# A robust optimization approach for dynamic traffic signal control with emission considerations

Ke Han[a,*], Hongcheng Liu[b], Vikash V. Gayah[c], Terry L. Friesz[b], Tao Yao[b]

[a]*Department of Civil and Environmental Engineering, Imperial College London, United Kingdom.*
[b]*Department of Industrial and Manufacturing Engineering, Pennsylvania State University, USA.*
[c]*Department of Civil and Environmental Engineering, Pennsylvania State University, USA.*

## Abstract

We consider an analytical signal control problem on a signalized network whose traffic flow dynamic is described by the Lighthill-Whitham-Richards (LWR) model (Lighthill and Whitham, 1955; Richards, 1956). This problem explicitly addresses traffic-derived emissions as constraints or objectives. We seek to tackle this problem using a mixed integer mathematical programming approach. Such class of problems, which we call LWR-Emission (LWR-E), has been analyzed before to certain extent. Since mixed integer programs are practically efficient to solve in many cases (Bertsimas et al., 2011b), the mere fact of having integer variables is not the most significant challenge to solving LWR-E problems; rather, it is the presence of the potentially nonlinear and nonconvex emission-related constraints/objectives that render the program computationally expensive.

To address this computational challenge, we proposed a novel reformulation of the LWR-E problem as a *mixed integer linear program* (MILP). This approach relies on the existence of a statistically valid macroscopic relationship between the aggregate emission rate and the vehicle occupancy on the same link. This relationship is approximated with certain functional forms and the associated uncertainties are handled explicitly using robust optimization (RO) techniques. The RO allows emissions-related constraints and/or objectives to be reformulated as linear forms under mild conditions. To further reduce the computational cost, we employ a link-based LWR model to describe traffic dynamics with the benefit of fewer (integer) variables and less potential traffic holding. The proposed MILP explicitly captures vehicle spillback, avoids traffic holding, and simultaneously minimizes travel delay and addresses emission-related concerns.

*Keywords:* signal control, emission consideration, robust optimization, mixed integer linear program

## 1. Introduction

Traffic signals tend to be the primary focus of urban traffic control and management strategies since they generally serve as the most frequent and restrictive bottlenecks on urban streets. Over time, the implementation of traffic signal control has evolved greatly: from simple fixed-time plans based on historical data and updated infrequently throughout the day to adaptive control systems that update continuously in response to real-time traffic information. The performance of a particular strategy depends on several factors: the optimization procedure employed to select signal

*Corresponding author
Email addresses: k.han@imperial.ac.uk (Ke Han), hql5143@psu.edu (Hongcheng Liu), gayah@engr.psu.edu (Vikash V. Gayah), tfriesz@psu.edu (Terry L. Friesz), tyy1@engr.psu.edu (Tao Yao)

timings, underlying model used to predict the evolution of traffic dynamics and the objective function considered in the optimization procedure.

Here, we distinguish between two types of optimization procedures: (1) heuristic approaches, such as those developed with feedback control, genetic algorithms and fuzzy logic (Koukol et al., 2015; Zhang et al., 2013) and (2) exact approaches, such as those arising from mathematical control theory and mathematical programming. Although useful for very large and complex optimization problems, heuristic approaches suffer from a failure to provide optimal solutions. Instead, these are more appropriate when exact approaches are computationally intractable. However, exact *mixed integer programs* (MIPs) have been used extensively in the signal control literature and are of particular interest due to their tractability for smaller networks. For example, Improta and Cantarella (1984) formulated and solved the traffic signal control problem for a single road junction as a mixed binary integer program. Lo (1999) and Lo (1999b) formulated the network-level signal control problem as a mixed integer linear program using the *cell transmission model* (CTM) (Daganzo, 1994, 1995). In these papers, time-varying traffic demand patterns were incorporated by adopting dynamic signal timing plans. Such methods were later extended in Lin and Wang (2004) to capture more realistic features of signalized junctions such as the total number of vehicle stops and signal preemption in the presence of emergency vehicles. Building upon this solid foundation that exists in the literature, this MIP approach will be adopted here.

In most works, the model of traffic dynamics is taken as fixed, which is reasonable for models that accurately describe the critical phenomena observed. In this paper, we consider the Lighthill-Whitham-Richards model (Lighthill and Whitham, 1955; Richards, 1956), also known as kinematic wave theory, to describe traffic dynamics on individual links and through signalized junctions. This well-known model is employed as it is one of the most used and trusted traffic flow models currently being used in network optimization procedures today (Aziz and Ukkusuri, 2012; Chitour and Piccoli, 2005; Han et al., 2014b; Lin and Wang, 2004; Liu et al., 2015; Lo, 1999c,b; Zhang et al., 2013). In particular, we employ a *link-based kinematic wave model* (LKWM) proposed in Han et al. (2012) to capture queue dynamics, shock waves and vehicle spillback, while integrating it with signalized junction models. In contrast to the cell-based math programming approaches reviewed above, the link-based approach requires fewer spatial variables and eliminates the problem of *traffic holding* that arises between two adjacent cells (Ziliaskopoulos, 2000) without using additional binary variables (Lo, 1999). In this way, the resulting mathematical program is more computationally efficient than the more traditional cell-based approach.

As for objective functions, the majority of adaptive traffic signal control schemes update signal timings to minimize total vehicular delays. Representatives of such signal-control systems are OPAC (Gartner, 1983), RHODES (Mirchandani and Head, 2000), SCAT (Sims and Dobinson, 1980) and SCOOT (Hunt et al., 1982). Other control strategies seek to minimize delays to a subset of vehicles; e.g., the goal of transit signal priority strategies is to reduce delays for transit vehicles, often to the detriment of those remaining, see Skabardonis (2000). More recently, a transit signal priority strategy was proposed to minimize total person delay, which essentially considers a weighted average of vehicular delay using the passenger occupancies of each vehicle as the weights (Christofa et al., 2013).

Relatively less attention has been given to vehicular emissions in the optimization of traffic signal timings. The earliest study that includes emissions in signal timing optimization appears to be Robertson et al. (1980), but this work relies on macroscopic simulations that do not accurately account for vehicle dynamics at intersections. The efforts that followed either relied on combining detailed emissions models with outputs from microscopic simulations or models (Stevanovic et al., 2009; Li and Shimamoto, 2011; Lin et al., 2010; Lv et al., 2013) or macroscopic emissions models estimated from data (Aziz and Ukkusuri, 2012; Zhang et al., 2013). The former approach is more accurate, but relies on computationally intensive simulation-based optimization methods. The latter is useful but as pointed out by a survey paper (Szeto et al., 2012) and the literature therein, the environmental considerations typically result in highly nonlinear and nonconvex constraints and objective functions in the mathematical programming formulation, which also imposes tremendous computational burdens. As a result, heuristic methods, such as one found in Ferrari (1995); Zhang et al. (2013), have been used to solve these types of problems. Classical methods such as the inner penalty technique (Yang and Bell, 1997) and augmented Lagrangian multiplier technique (Yang et al., 2010) have also been used, but well-defined exact approaches that account for these non-linear and (potentially) stochastic relationships currently do not exist.

This paper presents a novel approach to circumvent the aforementioned computational challenges by combining traditional objective functions (i.e., minimizing vehicular delays or maximizing vehicle throughput) with emissions considerations. The latter are incorporated using constraints in the optimization procedure (e.g., maximizing through-

put subject to some emissions standard that must be met) and these are reformulated as linear functions through the use of numerical experimentation and robust optimization. This method is made possible by leveraging observed relationships between aggregated emissions rates and vehicles occupancies on a link that arise when certain detailed emission models are employed (e.g., see Shabihkhani and Gonzales (2013)). Such empirical observations are supported by extensive numerical simulations, as we shall demonstrate below. Detailed description of the simulation and synthetic data is presented in Section 3.

Unfortunately, despite the strong correlation between the aggregated emission rate and certain macroscopic traffic quantities (e.g. link occupancy), there are non-negligible errors associated with such approximation. Of course, errors and perturbations to a deterministic model can render an optimal solution in the ideal case suboptimal in implementation. A natural approach to capture uncertainty is by assuming that unknown parameters follow certain probability distributions and by employing the notions and methodologies in stochastic programming. However, such an approach has two main limitations: (1) exact knowledge of error distributions is often difficult to acquire, and (2) stochastic programming is recognized as highly intractable to solve even with linear objective function and linear constraint functions. In view of these challenges, we propose to handle uncertainty in the perspective of robust optimization.

A robust optimization is a distribution-free uncertainty set approach that seeks to minimize the worst-case cost and/or to remain feasible in the worst scenario. Compared to stochastic programming, robust optimization makes no assumption on the underlying distribution of uncertain parameters. Moreover, it has been shown to work as a powerful approximation to stochastic programming and even probabilistic models with significantly reduced computational cost (Ben-Tal and Nemirovski, 1998, 1999, 2000; Bertsimas et al., 2011a,c; Bandi and Bertismas, 2012; Rikun, 2011). Although solutions to robust optimization problems can be relatively conservative, the conservatism is adjustable with the flexibility of choosing uncertainty sets (Bertsimas and Sim, 2004). A comprehensive review of robust optimization is provided by Bertsimas et al. (2011a).

Due to the nonlinear and nonconvex nature of emission-related constraints and/or objectives, signal optimization problems with emission considerations are very difficult to solve when formulated as mathematical programs. This paper proposes a practical and effective way to reformulate this problem by invoking a robust optimization approach based on the macroscopic emission models. Through numerical simulations we uncover well-defined macroscopic relationships between the link aggregate emission rate and the link occupancy, which are then utilized to reformulate emission-rated constraints/objectives into mathematically tractable forms. We show that a fair general class of emission-related constraints and objectives can be re-formulated as linear constraints with the utilization of dual variables. Effectively, the signal optimization problem with emission considerations are formulated as mixed integer linear programs (MILPs). These MILPs not only capture realistic traffic dynamics that exist on signalized networks such as shock waves and car spillback, but also address nonlinear and nonconvex emissions constraints/objectives in a mathematically tractable way. Moreover, they can be solved by commercial solvers fairly efficiently in a nearly mathematically tractable way (Bertsimas et al., 2011b). The proposed solution method is tested using a synthetic experiment to demonstrate its performance.

The methodological framework proposed in this paper is potentially transferrable to traffic control problems formulated as mathematical programs, in which vehicle emission, fuel consumption, or safety are within the purview of the traffic operator. For example, vehicle fuel consumption can be modeled in a similar way as emissions based on various operational modes of a moving car such as cruise, acceleration, deceleration, and idle (Barth et al., 1996). On the other hand, traffic risk index has been derived based on statistical analysis of historical information on accidents and macroscopic traffic data including traffic flow, occupancy and speed (Haj Salem et al., 2006). These problems may be similarly formulated as MILPs if the underlying macroscopic relationship can be approximated by piecewise affine functions.

The rest of this paper is organized as follows. Section 2 presents a mixed integer linear programming formulation of network signal optimization problems, without any emission considerations. In Section 3, we present details of the numerical simulations that uncover the macroscopic relationships between the aggregated emission rate and the link occupancy. In Section 4 we utilize the findings made in Section 3 to systematically and explicitly derive the robust counterpart of the LWR-E problem based on fairly general assumptions made on the macroscopic relationship. Section 5 discusses two generalizations of the LWR-E problem. Section 6 presents a numerical study of the proposed formulation. Finally, Section 7 provides some concluding remarks.

## 2. MILP approach for signal optimization

This section presents the mixed integer linear program (MILP) formulation of the dynamic traffic signal control problem without any consideration for emission, while Sections 4 and 5 will address emission-related constraints/objectives in detail. As we mentioned in the introduction, a link-based LWR model will be considered, the derivation of which employs the variational method (Daganzo, 2005). For the conciseness of our presentation, we only recap the key results below and refer the reader to Han et al. (2012) for a detailed derivation and analysis.

The LWR model is based on the following kinematic wave equation:

$$\partial_t \rho(t, x) + \partial_x f(\rho(t, x)) = 0 \tag{2.1}$$

where $\rho(t, x)$ denotes the vehicle density in a spatial-temporal domain; $f(\cdot)$ is the fundamental diagram that describes the macroscopic relationship between vehicle density and flow on the link. Throughout this paper, $f(\cdot)$ is assumed to be triangular of the following form [1]:

$$f(\rho) = \begin{cases} v\rho & \rho \in [0, \rho^c] \\ -w(\rho - \rho^{jam}) & \rho \in (\rho^c, \rho^{jam}] \end{cases} \tag{2.2}$$

where $\rho$ denotes vehicle density; $v$ and $w$ are respectively the speeds of the forward- and backward-propagating kinematic waves; $\rho^c$ is the critical density at which the flow is maximized; and $\rho^{jam}$ denotes the jam density. Moreover, we let $C$ be the flow capacity and $L$ be the length of the link. We allow all these quantities and variables to depend on a specific link $I_i$, and will use subscript '$i$' to indicate such a dependence.

We consider a given link expressed as a spatial interval $[a, b]$, and ignore for now the subscript '$i$' for notation convenience. We define a binary variable $\bar{r}(t)$, which indicates the traffic state at the entrance of the link $x = a$: $\bar{r}(t) = 0$ if traffic is free flow and $\bar{r}(t) = 1$ if traffic is congested. A similar notation $\hat{r}(t)$ is used for the exit of the link. We also define the link inflow $\bar{q}(t)$ and the link exit flow $\hat{q}(t)$. The key results of the link-based kinematic wave model, derived from the variational theory, are as follows.

$$\bar{r}(t) = \begin{cases} 1, & \text{if} \quad \int_0^t \bar{q}(\tau)\,d\tau = \int_0^{t-\frac{L}{w}} \hat{q}(\tau)\,d\tau + \rho^{jam}L \\ 0, & \text{if} \quad \int_0^t \bar{q}(\tau)\,d\tau < \int_0^{t-\frac{L}{w}} \hat{q}(\tau)\,d\tau + \rho^{jam}L \end{cases} \tag{2.3}$$

$$\hat{r}(t) = \begin{cases} 0, & \text{if} \quad \int_0^{t-\frac{L}{v}} \bar{q}(\tau)\,d\tau = \int_0^t \hat{q}(\tau)\,d\tau \\ 1, & \text{if} \quad \int_0^{t-\frac{L}{v}} \bar{q}(\tau)\,d\tau > \int_0^t \hat{q}(\tau)\,d\tau \end{cases} \tag{2.4}$$

In addition, we introduce the notions of *demand* and *supply* of a link (Lebacque and Khoshyaran, 1999) which, under the assumption of a triangular fundamental diagram, reduce to the following:

$$D(t) = \begin{cases} C & \text{if } \hat{r}(t) = 1 \\ \bar{q}\left(t - \frac{L}{v}\right) & \text{if } \hat{r}(t) = 0 \end{cases} \qquad S(t) = \begin{cases} C & \text{if } \bar{r}(t) = 0 \\ \hat{q}\left(t - \frac{L}{w}\right) & \text{if } \bar{r}(t) = 1 \end{cases} \tag{2.5}$$

### 2.1. Discrete-time formulation of the link dynamic

Let us introduce some key discrete-time notations employed in this paper, where the subscript '$i$' indicates association with link $I_i$, and the superscript '$k$' indicates the $k$-th time step.

| | |
|---|---|
| $\bar{q}_i^k$ | The flow at which vehicles enter link $I_i$ |
| $\hat{q}_i^k$ | The flow at which vehicles exit link $I_i$ |

---

[1] The triangular FD is the key to a much simplified variational representation of the solution and to the link-based LWR model considered in this paper.

| | |
|---|---|
| $\bar{r}_i^k$ | The binary variable indicating the traffic state at the entrance of $I_i$ |
| $\hat{r}_i^k$ | The binary variable indicating the traffic state at the exit of $I_i$ |
| $S_i^k$ | The supply of link $I_i$ |
| $D_i^k$ | The demand of link $I_i$ |
| $u_i^k$ | The binary signal control variable for link $I_i$ |

For a fixed time step size $\delta t$, we define $\Delta_i^f \doteq \left[\frac{L_i}{v_i \delta t}\right]$, $\Delta_i^b \doteq \left[\frac{L_i}{w_i \delta t}\right]$, where $[x]$ rounds the real number $x$ to the nearest integer. [2] We are now ready to state the discrete versions of (2.3) and (2.4) as follows.

