Mitigating the Impact of Light Rail on Urban Traffic Networks using Mixed Integer Linear Programming

Iain Guilliard, Felipe Trevizan, and Scott Sanner

Abstract

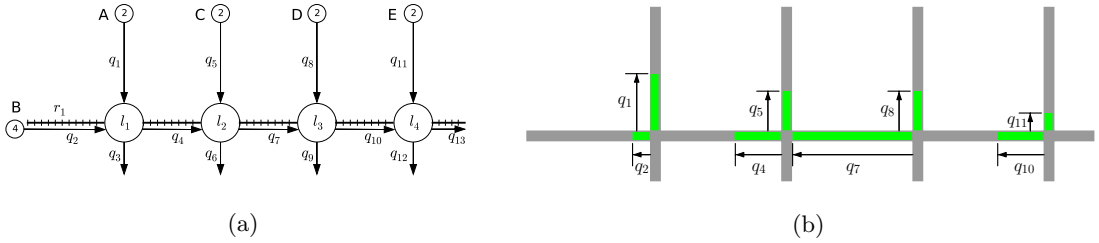
As urban traffic congestion is on the increase worldwide, many cities are increasingly looking to inexpensive public transit options such as light rail that operate at street-level and require coordination with conventional traffic networks and signal control. A major concern in light rail installation is whether enough commuters will switch to it to offset the additional constraints it places on traffic signal control and the resulting decrease in conventional vehicle traffic capacity. In this paper, we study this problem and ways to mitigate it through a novel model of optimized traffic signal control subject to light rail schedule constraints solved in a Mixed Integer Linear Programming (MILP) framework. Our key results show that while this MILP approach provides a novel way to optimize fixed-time control schedules subject to light rail constraints, it also enables a novel optimized adaptive signal control method that virtually nullifies the impact of the light rail presence, reducing average delay times in microsimulations by up to 58.7% vs. optimal fixed-time control.

1 Introduction

As urban traffic congestion is on the increase worldwide with estimated productivity losses in the hundreds of billions of dollars in the U.S. alone and immeasurable environmental impact [1], many cities are increasingly looking to public transit options such as light rail that are less expensive and often more reliable than heavy rail in order to reduce the number of conventional traffic commuters [2]. Since light rail often operates at street-level with exclusive right-of-way and requires coordination with conventional traffic networks and signal control, a major concern in light rail installation is whether enough commuters will switch to it to offset the additional constraints it places on traffic signal control.

Unfortunately, many large cities still use some degree of *fixed-time* control [3] even if they also use *actuated* or *adaptive* control methods such as SCATS [4] or SCOOT [5]; while these methods may support signal pre-emption for light rail crossing, they are unable to autonomously adapt the signal plan to the light rail schedule, hence posing problems for their effective integration with conventional traffic signal control. A more recent trend in the traffic signal control literature proposes the use of *optimized* controllers (that incorporate elements of both adaptive and actuated control) as evidenced in a variety of approaches including mixed integer linear programming (MILPs) [6, 7, 8, 9, 10, 11, 12, 13, 14], heuristic search [15, 16], queuing delay optimization [17, 18], scheduling-driven control [19, 20], and reinforcement learning [3]. While these approaches hold out the promise of more highly optimized traffic control methods, to date, *none* have studied the optimal integration of light rail schedule constraints with their respective methods nor the impact that such integration would have on traffic delay.

Nonetheless, the sub-optimal integration of traffic signal optimization and light rail schedules has been done before, such as in [21], which uses a genetic algorithm coupled with a microsimulator to optimize a subset of traffic signal and transit priority request parameters. Due to the usage of a genetic algorithm, this approach does not necessarily find the global optimum. Another example of signal plan optimization taking into account a schedule of transit priority requests is [22] which represents this problem as a MILP that minimizes a multi-modal delay objective. Additionally, this approach incorporates virtual priority requests to represent vehicle platoon arrivals in order to improve coordination between intersections. This approach is also not globally optimal since each intersection is solved separately. Moreover, while



(a) (b)

Figure 1: (a) Network 1, an arterial road with parallel light rail, used later to evaluate performance. (b) Example of a traffic flow through Network 1 modelled with QTM, showing the state of the queue variable for each link in green. QTM captures the physical queues that form at traffic signals, as well as spill back, where here, q_7 has spilled back into the upstream link forming q_4.

the model includes queuing delay and clearance times, it does not consider start up lost time. Another non-optimal MILP-based approach is introduced in [23] and their MILP represents platoon-based flows to optimize traffic signals with transit priority requests, but not the signal preemption associated with rail transit. The obtained MILP is then solved progressively on pairs of intersections along an arterial route and does not necessarily find the global optimum.

To address the deficiencies of these sub-optimal approaches, we introduce a MILP-based algorithm for optimizing traffic signals constrained by light rail schedules in which it is computationally feasible to find the optimal solution. To the best of our knowledge, this is the first algorithm capable of finding the optimal integration of light rail schedule constraints. In order to do so, we leverage the *Queue Transmission Model* (QTM) [13], a recent MILP model of traffic signal optimization where expected traffic queues and flows are continuous variables, traffic signals are discrete variables, and the overall optimization objective is to minimize delay. Among alternative MILP-based control methods cited previously, the advantages of QTM are high scalability and a focus on the accurate modeling of travel delay between intersections critical for prioritizing light rail arrivals.

Our approach to modelling light rail schedules also allow us to compute optimal fixed-time control policies with light rail constraints and is capable of finding the optimal splits, offsets and cycle time. This is an improvement on previous approaches to this problem (e.g., [6, 9, 14]) which require the cycle time to be fixed exogenously (i.e., a parameter of the algorithm) and are only able to optimize the splits and offsets. Computing the optimal cycle times has an impact beyond fixed-time control since adaptive controllers (e.g., [4]) also require cycle times to be known a priori and our method can be directly used there.

We make the follow key contributions:

1. The first method to globally optimize traffic signals integrated with light rail schedule constraints.
2. We provide a novel fixed-time controller to optimize cycle times, phase splits and offsets. The fixed-time control schedules can include light rail schedule constraints and common cycle length constraints, and can be incorporated *immediately* into existing fixed-time traffic controller infrastructure.
3. We provide a novel way to model lost time directly as a signal timing constraint and we show that it is critical to finding optimized signal plans.
4. We run a comprehensive suit of experiments using a microsimulator to validate the effectiveness of these contributions, both quantitatively and through visual inspection of the simulation results. Gaining insights into the optimal solution's properties can also help to further improve existing control strategies and provides a benchmark.

Our experiments show that optimal adaptive control can reduce traffic delay by up to 58.7% over optimal fixed-time control when light rail is introduced, and virtually nullifies its impact when compared

to using fixed-time control before the introduction of light rail. Ultimately, these results demonstrate a win-win situation where both vehicle traffic and light rail commuters benefit through the application of MILP-based optimization to jointly manage both light rail schedule priority and traffic networks.

2 The Queue Transmission Model (QTM)

To investigate the impact of light rail schedules on conventional traffic networks we need a model of both traffic flow and light rail constraints. As a model of traffic flow, we leverage the Queue Transmission Model (QTM) [13]. Informally, we show an example of a traffic network and the state of the queue variables in a QTM model in Figure 1. Formally a QTM is a tuple $(\mathcal{Q}, \mathcal{L}, \vec{\Delta t}, \mathbf{I})$, where $\mathcal{Q}$ and $\mathcal{L}$ are, respectively, the set of queues and lights; $\vec{\Delta t}$ is a vector of size N representing the discretization of the problem horizon $[0, \mathrm{T}]$ and the duration in seconds of the n-th time interval is denoted as Δt_n; and $\mathbf{I}$ is a matrix $|\mathcal{Q}| \times \mathrm{T}$ in which $I_{i,n}$ represents the flow of cars requesting to enter queue i from the outside of the network at time n.

A **traffic light** $\ell \in \mathcal{L}$ is defined as the tuple $(\Psi_\ell^{\min}, \Psi_\ell^{\max}, \mathcal{P}_\ell, \vec{\Phi}_\ell^{\min}, \vec{\Phi}_\ell^{\max})$, where:

- $\mathcal{P}_\ell$ is the set of phases of ℓ;
- $\Psi_\ell^{\min}$ ($\Psi_\ell^{\max}$) is the minimum (maximum) allowed cycle time for ℓ; and
- $\vec{\Phi}_\ell^{\min}$ ($\vec{\Phi}_\ell^{\max}$) is a vector of size $|\mathcal{P}_\ell|$ and $\Phi_{\ell,k}^{\min}$ ($\Phi_{\ell,k}^{\max}$) is the minimum (maximum) allowed time for phase $k \in \mathcal{P}_\ell$.

A **queue** $i \in \mathcal{Q}$ represents a segment of road that vehicles traverse at free flow speed; once traversed, the vehicles are vertically stacked in a stop line queue. Formally, a queue i is defined by the tuple $(\mathrm{Q}_i, \mathrm{T}_i^{\mathrm{p}}, \mathrm{F}_i^{\mathrm{out}}, \vec{F}_i, \vec{Pr}_i, \mathcal{Q}_i^{\mathcal{P}})$ where:

- Q_i is the maximum capacity of i;
- $\mathrm{T}_i^{\mathrm{p}}$ is the time required to traverse i and reach the stop line;
- $\mathrm{F}_i^{\mathrm{out}}$ represents the maximum traffic flow from i to the outside of the modeled network;
- $\vec{F}_i$ and $\vec{Pr}_i$ are vectors of size $|\mathcal{Q}|$ and their j-th entry (i.e., $\mathrm{F}_{i,j}$ and $\mathrm{Pr}_{i,j}$) represent the maximum flow from queue i to j and the turn probability from i to j ($\sum_{j \in \mathcal{Q}} \mathrm{Pr}_{i,j} = 1$), respectively; and
- $\mathcal{Q}_i^{\mathcal{P}}$ denotes the set of traffic light phases controlling the outflow of queue i.

2.1 Computing Traffic Flows with QTM

In this section, we review the QTM and how to compute traffic flows using QTM as a Linear Program (LP). We assume for the remainder of this section that a *valid* control plan for all traffic lights is fixed and given as a parameter; formally, for all $\ell \in \mathcal{L}$, $k \in \mathcal{P}_\ell$, and interval $n \in \{1, \ldots, N\}$, the binary variable $p_{\ell,k,n}$ is known a priori and indicates if phase k of light ℓ is active (i.e., $p_{\ell,k,n} = 1$) or not on interval n.

