$ : In this case, there is low vehicle traffic demand for queues in the current phase  $p$  and high demand for the other queues, thus the system should switch phases to accommodate the excess demand:

$$u_p(t) = \begin{cases} 1, & \text{if } z_p(t) \in (0, \theta_p^{min}) \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

Note that when the traffic light for movements in the current phase  $p$  switches from GREEN to RED, the movements for the next phase in the predefined phase sequence  $P_n = (p_{n,1}, p_{n,2}, \dots, p_{n,k})$  switches from RED to GREEN accordingly.

**Events.** The state transitions in the model defined through (2), (1), and (4) under the controller (6), (7), (8) and (9) are dictated by ten possible events defined as follows (see also Table I).

**Basic events.** For  $n = 1, \dots, N, q \in Q_n$ : (a)  $[x_q \downarrow 0]$ :  $x_q(t)$  reaches 0 from above, (b)  $[x_q \uparrow 0]$ :  $x_q(t)$  becomes positive from 0, (c)  $[x_q \downarrow s_p]$ :  $x_q(t)$  reaches  $s_p$  from above, (d)  $[x_q \uparrow s_p]$ :  $x_q(t)$  reaches  $s_p$  from below, (e)  $[x_q \downarrow c_q]$ :  $x_q(t)$  reaches  $c_q$  from above, which indicates the end of a blocking interval, (f)  $[x_q \uparrow c_q]$ :  $x_q(t)$  reaches  $c_q$  from below, which indicates the start of a blocking interval, (g)  $[z_p \uparrow \theta_p^{min}]$ :  $z_p(t)$  reaches its lower bound, (h)  $[z_p \uparrow \theta_p^{max}]$ :  $z_p(t)$  reaches its upper bound, (i)  $[\alpha_q \uparrow 0]$ :  $\alpha_q$  becomes positive from 0, (j)  $[\alpha_q \downarrow 0]$ :  $\alpha_q$  reaches 0 from above.

**Light switching events.** It is convenient to define the following light switching events which are induced by basic events according to the control rules. For  $n = 1, \dots, N, q \in Q_n$ :

- $R2G_q$ : traffic light for controllable movements in queue  $q$  at intersection  $n$  switches from RED to GREEN.
- $G2R_q$ : traffic light for controllable movements in queue  $q$  at intersection  $n$  switches from GREEN to RED.

#### A. Flow Burst Modeling and Analysis

A major consideration in studying a multi-intersection system is the role of a *flow burst* which is generated at each intersection and has impact on the downstream intersection. This was first studied in [19] for only two intersections and only straight flows using a complicated sequence of sub-processes. Here, we provide a more direct way to model such flow bursts that easily extends to multiple intersections and includes turning flows. Specifically, when a RED light switches to GREEN for queue  $q$ , a new flow burst is generated consisting of all vehicles queued at  $q$  that are released. The impact this flow burst has on any downstream queue  $q^d \in Q_q^D$  is caused by the start and end of this flow generation during each GREEN phase.

We start by defining event  $G_{q,m}$  for intersection  $n = 1, \dots, N, q \in Q_n, m = 1, 2, \dots$  causing the generation of a flow burst from queue  $q$  during its  $m$ th GREEN cycle. This event is induced at time  $t$  by one of two already defined events: (a)  $R2G_q$  if  $x_q(t) > 0$ , or (b) the first  $[\alpha_q \uparrow 0]$  event within the  $m$ th GREEN cycle if  $x_q(t) = 0$ .

When  $G_{q,m}$  occurs,  $\beta_q(t)$  increases from zero to positive and the corresponding flow burst will join the downstream queue  $q^d \in Q_q^D$  (empty or not) following a time delay, which then results in the downstream arrival rate increasing from 0 to positive, thus creating an interdependence of flow rates between neighboring intersections. The delay depends on (a) the average speed of the flow burst, denoted by  $f_q$ , affected by road quality and driver behavior, (b) the road length  $L_{q^d}$  where downstream queue  $q^d$  is located, and (c) the average length  $l$  of a vehicle (including a safe distance between two vehicles). Then, the flow process relationship between any two adjacent intersections caused by such delay can be estimated as:

$$\alpha_{q^d}(t) = \sum_{q \in Q_{q^d}^U} \gamma_{q,q^d} \beta_q(t - \frac{L_{q^d} - x_{q^d}(t) * l}{f_q}) \quad (10)$$

where  $\gamma_{q,q^d}$  is the ratio of flow from  $q$  that is directed to  $q^d$ . Note that  $L_{q^d} - x_{q^d}(t) * l \geq 0$  always holds even when blocking happens, which implies that  $\alpha_{q^d}(t)$  is always dependent on the upstream output process. For convenience, set  $\Delta_{q,q^d}(t) = \frac{L_{q^d} - x_{q^d}(t) * l}{f_q}$  to be the delay between queues  $q$  and  $q^d$  in (10) and note that  $\Delta_{q,q^d}(t) \in [0, \frac{L_{q^d}}{f_q}]$ . Thus, (10) provides the relationship for coupling the flow rates of two consecutive intersections, which allows us to identify events and their corresponding event time derivatives with respect to the controllable parameters  $\theta \in \Theta$ , as shown in the next section, without needing to know the value of  $\gamma_{q,q^d}$ . We also assume that any two flow bursts generated from the same intersection but different GREEN cycles will not join each other before the first one joins the downstream queue.

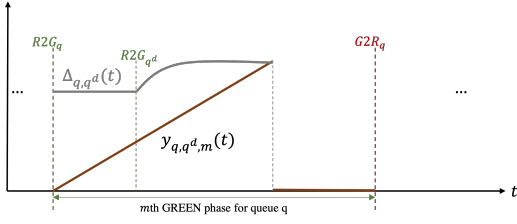


Fig. 4: Example for trajectories of  $y_{q,q^d,m}(t)$  and  $\Delta_{q,q^d}(t)$

We now define a second clock variable  $y_{q,q^d,m}(t) \in \mathbb{R}_0^+$  for intersection  $n = 1, \dots, N$ ,  $q \in Q_n$ ,  $q^d \in Q_n^D$ ,  $m = 1, 2, \dots$ , similar to (4), which denotes the time elapsed since  $G_{q,m}$  occurs. Its dynamics are

$$\dot{y}_{q,q^d,m}(t) = \begin{cases} 1, & \text{if } y_{q,q^d,m}(t) \in (0, \Delta_{q,q^d}(t)) \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

The clock variable  $y_{q,q^d,m}(t)$  has a *single* cycle for each flow burst  $m$ . When  $G_{q,m}$  occurs, we set  $y_{q,q^d,m}(t) = 0$  and initialize (11) so that  $y_{q,q^d,m}(t^+) > 0$ . As soon as  $y_{q,q^d,m}(t) = \Delta_{q,q^d}(t)$ , we set  $y_{q,q^d,m}(t^+) = 0$ . An example is shown in Fig. 4. Note that  $y_{q,q^d,m}(t) > 0$  is possible for multiple different  $m$ , i.e., several flow bursts may be active generated from the same intersection but during different GREEN cycles. This may happen when  $L_{q^d}$  is large and the flow bursts have yet to reach the downstream queue. Also note that when  $y_{q,q^d,m}(t) = \Delta_{q,q^d}(t)$ , this implies that the head of a flow burst generated by event  $G_{q,m}$  joins the downstream queue  $q^d$  at time  $t$ . This is the first instant when this flow burst can have an impact on the downstream queue. Therefore, we define it as an event  $J_{q,q^d,m}$ , which also induces an event  $[\alpha_{q^d} \uparrow 0]$ , i.e., the input flow for queue  $q^d$  becomes positive again.