$$\begin{cases} \delta t \sum_{k=1}^{l-\Delta_i^b} \hat{q}_i^k - \delta t \sum_{k=1}^{l} \bar{q}_i^k + \rho_i^{jam} L_i \ \leq \ \mathcal{M}(1 - \bar{r}_i^l) + \varepsilon \\ \delta t \sum_{k=1}^{l-\Delta_i^b} \hat{q}_i^k - \delta t \sum_{k=1}^{l} \bar{q}_i^k + \rho_i^{jam} L_i \ > \ -\mathcal{M}\bar{r}_i^l + \varepsilon \end{cases} \qquad \forall i, \ \forall l \qquad (2.6)$$

$$\begin{cases} \delta t \sum_{k=1}^{l-\Delta_i^f} \bar{q}_i^k - \delta t \sum_{k=1}^{l} \hat{q}_i^k \ \leq \ \mathcal{M}\hat{r}_i^l + \varepsilon \\ \delta t \sum_{k=1}^{l-\Delta_i^f} \bar{q}_i^k - \delta t \sum_{k=1}^{l} \hat{q}_i^k \ > \ \mathcal{M}(\hat{r}_i^l - 1) + \varepsilon \end{cases} \qquad \forall i, \ \forall l \qquad (2.7)$$

where $\rho_i^{jam}$ and $L_i$ denote respectively the jam density and length of link $I_i$. $\mathcal{M} > 0$ is a large constant, and $\varepsilon > 0$ is a small constant serving as a cut-off threshold; see the end of Section 2 for more discussion on these constants. Moreover, the demand $D_i^k$ and the supply $S_i^k$, whose continuous-time expressions are given by (2.5), are determined via the following linear inequalities, where $C_i$ denotes the flow capacity of link $I_i$:

$$\begin{cases} C_i + \mathcal{M}(\hat{r}_i^k - 1) \ \leq \ D_i^k \ \leq \ C_i \\ \bar{q}_i^{k-\Delta_i^f} - \mathcal{M}\hat{r}_i^k \ \leq \ D_i^k \ \leq \ \bar{q}_i^{k-\Delta_i^f} + \mathcal{M}\hat{r}_i^k \end{cases} \qquad \forall i, \ \forall k \qquad (2.8)$$

$$\begin{cases} C_i - \mathcal{M}\bar{r}_i^k \ \leq \ S_i^k \ \leq \ C_i \\ \hat{q}_i^{k-\Delta_i^b} + \mathcal{M}(\bar{r}_i^k - 1) \ \leq \ S_i^k \ \leq \ \hat{q}_i^{k-\Delta_i^b} - \mathcal{M}(\bar{r}_i^k - 1) \end{cases} \qquad \forall i, \ \forall k \qquad (2.9)$$

### 2.2. Dynamics at signalized junctions

In general, signalized junction models vary according to detailed intersection geometry and phasing schemes. For simplicity yet without loss of generality, we consider in this paper the simple junction shown in Figure 1 while referring the reader to Han and Gayah (2015) for an elaborated treatment of detailed junction layout, vehicle movements and multiple signal phases.

This junction has two incoming links ($I_1$, $I_2$) and two outgoing links ($I_3$, $I_4$). The vehicle turning percentages, $\alpha_{1,3}$, $\alpha_{1,4}$, $\alpha_{2,3}$, $\alpha_{2,4}$ as shown in Figure 1, are assumed to be known a priori, and can be usually estimated through historical turn-by-turn vehicle counts. The discrete-time junction dynamic may be written as

$$\hat{q}_i^k \ = \ \min\left\{ D_i^k, \ u_i^k \cdot \min\left\{ \frac{S_3^k}{\alpha_{i,3}}, \ \frac{S_4^k}{\alpha_{i,4}} \right\} \right\} \qquad i = 1, 2, \quad \forall k$$

$$\bar{q}_j^k \ = \ \alpha_{1,j} \hat{q}_1^k + \alpha_{2,j} \hat{q}_2^k \qquad j = 3, 4, \quad \forall k$$

---

[2] Rounding these quantities to integers lead to certain numerical errors due to time discretization. The impact of such an approximation (rounding effect) on the solution quality has been investigated in Han et al. (2014a).
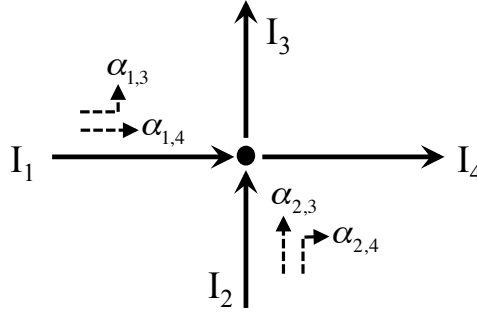
Figure 1: Two signalized junctions

**Remark 2.1.** *In the above expression, if certain vehicle turning percentage, say $\alpha_{i,3}$, is zero, then the term $\frac{S_3^k}{\alpha_{i,3}}$ is infinity and will be dropped from the "min" operator. This corresponds to the situation where vehicles discharged from $I_i$ do not enter the downstream link $I_3$, and thus $I_3$ is effectively removed from the junction as far as vehicles from $I_i$ are concerned. Thus, for what follows we always assume positive vehicle turning percentages. Moreover, although we assume here that these percentages are constants, it is a trivial extension to allow them to depend on time ('k').*

For $i = 1, 2$, let us define $\zeta_i^k = \min\left\{D_i^k, \frac{S_3^k}{\alpha_{i,3}}, \frac{S_4^k}{\alpha_{i,4}}\right\}$; then we have $\hat{q}_i^k = u_i^k \cdot \zeta_i^k$, which may be expressed as linear constraints:

$$\begin{cases} 0 \le \hat{q}_i^k \le \mathcal{M} u_i^k \\ \zeta_i^k + \mathcal{M}(u_i^k - 1) \le \hat{q}_i^k \le \zeta_i^k \end{cases} \qquad i = 1, 2, \quad \forall k \tag{2.10}$$

Moreover, $\zeta_i^k$, which is the minimum of three endogenous variables, may be expressed via linear constraints using two additional binary variables, $\xi_i^k$ and $\eta_i^k \in \{0, 1\}$:

$$\begin{cases} D_i^k - \mathcal{M}\eta_i^k \le \zeta_i^k \le D_i^k \\ \frac{S_3^k}{\alpha_{i,3}} - \mathcal{M}\xi_i^k - \mathcal{M}(1-\eta_i^k) \le \zeta_i^k \le \frac{S_3^k}{\alpha_{i,3}} \\ \frac{S_4^k}{\alpha_{i,4}} - \mathcal{M}(1-\xi_i^k) - \mathcal{M}(1-\eta_i^k) \le \zeta_i^k \le \frac{S_4^k}{\alpha_{i,4}} \end{cases} \qquad i = 1, 2 \qquad \forall k \tag{2.11}$$

It is easy to check that $\zeta_i^k = D_i^k$ when $\eta_i^k = 0$; $\zeta_i^k = \frac{S_3^k}{\alpha_{i,3}}$ when $\eta_i^k = 1$ and $\xi_i^k = 0$; $\zeta_i^k = \frac{S_4^k}{\alpha_{i,4}}$ when $\eta_i^k = 1$ and $\xi_i^k = 1$. Finally, to prevent conflicting traffic streams to be discharged at the same time, we stipulate that

$$u_1^k + u_2^k = 1 \qquad \forall k \tag{2.12}$$

Regarding the objective function one has a lot of flexibility in selecting its form as long as linearity is maintained. The following linear form is selected in this paper.

$$\max \sum_{k=1}^{M} \frac{1}{k+1} \sum_{I_i \in \mathcal{I}} \hat{q}_i^k \tag{2.13}$$

where $M$ is the total number of time intervals, and $\mathcal{I}$ is a prescribed set of links. For example, $\mathcal{I}$ may be selected to be the set of outgoing links of the network of interest; or $\mathcal{I}$ may be specified as the set of links that one wishes to prioritize in a congested network. Choosing such an objective function ensures that the throughputs on these links are maximized at any instance of time. Again, we emphasize that any type of linear objective function can be selected, depending on the specific application, without affecting the MILP formulation. In summary, the proposed MILP consists of the objective function (2.13) and constraints (2.6) through (2.12).

**Remark 2.2.** *No explicit constraints besides (2.12) are made on the signal control binary variables, which allows the signals to operate using splits and cycle lengths that change dynamically during the course of the control period. In this way, the signal splits and cycle lengths will likely vary with time as traffic flow patterns change. Note that this is the most flexible strategy, but might not be the most realistic in some networks. The methodology can be easily extended to more restrictive but realistic scenarios (e.g., fixed-cycle-and-split, fixed-cycle-dynamic-split, dynamic-cycle-fixed-split, etc.) through the introduction of additional constraints on signal timing parameters. This will not affect the main formulation for minimizing delays subject to emissions constraints.*

In the MILP formulation, the constant $\mathcal{M}$ can be chosen to be

$$\mathcal{M} \;\geq\; \max\left\{\max\{1,\,T\}\cdot\bar{C},\,\frac{\bar{C}}{\bar{\alpha}}\right\}$$

where $T$ is the length of the time horizon, $\bar{C}$ is the maximum link flow capacity in the network, $\bar{\alpha}$ is the smallest vehicle turning percentage in the network (notice that all turning percentages are assumed positive). The cut-off threshold $\varepsilon$ in (2.6) and (2.7) should be no greater than $\delta t \cdot \min_i C_i$.

## 3. Macroscopic relationship between the emission rate and traffic quantities

As mentioned earlier in Section 1, the key ingredient of the proposed MILP formulation for the LWR-E problem is the existence of a macroscopic relationship between the link's aggregate emission rate (AER) and its vehicle occupancy. In this section we will investigate such a relationship through analytical computations. This requires the modeling of vehicle movements within a link subject to signal controls, and a vehicle emission model that calculates the AER in a way consistent with the vehicle flow dynamics. More specifically, we employ the LWR model with a triangular fundamental diagram to predict and describe the evolution of vehicle density (see Section 2); on the emission side a modal emission model is considered and detailed below in Section 3.2.

### 3.1. Simulation setup

The hypothesized macroscopic relationship is investigated through a battery of simulations that employ the aforementioned LWR model and emission model. The simulations are conducted as follows. We consider an arbitrary link with given length and triangular fundamental diagram. In order to account for various levels of congestion, we randomly generate the demand and supply profiles at the upstream and downstream ends of the link, respectively. The values of the demand and supply are uniformly distributed between zero and the link flow capacity. Moreover, in order to incorporate the effect of the signal timing, we randomly generate sequences of green and red phases to control the link's discharge flow. Such variability in the signal controls is crucial in the simulation since the resulting macroscopic relationship (if any) will be used across all possible scenarios involving different signal control parameters. With given signal control and upstream demand and downstream supply profiles, we solve the LWR PDE (2.1) by applying the variational method (Han et al., 2012). Alternatively, one may solve the same PDE using the Godunov scheme (Godunov, 1959) or the cell transmission model (Daganzo, 1994). Once vehicle densities are available in space and time, we then apply a specific emission model to compute the total emission rate and the link occupancy $N(t)$:

$$N(t) \;=\; \int_a^b \rho(t,\,x)dx \qquad \text{or} \qquad N(t) \;=\; \int_0^t \bar{q}(\tau)d\tau - \int_0^t \hat{q}(\tau)d\tau \tag{3.14}$$

where the link of interest is expressed as the spatial interval $[a,\,b]$; $\rho(t,\,x)$ is the solution of the LWR PDE (2.1), and $\bar{q}(\cdot)$ and $\hat{q}(\cdot)$ are the link inflow and exit flow, respectively.

In the next subsection, we will present in detail the modal emission model and the resulting macroscopic relationships obtained from the simulation. The following link parameters are employed in our simulation.

$$v = 40/3 \text{ meter/s}, \quad w = 40/9 \text{ meter/s}, \quad \rho^{jam} = 0.4 \text{ vehicle/meter}, \quad C = 4/3 \text{ vehicle/second}$$

where $v$ and $w$ are the forward and backward wave speeds respectively, $\rho^{jam}$ denotes the jam density, and $C$ is the flow capacity. The link length is set to be $L = 400$ m. Note should be taken on the following fact: the macroscopic

relationship, as well as the uncertainty set calibrated for the robust optimization presented later, depend on link-specific parameters such as the fundamental diagram and link length. However, the most relevant fact that we rely upon here is that the relationship or the uncertainty region holds for a range of endogenous factors; this includes: vehicle arrival and discharge rate for the link (both of which are accounted for within our simulation) and most importantly the signal timing at the intersection. The latter fact is especially useful as it allows us to use these relationships to derive optimal signal timing plans that both minimize delay and account for emissions within the network or on individual links.

It should be noted that although this paper considers only one specific emission model concerning hydrocarbon due to space limitation, the proposed analytical framework can potentially accommodate a wider range of emission models that are similarly based on vehicle speed, acceleration and deceleration.

### 3.2. The modal emission model

We consider a modal emission model, which is based on the operational modes of a moving vehicle, including idle, steady-state cruise, acceleration and deceleration. This model relies on vehicle dynamics estimated from the LWR model. More specifically, let $\rho(t, x)$ be the solution of the LWR equation (2.1) on the link of interest; the vehicle speed $v(t, x)$ is computed as

$$v(t, x) = f(\rho(t, x))/\rho(t, x) \tag{3.15}$$

The acceleration/deceleration, $a(t, x)$, viewed as the derivative of the speed along the trajectories of moving vehicle, is computed as the material derivative in the Eulerian coordinates:

$$a(t, x) = \frac{D}{Dt}v(t, x) = \partial_t v(t, x) + v(t, x) \cdot \partial_x v(t, x)$$

Since the quantities $\rho(t, x)$ and $v(t, x)$ are in general non-differentiable, they are approximated in discrete time using finite differences. Let $\{t_i\}$ and $\{x_j\}$ be discrete temporal and spatial grid points, then we have

$$a(t_i, x_j) = \frac{v(t_{i+1}, x_j) - v(t_{i-1}, x_j)}{2\delta t} + v(t_i, x_j) \cdot \frac{v(t_i, x_{j+1}) - v(t_i, x_{j-1})}{2\delta x} \tag{3.16}$$

where $\delta t$ and $\delta x$ are the time and spatial steps, respectively.

