Macroscopic modelling and analysis of rush-hour congestion

Dieter Fiems, Balakrishna Prabhu

Macroscopic modelling and analysis of rush-hour congestion

Dieter Fiems[1] and Balakrishna Prabhu[2]
[1]Ghent University, Dep. TELIN, Ghent, Belgium
[2]LAAS-CNRS, Université de Toulouse, CNRS, Toulouse, France

Abstract

We consider a Markovian queueing model for computing the traffic density and travel times in a city at a macroscopic scale during rush hour. Accounting for the speed/density relation of the macroscopic fundamental diagram of traffic flow, we assume that the service rates of the queueing model at hand are state-dependent. We focus on the fluid limit and obtain a set of differential equations that describe the evolution of the traffic density at the level of neighbourhoods. We also calculate the time dependent travel times and consider the rational time-dependent choice between public and private transport, assuming that there is a congestion-free public alternative to private transportation. Numerical examples reveal that a small reduction in peak traffic can significantly reduce the average travel times.

1 Introduction

Most major cities suffer from traffic congestion during the morning and evening commutes. Apart from a waste of time, congestion is a major source of pollution and a health hazard [15]. Several types of remedial measures have been adopted by authorities in various cities to reduce congestion and its impact on the society. These measures take the form of a congestion charge like is the case in cities such as London or Stockholm, or, as is done in France, a restriction on the entry of old or more polluting vehicles during episodes of peak pollution.

Another mechanism for reducing congestion at bottlenecks is to take alternative paths. This has now become possible in real-time thanks to several online applications that can orient drivers towards paths of least travel times.

A third mechanism is to incentivise public transport. For such schemes to work, the public transport network has to be dense, reliable, and, more importantly, better in terms of travel times and cost. A properly dimensioned public transport network can provide an eco-friendly alternative to personal vehicle usage.

For the second and third mechanisms to work, it is important to be able to accurately estimate the travel times between major hubs. If a user can see in real-time that it will be faster (or less costly) to travel by public transport, it will give an additional incentive to use the public transport network. Further, if we assume that users are rational and choose the alternative with the least cost, we can estimate how many users will take the public transport. Such estimates can then be used to dimension the public transport network by appropriately choosing the frequencies and the vehicle sizes on different bus routes.

1.1 Contributions

The main contribution of this paper is a queueing-theory-based method for the computation of travel times in road traffic networks. While the link between queueing theory and road traffic networks has been made previously [17, 4, 2], it was only recently that this connection was used to compute travel times at the level of a single neighbourhood [7]. In this paper, we generalise the travel-time analysis of [7] to a network of neighbourhoods. In queueing theory parlance, this paper generalises the single server analysis in [7] to a network of queues. This is similar to how Jackson networks [10] or BCMP networks [1] extended the single server models.

As in [7], based on the travel-time analysis, we also study a public transport game in which commuters can choose between using their own vehicles or taking public transport. The main assumption in this game is that public transport networks are not impacted by congestion and their travel times are thus easier to estimate. This assumption is certainly true for subways and tramways, and is becoming more realistic for buses in cities with dedicated bus corridors. For a given fixed cost of travel by the public transport network, the time-varying Wardrop equilibrium is computed in order to determine the fraction of commuters in each neighbourhood that will switch from private vehicles to public transport. One can then compute the influence of pricing on the congestion and, as a consequence, on the travel times and the pollution. In addition, for a certain set of parameters, we also numerically compute the Price of Anarchy (PoA) which is the ratio of the social cost at the Wardrop equilibrium to the optimum social cost. For this, an optimal control problem is first formally defined. Since this problem does not have a closed-form solution, the PoA is obtained by using numerical solvers.

Finally, we show how to compute a congestion-based cost of the public transport network so that the Wardrop equilibrium matches the social optimum in the fixed cost setting. For time-invariant congestion games, it is known that tolls or taxes can be imposed to achieve a socially optimum behaviour in a game setting (see, e.g., [5]). We derive an analogous result for the time-varying case and for the specific traffic game considered here. This is something that was not done in [7] for the single neighbourhood model.

1.2 The macroscopic fundamental diagram

To compute the travel time of a vehicle in a road traffic network, we shall see the vehicle as a customer in a queueing network, and neighbourhoods will be assigned the role of servers. The specificity of road traffic networks is that the arrival rates of vehicles are time-dependent and the departure rates of vehicles depend on the congestion (or the state) of the network. The higher the congestion, the lower is the rate at which vehicles exit the network. This relationship between the exit rate (or flow in vehicles per minute) and the vehicle density (vehicles per unit length) is captured by what is known as the macroscopic fundamental diagram (MFD) which was first introduced in [6] for single-lane traffic. Here macroscopic refers to the spatial scale of neighbourhoods. This is in contrast to the microscopic fundamental diagram which gives this relationship at the spatial scale of road segments [14, 11, 13]. These macroscopic diagrams have in fact been validated using measurements carried out, for example, in Yokohama [9] and Toulouse [3]. The rescaled and interpolated MFDs for these two cities are shown in Figure 1.

A typical MFD has an inverted cup shape (like a concave function) as shown in Figure 1. When the vehicular density is small (towards the left edge of the curve), the flow initially increases with the density because congestion is yet to form and vehicles are travelling at
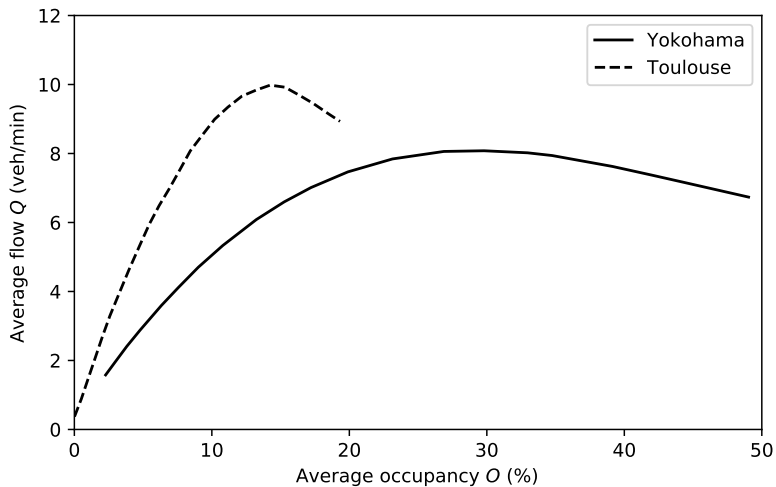
Figure 1: Macroscopic fundamental diagram

the maximum permissible speed. As the density increases further, congestion sets in, and the vehicular speed reduces. This leads to an inversion in the average flow which starts to decrease and becomes very small in case of traffic jams. This type of rate curves are not very typical in classic queueing networks.