Similarly, in order to model the end of a flow burst, we define an event  $G_{q,m}^e$ , which indicates the end of a flow burst generation from queue  $q$  during the  $m$ th GREEN cycle. This can be induced by (a)  $G2R_q$  if either  $\alpha_q(t) > 0$  or  $x_q(t) > 0$ , or (b) the last  $[\alpha_q \downarrow 0]$  event inside the  $m$ th GREEN cycle if  $x_q(t) = 0$ . This leads to the definition of another clock state variable  $r_{q,q^d,m}(t) \in \mathbb{R}_0^+$  for intersection  $n$  where  $q \in Q_n$ ,  $q^d \in Q_n^D$  and  $m = 1, 2, \dots$ , which measures the time elapsed since  $G_{q,m}^e$  occurs. Its dynamics are

$$\dot{r}_{q,q^d,m}(t) = \begin{cases} 1, & \text{if } r_{q,q^d,m}(t) \in (0, \Delta_{q,q^d}(t)) \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

When  $G_{q,m}^e$  occurs, we set  $r_{q,q^d,m}(t) = 0$  and initialize (12) so that  $r_{q,q^d,m}(t^+) > 0$ . As soon as  $r_{q,q^d,m}(t) = \Delta_{q,q^d}(t)$ , we set  $r_{q,q^d,m}(t^+) = 0$ . Similar to  $y_{q,q^d,m}(t)$ , it is possible that  $r_{q,q^d,m}(t) > 0$  for multiple different  $m$ . Note that when  $r_{q,q^d,m}(t) = \Delta_{q,q^d}(t)$ , this implies that the flow burst generated from queue  $q$  at the  $m$ th GREEN cycle ceases to have any impact on the downstream queue  $q^d$ . Therefore, we define it as an event  $J_{q,q^d,m}^e$ , which also induces an event  $[\alpha_{q^d} \downarrow 0]$ . Since we consider a single flow burst generated by the same GREEN phase, any  $[\alpha_{q^d} \uparrow 0]$  or  $[\alpha_{q^d} \downarrow 0]$  between

TABLE I: List of Events

|                           |                                                                                                                                                                                                                                                  |
|---------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Basic Events              | $[x_q \downarrow 0], [x_q \uparrow 0], [x_q \downarrow s_p], [x_q \uparrow s_p]$<br>$[x_q \downarrow c_q], [x_q \uparrow c_q], [z_p \uparrow \theta_p^{min}]$<br>$[z_p \uparrow \theta_p^{max}], [\alpha_q \uparrow 0], [\alpha_q \downarrow 0]$ |
| Light Switching Events    | $R2G_q, G2R_q$                                                                                                                                                                                                                                   |
| Flow Burst Tracing Events | $G_{q,m}, J_{q,q^d,m}, G_{q,m}^e, J_{q,q^d,m}^e$                                                                                                                                                                                                 |

$J_{q,q^d,m}$  and  $J_{q,q^d,m}^e$  is exogenous and has no impact on the downstream queue state.

In summary, the generation and impact of a flow burst is captured through the four additional *flow burst tracing events*  $G_{q,m}, J_{q,q^d,m}, G_{q,m}^e, J_{q,q^d,m}^e$ , which facilitate the IPA gradient evaluation in Section III. The full list of events is shown in Table I.

The multi traffic light intersection system can be viewed as a hybrid system in which the time-driven dynamics are given by (1), (4), (11), (12) and (2), while the event-driven dynamics are dictated by the basic events in Table I; these induce associated light switching and flow burst tracing events (defined for convenience). Although the dynamics are based on knowledge of the instantaneous flow processes  $\{\alpha_q(t)\}$  and  $\{\beta_q(t)\}$ , we will show that the IPA-based adaptive controller we design *does not require such knowledge* and depends only on estimating some rates in the vicinity of certain critical observable events.

### B. TLC Optimization Problem

With the parameterized controller defined through  $X_{n,0}, \dots, X_{n,6}$  and (6)-(9), our aim is to optimize a performance metric for the intersection operation with respect to the controllable parameters that comprise  $\Theta$  defined through (5). We choose our performance metric to be the weighted mean of all queue lengths over a fixed time interval  $[0, T]$ :

$$L(\Theta; x(0), z(0), T) = \frac{1}{T} \sum_{n=1}^N \sum_{q \in Q_n} \int_0^T \omega_q x_q(t; \Theta) dt \quad (13)$$

where  $\omega_q$  is a weight associated with queue  $q$ . In order to focus on the structure of a typical sample path of the hybrid system, observe that the sample path of any flow queue content  $\{x_q(t)\}$  consists of alternating *Non-empty Periods* (NEPs) and *Empty Periods* (EPs), which correspond to time intervals when  $x_q(t) > 0$  and  $x_q(t) = 0$  respectively, as shown in Fig. 5. We define two additional events:  $S_q$  for starting NEPs and  $E_q$  for ending them, both induced by basic events defined earlier. Moreover, we denote the  $k$ th NEP of queue  $q$  by  $[\xi_{q,k}, \eta_{q,k}]$  where  $\xi_{q,k}, \eta_{q,k}$  are the occurrence times of the  $k$ th  $S_q$  event and  $k$ th  $E_q$  event respectively. Inside NEPs, we further define *Blocking Periods* (BPs) when  $x_q(t) = c_q$ , indicating blocking happens. Since  $x_q(t) = 0$  during EPs of queue  $q$ , the sample function  $L(\Theta; x(0), z(0), T)$  in (13) can be rewritten as

$$L(\Theta; x(0), z(0), T) = \frac{1}{T} \sum_{n=1}^N \sum_{q \in Q_n} \sum_{k=1}^{K_q} \int_{\xi_{q,k}}^{\eta_{q,k}} \omega_q x_q(t; \Theta) dt \quad (14)$$

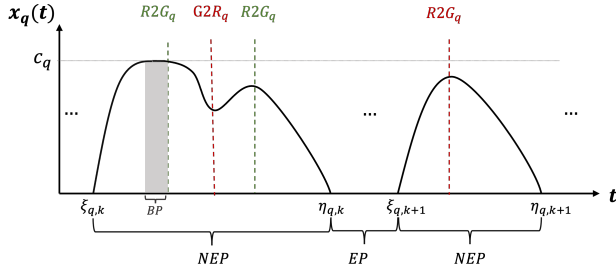


Fig. 5: Typical sample path of a traffic queue

where  $K_q$  is the (random) total number of NEPs during the sample path of queue  $q$  over  $[0, T]$ .

Thus, our goal is to determine  $\Theta$  that minimizes the expected weighted mean queue length:

$$J(\Theta; x(0), z(0), T) = E[L(\Theta; x(0), z(0), T)] \quad (15)$$

We note that it is not possible to derive a closed-form expression of  $J(\Theta; x(0), z(0), T)$  even if we had full knowledge of the processes  $\{\alpha_q(t)\}$  and  $\{\beta_q(t)\}$ . Therefore, a closed-form expression for the gradient  $\nabla J(\Theta)$  is also infeasible. The role of IPA is to obtain an *unbiased* estimate of  $\nabla J(\Theta)$  based on the sample function gradient  $\nabla L(\Theta)$  which can be evaluated based only on data directly observable along a single sample path such as Fig. 5, as will be shown in the next section. The unbiasedness of  $\nabla L(\Theta)$  is ensured under mild conditions on  $L(\Theta)$  (see [21]) and assuming that each  $\alpha_q(t)$  is piecewise continuously differentiable in  $t$  w.p. 1. In particular, we emphasize that no explicit knowledge of  $\alpha_q(t)$  is necessary to estimate  $\nabla J(\Theta)$  through  $\nabla L(\Theta)$ .

We can now invoke a gradient-based algorithm of the form

$$\theta_{p,i,l+1} = \theta_{p,i,l} - \rho_l \left[ \frac{dJ}{d\theta_{p,i,l}} \right]_{IPA} \quad (16)$$

where  $\theta_{p,i,l}$  is the  $i$ th parameter of  $\theta_p$  at the  $l$ th iteration ( $i \in \{1, 2, 3\}$ ),  $\rho_l$  is the stepsize at the  $l$ th iteration, and  $\left[ \frac{dJ}{d\theta_{p,i,l}} \right]_{IPA}$  is the IPA estimator of  $\frac{dJ}{d\theta_{p,i,l}}$ , which will be derived in the next section.

### III. INFINITESIMAL PERTURBATION ANALYSIS

We begin with a brief review of the IPA framework in [21]. Consider a sample path over  $[0, T]$  and denote the occurrence time of the  $k$ th event (of any type) by  $\tau_k$ . Let  $x'(\theta, t)$ ,  $\tau'_k(\theta)$  be the derivatives of  $x(\theta, t)$ ,  $\tau_k(\theta)$  over the scalar controllable parameter of interest  $\theta$  respectively. We omit the dependence on  $\theta$  for ease of notation hereafter. The dynamics of  $x(t)$  are fixed over any inter-event interval  $[\tau_k, \tau_{k+1})$ , represented by  $\dot{x}(t) = f_k(t)$ . Then, the state derivative satisfies

$$\frac{d}{dt} x'(t) = \frac{\partial f_k(t)}{\partial x} x'(t) + \frac{\partial f_k(t)}{\partial \theta} \quad (17)$$

with boundary condition (see [21]):

$$x'(\tau_k^+) = x'(\tau_k^-) + [f_{k-1}(\tau_k^-) - f_k(\tau_k^+)] \tau'_k \quad (18)$$

In order to evaluate (18),  $\tau'_k$  must be determined, which depends on the type of event taking place at  $\tau_k$ . For exogenous events (events causing a discrete state transition

that is independent of any controllable parameter), we have  $\tau'_k = 0$ . For endogenous events (events that occur when there exists a continuously differentiable function  $g_k$  such that  $\tau_k = \min\{t > \tau_{k-1} : g_k(x(\theta, t), \theta) = 0\}$  with guard condition  $g_k = 0$  (see [21]):

$$\tau'_k = -\left[ \frac{\partial g_k}{\partial x} f_k(\tau_k^-) \right]^{-1} \left( \frac{\partial g_k}{\partial \theta} + \frac{\partial g_k}{\partial x} x'(\tau_k^-) \right) \quad (19)$$

This framework captures how system states and event times change with respect to controllable parameters. Our goal is to estimate  $\nabla J(\Theta)$  through  $\nabla L(\Theta)$ , and, according to (14), the performance metric expression is a function of event time and system state variables. Thus, we apply the IPA framework to the TLC problem and evaluate how a perturbation in  $\theta$  would affect performance metrics.