Following the power-demand emission model proposed by Post et al. (1984), the overall instantaneous total power demand $Z$ (in kilowatt) for a vehicle with mass $m$ in (kilogram) is given by

$$Z = (0.04\,v + 0.5 \times 10^{-3}v^2 + 10.8 \times 10^{-6}v^3) + \frac{m}{1000}\frac{v}{3.6}\left(\frac{a}{3.6} + 9.81\sin\theta\right) \tag{3.17}$$

where the above quantity is in kilowatts and $\theta$ denotes the road grade. The reader is also referred to Barth et al. (1996) for an alternative description of the power demand function based on velocity and acceleration. Post et al. (1984) also propose the following model of hydrocarbon emission rate for vehicles based on field experiments:

$$r(t) = \begin{cases} 52.8 + 4.2Z & Z > 0 \\ 52.8 & Z \leq 0 \end{cases} \tag{3.18}$$

where the emission rate $r(t)$ is in grams/hour, and $Z$ is in kilowatt. Following the previous discussion, we now compute the aggregate emission rate (AER) in discrete time as

$$AER(t_i) = \delta x \sum_j \rho(t_i, x_j)r(t_i, x_j)$$

where $r(t_i, x_j)$ is the emission rate of a vehicle corresponding to the $j$-th cell at time $t_i$, calculated using (3.15)-(3.18). The resulting scatter plot of the link occupancy (LO) vs. the aggregate emission rate (AER) is shown in Figure 2, which is based on 42,000 simulation runs. From this figure we can observe a well-defined macroscopic relationship between these two quantities, although the points are relatively scattered, indicating potential errors associated with such an approximation. Overall, Figure 2 shows a positive correlation between the AER and the LO. This can be intuitively explained using the following two observations. (1) The total emission rate on the link level is in general

expected to show a growing trend with an increased number of vehicles on this link. Although vehicle dynamics do matter in this case and will play a role to some extent, they are not significant enough to overturn this trend, at least not for the emission model considered by this paper. (2) When the link is controlled by a signal light, a higher link occupancy indicates more severe congestion and a longer queue. As a result more vehicle stop-and-go movements are expected to occur and contribute to the total emission amount.
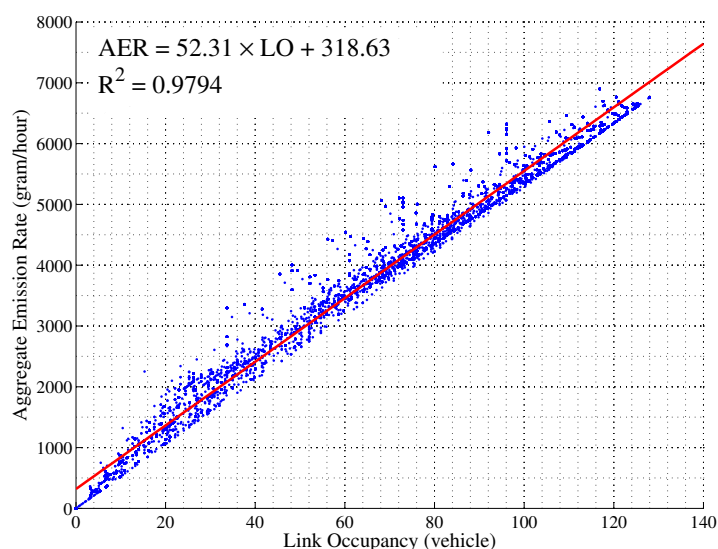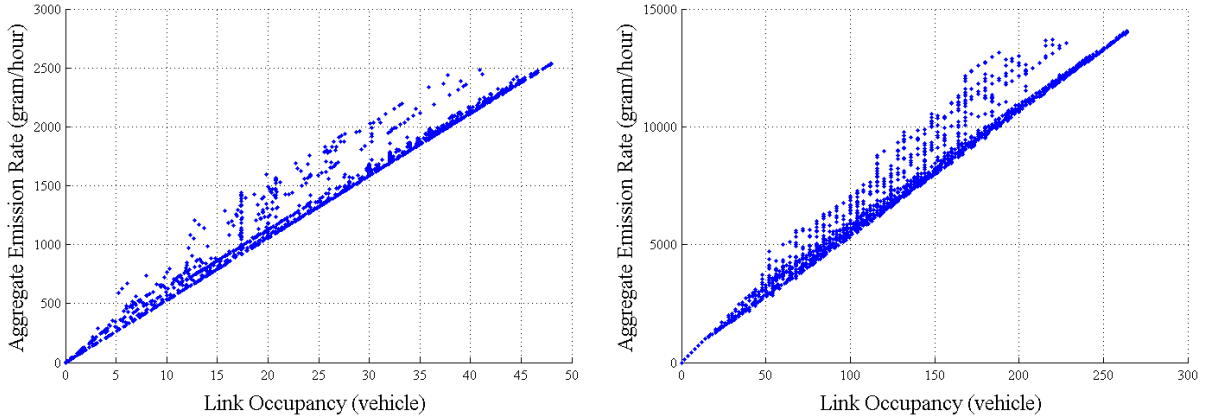


Figure 2: Scatter plot of link occupancy vs. AER.

The observed relationship between LO and AER may be approximated in a number of ways, e.g. using linear, piecewise linear, or piecewise smooth curve fittings. In Figure 2 we show the fitting result using linear approximation, with $R^2 = 0.9794$. We note, however, that this does not mean that the linear fit is the best choice; and other functional forms may be more appropriate. Nevertheless, there is a trade-off between the sophistication/goodness of the curve fitting and the simplicity and computational tractability of the resulting optimization formulation, as we subsequently show in Section 4. Such a trade-off needs to be taken into account when one formulates the uncertainty set based on curve fitting.

Some of the significant deviations from the linear fit, as shown in the figure, are caused by the highly nonlinear vehicle dynamics (i.e. acceleration and deceleration) captured by the LWR model under the control of signal light. More specifically, close to a signalized intersection the shock waves and vehicle stop-and-go movements influenced by the signal timings generate vehicle acceleration/deceleration profiles that significantly contribute to the emission of pollutants. Moreover, the emission amount are related to the actual spatial configuration of vehicle densities and speeds on the link, which cannot be adequately captured by the link occupancy alone.

In order to show that similar macroscopic relationships exist for some other choices of link parameters, say length, we show in Figure 3 two additional simulation results for $L = 200$ m and $L = 800$ m (the result shown in Figure 2 is based on $L = 400$ m). Although similar macroscopic relationships exist in these two additional cases, the uncertainty sets need to be calibrated separately.

The macroscopic relationships depicted in Figure 2 and Figure 3 are quite interesting from a robust optimization perspective. The reasons is that although the points are spread out in certain region, they are sparse in some places while dense in some other regions. An uncertainty region that simply covers all these points may be too conservative; and a more effective approach requires a more refined and sophisticated calibration of the uncertainty set. Detailed calibration of the uncertainty set associated with Figure 2 will be presented in Section 6.2.

9

Figure 3: Scatter plots of link occupancy vs. AER. Left: $L$ = 200 m; right: $L$ = 800 m.

## 4. Signal control with emission constraints

This section provides a general mathematical framework for incorporating emission-related side constraints, while relying on fairly general assumptions on the macroscopic relationship between the aggregate emission rate and the link occupancy. We also present explicit reformulations of the emission side constraints for three special cases, namely, when the relationship is (1) affine; (2) convex piecewise affine; and (3) concave piecewise affine. Notably, all the three reformulations lead to linear constraints, which, when combined with the signal control formulation presented earlier, do not alter the nature of the mixed integer linear program.

We propose a set of emission-related constraints constructed in a data-driven manner, which employs robust optimization techniques to handle prediction errors arising from the macroscopic relationship. In addition, we show that when this relationship is either affine or convex/concave piecewise affine, the resulting emission-related side constraints are still linear.

For an arbitrary link in the network, we let $N(t)$ be the occupancy of this link at time $t$. Here, for notation convenience the subscript '$i$' indicating the link ID is dropped in this section. The time-varying aggregate emission rate (AER) on this link is denoted $AER(t)$. In the following derivation of the emission constraints we assume a general polynomial form relating $N(t)$ to $AER(t)$. As a result, the proposed framework can handle a wide range of macroscopic relationships.

We assume the macroscopic relationship is approximated by a polynomial with degree $\mathcal{L}$, where the approximation is based on regression analysis or other types of curve-fitting techniques:

$$AER(N(t); \boldsymbol{a}) \doteq \sum_{l=0}^{\mathcal{L}} a_l (N(t))^l = \boldsymbol{a}^T \boldsymbol{N}(t) \qquad \forall t \in [0, T] \tag{4.19}$$

Here $\boldsymbol{a} \doteq (a_0, a_1, \ldots, a_{\mathcal{L}})^T \in \mathbb{R}^{\mathcal{L}}$ is a vector of coefficients in the polynomial, and $\boldsymbol{N}(t) \doteq \left((N(t))^0, \ldots, (N(t))^{\mathcal{L}}\right)^T$ is a vector-valued function of time. With these notations, we consider the first type of side constraint, which stipulates that the total emission amount on the subject link is constrained by a prescribed level $E$.

$$\int_0^T AER(N(t); \boldsymbol{a}) \, dt \leq E \tag{4.20}$$

Notice that, for now, we are simply replacing the macroscopic relationship with a polynomial function without any consideration of errors associated with this approximation. Later in the next subsection, we will take into account approximation errors and handle them using robust optimization techniques.

Side constraints of the form (4.20) can be easily generalized to address other types of environmental considerations including the following.

- The total emission on a subset of the network is bounded by some given value.

- The total emission on a subset of the network is minimized.

- The differences among the total emissions on a subset of links are bounded and/or minimized.

The first case is a trivial extension of the constraint (4.20), and will not be elaborated in this paper. The second case can be easily handled by invoking the "epigraph reformulation". The third case ensures that no link (or nearby environment) suffers much more than other parts of the network in terms of pollutant emission. This issue is identified by Benedek and Rilett (1998) as *environmental equity*. Detailed treatment of the last two environmental considerations will be presented in Section 5.

### 4.1. Emissions side constraints: a general formulation based on robust optimization

Inequality (4.20) involves a polynomial approximation of the relationship between the link occupancy and the aggregate emission rate. In order to ensure that the emission constraint is still satisfied in the presence of approximation errors, we consider the following constraint instead:

$$\int_0^T AER(N(t); \boldsymbol{a}(t)) \, dt \; \leq \; E \qquad \forall \boldsymbol{a}(\cdot) \in \boldsymbol{\eta}_a \tag{4.21}$$

where the coefficients in the polynomial are allowed to vary over time; that is, $\boldsymbol{a}(t) = (a_l(t) : 0 \leq l \leq \mathcal{L})$, and

$$AER(N(t); \boldsymbol{a}(t)) \; = \; \sum_{l=0}^{\mathcal{L}} a_l(t)(N(t))^l \tag{4.22}$$

Moreover, $\boldsymbol{\eta}_a$ is specified as a budget-like uncertainty set, which is similar to what is proposed by Atamtürk and Zhang (2007):

$$\boldsymbol{\eta}_a \; = \; \left\{ \boldsymbol{a}(\cdot) : \; L_l \leq a_l(t) \leq U_l, \; \forall t \in [0, T], \; \forall 0 \leq l \leq \mathcal{L}; \; \sum_{l=1}^{\mathcal{L}} \int_0^T a_l(t) \, dt \; \leq \; \frac{T \sum_{l=1}^{\mathcal{L}} U_l}{\sigma} \right\} \tag{4.23}$$

where $L_l$ and $U_l$, $0 \leq l \leq \mathcal{L}$, are respectively the lower and upper bounds of the corresponding coefficient; $\sigma$, which is used to adjust the conservatism of the robust optimization, satisfies

$$\sigma \in \left[ 1, \; \frac{\sum_{l=0}^{\mathcal{L}} U_l}{\sum_{l=0}^{\mathcal{L}} L_l} \right] \tag{4.24}$$

In (4.24), the upper bound on $\sigma$ ensures that the uncertainty set expressed in (4.23) is nonempty. This can be easily seen by manipulating the last inequality of (4.23):

$$\sum_{l=1}^{\mathcal{L}} T \cdot L_l \; \leq \; \sum_{l=1}^{\mathcal{L}} \int_0^T a_l(t) \, dt \; \leq \; \frac{T \sum_{l=1}^{\mathcal{L}} U_l}{\sigma}$$

Thus $\sigma$ must not exceed $\sum_{l=1}^{\mathcal{L}} U_l / \sum_{l=1}^{\mathcal{L}} L_l$. The constraints described by (4.21)-(4.24) correspond to the notion of robust optimization in the sense that the emission amount on the link of interest is bounded from above with any possible realization of the uncertain parameter $\boldsymbol{a}(t)$.

**Remark 4.1.** *By writing (4.23) we have implicitly assumed that coefficients associated with the zeroth-order term are not correlated with the other coefficients (the summation starts from $l = 1$ instead of $l = 0$). This implicit assumption will result in a more risk-averse formulation. However, this assumption is easy to relax, and the consequent generalization of our formulation is a trivial extension.*

The uncertainty set in (4.23) consists of two types of constraints: (1) a box constraint prescribing the upper and lower bounds of the uncertain parameters; and (2) a constraint that stipulates the sums of uncertain coefficients to be bounded from above (the last constraint in (4.23)). With just the first type of constraints, the RO will generate the most conservative solution by predicting that all the uncertain parameters are realized at the extreme case against the decision maker. However, such a worst case occurs only with a very low probability in a realistic system, and this

conservative solution is most likely to compromise the performance of the resulting system. Use of the second type of constraint can reduce the conservatism by excluding some of the extreme and rare cases.

The second type of constraint is made flexible by adjusting the value of $\sigma$. Specifically, a higher value of $\sigma$ implies a smaller uncertainty set, which results in solutions that are more risk-prone (less conservative). In the most conservative case, i.e. $\sigma = 1$, the last constraint in (4.23) is out of effect. Bandi and Bertismas (2012) and Bertsimas et al. (2014) provide data-driven approaches based on probability theory and statistical tests to determine the parameterization of the uncertainty sets in accordance with the observed data. Those approaches provide theoretical guarantee for the satisfaction of constraints in a probabilistic sense. We would like to further remark that the second type of constraint may also capture potential correlations among $a_l(t)$, $1 \leq l \leq \mathcal{L}$. However, due to space limitation this aspect of research will not be elaborated in this paper.

### 4.2. Time-discretization and explicit reformulation

Here, we adopt the same notation convention as in Section 2.1 by using superscript '$k$' to indicate association with the $k$-th discrete time step, $1 \leq k \leq M$. The constraint (4.21) can be time-discretized into the following form, where $\delta t$ denotes the time step size.

$$\sum_{k=1}^{M} \sum_{l=0}^{\mathcal{L}} a_l^k (N^k)^l \delta t \leq E \qquad \forall \hat{\boldsymbol{a}} \in \hat{\boldsymbol{\eta}}_a \tag{4.25}$$

where $a_l^k$ corresponds to $a_l(t_k)$ at the $k$-th time step, and $\hat{\boldsymbol{a}} \doteq (a_l^k : 0 \leq l \leq \mathcal{L}, \ 1 \leq k \leq M)$. The discrete-time version of the uncertainty set is

$$\hat{\boldsymbol{\eta}}_a = \left\{ \hat{\boldsymbol{a}} : L_l \leq a_l^k \leq U_l, \ \forall 1 \leq k \leq M, \ \forall 0 \leq l \leq \mathcal{L}; \ \sum_{k=1}^{M} \sum_{l=1}^{\mathcal{L}} a_l^k \delta t \leq \frac{T \sum_{l=1}^{\mathcal{L}} U_l}{\sigma} \right\} \tag{4.26}$$

The constraint (4.25) is in fact a semi-infinite constraint with an infinite index set $\hat{\boldsymbol{\eta}}_a$, which means that the inequality needs to be satisfied for infinitely many $\hat{\boldsymbol{a}}$. This makes it not directly computable. The following theorem provides a computable reformulation of (4.25) using dual variables.