For numerical examples, we shall use the MFD obtained from measurements in the city of Yokohama as the state-dependent service rate of servers.

1.3   Organisation

The rest of the paper is organised as follows. In Section 2, we introduce a Markovian model for vehicles in a road network. A set of partial differential equations for the distribution of the travel times are derived first when the number of vehicles in the network is small. We then obtain the corresponding equations in the fluid limit when the number of cars becomes large. This regime also leads to a simpler set of differential equations for the mean travel times. Section 3 presents the Wardrop equilbrium for the game with public transport as an alternative. The Wardrop equilibrium is then compared to the socially optimal travel choice and the PoA is calculated in the Section 4. Finally, we close with some conclusions in Section 5

2   Macrosocopic queueing model

We propose a Markovian queueing network with random routing for modelling congestion in a city at a macroscopic level. To be more precise, we model a city as a finite set of interconnected queues, where each queue represent the number of cars in a neighbourhood of the city. In contrast to classic Jackson networks, the service rate of each queue depends on the queue size, to reflect the relation between the speed and density of cars as described by the macroscopic fundamental diagram.

2.1 Markov model

Consider a city and a set of neighbourhoods $\mathcal{N} = \{1, 2, \ldots, N\}$, for which the arrival rate statistics and the macroscopic fundamental diagram are known. A neighbourhood could, in theory, be the entire city itself. We shall assume that new vehicles arrive at the nth neighbourhood in accordance to a Poisson process with time dependent rate $\lambda_n(t)$, $n \in \mathcal{N}$.

The duration of a journey depends on the length of the journey as well as on the speed during the journey. We assume that the lengths are independently and identically exponentially distributed with rate 1, taking some other value being equivalent to rescaling time. The speed of vehicles inside the different neighbourhoods at a given time instant will depend upon the average density within the neighbourhood in accordance with the macroscopic fundamental diagram (MFD). Moreover, the number of vehicles that either leave the neighbourhood or reach their destination will also depend on the density, or equivalently, on the number of vehicles in the area. Let α_{mn} denote the probability that a vehicle in neighbourhood m leaves for another neighbourhood n, and let α_{m0} denote the probability that a vehicle in neighbourhood m reaches its destination in neighbourhood m. To simplify notation, we will further assume that $a_{mm} = 0$ for all $m \in \mathcal{N}$.

Let $Q_n(t)$ denote the number of vehicles (moving around) in neighbourhood n at time t, let $\mathbf{Q}(t) = [Q_n(t)]_{n \in \mathcal{N}}$, and let $\pi(\mathbf{i}; t) = \Pr[Q_1(t) = i_1, \ldots, Q_N(t) = i_N] = \Pr[\mathbf{Q}(t) = \mathbf{i}]$ denote the probability that there are i_1 vehicles in neighbourhood 1 at time t, i_2 vehicles in neighbourhood 2 at time t, $\ldots$, and i_N vehicles in neighbourhood N at time t, with $\mathbf{i} = [i_1, \ldots, i_N] \in \mathbb{N}^N$. The density in a neighbourhood being proportional to the queue size in that neighbourhood, and the number of departures from the neighbourhood being proportional to the density, the departure rate in neighbourhood m equals,

$$\mu_m(Q_m(t)) = \theta_m F_m(\Theta_m Q_m(t)),$$

where θ_m and Θ_m are given constants and F_m denotes the macroscopic fundamental diagram of neighbourhood m.

In view of the assumptions above, we have the following transitions from state $\mathbf{i} \in \mathbb{N}^N$:

- There is a new arrival in neighbourhood $n \in \mathcal{N}$ with rate $\lambda_n(t)$. Such an arrival induces a transition to state $\mathbf{i} + \mathbf{e}_n$. Here, , $\mathbf{e}_m = [\mathbb{1}_{\{m=n\}}]_{n \in \mathcal{N}}$ is a vector of zeroes, apart from its mth element which is one.

- A vehicle moves from neighbourhood m to neighbourhood n with rate $\mu_m(i_m)\alpha_{mn}$, the new state being $\mathbf{i} + \mathbf{e}_n - \mathbf{e}_m$.

- A vehicle reaches its destination in neighbourhood m with rate $\mu_m(i_m)\alpha_{m0}$, the new state being $\mathbf{i} - \mathbf{e}_m$.

The description of the state transitions then immediately yields the following set of forward Chapmann-Kolmogorov equations,

$$\frac{\partial}{\partial t}\pi(\mathbf{i}; t) = \sum_{m \in \mathcal{N}} \sum_{n \in \mathcal{N}^*} \pi(\mathbf{i} - \mathbf{e}_n + \mathbf{e}_m; t)\mu_m(i_m + 1)\alpha_{mn}$$

$$+ \sum_{m \in \mathcal{N}} \pi(\mathbf{i} - \mathbf{e}_m; t)\lambda_m(t)$$

$$- \sum_{m \in \mathcal{N}} \pi(\mathbf{i}; t)\mu_m(i_m) - \sum_{m \in N} \lambda_m(t)\pi(\mathbf{i}; t),$$

with $\mathcal{N}^* = \mathcal{N} \cup \{0\}$, and with $\pi(\mathbf{i};t) = 0$ for $\mathbf{i} \notin \mathbb{N}^N$. Note that the notation implies that $\mathbf{e}_0 = \mathbf{0}$.