#### A. State Derivatives.

We define the derivatives of the state variable  $x_q(t)$ ,  $z_q(t)$ ,  $y_{q,q^d,m}(t)$ ,  $r_{q,q^d,m}(t)$  and event times  $\tau_k$  with respect to parameter  $\Theta_i$  ( $i = 1, \dots, |\Theta|$ ) as follows:

$$\begin{aligned} x'_{q,i} &\equiv \frac{\partial x_q(t)}{\partial \Theta_i}, \quad z'_{q,i} \equiv \frac{\partial z_q(t)}{\partial \Theta_i}, \quad y'_{q,q^d,m,i} \equiv \frac{\partial y_{q,q^d,m}(t)}{\partial \Theta_i}, \\ r'_{q,q^d,m,i} &\equiv \frac{\partial r_{q,q^d,m}(t)}{\partial \Theta_i}, \quad \tau'_{k,i} \equiv \frac{\partial \tau_k}{\partial \Theta_i} \end{aligned} \quad (20)$$

For ease of notation, we denote the state dynamics in (1), (4), (11), (12) over an inter-event interval  $[\tau_k, \tau_{k+1})$  as follows:

$$\begin{aligned} \dot{x}_q(t) &= f_{q,k}^x(t), \quad \dot{z}_p(t) = f_{p,k}^z(t), \quad \dot{y}_{q,q^d,m}(t) = f_{q,q^d,m,k}^y(t), \\ \dot{r}_{q,q^d,m}(t) &= f_{q,q^d,m,k}^r(t), \quad q \in Q_n, \quad q^d \in Q_q^D, \quad n = 1, \dots, N \end{aligned} \quad (21)$$

Combining the dynamics in (1), (4), (11) and (12) with (17), similar to the analysis in [4] we can easily conclude that the state derivative of any queue is unaffected within any inter-event time interval, i.e., for  $t \in [\tau_k, \tau_{k+1})$ :

$$\begin{aligned} x'_{q,i}(t) &= x'_{q,i}(\tau_k^+), \quad z'_{q,i}(t) = z'_{q,i}(\tau_k^+), \\ y'_{q,q^d,m,i}(t) &= y'_{q,q^d,m,i}(\tau_k^+), \quad r'_{q,q^d,m,i}(t) = r'_{q,q^d,m,i}(\tau_k^+) \end{aligned} \quad (22)$$

Next, for any discrete event time  $\tau_k$ , we evaluate queue content derivatives for any possible event occurring to start/end an EP/NEP or within any EP/NEP, and for any controllable parameter  $\Theta_i$  ( $i = 1, \dots, |\Theta|$ ):

- 1) Event inside EP: Since  $x_q(t) = 0$  throughout the whole EP, it immediately follows that

$$x'_{q,i}(\tau_k^+) = 0 \quad (23)$$

- 2) Event starting EP ( $E_q$ ): This is induced by the basic event  $[x_q \downarrow 0]$ . The state dynamics change from  $f_{q,k-1}^x(\tau_k^-) = \alpha_q(\tau_k^-) - h_q(\tau_k^-)$  to  $f_{q,k}^x(\tau_k^+) = \alpha_q(\tau_k^+) - \alpha_q(\tau_k^+) = 0$ . Then, from (18),

$$x'_{q,i}(\tau_k^+) = x'_{q,i}(\tau_k^-) + (\alpha_q(\tau_k^-) - h_q(\tau_k^-)) \tau'_{k,i} \quad (24)$$

- 3) Event starting NEP ( $S_q$ ) from EP: This can be induced in three possible ways:

- 3.1)  $S_q$  induced by light switching to RED ( $G2R_q$ ) when  $\alpha_q(\tau_k) > 0$ . We have  $f_{q,k-1}^x(\tau_k^-) = 0$  and  $f_{q,k}^x(\tau_k^+) = \alpha_q(\tau_k^+)$ . Based on (23) and (18):

$$x'_{q,i}(\tau_k^+) = -\alpha_q(\tau_k^+) \tau'_{k,i} \quad (25)$$

- 3.2)  $S_q$  induced by  $J_{q^u,q,m}$  where  $q^u \in Q_q^U$ . The state dynamics are  $f_{q,k-1}^x(\tau_k^-) = 0$ ,  $f_{q,k}^x(\tau_k^+) = \alpha_q(\tau_k^+) - \beta_q(\tau_k^+)$ , where  $\beta_q(\tau_k)$  is defined in (2) and  $\alpha_q(\tau_k^+) - \beta_q(\tau_k^+) > 0$  in order to induce  $S_q$ . Based on (18) we get

$$x'_{q,i}(\tau_k^+) = (\beta_q(\tau_k^+) - \alpha_q(\tau_k^+)) \tau'_{k,i} \quad (26)$$

- 3.3)  $S_q$  induced by an exogenous change in  $\alpha_q(\tau_k)$  when  $Q_q^U = \emptyset$ . In this case,  $\tau'_{k,i} = 0$ , so that we have

$$x'_{q,i}(\tau_k^+) = x'_{q,i}(\tau_k^-) = 0 \quad (27)$$

- 4) Event inside NEP. The following are all possible cases:

- 4.1)  $G2R_q$ : the state dynamics are  $f_{q,k-1}^x(\tau_k^-) = \alpha_q(\tau_k^-) - h_q(\tau_k^-)$  and  $f_{q,k}^x(\tau_k^+) = \alpha_q(\tau_k^+)$ . Note  $\alpha_q(t)$  is continuous except when joining events happen at  $q$ . Therefore,

$$x'_{q,i}(\tau_k^+) = x'_{q,i}(\tau_k^-) - h_q(\tau_k^-) \tau'_{k,i} \quad (28)$$

- 4.2)  $R2G_q$ : the state dynamics are  $f_{q,k-1}^x(\tau_k^-) = \alpha_q(\tau_k^-)$  and  $f_{q,k}^x(\tau_k^+) = \alpha_q(\tau_k^+) - h_q(\tau_k^+)$ . Therefore,

$$x'_{q,i}(\tau_k^+) = x'_{q,i}(\tau_k^-) + h_q(\tau_k^+) \tau'_{k,i} \quad (29)$$

- 4.3)  $J_{q^u,q,m}$  with  $q^u \in Q_q^U$ : the state dynamics are:  $f_{q,k-1}^x(\tau_k^-) = -\beta_q(\tau_k^-)$ ,  $f_{q,k}^x(\tau_k^+) = \alpha_q(\tau_k^+) - \beta_q(\tau_k^+)$  where  $\beta_q(\tau_k)$  follows (2), and it is continuous except when a light switching event happens at  $q$ . Thus,

$$x'_{q,i}(\tau_k^+) = x'_{q,i}(\tau_k^-) - \alpha_q(\tau_k^+) \tau'_{k,i} \quad (30)$$

- 4.4)  $J_{q^u,q,m}^e$  with  $q^u \in Q_q^U$ : similarly, the state dynamics are:  $f_{q,k-1}^x(\tau_k^-) = \alpha_q(\tau_k^-) - \beta_q(\tau_k^-)$ ,  $f_{q,k}^x(\tau_k^+) = -\beta_q(\tau_k^+)$ . Therefore,

$$x'_{q,i}(\tau_k^+) = x'_{q,i}(\tau_k^-) + \alpha_q(\tau_k^-) \tau'_{k,i} \quad (31)$$

- 4.5) Other exogenous events. Those events would not affect state derivatives, so that:

$$x'_{q,i}(\tau_k^+) = x'_{q,i}(\tau_k^-) \quad (32)$$

- 5) Event starting BP. Blocking is triggered by  $[x_q \uparrow c_q]$ . When queue  $q$  starts to block, in addition to the current state  $x_q(t)$ , states from all upstream queues  $q^u \in Q_q^U$  are also affected. For queue  $q$ , the state dynamics change from  $f_{q,k-1}^x(\tau_k^-) = \alpha_q(\tau_k^-) - \beta_q(\tau_k^-)$  to  $f_{q,k}^x(\tau_k^+) = 0$ . Therefore,

$$x'_{q,i}(\tau_k^+) = x'_{q,i}(\tau_k^-) + (\alpha_q(\tau_k^-) - \beta_q(\tau_k^-)) \tau'_{k,i} \quad (33)$$

