A Three-Stream Model for Arterial Traffic

Constant Bails, Aude Hofleitner, Yiguang (Ethan) Xuan, Alexandre Bayen

Submitted to
the 91st Annual Meeting of the Transportation Research Board
August 1, 2011

Word Count:

Number of words: 6230
Number of figures: 5 (250 words each)
Number of tables: 0 (250 words each)
Total: 7480

*Ecole Polytechnique, Department of Applied mathematics, constant.bails@polytechnique.edu
†Corresponding Author, Department of Electrical Engineering and Computer Science, University of California, Berkeley, aude.hofleitner@polytechnique.edu, and UPE/IFSTTAR/GRETTIA, France
‡Department of Civil and Environmental Engineering, University of California, Berkeley, xuanyg@berkeley.edu
§Department of Electrical Engineering and Computer Science and Department of Civil and Environmental Engineering, University of California, Berkeley, bayen@berkeley.edu
Abstract

In this article, we propose a new analytical traffic flow model for traffic dynamics at signalized intersections. During each cycle, both the arrival and the departure traffic are approximated by three distinct traffic streams with uniform density. Because of the similar representation of the arrival and the departure traffic, the results from a single intersection can easily be extended to a series of intersections. The number of parameters of the model is tractable, leading to analytical solutions of the problem. We prove that the total delay of one-way traffic is a quasi-convex function in the offset between consecutive traffic cycles and derive analytically the optimal signal control corresponding to different profiles of arrival densities. This allows timely adjustments of the control as congestion evolves throughout the day. We also study how different density profiles evolve in a corridor, from one intersection to the downstream one, if there is no traffic from/to side streets. We find that all density profiles eventually lead to one profile after a few intersections. This corresponds to a green wave, in agreement with physical intuition. Finally, we test the model against data from microsimulation using CORSIM. Vehicle delay predicted by the model is shown to be close to that from the microsimulation.
1 INTRODUCTION AND RELATED WORK

Urban transportation systems are the source of numerous inefficiencies and negative externalities. It is estimated that the amount of gasoline wasted in 2007 due to traffic congestion is 3.9 billion gallons and the time lost because of delays is 4.2 billion hours [31]. For an average car, congestion cost is estimated to be about 0.13 US$ per vehicle mile [22].

To reduce externalities and improve efficiency, it is important to understand traffic dynamics in a controlled environment and to identify optimal control strategies which could help alleviate the problem. For example, continuum models [21, 27] and later cell transmission models [8, 9, 20] have been proposed to model traffic dynamics on highways. Control strategies including ramp metering [18, 3] and variable speed limits [32, 24] have also been studied extensively.

This article focuses on the case of arterial traffic, which is more difficult to study than highway traffic because of frequent interventions of traffic signals and cross traffic. The majority of the studies on arterial traffic use numerical algorithms for signal optimization [25, 23, 15, 12], or rely on simulation. These methods can handle scenario analysis of complex systems and can generate the desired signal control numerically. However, the complexity of the solution process grows rapidly with the size of the problem [6], in addition to the fact that the amount of information needed for the optimization is large and tedious to obtain for large networks. In addition, numerical solutions might not provide physical insight on the traffic patterns controlled by such schemes. Analytical solutions provide a deeper understanding of traffic flow dynamics. The purpose of analytical methods is generally not to provide detailed solutions to specific problems, but to generate general principles to solve the problem, by making specific assumptions to reduce the number of parameters and the complexity of the problem. For example, [33] derives expressions for delays at signalized intersections assuming platoon inflow. The present article considers platoon traffic and ignores secondary traffic. This is complemented by [30] which considers both platoon traffic and secondary traffic.

In this article, we focus on analytical methods, but propose a different model, relying on hydrodynamic traffic models [4]. In the present model, the arrival and departure of traffic flows at each signalized intersection are represented by three streams of traffic during each cycle. Each traffic stream is characterized by its flow and duration (the time it takes for all the traffic within the stream to go through a point in space). This is realistic if one inspects the downstream of an intersection, where there are mainly three streams of traffic: no traffic during the red time, saturation flow during the beginning of the green time (as the queue dissipates), and less than saturation flow (if undersaturated) during the end of the green time, once the queue is fully dissipated. The present model approximates the third traffic stream with a constant flow. When both the arrival and the departure traffic are modeled in this way, the results from a single signalized intersection are automatically applicable to a corridor including multiple signalized intersections. In addition, the number of parameters is limited and only grows linearly with the size of the network, facilitating analytical solutions.

The rest of the article is organized as follows. Section 2 presents the model, which characterizes the departure traffic streams based on the arrival traffic streams. In Section 3, a single signalized intersection is studied. We prove that total delay is a quasi-convex function in the traffic light offset and derive the optimal offset under
different scenarios. The spatial evolution of these scenarios is also studied. The model is compared against microsimulation data in Section 4. Section 5 discusses the generality of the model and provides conclusion about the benefits of the method.

2 MODELING TRAFFIC FLOWS THROUGH A SINGLE SIGNALIZED INTERSECTION

In this section, we develop a model of traffic dynamics through a single intersection. The model treats the arrival traffic as inputs and the departure traffic as outputs. The model describes traffic flow at each intersection with a limited number of parameters, which does not grow with the complexity of the network. This property, referred to as parameter efficiency, facilitates analytical solutions. The model is structured so that results from a single intersection can easily be extended to a series of intersections.