**Theorem 4.2. (Robust reformulation with polynomial relationship)** *Let constants $L_l < U_l$, $0 \leq l \leq \mathcal{L}$ and $\sigma$ be given. If $\hat{\boldsymbol{\eta}}_a$ has nonempty interior, the semi-infinite constraint (4.25) is equivalent to the following set of constraints:*

$$\sum_{l=1}^{\mathcal{L}} \sum_{k=1}^{M} \beta_l^k U_l - \sum_{l=1}^{\mathcal{L}} \sum_{k=1}^{M} \gamma_l^k L_l + \frac{M \sum_{l=1}^{\mathcal{L}} U_l}{\sigma} \theta + M U_0 \delta t \leq E \tag{4.27}$$

$$s.t. \quad \beta_l^k - \gamma_l^k + \theta = (N^k)^l \delta t \qquad \forall \ 1 \leq l \leq \mathcal{L}, \qquad 1 \leq k \leq M \tag{4.28}$$

$$\beta_l^k, \ \gamma_l^k, \ \theta \geq 0 \qquad \forall \ 1 \leq l \leq \mathcal{L}, \qquad 1 \leq k \leq M \tag{4.29}$$

*where $\beta_l^k$, $\gamma_l^k$ and $\theta$ are dual variables, $1 \leq l \leq \mathcal{L}$, $1 \leq k \leq M$.*

*Proof.* The constraint (4.25) can be trivially rewritten as:

$$\max_{\hat{\boldsymbol{a}} \in \hat{\boldsymbol{\eta}}_a} \sum_{k=1}^{M} \sum_{l=1}^{\mathcal{L}} a_l^k (N^k)^l \delta t + \max_{\hat{\boldsymbol{a}} \in \hat{\boldsymbol{\eta}}_a} \sum_{k=1}^{M} a_0^k (N^k)^0 \delta t \leq E, \tag{4.30}$$

which is equivalent to

$$\max_{\hat{\boldsymbol{a}} \in \hat{\boldsymbol{\eta}}_a} \sum_{k=1}^{M} \sum_{l=1}^{\mathcal{L}} a_l^k (N^k)^l \delta t \leq E - \sum_{k=1}^{M} U_0 \delta t \tag{4.31}$$

The left hand side of (4.31) involves solving a parametric problem of the form:

$$\max \sum_{k=1}^{M} \sum_{l=1}^{\mathcal{L}} a_l^k (N^k)^l \delta t \tag{4.32}$$

$$s.t. \quad a_l^k \leq U_l \quad \forall 0 \leq l \leq \mathcal{L}, \ 1 \leq k \leq M \tag{4.33}$$

$$a_l^k \geq L_l \quad \forall 0 \leq l \leq \mathcal{L}, \ 1 \leq k \leq M \tag{4.34}$$

$$\sum_{k=1}^{M} \sum_{l=1}^{\mathcal{L}} a_l^k \delta t \leq \frac{T \sum_{l=1}^{\mathcal{L}} U_l}{\sigma} \tag{4.35}$$

where we treat each $N^k$ as a constant parameter, $1 \leq k \leq M$. The program (4.32)-(4.35) corresponds to the following dual problem:

$$\min \sum_{l=1}^{\mathcal{L}} \sum_{k=1}^{M} \beta_l^k U_l - \sum_{l=1}^{\mathcal{L}} \sum_{k=1}^{M} \gamma_l^k L_l + \frac{T \sum_{l=1}^{\mathcal{L}} U_l}{\sigma \delta t} \theta \tag{4.36}$$

$$s.t. \quad \beta_l^k - \gamma_l^k + \theta = (N^k)^l \delta t \quad \forall \ 1 \leq l \leq \mathcal{L}, \quad 1 \leq k \leq M \tag{4.37}$$

$$\beta_l^k, \gamma_l^k, \theta \geq 0 \quad \forall \ 1 \leq l \leq \mathcal{L}, \quad 1 \leq k \leq M \tag{4.38}$$

where $\beta_l^k$, $\gamma_l^k$ and $\theta$ are dual variables corresponding to constraints (4.33), (4.34) and (4.35), respectively. Under the assumption that $\hat{\eta}_a$ has nonempty interior, by noticing the compactness of $\hat{\eta}_a$, the primal program (4.32)-(4.35) and the dual program (4.36)-(4.38) have finite solutions and zero duality gap. Moreover, by duality, the objective value of any feasible solution of the dual problem (4.36)-(4.38) provides an upper bound of the primal problem (4.32)-(4.35). Therefore, if there exist $\beta_l^k$, $\gamma_l^k$ and $\theta$ such that (4.37) and (4.38) are satisfied, then if $N^k$ satisfies (4.27), (4.31) is also satisfied. On the other hand, if there exists $N^k$ that satisfies (4.31), then the objective value of the optimal solution to the parametric problem (4.32)-(4.35) is bounded above and thus there exists $\beta_l^k$, $\gamma_l^k$ and $\theta$ such that (4.37) and (4.38) are satisfied. This shows the desired equivalence result. □

**Remark 4.3.** *Theorem 4.2 is based on the discussions in Ben-Tal and Nemirovski (1999) and Bertsimas et al. (2011a). With this reformulation, the original semi-infinite constraint is now computable with standard nonlinear programming techniques. We will call the reformulation with constraints* (4.27)-(4.29) *the robust counterpart of the original robust problem* (4.25).

The nonlinearity of constraints (4.27)-(4.29) is caused by the powers of $N^k$ appearing in the right hand side of (4.28), which obviously stem from the polynomial approximation (4.22) with degree greater than one. It is not difficult to see that linearity will be retained if $\mathcal{L} = 1$, i.e. when the macroscopic relationship is approximated with an affine function. This observation leads to the simplified reformulation discussed below.

### 4.3. A special case when the macroscopic relationship is affine

In a special case where the relationship between the link occupancy (LO) and the aggregate emission rate (AER) is approximately affine, the robust reformulation discussed previously is considerably simplified. Indeed, we can reduce (4.21) to

$$\int_0^T \left[ a_1(t) N(t) + a_0(t) \right] dt \leq E \quad \forall (a_0(\cdot), a_1(\cdot)) \in \eta_a \tag{4.39}$$

where the set $\eta_a$ reduces to

$$\eta_a = \left\{ (a_0(\cdot), a_1(\cdot)) : L_l \leq a_l(t) \leq U_l, \ \forall t, \ l = 0, 1; \ \int_0^T a_1(t) dt \leq \frac{T U_1}{\sigma} = \frac{M \delta t U_1}{\sigma} \right\} \tag{4.40}$$

With time-discretization, we have the following formulation:

$$\delta t \sum_{k=1}^{M} \left( a_1^k N^k + a_0^k \right) \leq E \quad \forall (a_0, a_1) \in \hat{\eta}_a \tag{4.41}$$

where $a_0 \doteq (a_0^k : 1 \le k \le M)$, $a_1 \doteq (a_1^k : 1 \le k \le M)$. The uncertainty set is given as

$$\hat{\boldsymbol{\eta}}_a = \left\{ (a_0, a_1) : L_0 \le a_0^k \le U_0, \ L_1 \le a_1^k \le U_1, \ \forall 1 \le k \le M; \ \sum_{k=1}^{M} a_1^k \le \frac{MU_1}{\sigma} \right\} \tag{4.42}$$

With this uncertainty set, we have the following result concerning the robust reformulation, which is a special case of Theorem 4.2.

**Corollary 4.4. (Robust reformulation with affine relationship)** *If $\hat{\boldsymbol{\eta}}_a$ has nonempty interior, then the semi-infinite constraint* (4.41) *is equivalent to the following set of constraints.*

$$\sum_{k=1}^{M} U_1 \beta^k - \sum_{k=1}^{M} L_1 \gamma^k + \frac{MU_1}{\sigma} \theta + MU_0 \delta t \ \le \ E \tag{4.43}$$

$$s.t. \quad \beta^k - \gamma^k + \theta \ = \ \delta t N^k \qquad \forall 1 \le k \le M \tag{4.44}$$

$$\beta^k, \ \gamma^k, \ \theta \ \ge \ 0 \qquad \forall 1 \le k \le M \tag{4.45}$$

*where $\beta^k, \gamma^k, \theta$ are dual variables.*

Once the semi-infinite constraint (4.41) is reformulated as a finite set of linear constraints according to Corollary 4.4, the LWR-E problem with an affine macroscopic relationship can be formulated and solved as a mixed integer linear program. Such an MILP will be numerically tested in Section 6.

### 4.4. Piecewise affine macroscopic relationship

This section formulates the emission constraints for a class of more general macroscopic relationships, namely, piecewise affine functions. Since any continuous function can be approximated by a piecewise affine function, the formulation provided below will accommodate a wider range of macroscopic relationships not yet presented in this paper. In order to retain linearity in the constraints, we consider two types of piecewise affine relationships: convex piecewise affine and concave piecewise affine. Given the same objective to be minimized, convexity and concavity in the macroscopic relationship render completely different structures of the robust optimization and the MILP reformulation. More specifically, the convex piecewise affine constraint essentially means that <u>all</u> affine pieces can be bounded from above by a prescribed level in the worst case scenario. In contrast, the concave piecewise affine constraint means that <u>at least one</u> affine piece can be bounded from above in the worst case scenario. Such a difference looms large, especially, when we consider a computable reformulation in the form of a set of mixed integer linear constraints. In the following, we will discuss the convex case in Section 4.4.1 and the concave case in Section 4.4.2.

#### 4.4.1. Convex piecewise affine relationship

It will be seen in this section that the previously presented reformulation can be extended to treat the convex piecewise affine case. We denote by $\mathbb{M}$ the set of affine pieces that constitute the piecewise affine function. For each affine piece $m \in \mathbb{M}$, we let $b_{0,m}(t)$ and $b_{1,m}(t)$ be the zeroth- and first-order coefficients, respectively. We further employ the notation $\boldsymbol{b}(t) \doteq (b_{\tau,m}(t) : \tau \in \{0, 1\}, m \in \mathbb{M})$ for $t \in [0, T]$.

With these notations, the convex piecewise affine approximation of the AER is expressed as:

$$AER(N(t); \boldsymbol{b}(t)) \ \doteq \ \max_{m \in \mathbb{M}} \left\{ b_{1,m}(t) \cdot N(t) + b_{0,m}(t) \right\} \qquad t \in [0, T] \tag{4.46}$$

Again, we consider the following emission constraint:

$$\int_0^T AER(N(t); \boldsymbol{b}(t)) \, dt \ \le \ E \qquad \forall \boldsymbol{b}(\cdot) \in \boldsymbol{\eta}_b \tag{4.47}$$

where the budget-like uncertainty set is

$$\eta_b = \left\{ \boldsymbol{b}(\cdot): \; L_{\tau,m} \le b_{\tau,m}(t) \le U_{\tau,m}, \; \forall \tau \in \{0,1\}, \; \forall t \in [0,T], \; \forall m \in \mathbb{M}; \right.$$

$$\left. \sum_{m \in \mathbb{M}} \int_0^T b_{1,m}(t)\, dt \;\le\; \frac{T \sum_{m \in \mathbb{M}} U_{1,m}}{\sigma} \right\} \quad (4.48)$$

The constraint (4.47) can be immediately discretized as:

$$\delta t \sum_{k=1}^{M} \max_{m \in \mathbb{M}} \left\{ b_{1,m}^k \cdot N^k + b_{0,m}^k \right\} \;\le\; E \qquad \forall \hat{\boldsymbol{b}} \in \hat{\eta}_b \quad (4.49)$$

where $\hat{\boldsymbol{b}} \doteq (b_{\tau,m}^k : \; \tau \in \{0,1\}, \; 1 \le k \le M, \; m \in \mathbb{M})$ is the discrete-time version of $\boldsymbol{b}(\cdot)$, and the time-discretized uncertainty set is:

$$\hat{\eta}_b = \left\{ \hat{\boldsymbol{b}}: \; L_{\tau,m} \le b_{\tau,m}^k \le U_{\tau,m}, \; \forall \tau \in \{0,1\}, \; \forall 1 \le k \le M, \; \forall m \in \mathbb{M}; \right.$$

$$\left. \sum_{m \in \mathbb{M}} \delta t \sum_{k=1}^{M} b_{1,m}^k \;\le\; \frac{T \sum_{m \in \mathbb{M}} U_{1,m}}{\sigma} \right\} \quad (4.50)$$

Evidently, (4.49) is equivalent to

$$\max_{\hat{\boldsymbol{b}} \in \hat{\eta}_b} \delta t \sum_{k=1}^{M} \max_{m \in \mathbb{M}} \left\{ b_{1,m}^k \cdot N^k + b_{0,m}^k \right\} \;\le\; E \quad (4.51)$$

Even though the piecewise affine function is convex, the robust counterpart (4.51) involves a convex maximization (or equivalently, concave minimization) problem, which is nonconvex and intractable in general. The following theorem shows that the robust constraint (4.51) can be equivalently rewritten as a set of linear constraints with the aid of dual variables.

**Theorem 4.5. (Robust reformulation with convex piecewise affine relationship)** *Let constants $L_{\tau,m} < U_{\tau,m}$ and $\sigma$ be given, $\tau \in \{0,1\}$, $m \in \mathbb{M}$. Define a set of matrices*

$$\mathbb{V} \doteq \left\{ (v_{m,k}: \; m \in \mathbb{M}, \; 1 \le k \le M) \in \{0,1\}^{|\mathbb{M}| \times M} : \; \sum_{m \in \mathbb{M}} v_{m,k} = 1, \; \forall \, 1 \le k \le M \right\}$$

*If $\hat{\eta}_b$ has nonempty interior, then the semi-infinite constraint (4.49) is equivalent to the following finite set of linear constraints:*

$$\left. \begin{aligned} & \sum_{m \in \mathbb{M}} \sum_{k=1}^{M} \left( -L_{1,m} \gamma_{1,m}^{s,k} + \beta_{1,m}^{s,k} U_{1,m} + U_{0,m} \beta_{0,m}^{s,k} - L_0 \gamma_{0,m}^{s,k} \right) + \frac{M \sum_{m \in \mathbb{M}} U_{1,m}}{\sigma} \theta^s \;\le\; E \\ & -\gamma_{1,m}^{s,k} + \beta_{1,m}^{s,k} + \theta^s \;=\; N^k v_{m,k}^s \delta t \qquad \forall 1 \le k \le M, \; m \in \mathbb{M} \\ & -\gamma_{0,m}^{s,k} + \beta_{0,m}^{s,k} \;=\; v_{m,k}^s \delta t \qquad \forall 1 \le k \le M, \; m \in \mathbb{M} \\ & \gamma_{1,m}^{s,k} \ge 0, \; \beta_{1,m}^{s,k} \ge 0, \; \gamma_{0,m}^{s,k} \ge 0, \; \beta_{0,m}^{s,k} \ge 0 \qquad \forall 1 \le k \le M, \; m \in \mathbb{M} \\ & \theta^s \;\ge\; 0 \end{aligned} \right\} \quad \forall v_{m,k}^s \in \mathbb{V}$$

*where $\beta_{0,m}^{s,k}$, $\gamma_{0,m}^{s,k}$, $\beta_{1,m}^{s,k}$, $\gamma_{1,m}^{s,k}$ and $\theta^s$ are dual variables.*

*Proof.* The proof is moved to Appendix A for the conciseness of our presentation. □

Similar to Corollary 4.4 which handles the affine relationship, Theorem 4.5 provides a computable formulation that preserves linearity in the constraints even though the relationship between AER and LO is in a more general functional form. Notice that, however, the set of constraints shown in Theorem 4.5 have to be satisfied for each and every matrix $\upsilon_{m,k}^s$ in $\mathbb{V}$. In other words, there are $|\mathbb{V}|$ copies of such set of constraints, and each copy is parameterized by $s$. The cardinality $|\mathbb{V}|$ of the finite set $\mathbb{V}$ is $|\mathbb{M}|^M$. This means that the number of linear constraints grows exponentially. Nonetheless, these exponentially many constraints are all linear and do not involve any integer variables. We also make note of the fact that the presence of exponentially many constraints is not rare in combinatorial optimization problems; examples include the traveling salesman problem and the minimum spanning tree problem. In addition, algorithms such as the cutting plane method may be employed to systematically enumerate these constraints, which may lead to reduction in the computational cost. (For an example of this type of algorithms, see Nemhauser and Sigismondi, 1992).

### 4.4.2. Concave piecewise affine relationship

We show in this section that, when the relationship is piecewise affine and concave, one can maintain the linearity in the constraints by modifying the uncertainty sets and introducing additional integer variables. With the same notations introduced in Section 4.4.1, we define the concave and piecewise affine relationship as

$$AER(N(t);\, \boldsymbol{b}(t)) \;\doteq\; \min_{m\in\mathbb{M}}\left\{b_{1,m}(t)\cdot N(t) + b_{0,m}(t)\right\} \qquad t\in[0,\,T] \tag{4.52}$$

Notice that the "max" operator from the convex case has now been changed to "min". The emission constraint reads:

$$\int_0^T AER(N(t);\, \boldsymbol{b}(t))\,dt \;\leq\; E \qquad \forall \boldsymbol{b}(\cdot)\in\boldsymbol{\eta}_b \tag{4.53}$$

Here we have modified the uncertainty set $\boldsymbol{\eta}_b$ to be a product of sets: $\boldsymbol{\eta}_b = \eta_{b,1}\times\eta_{b,2}\times\cdots\times\eta_{b,m}\times\cdots\times\eta_{b,|\mathbb{M}|}$ with

$$\eta_{b,m} = \left\{\boldsymbol{b}_m(\cdot):\; L_{\tau,m}\leq b_{\tau,m}(t)\leq U_{\tau,m},\; \forall\tau\in\{0,1\},\; \forall t\in[0,\,T];\; \int_0^T b_{1,m}(t)\,dt \;\leq\; \frac{TU_{1,m}}{\sigma_m}\right\} \tag{4.54}$$

where $\boldsymbol{b}_m(\cdot) \doteq (b_{\tau,m}(\cdot):\; \tau\in\{0,1\})$. Unlike the uncertainty set defined in (4.48), in this case the uncertain parameters associated with different affine pieces are uncorrelated. In other words, the constraints related to the uncertain parameters of different affine pieces in the uncertainty set are decoupled. A time-discretization of constraint (4.53) is given as following:

$$\sum_{k=1}^M \min_{m\in\mathbb{M}}(b_{1,m}^k\cdot N^k + b_{0,m}^k)\delta t \;\leq\; E \qquad \forall\hat{\boldsymbol{b}}\doteq(\hat{\boldsymbol{b}}_1,\ldots,\hat{\boldsymbol{b}}_{|\mathbb{M}|})\in\hat{\boldsymbol{\eta}}_b\doteq\hat{\eta}_{b,1}\times\cdots\times\hat{\eta}_{b,|\mathbb{M}|} \tag{4.55}$$

where $\hat{\boldsymbol{b}}_m\doteq(b_{\tau,m}^k:\; \tau\in\{0,1\},\; 1\leq k\leq M)$, $\forall m\in\mathbb{M}$. The uncertainty set $\hat{\eta}_{b,m}$ is defined as

$$\hat{\eta}_{b,m} = \left\{\hat{\boldsymbol{b}}_m:\; L_{\tau,m}\leq b_{\tau,m}^k\leq U_{\tau,m},\; \forall\tau\in\{0,1\},\; \forall 1\leq k\leq M,\; \sum_{k=1}^M b_{1,m}^k\delta t \;\leq\; \frac{TU_{1,m}}{\sigma_m}\right\} \qquad m\in\mathbb{M} \tag{4.56}$$

The following theorem presents a reformulation of the robust constraint as a set of mixed integer linear constraints.