For upstream queue  $q^u \in Q_q^U$ , the state dynamics change from  $f_{q^u,k-1}^x(\tau_k^-) = \alpha_{q^u}(\tau_k^-) - \beta_{q^u}(\tau_k^-)$  to  $f_{q^u,k}^x(\tau_k^+) = \alpha_{q^u}(\tau_k^+)$ . Therefore,

$$x'_{q^u,i}(\tau_k^+) = x'_{q^u,i}(\tau_k^-) - \beta_{q^u}(\tau_k^-) \tau'_{k,i} \quad (34)$$

- 6) Event inside BP. Since the state dynamics do not change during blocking, the state derivatives do not change:

$$x'_{q,i}(\tau_k^+) = x'_{q,i}(\tau_k^-) \quad (35)$$

- 7) Event starting NEP from BP. This is triggered by  $[x_q \downarrow c_q]$ , and, similar to case 5, it affects the state of all upstream queues. For queue  $q$ , the state dynamics are:  $f_{q,k-1}^x(\tau_k^-) = 0$ ,  $f_{q,k}^x(\tau_k^+) = \alpha_q(\tau_k^+) - \beta_q(\tau_k^+)$ . Therefore:

$$x'_{q,i}(\tau_k^+) = x'_{q,i}(\tau_k^-) + (\beta_q(\tau_k^+) - \alpha_q(\tau_k^+)) \tau'_{k,i} \quad (36)$$

For upstream queue  $q^u \in Q_q^U$ , the state dynamics are:  $f_{q^u,k-1}^x(\tau_k^-) = \alpha_{q^u}(\tau_k^-)$  to  $f_{q^u,k}^x(\tau_k^+) = \alpha_{q^u}(\tau_k^+) - \beta_{q^u}(\tau_k^+)$ . Therefore,

$$x'_{q^u,i}(\tau_k^+) = x'_{q^u,i}(\tau_k^-) + \beta_{q^u}(\tau_k^+) \tau'_{k,i} \quad (37)$$

Observe that whenever  $Q_q^U \neq \emptyset$ , the value of  $\alpha_q$  is given by (10), thus capturing the interdependence of queue content derivatives between adjacent queues  $q^u, q$ . Also note that most of the queue content derivative expressions involve the event time derivative  $\tau'_{k,i}$ . Therefore, to complete our analysis we need to derive these expressions by applying (19) as shown next and use them in (24), (25), (26), (28), (29), (30), (31), (33), (34), (36) and (37).

## B. Event Time Derivatives

In this section, we derive the event time derivatives with respect to each of the controllable parameters  $\Theta_i$ ,  $i = 1, \dots, |\Theta|$  as formulated in (5). For queue  $q \in Q_n$  within intersection  $n = 1, \dots, N$ :

- 1) Event  $E_q$  occurs at  $\tau_k$ . This is induced by  $[x_q \downarrow 0]$  so that the guard condition is  $g_k = x_q - 0 = 0$ , which gives  $\frac{\partial g_k}{\partial x_q} = 1$ ,  $\frac{\partial g_k}{\partial \Theta_i} = 0$ . The dynamics are  $f_{q,k-1}^x(\tau_k^-) = \alpha_q(\tau_k^-) - h_q(\tau_k^-)$ ,  $f_{q,k}^x(\tau_k^+) = 0$ . Then, it follows from (19) that:

$$\tau'_{k,i} = \frac{-x_{q,i}'(\tau_k^-)}{\alpha_q(\tau_k^-) - h_q(\tau_k^-)} \quad (38)$$

This completes equation (24) and yields  $x_{q,i}'(\tau_k^+) = 0$  for the state derivative in case 2.

- 2) Event  $G2R_q$  occurs at  $\tau_k$ . This happens when there is a phase switch and queue  $q$  is part of the previous enabled phase  $p$  ( $q \in Q_p$ ) but not part of the next enabled phase  $p'$  ( $q \notin Q_{p'}$ ). Such a phase switch can be triggered by the following basic events:

- 2.1)  $[z_p \uparrow \theta_p^{max}]$  while  $x_p^{max}(\tau_k) > 0$ . The guard condition is  $g_k = z_p - \theta_p^{max}$  with  $\frac{\partial g_k}{\partial z_p} = 1$  and  $\frac{\partial g_k}{\partial \theta_p^{max}} = -1$ , and all other partial derivatives equals to zero. We also have  $f_{p,k}^z(\tau_k^-) = 1$ . Similar to the analysis in [4] we get:

$$\tau'_{k,i} = \tau'_{k,s,i} + \mathbb{1}_{\Theta_i = \theta_p^{max}} \quad (39)$$

where  $k_s$  is the index of the last light switching event ( $R2G_q$ ) prior to event  $k$  and  $\mathbb{1}$  is the indicator function. This completes equations (25) and (28).

2.2) [ $z_p \uparrow \theta_p^{min}$ ] while  $x_p^{max}(\tau_k) < s_p$  and  $x_p^{max}(t) > s_p$ . The guard condition is  $g_k = z_p - \theta_p^{min}$ . Similar to the last case, (19) can be reduced to:

$$\tau'_{k,i} = \tau_{k_s,i}' + \mathbb{1}_{\Theta_i = \theta_p^{min}} \quad (40)$$

2.3) [ $x_p^{max} \downarrow s_p$ ] while  $z_p \geq \theta_p^{min}$  and  $x_p^{max}(\tau_k) \geq s_p$ . The guard condition is  $g_k = x_{q^*}(\tau_k) - s_p$  with  $\frac{\partial g_k}{\partial x_{q^*}} = 1$  and  $\frac{\partial g_k}{\partial s_p} = -1$ , and all other partial derivatives equal to zero, where  $q^* = \operatorname{argmax}_{q \in Q_p} x_q(\tau_k)$ . Due to light switching, the state dynamics change from  $f_{q,k-1}^x(\tau_k^-) = \alpha_q(\tau_k^-) - \beta_q(\tau_k^-)$  to  $f_{q,k}^x(\tau_k^+) = \alpha_q(\tau_k^+)$ . As a result, (19) reduces to:

$$\tau'_{k,i} = \frac{\mathbb{1}_{\Theta_i = s_p} - x'_{q^*,i}(\tau_k^-)}{\alpha_{q^*}(\tau_k^-) - \beta_{q^*}(\tau_k^-)} \quad (41)$$

2.4) [ $x_p^{max} \uparrow s_p$ ] while  $z_p \geq \theta_p^{min}$  and  $x_p^{max}(\tau_k) < s_p$ . The guard condition is  $g_k = x_{\bar{q}^*}(\tau_k) - s_p$  with  $\frac{\partial g_k}{\partial x_{\bar{q}^*}} = 1$  and  $\frac{\partial g_k}{\partial s_p} = -1$ , and all other partial derivatives equal to zero, where  $\bar{q}^* = \operatorname{argmax}_{\bar{q} \in Q_n \setminus Q_p} x_{\bar{q}}(\tau_k)$ . Similar to the previous case:

$$\tau'_{k,i} = \frac{\mathbb{1}_{\Theta_i = s_p} - x'_{\bar{q}^*,i}(\tau_k^-)}{\alpha_{\bar{q}^*}(\tau_k^-)} \quad (42)$$

When it comes to event  $R2G_q$ , note that this is triggered under similar conditions to those of  $G2R_q$ . In particular,  $R2G_q$  occurs at phase switching times when queue  $q$  is not part of the previous enabled phase  $p$  ( $q \notin Q_p$ ) but part of the next enabled phase  $p'$  ( $q \in Q_{p'}$ ). Thus, we omit its analysis here. The combined four cases of these events together complete the state derivative calculation of (25), (29), and (28).