2.1 Three-Stream Model

Vehicular flow is modeled as a continuum and represented with macroscopic variables of flow \( q(x,t) \) (veh/s), density \( k(x,t) \) (veh/m) and velocity \( v(x,t) \) (m/s). The definition of flow implies the following relation between these three variables: 

\[ q(x,t) = k(x,t) \cdot v(x,t) \]

We assume that flow and density are linked by the fundamental diagram, as commonly done in traffic modeling [21, 27]. For arterial traffic, it is common to assume that this dependency is piecewise linear, leading to the assumption of a triangular fundamental diagram [10, 23]. The triangular fundamental diagram is fully characterized by three parameters: \( v_f \), the free flow speed (m/s); \( k_{\text{max}} \), the jam (or maximum) density (veh/m); and \( q_{\text{max}} \), the capacity (veh/s). We denote by \( k_c \) the critical density. It is the boundary density value between (i) free flowing conditions for which cars have the same velocity and do not interact and (ii) saturated conditions for which the density of vehicles forces them to slow down and the flow to decrease.

**Definition 1 (Stream of vehicles of density \( k \) and duration \( T \)).** A stream of vehicle of density \( k \) and duration \( T \) is a group of vehicles characterized by a uniform density \( k \). As the arrival or departure streams always travel at free flow speed \( v_f \), the flow within the stream is also uniform. The duration \( T \) of the stream is the time it takes for all vehicles within the stream to go through a point in space.

**Definition 2 (Undersaturated/saturated regime).** The presence of traffic signals leads to the formation of queues during the red time which start to dissipate as the signal turns green. If the queue fully dissipates before the end of the green time, we say the traffic conditions are undersaturated. Otherwise, we say that the regime is saturated.

**Definition 3 (Residual green time).** In an undersaturated arterial link, the residual green time is the period of time between the end of the queue dissipation and the beginning of the red time.

With a triangular fundamental diagram and uniform arrival of traffic, we can construct the time-space diagram, as shown in Figure 1. Note the three distinct streams of the departure flow in this figure: (1) the red time with flow zero and duration \( R \), (2)
the queue dissipation time with flow at capacity $q_{\text{max}} = k_c v_f$ and duration $G_q$, and (3) the residual green time with flow equal to the arrival flow and duration $C - R - G_q$, where $C$ is the cycle length. Note that in the saturated regime, the duration of the third state is zero since there is no residual green time. Also note that the speed of the back propagating wave for queue dissipation is denoted by $w$, and that for queue formation is denoted by $w_a$.

![Space time diagram of vehicles trajectories under uniform arrivals of density $k_o$ for an undersaturated regime.](image)

Figure 1. Space time diagram of vehicles trajectories under uniform arrivals of density $k_o$ for an undersaturated regime.

As the departure streams of a link correspond to the arrival streams of its downstream link, we propose to also model the arrival traffic as three streams, characterized by their density $k_i$ and their duration $T_i$, $i \in \{1, 2, 3\}$. However, if the arrival traffic includes three streams, the departure traffic is not necessarily three streams, as shown in the Figure 2d. The density of traffic during the residual green time may come from different streams and may not be uniform. To reduce the number of parameters to describe the system and make the model tractable, we assume that the density of traffic during the residual green time is uniform, with the average density derived next. This assumption is appropriate if street segments are long and vehicle streams of different densities merge into one stream with uniform density [28, 14].

The conservation of vehicles yields the following equation for the average density $k_f$ of the last departure stream (of duration $C - R - G_q$):
\[
\sum_{i=1}^{3} v_f k_i T_i = 0 \cdot R + v_f k_c G_q + v_f k_f (C - R - G_q).
\]

Note that the triangular fundamental diagram yields a simple relation between the flow \( q \) and the density \( k \) as \( q = v_f k \). We obtain the following expression for \( k_f \)

\[
k_f = \frac{\sum_{i=1}^{3} k_i T_i - k_c G_q}{C - R - G_q}.
\]

We denote by \( \Delta_i \) the delay experienced by the first vehicle of stream \( i \), where \( \Delta_1 = R \), the duration of the red light. If the arrival flow is uniform, the speed of queue formation is constant and is denoted \( w_i \). The speed of queue dissipation, \( w \), is also constant. They can be derived from the Rankine-Hugoniot [26] jump conditions as

\[
w_i = \frac{k_i v_f}{k_{\text{max}} - k_i} \quad \text{and} \quad w = \frac{k_c v_f}{k_{\text{max}} - k_c}.
\]
Figure 2. Dynamic of streams of vehicles through an intersection. Figure 2a: All the vehicles of the stream go through the intersection without stopping. Figure 2b: The first few vehicles of the stream do not stop at the intersection, they represent a fraction $\alpha$ of the vehicles of the stream. Figure 2c: All the vehicles of the stream stop at the intersection. Figure 2d: The last few vehicles of the stream do not stop at the intersection, they represent a fraction $\alpha$ of the vehicles of the stream.
We have \( w \geq w_i \) and thus the delay decreases linearly among the vehicles of the stream. If the queue does not fully dissipates as the last vehicle in stream \( i \) arrives (Cases 2 and 3), this last vehicle will experience a delay \( \Delta_{i+1} = \Delta_i - T_i (1 - \frac{k_i}{k_c}) \) (see Figure 2b). This expression is valid if and only if \( \Delta_i \geq T_i (1 - \frac{k_i}{k_c}) \). If this condition is not satisfied (Case 4, Figure 2d), the queue dissipates before the end of stream \( i \) and the last vehicles of the stream do not experience delay. The general expression for \( \Delta_{i+1} \) is

\[
\Delta_{i+1} = \max\left(0, \Delta_i - T_i \left(1 - \frac{k_i}{k_c}\right)\right).
\]

We introduce \( \tau_i \) such that \( \tau_i/T_i \) represents the fraction of stream \( i \) which stops at the intersection and have

\[
\tau_i = \min\left(\Delta_i, \frac{k_c}{k_i} - \frac{1}{k_c} T_i w\right).
\]

### 2.3 Characterization of the Departure Streams

We now extend the discussion to the entire cycle, and derive analytical expressions for the densities and durations of the departure streams, parameterized by the characteristics of the arrival streams. Without loss of generality, we assume that the signal turns red at \( t = 0 \), and stream 1 hits the red light at the beginning of the cycle.