Let $W_n(\tau)$ denote the remaining travel time of a vehicle in neighbourhood n at time τ, and let $\nu_n(\tau, t | \mathbf{i}) = \Pr[W_n(\tau) > t | \mathbf{Q}(\tau) = \mathbf{i}]$ denote the complementary distribution function of this vehicle, conditioned on having $\mathbf{Q}(\tau) = \mathbf{i}$ vehicles in the different neighbourhoods at time τ. Conditioning on the possible events in the interval $(\tau, \tau + h]$, we can express $\nu_n(\tau, t + h | \mathbf{i})$ as follows,

$$
\begin{aligned}
\nu_n(\tau, t + h | \mathbf{i}) = &\sum_{m \in \mathcal{N}} \nu_n(\tau + h, t | \mathbf{i} + \mathbf{e}_m) \lambda_m(t) h \\
&+ \sum_{m \in \mathcal{N} \setminus \{n\}} \sum_{k \in \mathcal{N}^*} \nu_n(\tau + h, t | \mathbf{i} - \mathbf{e}_m + \mathbf{e}_k) \mu_m(i_m) \alpha_{mk} h \\
&+ \sum_{m \in \mathcal{N}^*} \nu_n(\tau + h, t | \mathbf{i} - \mathbf{e}_n + \mathbf{e}_m) \mu_n(i_n) \alpha_{nm} \left(1 - \frac{1}{i_n}\right) h \\
&+ \sum_{m \in \mathcal{N}} \nu_m(\tau + h, t | \mathbf{i} - \mathbf{e}_n + \mathbf{e}_m) \mu_n(i_n) \alpha_{nm} \frac{1}{i_n} h \\
&+ \nu_n(\tau + h, t | \mathbf{i}) \left(1 - \sum_{m \in \mathcal{N}} \lambda_m(t) h - \sum_{m \in \mathcal{N}} \mu_m(i_m) h\right) + o(h),
\end{aligned}
$$

Here we assume that each vehicle in the neighbourhood is equally likely to depart from the neighbourhood (either by arriving at the destination or by leaving for another neighbourhood). Hence, in queueing terms, we assume that the "server" uses a processor sharing or random order of service discipline. Sending $h \to 0$, we further have,

$$
\begin{aligned}
\frac{\partial}{\partial t} \nu_n(\tau, t | \mathbf{i}) - \frac{\partial}{\partial \tau} \nu_n(\tau, t | \mathbf{i}) = &\sum_{m \in \mathcal{N}} \nu_n(\tau, t | \mathbf{i} + \mathbf{e}_m) \lambda_m(t) \\
&+ \sum_{m \in \mathcal{N} \setminus \{n\}} \sum_{k \in \mathcal{N}^*} \nu_n(\tau, t | \mathbf{i} - \mathbf{e}_m + \mathbf{e}_k) \mu_m(i_m) \alpha_{mk} \\
&+ \sum_{m \in \mathcal{N}^*} \nu_n(\tau, t | \mathbf{i} - \mathbf{e}_n + \mathbf{e}_m) \mu_n(i_n) \alpha_{nm} \left(1 - \frac{1}{i_n}\right) \\
&+ \sum_{m \in \mathcal{N}} \nu_m(\tau, t | \mathbf{i} - \mathbf{e}_n + \mathbf{e}_m) \mu_n(i_n) \alpha_{nm} \frac{1}{i_n} \\
&- \nu_n(\tau, t | \mathbf{i}) \left(\sum_{m \in \mathcal{N}} \lambda_m(t) + \mu_m(i_m)\right).
\end{aligned}
$$

Finally, the expected travel time of a vehicle that arrives at time τ in neighbourhood n is,

$$
\bar{W}_n(\tau) = \sum_{\mathbf{i} \in \mathbb{N}^N} \pi(\mathbf{i};t) \int_0^\infty \nu_n(\tau, t | \mathbf{i} + \mathbf{e}_n) dt.
$$

Here we used the observation that the mean travel time of a vehicle arriving at time τ in neighbourhood n while there are $\mathbf{i}$ vehicles in the different neighbourhoods equals the mean remaining travel time of a vehicle in nieghbourhood n when there are $\mathbf{i} + \mathbf{e}_n$ vehicles in the different neighbourhoods.

2.2 Fluid limit

As the fundamental diagram translates to state-dependent service rates, there is no simple analytic solution for the infinite systems of differential equations for $\pi(\mathbf{i};t)$ and $w(\tau,t|\mathbf{i})$. Moreover, numerically solving these equations is computationally demanding. We however can study the fluid limit of the system, that is, we consider a sequence of models $Q^K(t)$, the Kth model having arrival rate $\lambda_n(t)K$ in the nth neighbourhood and study $q_n(t) = \lim_{K\to\infty} Q_n^K(t)K^{-1}$. The parameter K is the scale at which we are making measurements. As the scale becomes large, the number of vehicles also grows proportionally, and hence we have to divide Q_n by K to obtain a quantity which goes to a finite limit. This quantity can be thought of as the density of the vehicles. By the functional law of large numbers, the randomness of the movement of individual vehicles is washed out by the scaling. This leads to a considerable reduction of the complexity of the analysis since we are no longer dealing with individual cars but instead deal with a large number of cars which behave in a similar way.

Let $\bar{\mu}_n(x)$ denote the departure rate from the nth neighbourhood in the fluid regime, that is,

$$\bar{\mu}_n(x) = \lim_{K\to\infty} \frac{\mu_n(\lceil Kx \rceil)}{K}.$$

Then, by writing the evolution of $Q_n(t)$ as the difference of randomly time-changed Poisson processes and applying the functional strong law of large numbers for Poisson processes, we find that the limiting process $q_n(t)$ adheres the following ordinary differential equation,

$$\dot{q}_n(t) = \lambda_n(t) - \bar{\mu}_n(q_n(t)) + \sum_{m\in\mathcal{N}} \bar{\mu}_m(q_m(t))\alpha_{mn}. \tag{1}$$