3) Event  $J_{q^u,q,m}$  occurs at  $\tau_k$ , where  $q^u \in Q_q^U$ . This is an endogenous event triggered by  $y_{q^u,q,m}(\tau_k) = \Delta_{q^u,q}(\tau_k)$ , so that the associated guard condition is  $g_k = y_{q^u,q,m}(\tau_k) - \frac{L_{q-q^u}(\tau_k) * l}{f_{q^u}} = 0$ . Since there are two state variables present in this guard condition, we take derivatives of  $g_k$  with respect to parameter  $\Theta_i$  first. Since the event time variable  $\tau_k$  is directly affected by  $\Theta_i$ , while state variables  $x_q(\tau_k)$  and  $y_{q^u,q,m}(\tau_k)$  are both directly and indirectly affected by  $\Theta_i$  through  $\tau_k$ , using the chain rule yields the derivative  $g'_{k,i}$  as follows:

$$g'_{k,i} = y'_{q^u,q,m,i}(\tau_k^-) + f_{q^u,u,m,k-1}^y(\tau_k^-) \tau'_{k,i} + \frac{l}{f_{q^u}} (x'_{q,i}(\tau_k^-) + f_{q,k-1}^x(\tau_k^-) \tau'_{k,i}) = 0 \quad (43)$$

The last event before  $\tau_k$  that would cause the change of state  $y_{q^u,q,m}$  is  $G_{q^u,q,m}$  (start of flow burst generation). Denoting its occurrence time by  $\tau_{k_g}$ , we have  $f_{q^u,q,m,k_g-1}^y(\tau_{k_g}^-) = 0$  and  $f_{q^u,q,m,k_g}^y(\tau_{k_g}^+) = 1$ . Also, since  $y_{q^u,q,m}(t) = 0$  right before  $\tau_{k_g}$ , we have  $y'_{q^u,q,m,i}(\tau_{k_g}^-) = 0$  for all  $i = 1, \dots, |\Theta|$ . Therefore, by applying (18) and (22), we get:

$$\begin{aligned} y'_{q^u,q,m,i}(\tau_k^-) &= y'_{q^u,q,m,i}(\tau_{k_g}^+) \\ &= y'_{q^u,q,m,i}(\tau_{k_g}^-) + (0 - 1) \tau'_{k_g,i} \\ &= -\tau'_{k_g,i} \end{aligned} \quad (44)$$

The state dynamics can be calculated from (11) and (1) with  $\alpha_q(\tau_k^-) = 0$ :

$$f_{q^u,q,m,k-1}^y(\tau_k^-) = 1 \quad (45)$$

$$f_{q,k-1}^x(\tau_k^-) = -\beta_q(\tau_k^-) \quad (46)$$

Then, making use of (44), (45) and (46) in (43) finally yields:

$$\tau'_{k,i} = \frac{f_{q^u}}{f_{q^u} - l\beta_q(\tau_k^-)} (\tau'_{k_g,i} - \frac{l}{f_{q^u}} x_{q,i}'(\tau_k^-)) \quad (47)$$

where  $\tau'_{k_g,i}$  is the event time derivative when  $G_{q^u,q,m}$  occurs with  $[\beta_{q^u} \uparrow 0]$ . This can be induced by: (a)  $R2G_{q^u}$  when  $x_{q^u}(\tau_{k_g,i}) > 0$ , in which case  $\tau'_{k_g,i}$  follows (39), (40), (41) or (42); (b)  $J_{q^{uu},q^u,\hat{m}}$  when  $x_{q^u}(\tau_{k_g,i}) = 0$  and  $u_{q^{uu}} = 1$ , where  $q^{uu}$  is an upstream queue of  $q^u$  ( $q^{uu} \in Q_{q^u}^U$ ), and  $\hat{m}$  is the GREEN cycle index of  $q^{uu}$  at  $\tau_{k_g}$ . In this case,  $\tau'_{k_g,i}$  follows (47); (c) an exogenous change of  $\alpha_{q^u}$  where  $\tau'_{k_g,i} = 0$ . This completes the state derivatives equations (26) and (30).

4) Event  $J_{q^u,q,m}^e$  occurs at  $\tau_k$  where  $q^u \in Q_q^U$ . As in the last case, this is an endogenous event triggered by  $r_{q^u,q,m}(\tau_k) = \Delta_{q^u,q}(\tau_k)$ , with guard condition  $g_k = r_{q^u,q,m}(\tau_k) - \frac{L_{q-q^u}(\tau_k) * l}{f_{q^u}} = 0$ . Similar to the analysis in the previous case, the derivative of the guard condition with respect to  $\Theta_i$  is:

$$g'_{k,i} = r'_{q^u,q,m,i}(\tau_k^-) + f_{q^u,u,m,k-1}^r(\tau_k^-) \tau'_{k,i} + \frac{l}{f_{q^u}} (x'_{q,i}(\tau_k^-) + f_{q,k-1}^x(\tau_k^-) \tau'_{k,i}) = 0 \quad (48)$$

The last event before  $\tau_k$  that would cause the change of state  $r_{q^u,q,m}$  is  $G_{q^u,q,m}^e$  (end of flow burst generation). Denoting its occurrence time by  $\tau_{k_e}$ , we have  $f_{q^u,q,m,k_e-1}^r(\tau_{k_e}^-) = 0$  and  $f_{q^u,q,m,k_e}^r(\tau_{k_e}^+) = 1$ . Therefore,

$$\begin{aligned} r'_{q^u,q,m,i}(\tau_k^-) &= r'_{q^u,q,m,i}(\tau_{k_e}^+) \\ &= r'_{q^u,q,m,i}(\tau_{k_e}^-) + (0 - 1) \tau'_{k_e,i} \\ &= -\tau'_{k_e,i} \end{aligned} \quad (49)$$

Also by (12) and (1):

$$f_{q^u,q,m,k-1}^r(\tau_k^-) = 1 \quad (50)$$

$$f_{n,d,k-1}^x(\tau_k^-) = \alpha_q(\tau_k^-) - \beta_q(\tau_k^-) \quad (51)$$

where the value of  $\beta_q(\tau_k^-)$  follows (2). Then, using (49), (50) and (51) in (48) we finally obtain:

$$\begin{aligned} \tau'_{k,i} &= \frac{f_{q^u}}{f_{q^u} + (\alpha_q(\tau_k^-) - \beta_q(\tau_k^-))l} * \\ &\quad (\tau'_{k_e,i} - \frac{l}{f_{q^u}} x_{q,i}'(\tau_k^-)) \end{aligned} \quad (52)$$

where  $\tau'_{k_e,i}$  is the event time derivative when  $G_{q^u,q,m}^e$  occurs with  $[\beta_{q^u} \downarrow 0]$ . This can be induced by: (a)  $G2R_{q^u}$  when  $\alpha_{q^u}(\tau_{k_e}) > 0$  or  $x_{q^u}(\tau_{k_e}) > 0$ , in which case  $\tau'_{k_e,i}$  follows (39), (40), (41) or (42); (b)  $J_{q^{uu},q^u,\hat{m}}$  where  $q^{uu}$  is an upstream queue of  $q^u$  ( $q^{uu} \in Q_{q^u}^U$ ),

and  $\hat{m}$  is the GREEN cycle index of  $q^{uu}$  at  $\tau_{k,q}$ . In this case,  $\tau'_{k,e,i}$  follows (52); (c) an exogenous change of  $\alpha_{q^u}$  where  $\tau'_{k,e,i} = 0$ . This completes the state derivative equation (31).

- 5) Event  $[x_q \uparrow c_q]$  occurs at  $\tau_k$  and induces blocking for queue  $q$ . The guard condition is  $g_k = x_q - c_q = 0$ . Similar to the calculation in case 1:

$$\tau'_{k,i} = \frac{-x_{q,i}'(\tau_k^-)}{\alpha_q(\tau_k^-) - \beta_q(\tau_k^-)} \quad (53)$$

This completes equations (33) and (34).

- 6) Event  $[x_q \downarrow c_q]$  occurs at  $\tau_k$  and releases blocking for queue  $q$ . Note that we cannot directly calculate  $\tau_k$  through (19) since  $f_{q,k-1}^x(\tau_k^-) = 0$ . However,  $[x_q \downarrow c_q]$  can only be induced by: (a)  $R2G_q$ , in which case  $\tau_k$  follows (39), (40), (41) or (42); (b)  $[\alpha_q \downarrow 0]$ , in which case  $\tau_k = 0$  if  $Q_q^U = \emptyset$  or  $\tau_k$  follows (52) otherwise. This completes equations (36) and (37).