**Theorem 4.6. (Robust reformulation with concave piecewise affine relationship)** *Let constants $L_{\tau,m} < U_{\tau,m}$ and $\sigma_m$ be given for all $\tau\in\{0,1\}$ and $m\in\mathbb{M}$. If $\hat{\boldsymbol{\eta}}_b$ has nonempty interior, the semi-infinite constraint (4.55) is equivalent*

16

*to the following set of linear constraints:*

$$\sum_{m \in \mathbb{M}} \left\{ -\sum_{k=1}^{M} L_{1,m}\gamma_{1,m}^k + \sum_{k=1}^{M} \beta_{1,m}^k U_{1,m} + \theta_m \frac{MU_{1,m}}{\sigma_m} + \sum_{k=1}^{M} U_{0,m}\beta_{0,m}^k - \sum_{k=1}^{M} L_0\gamma_{0,m}^k \right\} \leq E$$

$$N^k \delta t - (1 - \upsilon_{m,k})\mathcal{M} \leq -\gamma_{1,m}^k + \beta_{1,m}^k + \theta_m \leq (1 - \upsilon_{m,k})\mathcal{M} + N^k \delta t \qquad \forall 1 \leq k \leq M, \; m \in \mathbb{M}$$

$$-\upsilon_{m,k}\mathcal{M} \leq -\gamma_{1,m}^k + \beta_{1,m}^k + \theta_m \leq \upsilon_{m,k}\mathcal{M} \qquad \forall 1 \leq k \leq M, \; m \in \mathbb{M}$$

$$-\gamma_{0,m}^k + \beta_{0,m}^k = \upsilon_{m,k}\delta t \qquad \forall 1 \leq k \leq M, \; m \in \mathbb{M} \qquad\qquad (4.57)$$

$$\gamma_{1,m}^k \geq 0, \; \beta_{1,m}^k \geq 0, \; \gamma_{0,m}^k \geq 0, \; \beta_{0,m}^k \geq 0 \qquad \forall 1 \leq k \leq M, \; m \in \mathbb{M}$$

$$\upsilon_{m,k} \in \{0, \, 1\} \qquad \forall 1 \leq k \leq M, \; m \in \mathbb{M}$$

$$\theta_m \geq 0 \qquad m \in \mathbb{M}$$

$$\sum_{m \in \mathbb{M}} \upsilon_{m,k} = 1 \qquad \forall 1 \leq k \leq M$$

*where $\beta_{1,m}^k$, $\gamma_{1,m}^k$, $\beta_{0,m}^k$, $\gamma_{0,m}^k$ and $\theta_m$ are dual variables and $\mathcal{M} > 0$ is sufficiently large.*

*Proof.* The proof is postponed until Appendix B. □

The concave piecewise affine relationship usually results in nonconvex constraints, which can bring substantial computational challenges. In fact, these nonconvex constraints cannot be well handled by existing commercial solvers including CPLEX and Gurobi. Nonetheless, with Theorem 4.6, we can again retain linearity by introducing additional integer variables. We further note that due to the last constraint in (4.57), these integer variables are highly correlated so that the search space of the problem grows only on an order of $M$. Therefore, our reformulation significantly reduces the computational ramifications of handling a set of nonconvex constraints.

## 5. Generalization of the robust optimization approach

As mentioned right before Section 4.1, the proposed robust optimization formulation can be easily extended to capture two additional types of environmental considerations: (1) to minimize the total emission on a subset of the network; and (2) to minimize the differences in the total emissions on some relevant links, or, to "equalize" the emissions among the links (environmental equity). These two generalizations will be discussed in this section.

### 5.1. Minimizing emissions

The proposed model is easily generalizable to the case where the traffic-driven emissions are subject to minimization rather than side constraints. In this case, we conveniently invoke the "epigraph reformulation". For example, if the objective is to minimize the total emission amount on some subset of links $\mathbb{I}_S \subset \mathbb{I}$ where $\mathbb{I}$ denotes the set of links in the network; that is,

$$\min \sum_{I_i \in \mathbb{I}_S} \int_0^T AER_i(N_i(t); \, \boldsymbol{a}_i(t)) \, dt$$

We may simply minimize a dummy variable $z$ with one additional constraint

$$\sum_{I_i \in \mathbb{I}_S} \int_0^T AER_i(N_i(t); \, \boldsymbol{a}_i(t)) \, dt \leq z$$

This additional constraint reduces to the form of (4.20), and our previously presented results apply here.

*5.2. Environmental equity*

The equity constraint is relevant when the network planner needs to enforce emissions to be distributed relatively evenly across all or part of the network. To address this issue, we consider the constraint of the following form:

$$\int_0^T AER_i(N_i(t);\, \boldsymbol{a}_i(t))\, dt - \int_0^T AER_j(N_j(t);\, \boldsymbol{a}_j(t))\, dt \ \le\ E_{ij} \qquad \forall I_i,\, I_j \in \mathbb{I}$$

where $E_{ij}$ is some prescribed threshold. We assume for now no uncertainty in the parameters $\boldsymbol{a}_i(\cdot)$ and $\boldsymbol{a}_j(\cdot)$. By assuming an affine relationship between the emission rate and link occupancy without loss of generality, this constraint can be rewritten as

$$\int_0^T \left\{ a_{i,1}(t)N_i(t) + a_{i,0}(t) \right\} dt - \int_0^T \left\{ a_{j,1}(t)N_j(t) + a_{j,0}(t) \right\} dt \ \le\ E_{ij} \qquad \forall I_i,\, I_j \in \mathbb{I}$$

Now we allow $\boldsymbol{a}_i(\cdot)$ and $\boldsymbol{a}_j(\cdot)$ to be uncertain and the uncertainty for each is described by budget uncertainty sets $\boldsymbol{\eta}_{a_i}$ and $\boldsymbol{\eta}_{a_j}$ similar to (4.23). Then we have the following robust constraint:

$$\int_0^T \left\{ a_{i,1}(t)N_i(t) + a_{i,0}(t) \right\} dt - \int_0^T \left\{ a_{j,1}(t)N_j(t) + a_{j,0}(t) \right\} dt \ \le\ E_{ij}$$

$$\forall \boldsymbol{a}_i(\cdot) \doteq (a_{i,1}(\cdot),\, a_{i,0}(\cdot)) \in \boldsymbol{\eta}_{a_i}, \quad \boldsymbol{a}_j(\cdot) \doteq (a_{j,1}(\cdot),\, a_{j,0}(\cdot)) \in \boldsymbol{\eta}_{a_j} \quad \forall I_i,\, I_j \in \mathbb{I}$$

(5.58)

This is equivalent to the following: $\forall I_i,\, I_j \in \mathbb{I}$,

$$\sup_{\boldsymbol{a}_i(\cdot)\in\boldsymbol{\eta}_{a_i}} \left\{ \int_0^T \left\{ a_{i,1}(t)N_i(t) + a_{i,0}(t) \right\} dt \right\} + \sup_{\boldsymbol{a}_j(\cdot)\in\boldsymbol{\eta}_{a_j}} \left\{ -\int_0^T \left\{ a_{j,1}(t)N_j(t) + a_{j,0}(t) \right\} dt \right\} \ \le\ E_{ij}$$

We proceed as before to time-discretize the problem. An explicit reformulation of the discretized constraint can be easily derived by invoking the dual formulations for each of the two maximization problems involved in the constraint. These dual formulations differ little from those discussed in Section 4.3 and thus are omitted here. As a result, linear constraints can be derived to replace the robust constraint (5.58), and the MILP formulation for the environmental equity problem follows.

Finally, one may also try to minimize the difference $E_{ij}$ between any two links $I_i$ and $I_j$, by adopting the same procedure discussed in Section 5.1 through the "epigraph reformulation".

## 6. Numerical study

*6.1. Network setup*

In this section, we consider a hypothetical network consisting of four intersections, three of which are signalized; see Figure 4. The ten links in this network are assumed to have the same triangular fundamental diagram with the following parameters.

$$v = 40/3 \text{ (meter/s)}, \quad \rho^{jam} = 0.4 \text{ veh/meter}, \quad \rho^c = 0.1 \text{ (veh/meter)}, \quad C = 4/3 \text{ (veh/s)}$$

In addition, all links have the same length of $L = 400$ m.

For simplicity, we assume that routing is such that vehicles at all intersections have a fixed probability of selecting either of the two downstream approaches; and the turning ratios are specified as follows.

$$\alpha_{1,5} = 0.50, \quad \alpha_{3,5} = 0.40, \quad \alpha_{2,6} = 0.30, \quad \alpha_{5,6} = 0.50, \quad \alpha_{10,3} = 0.53$$

Note that this assumption is not essential to our formulation or computation, but was made to simplify the presentation of results. In reality these turning percentages may be estimated based on turn-by-turn vehicle counts at intersections. The time horizon of our numerical example is a 15-min time period, with a time step of 10 s.
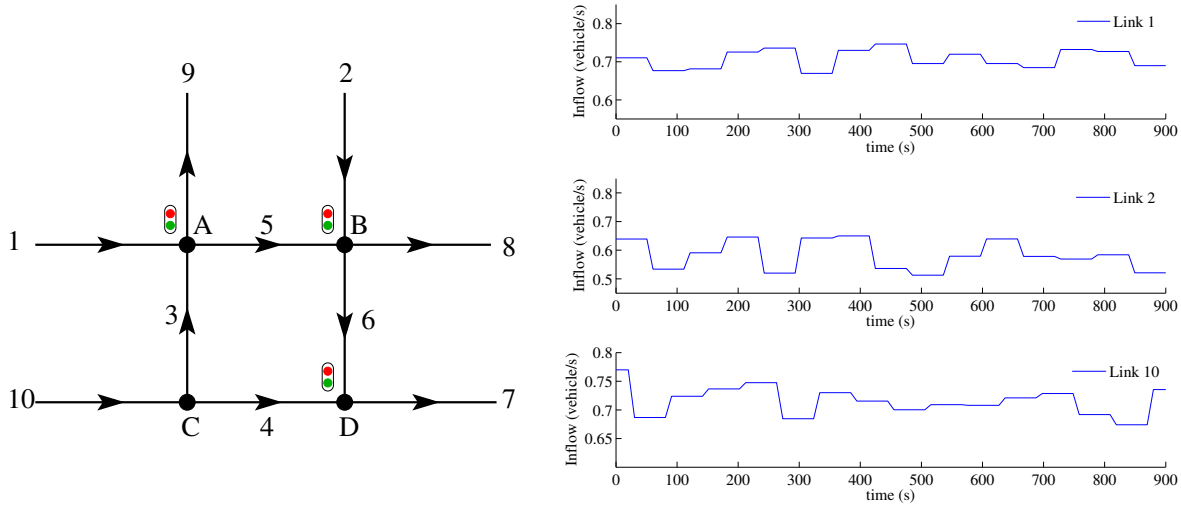
Figure 4: The test network (left) and the demand profile corresponding to Scenario I (right).

|         | Scenario I (light) | Scenario II (medium) | Scenario III (heavy) |
|---------|--------------------|----------------------|----------------------|
| Link 1  | 53.1 %             | 60.6 %               | 64.4 %               |
| Link 2  | 43.7 %             | 51.2 %               | 54.9 %               |
| Link 10 | 51.6 %             | 68.4 %               | 83.38 %              |

Table 1: Demand levels in the three scenarios with different traffic loads. The values in the table represent ratios between the average link inflows and the link flow capacity.

In order to test the effectiveness of the proposed signal timing in reducing both congestion and emission under different levels of congestion, we consider three scenarios with three different levels of demand at the upstream ends of all boundary links. The demand profile in the first scenario, which is shown in Figure 4, corresponds to the lightest traffic load. Table 1 shows, for each of the three scenarios, the ratios between the average link inflows and the link flow capacity.

Throughout this numerical study, the MILPs were solved with ILOG Cplex 12.1.0, which ran with Intel Xeon X5675 Six-Core 3.06 GHz processor provided by the Penn State Research Computing and Cyberinfrastructure.

### 6.2. Calibration of the uncertainty set

In this section, the macroscopic relationship between a link's aggregate emission rate (AER) and its link occupancy (LO) will be analyzed to inform the uncertainty set used in the robust optimization. Since the link parameters in this numerical example are identical to those employed in Section 3.1, we may focus on the macroscopic relationship depicted in Figure 2 without conducting further numerical experiments.

The macroscopic relationship shown in Figure 2 is approximately affine with the following regression coefficients

$$AER \approx a_1 \times LO + a_0 = 52.31 \times LO + 318.63 \tag{6.59}$$

where $AER$ (in gram/hour) denotes the aggregated emission rate of hydrocarbon on a link level; $LO$ (in number of vehicles) denotes the link occupancy.

In order to construct the uncertainty set of the form (4.42) for the robust optimization, we select the following lower and upper bounds for the affine coefficients:

$$L_1 \leq a_1 \leq U_1, \qquad L_0 \leq a_0 \leq U_0 \tag{6.60}$$

where

$$L_0 = 0, \quad U_0 = 400, \quad L_1 = 53.3, \quad U_1 = 66 \tag{6.61}$$
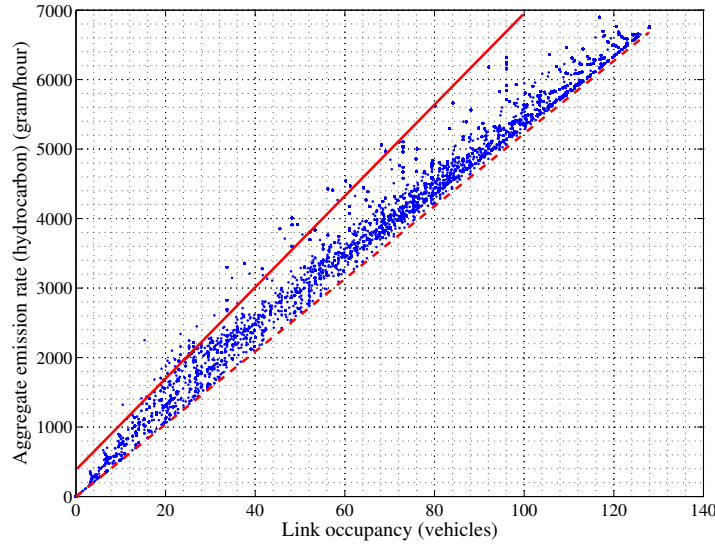
19

Figure 5: Scatter plot of the macroscopic relationship obtained from simulation. The solid line represents the upper envelop of the uncertainty region; the dashed line represent the lower envelop of the uncertainty region.