We represent the problem of finding the maximal flow between capacity-constrained queues as an LP over the following variables defined for all intervals $n \in \{1, \ldots, \mathrm{N}\}$ and queues i and j:

- $q_{i,n} \in [0, \mathrm{Q}_i]$, traffic volume waiting in the stop line of queue i at the beginning of interval n;
- $f_{i,n}^{\mathrm{in}} \in [0, I_{i,n}]$, inflow to the network via queue i during interval n;
- $f_{i,n}^{\mathrm{out}} \in [0, \mathrm{F}_i^{\mathrm{out}}]$, outflow from the network via queue i during interval n; and
- $f_{i,j,n} \in [0, \mathrm{F}_{i,j}]$, flow from queue i into queue j during interval n.

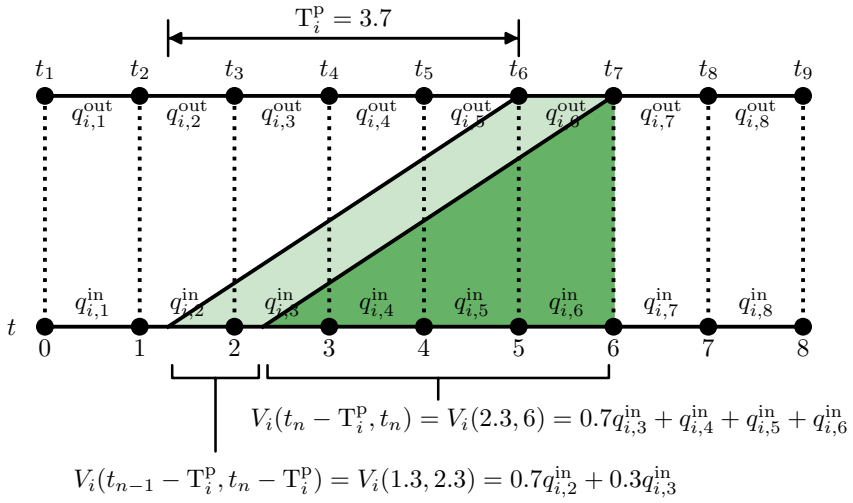


Figure 2: An example of using $V_i(x, y)$ for a queue i, to find the volume of traffic arriving at the stop line during interval 6, and the total volume of traffic within the link.

The maximum traffic flow from queue i to queue j is enforced by constraints (C1) and (C2). (C1) ensures that only the fraction $\mathrm{Pr}_{i,j}$ of the total internal outflow of i goes to j, and (C2) forces the flow from i to j to be zero if all phases controlling i are inactive. (i.e., $p_{\ell,k,n} = 0$ for all $k \in \mathcal{Q}_i^{\mathcal{P}}$).

$$f_{i,j,n} \leq \mathrm{Pr}_{i,j} \sum_{k=1}^{|\mathcal{Q}|} f_{i,k,n} \tag{C1}$$

$$f_{i,j,n} \leq \mathrm{F}_{i,j} \sum_{p_{\ell,k,n} \in \mathcal{Q}_i^{\mathcal{P}}} p_{\ell,k,n} \tag{C2}$$

To simplify the presentation of the remainder of the LP, we define $q_{i,n}^{\mathrm{in}}$ (C3) and $q_{i,n}^{\mathrm{out}}$ (C4) to, respectively, represent the volume of traffic to enter and leave queue i during interval n.

$$q_{i,n}^{\mathrm{in}} = \Delta t_n (f_{i,n}^{\mathrm{in}} + \sum_{j=1}^{|\mathcal{Q}|} f_{j,i,n}) \tag{C3}$$

$$q_{i,n}^{\mathrm{out}} = \Delta t_n (f_{i,n}^{\mathrm{out}} + \sum_{j=1}^{|\mathcal{Q}|} f_{i,j,n}) \tag{C4}$$

The volume of traffic that entered queue i between two time points x and y is denoted as $V_i(x, y)$ and defined in equation (1), where m and w are the indexes of the time intervals s.t. $t_m \leq x < t_{m+1}$ and $t_w \leq y < t_{w+1}$, and is obtained by summing the segments of $q_{i,n}^{\mathrm{in}}$ overlapping the interval $[x, y]$.

$$V_i(x, y) = (t_{m+1} - x) \frac{q_{i,m}^{\mathrm{in}}}{\Delta t_m} + \left(\sum_{k=m+1}^{w-1} q_{i,k}^{\mathrm{in}} \right) + (y - t_w) \frac{q_{i,w}^{\mathrm{in}}}{\Delta t_w} \tag{1}$$

Using these helper variables, (C5) represents the flow conservation principle for queue i where $V_i(t_{n-1} - \mathrm{T}_i^{\mathrm{p}}, t_n - \mathrm{T}_i^{\mathrm{p}})$ is the volume of cars that reached the stop line during Δt_{n-1}, and t_n represents the time elapsed since the beginning of the problem until the start of interval n such that $t_n = t_{n-1} + \Delta t_{n-1}$, and $t_1 = 0$. Notice that (C5) represents a non-first order Markovian update because the update considers

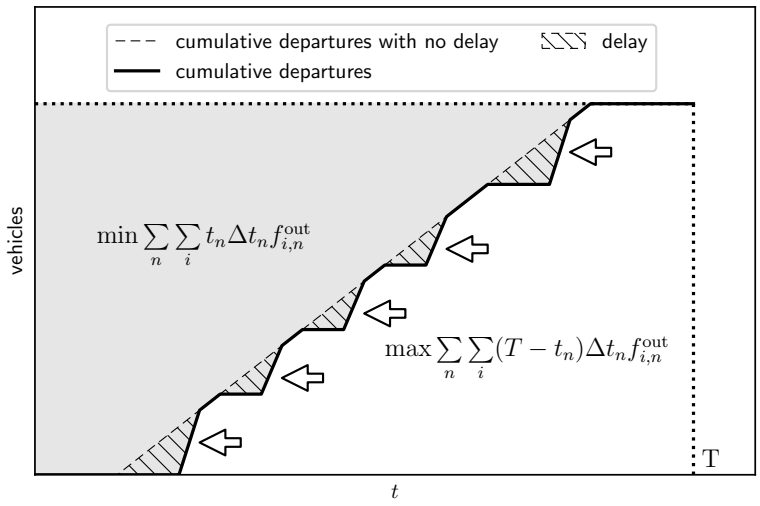


Figure 3: Example of a total cumulative departure curve for a network. Total delay in vehicle seconds is represented by the hatched region between the departure curve and its upper bound, the departure curve with no delay. The first term of the objective function (O1) seeks to maximize the area under the departure curve. This is equivalent to minimizing the area *above* the departure curve (shaded region) and therefore (O1) minimizes delay.

the number of time steps spanned by $\mathrm{T}_i^{\mathrm{p}}$. To ensure that the total volume of traffic traversing i (i.e., $V_i(t_n - \mathrm{T}_i^{\mathrm{p}}, t_n)$) and waiting at the stop line does not exceed the capacity of the queue, (C6) is enforced. See Fig. 2 for an example of the volume functions.

$$q_{i,n} = q_{i,n-1} - q_{i,n-1}^{\text{out}} + V_i(t_{n-1} - \mathrm{T}_i^{\mathrm{p}}, t_n - \mathrm{T}_i^{\mathrm{p}}) \tag{C5}$$

$$V_i(t_n - \mathrm{T}_i^{\mathrm{p}}, t_n) + q_{i,n} \leq \mathrm{Q}_i \tag{C6}$$

QTM uses the objective function (O1) to minimize total delay in the network. By maximizing each flow, $f_{i,n}^{\text{out}}$, $f_{i,n}^{\text{in}}$ and $f_{i,j,n}$ against its upper bound, weighted by the time remaining until the end of the problem horizon T, the optimizer is forced to allow as much traffic volume as possible into the network and move traffic to the outside of the network as soon as possible.

$$\begin{aligned} \max & \sum_{n=1}^{\mathrm{N}} \sum_{i=1}^{|\mathcal{Q}|} (\mathrm{T} - t_n) \Delta t_n f_{i,n}^{\text{out}} + \sum_{n=1}^{\mathrm{N}} \sum_{i=1}^{|\mathcal{Q}|} (\mathrm{T} - t_n) \Delta t_n f_{i,n}^{\text{in}} \\ & + \beta \sum_{n=1}^{\mathrm{N}} \sum_{i=1}^{|\mathcal{Q}|} (\mathrm{T} - t_n) \sum_{j=1}^{|\mathcal{Q}|} \Delta t_n f_{i,j,n}) \end{aligned} \tag{O1}$$

The first term of objective (O1) corresponds to minimizing delay. To see this, consider the objective transformation $\max J = -\min -J$ and, after expanding, we have the equivalent minimization (2). Without loss of generality, if we consider the case where all traffic clears the network, then the sum $\sum_n \sum_i T \Delta t_n f_{i,n}^{\text{out}}$ is a constant and (2) can be reduced to (3). Since Δt_n is implicitly 1 in CTM, (3) is equal to the objective function given in [9], which was shown to be the minimization of total delay.

Fig. 3 provides a graphical interpretation of (2) as the difference of areas, where the first term of (O1) is the area below the cumulative departure curve for the network, the sum $\sum_n \sum_i T \Delta t_n f_{i,n}^{\text{out}}$ is the area

of the dotted rectangle enclosing the curve, and the shaded area above the curve is given by (3).

$$-\min \sum_{n=1}^{\mathrm{N}} \sum_{i=1}^{|\mathcal{Q}|} t_n \Delta t_n f_{i,n}^{\text{out}} - \sum_{n=1}^{\mathrm{N}} \sum_{i=1}^{|\mathcal{Q}|} T \Delta t_n f_{i,n}^{\text{out}} \tag{2}$$

$$\min \sum_{n=1}^{\mathrm{N}} \sum_{i=1}^{|\mathcal{Q}|} t_n \Delta t_n f_{i,n}^{\text{out}} \tag{3}$$

The second term of (O1) maximizes the inflow at the rate given by $\mathbf{I}$, but allows for elasticity in the case of any queue spill back that blocks an input.

The third term of (O1) ensures that the optimizer always moves vehicles from i to j when the associated traffic phase is active and j is not full. As described in [9], the value of β should be sufficiently small to avoid interfering with the main objective by giving too much priority to the internal flows.