In summary, if  $G2R_q$  occurs at time  $\tau_k$ :

$$\tau'_{k,i} = \begin{cases} \tau_{k,s,i}' + \mathbb{1}_{\Theta_i = \theta_p^{max}}, & \text{if induced by } [z_p \uparrow \theta_p^{max}] \\ \tau_{k,s,i}' + \mathbb{1}_{\Theta_i = \theta_p^{min}}, & \text{if induced by } [z_p \uparrow \theta_p^{min}] \\ \frac{\mathbb{1}_{\Theta_i = s_p} - x_{q^*,i}'(\tau_k^-)}{\alpha_{q^*}(\tau_k^-) - \beta_{q^*}(\tau_k^-)}, & \text{if induced by } [x_p^{max} \downarrow s_p] \\ \frac{\mathbb{1}_{\Theta_i = s_p} - x_{\bar{q}^*,i}'(\tau_k^-)}{\alpha_{\bar{q}^*}(\tau_k^-)}, & \text{if induced by } [x_p^{max} \uparrow s_p] \end{cases} \quad (54)$$

where  $q^* = \operatorname{argmax}_{q \in Q_p} x_q(\tau_k)$ , and  $\bar{q}^* = \operatorname{argmax}_{\bar{q} \in Q_n \setminus Q_p} x_{\bar{q}}(\tau_k)$  as defined in Section III-B. Otherwise:

$$\tau'_{k,i} = \begin{cases} \frac{-x_{q,i}'(\tau_k^-)}{\alpha_q(\tau_k^-) - h_q(\tau_k^-)}, & \text{if } E_q \text{ occurs at } \tau_k \\ \frac{f_{q^u}^*}{f_{q^u} - l\beta_q(\tau_k^-)} \left( \tau'_{k,g,i} - \frac{l}{f_{q^u}} x_{q,i}'(\tau_k^-) \right), & \text{if } J_{q^u,q,m} \text{ occurs at } \tau_k \\ \frac{f_{q^u} + (\alpha_q(\tau_k^-) - \beta_q(\tau_k^-))l}{f_{q^u}} \left( \tau'_{k,g,i} - \frac{l}{f_{q^u}} x_{q,i}'(\tau_k^-) \right), & \text{if } J_{q^u,q,m}^e \text{ occurs at } \tau_k \\ \frac{-x_{q,i}'(\tau_k^-)}{\alpha_q(\tau_k^-) - \beta_q(\tau_k^-)}, & \text{if } [x_q \uparrow c_q] \text{ occurs at } \tau_k \end{cases} \quad (55)$$

Note that the event time derivative for  $R2G_q$ ,  $G_{q^u,q,m}$ ,  $G_{q^u,q,m}^e$  and  $[x_q \downarrow c_q]$  can be derived from the above equation based on different cases with all events as defined in Section II.

### C. Cost Derivatives

With all state derivatives obtained as shown in the previous section, we can obtain the IPA cost gradient estimator as the derivative of  $L(\Theta; x(0), z(0), T)$  in (14). Similar to the derivation provided in [4], the IPA estimator consisting of  $dL(\Theta)/d\Theta_i$ ,  $i = 1, \dots, |\Theta|$  is given by

$$\frac{dL(\Theta)}{d\Theta_i} = \frac{1}{T} \sum_{n=1}^N \sum_{q \in Q_n} \sum_{k=1}^{K_q} \omega_q \frac{dL_{q,k}(\Theta)}{d\Theta_i} \quad (56)$$

where

$$L_{q,k}(\Theta) = \int_{\xi_{q,k}}^{\eta_{q,k}} x_q(\Theta, t) dt \quad (57)$$

$$\begin{aligned} \frac{dL_{q,k}(\Theta)}{d\Theta_i} &= x'_{q,i}(\xi_{q,k}^+)(t_{q,k}^1 - \xi_{q,k}) \\ &\quad + x'_{q,i}(t_{q,k}^{O_{q,k}^+})(\eta_{q,k} - t_{q,k}^{O_{q,k}}) \\ &\quad + \sum_{o=2}^{O_{q,k}} x'_{q,i}(t_{q,k}^{O_{q,k}^{o-1}+})(t_{q,k}^o - t_{q,k}^{O_{q,k}^{o-1}}) \end{aligned} \quad (58)$$

where  $O_{q,k}$  is the total number of events on the observed sample path at queue  $q$  within the  $k$ th NEP,  $t_{q,k}^o$  is the observed time of the  $o$ th event in that NEP, and  $\xi_{q,k}$ ,  $\eta_{q,k}$  are the observed occurrence times of the start and end of the  $k$ th NEP respectively.

It is clear from (56) and (58) that each IPA derivative is basically the accumulation of measurable inter-event times (timers) multiplied by a corresponding state derivative. The full information needed to evaluate this IPA estimator consists of (a) event time data, which are easy to record by directly observing events, and (b) state derivatives at those times, which are given by the simple iterative expressions derived in Sections III-A and III-B. Some state derivatives evaluated at an event time  $\tau_k$  involve flow rates  $\alpha_q(\tau_k)$ ,  $h_q(\tau_k)$ ; however, these are *only needed at specific events* (e.g., at the time when  $S_q$  is induced by a light switching event in (25)), making them easy to estimate in practice, as described in the next section. Regarding the unconstrained departure rate  $h_q(t)$ , considering it as a constant also makes it easy to estimate through simple offline counting methods.

Returning to our observation in (10) that the  $\gamma_{q,qd}$  information is not needed, observe that this is indeed not involved in the derivation the IPA derivatives.

Finally, using simple online gradient-based algorithms as in (16), we can adjust the controllable parameters to improve the overall performance and attain possibly local optima if the operating conditions do not change substantially; see also Section IV.

### D. Perturbation Propagation

The IPA method is able to adjust the controllable parameters automatically in an online data-driven manner. When applying it to a multi-intersection traffic system, the adjustment of the controllable parameters can synchronize the traffic lights and naturally induce a “green wave” effect by propagating state perturbations to downstream intersections. Such propagation is induced by the flow burst joining events and affects downstream state derivatives, thus capturing how a perturbation in a TLC parameter  $\Theta_i$  causes a perturbation in the queue content  $x_q(t)$ . For example, if an  $S_q$  event occurs induced by a  $J_{q^u,q,m}$  event (where  $q^u \in Q_q^U$ ) when  $v_q = 0$  (the light is RED), then by combining (26) and (55), the state derivative becomes

$$\begin{aligned} x'_{q,i}(\tau_k^+) &= (\beta_q(\tau_k^+) - \alpha_q(\tau_k^+)) * \\ &\quad \left( \tau'_{k,g,i} - \frac{l}{f_{q^u}} x'_{q,i}(\tau_k^-) \right) \frac{f_{q^u}}{f_{q^u} - l\beta_q(\tau_k)} \end{aligned} \quad (59)$$

Since  $\beta_q(\tau_k^+) = 0$  when  $v_q = 0$  and  $x'_{q,i}(\tau_k^-) = 0$ , this is reduced to  $x'_{q,i}(\tau_k^+) = -\alpha_q(\tau_k)\tau'_{kq,i}$ , where the flow rate  $\alpha_q(\tau_k)$  can be traced to the upstream departure rate based on (10) and the event time derivative at  $G_{q^u,q,m}$  also depends on upstream information: (a) if this is triggered by  $R2G_{q^u}$ , then  $\tau'_{kq,i}$  is an accumulation of light switching counts at the upstream intersection, whereas (b) if it is triggered by  $J_{q^{uu},q^u,\hat{m}}$  this again depends on the event time derivative at  $G_{q^{uu},q^u,\hat{m}}$ , where  $q^{uu} \in Q_{q^u}^u$ . Thus, recursively,  $x'_{q,i}(\tau_k^+)$  is affected by all upstream states capturing how a perturbation can be propagated. A similar process takes place under other flow burst joining conditions.

It is also worth observing that combining (24) and (55) gives  $x'_{q,i}(\tau_k^+) = 0$  when event  $E_q$  occurs at  $\tau_k$ . This implies that any parameter perturbation affecting  $x_q(t)$  is reset to zero after such events at queue  $q$ . Therefore, for any perturbation at  $q^u$  to propagate beyond  $q$  requires that after a  $J_{q^u,q,m}$  event occurs it must be followed by a  $E_q$  event before a  $G2R_q$  event occurs so that the vehicles from  $q^u$  benefiting from a positive GREEN cycle perturbation have the chance to get through  $q$  without stopping; otherwise, the perturbation at  $q^u$  is “cancelled” by  $E_q$ . Note that such propagation only occurs through  $J_{q^u,q,m}$  events, which limits computation to simple derivative updates at selected events and makes this propagation analysis scalable. This also allows us to easily track how TLC parameters affect “green waves” along a series of intersections.