A fraction \( 1 - \alpha \) of the vehicles of stream 1 reaches the intersection after the signal turns red whereas the remaining vehicles reach the intersection before the signal turns red. As we consider the signal dynamics as periodic, we can also consider that the remaining vehicles reach the intersection at the end of the cycle. To simplify the notations in the derivation, we choose this second representation, the arrival streams are thus modeled as four streams with densities \( k_i \) and duration \( \tilde{T}_i \) with \( \tilde{T}_1 = (1 - \alpha)T_1 \), \( \tilde{T}_2 = T_2 \), \( \tilde{T}_3 = T_3 \), \( \tilde{T}_4 = \alpha T_1 \) and \( k_4 = k_1 \).

In a corridor with several signalized intersections, \( \alpha \) is determined by the offset between consecutive signals. The delay experienced by the first vehicle that stops at the signal is \( \Delta_1 = R \).

The expressions of \( (\Delta_i)_{i=1:5} \) and \( (\tau_i)_{i=1:4} \) are computed for the four streams according to equations (3) and (4), with the initialization \( \Delta_1 = R \) (see Figure 3). We have

\[
\Delta_1 = R \\
\Delta_2 = \max\left(0, R - \tilde{T}_1 (1 - \frac{k_1}{k_c})\right) \\
\Delta_3 = \max\left(0, R - \tilde{T}_1 (1 - \frac{k_1}{k_c}) - \tilde{T}_2 (1 - \frac{k_2}{k_c})\right) \\
\Delta_4 = \max\left(0, R - \tilde{T}_1 (1 - \frac{k_1}{k_c}) - \tilde{T}_2 (1 - \frac{k_2}{k_c}) - \tilde{T}_3 (1 - \frac{k_3}{k_c})\right) \\
\Delta_5 = \max\left(0, R - \tilde{T}_1 (1 - \frac{k_1}{k_c}) - \tilde{T}_2 (1 - \frac{k_2}{k_c}) - \tilde{T}_3 (1 - \frac{k_3}{k_c}) - \tilde{T}_4 (1 - \frac{k_4}{k_c})\right)
\]

and
\[
\begin{align*}
\tau_1 &= R \frac{k_c}{k_c - k_1} \\
\tau_2 &= \max \left( 0, R - \frac{\tilde{T}_1}{k_c} \left( 1 - \frac{k_1}{k_c} \right) \right) \frac{k_c}{k_c - k_2} \\
\tau_3 &= \max \left( 0, R - \frac{\tilde{T}_1}{k_c} \left( 1 - \frac{k_1}{k_c} \right) - \frac{\tilde{T}_2}{k_c} \left( 1 - \frac{k_2}{k_c} \right) \right) \frac{k_c}{k_c - k_3} \\
\tau_4 &= \max \left( 0, R - \frac{\tilde{T}_1}{k_c} \left( 1 - \frac{k_1}{k_c} \right) - \frac{\tilde{T}_2}{k_c} \left( 1 - \frac{k_2}{k_c} \right) - \frac{\tilde{T}_3}{k_c} \left( 1 - \frac{k_3}{k_c} \right) \right) \frac{k_c}{k_c - k_1}
\end{align*}
\]

(6)

The intersection modifies the structure of the three arrival streams into several departure streams as follows:

\[
\begin{array}{c}
\text{Arrival streams} \\
\{(k_1, T_1), (k_2, T_2), (k_3, T_3)\} \\
\end{array} \quad \rightarrow \quad \begin{cases}
(0, R) \\
\left(k_c, \min(T_1, \tau_1) \frac{k_2}{k_c}\right) \\
\left(k_1, \max(0, T_1 - \tau_1) \frac{k_2}{k_c}\right) \\
\left(k_1, \max(0, T_1 - \tau_1) \frac{k_2}{k_c}\right) \\
\left(k_2, \min(T_2, \tau_2) \frac{k_3}{k_c}\right) \\
\left(k_2, \max(0, T_2 - \tau_2) \frac{k_3}{k_c}\right) \\
\left(k_3, \min(T_3, \tau_3) \frac{k_4}{k_c}\right) \\
\left(k_3, \max(0, T_3 - \tau_3) \frac{k_4}{k_c}\right) \\
\left(k_c, \min(T_4, \tau_4) \frac{k_4}{k_c}\right) \\
\left(k_1, \max(0, T_4 - \tau_4) \frac{k_4}{k_c}\right)
\end{cases}
\]

(7)

In this article, we assume that the traffic from/to the side streets does not affect the dynamic of the corridor. The present derivations may be generalized to take into account the effect of side street traffic. The analysis of the effect of side street traffic is out of the scope of this article but we discuss how it could be integrated to the present approach. For example, one may consider that side street traffic has a constant arrival density \(\epsilon \in [0, 1]\). With these considerations, the density of the first departure stream would then be modified from 0 to \(k_{ss}^i\) and all the densities of the following streams would be multiplied by \((1 - \epsilon^i)\). In Section 3, we show that side traffic does not perturb the optimization of a single intersection but may be relevant for corridor optimization when side streets traffic has important interactions with the traffic on the corridor.

