Probability distributions of travel times on arterial networks: a traffic flow and horizontal queuing theory approach

Aude Hofleitner[*]   Ryan Herring[†]   Alexandre Bayen[‡]

91st Annual Meeting of the Transportation Research Board
January 22-26, 2012, Washington D.C.

Initial submission: August 1, 2012
Final submission: November 15, 2012

Word Count:

Number of words:	5967
Number of figures:	6 (250 words each)
Number of tables:	1 (250 words each)
Total:	7717

[*]Corresponding Author, Department of Electrical Engineering and Computer Science, University of California, Berkeley, aude.hofleitner@polytechnique.edu, and UPE/IFSTTAR/GRETTIA, France

[†]Apple Inc. Affiliation during redaction of the article: California Center for Innovative Transportation, Berkeley, CA, ryanherring@berkeley.edu

[‡]Department of Electrical Engineering and Computer Science and Department of Civil and Environmental Engineering, University of California, Berkeley, bayen@berkeley.edu

1

Abstract

In arterial networks, traffic flow dynamics are driven by the presence of traffic signals, for which precise signal timing is difficult to obtain in arbitrary networks or might change over time. A comprehensive model of arterial traffic flow dynamics is necessary to capture its specific features in order to provide accurate traffic estimation approaches. From hydrodynamic theory, we model arterial traffic dynamics under specific assumptions standard in transportation engineering. We use this flow model to develop a statistical model of arterial traffic. The statistical approach is essential to capture the variability of travel times among vehicles: (1) the delay experienced by a vehicle depends on the time when it enters the link (in relation to the signal green/red phases) and this entrance time can occur at any random time during the cycle and (2) the free flow speed of a vehicle depends both on the driver and on external factors (jaywalking, double parking, etc.) and is another source of uncertainty. These two sources of uncertainty are captured by deriving the probability distribution of delays (from hydrodynamic theory) and modeling the nominal free flow travel time as a random variable (which encodes variability in driving behavior). We derive an analytical expression for the probability distribution of travel times between any two locations on an arterial link, parameterized by traffic parameters (cycle time, red time, free flow speed distribution, queue length and queue length at saturation).

We validate the model using probe vehicle data collected during a field test in San Francisco, as part of the *Mobile Millennium* system. The numerical results show that the new distribution derived in this article more accurately represents the actual distribution of travel times than other distributions that are commonly used to represent travel times (normal, log-normal and Gamma distributions). We also show that the model performs particularly well when the amount of data available is small. This is very promising as the volume of probe vehicle data available in real time to most traffic information systems today remains sparse.

2

1 Introduction and related work

Traffic congestion comes with important external costs due to added travel time, wasted fuel and increased traffic accidents [33]. An accurate, reliable system for estimating and forecasting traffic conditions is essential to operations and planning. Historically, the design of highway traffic monitoring systems relied mostly on dedicated sensing infrastructure (loop detectors, radars, video cameras). When properly deployed, these data feeds provide sufficient information to reconstruct macroscopic traffic variables (flow, density, velocity) using traffic flow models developed in the literature [26, 31, 12]. However, for the secondary network or highways not covered by this infrastructure, traffic estimation relies on probe vehicle data, which comes from various sources (fleets, smartphones, RFID tags), each with their own specific issues (sparsity, bias, noise, coverage).

Proof of concept studies have demonstrated the feasibility of designing highway traffic monitoring systems relying on probe data only [19, 36]. However, arterials come with additional challenges: the underlying flow physics which governs them is more complex and highly variable (traffic lights with unknown cycles, turn movements, pedestrian traffic). Microscopic models have mainly focused on modeling single intersections (or a few intersections) relying on significant data availability assumptions (including signal timing, vehicle counts or a high penetration rate of travel time measurements) [6]. While macroscopic flow models exist for the secondary network [17, 32], their parameters require site-specific calibration experiments. In addition, even if they were known, the complexity and statistical variability of the underlying flows make it challenging to perform estimation of the full macroscopic state of the system at low penetration rates of probe vehicles (which appears to be one of the few available data sources for arterial networks in the near future, at a global scale).

An important challenge in arterial traffic estimation is the characterization of travel time distributions, which was first studied with the emergence of flow-based traffic engineering [8]. Previous research uses *vertical queueing theory* to study the probability distribution of delays and queue lengths under stationary assumptions. Vertical queues presume that vehicles do not back up over the length of the roadway, but rather stack up upon one another at the stop line of a traffic signal. Under fixed cycle assumption, the derivation of the average delay and queue length at the end of the green time is derived using analytical expressions and numerical simulations for Poisson arrivals [35, 29] and more general arrival distributions [14, 11, 25]. The characterization of the stationary delay distribution was derived under simplified assumptions [5, 18], using numerical methods [30]. Recent work proposes a method to dynamically estimate the mean and variance of delay but does not characterize the entire distribution [34]. Vertical queuing theory does not model how vehicles physically queue over the length of the roadway. To address this, we propose using a *horizontal queuing theory* based approach in this article. For practitioners, analytical formulas of the mean delay are given in the Highway Capacity Manual [3] and related work [16]. They rely on static parameters of the road (number of lanes, average flow, cycle timing), rarely accessible on large scale networks.

We use the physics of traffic flows as a basis for designing probability distributions of the traffic variables. This work provides a hydrodynamic theory based statistical model of arterial traffic. We formulate specific assumptions on the physics of traffic flow to make the problem tractable while keeping it realistic. We derive analytical expressions for the probability distributions of travel times between arbitrary points of a link of the network. These distributions are characterized by a small set of parameters with direct physical interpretation (signal timing, queue length). When travel time measurements are available, *e.g.* from sparsely sampled probe vehicles representative of today's available data, one can estimate these parameters and thus estimate the probability distribution of travel times. Ultimately, our approach estimates parameters that were assumed given in previous research and these parameters represent valuable information for traffic management entities.

3

The rest of this article is organized as follows. In Section 2, we present traffic theory results derived from hydrodynamic models and horizontal queuing theory. We use these results in Section 3 to derive parametric delay distributions between two points on an arterial link (Section 3). We discuss the estimation capabilities of the parameters of the model (signal timing, queue length) depending on the sampling scheme. Noticing that the travel time is the sum of the delay and the free flow travel time, we derive the probability distribution of travel times in Section 4 and study how we can learn its parameters (in particular the red time, the queue length and the congestion level) using sparsely sampled probe vehicles. We study the estimation capabilities of the algorithm in Section 5 using high-frequency probe vehicle data collected by the *Mobile Millennium* system.

2   Traffic flow modeling and horizontal queueing theory

2.1   Traffic model

In traffic flow theory, it is common to model vehicular flow as a continuum and represent it with macroscopic variables of *flow* $q(x,t)$ (veh/s), *density* $\rho(x,t)$ (veh/m) and *velocity* $v(x,t)$ (m/s). The definition of flow gives the following relation between these three variables [26, 31]:

$$q(x,t) = \rho(x,t)\, v(x,t). \tag{1}$$

Experimental data has shown a relationship between flow and density known as the *fundamental diagram* (FD) of traffic flows [12], represented by the relation $q = \psi(\rho)$. In this article, we make the common assumption in arterial traffic of a triangular FD [13, 27]. It is fully characterized by: v_f, the free flow speed (m/s); $\rho_{\max}$, the jam (or maximum) density (veh/m); and $q_{\max}$, the capacity (veh/m). Its analytical expression is given by:

$$\psi(\rho) = \begin{cases} v_f \rho & \text{if } \rho \in [0, \rho_c] \\ q_{\max} - w(\rho - \rho_c) & \text{if } \rho \in [\rho_c, \rho_{\max}] \end{cases}, \text{ with } \rho_c = \frac{q_{\max}}{v_f} \text{ and } w = \frac{\rho_c v_f}{\rho_{\max} - \rho_c}.$$

We note that ρ_c represents 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. When a queue dissipates, vehicles are released from the queue with the maximum flow—capacity $q_{\max}$—which corresponds to the critical density $\rho_c = q_{\max}/v_f$.

For a given road segment of interest, the arrival rate at time t, i.e. the flow of vehicles entering the link at t, is denoted by $q_a(t)$. Equation (1) relates it to the arrival density $\rho_a(t) = q_a(t)/v_f$.

2.2   Traffic flow modeling assumptions

We make the following assumptions on the dynamics of traffic flow:

1. *Triangular fundamental diagram*: standard assumption in transportation engineering [13].

