This item was submitted to [Loughborough's Research Repository](#) by the author.
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

Analytical evaluation of flexible-sharing strategies on multimodal arterials

PLEASE CITE THE PUBLISHED VERSION

<https://doi.org/10.1016/j.tra.2018.01.038>

PUBLISHER

Elsevier

VERSION

AM (Accepted Manuscript)

PUBLISHER STATEMENT

This paper was accepted for publication in the journal Transportation Research Part A: Policy and Practice and the definitive published version is available at <https://doi.org/10.1016/j.tra.2018.01.038>.

LICENCE

CC BY-NC-ND 4.0

REPOSITORY RECORD

He, Haitao, Monica Menendez, and S Ilgin Guler. 2019. "Analytical Evaluation of Flexible-sharing Strategies on Multimodal Arterials". figshare. <https://hdl.handle.net/2134/11279792.v1>.

Analytical evaluation of flexible sharing strategies on multi-modal arterials

Haitao He^{a,*}, Monica Menendez^a, S. Ilgin Guler^b

^a*Traffic Engineering Group, Institute for Transport Planning and Systems, ETH Zurich, Switzerland*

^b*Department of Civil and Environmental Engineering, The Pennsylvania State University, USA*

Abstract

This paper examines strategies that allocate road space capacity in a flexible manner at a link or a node of an arterial to move buses in front of the car queues without continuously banning cars from using one full lane. These strategies are collectively referred to as flexible sharing strategies, and are modelled parsimoniously using a single parameter to describe their effectiveness. This parameter can be estimated analytically, empirically, or with simulations. We then build an analytical framework with input-output systems and fixed set theory to evaluate bi-modal arterials with any kind of flexible sharing strategies implemented at multiple locations. Using this framework, the performance of the bi-modal arterial is quantified by a set of vehicle throughput. This set is convex with linear constraints. Moreover, the requirements to efficiently install flexible sharing strategies along the arterial are mathematically established. Lastly, one particular flexible sharing strategy, the pre-signal, is illustrated as an example both mathematically and with a calibrated simulation.

Keywords: Multi-modal; Capacity; Bus priority; Arterial; Traffic flow; Pre-signal

1. Introduction and background

Improving public transportation operations is important because more people can be moved within limited road space. Buses are promoted with various bus priority strategies in many cities around the world to encourage people to switch from cars. This could lead to a more sustainable transportation system in the long run. One common bus priority strategy is to implement dedicated bus lanes. However, the discrete nature of traditional dedicated bus lanes (i.e. either one full lane is reserved for buses or none) is problematic because such allocation of road space capacity between buses and cars is likely sub-optimal. Bus flows in most urban arterials are much lower than the capacity of the dedicated lane, while car flows could often be higher than the capacity of the remaining lanes. This is especially detrimental at bottleneck locations, where growing car queues during peak hours could eventually spillback to upstream intersections (Daganzo (2007); Liu and Chang (2011)). This can impact the traffic performance of the whole network. Moreover, allocating capacity between modes in such a fixed manner cannot respond to the variation in traffic demand throughout the day. Therefore, a continuously dedicated bus lane along an arterial is often undesirable, especially when the bus flow is low or the car demand is high.

This paper examines strategies that allocate road space capacity in a flexible manner at a given element (i.e. a link or a node) of an arterial to move buses in front of the car queues without continuously banning

*Corresponding author
Phone: +41 /(0)44-633-3195
Email address: haitao.he@ivt.baug.ethz.ch (Haitao He)

cars from using one full lane. We refer to these strategies as flexible sharing strategies, because the capacity allocation is not fixed. These strategies use flexible space allocation to combine the advantages of mixed lanes (the elements' capacity can be fully utilized if buses are not present) and dedicated bus lanes (the buses receive priority). There are different kinds of flexible sharing strategies. On links, this includes any form of shared-use bus lane, e.g. sharing with taxis or high-occupancy cars (Agrawal et al. (2013); Guler and Cassidy (2012)). On nodes, this includes innovative intersection designs such as queue jumper lanes and pre-signals (Nowlin and Fitzpatrick (1997); Wu and Hounsell (1998); Guler and Menendez (2014a,b); He et al. (2016)). Excluded from the list are pre-signals which use the opposite direction lane (Guler et al. (2016)), dynamic bus lanes which must be implemented on multiple elements along the arterial (Eichler and Daganzo (2006); Chiabaut et al. (2013)), and transit signal priority which changes how capacity is allocated at the intersection between different approaches (Baker et al. (2002); Smith et al. (2005)).