The corresponding lower envelop (by letting $a_1 = L_1$, $a_0 = L_0$) and upper envelop (by letting $a_1 = U_1$, $a_0 = U_0$) of the uncertain region are now shown in Figure 5. Notice that one has a lot of freedom in choosing these upper and lower bounds; however, they do affect the performance of the resulting robust optimization. As we commented in Section 4.1, the upper envelop sets a worse-case value for the emission rates, and it is likely to overestimate the emission rates for some/most scenarios. From Figure 5 we see that while the upper envelop provides a tight bound on the emission rates when the traffic is relatively light (i.e., $0 \leq LO \leq 40$ vehicles), it tends to mostly overestimate the emission rates when the traffic volume grows (i.e., $LO \geq 40$ vehicles). To avoid the robust constraints being too conservative, we invoke the parameter $\sigma$ introduced in (4.42) to adjust the conservativeness by allowing some realized coefficients $a_1^k$ to be strictly less than the upper bound $U_1$, that is,

$$\sum_{k=1}^{M} a_1^k \leq M \frac{U_1}{\sigma} \qquad \text{for some } \sigma \in \left[1, \frac{U_1}{L_1}\right] \tag{6.62}$$

where $a_1^k$ can be interpreted as the first-order coefficient in an actual (realized) instance of the relationship between $LO$ and $AER$ at the $k$-th time step. A general rule of thumb is that the more the upper bound seems to overestimate, the larger $\sigma$ should be, although ideally such a choice should be specifically quantified and even optimized based on available data. Interested readers are referred to Bandi and Bertsimas (2012) and Bertsimas et al. (2014) for some discussions on data-driven calibration of the uncertainty set. In our particular example, we choose $\sigma = 1.2$. Other choices of $\sigma$ can be also considered but will not be elaborated in this paper.

Notice that the region formed by the lower and upper envelops does not contain all the points shown in the figure. However, the majority (96.06%) of these points fall within this region, and we treat the rest as outliers. The reason for ignoring these outliers is that they only make up 3.94% of the total dataset, and are sparsely distributed outside (and consistently above) the uncertainty region. Including these points in the uncertainty set would make our estimation too conservative by considering the (very small) chance that these emission rates occur.

The uncertainty set for the robust optimization, according to (4.42), is therefore constructed as follows.

$$\hat{\eta}_a = \left\{ (a_0^k, a_1^k) : 0 \leq a_0^k \leq 400,\ 53.3 \leq a_1^k \leq 66,\ \forall 1 \leq k \leq M,\ \sum_{k=1}^{N} a_1^k \leq 55M \right\} \tag{6.63}$$

20

for all the links in the network, where $M$ is the number of time intervals in this problem. Notice that the same uncertainty set (6.63) is used for all the links in the network because they have identical parameters; otherwise, the uncertainty sets need to be calibrated separately.

## 6.3. The base case

For comparison purposes, we first consider a base case where the traffic signal timing is optimized, but without any emission considerations. This is achieved by simply solving the mixed integer linear program introduced in Section 2. In the base case we are mainly concerned with maximizing network throughput and minimizing delays; thus in view of (2.13) we adopt the following objective function:

$$\max \sum_{k=1}^{M} \frac{1}{1+k} \left( \hat{q}_7^k + \hat{q}_8^k + \hat{q}_9^k \right) \tag{6.64}$$

where $\hat{q}_7^k$, $\hat{q}_8^k$ and $\hat{q}_9^k$ are the exit flows at the $k$-th time step of the outgoing links 7, 8, and 9, respectively. An objective function of the form (6.64) tends to maximize the network throughput at any instance of time, and is similarly considered by Han et al. (2013, 2014a,b).

The base case results corresponding to the three demand levels are summarized in Table 2. The table also includes the hydrocarbon emission amount on each link, which is calculated from the MILP solution and the detailed modal emission model elaborated in Section 3.2. The purpose of this table is twofold: (1) to enable a comparison between the base case and the emission-constrained case presented later; and (2) to suggest appropriate upper bounds on the emission amount for the emission-constrained case. In the presentation of our results, we only consider links 1 through 6, since the rest of the links are not directly controlled by the signals under consideration.

| Scenario I | Objective value | 5.331 | | | | | |
|---|---|---|---|---|---|---|---|
| | HC emissions (gram) | link 1 | link 2 | link 3 | link 4 | link 5 | link 6 |
| | | 390.5 | 309.1 | 208.6 | 158.7 | 343.4 | 210.2 |
| | Total emission (gram) | 1620.5 | | | | | |

| Scenario II | Objective value | 6.615 | | | | | |
|---|---|---|---|---|---|---|---|
| | HC emissions (gram) | link 1 | link 2 | link 3 | link 4 | link 5 | link 6 |
| | | 558.0 | 400.4 | 263.5 | 214.3 | 514.8 | 252.9 |
| | Total emission (gram) | 2203.9 | | | | | |

| Scenario III | Objective value | 7.142 | | | | | |
|---|---|---|---|---|---|---|---|
| | HC emissions (gram) | link 1 | link 2 | link 3 | link 4 | link 5 | link 6 |
| | | 1359.2 | 430.6 | 560.1 | 269.1 | 553.0 | 252.7 |
| | Total emission (gram) | 3424.7 | | | | | |

Table 2: Objective values and hydrocarbon (HC) emissions in the base case. The three traffic scenarios with increasing demand levels are considered. The objective value is expressed by (6.64).

## 6.4. Simultaneous control of traffic and hydrocarbon emission

In this subsection, we solve the signal optimization problem with emission side constraints (LWR-E). This is achieved by solving the MILP for the base case with additional emission-related robust counterpart expressed by (4.43)-(4.45), where the detailed calibration of the uncertainty set is presented in Section 6.2. In view of the base case summarized in Table 2, we chose the following upper bounds (in gram) on the emission amount for each link, where bounds strictly below the corresponding emissions in the base case are underlined.

- Scenario I:    $E_1 = \underline{390}$,  $E_2 = 310$,  $E_3 = 210$,  $E_4 = 160$,  $E_5 = \underline{310}$,  $E_6 = 240$;

- Scenario II:    $E_1 = 600$,  $E_2 = \underline{380}$,  $E_3 = 300$,  $E_4 = \underline{210}$,  $E_5 = \underline{490}$,  $E_6 = \underline{250}$;

- Scenario III:    $E_1 = \underline{1100}$,  $E_2 = 440$,  $E_3 = 750$,  $E_4 = 300$,  $E_5 = 600$,  $E_6 = 300$.

Notice that we did not set all the bounds to be strictly below the actual emissions in the base case, because of the apparent trade-off of vehicle throughput and emission among links connected to the same intersection. Thus, under the same network load, a signal control strategy is unlikely to simultaneously reduce the emissions on all the links. In fact, setting all the bounds to be strictly below the actual emissions in the base case is likely to yield infeasibility in most of our calculations.

| Scenario I | Objective value | 5.330 (**0.02% less than the base case**) | | | | | |
|---|---|---|---|---|---|---|---|
| | HC emissions (gram) (upper bound) | link 1 | link 2 | link 3 | link 4 | link 5 | link 6 |
| | | 390.2 | 313.3 | 204.8 | 161.3 | 289.8 | 229.0 |
| | | (390) | (310) | (210) | (160) | (310) | (240) |
| | Total emission (gram) | 1588.4 grams (**1.98% less than the base case**) | | | | | |

| Scenario II | Objective value | 6.612 (**0.05% less than the base case**) | | | | | |
|---|---|---|---|---|---|---|---|
| | HC emissions (gram) (upper bound) | link 1 | link 2 | link 3 | link 4 | link 5 | link 6 |
| | | 598.5 | 369.2 | 303.8 | 214.6 | 443.8 | 247.1 |
| | | (600) | (380) | (300) | (210) | (490) | (250) |
| | Total emission (gram) | 2157.0 grams (**2.12% less than the base case**) | | | | | |

| Scenario III | Objective value | 7.058 (**1.18% less than the base case**) | | | | | |
|---|---|---|---|---|---|---|---|
| | HC emissions (gram) (upper bound) | link 1 | link 2 | link 3 | link 4 | link 5 | link 6 |
| | | 1160.5 | 434.6 | 756.4 | 272.1 | 622.2 | 294.5 |
| | | (1100) | (440) | (750) | (300) | (600) | (300) |
| | Total emission (gram) | 3540.3 grams (**3.38% more than the base case**) | | | | | |

Table 3: Objective value and hydrocarbon (HC) emissions in the LWR-E case. The numbers in the parentheses are the prescribed emission upper bounds.

The results of the proposed MILP formulation for the emission-constrained signal optimization are summarized in Table 3. We see that for all the three scenarios our proposed signal optimization scheme effectively keeps the emissions below the prescribed level; and this is done at a relatively small cost to the overall throughput of the network; that is, compared with the base case, the objective values in the LWR-E case decrease by only 0.02%, 0.05% and 1.18% in Scenarios I, II and III, respectively (recall that while lower total emissions are desired, the objective function is such that a lower value represents a worse performance). In addition, for Scenarios I and II, bounding the emission amount on certain links successfully reduces the total emission on the entire network, by 1.98% and 2.12% respectively.

In Scenario III, where the traffic load is the heaviest, we observe that, despite a significant local emission reduction (on link 1 where emission has been reduced from 1359 g to 1160 g), the overall network emission increases by 3.38%. Moreover, the total network throughput suffers more than the previous two scenarios. This further highlights the potential trade-off not only between traffic delay and emission, but also between the local and global emission amount. And such a trade-off is likely to be significant when the network is heavily congested (demand Scenario III). Nevertheless, in certain cases a trade-off between local and global emissions, such as that shown in Scenario III of Table 3, is desirable as some links may be closer to densely populated areas than others, and shifting the emissions on those links to other parts of the network results in a lower impact on population exposure and public health. Therefore, this numerical example highlights the need for a well-planned, multi-criteria signal optimization strategy based on well-defined *key performance indicators*, and additional tools for the modeling of pollutant dispersion and public exposure. These are, however, beyond the scope of this paper.

A visualization of the emission profiles for all the three test scenarios is provided in Figure 6.

We also see from Table 3 that a few emission constraints are slightly violated; this is expected for the following two reasons: (1) the calibration process for the uncertainty set presented in Section 6.2 ignores part of the data points that lie outside of the uncertainty region; (2) we chose the parameter $\sigma$ to be larger than 1 (see (6.62)), which means that the approach taken to handle the uncertainty in the AER is relatively less conservative and could lead to, with a small probability, violation of the constraints. We present a quantification of the violations in Table 4 for all the links
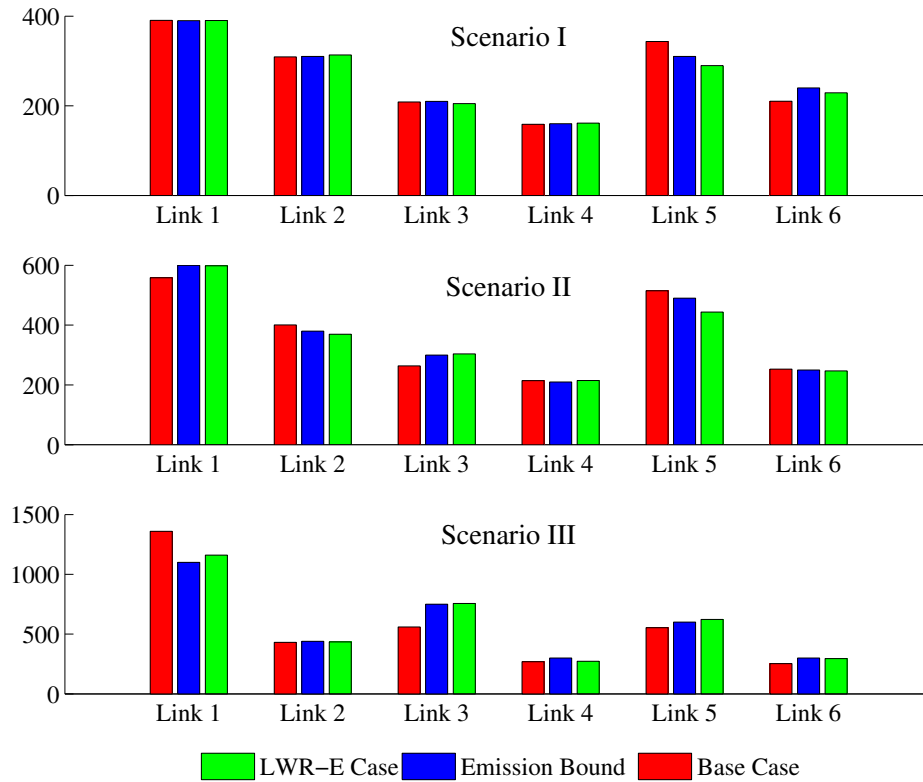
Figure 6: Comparisons of link-specific HC emissions in the base cases and the emission-constrained (LWR-E) cases.

in all the three scenarios, which shows that the violation of the constraints is within an acceptable range.

|  | link 1 | link 2 | link 3 | link 4 | link 5 | link 6 |
|---|---|---|---|---|---|---|
| Scenario I | 0.05% | 1.06% | - | 0.81% | - | - |
| Scenario II | - | - | 1.27% | 2.19% | - | - |
| Scenario III | 5.45% | - | 0.85% | - | 3.70% | - |

Table 4: Violations of the emission constraints. "-" means that the actual emission is below the bound.

We investigate a specific intersection $B$ to illustrate the effects of the proposed signal control strategy. The base case is compared with the LWR-E case under Scenario II. As can be seen from Table 2 and Table 3, the total emission amount on the two incoming links, 2 and 5, are respectively 400 g and 515 g (base), and 369 g and 444 g (LWR-E). Moreover, the emission upper bounds imposed by the LWR-E problem are 380 g and 490 g. As we see from Figure 5 that the aggregate emission rate on a link is highly correlated to the level of congestion (i.e. the link occupancy) on that link. Thus, in order to simultaneously reduce the total emissions on links 2 and 5 to a point below their respectively upper bounds, the signal controls in the LWR-E case reduced the inflow of link 5 from the upstream node $A$ to alleviate its congestion, and the signal at node $B$ acted accordingly to reduce the congestion on link 2 by allocating more green time to that link. As a result, the congestion and emission amount on links 2 and 5 were reduced simultaneously. However, this improvement in the LWR-E case was offset by the fact that the congestion was accumulated on links 1 and 3 as a result of reduced inflow of link 5, causing the emission on these links to increase: respectively from 558 g to 598 g (link 1) and 264 g to 304 g (link 3).

To visualize the effects of the LWR-E signal controls, we show the Moskowitz functions of link 5 in the base and LWR-E cases in Figure 7 (similar trends on link 2 will not be show here due to space limitation). It can be seen that

23

the Moskowitz surface is separated by a "shock wave", which is shown as a kink, into two domains: the uncongested region (on the side of the link entrance) and the congested region (on the side of the link exit). The separating shock wave travels back and forth as a result of the changing downstream boundary conditions caused by the signal control, indicating the growth and dissipation of queues. From Figure 7 we see that the LWR-E case yields less queuing on link 5 than the base case. The comparison between the base case and the LWR-E case at junction B suggests that the LWR-E case yields less queuing near this intersection, which reduces vehicle acceleration and deceleration, both of which contribute significantly to the emissions near the stop line.

To further illustrate the effect on emission of vehicle queuing near the intersection, we show in Figure 8 the contour lines of the Moskowitz function of link 5, which represent vehicle trajectories in the space-time diagram. We can clearly observe that the LWR-E case on average yields fewer vehicle stops than the base case. To further quantify this, we perform the following simple and intuitive calculation: the number of stops (represented as the horizontal line segments in the trajectories) in the base case is approximately 76, while the number of stops in the LWR-E case is roughly 52. Given that there are 50 contour lines in each figure, we estimate that the average number of stops per vehicle is 1.52 in the base case, and 1.04 in the LWR-E case. Similarly, the average number of vehicle stops on link 2 is 0.34 in the base case and 0.1 in the LWR-E case, although the corresponding contour lines are not show here. This explains the reduction of emission on these links as vehicle stops and acceleration/deceleration have been reduced by the signal controls in the LWR-E case.