2. *Stationarity of traffic*: during each estimation interval, the parameters of the light cycles (red time R and cycle time C) are constant and the arrival rate of vehicles is C-periodic. This applies in particular to fixed time signals or signal plans which remain constant for several cycles. Moreover, we assume that there is no consistent increase or decrease in the length of the queue, nor instability. With these assumptions, the traffic dynamics are periodic with period C (length of the light cycle). This work is primarily focused on estimating travel time distributions for cases in which measurements are *sparse*. The assumption of stationary quantities for a limited period of time does not limit the derivations of the model because we are interested in trends rather than fluctuations. The duration of time intervals during

4

which traffic is assumed stationary may depend on the time of the day as conditions may change more rapidly at the beginning and at the end of rush hour periods, as congestion forms and dissipates. Note that an algorithm which detects changes in traffic conditions [21] may be run in parallel to dynamically update estimation intervals depending on the traffic conditions.

3. *Uniform arrivals*: the desire to derive an analytical model of arterial traffic leads to the simplifying assumption of constant arrival density ρ_a for each estimation interval. Note that constant arrivals are periodic with period C (for any C) and thus, traffic dynamics remain stationary under this assumption. We will discuss how to relax this assumption in the remainder of this article.

4. *Model for differences in driving behavior*: the free flow pace (inverse of the free flow speed) is not the same for all vehicles: it is modeled as a random variable with parameter vector θ_p—*e.g.* the free flow pace has a Gaussian or Gamma distribution with parameter vector $\theta_p = (\bar{p}_f, \sigma_p)^T$ where $\bar{p}_f$ and σ_p are respectively the mean and the standard deviation of the random variable.

Remark (**Multi-lane arterials**). *We do not take into account lane changes, passing or merging in this model. For an arterial link with several lanes, we assume that there is one queue per lane, with its own dynamics. The parameters of the road network and the level of congestion may be different on each lane (*e.g.* to model turning movements) or equal (to limit the number of parameters of the model). In the numerical implementation presented in this article, we consider that all lanes have the same queue length and do not model the different phases of traffic signals due to dedicated turns.*

2.3   Arterial traffic dynamics

In arterial networks, traffic is driven by the formation and the dissipation of queues at intersections. The dynamics of queues are characterized by shocks, which are formed at the interface of traffic flows with different densities.

We define two discrete traffic regimes: *undersaturated* and *congested*, which represent different dynamics of the arterial link depending on the presence (respectively the absence) of a remaining queue when the light switches from green to red. Figure 1 illustrates these two regimes under the assumptions made in Section 2.2. The speed of formation and dissolution of the queue are respectively called v_a and w. Their expression is derived from the Rankine-Hugoniot [15] jump conditions and given by

$$v_a = \frac{\rho_a v_f}{\rho_{\max} - \rho_a} \quad \text{and} \quad w = \frac{\rho_c v_f}{\rho_{\max} - \rho_c}. \tag{2}$$

Undersaturated regime.   In this regime, the queue fully dissipates within the green time. This queue is called the *triangular queue* (from its triangular shape on the space-time diagram of trajectories). It is defined as the spatio-temporal region where vehicles are stopped on the link. Its length is called the maximum queue length, denoted $l_{\max}$, which can also be computed from traffic theory:

$$l_{\max} = R\frac{wv_a}{w - v_a} = R\frac{v_f}{\rho_{\max}}\frac{\rho_c\rho_a}{\rho_c - \rho_a}. \tag{3}$$

Congested regime.   In this regime, there exists a part of the queue downstream of the triangular queue called the *remaining queue* with length l_r corresponding to vehicles which must stop multiple times before going through the intersection.

All notations introduced up to here are illustrated for both regimes in Figure 1.

Stationarity of the two regimes.   Assumption 2 made earlier implies the periodicity of these queue evolutions. In particular, there is not a consistent increase or decrease in the length of the

5

queue for the duration of the time interval. This assumes that the congested regime is exactly at *saturation*: the numbers of vehicles entering and exiting the link during a cycle are equal. At saturation, the arrival density is $\rho_a^s = \frac{C-R}{C}\rho_c$. The triangular queue length at saturation is computed by replacing $\rho_a = \rho_a^s$ in equation (3) or by noticing that the number of vehicles that stop in the queue ($l_{\max}^s \rho_{\max}$) is equal to the number of vehicles that exit the link in the duration of a cycle ($(C-R)v_f\rho_c$):

$$l_{\max}^s = v_f\rho_c(C-R)/\rho_{\max} \tag{4}$$

Note that saturation is an idealized notion that we assume valid for small time intervals. In the following, x is used to denote the distance from a location on a link to the downstream intersection.

The undersaturated and congested regimes are labeled u and c respectively. A probabilistic model based on the assumptions formulated in this section provides the *probability distribution function* pdf of delays δ_{x_1,x_2} and travel times y_{x_1,x_2} between two locations x_1 and x_2 on a link of the network. They are denoted $h(\delta_{x_1,x_2})$ (Section 3) and $g(y_{x_1,x_2})$ (Section 4) respectively. These pdf are parameterized by the traffic parameters: the free flow pace p_f with pdf φ_p (parameterized by θ_p), the cycle time C, the red time R, the queue length at saturation $l_{\max}^s$ and the queue length l. Note that $l = l_{\max}$, length of the triangular queue in the undersaturated regime and $l = l_{\max}^s + l_r$, sum of the length of the triangular queue at saturation and the remaining queue in the congested regime. This set of variables is sufficient to characterize the distribution of travel times resulting from the modeling assumptions.

3 Modeling the probability distribution of stopping time

The delay experienced by vehicles traveling on arterial networks is conditioned on two factors. First, the traffic conditions dictate the state of traffic experienced by all vehicles entering the link. Second, the time (in relation to the beginning of the signal's cycle) at which each vehicle enters a link determines how much delay will be experienced in the queue due to the presence of the traffic signal. Under similar traffic conditions, drivers experience different delays depending on their arrival time. Using the assumption that the arrival density (and thus the arrival rate) is constant, arrival times are uniformly distributed on the duration of the light cycle. This allows us to derive the analytical expression, $h^s(\delta_{x_1,x_2})$, $s \in \{u,c\}$ of the pdf of stopping time δ_{x_1,x_2} between locations x_1 and x_2.

In this work, we assume that we receive travel time measurements from vehicles traveling on the network. The vehicles are sampled uniformly in time, as is commonly done with fleets, and they send tuples of the form (x_1,t_1,x_2,t_2) where x_1 is the location of the vehicle at t_1, x_2 is the position of the vehicle at t_2 and $t_2 - t_1$ represents the sampling interval (usually constant from one measurement to another). This is representative of fleets which typically send data every minute in urban networks. We consider the tuples sent by the vehicles as independent. For example, we assume that the sampling strategy is such that we cannot reconstruct the trajectories of vehicles from the tuples (*e.g.* at each sampling time, the vehicles send tuples with a defined probability).

3.1 Pdf of stopping time in the undersaturated regime

In the undersaturated regime, we call η_{x_1,x_2}^u the fraction of the vehicles entering the link during a cycle that experience a delay between x_1 and x_2. The remaining vehicles entering the link in a cycle travel from x_1 to x_2 without experiencing any delay. The proportion η_{x_1,x_2}^u is computed as the ratio of vehicles joining the queue between x_1 and x_2 over the total number of vehicles entering the link in one cycle (Figure 2). The number of vehicles joining the queue between x_1 and x_2 is the number of vehicles stopped between x_1 and x_2: $\left(\min(l_{\max},x_1) - \min(l_{\max},x_2)\right)\rho_{\max}$. The number of vehicles entering the link is $v_f C \rho_a$. The proportion of vehicles delayed between
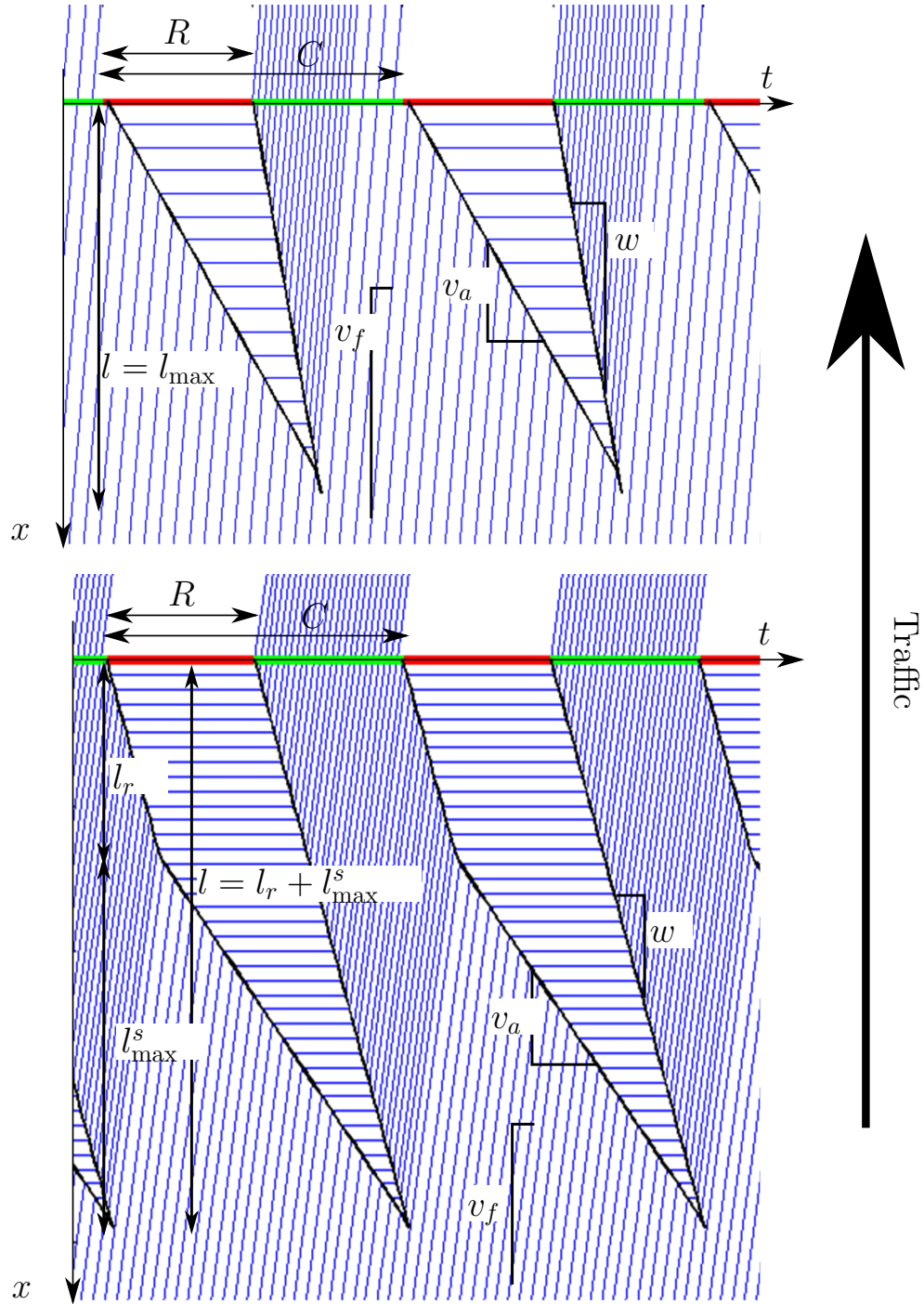
6