As seen in (7), the number of departure streams can be more than three. Indeed, each stream \(i, i \in \{1, \ldots, 4\}\) leads to up to two streams: a stream representing the queue discharge if \(\Delta_i > 0\) (otherwise this stream has duration zero) and a stream representing the vehicles which do not stop if \(\Delta_i + 1 = 0\) (otherwise this stream has duration zero). This leads to up to eight streams to which we add the red phase of the signal which creates a ninth stream of density 0 and duration \(R\). To limit the number of parameters and control the complexity of the model, we approximate the departure streams listed above by three departure streams, corresponding to the red time, the queue dissipation time, and the residual green time. The red time leads to a stream of density 0 and duration \(R\). The queue dissipation leads to a stream of density \(k_c\) and
Figure 3. Top: Arrival streams of vehicles. The stream that reaches the signal as the traffic light turns red is split between two streams denoted stream 1 and stream 4. Stream 1 has duration $(1 - \alpha)T_1 = \tilde{T}_1$. It reaches the intersection as the signal turns red. Stream 4 has duration $\alpha T_1 = \tilde{T}_4$. It reaches the intersection at the end of the cycle. The waiting times of the first and last vehicles of stream $i$ are denoted $\Delta_i$ and $\Delta_{i+1}$. Note that the $\Delta_i$ can be null. In particular, in an undersaturated regime, we have $\Delta_5 = 0$ since the queue fully dissipates as the signal turns red. Bottom: Dynamic of three arrival streams through a signalized intersection, illustrating equation (7)
duration $G_q$ and we approximate the multiple streams of the residual green time as a single stream of density $k_f$ and duration $C - (R + G_q)$, as derived in (1). The densities and durations of the three departure streams are given by

$$\text{Arrival streams} \mapsto \text{Averaged departure streams}$$

$$\left\{ \begin{array}{l}
(k_1, T_1) \\
(k_2, T_2) \\
(k_3, T_3)
\end{array} \right\} \mapsto \left\{ \begin{array}{l}
(0, R) \\
(k_c, G_q) \\
(k_f, C - R - G_q)
\end{array} \right\}$$

(8)

with

- $G_q = \min(\alpha T_1, \tau_1)\frac{k_1}{k_c} + \min(T_2, \tau_2)\frac{k_2}{k_c} + \min(T_3, \tau_3)\frac{k_3}{k_c} + \min((1 - \alpha)T_1, \tau_4)\frac{k_1}{k_c}$ the duration of the queue dissipation,

- $k_f$ the merging density which only depends on $G_q$ and the parameters of the intersection as computed in (1).

### 3 APPLICATION TO THE OPTIMIZATION OF TRAFFIC SIGNALS

The model described in Section 2 provides a framework to analyze the dynamics of traffic flows through an arterial corridor. The assumptions lead to analytical derivations and a better understanding of the dynamics, providing insight for the control of arterial networks. In this section, we use this framework to analyze the well studied problem of one way corridor signal optimization. We provide analytical optimal control strategies for different scenarios of the arrival streams. This allows for timely adjustments of the control strategy in real time as congestion changes throughout the day.

#### 3.1 Problem Setting

We choose to minimize the total delay $D$ experienced at an intersection, given by

$$D = \int_0^C W(t)q(t)dt = \int_0^C W(t)vfk(t)dt,$$

(9)

where $W(t)$ is the delay experienced by the flow entering at time $t$, $q(t)$ and $k(t)$ are the flow and the density of the stream that enters at time $t$. $C$ is the cycle length assumed to have the same value for all signals.

We consider the optimization of the total delay because it finds a compromise between the duration of the delay experienced by the stopping vehicles and the proportion of vehicles that go through the intersection without stopping. Other choices of optimization problems are possible such as the maximization of the number of vehicles going through the intersection without stopping or the minimization of the maximal delay.

We derive the analytical expression of the objective function, assuming that vehicles arrive from an upstream intersection with a three-stream structure. We notice that
the cost function is additive and that we can compute the contribution of each stream independently.

As derived in Section 2.2, the delay decreases linearly among the stopping vehicles of a stream \( i \) (from the first stopping vehicle with delay \( \Delta_i \) to the last stopping vehicle with delay \( \Delta_{i+1} \)). The total delay experienced by the vehicles of a stream is the average delay of the stopping vehicles times the number of stopping vehicles. According to the definition of \( \tau_i \), the number of vehicles stopping in the queue is \( k_i v_f \tau_i \) and the minimum and maximum delays of the stopping vehicles of stream \( i \) are given by \( \Delta_{i+1} \) and \( \Delta_i \) respectively (see Figure 3).

Remark (Control variables). In traffic signal optimization, we control the duration of the red light and the offset between the two traffic signals. In a one way corridor, it is not relevant to minimize according to the duration of the red time because, without any constraints, the optimal value of the objective function is zero, corresponding to a red time equal to zero. We only control the actual offset \( \Theta \) between the two traffic signals. We introduce the standardized offset \( t_0 = \Theta - \frac{L}{v_f} \), which takes into account the free flow travel time of vehicles along the link. Here, \( L \) represents the length of the link between the two intersections.