3 Traffic Control with MILP-encoded QTM

In this section, we show how to compute the optimized adaptive control plan by extending the LP (O1, C1–C6) into an Mixed-Integer LP (MILP). Formally, for all $\ell \in \mathcal{L}$, $k \in \mathcal{P}_\ell$, and interval $n \in \{1, \ldots, N\}$, the phase activation parameter $p_{\ell,k,n} \in \{0, 1\}$ becomes a free variable to be optimized. In order to obtain a valid control plan, only one phase of traffic light ℓ is allowed to be active at any interval n (C7). Moreover, phase changes follow a (cyclic) fixed ordered sequence indexed by k (C8), where (C8) assumes that $k+1$ equals 1 if $k = |\mathcal{P}_\ell|$.

$$\sum_{k=1}^{|\mathcal{P}_\ell|} p_{\ell,k,n} = 1 \tag{C7}$$

$$p_{\ell,k,n-1} \leq p_{\ell,k,n} + p_{\ell,k+1,n} \tag{C8}$$

Next, the minimum and maximum phase durations (i.e., $\Phi_{\ell,k}^{\min}$ and $\Phi_{\ell,k}^{\max}$) for each phase $k \in \mathcal{P}_\ell$ of traffic light ℓ is enforced. To encode these constraints, we use the helper variable $d_{\ell,k,n} \in [0, \Phi_{\ell,k}^{\max}]$, defined by constraint (C9), that: (i) holds the elapsed time since the start of phase k when $p_{\ell,k,n}$ is active; (ii) is constant and holds the duration of the last phase until the next activation when $p_{\ell,k,n}$ is inactive; and (iii) is restarted when phase k changes from inactive to active. (C9) is encoded using the *big-M* method and $\Phi_{\ell,k}^{\max}$ as the large constant since $d_{\ell,k,n} \leq \Phi_{\ell,k}^{\max}$ and $\Delta t_n \leq \Phi_{\ell,k}^{\max}$. Similarly, constraint (C10) ensures the minimum phase time of k and is not enforced while k is still active.

$$d_{\ell,k,n} = \begin{cases} d_{\ell,k,n-1} + \Delta t_{n-1}, & p_{\ell,k,n-1} = p_{\ell,k,n} = 1 \\ 0, & p_{\ell,k,n-1} = 0, p_{\ell,k,n} = 1 \\ d_{\ell,k,n-1}, & \text{otherwise} \end{cases} \tag{C9}$$

$$d_{\ell,k,n} \geq \Phi_{\ell,k}^{\min}(1 - p_{\ell,k,n}) \tag{C10}$$

Lastly, QTM constrains the sum of all the phase durations for light ℓ to be within the cycle time limits $\Psi_\ell^{\min}$and $\Psi_\ell^{\max}$ (C11). Additionally, (C11) is enforced right after the end of the each cycle, i.e., when its first phase is changed from inactive to active.

$$\Psi_\ell^{\min}(p_{\ell,1,n} - p_{\ell,1,n-1}) \leq d_{\ell,1,n-1} + \sum_{k=2}^{|\mathcal{P}_\ell|} d_{\ell,k,n} \leq \Psi_\ell^{\max} \tag{C11}$$

In order to allow the controller to optimize an initial phase offset at the start of the plan, C9 and C11 are only applied for $n > 1$.

The MILP (O1, C1–C11) encodes the problem of finding the optimized adaptive traffic control plan in a QTM network without light rail.

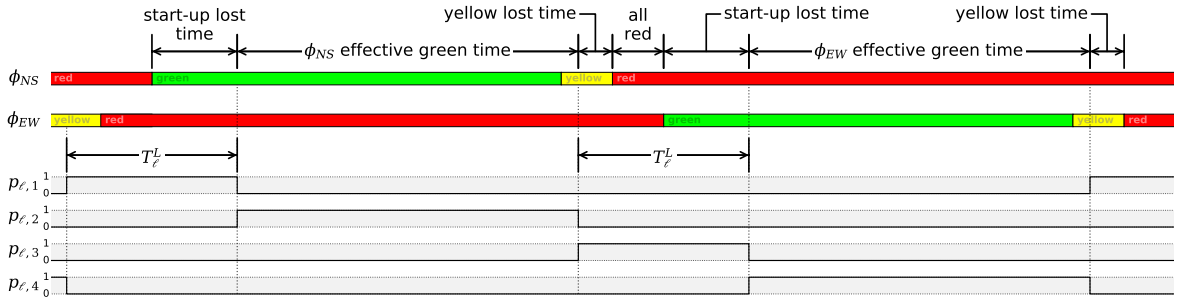


Figure 4: Signal timing for a crossroad intersection with two phases ϕ_{NS} and ϕ_{EW} controlling the north-south and east-west approaches respectively. The intersection is modeled using a QTM light ℓ, with four consecutive phases: $\mathcal{P}_\ell = \{p_{\ell,1}, p_{\ell,2}, p_{\ell,3}, p_{\ell,4}\}$ with $\mathcal{Q}^{\mathcal{P}}_{NS} = \{p_{\ell,2}\}$, $\mathcal{Q}^{\mathcal{P}}_{EW} = \{p_{\ell,4}\}$ and $p_{\ell,1}, p_{\ell,3}$ are the lost time phases of fixed duration T^L_ℓ. The active states of $p_{\ell,2}$ and $p_{\ell,4}$ represent the *effective* green time of their corresponding approaches, where traffic flows at the free flow speed. The lost time is represented by the active states of $p_{\ell,1}$ and $p_{\ell,3}$, which inhibit all traffic flow, forcing the solver to clear the intersection between each signal change and incur lost time delay.

3.1 Lost Time

An additional delay experienced by traffic within a signalized network is the time lost during signal changes. This lost time is made up of several components:

- **Start up lost time**: the time require for a driver waiting at the stop line to react to a green signal and accelerate up to the free flow speed.
- **Yellow lost time**: the remaing time of a yellow signal where drivers react and come to a stop.
- **All red time**: the time preceding the start of every green signal where all approaches are held red to allow vehicles from the previous phase to clear the intersection.

As with other MILP formulations, the active states of QTM phase variables represent the *effective* green time of the associated traffic signal phases, i.e, the time during which vehicles are flowing at the free flow speed, but switch instantaneously between phases in the cycle, without consideration for the lost time associated with the signal change [24, 25]. We extend QTM to model lost time by inserting additional fixed duration phase variables into the cycle at each signal change that inhibit the flow of traffic when active. If the lost time per signal change for light ℓ is T^L_ℓ, then we fix the duration of lost time phase k with $\Phi^{\min}_{\ell,k} = \Phi^{\max}_{\ell,k} = T^L_\ell$, and the solver cannot transition from one signal state to the next without first incurring a delay of T^L_ℓ. In general, a total of n additional phase variables are needed per cycle, where n is the number of signal phases in the cycle.

To obtain a signal plan using only the original n phases, the solution of a QTM with lost time network is post-processed by removing the lost time phases and adjusting the start (end) of each green time by the start up (yellow) lost time, leaving an all red time between signal changes. Fig. 4 shows a signal plan for a crossroad with two phases ϕ_{NS} and ϕ_{EW}, modeled using a QTM light ℓ with four phases. The durations of $p_{\ell,2}$ and $p_{\ell,4}$ represent the effective green time of signal phases ϕ_{NS} and ϕ_{EW}, and are optimized by the solver. The durations of $p_{\ell,1}$ and $p_{\ell,3}$ are fixed and represent the lost time associated with the signal changes.

3.2 QTM as a Fixed-Time Controller

We can further extend QTM to compute an optimized control plan with fixed phase durations. For all $\ell \in \mathcal{L}$, $k \in \mathcal{P}_\ell$, we introduce the new variable $\phi^{\text{fixed}}_{\ell,k} \in [\Phi^{\min}_{\ell,k}, \Phi^{\max}_{\ell,k}]$ and replace the bounds constraints

on $d_{\ell,k,n}$ (that is, $d_{\ell,k,n} \leq \Phi^{\max}_{\ell,k}$ and C10) with fixed duration constraints (C12) and (C13).

$$d_{\ell,k,n} \leq \phi^{\text{fixed}}_{\ell,k} \tag{C12}$$
$$d_{\ell,k,n} \geq \phi^{\text{fixed}}_{\ell,k} - \Phi^{\max}_{\ell,k} p_{\ell,k,n} \tag{C13}$$

Similarly to the variable phase duration constraints, the *big-M* method is employed in (C12) and (C13), using $\Phi^{\max}_{\ell,k}$ as the constant, to enforce $d_{\ell,k,n} = \phi^{\text{fixed}}_{\ell,k}$ only while the phase is inactive. The constraints (C9, C11, C12, C13), allow the fixed-time controller to optimize the phase splits, cycle length and offset for each light ℓ.

A further utility to aid coordination between intersections, is to enforce a common cycle length among a set of lights. To force a common cycle length optimized by QTM, we introduce the new variable $\psi^{\text{fixed}} \in [\max_\ell \{\Psi^{\min}_\ell\},\ \min_\ell \{\Psi^{\max}_\ell\}]$. We can then replace (C11) with the new constraints (C14) and (C15).

$$d_{\ell,1,n-1} + \sum_{k=2}^{|\mathcal{P}_\ell|} d_{\ell,k,n} \leq \psi^{\text{fixed}} \tag{C14}$$
$$d_{\ell,1,n-1} + \sum_{k=2}^{|\mathcal{P}_\ell|} d_{\ell,k,n} \geq \psi^{\text{fixed}} - \mathcal{M}(1 - p_{\ell,1,n} + p_{\ell,1,n-1}) \tag{C15}$$

(C14) and (C15) constrain the sum of all phase durations for light ℓ to equal ψ^{fixed}. Similar to (C11), constraint (C15) is enforced right at the end of each cycle using the *big-M* method, where $\mathcal{M} = \max_\ell \{\Psi^{\max}_\ell\}$. We consider here a single global ψ^{fixed} for all $\ell \in \mathcal{L}$, however it would be trivial to have disjoint subsets of $\mathcal{L}$ corresponding to different regions of the network, each with its own localized ψ^{fixed}.

With the addition of constraints (C12) to (C15), We now have four different controllers available to us:

1. MILP (O1, C1–C11), a fully optimized adaptive controller.
2. MILP (O1, C1–C9, C11–C13), a fixed-time controller able to optimize the phase splits, cycle length and offset for each light ℓ.
3. MILP (O1, C1–C9, C12–C15), a fixed-time controller, but with the additional constraint of a common cycle length between lights.
4. MILP (O1, C1–C10, C14, C15) an optimized adaptive controller but also with a fixed, common cycle length.

3.3 Light Rail Constraints

As a novel extension of the QTM to incorporate a fixed-schedule light rail, we add constraints to the MILP model to fix the free variable $p_{\ell,k,n}$ for all n s.t. the light rail uses phase k of ℓ at time n. Formally, given a schedule $S_\ell(k,n) \in \{0,1\}$ where 1 represents that the light rail uses phase k of ℓ at time n, we replace (C9) and (C10) by (C16) and (C17) when $\sum_{k \in \mathcal{P}_\ell} S_\ell(k,n) > 0$.

$$p_{\ell,k,n} = S_\ell(k,n) \tag{C16}$$
$$d_{\ell,k,n} = d_{\ell,k,n-1} \tag{C17}$$

(C16) enforces that the correct phase k is active when the light rail reaches the traffic light ℓ, and (C17) ensures that the light rail can pass through ℓ even if more than the maximum phase time $\Phi^{\max}_{\ell,k}$ is necessary.