For the travel time calculations, we scale the arrival process and service process as before. However, we retain the randomness of the random order of service (or processor sharing) discipline. Hence, in the limit, the travel times are still random, let $W_n^K(\tau)$ denote the travel time from neighbourhood n at time τ for the Kth model, and $w_n(\tau,t)$ be the complimentary waiting time distribution in the fluid limit,

$$w_n(\tau,t) = \lim_{K\to\infty} \Pr[W_n^K(\tau) > t].$$

We obtain the following set of differential equations for these complimentary travel time distributions,

$$\frac{\partial}{\partial t}w_n(\tau,t) - \frac{\partial}{\partial \tau}w_n(\tau,t) =$$

$$-w_n(\tau,t)\frac{\bar{\mu}_n(q_n(\tau))}{q_n(\tau)} + \sum_{m\in\mathcal{N}} w_m(\tau,t)\alpha_{nm}\frac{\bar{\mu}_n(q_n(\tau))}{q_n(\tau)}.$$

Finally, the mean travel time of a vehicle arriving at time τ in neighbourhood n can then be computed by integrating $w_n(\tau,t)$ over t, that is,

$$\bar{w}_n(\tau) = \int_0^\infty w_n(\tau,t)dt.$$

Combining the former expressions also yields the following differential equation for the mean waiting times,

$$\frac{d}{d\tau}\bar{w}_n(\tau) = \bar{w}_n(\tau)\frac{\bar{\mu}_n(q_n(\tau))}{q_n(\tau)} - 1 - \sum_{m\in\mathcal{N}} \bar{w}_m(\tau)\alpha_{nm}\frac{\bar{\mu}_n(q_n(\tau))}{q_n(\tau)}. \tag{2}$$

In contrast to the system of differential equations for the traffic densities (1), the differential equation above cannot be used to numerically calculate the time-dependent travel times. Indeed, in order to numerically solve the system of ordinary differential equations, the initial expected travel times $w_n(0)$ are needed. Note that these do depend on the future evolution of the traffic densities. To overcome this difficulty, we will reverse time. Assuming that λ_n is constant for t sufficiently large, the traffic density in the different neighbourhoods (in the fluid regime) and the corresponding travel times will become approximately constant as well. We therefore choose a sufficiently large T and approximate the mean travel times at time T by the asymptotic mean travel times $w_n(T) \approx \lim_{\tau \to \infty} w_n(\tau)$. For ease of notation, introduce $\bar{\omega}_n(\tau) = \bar{w}_n(T - \tau)$. We then obtain the following system of ordinary differential equations for $\bar{\omega}_n$,

$$\frac{d}{d\tau}\bar{\omega}_n(\tau) = 1 - \bar{\omega}_n(\tau)\frac{\bar{\mu}_n(q_n(T-\tau))}{q_n(T-\tau)}$$
$$+ \sum_{m \in \mathcal{N}} \bar{\omega}_m(\tau)\alpha_{nm}\frac{\bar{\mu}_n(q_n(T-\tau))}{q_n(T-\tau)} \ .$$

In contrast to the system of equations (2), the initial value of the ode can now be calculated. Indeed, by assuming that $\lambda_n(t)$ is constant and by sending $t \to \infty$, we find that the initial value of the ode solves the linear system of equations,

$$\bar{\omega}_n(0)\frac{\bar{\mu}_n(q_n(T))}{q_n(T)} = 1 + \sum_{m \in \mathcal{N}} \bar{\omega}_m(0)\alpha_{nm}\frac{\bar{\mu}_n(q_n(T))}{q_n(T)} \ . \tag{3}$$

Remark 1 *The assumptions on the functions λ_n do not constrain the problems that can be solved by the method at hand. Indeed, if λ_n is not constant for large time values in the time interval under study, it is straightforward to extend the interval such that this is the case. This should not affect the results in the original time interval of interest. The time interval needs to be large enough such that any future traffic patterns do not affect the travel times in the region of interest.*

2.3 Numerical example

We now illustrate our approach by a numerical example. We consider a city divided into 4 neighbourhoods as depicted in Figure 2. Traffic only arrives in the outer areas 2 to 4. Let $\lambda_i(t)$ denote the arrival rate (AR) in area $i \in \{2, 3, 4\}$. Departures in each area are governed by the MFD of Yokahama, which is properly rescaled to reflect realistic travel times (around 15 minutes in total when there is no congestion). Vehicles leave for the centre neighbourhood 1 with probability ψ, to any of its neighbours with probability ϕ or park in the neighbourhood with probability $1 - \psi - 2\phi$.

Figure 3 shows the evolution of the traffic density and the expected waiting times in the fluid limit regime for $\phi = 0.1$ and $\psi = 0.4$ and for $\lambda_3(t) = \lambda_2(t)$ and $\lambda_4(t) = \lambda_2(t)$, with $\lambda_2(t)$ as depicted. Notice that the arrival rates in the fluid regime are very small. This is not unexpected: the fluid limit is a deterministic approximation for the city with arrival rates $K\lambda_i$, where K is sufficiently large.

The arrival rate curve reaches its maximum value at $t = 60$, while the peaks of the traffic density and expected travel times are somewhat later. This is most pronounced for the traffic density in the centre neighbourhood 1. This is not unexpected as all traffic towards
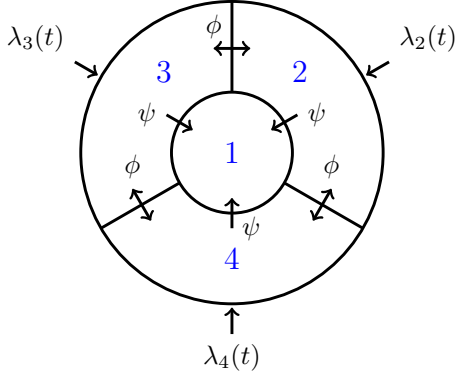
Figure 2: City with four neighbourhoods. There are arrivals in the outer areas, and only transitions between neighbouring areas, and towards the centre of the city.

the centre needs to cross the outside neighbourhoods, such that peak traffic is delayed. This is also reflected in the travel times in the centre area. Congestion in the centre starts when the congestion in the outside neighbourhoods is already resolving. Finally, it is readily seen that the maximum traffic density exceeds 30% in some cases. At such traffic densities, the average flow decreases for increasing densities, see Figure 1.