Figure 1: Space time diagram of vehicle trajectories with uniform arrivals in an undersaturated traffic regime (top) and a congested traffic regime (bottom).
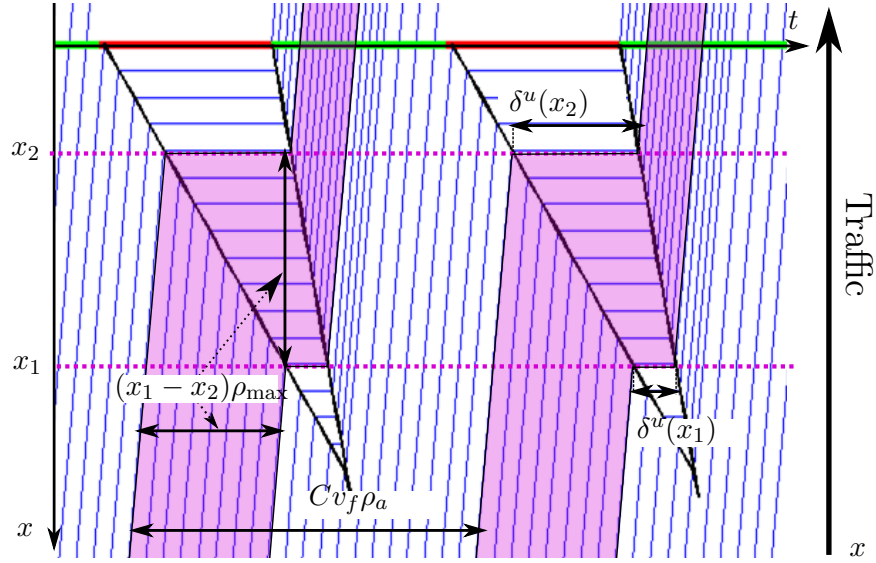
7

Figure 2: The proportion of delayed vehicles $\eta^u_{x_1,x_2}$ is the ratio between the number of vehicles joining the queue between x_1 and x_2 over the total number of vehicles entering the link in one cycle. The trajectories highlighted in purple represent the trajectories of vehicles delayed between x_1 and x_2.

x_1 and x_2 is thus $\eta^u_{x_1,x_2} = (\min(x_1, l_{\max}) - \min(x_2, l_{\max})) \frac{\rho_{\max}}{v_f C \rho_a}$. Multiplying the nominator and denominator by $l_{\max}$, using equation (3) to eliminate ρ_a and equation (4), we have the expression of $\eta^u_{x_1,x_2}$ in terms of the model parameters R, C and $l^s_{\max}$ and the state variable $l = l_{\max}$:

$$\eta^u_{x_1,x_2} = \frac{\min(x_1, l_{\max}) - \min(x_2, l_{\max})}{l_{\max}} \left(\frac{R}{C} + \left(1 - \frac{R}{C} \right) \frac{l_{\max}}{l^s_{\max}} \right).$$

The first factor scales the proportion of stopping vehicles as a function of the measurement locations. The second factor represents the proportion of stopping vehicles if x_1 is upstream of the queue and x_2 is at the intersection. As the queue length $l_{\max}$ tends to zero, the fraction of stopping vehicles tends to R/C. When the queue length increases, the fraction of stopping vehicles increases linearly until it reaches one at saturation ($l_{\max} = l^s_{\max}$).

The stopping time experienced when stopping at x is denoted by $\delta^u(x)$ for the undersaturated regime. Because the arrival of vehicles is homogenous, the delay $\delta^u(x)$ increases linearly with x. At the intersection ($x = 0$), the delay is maximal and equals the duration of the red light R. At the end of the queue ($x = l_{\max}$) and upstream of the queue ($x \geq l_{\max}$), the delay is null. Thus the expression of $\delta^u(x)$ is as follows:

$$\delta^u(x) = R \left(1 - \frac{\min(x, l_{\max})}{l_{\max}} \right).$$

Given that the arrival of vehicles is uniform in time, the distribution of the location where the vehicles reach the queue between x_1 and x_2 is uniform in space. For vehicles reaching the queue between x_1 and x_2, the probability to experience a delay between locations x_1 and x_2 is uniform. The uniform distribution has support $[\delta^u(x_1), \delta^u(x_2)]$, corresponding to the minimum and maximum delay between x_1 and x_2.

8

The stopping time of vehicles between x_1 and x_2 is a random variable with a mixture distribution with two components. The first component represents the vehicles that do not experience any stopping time between x_1 and x_2 (mass distribution in 0), the second component represents the vehicles reaching the queue between x_1 and x_2 (uniform distribution on $[\delta^u(x_1), \delta^u(x_2)]$). We note $\mathbf{1}_A$ the indicator function of set A,

$$\mathbf{1}_A(x) = \left\{ \begin{array}{l} 1 \text{ if } x \in A \\ 0 \text{ if } x \notin A \end{array} \right.$$

The Dirac distribution centered in a, used to represent a mass probability is denoted $\mathrm{Dir}_{\{a\}}(\cdot)$. The pdf of total delay between x_1 and x_2 (Figure 3, left) reads:

$$h^t(\delta_{x_1,x_2}) = (1 - \eta^u_{x_1,x_2})\mathrm{Dir}_{\{0\}}(\delta_{x_1,x_2}) + \frac{\eta^u_{x_1,x_2}}{\delta^u(x_2) - \delta^u(x_1)}\mathbf{1}_{[\delta^u(x_1),\delta^u(x_2)]}(\delta_{x_1,x_2})$$

The cumulative distribution function of total delay $H^t(\cdot)$ reads:

$$H^t(\delta_{x_1,x_2}) = \left\{ \begin{array}{ll} 0 & \text{if} \quad \delta_{x_1,x_2} < 0 \\ (1 - \eta^u_{x_1,x_2}) & \text{if} \quad \delta_{x_1,x_2} \in [0, \delta^u(x_1)] \\ (1 - \eta^u_{x_1,x_2}) + \eta^u_{x_1,x_2}\frac{\delta_{x_1,x_2} - \delta^u(x_1)}{\delta^u(x_2) - \delta^u(x_1)} & \text{if} \quad \delta_{x_1,x_2} \in [\delta^u(x_1), \delta^u(x_2)] \\ 1 & \text{if} \quad \delta_{x_1,x_2} > \delta^u(x_2) \end{array} \right.$$