We notice that the standardized offset \( t_0 \) is related to \( \alpha \) by \( t_0 = (1-\alpha)T_1 \). This gives the explicit expression of the total delay as a function of \( t_0 \), denoted \( D(t_0) \). Moreover, the offset \( t_0 \) determines which stream hits the signal first. This leads to an implicit dependence represented by the cyclic permutation between the streams, so that the stream that reaches the intersection as the signal turns red is denoted 1. We derive the analytical expression of the total delay \( D(t_0) \) by summing the contributions of the three arrival streams, using the previous derivations:

\[
D = v_f \left[ k_1 \min(\tau_1, T_1 - t_0) \frac{\Delta_1 + \Delta_2}{2} + k_2 \min(\tau_2, T_2) \frac{\Delta_2 + \Delta_3}{2}
+k_3 \min(\tau_3, T_3) \frac{\Delta_3 + \Delta_4}{2} + k_1 \min(\tau_4, t_0) \frac{\Delta_1 + \Delta_5}{2} \right] \tag{10}
\]

In the case of a saturated regime, all vehicles experience some delay. Let \( \Delta_{\min} \) represent the minimum delay experienced by the vehicles on the link, then the total delay is given by \( D_{\text{sat}} = D + \Delta_{\min} v_f \sum_{i=1}^3 k_i T_i \). Noticing that only the first term of the sum, \( D \), depends on \( t_0 \), it is equivalent to minimize \( D \) or \( D_{\text{sat}} \) and thus equation (10) is used to minimize the total delay in a saturated regime.

### 3.2 Convexity of the Cost Function

We notice from (6) that \( \forall i, \tau_i \leq \tau_{i-1} \). In particular, if there exists \( j \) such that \( \tau_j = 0 \), then \( \tau_m = 0 \) for \( m \geq j \). We also have \( \Delta_m = 0 \) for \( m \geq j \) since \( \tau_m = \frac{k_c}{k_c - k_m} \Delta_m \).

Proposition 1 (Analytical expression of \( D \)). In an undersaturated regime, \( \forall t_0, \exists \! m \in \{1, \ldots, 4\} \) such that \( 0 < \tau_m \leq T_m \) and we can simplify the expression of the cost function as follows:

\[
D = v_f \sum_{i=1}^{m-1} k_i \tilde{T}_i \frac{\Delta_i + \Delta_{i+1}}{2} + k_m \frac{k_c}{k_c - k_m} \frac{\Delta_m^2}{2} \tag{11}
\]
Proof. Intuitively, the index \( m \) represents the stream of vehicle from Case 4, for which the last vehicles of the stream do not stop on the queue. We will prove formally the existence and uniqueness of this index \( m \), beginning by two first intermediate results (Lemma 1 and 2).

Lemma 1. \( \forall i \geq 2, \tau_i > 0 \iff \tau_i - 1 > \bar{T}_{i-1} \).

Proof. Replacing \( \tau_i \) by its expression (Equation (4)), multiplying the strict inequality, \( \tau_i > 0 \), by the positive term \( \frac{k_e-k_i}{k_e} \) and rearranging the sum, we have

\[
R - \sum_{n=1}^{i-2} \bar{T}_n (1 - \frac{k_n}{k_c}) > \bar{T}_{i-1} (1 - \frac{k_{i-1}}{k_c}).
\]

Multiplying this inequality by \( \frac{k_i}{k_e-k_{i-1}} \), we have

\[
\left( R - \sum_{n=1}^{i-2} \bar{T}_n (1 - \frac{k_n}{k_c}) \right) \frac{k_c}{k_e-k_{i-1}} > \bar{T}_{i-1}
\]

and in particular \( \tau_i - 1 > \bar{T}_{i-1} > 0 \).

Lemma 2. \( \forall i \leq 3, \tau_i \leq \bar{T}_i \iff \tau_{i+1} \leq 0 \).

Proof. Replacing \( \tau_i \) by its expression (Equation (4)) and multiplying the inequality, \( \tau_i \leq \bar{T}_i \), by the positive term \( \frac{k_e-k_i}{k_e} \), we have

\[
R - \sum_{n=1}^{i} \bar{T}_n (1 - \frac{k_n}{k_c}) \leq 0
\]

We multiply the inequality by \( \frac{k_i}{k_e-k_{i-1}} \) and recognize the expression of \( \tau_{i+1} \) from (4). In addition, \( \tau_{i+1} \) is defined as being non negative and thus \( \tau_{i+1} = 0 \), and in particular \( \tau_{i+1} \leq \bar{T}_{i+1} \).

We prove the existence and uniqueness of \( m \): we prove that if such an \( m \) exists, it is necessarily unique and we then prove its existence

- **Uniqueness.** Let \( m \) be an index such that \( 0 < \tau_m \leq \bar{T}_m \). By induction, Lemma 1 and 2 imply that \( \forall j < m, \tau_j > \bar{T}_j > 0 \) and \( \forall j > m, \tau_j = 0 \leq \bar{T}_j \). This proves that if \( m \) exists, it is unique.

- **Existence.** We define \( j = \max \{ n \in \{0, \ldots, 4 \} | \tau_n > \bar{T}_n \} \), where \( \tau_0 \) and \( \bar{T}_0 \) are chosen arbitrarily such that \( \tau_0 > \bar{T}_0 \) and show that \( m = j + 1 \).

In an undersaturated regime, \( \tau_3 \leq t_0 \), so \( j \leq 3 \). The condition \( \tau_0 > \bar{T}_0 \) implies that \( j \geq 0 \) and thus the definition of \( j \) is proper (\( j \) is not infinite). The maximality of \( j \) implies that \( \tau_{j+1} \leq \bar{T}_{j+1} \). Using Lemma 2, we have \( \forall i \geq j + 2, \tau_i = 0 \). It remains to prove that \( \tau_{j+1} > 0 \). Reasoning by contradiction, we assume that \( \tau_{j+1} = 0 \).

- If \( j = 0 \), this implies that \( \forall n \in \{1, \ldots, 4 \}, \tau_n = 0 \) which means that no vehicle experiences delay and contradicts the assumption \( \tau_{j+1} = 0 \) as long as the red time is positive. and thus \( \forall i \geq j + 2, \tau_i = 0 \).