To assess the accuracy of the fluid approximation, Figure 4 again shows the time-dependent waiting times, and compares these curves with the expected waiting times, conditional on the system content. That is, for these simulation results, we draw a single sample path of the traffic densities over time, and then use Monte Carlo simulation to estimate the mean waiting time at various points in time for this sample path. For each time instant, we sampled the waiting time 5000 times (recall that the waiting time for a given evolution of the traffic densities is a random variable as the vehicle that leaves at a certain instant is randomly selected). We also calculated the 99% confidence intervals for the estimators of the conditional mean waiting times, but omitted these from the figure as their sizes are to small to be properly depicted. We considered 3 distinct scale values K. Practically, the values $K = 500$, $K = 1000$ and $K = 5000$ correspond to having a peak arrival intensity at the city of 35, 70 and 350 vehicles per minute, respectively.

For $K = 500$, the simulation results considerably deviate from the fluid limit. This reflects the remaining randomness in the traffic density process for smaller K. For $K = 1000$, the correspondence between simulation results and the fluid limit are already good, while the correspondence is very good for $K = 5000$. Even for $K = 5000$, the traffic intensity is not that high (at the scale of a city). Hence we can conclude that the fluid approximation works in parameter settings of interest.

3   Public transport game

The derivations above not only allow for estimating the travel times of the vehicles in the city. Our results can also be used for dimensioning the capacity of public transport systems, and for assessing the impact of the public transport cost on congestion. In this section, we consider a scenario in which users have access to travel time estimates both for commuting by personal vehicles as well as by public transport. If the public transport system (PTS)
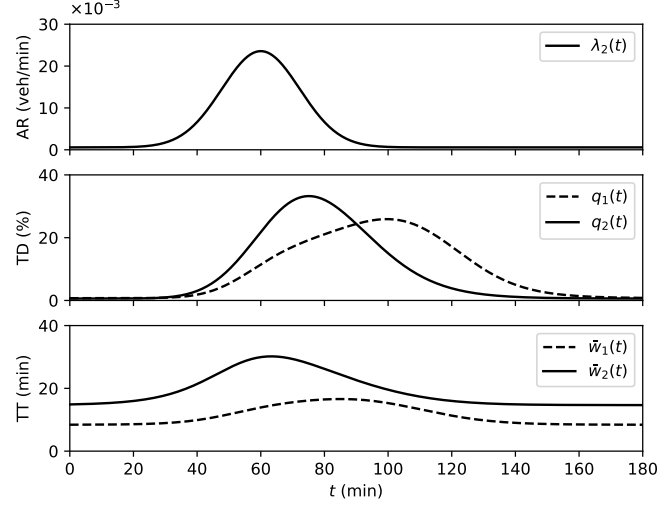
Figure 3: Evolution of the arrival rate (AR), traffic density (TD) and travel times (TT) during rush hour.
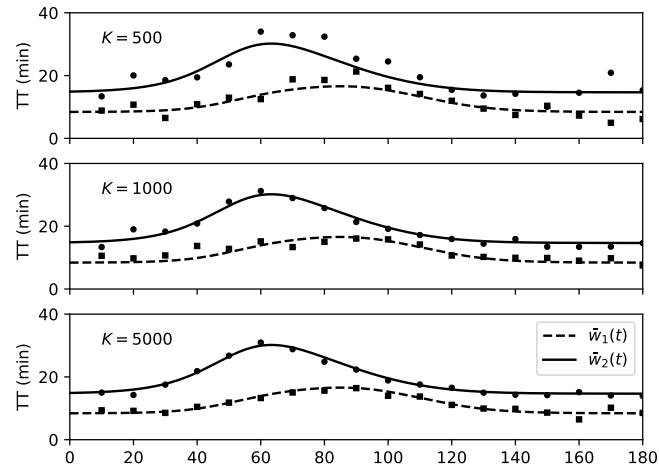

Figure 4: Accuracy of the mean travel times in the fluid approximation.

offers a faster ride, a certain number of users may switch from personal vehicles to the PTS. Assuming users are rational, that is, they choose the mode of transport that minimises their travel times, we use concepts from game theory to compute the number (or the fraction) of users that will switch to the PTS. This can be used to estimate the load on the PTS and dimension its capacity accordingly.

For simplicity, assume that taking the PTS incurs a cost, which may depend upon the neighbourhood and the time of departure, but not on the congestion in the road network. This cost includes the travel time in the public transport system, the travel time to the access points of the public transport system, the inconvenience of using public transport, costs related to possible disruptions of the service, etc. We assume that this cost can be expressed in terms of travel times: the cost of the public transportation system is the maximal travel time one is willing to undergo by private transport. For $n \in \mathcal{N}$, let $C_n(t)$ denote the cost of the PTS at time t for a vehicle starting in neighbourhood n.

Remark 2 *Note that the assumption that the travel cost does not depend on congestion is natural in case the public transportation system has dedicated lines. This can be either completely separated from the road network like a metro network, or consist of separate lanes embedded in the road network.*

3.1 Dynamic Wardrop equilibrium

We investigate the strategies of rational users when the choice of each user has a negligible impact on the travel times of the population, a solution concept introduced in game theory by Wardrop [18] in the context of choosing routes in road traffic networks. Known as the Wardrop equilibrium, it says that at the equilibrium choice of routes is such that the cost on each used route is the same and the cost on the unused ones is larger than that of the used ones. In contrast to the classic Wardrop equilibrium, our equilibrium explicitly depends on time. For this *dynamic Wardrop equilibrium*, the choice at time t not only depends on the choices of the other users at time t, but also on the choices of other users (both prior and later than time t).

Let $p_n^\circ(t)$ denote the fraction of users that opt for private transport at time t. The function p_n° is the time-dependent routing strategy for arrivals in neighbourhood n. The set of routing strategies constitutes a Wardrop equilibrium if, for each $n \in \mathcal{N}$ such that $\lambda_n(t) > 0$, we have the following constraints on the travel times:

$$\begin{cases} \bar{w}_n(t; p_n^\circ) < C_n(t) & \text{for } p_n^\circ(t) = 1 \,, \\ \bar{w}_n(t; p_n^\circ) = C_n(t) & \text{for } 0 < p_n^\circ(t) < 1 \,, \\ \bar{w}_n(t; p_n^\circ) > C_n(t) & \text{for } p_n^\circ(t) = 0 \,. \end{cases} \tag{4}$$

The first equation states that all users in neighbourhood n prefer private transport if the cost of the PTS exceeds the expected travel time. The second equation states that if a non-zero fraction of users opt for PTS and a non-zero fraction of users opt for private transport, the cost of both should be equal. Finally, the last equality states that everyone uses the PTS if the cost of private transport exceeds that of the PTS.