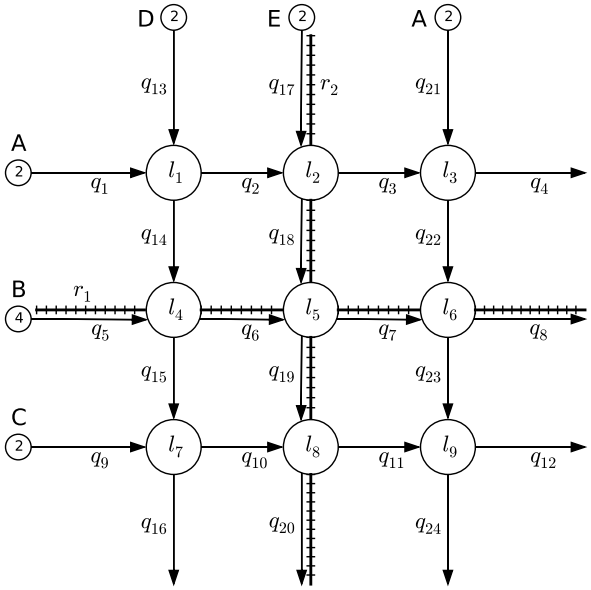


Figure 5: Network 2, an urban grid with crisscrossing streets and light rail.

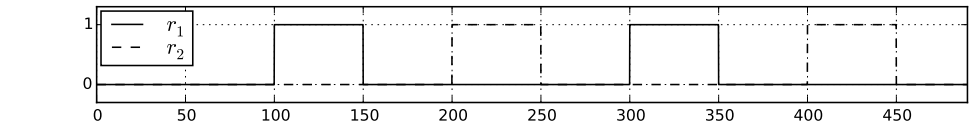
4 Empirical Evaluation

In this section we compare the solutions for traffic networks modeled using QTM with lost time before and after the introduction of a light rail. We consider both fixed-time control, i.e., a non-adaptive control plan, and optimized adaptive control obtained by solving the MILP (O1, C1–C11, C16, C17) for the optimized controller, and the MILP (O1, C1–C9, C11–C13, C16, C17) for the fixed controller. For comparison, we also use the optimized adaptive controller with common cycle time constraints, MILP (O1, C1–C10, C14–C17), and the fixed controller with common cycle time constraints, MILP (O1, C1–C9, C12–C17).

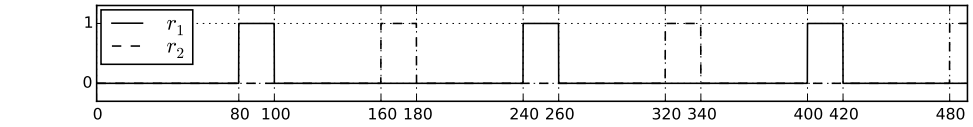
All the computed controllers are simulated using an Intellegent Driver Model (IDM) based microsimulator [26]. As a car following model, IDM will maintain a given safe time headway between vehicles, while also trying to achieve the given desired velocity. Vehicles encountering red signals or stationary vehicles will decelerate and stop. Upon a signal change from green to yellow, the simulator will apply a given braking deceleration to any vehicles estimated not to cross the stopline within the time before the red signal, if continuing at their current speed. The simulated total travel time and observed delay distribution of each controller are used as comparison metrics. Our hypothesis is that our optimized adaptive approach is able to mitigate the impact of introducing light rail w.r.t. both metrics.

Microsimulation Parameters: We choose IDM parameters similar to those suggested in [26], that give realistic values for urban traffic with a flow capacity of 0.5 vehicles/s and a jam density of 0.15 vehicles/m. To simulate the average conditions, we give all vehicles the same parameter values: length $l = 4.67$ m, desired velocity $v_0 = 15$ m/s, safe time headway $T = 1$ s, maximum acceleration $a = 2$ m/s^2, desired deceleration $b = 3$ m/s^2, acceleration exponent $\delta = 4$, and jam distances $s_0 = s_1 = 2$ m. Fig. 7 shows flow-density samples from an IDM microsimulation with these values, and how the QTM flow parameters used in the experiments are calibrated.

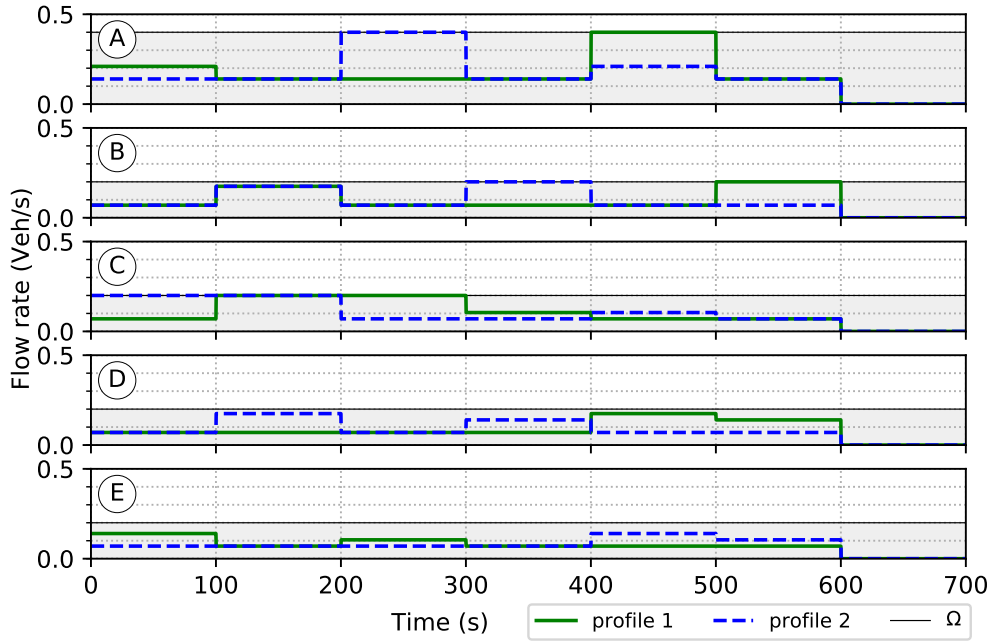
Network Parameters: We consider two networks of differing complexity: an arterial crossed by four side streets (Fig. 1a) and a 3-by-3 grid (Fig. 5). The queues receiving vehicles from outside of the network are marked in Figs. 1a and 5 and we refer to them as input queues. The maximum queue capacity (Q_i) is 60 vehicles for non-input queues and infinity for input queues to prevent interruption of the input demand due to spill back from the stop line. The free flow speed $v_f = 13.2$ m/s and the traversal time of each queue i (T_i^p) is set at 30s, except for the output queues on Network 1 where the traversal time is 10s. For each street, flows are defined from the head of each queue i into the tail of the next queue



(a)



(b)



(c)

Figure 6: Light rails schedules: (a) **slow** (long and infrequent) light rail; and (b) **fast** (short and frequent) light rail. (c) Example of two different demand profiles (inflow rates) applied to the Network 1 inputs corresponding to the letter labels in each plot, where Ω is the maximum inflow rate as annotated on each input in vehicles per Δt.

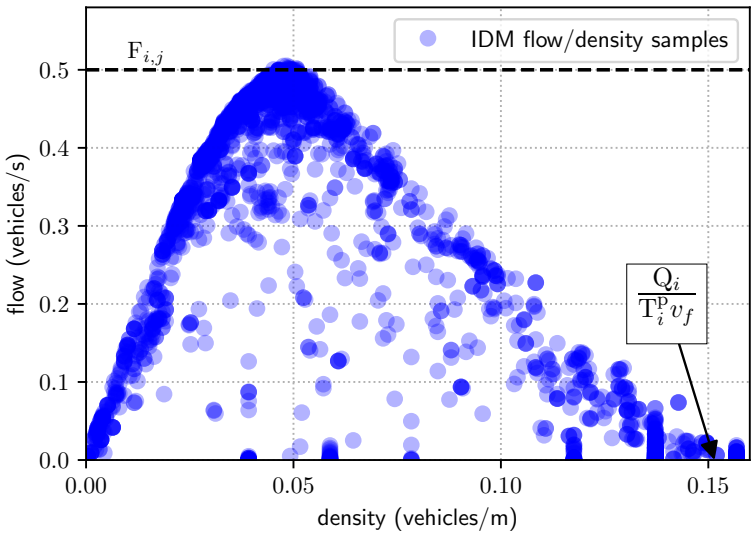


Figure 7: Flow-density samples from an IDM microsimulation of Network 2 along $q_{5,}$ to $q_{8,}$, showing how QTM parameters $F_{i,j} = 0.5$ and $Q_i = 60$ are calibrated to the simulation.

j; there is no turning traffic ($Pr_{i,j} = 1$), and the maximum flow rate between queues, $F_{i,j}$, is set at 0.5 vehicles/s. All traffic lights have two phases, north-south and east-west, and for each traffic light ℓ and phase k, $\Phi^{\min}_{\ell,k}$ is 10s, $\Phi^{\max}_{\ell,k}$ is 60s, $\Psi^{\min}_{\ell}$ is 40s, and $\Psi^{\max}_{\ell}$ is 140s. Whenever lost time is considered, we use $T^L_\ell = 10$s for all $\ell \in \mathcal{L}$, made up of 6s of startup lost time, 2s of yellow lost time and 2s of all red.

Demand Profiles:

Each network is evaluated at increasing demand levels up to the point where $f^{\text{in}}_{i,n}$ becomes saturated. For each demand level, traffic is injected into the network in bursts of 100s for a total of 600s, and the number of vehicles entering the network through i at time n is defined as $I_{i,n} = {}^1/_{\Delta t_n} \max(\xi \Omega_i w_i(t_n), \Omega_i)$ where: w_i is the weight function corresponding to the letter label of i in Figs. 1a and 5, and defined in Appendix A; Ω_i is the maximum inflow rate in vehicles per Δt_n, as annotated at the start of queue i in Figs. 1a and 5; and $\xi \in (0, 2]$ is the scaling factor for the demand level being evaluated.