- If \( j \geq 1 \), then Lemma 2 implies that \( \tau_j \leq \bar{T}_j \), which contradicts the maximality of \( j \).
We conclude that \( \tau_{j+1} > 0 \) and thus \( m = j + 1 \) is the unique index such that
\[ 0 \leq \tau_m \leq \overline{T}_m. \]

\( \square \)

**Remark.** The index \( m \) is piecewise constant in \( t_0 \) and thus the expression of \( D \) holds on each of these intervals. Physically, \( m \) represents the index of the first stream in which some vehicles go through the intersection without stopping. Moreover, the expression holds in the case of a saturated regime, with \( m = 5 \) and the convention \( k_5 = 0 \).

**Proposition 2 (Property of \( D \)).** The function \( t_0 \mapsto D(t_0) \) is piecewise quadratic.

**Proof.** We study the cost function \( D() \) over an interval in which \( m \) is constant and use the expression of \( D(t_0) \) computed in Proposition 1. Both the \( \Delta_s \) and \( \overline{T}_1 \) are linear in \( t_0 \). All the terms of the sum from \( i = 2 \) to \( m - 1 \) are linear in \( t_0 \). The first term of the sum is quadratic in \( t_0 \). Therefore, \( D \) is the sum of a quadratic term and of linear terms and is quadratic on each interval in which \( m \) is constant. On each of these intervals, we have \( D(t_0) = at_0^2 + bt_0 + c \) with

\[
a = \frac{(k_c - k_1)(k_m - k_1)}{2(k_c - k_m)} \quad (12)
\]

\[
b = -\frac{Rk_c(k_1 - k_m)}{2} - \frac{\sum_{i=1}^{k-1} T_i(k_c - k_1)(k_i - k_m)}{k_c - k_m} \quad (13)
\]

and the optimum (either a minimum or a maximum according to the sign of \( a \)) is reached in:

\[
-\frac{b}{2a} = \sum_{i=1}^{m-1} T_i \frac{k_m - k_i}{k_m} - R \frac{k_c}{k_c - k_1} \quad (14)
\]

\( \square \)

Since \( D \) is piecewise quadratic, we study its monotony on each interval where \( m \) is constant in order to determine where the global optimum is. Such a study leads to the following property.

**Definition 4 (Quasi-convex function [7]).** A function \( f : D_f \to \mathbb{R} \) is called quasi-convex if its domain \( D_f \) and all its sublevel sets \( S_{f, \alpha} = \{ x \in D_f : f(x) \leq \alpha \} \) for \( \alpha \in \mathbb{R} \) are convex. In particular, a function is quasi-convex if one of the following conditions hold: (1) \( f \) is non decreasing, (2) \( f \) is non increasing, (3) \( \exists c \in D_f \) such that for \( t \leq c \) (and \( t \in D_f \)), \( f \) is nonincreasing, and for \( t \geq c \) (and \( t \in D_f \)), \( f \) is nondecreasing.

**Proposition 3 (Quasi-convexity property).** If we choose the time initialization such that \( t = 0 \) as the beginning of the stream with the highest density enters the link, \( t_0 \mapsto D(t_0) \) is a quasi-convex function on \([0, C]\).

**Proof.** Sketch of the proof, the full proof is available in [5].

We study the monotonicity of \( D \) over each interval corresponding to the three arrival streams - i.e. over \([0, T_1], [T_1, T_1 + T_2], [T_1 + T_2, C]\) and prove that there exists \( t_c \) such that the function \( t_0 \mapsto D(t_0) \) is non increasing for \( t_0 \in [0, t_c] \) and non decreasing for \( t_0 \in [t_c, C] \). The cost function is nonincreasing over the interval corresponding to the arrival stream with the highest density, it is nondecreasing over the interval corresponding to arrival stream with the lowest density and the behavior over the
last interval is such that the minimum is either reached over this interval or at the bounds of this interval. It may be reached outside of this interval if the interval over which the cost function is nonincreasing and the interval over which the cost function is nondecreasing are consecutive. Eventually, after enumerating all possible cases, we prove the quasi-convexity of the function.

3.3 Optimization of a One-Way Corridor

Given the variations of \( D(\cdot) \) on \([0, C]\), derived in the proof of Proposition 3, we can compute the optimal control (choice of the offset \( t_0 \)) analytically. We define two families of control solutions: (1) the corner solutions in which \( t_0 \) corresponds to the beginning of a stream and (2) the solutions in which \( t_0 \) lies inside the arrival time of a stream. The latter solutions only exist if the optimal \( t_0 \) is such that the first stream which stops at the signal is the one with the intermediate density, (see [5] for details). We index this intermediate density by 1. In the following, we use the convention \( k_2 \leq k_1 \leq k_3 \).

The optimal \( t_0 \) is denoted \( t_0^* \).

In corridor optimization, we optimize the offset of traffic signals over several consecutive intersections. Optimizing the sum of the total delays at each intersection over each offset is a difficult problem to solve analytically. Instead, we solve an optimization problem for each intersection. Given the departure streams resulting from the optimal control at intersection \( i \) (arrival streams of the downstream intersection \( i + 1 \)), we compute the optimal control to be applied at intersection \( i + 1 \). We define a scenario as a class of arrival streams leading to a specific choice of \( t_0^* \), denoted control strategy. A scenario \( s \) is unstable if it leads to a different choice of \( t_0^* \), denoted control strategy. A scenario \( s \) is stationary if, once this scenario occurs at an intersection, it will occur at all the downstream intersections. In the following, we identify the conditions, on the arrival streams, for each control strategy to be the optimal one. We also summarize conditions for these conditions to hold at the downstream intersection, making this scenario stationary. The details of the derivations can be found in [5] and we focus on the interpretation of these results.