To the best of our knowledge, the functions p_n° ($n \in \mathcal{N}$) have no nice and easy analytical solution. Therefore, we resort to the following iterative numerical algorithm for their computation. Given the arrival curves $\lambda_n(t)$, the fundamental diagrams $\bar{\mu}_n(x)$ and the routing

probabilities α_{mn}, we set $p_n(t) = 1$, and then update it according to

$$p_n(t) \leftarrow \min \left(1, p_n(t) e^{-\beta_n (\omega(t;p) - C_n(t))}\right) ,$$

for $\lambda_n(t) \neq 0$, such that $p_n \to p_n^\circ$. Here β_n is a parameter which determines the speed of convergence of the recursion. Small β_n corresponds to conservative updates, such that convergence is slow. In contrast, for large β_n, the update rule may never converge. Finally, note that for $\lambda_n(t) = 0$, the value $p_n^\circ(t)$ has no meaning as there is no choice to be made in the absence of traffic.

3.2   Numerical example

We now calculate the Wardop equilibrium for some traffic scenarios. We retain the assumptions on the arrival traffic and the random routing within the city of section 2.3 and consider two different public transport offerings. In the first scenario, there is a PTS offering in all neighbourhoods with a fixed cost of 25 minutes. The traffic densities and travel times for this symmetric scenario are depicted in Figure 5. Note that we only depict results for neighbourhoods 1 and 2, as the results for neighbourhoods 2, 3 and 4 are identical by symmetry. In the second scenario, there is a PTS offering in neighbourhoods 2 and 3, but none in neighbourhood 4. Moreover, the PTS cost differs: $C_2(t) = 20$ minutes and $C_3(t) = 25$ minutes. The results for this asymmetric scenario are depicted in Figure 6.

In both scenario's we can easily verify that the travel times never exceed the PTS cost. Moreover, once the travel times hit the PTS cost, a considerable fraction of users opt for the PTS. This reduction of traffic is not only beneficial for the travel times from the outer neighbourhoods. Also the travel time within the centre decreases as there is also a reduction of traffic there. To facilitate comparisons, the traffic densities and travel times without the availability of the PTS are also depicted in Figure 5 (by dotted curves). Finally, in Figure 6, we also observe a reduction in travel times for neighbourhood 4 where there is no PTS offering. This is not unexpected, as users from other neighbourhoods that opt for the PTS reduce the overall traffic in neighbourhood 4 as well.

4   Social optimum

When users choose selfishly between two options, it often leads to inefficiencies in terms of the overall cost function. In this section, we quantify these inefficiencies using the concept of Price of Anarchy (PoA) [12]. The PoA is defined as the ratio of the cost at the Wardrop equilibrium and the cost at the social optimum. A higher value of PoA implies that the cost of allowing for individual decision making is higher, which can be interpreted as the price that is paid for decentralised decision making.

4.1   Optimal control problem

Let $p_n(t)$ be the probability with which a user entering in neighbourhood n uses her vehicle. Then, the mean transportation cost $K_n(t)$ at time t of a user entering neighbourhood n assuming policy $\mathbf{p}$ equals

$$K_n(t; \mathbf{p}) := \sum_i \bar{w}_n(t; \mathbf{p}) p_n(t) + (1 - p_n(t)) C(t) . \tag{5}$$
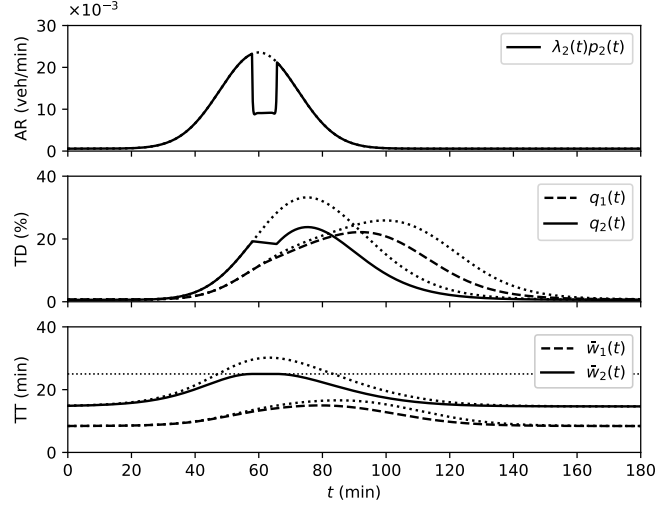
11

Figure 5: Arrival rate (AR), traffic density (TD) and travel times (TT) during rush hour in Wardrop equilibrium (symmetric public transport costs).
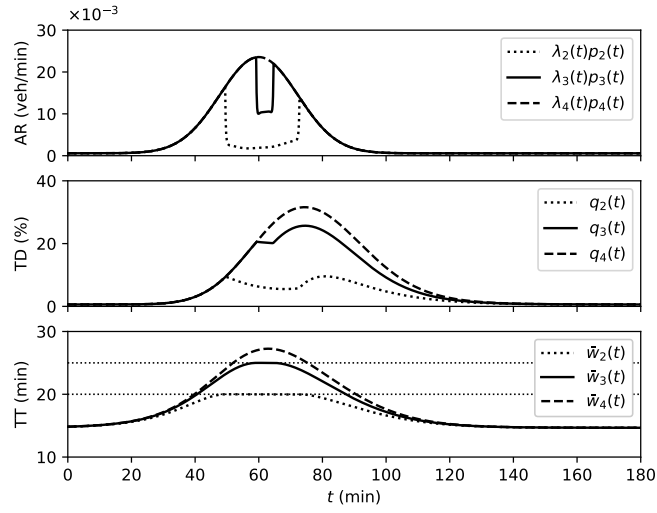

Figure 6: Arrival rate (AR), traffic density (TD) and travel times (TT) during rush hour in Wardrop equilibrium (symmetric public transport costs).