There are various measures to assess the operation of a transportation system, for example capacity, accessibility, equality, reliability, economical costs, and environmental impacts. Among them, capacity is one of the most important parameters which describes the maximum sustainable quantity of vehicles that can pass through the system over a period of time. Systems with lower capacities are more prone to congestion, which results in decrease of labor productivity, reduction of system reliability, increase in transportation cost of goods, and higher environmental impacts. Therefore, this paper aims to provide a theoretical framework to systematically analyze the capacity of a multi-modal arterial, which could aid in policy-making. Single-mode transportation systems are often assessed in terms of capacity on the link/node level (Green-shields (1935); Lighthill and Whitham (1955); Allsop (1972); Guler and Cassidy (2012)), on the arterial level (Transportation Research Board (1997); Eichler and Daganzo (2006); Wu et al. (2011)), or on the network level (Mahmassani et al. (1984, 1987); Geroliminis and Daganzo (2008); Daganzo and Geroliminis (2008); Geroliminis and Sun (2011); Mahmassani et al. (2013)).

When multi-modal transportation systems are assessed in terms of capacity, the difficulty arises because bus and car performance need to be respectively analyzed and compared. Therefore, most literature analyzes the car capacity at specific levels of bus demands (e.g. Boyac and Geroliminis (2011); Geroliminis et al. (2014); Arnet et al. (2015)). The main drawback of analyzing bus and cars separately is that the improvement of one mode usually comes at the cost of another. To overcome this, researchers have tried to compute a single representative value for the system, for example the passenger capacity (Geroliminis et al. (2014); Chiabaut et al. (2014); Chiabaut (2015); Loder et al. (2017)), or the average person travel time (Zheng and Geroliminis (2013); Christofa et al. (2016)). The problem with analyzing transportation systems in terms of travel times is that it requires either elaborate calculations or simulations. This is especially problematic because many different cases need to be considered when more innovative strategies are introduced. Therefore, most works on innovative bus priority strategies have been restricted to the isolated intersection level (Nowlin and Fitzpatrick (1997); Wu and Hounsell (1998); Guler and Menendez (2014a,b); He et al. (2016)). Some have tried to analyze the problem at the arterial level, but with some assumptions which might limit their generality and applicability in practice (Eichler and Daganzo (2006); Truong et al. (2015); Christofa et al. (2016)). Moreover, when various strategies are combined on an arterial, or different systems need to be compared, the analysis becomes very tedious.

In order to add to the set of existing tools for the quantitative evaluation of multi-modal arterials, in this paper we build an analytical framework which is: i) robust: consistent with single-mode interpretations yet easily expandable to more modes, ii) versatile: applicable to both intuitive cases when the arterial capacity is allocated in a fixed manner, and non-intuitive cases when flexible sharing strategies are installed, iii) generalized: applicable to all flexible sharing strategies, iv) parsimonious: with very few additional parameters, v) practical: all parameters are easy to calculate in practice. This mathematical framework is formulated for a bi-modal arterial in Section 2. Based on this framework, mathematical definitions for improving bi-modal arterial capacity are discussed in Section 3. These two sections build the theoretical foundation for Section 4, where graphical tools and mathematical principles are established to improve the operation of a bi-modal arterial using flexible sharing strategies. In Section 5, a pre-signal strategy is

examined as an example of a flexible sharing strategy. Improvements in traffic operation, when pre-signals are installed according to our principles, are illustrated both mathematically and with a calibrated simulation. Section 6 offers some concluding remarks.

2. Mathematical formulation for the operation of a bi-modal arterial

For a given arterial with elements (links or nodes), denote these elements in sequence by $\{i|i \in I\}$, where 1 is the most upstream element and $|I|$ is the most downstream element. Denote the respective capacities of the elements with $\{Q_i|i \in I\}$ (in car equivalent unit/hour), which includes effects of the width of the road, the signal control of the intersection, the location of bus stops, and the speed limit of the arterial. Here, we assume that the capacities Q_i of the elements on the arterial are fixed. This is the case, for example, when an extra lane could not be added due to space restrictions, a bus stop cannot be relocated due to planning constraints, or green signals could not be extended due to high traffic demand from other approaches. Hence, the goal is to improve the operation of cars and buses on this arterial by implementing flexible sharing strategies between the two modes to better utilize the existing capacities of the arterial elements (including those determined by the existing signal controls).

When the sharing of capacity is achieved by dedicating lanes to each mode, the capacities for buses and cars are expressed as fixed portions of the element's total capacity. If out of the n_i lanes of element i , there are n_i^b bus lanes and n_i^c car lanes ($n_i^b + n_i^c = n_i$), the capacity Q_i of this element is separately reserved for buses $Q_i^b = n_i^b Q_i / n_i$ and cars $Q_i^c = n_i^c Q_i / n_i$, where $Q_i^b + Q_i^c = Q_i$. Note that here Q_i^b should be understood as the capacity equivalence (in car equivalent unit/hour), not the maximum flow. If the car equivalence of a bus is p , the constraints on bus input (q^b) and car input (q^c) such that no congestion is created are respectively $q^b \leq Q_i^b / p$ and $q^c \leq Q_i^c$. Therefore, there are various capacity measures used to describe the system: the maximum flow of cars (Q_i^c), the maximum flow of buses (Q_i^b / p), and the maximum flow of all vehicles ($Q_i^b / p + Q_i^c$). None of the above, alone, adequately captures the performance of the arterial. This problem becomes even more contentious for elements with mixed lanes or flexible sharing strategies, because the capacity of one mode usually depends on the input of the other mode. This makes any systematic comparison between bi-modal transportation systems difficult from the perspective of capacity. Therefore, the existing methodologies for single-mode transportation systems are no longer adequate to quantify capacity sharing between modes.