3.2   Pdf of stopping time in the congested regime

For the congested regime, as for the undersaturated regime, the pdf of stopping time is computed by deriving the delay experienced between x_1 and x_2 for each arrival time in a cycle. The distance traveled by vehicles in the queue in the duration of a light cycle is $l^s_{\max}$. We call n_s the maximum number of stops experienced by the vehicles in the remaining queue between the locations x_1 and x_2:

$$n_s = \left\lceil \frac{\min(x_1, l_r) - \min(x_2, l_r)}{l^s_{\max}} \right\rceil .$$

In this article, we do not model specifically queue spill-over to upstream links. However, this vital component is indirectly taken into account by the flexibility of a statistical model. Indeed, a queue spill-over has the effect to reduce the flow that can exit the upstream links and change the behavior on these links accordingly. The statistical model will automatically learn from the data the new parameters of the dynamics. The delay experienced at location x when reaching the triangular queue at x is readily derived from the expression of the delay in the undersaturated regime, noticing that the delay for x o, the remaining queue is R:

$$\delta^c(x) = \left\{ \begin{array}{ll} R & \text{if } x \leq l_r \\ R\frac{l_r + l^s_{\max} - x}{l^s_{\max}} & \text{if } x \in [l_r, l_r + l^s_{\max}] \\ 0 & \text{if } x \geq l_r + l^s_{\max} \end{array} \right.$$

The details of the derivation are given in [22] (Section 4.3 and Appendix A). We summarize the derivations, classified depending on the location of the positions x_1 and x_2 with respect to the remaining (l_r) and saturation ($l^s_{\max}$) queue lengths. The analytical results are also summarized in Table 1.

1. x_1 *Upstream* $-$ x_2 *Remaining* ($x_1 \geq l_r + l^s_{\max}$, $x_2 \leq l_r$): We define the critical location x_c by $x_c = x_2 + n_s l^s_{\max}$. Vehicles reaching the triangular queue upstream (resp. downstream) of x_c stop n_s (resp. $n_s - 1$) times in the remaining queue on the road segment $[x_1, x_2]$. The vehicles

9

experience a delay uniformly distributed on $[\delta_{\min}, \delta_{\max}]$ with $\delta_{\min} = (n_s - 1)R + \delta^c(x_c)$ and $\delta_{\max} = n_s R + \delta^c(x_c) = \delta_{\min} + R$. The pdf of stopping time reads:

$$h^t(\delta_{x_1, x_2}) = \frac{1}{\delta_{\max} - \delta_{\min}} \mathbf{1}_{[\delta_{\min}, \delta_{\max}]}(\delta_{x_1, x_2}), \qquad \begin{array}{l} \delta_{\min} = \delta^c(x_c) + (n_s - 1)R, \\ \delta_{\max} = \delta^c(x_c) + n_s R \end{array} .$$

2. x_1 *Triangular* – x_2 *Triangular* $(x_1, x_2 \geq l_r)$: Given that the path is upstream of the remaining queue, this case is similar to the undersaturated regime, where derivations are updated to account for the fact that the triangular queue starts at $x = l_r$. We adapt the notation from Section 3.1 and denote by $\eta^c_{x_1, x_2}$ the fraction of the vehicles entering the link in a cycle that experience delay between locations x_1 and x_2.

$$\eta^c_{x_1, x_2} = \frac{\min(x_1 - l_r, l^s_{\max}) - \min(x_2 - l_r, l^s_{\max})}{l^s_{\max}}.$$

This delay is uniformly distributed on $[\delta^c(x_1), \delta^c(x_2)]$. The remainder do not stop between x_1 and x_2. The pdf of stopping time reads:

$$h^t(\delta_{x_1, x_2}) = (1 - \eta^c_{x_1, x_2})\mathrm{Dir}_{\{0\}}(\delta_{x_1, x_2}) + \frac{\eta^c_{x_1, x_2}}{\delta^c(x_2) - \delta^c(x_1)} \mathbf{1}_{[\delta^c(x_1), \delta^c(x_2)]}(\delta_{x_1, x_2}).$$

3. x_1 *Remaining* – x_2 *Remaining* $(x_1, x_2 \leq l_r)$: We define the critical location x_c by $x_c = x_2 + (n_s - 1)l^s_{\max}$. The vehicles reaching the queue between x_1 and x_c stop n_s times in the remaining queue between x_1 and x_2, their stopping time is $n_s R$. The reminder of the vehicles stop $n_s - 1$ times in the remaining queue and their stopping time is $(n_s - 1)R$. The pdf of stopping time reads:

$$h^t(\delta_{x_1, x_2}) = \frac{x_1 - x_c}{l^s_{\max}} \mathrm{Dir}_{\{n_s R\}}(\delta_{x_1, x_2}) + \left(1 - \frac{x_1 - x_c}{l^s_{\max}}\right)\mathrm{Dir}_{\{(n_s - 1)R\}}(\delta_{x_1, x_2}).$$

4. x_1 *Triangular* – x_2 *Remaining* $(x_1 \in [l_r, l_r + l^s_{\max}],\ x_2 \leq l_r)$: We define the critical location x_c by $x_c = x_2 + n_s l^s_{\max}$.

  $\diamond$ If $x_1 \geq x_c$, a fraction $(x_1 - x_c)/l^s_{\max}$ of the vehicles entering the link in a cycle join the triangular queue between x_1 and x_c. They stop once in the triangular queue and n_s times in the remaining queue. Among these vehicles, the stopping time is uniformly distributed on $[\delta^c(x_1) + n_s R,\ \delta^c(x_c) + n_s R]$. A fraction $(x_c - l_r)/l^s_{\max}$ of the vehicles entering the link in a cycle join the triangular queue between x_c and $l^s_{\max}$. Among these vehicles, the stopping time is uniformly distributed on $[\delta^c(x_c) + (n_s - 1)R,\ n_s R]$. The remainder of the vehicles reach the remaining queue between l_r and $x_1 - l^s_{\max}$ and their stopping time is $n_s R$. The pdf of stopping time reads:

$$\begin{aligned} h^t(\delta_{x_1, x_2}) = &\ \frac{x_1 - x_c}{l^s_{\max}} \frac{\mathbf{1}_{[\delta^c(x_1) + n_s R,\ \delta^c(x_c) + n_s R]}(\delta_{x_1, x_2})}{\delta^c(x_c) - \delta^c(x_1)} && \text{stop between } x_1 \text{ and } x_c \\ &+ \frac{x_c - l_r}{l^s_{\max}} \frac{\mathbf{1}_{[\delta^c(x_c) + (n_s - 1)R,\ n_s R]}(\delta_{x_1, x_2})}{R - \delta^c(x_c)} && \text{stop between } x_c \text{ and } l_r \\ &+ \left(1 - \frac{x_1 - l_r}{l^s_{\max}}\right)\mathrm{Dir}_{\{n_s R\}}(\delta_{x_1, x_2}). && \text{stop between } l_r \text{ and } x_1 - l^s_{\max} \end{aligned}$$

  $\diamond$ If $x_1 \leq x_c$, a fraction $(x_1 - l_r)/l^s_{\max}$ of the vehicles entering the link in a cycle join the triangular queue between x_1 and l_r. They stop once in the triangular queue and $n_s - 1$ times in the remaining queue. Among these vehicles, the stopping time is uniformly distributed on $[\delta^c(x_1) + (n_s - 1)R,\ n_s R]$. A fraction $1 - (x_c - l_r)/l^s_{\max}$ of the vehicles entering the link

10

in a cycle join the remaining queue between l_r and $x_c - l^s_{\max}$. The stopping time of these vehicles is $n_s R$. The remainder of the vehicles experiences a stopping time of $(n_s - 1)R$. The pdf of stopping time reads:

$$
\begin{aligned}
h^t(\delta_{x_1,x_2}) = {} & \frac{x_1 - l_r}{l^s_{\max}} \frac{\mathbf{1}_{[\delta^c(x_1)+(n_s-1)R,\, n_s R]}(\delta_{x_1,x_2})}{R - \delta^c(x_1)} && \text{stop between } x_1 \text{ and } l_r \\
& + \left(1 - \frac{x_c - l_r}{l^s_{\max}}\right) \mathrm{Dir}_{\{n_s R\}}(\delta_{x_1,x_2}) && \text{stop between } l_r \text{ and } x_c - l^s_{\max} \\
& + \frac{x_c - x_1}{l^s_{\max}} \mathrm{Dir}_{\{(n_s-1)R\}}(\delta_{x_1,x_2}). && \text{stop between } x_c - l^s_{\max} \text{ and } x_1 - l^s_{\max}
\end{aligned}
$$

4 Probability distributions of travel times and estimation from sparsely sampled probe vehicles

4.1 Travel time distributions

Along the path between x_1 and x_2, the travel time y_{x_1,x_2} is a random variable, written as the sum of two independent random variables: the delay δ_{x_1,x_2} experienced between x_1 and x_2 and the free flow travel time of the vehicles $y_{f;\,x_1,x_2}$. The free flow travel time is proportional to the distance of the path and the free flow pace p_f such that $y_{f;\,x_1,x_2} = p_f(x_1 - x_2)$. We have $y_{x_1,x_2} = \delta_{x_1,x_2} + y_{f;\,x_1,x_2}$.