## 7. Conclusion

We propose a signal optimization formulation to simultaneously minimize expected vehicle delays while accounting for constraints/objectives related to vehicular emissions throughout a network. The latter is incorporated by developing macroscopic relationships that describe the range of expected emissions on a link as a function of the occupancy of that link. Such reduced models allow emissions to be easily calculated in a network traffic model framework, as opposed to traditional techniques that require the trajectory of individual drivers in the network. Furthermore, we specifically account for the errors that exist as a result of the reduced and approximate emissions models through the use of robust optimization. This formulation entails the following theoretical advantages: (1) it is computationally tractable as a result of linearity maintained through our (piecewise) linear approximation to emission constraints; (2) it is robust to model inaccuracy through the use of robust optimization; and (3) it appropriately captures queue spillbacks to upstream links through a set of mixed integer linear constraints.

In contrast, without the RO, the only existing alternative to our knowledge is incorporating emission-related constraints as highly nonlinear and nonconvex side constraints. The signal control model then becomes a mixed integer nonlinear program (MINLP), for which there are limited theories for analysis and scarce solution methods for computation. Existing solvers for MINLP are generally slow and a global optimal solution cannot be guaranteed. In contrast, our proposed approach results in a mixed integer linear program (MILP) for the same problem, which admits effective and well-developed solution schemes that guarantee global optimality. Moreover, state-of-the-art algorithms implemented within commercial solvers highly exploit the linear structures of MILPs and thus enjoy much improved and even close-to-tractable computational efficiency (Bertsimas et al., 2011b).

A numerical study on a hypothetical network demonstrates the effectiveness of the proposed procedure in bounding/reducing emissions, while maintaining a satisfactory performance of the signalized network in terms of throughputs. However, as pointed out by Scenario III in the study, the local bounding/minimization of emissions may lead to network-wide degradation in terms of both air quality and throughput. Thus, a well-balanced signal timing plan with environmental concerns requires multiple criteria and proper weighting of different objectives.

The following data would be needed to effectively apply this methodological framework in practice: knowledge of network geometries (e.g., link lengths and fundamental diagrams), measures of traffic flows at the boundaries of the network at regular intervals (perhaps through fixed-location induction loop detectors commonly available in urban networks), anticipated or historical turning fractions at all intersections (e.g. using turn-by-turn vehicle counts), and operation constraints related to signal phases, green/red times, and signal offsets. An immediate follow-up of this research will focus on an extension of the framework to account for more realistic signal control scenarios (e.g., fixed-cycle-dynamic-split, dynamic-cycle-fixed-split, etc.) through the introduction of additional constraints on signal timing parameters. The proposed linearization of the emission constraints is also applicable for on-line signal controls based on real-time information by following the general framework proposed by Liu et al. (2015).
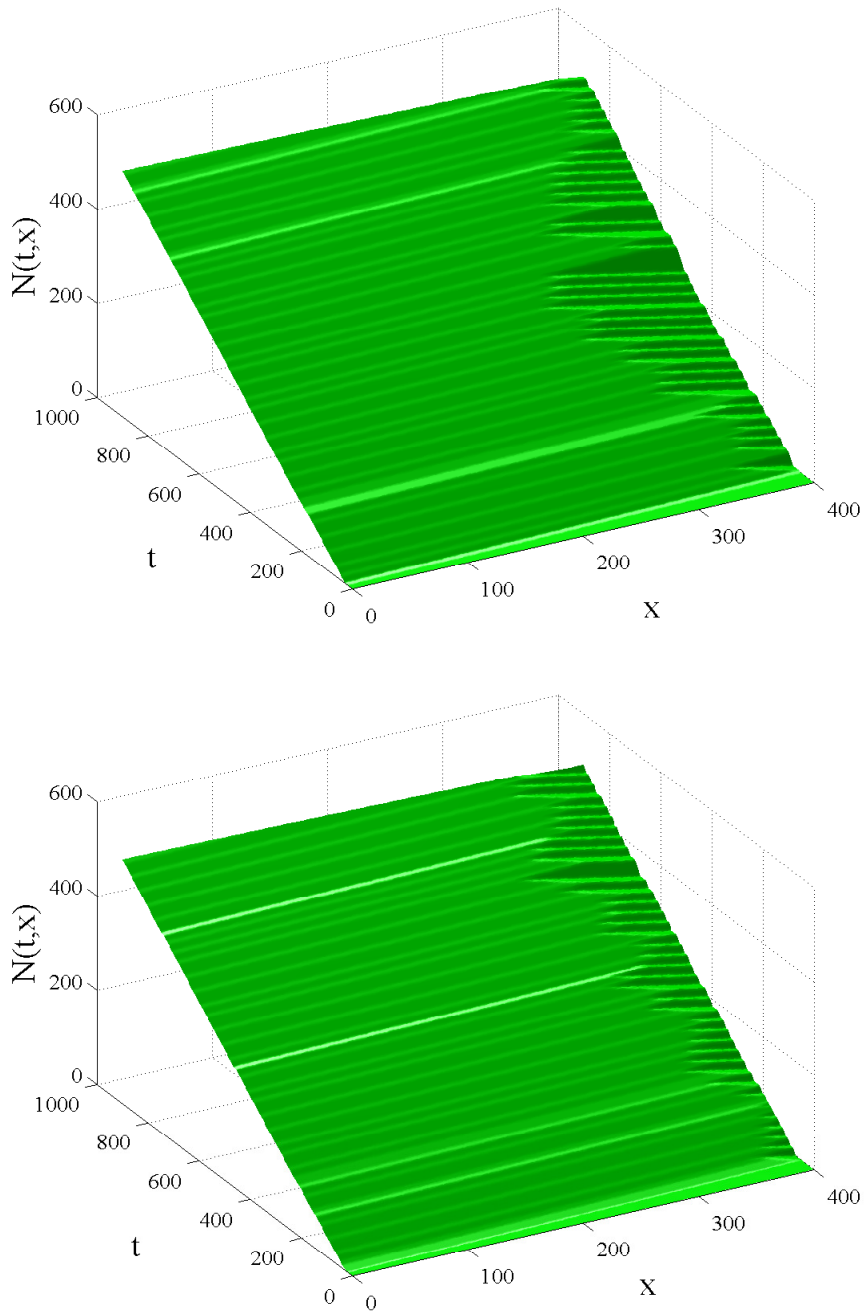
Figure 7: The Moskowitz functions of link 5 in the base case (top) and the LWR-E case (bottom).

## Appendix  A.  Proof of Theorem 4.5

*Proof.* Notice that (4.49) can be immediately rewritten as

$$\max_{\hat{\boldsymbol{b}} \in \hat{\boldsymbol{\eta}}_b} \delta t \sum_{k=1}^{M} \max_{m \in \mathbb{M}} \left( b_{1,m}^k \cdot N^k + b_{0,m}^k \right) \ \leq \ E,$$
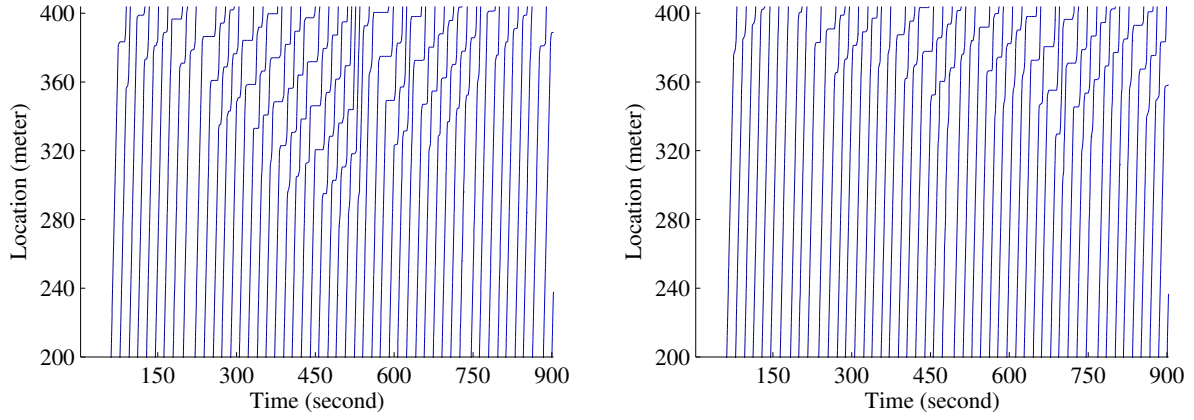
25

Figure 8: Contour lines of the Moskowitz functions of link 5, in the base case (left), and the LWR-E case (right).

which is equivalent to

$$\max_{\hat{\boldsymbol{b}} \in \hat{\eta}_b,\, (\upsilon_{m,k}) \in \mathbb{V}} \delta t \sum_{k=1}^{M} \sum_{m \in \mathbb{M}} \upsilon_{m,k} \left( b_{1,m}^k \cdot N^k + b_{0,m}^k \right) \ \leq \ E \tag{A.1}$$

Here, $\mathbb{V} = \left\{ (\upsilon_{m,k} : m \in \mathbb{M},\, 1 \leq k \leq M) \in \{0,\,1\}^{|\mathbb{M}| \times M},\ \sum_{m \in \mathbb{M}} \upsilon_{m,k} = 1,\ \forall 1 \leq k \leq M \right\}$. Consider the maximization problem for any given $(\upsilon_{m,k}) \in \mathbb{V}$ and feasible vector $\{N^k :\ 1 \leq k \leq M\}$:

$$\max_{\hat{\boldsymbol{b}} \in \hat{\eta}_b} \sum_{k=1}^{M} \sum_{m \in \mathbb{M}} \upsilon_{m,k} \left( b_{1,m}^k \cdot N^k + b_{0,m}^k \right) \delta t$$

It is easy to check that the dual problem is formulated as follows:

$$\min \ \sum_{m \in \mathbb{M}} \sum_{k=1}^{M} \left( -L_{1,m} \gamma_{1,m}^k + \beta_{1,m}^k U_{1,m} + U_{0,m} \beta_{0,m}^k - L_0 \gamma_{0,m}^k \right) + \theta \frac{M \sum_{m \in \mathbb{M}} U_{1,m}}{\sigma}$$

$$s.t. \ -\gamma_{1,m}^k + \beta_{1,m}^k + \theta \ = \ N^k \upsilon_{m,k} \delta t \qquad \forall 1 \leq k \leq M,\ m \in \mathbb{M}$$

$$-\gamma_{0,m}^k + \beta_{0,m}^k \ = \ \upsilon_{m,k} \delta t \qquad \forall 1 \leq k \leq M,\ m \in \mathbb{M}$$

$$\gamma_{1,m}^k \ \geq \ 0,\ \beta_{1,m}^k \ \geq \ 0,\ \gamma_{0,m}^k \ \geq \ 0,\ \beta_{0,m}^k \ \geq \ 0 \qquad \forall 1 \leq k \leq M,\ m \in \mathbb{M}$$

$$\theta \ \geq \ 0$$

Therefore, the satisfaction of (A.1) can be restated as follows:

$$\sum_{m \in \mathbb{M}} \sum_{k=1}^{M} \left( -L_{1,m} \gamma_{1,m}^k + \beta_{1,m}^k U_{1,m} + U_{0,m} \beta_{0,m}^k - L_0 \gamma_{0,m}^k \right) + \theta \frac{M \sum_{m \in \mathbb{M}} U_{1,m}}{\sigma} \ \leq \ E$$

$$-\gamma_{1,m}^k + \beta_{1,m}^k + \theta \ = \ N^k \upsilon_{m,k} \delta t \qquad \forall 1 \leq k \leq M,\ m \in \mathbb{M}$$

$$-\gamma_{0,m}^k + \beta_{0,m}^k \ = \ \upsilon_{m,k} \delta t \qquad \forall 1 \leq k \leq M,\ m \in \mathbb{M}$$

$$\gamma_{1,m}^k \ \geq \ 0,\ \beta_{1,m}^k \ \geq \ 0,\ \gamma_{0,m}^k \ \geq \ 0,\ \beta_{0,m}^k \ \geq \ 0 \qquad \forall 1 \leq k \leq M,\ m \in \mathbb{M}$$

$$\theta \ \geq \ 0$$

This should hold for all possible choices of $(\upsilon_{m,k}) \in \mathbb{V}$, which immediately leads to the desired result. $\qquad\square$

### Appendix B. Proof of Theorem 4.6

*Proof.* Constraint (4.55) is equivalent to stipulating that

$$\max_{\hat{\boldsymbol{b}}\in\hat{\eta}_b}\left\{\sum_{k=1}^{M}\min_{m\in\mathbb{M}}\left(b_{1,m}^k\cdot N^k+b_{0,m}^k\right)\delta t\right\}=\max_{\hat{\boldsymbol{b}}\in\hat{\eta}_b}\min_{\boldsymbol{v}\in V}\underbrace{\left\{\sum_{k=1}^{M}\sum_{m\in\mathbb{M}}\upsilon_{m,k}(b_{1,m}^k\cdot N^k+b_{0,m}^k)\delta t\right\}}_{\mathcal{F}(\hat{\boldsymbol{b}},\,\boldsymbol{v})}\leq E \tag{B.1}$$

where $V\doteq\{\boldsymbol{v}=(\upsilon_{m,k}:\ m\in\mathbb{M},\ 1\leq k\leq M)\in[0,1]^{|\mathbb{M}|\times M}:\ \sum_{m\in\mathbb{M}}\upsilon_{m,k}=1,\ \forall 1\leq k\leq M\}$. Since $\mathcal{F}(\hat{\boldsymbol{b}},\,\boldsymbol{v})$ is convex in $\boldsymbol{v}$ and concave in $\hat{\boldsymbol{b}}$, we can switch the "max" and "min" operator to obtain a dual problem without duality gap, i.e.,

$$\max_{\hat{\boldsymbol{b}}\in\hat{\eta}_b}\min_{\boldsymbol{v}\in V}\mathcal{F}(\hat{\boldsymbol{b}},\,\boldsymbol{v})\ =\ \min_{\boldsymbol{v}\in V}\max_{\hat{\boldsymbol{b}}\in\hat{\eta}_b}\mathcal{F}(\hat{\boldsymbol{b}},\,\boldsymbol{v}) \tag{B.2}$$

for any feasible vector $(N^k:\ 1\leq k\leq M)$. Combining (B.1) and (B.2), we have an alternative formulation of constraint (4.55) given as

$$
\begin{aligned}
E\ &\geq\ \max_{\hat{\boldsymbol{b}}\in\hat{\eta}_b}\min_{\boldsymbol{v}\in V}\mathcal{F}(\hat{\boldsymbol{b}},\,\boldsymbol{v})\\
&=\ \min_{\boldsymbol{v}\in V}\max_{\hat{\boldsymbol{b}}\in\hat{\eta}_b}\left\{\sum_{k=1}^{M}\sum_{m\in\mathbb{M}}\upsilon_{m,k}(b_{1,m}^k\cdot N^k+b_{0,m}^k)\delta t\right\}\\
&=\ \min_{\boldsymbol{v}\in V}\left\{\sum_{m\in\mathbb{M}}\left[\max_{\hat{\boldsymbol{b}}\in\hat{\eta}_b}\sum_{k=1}^{M}\upsilon_{m,k}(b_{1,m}^k\cdot N^k+b_{0,m}^k)\delta t\right]\right\}
\end{aligned} \tag{B.3}
$$

Consider the inner problem of (B.3) for given $\boldsymbol{v}\in V$ and $m\in\mathbb{M}$:

$$\max_{\hat{\boldsymbol{b}}\in\hat{\eta}_b}\sum_{k=1}^{M}\upsilon_{m,k}(b_{1,m}^k\cdot N^k+b_{0,m}^k)\delta t.$$

We readily see that its dual formulation is given as (note that $\boldsymbol{v}$ and $m$ are fixed here):

$$
\begin{aligned}
G^*(\boldsymbol{v},\,m)\ \doteq\ \min\quad &\sum_{k=1}^{M}\left(-L_{1,m}\gamma_{1,m}^k+\beta_{1,m}^kU_{1,m}+U_{0,m}\beta_{0,m}^k-L_0\gamma_{0,m}^k\right)+\theta_m\frac{MU_{1,m}}{\sigma_m}\\
s.t.\quad &-\gamma_{1,m}^k+\beta_{1,m}^k+\theta_m\ =\ N^k\upsilon_{m,k}\delta t\qquad\forall 1\leq k\leq M\\
&-\gamma_{0,m}^k+\beta_{0,m}^k\ =\ \upsilon_{m,k}\delta t\qquad\forall 1\leq k\leq M\\
&\gamma_{1,m}^k\ \geq\ 0,\ \beta_{1,m}^k\ \geq\ 0,\ \gamma_{0,m}^k\ \geq\ 0,\ \beta_{0,m}^k\ \geq\ 0\qquad\forall 1\leq k\leq M\\
&\theta_m\ \geq\ 0
\end{aligned}
$$