#### IV. SIMULATION RESULTS

We use the Eclipse SUMO (Simulation of Urban MOBility) simulator to build a simulation environment for traffic through a  $m * n$  traffic light systems. We initially consider a  $2 * 3$  grid shape, so that  $N = 6$ . The length of each road segment between two neighboring intersections is set to be  $300m$ . For each individual intersection, we use the traffic structure shown in Fig. 2: each road segment contains two queues - one for left-turn movements and one for combined straight and right-turn movements. Four phases are designed as described in Sec.II: (a) [W-s],[E-s]; (b)[W-l],[E-l];(c)[S-s],[N-s]; (d)[S-l],[N-l]. Although IPA is independent of the arrival processes, we select Poisson arrival processes in SUMO to capture the traffic demand for each origin-destination (OD) pair, where the origins and destinations are different and correspond to boundary edges of the traffic network. To simplify traffic demand representation, OD pairs are grouped into four categories: row-to-row, row-to-column, column-to-row, and column-to-column. Let  $\bar{\alpha} = [\bar{\alpha}_{rr}, \bar{\alpha}_{rc}, \bar{\alpha}_{cr}, \bar{\alpha}_{cc}]$  denote the Poisson arrival rates for each OD pair group, where subscripts indicate directionality. For example,  $\bar{\alpha}_{rr}$  represents Poisson rate for OD pairs with horizontal departure and arrival edges. This abstraction allows the aggregated modeling of demand patterns between boundary edges. Since we only need the real-time arrival flow rate at certain event times, we can estimate an instantaneous arrival rate through  $\alpha_q(\tau_k) = N_q/t_w$ , where  $N_q$  denotes the number of vehicles joining queue  $q$  during a time window of size  $t_w$  before event time  $\tau_k$ ; this is easy to detect and record

in SUMO. We also estimate the maximum departure rate as a constant value  $h_q(t) = H$  for all queues (where  $H$  is determined through a separate offline analysis). We set  $H = 1.3 \text{ vehicles/s}$ ,  $v = 10 \text{ m/s}$ , and equal weights for all flows ( $w_q = 1$ ) throughout this section. With this setting, we have performed simulation experiments to demonstrate improvements in mean waiting times, synchronization leading to “green waves”, adaptivity, and scalability.

##### A. Online TLC Implementation for Uncongested Traffic

The ultimate goal of TLC is to operate on line, i.e., observe real-time traffic data and adaptively adjust the controllable parameters. We simulate this process by creating a single long sample path and updating IPA derivatives in (58) with every observed event occurrence (assuming some sensing capabilities for detecting events and event times). The results are accumulated and the parameters are updated every 1000s using the data collected during the most recent time window. Two metrics are used to evaluate the performance of the TLC method: (a) Mean waiting time, which measures the time of stopping at a queue, providing an indication of delay experienced behind traffic signals; (b) Time-distance ratio, which calculates the average time taken to travel a fixed distance. This incorporates both stopping time and travel speed. A higher ratio indicates more undesirable traffic conditions associated with congestion. We set the initial controllable parameters for each phase of each traffic light as  $\theta_p = [20, 40, 10]$ ,  $p \in P_n, n = 1, \dots, 6$ , indicating minimum GREEN time threshold, maximum GREEN time threshold, and queue content threshold respectively.

We start with different relatively small  $\bar{\alpha}$  values that would not lead to congested traffic, and record performances as shown in Table II. Observe that under different unbalanced traffic demands, both performance metrics show improvements from 0.75% to 46%. Typical sample cost trajectories for different flows under  $\bar{\alpha} = [0.02, 0.01, 0.02, 0.01]$  are shown in Fig. 6. Observe that a left-turn queue always has a higher waiting time than a straight & right-turn queue, while the gap becomes smaller as both values decrease.

We also compare the performance of our method against Webster’s Method [22], which is a rational approach for designing fixed-cycle traffic signals based on Webster’s formula. Webster’s Method incorporates static information such as flow ratios and lost time to optimize predetermined cyclic plans rather than adapting in real-time. As shown in Fig. 7, our adaptive TLC method initially has higher waiting times before optimizing the signal timings. However, it is then able to self-adjust the parameters, while the fixed cycle timing from Webster’s method is prone to slight demand fluctuations, leading to consistently higher overall wait times. It should be noted that under uncongested traffic conditions, fixed-cycle policies, like Webster’s method, can deliver relatively stable and acceptable performance, despite some fluctuations. However, such predetermined signal plans becomes ineffective in scenarios where congestion emerges, as demonstrated next.

TABLE II: Simulation results for different traffic demands

| $\bar{\alpha}$        | Average Waiting Time |       |                 | Time-Distance Ratio |       |                 |
|-----------------------|----------------------|-------|-----------------|---------------------|-------|-----------------|
|                       | init                 | opt   | reduction ratio | init                | opt   | reduction ratio |
| [0.02,0.01,0.01,0.01] | 12.03                | 6.42  | 46.63%          | 0.122               | 0.115 | 5.74%           |
| [0.02,0.02,0.01,0.01] | 19.06                | 10.81 | 43.28%          | 0.141               | 0.131 | 7.09%           |
| [0.02,0.01,0.02,0.01] | 16.20                | 10.71 | 42.96%          | 0.133               | 0.132 | 0.75%           |
| [0.02,0.01,0.01,0.02] | 23.17                | 12.37 | 46.61%          | 0.152               | 0.135 | 11.18%          |

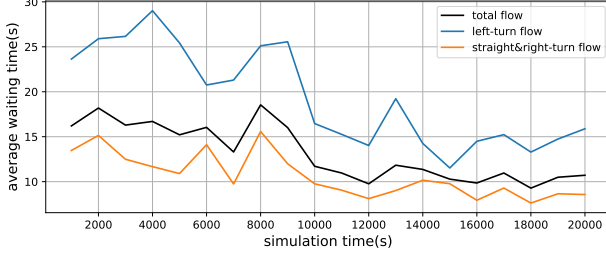
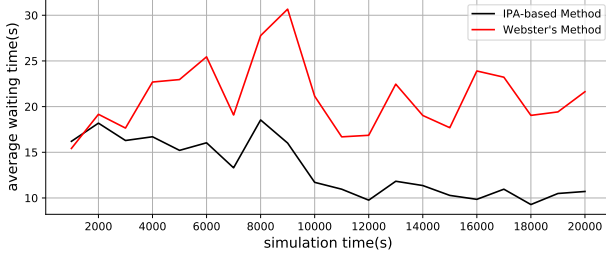

 Fig. 6: Sample cost trajectory for online implementation under uncongested traffic with  $\bar{\alpha} = [0.02, 0.01, 0.02, 0.01]$ 


Fig. 7: Sample cost trajectories comparison with Webster's Method under uncongested traffic

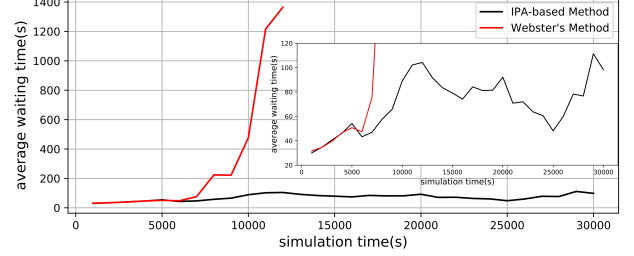


Fig. 8: Sample cost trajectories comparison with Webster's Method under potential congested traffic

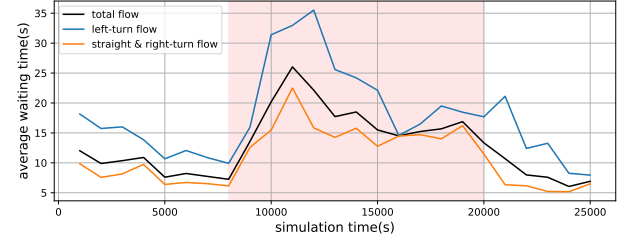


Fig. 9: Sample cost trajectories with traffic demand perturbation

### B. Online TLC Implementation for Congested Traffic

We next consider scenarios where traffic demand is high enough that improper signal timing can lead to congestion and blocking. In such cases, our adaptive TLC method can detect emerging blockages and adjust the signal parameters accordingly. This allows responding to congestion in real time, which can be challenging for static traffic light schemes like Webster's Method that rely on flow ratios to determine signal cycles.

We use the same simulation setting as the uncongested case, with traffic demand set as  $\bar{\alpha} = [0.02, 0.02, 0.02, 0.011]$ . A typical comparison of sample trajectories is shown in Fig. 8. When traffic congestion builds up early in the simulation, our method can adapt parameters based on observed blocking to help mitigate gridlock. In contrast, the average waiting time with Webster's method increases rapidly and quickly leads to unstable conditions.

### C. TLC Adaptivity

Our TLC method is designed to adapt to changing traffic conditions. We illustrate this property by observing how performance changes when traffic demand is perturbed. The initial traffic demand is set through  $\bar{\alpha} =$

$[0.01, 0.01, 0.01, 0.015]$ , with the same initial parameters as in previous cases. We add traffic perturbations by doubling the Poisson rate of the flow for row-to-column OD pairs at  $t = 8000s$  and then return to the original rate at  $t = 20000s$ . The cost trajectory is shown in Fig. 9 where the shaded area corresponds to the time interval over which traffic demand was increased. We can see that the waiting time initially decreases due to proper parameter adjustments. When traffic demand abruptly increases, the waiting time increases as expected, since the previously optimized parameters no longer apply to the new traffic demand. Nonetheless, they gradually adjust and converge to new optimal values after several iterations, including the time interval after the traffic demand is returned to its original value.