We model the differences in traffic behavior by considering the free flow pace p_f as a random variable with distribution φ^p and domain of definition $\mathcal{D}_{\varphi^p}$. For convenience, we define the extension of φ as being zero on $\mathbb{R}\backslash\mathcal{D}_{\varphi^p}$ (the complement of the domain of definition), which do not change the distribution of the random variable p_f. With a slight abuse of notation, we still call this extension φ.

Using a linear change of variables, we derive the probability distribution $\varphi^y_{x_1,x_2}$ of free flow travel time $y_{f;\,x_1,x_2}$ between x_1 and x_2:

$$
p_f \sim \varphi^p(p_f) \Rightarrow \varphi^y_{x_1,x_2}(y_{f;\,x_1,x_2}) = \varphi^p\left(\frac{y_{f;\,x_1,x_2}}{x_1 - x_2}\right) \frac{1}{x_1 - x_2}
$$

To derive the pdf of travel times we use the following fact:

Fact 1 (**Sum of independent random variables**). *If X and Y are two independent random variables with respective pdf f_X and f_Y, then the pdf f_Z of the random variable $Z = X + Y$ is given by the convolution product of f_X and f_Y, denoted $f_Z(z) = (f_X * f_Y)(z)$ and defined as $f_Z(z) = \int_{\mathbb{R}} f_X(t) f_Y(z - t)\, dt$.*

This classical result in probability is derived by computing the conditional pdf of Z given X and then integrating over the values of X according to the total probability law.

For each regime $s \in \{u, c\}$, the probability distribution of travel times reads:

$$
g^s(y_{x_1,x_2}) = \left(h^s * \varphi^y_{x_1,x_2}\right)(y_{x_1,x_2}).
$$

We derive the general expression of the travel time distributions when vehicles experience a delay with mass probability in Δ and when vehicles experience a delay with uniform distribution on $[\delta_{\min}, \delta_{\max}]$. These expressions enable the computation of the pdf of travel times using the linearity of the convolution operator.

Travel time distribution when the delay has a mass probability in Δ:

11

Case	Trajectories	Weight	Dist.	Support
<u>Case 1</u> $x_1 \geq l_r + l^s_{\max}$, $x_2 \leq l_r$, $x_c = x_2 + n_s l^s_{\max}$	All	1	Unif.	$[(n_s - 1)R + \delta^c(x_c),$ $n_s R + \delta^c(x_c)]$
<u>Case 2</u> $x_1 \geq l_r$, $x_2 \geq l_r$	No stop between x_1 and x_2	$\mathcal{P}(\bar{s}_{x_1}, \bar{s}_{x_2}) \times$ $(1 - \eta^c_{x_1,x_2})$	Mass	$\{0\}$
	Reach the (triangular) queue between x_1 and x_2	$\mathcal{P}(\bar{s}_{x_1}, \bar{s}_{x_2}) \times$ $\eta^c_{x_1,x_2}$	Unif.	$[\delta^c(x_2), \delta^c(x_1)]$
<u>Case 3</u> $x_1 \leq l_r$, $x_2 \leq l_r$, $x_c = x_2 + (n_s - 1)l^s_{\max}$	Reach the (remaining) queue between x_1 and x_c	$\mathcal{P}(\bar{s}_{x_1}, \bar{s}_{x_2}) \times$ $\dfrac{x_1 - x_c}{l^s_{\max}}$	Mass	$\{n_s R\}$
	Reach the (remaining) queue between x_c and $x_1 - l^s_{\max}$	$\mathcal{P}(\bar{s}_{x_1}, \bar{s}_{x_2}) \times$ $\dfrac{x_c - x_1 + l^s_{\max}}{l^s_{\max}}$	Mass	$\{(n_s - 1)R\}$
<u>Case 4a</u> $x_1 \in [l_r, l_r + l^s_{\max}]$, $x_2 \leq l_r$, $x_c = x_2 + n_s l^s_{\max}$, $x_c \leq x_1$	Reach the (triangular) queue between x_1 and x_c	$\mathcal{P}(\bar{s}_{x_1}, \bar{s}_{x_2}) \times$ $\dfrac{x_1 - x_c}{l^s_{\max}}$	Unif.	$[n_s R + \delta^c(x_1),$ $n_s R + \delta^c(x_c)]$
	Reach the (triangular) queue between x_c and l_r	$\mathcal{P}(\bar{s}_{x_1}, \bar{s}_{x_2}) \times$ $\dfrac{x_c - l_r}{l^s_{\max}}$	Unif.	$[(n_s - 1)R + \delta^c(x_c),$ $n_s R]$
	Reach the (remaining) queue between l_r and $x_1 - l^s_{\max}$	$\mathcal{P}(\bar{s}_{x_1}, \bar{s}_{x_2}) \times$ $\dfrac{l_r - x_1 + l^s_{\max}}{l^s_{\max}}$	Mass	$\{n_s R\}$
<u>Case 4b</u> $x_1 \in [l_r, l_r + l^s_{\max}]$, $x_2 \leq l_r$, $x_c = x_2 + n_s l^s_{\max}$, $x_c \geq x_1$	Reach the (triangular) queue between x_1 and l_r	$\mathcal{P}(\bar{s}_{x_1}, \bar{s}_{x_2}) \times$ $\dfrac{x_1 - l_r}{l^s_{\max}}$	Unif.	$[(n_s - 1)R + \delta^c(x_1),$ $n_s R]$
	Reach the (remaining) queue between l_r and $x_c - l^s_{\max}$	$\mathcal{P}(\bar{s}_{x_1}, \bar{s}_{x_2}) \times$ $\dfrac{l_r - x_c + l^s_{\max}}{l^s_{\max}}$	Mass	$\{n_s R\}$
	Reach the (remaining) queue between $x_c - l^s_{\max}$ and $x_1 - l^s_{\max}$	$\mathcal{P}(\bar{s}_{x_1}, \bar{s}_{x_2}) \times$ $\dfrac{x_c - x_1}{l^s_{\max}}$	Mass	$\{(n_s - 1)R\}$

Table 1: The pdf of measured delay is a mixture distribution. The different components and their associated weight depend on the location of stops of the vehicles with respect to the queue length and sampling locations.

12

The stopping time is Δ. This corresponds to trajectories with n_s stops ($n_s \geq 0$) in the remaining queue. This includes the non stopping vehicle in the undersaturated regime, when the remaining queue has length zero. The travel time distribution is derived as

$$
\begin{aligned}
g(y_{x_1,x_2}) &= \left(\mathrm{Dir}_{\{\Delta\}} * \varphi^y_{x_1,x_2}\right)(y_{x_1,x_2}) \\
&= \varphi^y_{x_1,x_2}(y_{x_1,x_2} - \Delta).
\end{aligned}
\tag{5}
$$

Travel time distribution when the delay is uniformly distributed on $[\delta_{\min}, \delta_{\max}]$:

Vehicles experience a uniform delay between a minimum and maximum delay respectively denoted $\delta_{\min}$ and $\delta_{\max}$. The probability of observing a travel time y_{x_1,x_2} is given by

$$
g(y_{x_1,x_2}) = \frac{1}{\delta_{\max} - \delta_{\min}} \int_{-\infty}^{+\infty} \mathbf{1}_{[\delta_{\min}, \delta_{\max}]}(y_{x_1,x_2} - z)\, \varphi^y_{x_1,x_2}(z)\, dz.
\tag{6}
$$

The integrand is not null if and only if $y_{x_1,x_2} - z \in [\delta_{\min}, \delta_{\max}]$, *i.e.* if and only if $z \in [y_{x_1,x_2} - \delta_{\max}, y_{x_1,x_2} - \delta_{\min}]$. Since $\varphi^y_{x_1,x_2}(z)$ is equal to zero for $z \in \mathbb{R} \setminus \mathcal{D}_\varphi$, the integrand is not null if and only if $z \in [y_{x_1,x_2} - \delta_{\max}, y_{x_1,x_2} - \delta_{\min}] \bigcap \mathcal{D}_\varphi$.