3.3.1 The Optimal Control Is a Corner Solution

The optimal control \( t_0^* \) is either \( 0, T_1 \), or \( T_1 + T_2 \). One of the three streams is coordinated such that its first car reaches the signal at the beginning of the red time. Intuitively, this stream should have the lowest density. However, the following stream, which may have a density close to \( k_c \), can join the queue before it fully dissipates, causing a rapid increase in the queue length and thus in the total delay. Depending on the densities of the streams and on how many streams join the queue, the corner solution can be either of the three possibilities. Figure 4 summarizes the different scenarios representing an optimal control strategy associated with a class of arrival streams.
3.3.2 The Optimal Control Is Not a Corner Solution

There is only one scenario in which the optimal control is not a corner solution, then
\[ t^*_0 = T_1 + T_2 \frac{k_3 - k_2}{k_3 - k_1} - R \frac{k_c}{k_c - k_1}. \]
In this scenario, the first stream, with the intermediate density, is split into: a stream which does not stop in the queue (stream 4) and a stream which reaches the intersection as the signal turns red (stream 1). As the offset increases, additional vehicles from the first stream experience long delays. These long delays are not compensated by the smaller number of vehicles from the third stream (with the highest density) which experience short delays. As the offset decreases, fewer vehicles from the first stream (intermediate density) experience delay. This reduction in the total delay for the first stream is overcompensated by the significant increase in the total delay experienced by the vehicles from the third stream (with the highest density). This illustrates a trade-off between having a few cars with long delays and a lot of cars with short delays.

3.4 Relations between the scenarios and convergence towards a unique stationary optimal control

Now that we have identified all the possible scenarios, we study the interactions between them and the transitions from one to another (Figure 4).

The green dotted arrows illustrate that different paths are possible from a scenario. This means that once this scenario occurs, different scenarios are possible at the downstream intersection. The scenario at the downstream intersection depends on the parameters of the arrival streams. The solid red arrows illustrate that only one path is possible from the scenario. This means that once the scenario occurs, there is a unique scenario possible at the downstream intersection. We notice that all the scenarios converge after a finite number of iterations towards the unique stationary scenario (bottom left of the figure).

Physically, this scenario corresponds to what is called a green wave [13]. A green wave is a flow of vehicles going through a series of intersections without stopping at any red light. This result is intuitive. Indeed, at each intersection, one of the departure stream has no vehicles, corresponding to the red light. Because of the conservation of vehicles, the two other streams have a higher density after each intersection, until it reaches the critical density \( k_c \).

In a green wave, vehicles are clustered in a single stream of critical density. They arrive at the intersection during the green time and do not experience any delay. This is possible as long as the regime is undersaturated, since the duration of the single stream, at critical density, must be inferior to the duration of the green time. This minimum, expected to be local because we only optimize each intersection individually and not the entire set of intersections at once, is actually a global minimum because the cost function is null, it is not possible to do better. If the regime is saturated, it is still optimal to do a green wave from a local point of view, but it is not sure if we optimize globally.

However, a green wave is not the ideal solution for synchronizing traffic lights because it is very sensitive to external factors. At critical density, the traffic dynamics may be unstable (showing the limits of the modeling of traffic flow with a fundamental diagram). A single incident on the network (jaywalking, parallel parking) or small
Figure 4. The figure represents the different control scenarios (optimal control strategy and corresponding class of arrival streams). It also shows the dynamics of the scenario in a corridor leading to a unique stationary scenario which corresponds to a green wave.
calibration errors may cause significant delays and the formation of queues.

To improve this situation, we can choose to apply the optimal control in real-time. Given the traffic conditions at the downstream intersection (from sensors for instance), we apply the optimal control and thus anticipate an incident which would have disrupted the green wave. This idea of real-time control traffic has already been studied with real-time computations [19, 2, 29]. Here, all computations can be done off-line and analytically, reducing the online computations to comparisons between parameters, which are quasi-instantaneous.

The presence of significant side traffic changes the values of the densities of the streams and makes the conservation of the number of vehicles not hold anymore. The value of $\bar{\rho}$ is not conserved along the corridor and this brings perturbations in the model described above and uncertainty in the evolution of the control scenarios. At the intersections where the side traffic is significant, the control scenario might go back instead of following the arrows of Figure 4, slowing the process of reaching the stationary control scenario. Although the details of the evolution of the control scenarios when the side traffic become too significant are not the purpose of this article, a real-time control using sensors could be implemented in such a case, because it would measure the departure streams of an intersection and transmit to the downstream intersection information on the arrival streams. Given the arrival streams, the traffic light applies the optimal control using the diagram of Figure 4. A limit for this is possible delays in the optimal control leading to unexpected feedback dynamics. Indeed, the control is applied once the last vehicle of the upstream link leaves the link. However, it is possible to consider piecewise constant controls which average the information of the upstream links for a given interval before applying the control, leading to a smoother feedback which integrates the past dynamics.

4  NUMERICAL ANALYSIS AND VALIDATION

In this section, we validate our model with microsimulation. Results predicted by the model are compared with results from CORSIM [11], and we find that the two results are very similar.