Light Rail Parameters: we use two different light rail schedules: a slow light rail with a crossing duration of 50s, a period of 200s, and a travel time of 100s between lights (Fig. 6a); and a fast light rail with a crossing duration of 20s, period of 160s, and travel time of 80s between lights (Fig. 6b). On Network 2, the North-South schedule is offset by 100s for the slow light rail and 80s for the fast light rail to avoid a collision at l_5.

Evaluation: We evaluate each network in two scenarios: before the introduction of light rail and after, and in each scenario using both a fixed-time controller and an optimized adaptive controller. For each experiment, we perform one or more *runs* where a run consists of: (i) generate a random demand profile P using the algorithm presented in Appendix A; (ii) compute the signal plan using QTM configured as either an optimized adaptive controller or a fixed-time controller for the demand profile P; and (iii) evaluate the obtained signal plan by microsimulation on the demand profile P using IDM. We use a problem horizon T large enough, typically in the range 1000s – 1500s, to allow all traffic to clear the network, that lets us measure the incurred delay in all the vehicles.

For the experiment on Network 1, we perform 10 runs and report their average delay, and we concatenate the observed delays and number of stops of all the 10 runs for the reported boxplots and cumulative distributions. For Network 2, we report 2 different experiments: the first done in a single run (i.e., using a single random demand profile for both the controller and microsimulation), and the second using a single run in which the signal plan is further evaluated by microsimulating with 9 additional random profiles.

The latter experiment allows us to evaluate the robustness of the controllers w.r.t. changes to the assumed input levels since the microsimulation will be performed using demand profiles that the controllers were not optimized for.

For lost time comparison, we use two different configurations of each controller: one with QTM incurring lost time delay and the other without. Before microsimulation, we adjust the green time of the signal plan from the lost time controller to account for the start up and yellow lost time, as illustrated in Fig. 4.

MILP Solver Parameters: For all experiments, we used Gurobi as the MILP solver with a MIP gap accuracy of 0.01% and $\beta = 0.0001$ in constraint (O1). If the solver execution time reaches 144 h then the solver is halted and the best solution found so far is used. Solution times range from typically real time (less than 200s) for optimized adaptive solutions of Network 1, to over 100 h for fixed-time plans of Network 2; however, once the fixed-time solution is found, it can be deployed indefinitely.

5 Results

What is the impact of modelling lost time delay? Fig. 8a shows, for Network 1 without light rail, the average delay per vehicle as a function of demand level under optimized adaptive control. The QTM predicted delay for the controller without lost time is considerably lower, but the policies found with QTM incurring lost time show improved performance under microsimulation and closely match the prediction. Figs. 8b and 8c show the microsimulation time-distance plots at demand level II for several links along the arterial of Network 1. The y-axis of these plot shows the distance along the street, and the x-axis shows the evolution over time. Each black trace represents the journey of a vehicle along the street. Traffic signals at fixed distances down the street appear as red horizontal dashed lines, where a solid bar represents that the phase is inactive (i.e., the light is red) and traffic cannot pass; otherwise the phase is active. The time-distance plots capture the queueing behaviour of the traffic as each vehicle decelerates when approaching congested traffic, or red or yellow light ahead. When a vehicle is stationary, its trace becomes horizontal. Green represents QTM traffic flow prediction, where darker shades represent regions of higher density. Fig. 8b shows the controller with QTM incurring lost time delay, and the prediction closely matches the microsimulation. However, in Fig. 8c the QTM policy without lost time is unrealizable and the microsimulation quickly diverges.

What is the impact of a common cycle time on the controller? Fig. 9 shows, for Network 1, with and without light rail, the average delay per vehicle as a function of demand for each of the four controllers. With no light rail (Fig. 9a) both the optimized adaptive controllers have very similar average delay, while the fixed controllers have the same average delay. An inspection of the fixed signal plans shows them to be identical. With the introduction of the fast light rail (Fig. 9b) the controllers without common cycle length constraints are able to reduce delay further by utilizing different cycle lengths along the arterial. The same outcome was observed on Network 2 without light rail, however with the introduction of the light rail, the controllers with common cycle length constraints were unable to find a feasible solution that also satisfied the light rail schedule.

Considering the lower plots that show the distribution of delay and give an indication of the quality of the solutions, we can see that at demand level II in Fig. 9a and Fig. 9b, both the optimized controllers find policies with similar average delay. But the box plots show that the optimized controller with common cycle length trades a lower median and upper quartile for a higher maximum delay.

These results corroborate with methods already employed by traffic engineers [25], that for regular traffic networks, using a common cycle lengths between adjacent intersections is a useful aid for achieving good coordination, especially with fixed-time control. However, when the network is not regular, or has addition constraints such as lightrail, improved solutions may be found with mixed cycle lengths.

Is it possible to mitigate the impact of light rail on delay? Figs. 10a and 10b show, for each network, the average delay per vehicle as a function of demand for both fixed-time and optimized adaptive control approaches in three scenarios: before the light rail and after the installation of light rail using the slow and the fast schedules. In all cases the controller models lost time. As we hypothesized,

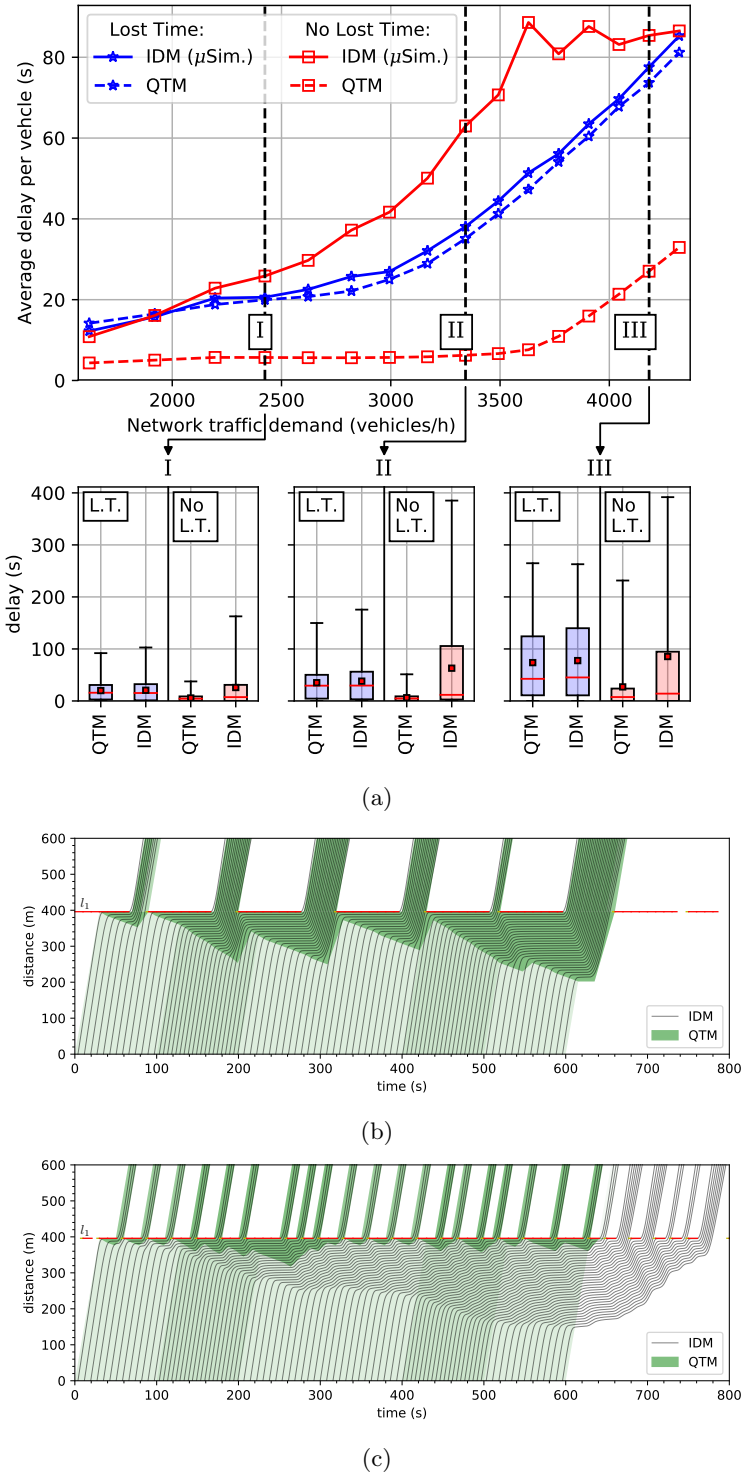


Figure 8: Lost time delay. (a) Upper plot shows QTM average delay, with and without lost time, compared to microsimulation of Network 1 (no light rail) at increasing demand levels. Lower plot: Box plots representing distribution of delay at three different demand levels. Policies found with QTM incurring lost time show improved performance under microsimulation. (b,c) Time-distance plots at demand level II, from q_2 to q_4. (b) QTM incurring lost time delay closely predicts microsimulation, but in (c) QTM policies without lost time are unrealizable and the microsimulation quickly diverges.

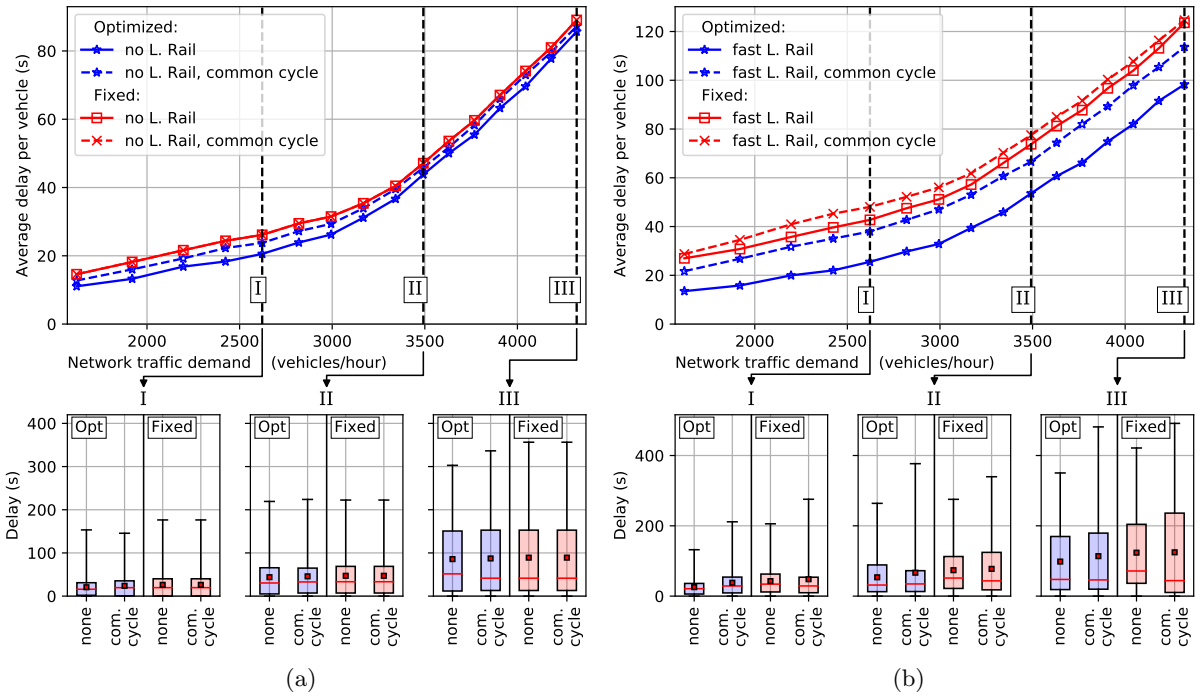


Figure 9: Comparison of the four controllers running on Network 1. (a) With no light rail: The optimized controllers are very close in performance, while both the fixed controllers find the same solution (b) With the fast light rail: The controllers without common cycle constraints are able to further reduce delay by using different cycle lengths along the arterial.

optimized adaptive control is able to mitigate the impact of the introduction of light rail and it marginally increases the average delay when compared with the average delay produced by the fixed-time controller **before** the light rail. Moreover, as shown in Figs. 10c and 10d, the optimized adaptive controller also produces better signal plans than the fixed-time controller, i.e., plans with smaller median, third quartile, and maximum delay.