As an illustration, we derive the probability distribution of travel times on a partial link in the undersaturated regime, for a pace distribution with *support* on $\mathbb{R}^+$. In this article, we denote by *support* the set of points where the function is not zero. We write the delay distribution as a mixture of mass probabilities and uniform distributions with two components. The first component, with weight $1 - \eta^u_{x_1,x_2}$, represents the delay distribution of the vehicles which do not stop between x_1 and x_2. It is a mass distribution in 0. The second component, with weight $\eta^u_{x_1,x_2}$, represents the delay distribution of the vehicles who do stop between x_1 and x_2. It is a uniform distribution with support $[\delta^u(x_1), \delta^u(x_2)]$. The probability distribution of delay is illustrated Figure 3 (left) with different line styles for the two components. We use the linearity of the convolution to convolve each component of the mixture with the probability distribution of free flow travel times $\varphi^y_{x_1,x_2}$. The pdf of travel times are computed according to (5) and (6) for the non-stopping and the stopping vehicles respectively and are illustrated Figure 3 (center). We sum the pdf of travel times of the non-stopping and the stopping vehicles with their respective weights ($1 - \eta^u_{x_1,x_2}$ and $\eta^u_{x_1,x_2}$) to obtain the pdf of travel times between locations x_1 and x_2 for all the vehicles entering the link in a cycle (Figure 3, right). The probability distribution of travel times reads:

$$
g^u(y_{x_1,x_2}) = \begin{cases}
0 & \text{if } y_{x_1,x_2} \leq 0, \\[4pt]
(1 - \eta^u_{x_1,x_2})\varphi^y_{x_1,x_2}(y_{x_1,x_2}) & \text{if } y_{x_1,x_2} \in [0, \delta^u(x_1)], \\[4pt]
\left\{ \begin{aligned} &(1 - \eta^u_{x_1,x_2})\varphi^y_{x_1,x_2}(y_{x_1,x_2}) \\ &+ \frac{\eta^u_{x_1,x_2}}{\delta^u(x_2) - \delta^u(x_1)} \int\limits_{0}^{y_{L,0} - \delta^u(x_1)} \varphi^y_{x_1,x_2}(z)\, dz \end{aligned} \right\} & \text{if } y_{x_1,x_2} \in [\delta^u(x_1), \delta^u(x_2)], \\[12pt]
\left\{ \begin{aligned} &(1 - \eta^u_{x_1,x_2})\varphi^y_{x_1,x_2}(y_{x_1,x_2}) \\ &+ \frac{\eta^u_{x_1,x_2}}{\delta^u(x_2) - \delta^u(x_1)} \int\limits_{y_{L,0} - \delta^u(x_2)}^{y_{L,0} - \delta^u(x_1)} \varphi^y_{x_1,x_2}(z)\, dz \end{aligned} \right\} & \text{if } y_{x_1,x_2} \geq \delta^u(x_2).
\end{cases}
$$

The derivations are similar in the congested regime: we convolve each component of the stopping time distribution with the pdf of free flow travel times $\varphi^y_{x_1,x_2}$. We recall that for the different cases described in Section 3.2, the delay is a mixture of mass probabilities and uniform distributions and thus either equation (5) or (6) is used on each component.
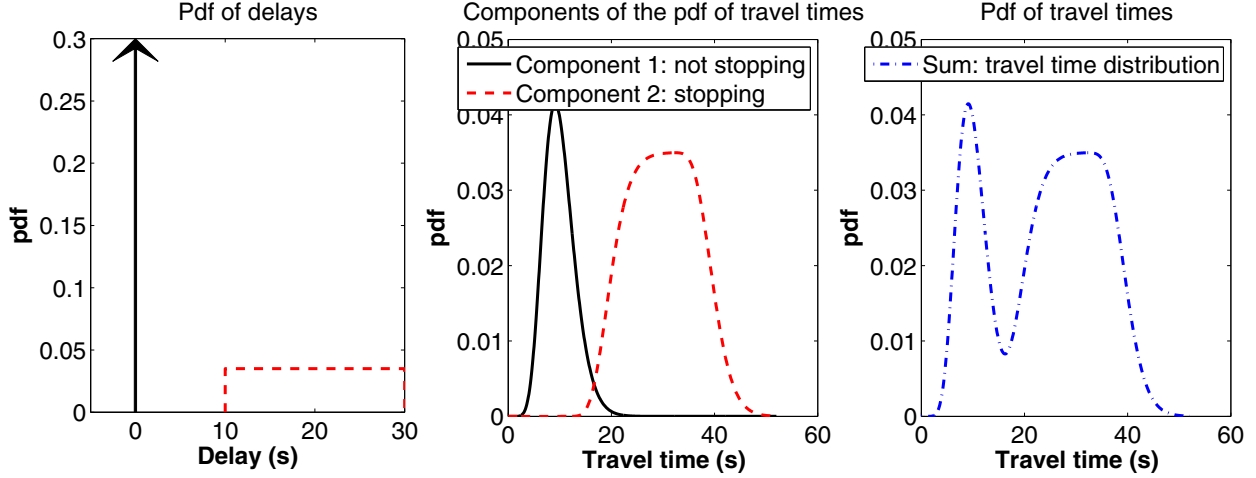
13

Figure 3: Probability distributions of travel times between arbitrary locations in the undersaturated regime. The figure represents the pdf of travel times between x_1 and x_2 where both x_1 and x_2 are in the triangular queue, with $\delta^u(x_1) = 10s$ and $\delta^u(x_2) = 30s$ and $\eta^u_{x_1,x_2} = 0.7$. The free flow travel time between x_1 and x_2 is a random variable with Gamma distribution. The mean free flow travel time is 10 seconds and the standard deviation is 3s.

For example, the probability distribution of link travel times (Case 1, $x_1 = L$ length of the link, $x_2 = 0$) is computed via equation (6) and reads

$$
g^c(y_{L,0}) = \begin{cases} 0 & \text{if} \quad y_{L,0} \leq \delta_{\min}, \\ \frac{1}{\delta_{\max}-\delta_{\min}} \int_0^{y_{L,0}-\delta_{\min}} \varphi^y_{L,0}(z)dz & \text{if} \quad y_{L,0} \in [\delta_{\min}, \delta_{\max}], \\ \frac{1}{\delta_{\max}-\delta_{\min}} \int_{y_{L,0}-\delta_{\max}}^{y_{L,0}-\delta_{\min}} \varphi^y_{L,0}(z)\,dz & \text{if} \quad y_{L,0} \geq \delta_{\max}. \end{cases}
$$

with $\delta_{\min} = \delta^c(n_s l^s_{\max}) + (n_s - 1)R$, $\delta_{\max} = \delta^c(n_s l^s_{\max}) + n_s R$ and $n_s = \left\lceil \frac{l_r}{l^s_{\max}} \right\rceil$.

4.2 Learning traffic conditions from sparsely sampled probe vehicles

From traffic flow theory, we derived a probability distribution of travel times between arbitrary locations on an arterial link. These distributions are parameterized by the network parameters (average red time R, average cycle time C, driving behavior θ_p and saturation queue length $l^s_{\max}$) and the level of congestion represented by the queue length $l_{\max}$. As probe vehicles report their location periodically in time, the duration between two successive location reports x_1 and x_2 represents a measurement of the travel time of the vehicle on its path from x_1 to x_2. We use these travel time observations from probe vehicles to learn the parameters of the travel time distributions.

Common sampling rates for probe vehicles are around one minute and probe vehicles typically traverse several links between successive location reports. It is possible to optimally decompose the path travel time to estimate the travel time spent on each link of the path [20]. In this article, we assume that this decomposition has already been achieved and we focus on the estimation of the pdf of travel times. Since probe vehicles may report their location at any point x_1 and x_2, they provide partial link travel time measurements that allow for the estimation of the independent parameters of each link: the red time R, the queue length $l_{\max}$, the fraction of stopping vehicles on the link among the vehicles entering the link in one cycle $\eta^u_{L,0}$ and the driving behavior θ_p. The estimation of the parameters of link i is done by maximizing the likelihood (or more conveniently the log-likelihood) of the (partial) link travel times of this link with respect to these parameters. Note that the parameters of the travel time distribution (R^i,

14

$\eta_{L,0}^u$, l^i and θ_p^i) do not depend on the locations x_1 and x_2 of measurement j. In particular, we learn the travel time distributions using travel time measurements which span different portions of the link, *i.e.* the locations x_1 and x_2 depend on the index of the measurement (j), even though we do not explicit this dependency for notational simplicity. Let $(y_{x_1,x_2}^j)_{j=1:J^i}$ represent the set of (partial link) travel times allocated to link i. The estimation problem is given by:

$$\underset{R^i,\eta_{L,0}^u,l^i,\theta_p^i}{\text{minimize}} \sum_{j=1}^{J^i} -\ln(g^i(y_{x_1,x_2}^j)) \tag{7}$$

$$\textbf{s.t. } \eta_{L,0}^u \in [0,1], \ l^i \in [0,L^i].$$

Additional constraints and bounds may be added to limit the feasible set to physically acceptable values of the parameters and improve the estimation when little data is available. The optimization problem (7) is not convex but it is a small scale optimization problem (feasible set of dimension five). Numerous optimization techniques can be used to solve this problem including global optimization algorithms [23, 37]. Moreover, since the parameters represent physical parameters, they can be bounded to limit the feasible set to a compact set (of dimension five). It is thus possible to do a *grid search*. The grid search algorithm defines a grid on the bounded feasible set and evaluates the objective function for each set of parameters defined by the grid. We keep the B best set of parameters, associated with the lowest values of the objective function and perform a first or a second order optimization algorithm [10] from this best set of parameters. In the implementation of the algorithm used to produce the results of Section 5, we set $B = 4$ and used the `active-set` algorithm in the Matlab [1] optimization toolbox, which is a second order optimization algorithm based on *Sequential Quadratic Programming* [9].