Here, we have assumed that the cost of public transport does not depend upon the point of entry and exit of the user.

The socially optimal policy is the one that minimises the total weighted mean cost given by

$$\mathbf{p}^* = \arg\inf_{\mathbf{p}} \int_{T_1}^{T_2} \sum_n \lambda_n(t) K_n(t; \mathbf{p}) dt \tag{6}$$

where $[T_1, T_2]$ is the rush hour interval. Note that the expression above includes the arrival rates $\lambda_n(t)$ as the travel cost impacts more users if there are many arrivals.

With $\mathbf{p}^\circ(t)$ as the solution of the set of equations (4). The PoA is then defined as:

$$\text{PoA} = \frac{\int_{T_1}^{T_2} \sum_n \lambda(t)_i K_n(t; \mathbf{p}^\circ) dt}{\int_{T_1}^{T_2} \sum_n \lambda_n(t) K_n(t; \mathbf{p}^*) dt}.$$

As was the case in [7], (6) is not an optimal control problem for a general $\lambda_n(t)$ general because $\bar{\mathbf{w}}(0)$ cannot be determined independently of $\mathbf{q}$ (see the discussion below (2)). In order to solve this problem, we shall make the same assumption as was made for obtaining (3). That is, we shall assume that $\lambda_n(t)$ goes to a constant as $t \to \infty$. Taking the T_2 to be sufficiently large so that the system is in stationarity, we can then approximate (6) by the solution of following optimal control problem:

$$\min_{\mathbf{p}(t) \in [0,1]} \int_{T_1}^{T_2} \sum_n \lambda_n(t) K_n(t; \mathbf{p}) dt, \tag{7}$$

subject to

$$\frac{d}{dt} q_n(t) = \lambda_n(t) - \bar{\mu}_n(q_n(t)) + \sum_{m \in \mathcal{N}} \bar{\mu}_m(q_m(t)) \alpha_{mn}, \tag{8}$$

$$\frac{d}{dt} \bar{w}_n(t) = \bar{w}_n(t) \frac{\bar{\mu}_n(q_n(t))}{q_n(t)} - 1 - \sum_{m \in \mathcal{N}} \bar{w}_m(t) \alpha_{nm} \frac{\bar{\mu}_n(q_n(t))}{q_n(t)}, \tag{9}$$

with the initial values of $q_n(0)$ computed using the steady state solution of (1) and $\bar{w}(T)$ obtained as the solution of (3).

4.2   Congestion-based pricing

In the previous subsection, the price (or cost) of public transport was taken to be the same as the one in the game of Section 3. The numerical example in Section 4.3 illustrates that the PoA in this setting is larger than 1. That is, selfish decision making induces inefficiencies from the point of view of the social cost.

One way to remove these inefficiencies is to impose these decisions on individuals. However, this forced choice can be seen as too authoritarian. Another method is impose tolls or taxes on certain routes or resources. Here, the individuals still make selfish decisions, but in a game with different costs. In static congestion games, that is games in which arrival rates are constant, it is known that by introducing appropriate tolls, one can obtain the social optimum in a game setting [5]. In this subsection, we follow this method and compute a congestion-based pricing of public transport such that the equilibrium of the dynamic

congestion game with modified costs will be same as the social optimum for the fixed cost setting that was computed in the subsection above.

Let $\hat{C}(t)$ be the congestion-based cost of the public transport network, and let $\hat{p}_n^\circ(t)$ be the equilibrium probability in neighbourhood n of the game (4) with the cost replaced by $\hat{C}$. For $\hat{p}_n^\circ$ to be equal to p_n^* (which is the solution of (6)), it needs to satisfy certain conditions that are determined by the Pontryagin's maximum principle [8].

When the maximum principle is applied to (6), it follows that (we skip the details)

$$\begin{cases} \bar{w}_n(t; p_n^*) < C_n - \gamma_n(t) & \text{for } p_n^*(t) = 1 \,, \\ \bar{w}_n(t; p_n^*) = C_n - \gamma_n(t) & \text{for } 0 < p_n^*(t) < 1 \,, \\ \bar{w}_n(t; p_n^*) > C_n - \gamma_n(t) & \text{for } p_n^*(t) = 0 \,, \end{cases} \qquad (10)$$

where $\gamma_n(t)$ is the adjoint (or the co-state) variable for $q_n(t)$ and is the multiplier of (8).

Comparing (10) and (4) in which $C_n(t) = \hat{C}_n(t)$, we obtain

$$\hat{C}_n(t) = C_n - \gamma_n(t) \,, \quad \forall n. \qquad (11)$$

That is, congestion-based pricing for the game to have a PoA of 1 is the sum of the fixed original cost C_n and a congestion-based cost $\gamma_n(t)$ in neighbourhood n. We can interpret $\gamma_n(t)$ as the toll (or subsidy if is negative) that the transport authority should impose to induce a PoA of 1.

4.3 Numerical example

We now calculate the social optimum policy, PoA, and the congestion-based pricing for the symmetric traffic scenario of Section 3.2. The assumptions on the arrival traffic and the routing within the city are the ones described in Section 2.3 and there is a PTS offering in all neighbourhoods with a fixed cost of 25 minutes. The optimal control problem (7) was solved using the software BOCOP which is a specialised numerical solver for such problems [16].

In Figure 7, we depict the optimal arrival rate in neighbourhood 2 (top figure), the traffic density in neighbourhoods 1 and 2 (middle figure), and the mean travel times in neighbourhoods 1 and 2 (bottom figure). Also, due to symmetry of the arrival rates, traffic densities and travel times in neighbourhoods 3 and 4 are identical to those in neighbourhood 2, and are therefore omitted.

The PoA for this particular example was found to be 1.081, that is the average social cost at the Wardrop equilibrium was 8% higher than at the social optimum.

Finally in Figure 8 we show the congestion-based pricing of (11) for this example that leads to the social optimum. The adjoint function $\gamma_n(t)$ was computed by BOCOP. Observe that, for this example, the pricing is in fact a subsidy since it involves reducing the cost of public transport at peak hours. Further, the subsidy is higher during periods of higher travel times (or higher congestion).