Table 1 shows the average delay in seconds of the optimized adaptive and fixed-time controllers for both networks. In the scenarios with a light rail, the improvement obtained by the optimized adaptive approach ranges from 20.4% to 58.7% w.r.t. the fixed-time approach. We can see that the optimized adaptive controller successfully nullifies the impact of adding a light rail to the networks since the average delay obtained by it is approximately the same as the average delay for the fixed-time controller with no light rail, with the average delay increased by at most 17.8 s.

What is the impact of unexpected arrivals on the controller? Fig. 11 shows the average delay per vehicle averaged over 10 demand profiles for Network 2, where the arrival rates in the simulation differ over time from the profile used to generate the signal plan. The overall conclusion is still the same: the optimized controller outperforms the fixed-time controller. Compared to the results in Fig. 10b, the optimized controller shows a slight increase in the average delay at lower demand levels and almost no change at higher demand levels while, for the fixed controller, the policy is more robust to the unexpected arrivals and there is minimal change across all demand levels. The fixed controller is more robust to unexpected arrivals because its policy is optimize for the average arrival rate as opposed to an exact demand profile, thus any demand profile with the same or similar average arrival rate will have little impact in the average delay. Since the optimized controller attempts to coordinate signal timings with expected changes in inflow rate to form platoons, the changes in the simulated demand profile has an

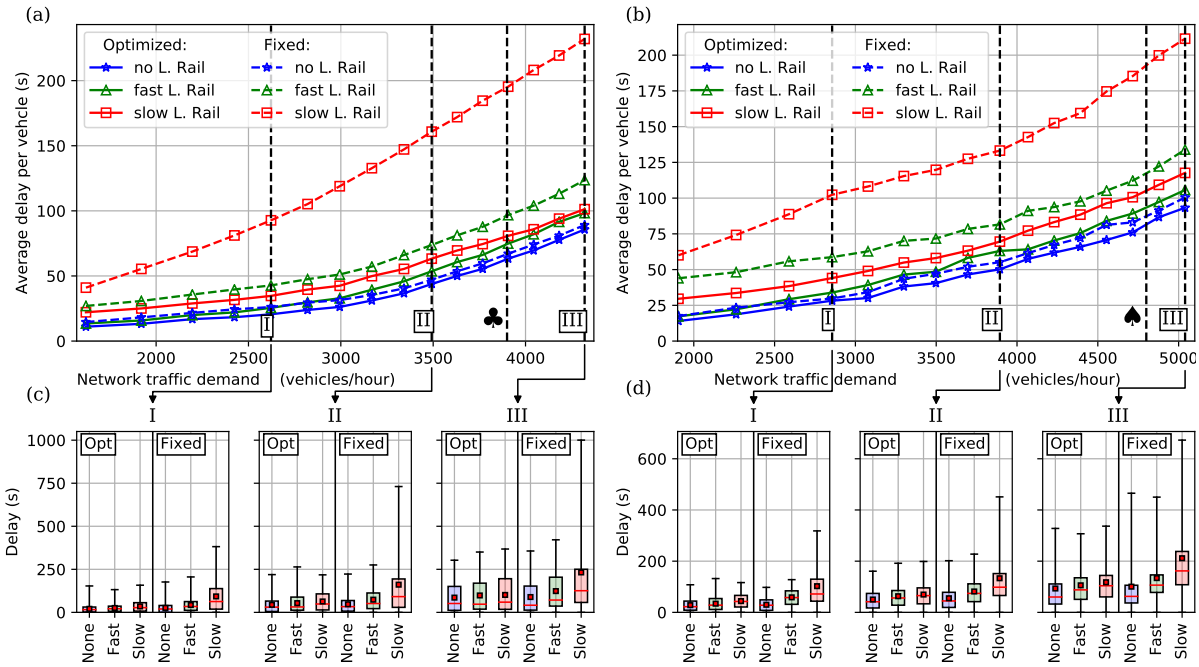


Figure 10: Microsimulation average delay by network demand for the Network 1 (a) and Network 2 (b). (c,d): Box plots representing the observed distribution of delay for 3 different values of demand for each network, comparing delay distribution without the light rail with the impact of the fast and slow schedules. Optimized adaptive control is able to mitigate the impact of light rail on average delay, while also producing better signal plans (lower median, third quartile and maximum delay).

impact the average delay. The impact of unexpected arrivals at lower demand levels is small because some vehicles may need to wait an additional cycle before joining a platoon coordinated by the policy, while at higher demand levels, the residual queue build up at the inputs buffers any mismatch resulting in a negligible impact in the average delay. In all cases, the box plots show the solution quality for each controller is not impacted by the unexpected arrivals.

How many drivers must switch to using light rail to maintain the same average delay? This is a question that will be asked by planners evaluating the impact of adding a light-rail to a traffic network, along with an upgrade to the signal control system. Fig. 12 shows the percentage reduction in average delay as a function of the percentage of vehicles who's drivers are switching to traveling on the light rail. In these plots, the demand level is fixed and higher values are better (i.e., there is a larger decrease in the average delay) and zero means that there is no change after installing light rail. For the three combinations of before and after policies presented, we can see that, while keeping the fixed-time controller requires from 14.2% to 47% of the drives to switch to light rail in order to obtain the same average delay as before its installation, the optimized adaptive approach requires only from 5.8% to 13.1% of the drivers to switch when already using optimized adaptive control **before** the light rail. When compared fixed-time before the light rail and optimized adaptive after, the gains are even greater with only 3% to 9.6% of the drives required to switch to the light rail.

How does the quality of optimized adaptive policies compare with fixed policies? To answer this we show in Figs. 13a to 13d the microsimulation time-distance plots for several streets in Network 1. Figs. 13a and 13b show that both controllers balance between establishing coordinated "green corridors" along the arterial, and servicing the side streets, where the combined density at times exceeds that of the arterial. However, the fixed controller is forced to find a single repeating policy sized for the

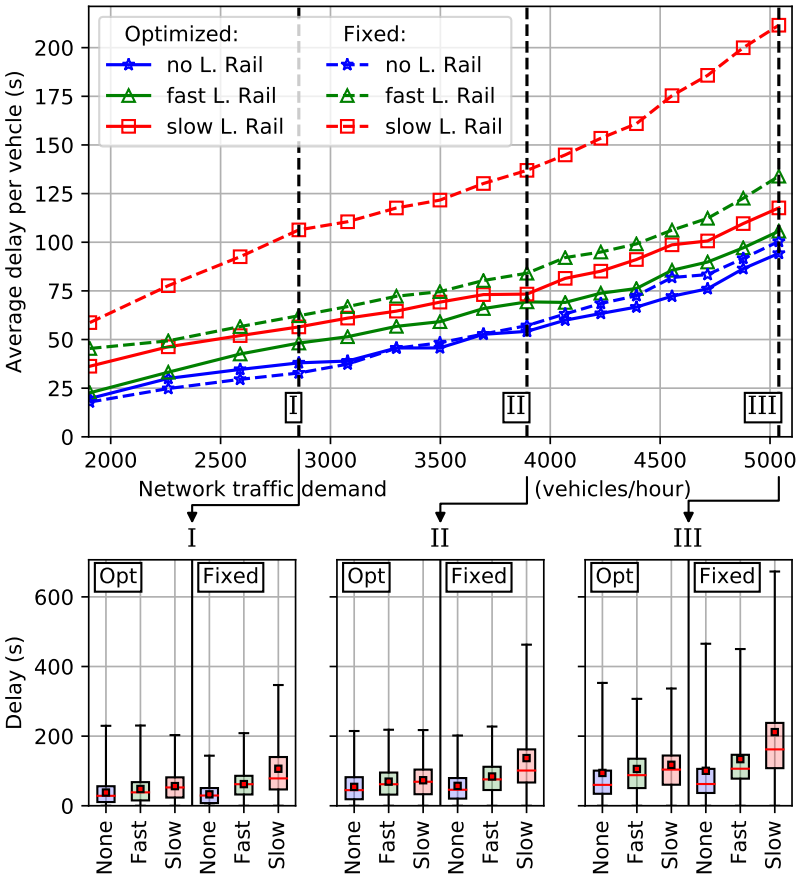


Figure 11: Microsimulation average delay for Network 2, averaged over 10 different demand profiles, where the arrival rates differ from the profile used by each controller. Comparing these results (unexpected arrivals) against Figs. 10b and 10d (arrivals as expected), we can see that the obtained policies are robust.

	Light Rail	Fixed	Opt. Adapt.	**Improv.**
Arterial (♣) @ 3900 Veh/h	None	66.8 s	62.9 s	**5.7%**
	Fast	96.3 s	74.5 s	**22.7%**
	Slow	194.2 s	80.6 s	**58.7%**
Grid (♠) @ 4800 Veh/h	None	87.3 s	81.5 s	**6.7%**
	Fast	117.4 s	93.4 s	**20.4%**
	Slow	192.9 s	105.1 s	**45.5%**

Table 1: Average delay computed via microsimulation in seconds and improvement of optimized adaptive controller over fixed-time controller for both networks and the three light rail scenarios. The demand level is fixed and correspond to the points ♣ and ♠ in Figs. 10a and 10b, respectively. The improvement obtained by our optimized adaptive approach when a light rail is introduced, ranges from 20.4% to 57.8% w.r.t. the fixed-time approach.