5   Numerical experiments and results

The model presented in this article relies on assumptions on the dynamics of traffic flows on each link of the network to derive probability distributions of travel times. The goal of this section is to show numerically that these travel time distributions, derived from the physics of traffic flows, represent the empirical distribution of travel times more accurately than classical distributions such as normal, log-normal or Gamma distributions.

We consider four classes of distributions: the *traffic* distribution derived in this article, the *normal* distribution, the *log-normal* distribution and the *Gamma* distribution. For each class of distributions, we test the hypothesis that link travel times are distributed according to this distribution (the complementary hypothesis is that the travel times are *not* distributed according to this distribution). We use data collected during a *field experiment* from the 29th of June to the 1st of July 2010 as part of the *Mobile Millennium* project [2, 7]. Twenty drivers, each carrying a GPS device, drove for 3 hours (3:15pm to 6:15pm) around two distinct loops in San Francisco. The first loop was 1.89 miles long and the second one 2.31 miles long. The GPS devices recorded the location of the vehicles every second and provided detailed information on the trajectories of the drivers. From this detailed data, we extract link travel times.

For each link of the network, we compute the maximum likelihood estimates of the distribution parameters for each class of distributions. This learning of distribution parameters is performed using a fraction of the link travel times collected by the drivers. We vary the percentage of available data used for the training of the distributions to study the influence of the amount of data required to learn the parameters accurately. For each class of distribution, we test the hypothesis H_0: *the link travel times are distributed according to the distribution* on the validation link travel times using the *Kolmogorov-Smirnov* test [28], also referred to as K-S test. The K-S test is a standard non-parametric test to state whether samples are distributed
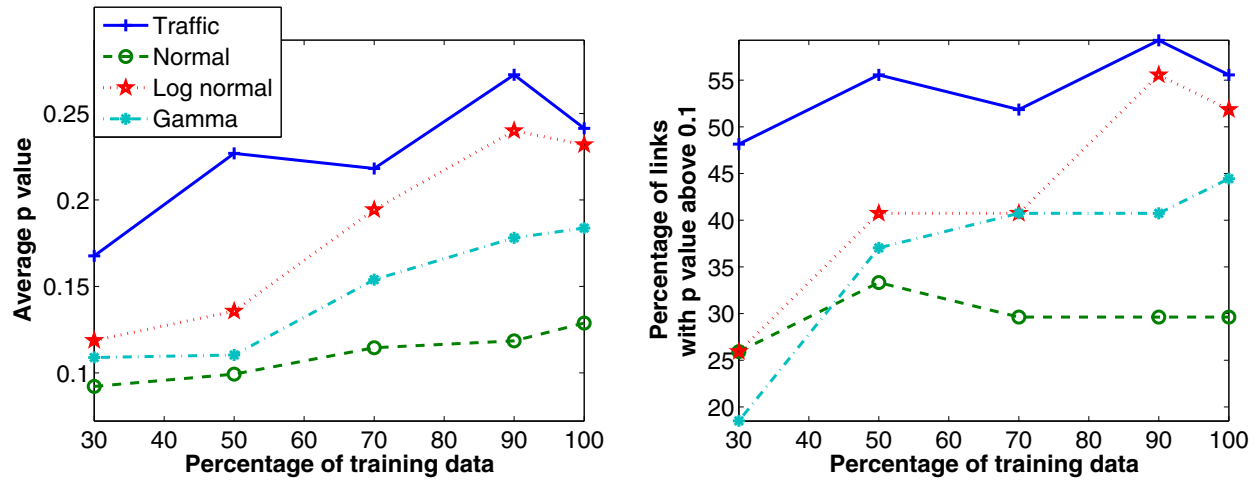
15

Figure 4: Experimental set-up of the field experiment representing the two distinct loops in downtown San Francisco, CA.

according to an hypothetical distribution (in opposition to other tests like the T-test that tests uniquely the mean, or the chi-squared test that assumes that the data is normally distributed). The test is based on the K-S statistics which is computed as the maximum difference between the empirical and the hypothetical cumulative distributions. The test provides a p-value which informs us on the goodness of the fit. Low p-values indicate that the data does not follow the hypothetical distribution. For each hypothetical distribution, Figure 5 (left) shows the average p-value of the links of the network as the percentage of training data increases. The hypothesis H_0 is rejected for p-values inferior to the significance level α. The significance level α corresponds to the percentage of Type-I error allowed by the test (rejecting the null hypothesis when it is actually true). Figure 5 (right) shows the evolution of the percentage of links that passes the K-S test at significance level $\alpha = 0.1$. Both figures show that the traffic distribution represents a better fit of the travel time distributions than any of the other distributions tested in this article. We notice that the relative superiority of the traffic model is more significant when little data is available. This may be a sign of the robustness of the model when little data is available (because of the intrinsic structure of the distributions representing the physical model). This is precisely the goal of the algorithm (and model), which was specifically created to handle low volumes of probe data. We also notice that the log-normal model performs well compared to the normal or the Gamma distribution.

We analyze the differences of the traffic and the log-normal by studying the empirical distribution of travel times. Figure 6 (left) shows the hypothetical and empirical distribution of travel times. The traffic distribution captures the specific characteristics of traffic dynamics. We can see a peak in the distribution representing the vehicles that do not stop on the link and travel at their free flow speed. For higher travel times, the distribution is approximately uniform, representing the vehicles that are delayed on the link, between a minimum delay (0) and a maximum delay (the duration of the red time). As for the log-normal distribution, it cannot capture these specifics of the travel time distribution and the parameters are harder to interpret.

Due to light synchronization, some links have arrivals with platoons, and thus do not follow the hypothesis of constant arrivals. On these links, delays are not uniformly distributed among

16

Figure 5: Goodness of fit of the model depending on the percentage of training data used to learn the parameters. (**Left**) Average p-value of the links of the network for the different hypothetical distributions (traffic, normal, log-normal and Gamma). (**Right**) Percentage of links that pass the KS test with a significance level $\alpha = 0.1$.
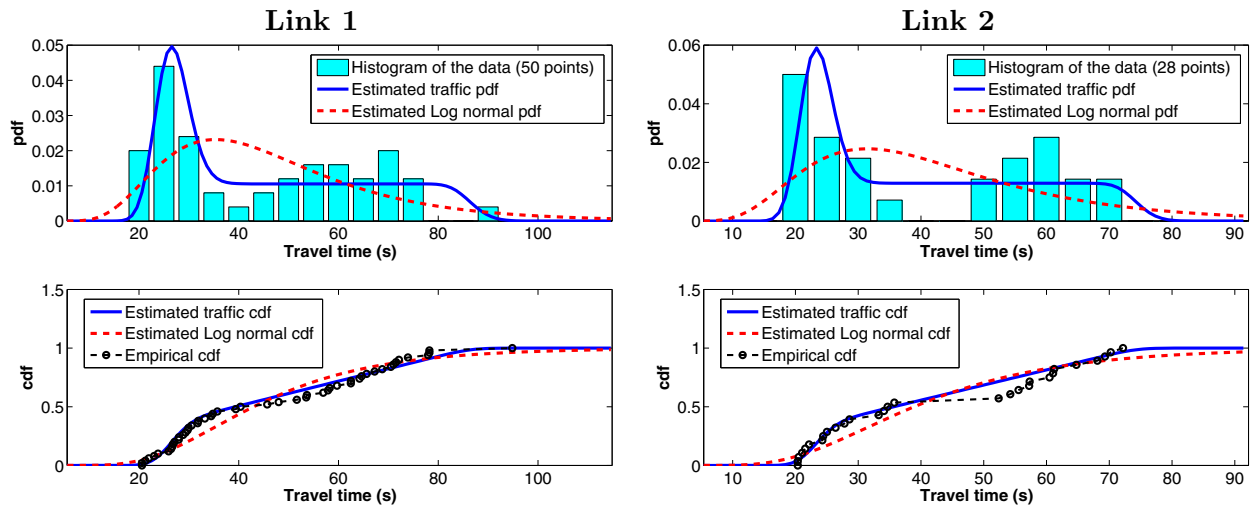

Figure 6: Comparison of the traffic and the log-normal distributions with the empirical distribution of travel times on two links of the network. The figure represents both the pdf and the cdf of the traffic (solid blue line) and log-normal (dashed red line) distributions. The histograms on the top figures represent interval counts of the probe travel times, normalized so that the area of the histogram sums to one. The cumulated histograms (bottom figures) are the cumulated distributions of the histograms. The black line with circles represents the *empirical cumulative distribution* (Kaplan-Meier estimate [24]) of the travel times collected by the probes. (**Left, link 1**) Both distributions capture the long tail of the distribution but only the traffic distribution is able to represent the peak in the pdf due to the non stopping vehicles and to estimate accurately the maximum delay. (**Right, link 2**) On this link, we notice very few travel times between 35 and 50 seconds, likely due to important synchronization with the upstream link. None of the traffic or log-normal distribution is able to capture this. However, the traffic distribution models accurately the peak due to the non stopping vehicles and estimate the maximum delay.