5 Conclusions

Drawing on queueing network theory and accounting for the macroscopic fundamental diagram of traffic flow, we have proposed a Markovian queueing network model for computing the time-dependent travel times during rush hour in a congested area. The model at hand
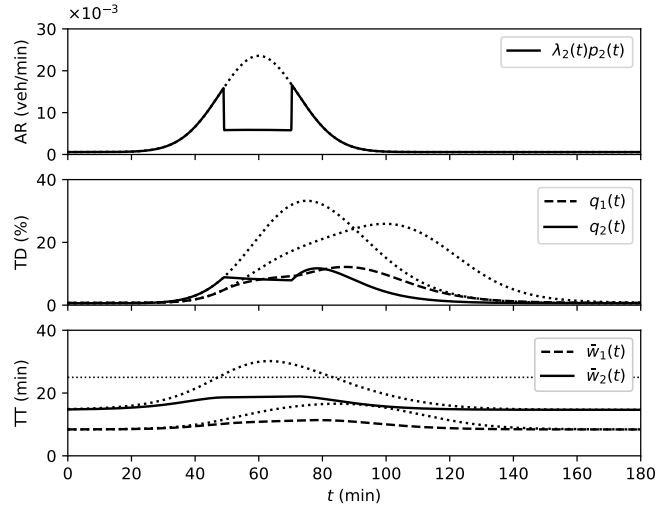
Figure 7: Socially optimal arrival rate (AR), traffic density (TD) and travel times (TT) during rush hour (symmetric public transport costs).
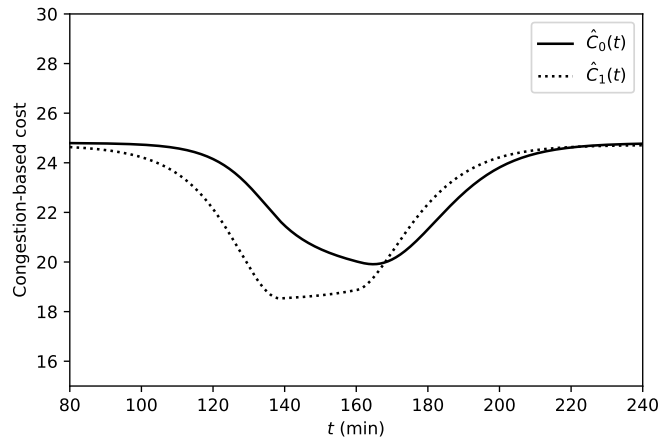

Figure 8: Congestion-based cost function that acheives a PoA of 1 (symmetric public transport costs).

was characterised by two non-typical properties: queue-dependent departure rates to reflect the macroscopic fundamental diagram and time-dependent arrival intensities which are typical in rush-hour scenarios. As the traffic intensity is high during rush hour, we study the fluid limit of the queueing network model at hand. Numerical experiments show that the travel time can be reduced significantly if a reasonable fraction of users switches to public transport. Relying on game theory, we show how our model can be used to assess the impact of public transport systems on congestion and compare the game-theoretical and socially optimal traffic mixes.

References

[1] Forest Baskett, K. Mani Chandy, Richard R. Muntz, and Fernando G. Palacios. Open, closed, and mixed networks of queues with different classes of customers. *J. ACM*, 22(2):248–260, April 1975.

[2] M. A. A. Boon, R. D. van der Mei, and E. M. M. Winands. Applications of polling systems. *Surveys in Operations Research and Management Science*, 16(2):67–82, 2011.

[3] C. Buisson and C. Ladier. Exploring the impact of homogeneity of traffic measurements on the existence of macroscopic fundamental diagrams. *Transportation Research Record*, 2124:127–136, 2009.

[4] X. Chen, Z. Li, L. Li, and Q. Shi. A traffic breakdown model based on queueing theory. *Networks and Spatial Economics*, 14(3):485–504, Dec 2014.

[5] Richard Cole, Yevgeniy Dodis, and Tim Roughgarden. Pricing network edges for heterogeneous selfish users. In *Proceedings of the Thirty-fifth Annual ACM Symposium on Theory of Computing*, STOC '03, pages 521–530, New York, NY, USA, 2003. ACM.

[6] C. F. Daganzo. Urban gridlock: Macroscopic modeling and mitigation approaches. *Transportation Research Part B*, 41:49–62, 2007.

[7] Dieter Fiems, Balakrishna Prabhu, and Koen De Turck. Travel times, rational queueing and the macroscopic fundamental diagram of traffic flow. *Physica A: Statistical Mechanics and its Applications*, 524:412 – 421, 2019.

[8] H. P. Geering. *Optimal Control with Engineering Applications.* Springer, 2007.

[9] N. Geroliminis and C. F. Daganzo. Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B*, 42:759–770, 2008.

[10] James R. Jackson. Jobshop-like queueing systems. *Management Science*, 10(1):131–142, 1963.

[11] B. Kerner. *The Physics of Traffic - Empirical Freeway Pattern Features, Engineering Applications, and Theory.* Springer, Berlin, 2004.

[12] E. Koutsoupias and C. H. Papadimitriou. Worst-case equilibria. In *Proceeding of STACS*, 1999.

[13] R. Kühne and N. H. Gartner, editors. *Transportation Research Circular E-C149 (75 Years of the Fundamental Diagram for Traffic Flow Theory: Greenshields Symposium)*. Transportation Research Board, 2011.

[14] H. Lehmann. Distribution function properties and the fundamental diagram in kinetic traffic flow theory. *Phys. Rev. E*, 54:6058–6064, Dec 1996.

[15] J.I. Levy, J.J. Buonocore, and K. von Stackelberg. Evaluation of the public health impacts of traffic congestion: a health risk assessment. *Environmental Health*, 9(Article 65), 2010.

[16] Team Commands, Inria Saclay. BOCOP: an open source toolbox for optimal control. `http://bocop.org`, 2017.

[17] N. Vandaele, T. Van Woensel, and A. Verbruggen. A queueing based traffic flow model. *Transportation Research Part D: Transport and Environment*, 5(2):121–135, 2000.

[18] J. G. Wardrop. Some theoretical aspects of road traffic research. *ICE Proceedings: Engineering Divisions*, 1(3):325–362, 1952.