By strong duality,

$$G^*(\boldsymbol{v},\,m)\ =\ \max_{\hat{\boldsymbol{b}}\in\hat{\eta}_b}\sum_{k=1}^{M}\upsilon_{m,k}(b_{1,m}^k\cdot N^k+b_{0,m}^k)\delta t \tag{B.4}$$

Therefore, by further invoking (B.3), constraint (4.55) is equivalent to

$$E\ \geq\ \max_{\hat{\boldsymbol{b}}\in\hat{\eta}_b}\min_{\boldsymbol{v}\in V}\mathcal{F}(\hat{\boldsymbol{b}},\,\boldsymbol{v})\ =\ \min_{\boldsymbol{v}\in V}\left\{\sum_{m\in\mathbb{M}}G^*(\boldsymbol{v},\,m)\right\} \tag{B.5}$$

In view of the equality in (B.1), we know that there exists $(\upsilon_{m,k}^*)\in\{0,\ 1\}^{|\mathbb{M}|\times M}\cap V$ such that

$$\max_{\hat{\boldsymbol{b}}\in\hat{\eta}_b}\min_{\boldsymbol{v}\in V}\mathcal{F}(\hat{\boldsymbol{b}},\,\boldsymbol{v})\ =\ \max_{\hat{\boldsymbol{b}}\in\hat{\eta}_b}\mathcal{F}(\hat{\boldsymbol{b}},\,(\upsilon_{m,k}^*)) \tag{B.6}$$

Combining (B.4), (B.5), and (B.6), we may continue to get

$$
\begin{aligned}
\min_{\boldsymbol{v} \in V} \left\{ \sum_{m \in \mathbb{M}} G^*(\boldsymbol{v}, m) \right\} \; &= \; \max_{\hat{\boldsymbol{b}} \in \hat{\eta}_b} \min_{\boldsymbol{v} \in V} \mathcal{F}(\hat{\boldsymbol{b}}, \boldsymbol{v}) \; = \; \max_{\hat{\boldsymbol{b}} \in \hat{\eta}_b} \mathcal{F}(\hat{\boldsymbol{b}}, (\upsilon_{m,k}^*)) \\
&= \; \max_{\hat{\boldsymbol{b}} \in \hat{\eta}_b} \left\{ \sum_{m \in \mathbb{M}} \sum_{k=1}^{M} \upsilon_{m,k}^*(b_{1,m}^k \cdot N^k + b_{0,m}^k)\delta t \right\} \; = \; \sum_{m \in \mathbb{M}} \max_{\hat{\boldsymbol{b}} \in \hat{\eta}_b} \left\{ \sum_{k=1}^{M} \upsilon_{m,k}^*(b_{1,m}^k \cdot N^k + b_{0,m}^k)\delta t \right\} \\
&= \; \sum_{m \in \mathbb{M}} G^*((\upsilon_{m,k}^*), m)
\end{aligned}
$$

Recall that $(\upsilon_{m,k}^*) \in \{0, 1\}^{|\mathbb{M}| \times M} \cap V$. Then the above equality indicates that

$$
\min_{\boldsymbol{v} \in V} \left\{ \sum_{m \in \mathbb{M}} G^*(\boldsymbol{v}, m) \right\} \; = \; \min_{\boldsymbol{v} \in V \cap \{0, 1\}^{|\mathbb{M}| \times M}} \left\{ \sum_{m \in \mathbb{M}} G^*(\boldsymbol{v}, m) \right\}
$$

Invoking (B.5) again, we equivalently write constraint (4.55) as

$$
\min_{\boldsymbol{v} \in V \cap \{0, 1\}^{|\mathbb{M}| \times M}} \left\{ \sum_{m \in \mathbb{M}} G^*(\boldsymbol{v}, m) \right\} \; \leq \; E \tag{B.7}
$$

which is the same as requiring that there exists some $\boldsymbol{v} \in V \cap \{0, 1\}^{|\mathbb{M}| \times M}$ such that $\sum_{m \in \mathbb{M}} G^*(\boldsymbol{v}, m) \leq E$. In addition, notice that $V \cap \{0, 1\}^{|\mathbb{M}| \times M} = \{(\upsilon_{m,k}) \in \{0, 1\}^{|\mathbb{M}| \times M} : \sum_{m \in \mathbb{M}} \upsilon_{m,k} = 1, \forall 1 \leq k \leq M\}$. Combining this with the definition of $G^*(\boldsymbol{v}, m)$, we see that constraint (4.55) is equivalent to the following system.

$$
\begin{aligned}
&\sum_{m \in \mathbb{M}} \left\{ -\sum_{k=1}^{M} L_{1,m}\gamma_{1,m}^k + \sum_{k=1}^{M} \beta_{1,m}^k U_{1,m} + \theta_m \frac{MU_{1,m}}{\sigma_m} + \sum_{k=1}^{M} U_{0,m}\beta_{0,m}^k - \sum_{k=1}^{M} L_0\gamma_{0,m}^k \right\} \; \leq \; E \\
&-\gamma_{1,m}^k + \beta_{1,m}^k + \theta_m \; = \; N^k \upsilon_{m,k}\delta t \qquad \forall 1 \leq k \leq M, \; m \in \mathbb{M} \\
&-\gamma_{0,m}^k + \beta_{0,m}^k \; = \; \upsilon_{m,k}\delta t \qquad \forall 1 \leq k \leq M, \; m \in \mathbb{M} \\
&\gamma_{1,m}^k \geq 0, \; \beta_{1,m}^k \geq 0, \; \gamma_{0,m}^k \geq 0, \; \beta_{0,m}^k \geq 0 \qquad \forall 1 \leq k \leq M, \; m \in \mathbb{M} \\
&\sum_{m \in \mathbb{M}} \upsilon_{m,k} \; = \; 1 \qquad \forall 1 \leq k \leq M \\
&\upsilon_{m,k} \in \{0, 1\} \qquad \forall 1 \leq k \leq M, \; m \in \mathbb{M} \\
&\theta_m \geq 0 \qquad m \in \mathbb{M}
\end{aligned} \tag{B.8}
$$

Observe that the second line of (B.8) involves a nonlinear term, which can be replaced with linear constraints with mixed integers using the "Big-M" method. The resulting mixed integer linear constraints are presented in the second and the third lines of (4.57). □

## References

Atamtürk, A., Zhang, M., 2007. Two-stage robust network flow and design under demand uncertainty. Operations Research 55 (4), 662-673.

Aziz, H.M.A., Ukkusuri, S.V., 2012. Integration of environmental objectives in a system optimal dynamic traffic assignment model. Computer-Aided Civil and Infrastructure Engineering 27, 494-511.

Bandi, C., Bertsimas, D., 2012. Tractable stochastic analysis in high dimensions via robust optimization. Mathematical Programming, Series B 134 (1), 23-70.

Barth, M., An, F., Norbeck, J., Ross, M., 1996. Modal emission modeling: a physical approach. Transportation Research Record 1520, 81-88.

Ben-Tal, A., Nemirovski, A., 1998. Robust convex optimization. Mathematics of Operations Research 23 (4), 769-805.

Ben-Tal, A., Nemirovski, A., 1999. Robust solutions of uncertain linear programs. Operations Research Letters 25 (1), 1-14.

Ben-Tal, A., Nemirovski, A., 2000. Robust solutions of linear programming problems contaminated with uncertain data. Mathematical Programming 88, 411-421.

Benedek, C.M., Rilett, L.R., 1998. Equitable traffic assignment with environmental cost function. Journal of Transportation Engineering 124, 16-22.

Bertsimas, D., Brown, D.B., Caramanis, C., 2011a. Theory and applications of robust optimization. SIAM Review 53 (3), 464-501.

Bertsimas, D., Chang, A., Rudin, C., 2011b. Integer optimization methods for supervised ranking. Available online at http://hdl.handle.net/1721.1/67362.

Bertsimas, D., Gamarnik, D., Rikun, A.A., 2011c. Performance analysis of queueing networks via robust optimization. Operations Research 59 (2), 455-466.

Bertsimas, D., Gupta, V., Kallus, N., 2014. Data-driven robust optimization. Working paper. Online at http://www.mit.edu/v̄gupta1/Papers/DataDrivenRobOptv1.pdf

Bertsimas, D., Sim, M., 2004. Price of robustness. Operations Research 52 (1), 35-53.

Chitour, Y., Piccoli, B., 2005. Traffic circles and timing of traffic lights for cars flow. Discrete and Continuous Dynamical Systems - Series B 5 (3), 599-630.

Christofa, E., Papmichail, I., Skabardonis, A., 2013. Person-based traffic responsive signal control optimization. IEEE Transactions on Intelligent Transportation Systems 14 (3), 1278-1289.

Daganzo, C.F., 1994. The cell transmission model. Part I: A simple dynamic representation of highway traffic. Transportation Research Part B: Methodological 28 (4), 269-287.

Daganzo, C.F., 1995. The cell transmission model. Part II: Network traffic. Transportation Research Part B: Methodological 29 (2), 79-93.

Daganzo, C.F., 2005. A variational formulation of kinematic waves: basic theory and complex boundary conditions. Transportation Research Part B: Methodological 39 (2), 187-196.

Ferrari, P., 1995. Road pricing and network equilibrium. Transportation Research Part B: Methodological 29 (5), 357-372.

Gartner, N.H., 1983. OPAC: a demand-responsive strategy for traffic signal control. Transportation Research Record 906, 75-81.

Godunov, S.K., 1959. A difference scheme for numerical solution of discontinuous solution of hydrodynamic equations. Math Sbornik 47 (3), 271-306.

Haj Salem, H., Ramananjoana, C., Lebacque, J.P., 2006. Risk index model for real-time prediction of potential traffic crash on urban motorway. Proceedings of the EWGT2006, Advances in Traffic and transportation Systems Analysis, BARI, Italy,

Han, K., Friesz, T.L., Yao, T., 2013. A partial differential equation formulation of Vickrey's bottleneck model, part II: Numerical analysis and computation. Transportation Research Part B: Methodological 49, 75-93.

Han, K., Friesz, T.L., Yao, T., 2014a. A variational approach for continuous supply chain networks. SIAM Journal on Control and Optimization 52 (1), 663-686.

Han, K., Gayah, V., 2015. Continuum signalized junction model for dynamic traffic networks: Offset, spillback, and multiple signal phases. Transportation Research Part B 77, 213-239.

Han, K., Gayah, V., Piccoli, B., Friesz, T.L., Yao, T., 2014. On the continuum approximation of the on-and-off signal control on dynamic traffic networks. Transportation Research Part B: Methodological, 61, 73-97.

Han, K., Piccoli, B., Szeto, W.Y., 2012. Continuous-time link based kinematic wave model: Formulation, solution existence and well posedness. ¡http://arxiv.org/abs/1208.5141¿

Hunt, P.B., Robertson, D.I., Bretherton, R.D., Royle, M.C., 1982. The scoot on-line traffic signal optimisation technique. Traffic Engineering and Control 23 (4), 190-192.

Improta, G., Cantarella, G.E., 1984. Control system design for an individual signalized junction. Transportation Research Part B: Methodological 18 (2), 147-167.

Koukol, M., Zajickova, L., Marek, L., Tucek, P., 2015. Fuzzy logic in traffic engineering: A review on signal control, Mathematical Problems in Engineering, Article ID 979160, 2015.

Lebacque, J., Khoshyaran, M., 1999. Modeling vehicular traffic flow on networks using macroscopic models, in: Finite Volumes for Complex Applications II, 551-558, Hermes Science Publications, Paris.

Li, C., Shimamoto, S. 2011. Dynamic traffic light control scheme for reducing $CO_2$ emissions employing ETC technology. International Journal of Managing Public Sector Information and Communication Technologies 2 (1), 1-12.

Lighthill, M., Whitham, G., 1955. On kinematic waves. II. A theory of traffic flow on long crowded roads. Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences 229 (1178), 317-345.

Lin, S., De Schutter, B., Zegeye, S.K., Hellendoorn, H., Xi, Y. 2010. Integrated urban traffic control for the reduction of travel delays and emissions. Proceedings of the 13th International IEEE Conference on Intelligent Transportation Systems.

Lin, W.H., Wang, C., 2004. An enhanced 0 −1 mixed-integer LP formulation for traffic signal control. IEEE Transactions on Intelligent transportation systems 5 (4), 238-245.

Liu, H., Han, K., Gayah, V., Friesz, T.L., Yao, T., 2015. Data-driven linear decision rule approach for distributionally robust optimization of on-line signal control. In 21st International Symposium on Transportation and Traffic Theory.

Lo, H., 1999. A novel traffic signal control formulation. Transportation Research Part A 33, 433-448.

Lo, H., 1999. A cell-based traffic control formulation: strategies and benefits of dynamic timing plans. Transportation Science 35 (2), 148-164.

Lo, H., A dynamic traffic assignment formulation that encapsulates the cell-transmission model. Transportation and Traffic Theory, 327-350.

Lv, J., Zhang, Y., Zietsman, J., 2013. Investigating emission reduction benefit from intersection signal optimization. Journal of Intelligent Transportation Systems: Technology, Planning, and Operations 17 (3), 200-209.

Nemhauser, G.L., Sigismondi, G., 1992. A strong cutting plane/branch-and-bound algorithm for node packing. The Journal of the Operational Research Society 43 (5), 443-457.

Mirchandani, P., Head, L., 2000. A real-time traffic signal control system: architecture, algorithms, and analysis. Transportation Research Part C: Emerging Technologies 9, 415-432.

Post, K., Kent, J.H., Tomlin, J., Carruthers, N., 1984. Fuel consumption and emission modelling by power demand and a comparison with other models. Transportation Research Part A: Policy and Practice 18 (3), 191-213.

Robertson, D.I., Lucas, C.F. and Baker, R.T. 1980. Coordinating traffic signals to reduce fuel consumption. TRL Report LR934. Transport Research Laboratory, Crowthorne, Berkshire, United Kingdom.

Richards, P.I., 1956. Shockwaves on the highway. Operations Research 4 (1), 42-51.

Rikun, A.A., 2011. Application of robust optimization to queueing and inventory systems. Ph.D. Dissertation. Sloan School of Management.

Massachusetts Institute of Technology.

Shabihkhani, R., Gonzales, E.J., 2013. Analytical model for vehicle emissions at a signalized intersection: Integrating traffic and microscopic emissions models. In Proceedings of the 92nd Annual Meeting of the Transportation Research Board, Washington, DC.

Sims, A.G., Dobinson, K.W., 1980. The Sydney coordinated adaptive traffic (SCAT) system philosophy and benefits. IEEE Transactions on Vehicular Technology 29 (2), 130-137.

Skabardonis, A., 2000. Control strategies for transit priority. Transportation Research Record 1727, 20-26.

Stevanovic, A., Stevanovic, J., Zhang, K., Batterman, S., 2009. Optimizing traffic control to reduce fuel consumption and vehicular emissions: Integrated approach with VISSIM, CMEM, and VISGAOST. Transportation Research Record 2128, 105-113.

Szeto, W.Y., Jaber, X., Wong, S.C., 2012. Road network equilibrium approaches to environmental sustainability. Transport Reviews: A Transnational Transdisciplinary Journal 32 (4), 491-518.

Yang, H., Bell, M.G.H, 1997. Traffic restraint, road pricing and network equilibrium. Transportation Research Part B: Methodological 31, 303-314.

Yang, H., Xu, W., He, B., Meng, Q., 2010. Road pricing for congestion control with unknown demand and cost functions. Transportation Research Part C: Emerging Technologies 18, 157-175.

Zhang, L., Yin, Y. and Chen, S. 2013. Robust signal timing optimization with environmental concerns. Transportation Research Part C: Emerging Technologies 29, 55-71.

Ziliaskopoulos, A.K., 2000. A linear programming model for the single destination system optimal dynamic traffic assignment problem, Transportation Science 34 (1), 37-49.