<center>17</center>

the stopping vehicles and the derivations of the queuing model have to be adapted [4]. Basically, the delay function $\delta^r(x)$ ($r \in \{u, c\}$, see Section 3) is piecewise linear and the derivations of the statistical distributions must be updated accordingly, adding parameters to the model. Figure 6 (right) represents the empirical and hypothetical distribution of travel times for a link with platoon arrivals. We can see that there are very few vehicles with a travel time between 30 and 50 seconds, representing a time interval during which there is very few arrivals on the links, likely when the upstream signal is red. We notice that the log-normal distribution does not capture this characteristics of the distribution either. Moreover, the traffic model provides an estimation of the red time, the free flow speed and the fraction of stopping vehicles (representing congestion) which is important information for traffic management and operations.

6   Conclusion

In this article, we derived a parametric probability distribution of travel times between arbitrary locations on an arterial link. This probability distribution is derived from hydrodynamic theory and represents the dynamics of traffic flow on arterial links. In particular, it captures the delay of vehicles due to the presence of a queue that forms and dissipates periodically because of the traffic signal. These distributions are parameterized by physical parameters: the red time, the cycle time, the parameters of the free flow pace, the queue length and the saturation queue length. Depending on the data available, these parameters may not be estimated independently, but we can always retrieve the duration of the red time, the level of congestion and the parameters of the free flow pace . The queue length can also be estimated from probe vehicles reporting their location at any location on the link.

The goodness of fit of the distributions was tested on probe data collected during a field test in San Francisco. The numerical results show the superiority of the traffic distribution to represent the distribution of travel times compared to "classic" distributions (normal, nognormal and Gamma distributions), commonly used to represent the distribution of travel times. The traffic distribution performs particularly well (in comparison with the other distributions) when little data is available.

The numerical analysis shows that the uniform arrivals is the most restrictive assumption on which this work is based, as it does not take into account signal synchronization. We are currently working on a generalization of the proposed approach in which vehicles arrive in platoons of homogeneous density. Note that this generalization does not invalidate the methodology presented in this article. In particular, the probability distribution of stopping times will remain a mixture of discrete mass probabilities and uniform distributions.

The probability distribution of travel times are finite mixture distributions [20]. Each component of the mixture corresponds to a type of delay: stopping or not stopping for the undersaturated regime or depending on the location of the vehicle (Table 1) in the congested regime. The estimation of transition probabilities representing the probability of a type of delay on a link given the type of delay on the upstream link would allow to compute route travel time distributions with a Markov chain approach.

Acknowledgements

References

[1] MATLAB. http://mathworks.com.

[2] The *Mobile Millennium* Project. http://traffic.berkeley.edu.

[3] *Highway Capacity Manual*. TRB, National Research Council, Washington, D.C., 2000.

[4] R. E. Allsop. An analysis of delays to vehicle platoons at traffic signals. In *Procedings of the fourth international symposium on the Theory of Traffic Flow, University of Karlsruhe, Germany*, 1968.

[5] R. E. Allsop. Delay at fixed time traffic signals-I: Theoretical analysis. *Transportation Science*, 6:280–285, 1972.

[6] X. Ban, R. Herring, P. Hao, and A. Bayen. Delay pattern estimation for signalized intersections using sampled travel times. In *Proceedings of the 88th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2009.

[7] A. Bayen, J. Butler, and A. Patire et al. Mobile Millennium final report. Technical report, University of California, Berkeley, CCIT Research Report UCB-ITS-CWP-2011-6, To appear in 2011.

[8] D. S. Berry and D. M. Belmont. Distribution of vehicle speeds and travel times. *Proc. 2nd Berkeley Sympos. Math. Statist. Probab.*, pages 589–602, 1951.

[9] P.T. Boggs and J.W. Tolle. Sequential quadratic programming. *Acta numerica*, 4(1):51, 1995.

[10] S.P. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge Univ Pr, 2004.

[11] M. S. Van Den Broek, J. S. H. Van Leeuwaarden, I. Adan, and O. J. Boxma. Bounds and approximations for the fixed-cycle traffic-light queue. *Transportation Science*, 40(4):484496, 2006.

[12] C. Daganzo. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research B*, 28(4):269–287, 1994.

[13] C. Daganzo and N. Geroliminis. An analytical approximation for the macroscopic fundamental diagram of urban traffic. *Transportation Research Part B: Methodological*, 42(9):771–781, 2008.

[14] J. N. Darroch. On the traffic-light queue. *The Annals of Mathematical Statistics*, 35(1):pp. 380–388, 1964.

[15] L. C. Evans. *Partial Differential Equations*. Graduate Studies in Mathematics, V. 19. American Mathematical Society, Providence, RI, 1998.

[16] D. Fambro and N. Rouphail. Generalized delay model for signalized intersections and arterial streets. *Transportation Research Record: Journal of the Transportation Research Board*, 1572:112–121, January 1997.

[17] N. Geroliminis and C. Daganzo. Macroscopic modeling of traffic in cities. In *Proceedings of the 86th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2007.

[18] D. Heidemann. Queue length and delay distributions at traffic signals. *Transportation Research Part B: Methodological*, 28(5):377–389, 1994.

[19] J. Herrera, D. Work, R. Herring, X. Ban, Q. Jacobson, and A. Bayen. Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment. *Transportation Research Part C: Emerging Technologies*, 18(4):568–583, August 2010.

[20] A. Hofleitner and A. Bayen. Optimal decomposition of travel times measured by probe vehicles using a statistical traffic flow model. *IEEE Intelligent Transportation System Conference (IEEE ITSC '11)*, October 2011.

[21] A. Hofleitner, L. El Ghaoui, and A. Bayen. Online least-squares estimation of time varying systems with sparse temporal evolution and application to traffic estimation. *50th Conference on Decision and Control (CDC 2011)*.

[22] A. Hofleitner, R. Herring, and A. Bayen. A hydrodynamic theory based statistical model of arterial traffic. *Technical Report UC Berkeley, UCB-ITS-CWP-2011-2*, *http://www. eecs. berkeley. edu/~aude/papers/traffic_ distributions. pdf*, January 2011.

[23] R. Horst and H. Tuy. Global optimization: deterministic approaches. *Journal of the Operational Research Society*, 45(5):595–596, 1994.

[24] E. L. Kaplan and Paul Meier. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282):pp. 457–481, 1958.

[25] J. S. H. Van Leeuwaarden. Delay analysis for the fixed-cycle traffic-light queue. *Transportation Science*, 40(2):189199, 2006.

[26] M. Lighthill and G. Whitham. 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, May 1955.

[27] H. K. Lo, E. Chang, and Y. C. Chan. Dynamic network traffic control. *Transportation Research Part A: Policy and Practice*, 35(8):721–744, 2001.

[28] F. J. Massey. The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association*, 46(253):68–78, 1951.

[29] A. J. Miller. Settings for Fixed-Cycle traffic signals. *Operations Research*, 14(4), December 1963.

[30] K. Ohno. Computational algorithm for a fixed cycle traffic signal and new approximate expressions for average delay. *Transportation Science*, 12(1):29–47, 1978.

[31] P. Richards. Shock waves on the highway. *Operations Research*, 4(1):42–51, February 1956.

[32] A. Skabardonis and N. Geroliminis. Real-time estimation of travel times on signalized arterials. In *Proceedings of the 16th International Symposium on Transportation and Traffic Theory*, University of Maryland, College Park, MD, July 2005.

[33] TTI. Texas Transportation Institute: Urban Mobility Information: 2007 Annual Urban Mobility Report. http://mobility.tamu.edu/ums/, 2007.

[34] F. Viti and H. J. Van Zuylen. The dynamics and the uncertainty of queues at fixed and actuated controls: A probabilistic approach. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, 13(1), 2009.

[35] F. V. Webster. *Traffic signal settings*. Road Research Technical Paper No. 39, Road Research Laboratory, England, published by HMSO, 1958.

[36] D. B. Work, S. Blandin, O-P Tossavainen, B. Piccoli, and A. M. Bayen. A traffic model for velocity data assimilation. *Applied Mathematics Research eXpress*, 2010(1):1–35, 2010.

[37] A. Zhigljavsky and A. Zilinskas. *Stochastic global optimization*. Springer, 2007.

20