### D. TLC Scalability

We have seen in Section III that the IPA gradient estimation process is entirely event-driven. Therefore, the computational complexity of the TLC method is linear in the number of events. This implies that our approach scales with the number of traffic lights in a network of interconnected intersections, since the presence of a new intersections involves the addition of a similar number of events as any other and the IPA process scales accordingly. We illustrate this property

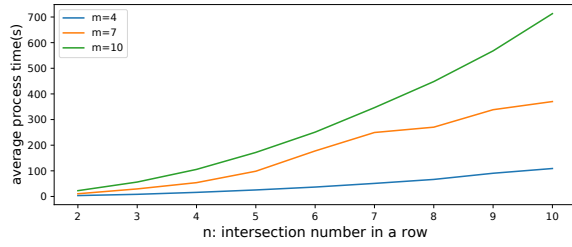


Fig. 10: Scalability: CPU time as a function of the number of intersections

by recording the CPU time of the IPA calculations as a function of the increasing number of intersections as shown in Fig. 10. With a fixed number of parallel row arteries ( $m$ ), increasing the intersections per row from 2 to 10 results in an approximately linear increase in average IPA computation time. Furthermore, the gap in process time between proportional  $m$  values increases evenly. This linear relationship indicates that the computational complexity scales well as more intersections are added. Note that in an actual traffic network the IPA calculations can be distributed over each intersection resulting in parallel processing. Given the observed linear scaling, the TLC approach can be extended to large networks while maintaining tractable computation times.

## V. CONCLUSION AND FUTURE WORK

We have studied a TLC problem for grid-like traffic systems considering straight, left-turn, and right-turn flows. We have considered a flexible traffic structure modeling framework that provides flexibility to model heterogeneous road structures, lane configurations, traffic directions, intersection designs, and traffic signal patterns. The analysis incorporates transit delays and blocking for vehicle movements between intersections. We have used a stochastic hybrid system model and derived IPA gradient estimators of a cost metric with respect to TLC parameters (lower and upper limits to GREEN cycles, as well as thresholds on queue lengths for dynamic adjustments), accounting for delays and blocking through flow burst definitions and burst joining logic. Based on gradient estimates, we adjust the parameters iteratively through an online gradient-based algorithm in order to improve overall performance, with the ability to automatically adapt to changing traffic conditions. Performance is measured through the mean waiting time and a time-distance ratio metric capturing speed information. Our next steps are to include additional traffic flows, such as bicycle and pedestrian traffic, similar to [6], and to develop adaptive methods to automatically determine optimal gradient step sizes that can result in faster convergence.

## REFERENCES

- [1] Y. Zhang, R. Su, K. Gao, and Y. Zhang, "Traffic light scheduling for pedestrians and vehicles," in *2017 IEEE Conference on Control Technology and Applications (CCTA)*. Mauna Lani Resort, HI, USA: IEEE, Aug. 2017, pp. 1593–1598.
- [2] T. Kaur and S. Agrawal, "Adaptive traffic lights based on hybrid of neural network and genetic algorithm for reduced traffic congestion," in *2014 Recent Advances in Engineering and Computational Sciences (RAECS)*. Chandigarh, India: IEEE, Mar. 2014, pp. 1–5.
- [3] C. Dong, K. Yang, J. Guo, X. Chen, H. Dong, and Y. Bai, "Analysis and Control of Intelligent Traffic Signal System Based on Adaptive Fuzzy Neural Network," in *2019 5th Intl. Conf. on Transportation Information and Safety (ICTIS)*. Liverpool, United Kingdom: IEEE, July 2019, pp. 1352–1357.
- [4] J. L. Fleck, C. G. Cassandras, and Y. Geng, "Adaptive Quasi-Dynamic Traffic Light Control," *IEEE Transactions on Control Systems Technology*, vol. 24, no. 3, pp. 830–842, May 2016.
- [5] R. Zhu, S. Wu, L. Li, P. Lv, and M. Xu, "Context-Aware Multiagent Broad Reinforcement Learning for Mixed Pedestrian-Vehicle Adaptive Traffic Light Control," *IEEE Internet of Things Journal*, vol. 9, no. 20, pp. 19 694–19 705, Oct. 2022.
- [6] Y. Chen and C. G. Cassandras, "Adaptive traffic light control for competing vehicle and pedestrian flows," *arXiv preprint arXiv:2303.08173*, 2023.
- [7] C. Gershenson, "Self-Organizing Traffic Lights," Feb. 2005, arXiv:nlin/0411066.
- [8] R. Hassin, "A Flow Algorithm for Network Synchronization," *Operations Research*, vol. 44, no. 4, pp. 570–579, 1996.
- [9] Z. He, L. Wang, D. Li, and L. Zhang, "Steady-State Signal Control for Urban Traffic Networks," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*. Gran Canaria, Spain: IEEE, Sept. 2015, pp. 463–470.
- [10] H. Wang, M. Zhu, W. Hong, C. Wang, G. Tao, and Y. Wang, "Optimizing Signal Timing Control for Large Urban Traffic Networks Using an Adaptive Linear Quadratic Regulator Control Strategy," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 1, pp. 333–343, Jan. 2022.
- [11] S. Lin, B. De Schutter, Y. Xi, and H. Hellendoorn, "Efficient network-wide model-based predictive control for urban traffic networks," *Transportation Research Part C: Emerging Technologies*, vol. 24, pp. 122–140, Oct. 2012.
- [12] W. Huang, L. Li, and H. K. Lo, "Adaptive traffic signal control with equilibrium constraints under stochastic demand," *Transportation Research Part C: Emerging Technologies*, vol. 95, pp. 394–413, Oct. 2018.
- [13] T. Le, H. L. Vu, Y. Nazarathy, B. Vo, and S. Hoogendoorn, "Linear-Quadratic Model Predictive Control for Urban Traffic Networks," *Procedia - Social and Behavioral Sciences*, vol. 80, pp. 512–530, June 2013.
- [14] A. L. C. Bazzan, "A distributed approach for coordination of traffic signal agents," *Autonomous agents and multi-agent systems*, vol. 10, no. 1, pp. 131–164, 2005.
- [15] S. El-Tantawy and B. Abdulhai, "Multi-Agent Reinforcement Learning for Integrated Network of Adaptive Traffic Signal Controllers (MARLIN-ATSC)," in *2012 15th International IEEE Conference on Intelligent Transportation Systems*. Anchorage, AK, USA: IEEE, Sept. 2012, pp. 319–326.
- [16] S. El-Tantawy, B. Abdulhai, and H. Abdelgawad, "Multiagent Reinforcement Learning for Integrated Network of Adaptive Traffic Signal Controllers (MARLIN-ATSC): Methodology and Large-Scale Application on Downtown Toronto," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 3, pp. 1140–1150, Sept. 2013.
- [17] X. Liang, X. Du, G. Wang, and Z. Han, "A Deep Reinforcement Learning Network for Traffic Light Cycle Control," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 1243–1253, Feb. 2019. [Online]. Available: <https://ieeexplore.ieee.org/document/8600382/>
- [18] Y. Wang, T. Xu, X. Niu, C. Tan, E. Chen, and H. Xiong, "STMARL: A Spatio-Temporal Multi-Agent Reinforcement Learning Approach for Cooperative Traffic Light Control," *IEEE Transactions on Mobile Computing*, vol. 21, no. 6, pp. 2228–2242, June 2022.
- [19] R. Chen and C. G. Cassandras, "Stochastic flow models with delays, blocking and applications to multi-intersection traffic light control," *Discrete Event Dynamic Systems*, vol. 30, no. 1, pp. 125–153, Mar. 2020.
- [20] Y. Chen and C. G. Cassandras, "Scalable adaptive traffic light control over a traffic network including transit delays," in *Proceedings of 2023 IEEE Intl. Intelligent Transportation Systems Conference*, 2023.
- [21] C. G. Cassandras, Y. Wardi, C. G. Panayiotou, and C. Yao, "Perturbation Analysis and Optimization of Stochastic Hybrid Systems," *European Journal of Control*, vol. 16, no. 6, pp. 642–661, Jan. 2010.

- [22] B. S. Reddy and N. V. H. Reddy, "Signal design for t-intersection by using webster's method in nandyal town kurnool district of andhra pradesh," *International Research Journal of Engineering and Technology (IRJET) Volume*, vol. 3, 2016.