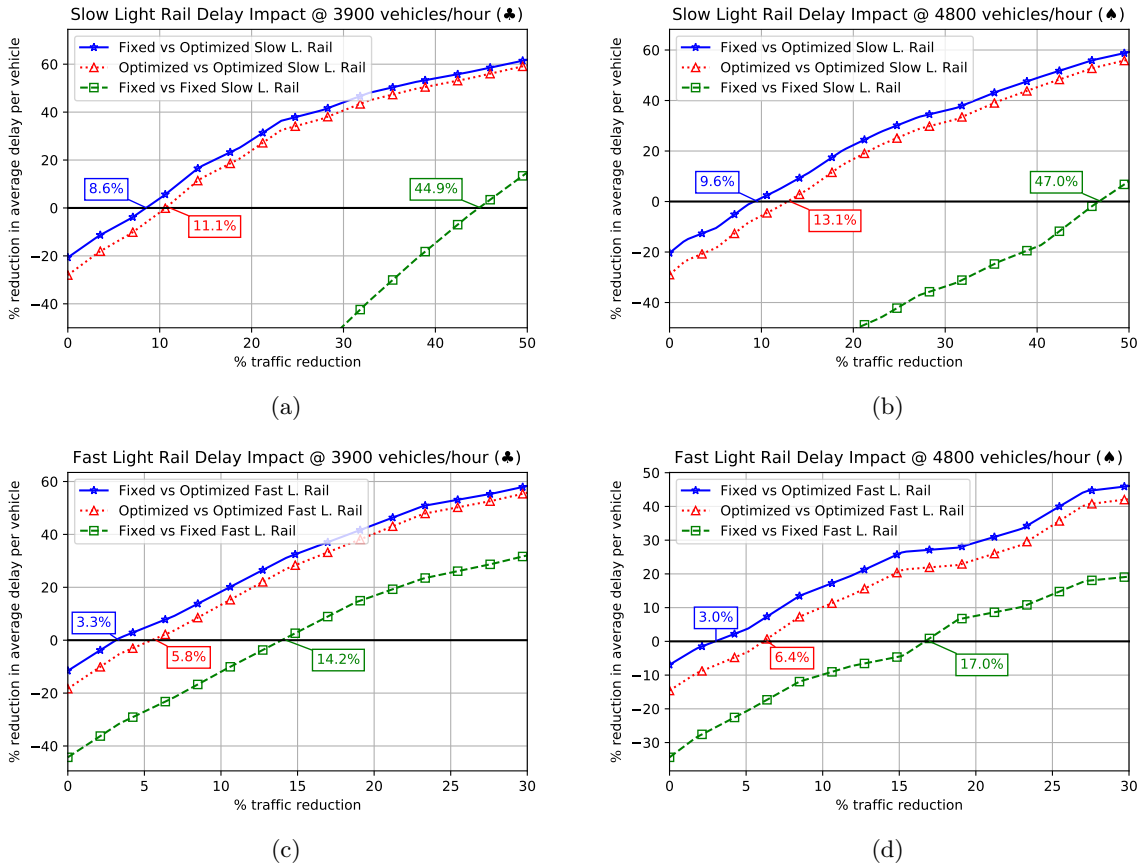


Figure 12: Impact on average delay for the Network 1 (first column) and Network 2 (second column) for both light rail schedules (rows) in different scenarios (curves) of traffic control system before and after installation of light rail. The x-axis is the percentage of vehicles switching to the public transportation and the y-axis is the % reduction in delay after the light rail is installed. Negative % represents an increase in average delay. The vehicle demand for (a-d) are marked as ♣ and ♠ in their respective plots in Fig. 10.

average traffic density in the network. As a byproduct, the side street (Fig. 13c) under the fixed-time controller suffers from accumulative queue build-up following each transit of the light rail. In Figs. 13b and 13d, we see that the optimized adaptive controller is able to clear out the queue build up in the side street by increasing the phase time of the side street for a cycle after the transit has passed through, and then returns to a schedule that prioritizes the arterial depending on the changes in traffic density. When the traffic density in the arterial is higher than the side streets, the optimized adaptive controller will coordinate "green corridors" along the arterial.

Fig. 14 provides more details on the behavior of the signal plans for demand level II (Figs. 10a and 10b) by showing the cumulative number of vehicles by number of observed stops. In all cases for Network 1 and 2 the optimized controller does better with less stops at higher frequencies. For Network 1, we see that both controllers choose to prioritize the side streets over the arterial, with less stops at higher frequencies in the side streets. But in the case of the slow light rail with the optimized controller, 94% of the vehicles experience three or less stops along the arterial while for the fixed controller 100% of the vehicles experience three stops or more.

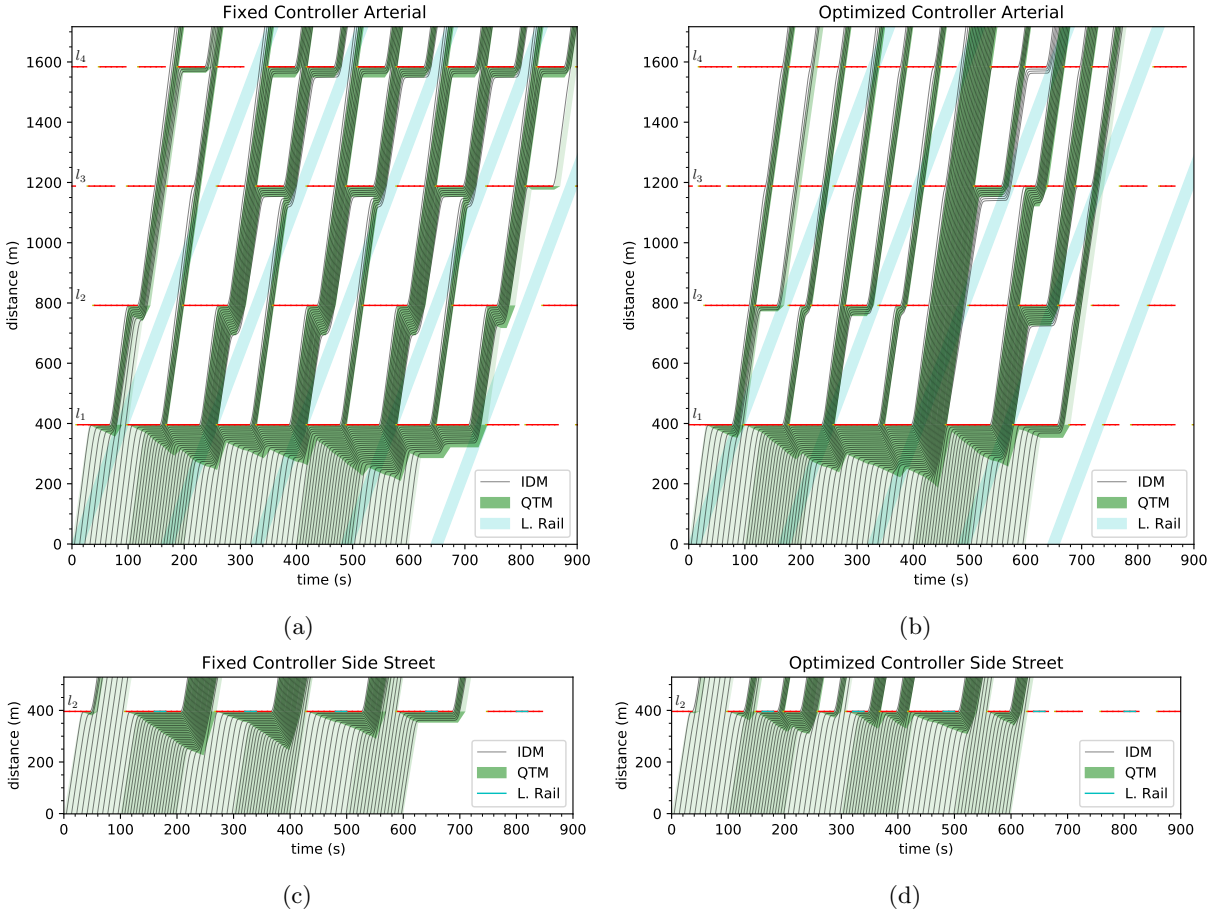


Figure 13: (a-b) Micro-simulation time-distance plots from Network 1 along the links from q_2 to q_{13}, showing that both controllers find well coordinated solutions where the timing of green signals along the link is offset at each intersection to maintain a continuous flow of traffic at the free flow speed, a solution well known to traffic engineers. However, the optimized controller is able to dynamically adjust the "width" of the bands to match the traffic volume along the link, allowing more green time to be allocated to cross traffic. (c-d): Microsimulation time-distance plots of side street q_5 to q_6. While both controllers can find coordinated policies along the arterial, the optimized adaptive controller is able to clear out the queues (horizontal flow lines) in the side streets following the transit of the light rail.

6 Conclusion

In this paper, we introduced a new method to generate an adaptive controller that optimizes traffic signals integrated with light rail schedule constraints. The obtained adaptive controllers are guaranteed to be globally optimal and, to the best of our knowledge, this is the first globally optimal algorithm capable of handling light rail schedule constraints. Our approach is based on the Queue Transmission Model of traffic signal control which we extended to incorporate light rail schedule constraints. We also provided a novel way to model lost time directly as a signal timing constraint and show that it is critical to finding optimized signal plans.

We also introduced a novel approach to compute fixed-time controller plans that optimize cycle times, phase splits and offsets. The obtained fixed-time controllers are also guaranteed to be globally optimal and they can handle both light rail schedule constraints and common cycle length constraints. The computed fixed-time control schedules can be incorporated *immediately* into existing fixed-time traffic

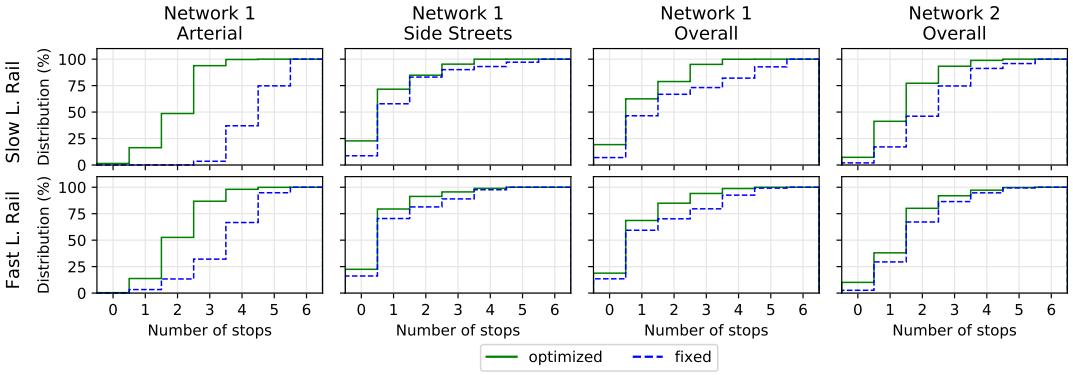


Figure 14: Impact of controller on number of stops as a cumulative distribution. Top row is for the slow light rail schedule. Bottom row is for the fast light rail schedule. In all cases the optimized controller does better with less stops at higher frequencies.

controller infrastructure, yielding important benefits for those municipalities that prefer not to migrate to a fully adaptive control.