We use CORSIM to simulate an arterial corridor equipped with four signalized intersections. To compare with the model, we focus on only one way of the traffic, heading east. As our model does not take into account traffic from/to side streets, the traffic flow is set in the simulation to be through only. Traffic from the side streets is through only as well. We denote the intersections by the indices 1 to 4 from West to East. The settings of the simulation are the following:

- The distance between two consecutive intersections is 500 feet (152.4 meters)
- The cycle has the same duration for every signal and lasts 60 seconds
- Every link is assumed to have one lane only
- Arrival flow upstream of the first intersection is 300 vehicles/hour
- Saturation flow is 2000 vehicles/hour

The arterial corridor is simulated for a range of values of the red time and the offset. For every simulation, the red time is common to every signal and the offset between two consecutive traffic lights is the same on each link. Each simulation is run
10 times for every set of values of the parameters and each simulation lasts 20 cycles. The comparison variable between the simulation and the model is the total delay of all the vehicles, experienced at an intersection, during a cycle. To avoid the effects of initialization, the total delays are averaged over the last 10 cycles of each simulation. We will compare the total delay per cycle over the three links between the intersection 1 and 4.

In the model, we consider that the arrival flow upstream of intersection 1 is uniform. Departure streams of each intersection are computed according to (8). The departure streams of intersection \( i \) are the arrival streams of intersection \( i + 1 \). At each intersection, we compute the total delay per cycle using equation (10).

We compare the total delays per cycle from the simulation and from the model in Figure 5. The left column represents the results computed between intersections 1 and 2. From top to bottom, the figure represents the total delay per cycle computed by the model, the microsimulation and the difference between the microsimulation and the model. The results are presented as functions of the red time \( R \) and the offset \( t_0 \). The model underestimates the total delay by about 20% on average. We notice that the two surfaces have extremely similar shapes. The total delays computed by the simulation and by the model exhibit a similar dependency on the parameters (red time and offset), which implies that the assumptions of the models are reasonable for signal control.

The model is relevant to obtain better understanding of traffic flow dynamics and study problems where absolute values are not as important as intuition on the response of a corridor to a change in the parameter values. The traffic signal optimization problem is a good application of our model because the key point of this problem is to obtain the value of the optimal control and not the one of the minimal total delay. Even though the minimal value of the total delay is underestimated by about 20% by our model, the optimal control derived by our model and by the simulation are close due to the similar shapes of the curves of the total delay.

In Figure 5, the right column represents the results computed between intersections 3 and 4. The model again underestimates the total delay by about 40% on average. The two surfaces remain very similar, though the difference is more notable compared with the delay between intersections 1 and 2. This result is expected, due to our approximation of the third stream made at each intersection. The error is thus increasing each time an approximation is made, corresponding to another intersection gone through.

From a hydrodynamical theory point of view, if we consider an intersection with uniform arrivals (a single stream of density \( k \) and duration \( C \)), there are exactly three streams downstream of the intersection (red time with density zero and duration \( R \), queue discharge with density \( k_c \) and duration \( G_q \) and residual green time with density \( k \) and duration \( C - (R + G_q) \)). The differences in the computation of the total delay between the model and the microsimulation do not result from the three-stream approximation. We estimate the error of 20% to be due to the triangular shape of the fundamental diagram and to the deterministic trajectories of the vehicles. We can consider this difference of 20% as a baseline error. The approximation of the model as a three-stream traffic flow lead to an underestimation of 40% of the total delay at intersection 4. To be used for delay or travel time estimation, the model needs to be improved to model traffic flows after several intersections.
Figure 5. Comparison of the total delay computed by the microsimulation and by the model. Top: Total delay per cycle computed by the model between intersections 1 and 2 (left) and between intersections 3 and 4 (right). Center: Total delay per cycle computed by the microsimulation between intersections 1 and 2 (left) and between intersections 3 and 4 (right). Bottom: Difference between the total delay per cycle computed by the microsimulation and by the model. The results are presented for the total delay between intersections 1 and 2 (left) and between intersections 3 and 4 (right).
5 DISCUSSION AND CONCLUSIONS

This work presents the derivations of a model of arterial traffic flow through signalized intersections. This model allows the traffic flow to be characterized by a small number of parameters. Moreover, the study of a corridor is made easier and analytical by the similar structure of the arrival and the departure flows at each intersection.

This model provides an analytical solution to the classic problem of traffic light coordination. We notice that the total waiting time of the vehicles during a cycle is a quasi-convex function of the offset between successive traffic signals. We use this quasi-convexity property to derive the optimal control analytically. For a corridor with multiple intersections, this analysis provides optimal control for the traffic signal at an intersection as a function of the departure streams of the upstream intersection. We analyze how the optimal control strategies evolve throughout the multiple intersections.

After a few intersections, our analysis shows that the choice of the optimal offset leads to a green wave, an intuitive optimization of the offset on a corridor. The results go beyond recalling that the formation of a green wave is the optimal control strategy on a corridor. They provide analytical optimal control strategies for the choice of the offsets, as a function of the arrival streams. This provides valuable information for a real-time implementation with timely adaptation of the control strategies as traffic conditions change, since it does not require additional computation. Given flow measurements from sensors, the traffic signals can compute the optimal offset from the analytical expressions derived in this article. In particular, no online optimization is necessary which is crucial to implement real-time control strategies. The implementation of such algorithms have become a realistic approach to real-time traffic signal control in the recent years, with the emergence of novel sensing technologies available for online control [1].

This model is not limited to the one-way synchronization problem and could be applied to model the flow in numerous arterial traffic situations. The two-way corridor can be studied with the same method and preliminary results are available in [5]. In a saturated regime, it is optimal to optimize the direction of traffic with the longer red light duration. This result does not hold under an undersaturated regime but it provides some model of the system behavior which can be adapted to further studies of the two-way problem. In addition to traffic lights synchronization, this model has potential applications to model the probability distribution of travel times on arterial corridor. In particular, we are interested in the additional accuracy provided by the three-stream approach compared to models which do not take into account light synchronization and assume constant arrival rates [17, 16].

References