Lastly, we have compared our optimal adaptive and fixed-time controllers in a comprehensive suit of experiments using microsimulation as a realistic, finer-grained, nonlinear model of traffic flow. Our results show that the optimal adaptive controller is able to minimize the impact of introducing light rail on conventional traffic networks on the average delay with respect to fixed-time signal control. The experiments also show that the adaptive controllers finds better quality solutions, i.e., solutions with substantially lower third quartile and maximum observed delay. Our key results demonstrate for the first time the potential of MILP-based QTM traffic signal control approaches to virtually nullify the impact of installing light rail on conventional traffic — our model can reduce traffic delay by up to 58.7% over optimal fixed-time control when light rail is introduced. Consequently, the use of MILP-based optimized adaptive controllers like QTM could remove the critical public concern of increased traffic delay resulting from light rail installation, and thus positively impact the environment, urban productivity, and commute time reductions for all commuters.

For future work, a key question to resolve is how large we can scale the traffic and light rail network before we need to investigate decomposition-based approaches to scaling the solution (e.g., MILP-based methods like dual decomposition or region-based traffic network partitioning schemes). Future work should also examine the (online) learnability of QTM parameters from different traffic sensor data, for instance, conventional inductive (double) loop counters, radar, and video feeds. Finally, noting that the nonlinear microsimulation model offers a higher-fidelity model of traffic behavior, future work should consider expanding the QTM to model nonlinear traffic flows [27, 28, 29, 30] and investigating the benefits of nonlinear optimization relative to the existing QTM.

References

[1] Bazzan, A.L.C., Klügl, F.: 'Introduction to Intelligent Systems in Traffic and Transportation'. Synthesis Lectures on Artificial Intelligence and Machine Learning. (Morgan & Claypool Publishers, 2013).

[2] Thompson, G.L. 'Defining an alternative future: Birth of the light rail movement in north america'. In: Transportation Research Board (TRB) Annual Meeting, Washington D.C., January 2003

[3] El.Tantawy, S., Abdulhai, B., Abdelgawad, H.: '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*, 2013, **14**, (3), pp. 1140–1150

[4] Sims, A.G., Dobinson, K.W.: 'The Sydney coordinated adaptive traffic (SCAT) system: Philosophy and benefits', *IEEE Transactions on Vehicular Technology*, 1980, **29**, (2), pp. 130–137

[5] Hunt, P.B., Robertson, D.I., Bretherton, R.D., Winton, R.I. 'SCOOT–a traffic responsive method of coordinating signals'. (Transportation Road Research Lab, Crowthorne, U.K., 1981)

[6] Lo, H.K.: 'A novel traffic signal control formulation', *Transportation Research Part A: Policy and Practice*, 1998, **33**, (6), pp. 433–448

[7] Gartner, N., Little, J.D., Gabbay, H. 'Optimization of traffic signal settings in networks by mixed-integer linear programming'. (DTIC Document, 1974)

[8] Gartner, N.H., Stamatiadis, C.: 'Arterial-based control of traffic flow in urban grid networks', *Mathematical and computer modelling*, 2002, **35**, (5), pp. 657–671

[9] Lin, W.H., Wang, C.: 'An enhanced 0-1 mixed-integer lp formulation for traffic signal control', *Intelligent Transportation Systems, IEEE Transactions on*, 2004, **5**, (4), pp. 238–245

[10] He, Q., Head, K.L., Ding, J.: 'Pamscod: Platoon-based arterial multi-modal signal control with online data', *Procedia-Social and Behavioral Sciences*, 2011, **17**, pp. 462–489

[11] Han, K., Friesz, T.L., Yao, T.: 'A link-based mixed integer LP approach for adaptive traffic signal control', *arXiv preprint arXiv:12114625*, 2012,

[12] Han, K., Liu, H., Gayah, V.V., Friesz, T.L., Yao, T.: 'A robust optimization approach for dynamic traffic signal control with emission considerations', *Transportation Research Part C: Emerging Technologies*, 2016, **70**, pp. 3–26

[13] Guilliard, I., Sanner, S., Trevizan, F.W., Williams, B.C. 'Nonhomogeneous time mixed integer linear programming formulation for traffic signal control.'. *Transportation Research Record: Journal of the Transportation Research Board*, 2016, **2595**, (1), pp. 128–138

[14] Wada, K., Usui, K., Takigawa, T., Kuwahara, M.: 'An optimization modeling of coordinated traffic signal control based on the variational theory and its stochastic extension', *Transportation Research Procedia*, 2017, **23**, pp. 624–644

[15] Lo, H.K., Chang, E., Chan, Y.C.: 'Dynamic network traffic control', *Transportation Research Part A: Policy and Practice*, 1999, **35**, (8), pp. 721–744

[16] He, Q., Lin, W.H., Liu, H., Head, K.L. 'Heuristic algorithms to solve 0–1 mixed integer LP formulations for traffic signal control problems'. In: Service Operations and Logistics and Informatics (SOLI), 2010 IEEE International Conference on., July 2010, pp. 118–124

[17] Varaiya, P.: 'Max pressure control of a network of signalized intersections', *Transportation Research Part C: Emerging Technologies*, 2013, **36**, pp. 177–195

[18] Li, J., Zhang, H.: 'Coupled linear programming approach for decentralized control of urban traffic', *Transportation Research Record: Journal of the Transportation Research Board*, 2014, **2439**, (1), pp. 83–93

[19] Xie, X.F., Smith, S.F., Barlow, G.J. 'Schedule-driven coordination for real-time traffic network control.'. In: ICAPS. 2012.

[20] Smith, S., Barlow, G., Xie, X.F., Rubinstein, Z. 'SURTRAC: Scalable Urban Traffic Control'. In: Transportation Research Board 92nd Annual Meeting Compendium of Papers., Transportation Research Board, Washington D.C., January 2013.

[21] Stevanovic, J., Stevanovic, A., Martin, P.T., Bauer, T.: 'Stochastic optimization of traffic control and transit priority settings in vissim', *Transportation Research Part C: Emerging Technologies*, 2008, **16**, (3), pp. 332 – 349.

[22] He, Q., Head, K.L., Ding, J.: 'Multi-modal traffic signal control with priority, signal actuation and coordination', *Transportation Research Part C: Emerging Technologies*, 2014, **46**, pp. 65–82

[23] Christofa, E., Ampountolas, K., Skabardonis, A.: 'Arterial traffic signal optimization: A person-based approach', *Transportation Research Part C: Emerging Technologies*, 2016, **66**, pp. 27–47

[24] Webster, F.V. 'Traffic signal settings', (London: H.M.S.O., 1958)

[25] Wolshon, B., Pande, A.: 'Traffic engineering handbook'. (John Wiley & Sons, 2016)

[26] Treiber, M., Hennecke, A., Helbing, D.: 'Congested traffic states in empirical observations and microscopic simulations', *Physical Review E*, 2000, **62**, (2), pp. 1805

[27] Lu, S., Dai, S., Liu, X.: 'A discrete traffic kinetic model–integrating the lagged cell transmission and continuous traffic kinetic models', *Transportation Research Part C: Emerging Technologies*, 2011, **19**, (2), pp. 196–205

[28] Muralidharan, A., Dervisoglu, G., Horowitz, R. 'Freeway traffic flow simulation using the link node cell transmission model'. In: American Control Conference, ACC'09, (IEEE 2009), pp. 2916–2921

[29] Kim, Y. 'Online traffic flow model applying dynamic flow-density relation'. In: Eleventh International Conference on Road Transport Information and Control, London, UK, March 2002

[30] Huang, K.C. 'Traffic Simulation Model for Urban Networks: CTM-URBAN', Masters Thesis, Concordia University, 2011

Appendix A

Demand Profiles

To simulate the effect of random arrivals, we use demand profiles with flow rates that vary randomly every 100s for a total duration of 600s:

$$I_{i,n} = \frac{1}{\Delta t_n}\max(\xi\Omega_i \mathrm{w}_i(t_n), \Omega_i)$$

$$w_i(t_n) = \begin{cases} \vec{\mathrm{r}}_k & \text{if } 100(k-1) \le t_n < 100k \text{ and } 1 \le k \le 6 \\ 0 & \text{if } t_n \ge 600 \end{cases}$$

$$\vec{\mathrm{r}} = \frac{1}{2} + \frac{1}{4}\textsc{RandIntVec}(6, 0\,..\,4, \mathrm{v}_i)$$

$$\mathrm{v}_i = \begin{cases} 5 & \text{if } i \text{ is labeled } \mathsf{A} \\ 9 & \text{if } i \text{ is labeled } \mathsf{B} \\ 7 & \text{if } i \text{ is labeled } \mathsf{C} \\ 3 & \text{if } i \text{ is labeled } \mathsf{D} \\ 5 & \text{if } i \text{ is labeled } \mathsf{E} \end{cases}$$

where:

- $\mathrm{w}_i(t_n)$ is the weight function for queue i at time t_n

- Ω_i is the maximum inflow rate in vehicles per Δt_n, as annotated at the start of queue i in Figs. 1a and 5; and
- $\xi \in (0, 2]$ is the scaling factor for the demand level being evaluated.
- $\mathrm{v_i}$ is a constant for queue i corresponding to the letter label of i in Figs. 1a and 5

The function RandIntVec returns a vector of N random integers in the range $[lo\,..\,hi]$, which sum to v. Forcing the random vector for queue i to sum to v_i ensures that across all experiments, the total number of vehicles entering i will be the same.

```
1: function RandIntVec( N, lo..hi , v)
2:     r⃗ ← 0⃗ of length N
3:     sum ← 0
4:     do
5:         for each i = 1..N do:
6:             r⃗_i ← random integer in the range [lo..hi]
7:             sum ← sum + r⃗_i
8:     while sum ≠ v
9:     return r⃗
```