To resolve this problem, we first introduce some notations. First, we consider the isolated elements. The inputs to each element i of the arterial are the bus input q_i^b and the car input q_i^c . The outputs of each element are the bus flow $f_i^b(q_i^b, q_i^c)$, and car flow $f_i^c(q_i^b, q_i^c)$, which are both multivariate functions. At signalized intersections, the output of the intersection element will evidently vary with time, even for a constant input. However, the output in our definition is the average output over one cycle for fixed-time signals, or the average output over a long period of time for actuated signals. Mathematically, the element i is considered to be performing a vector transformation of the input vector, $f_i(q_i^b, q_i^c) = (f_i^b(q_i^b, q_i^c), f_i^c(q_i^b, q_i^c))$.

Next, we consider an arterial which is composed of multiple elements. To aid in the theoretical development, we will assume the arterial is closed, i.e. there is only traffic inflow in the first element and traffic outflow from the last element¹. With the closed arterial assumption, the operation of an arterial could be considered as a serial input-output system. The output of element i is the input for the element $i + 1$, but shifted in time by the free-flow travel time. For simplicity, in the rest of the paper we will exclude the time variable from the notation. However, the reader should be aware that there is a time lag between elements. Then, mathematically, the arterial could be considered to be performing a serial vector transformation of the

¹The conclusions of this paper will not deviate too far if the arterial is open and the total inflow and total outflow at each node element are similar to each other. For a discussion on the formulation for an open arterial, the reader can refer to He et al. (2017)

input vector of the most upstream element (i.e. the demand for the arterial). Mathematically we have $f(q_1^b, q_1^c) = f_{|I|} \dots f_3 f_2 f_1(q_1^b, q_1^c) = (f^b(q_1^b, q_1^c), f^c(q_1^b, q_1^c))$, where $f(q_1^b, q_1^c)$ is a composite function representing the entire arterial.

Note that, since the output cannot be greater than the input for either mode, i.e. $f_i^b(q_i^b, q_i^c) \leq q_i^b, f_i^c(q_i^b, q_i^c) \leq q_i^c$, there are only two possibilities for the output flows of a given input. If the input is within the capacity, the output is equal to the input, i.e. $f(q_i^b, q_i^c) = (q_i^b, q_i^c)$. Mathematically this means the input (q_i^b, q_i^c) is a fixed point of the corresponding transformation. If the input exceeds the capacity, then the output is smaller than the input, i.e. $|f(q_i^b, q_i^c)| < |(q_i^b, q_i^c)|$ ². Based on these two possibilities, we then establish the following mathematical definitions for the concept of ‘potential throughput’.

Definition 1: The potential throughput of a bi-modal system is the fixed set (i.e. the set of all fixed points) of the vector transformation corresponding to the road element or the arterial. Mathematically, $\Omega_i = \{(q_i^b, q_i^c) | f_i(q_i^b, q_i^c) = (q_i^b, q_i^c)\}$ is the potential throughput for an element i , and $\Omega = \{(q_1^b, q_1^c) | f(q_1^b, q_1^c) = (q_1^b, q_1^c)\}$ is the potential throughput for an arterial.

Definition 2: The input exceeds the capacity if it does not belong to the potential throughput. Mathematically, (q_i^b, q_i^c) exceeds the capacity of a bi-modal system when $|f_i(q_i^b, q_i^c)| < |(q_i^b, q_i^c)|$ for an element i , and when $|f(q_1^b, q_1^c)| < |(q_1^b, q_1^c)|$ for an arterial.

It should be noted that, while the capacity of a single-mode arterial is discussed in terms of a single number, our formulation discusses the capacity of a bi-modal arterial in terms of a set of vehicle throughputs. This set is used to quantify the performance a bi-modal facility.

It could be easily verified that the transformation function for elements with dedicated lanes for buses and cars is,

$$f_i^b(q_i^b, q_i^c) = \begin{cases} q_i^b, & pq_i^b \leq Q_i^b \\ \frac{Q_i^b}{p}, & pq_i^b > Q_i^b \end{cases} \quad f_i^c(q_i^b, q_i^c) = \begin{cases} q_i^c, & q_i^c \leq Q_i^c \\ Q_i^c, & q_i^c > Q_i^c \end{cases} \quad (1)$$

where $Q_i^b + Q_i^c = Q_i$. The potential throughput for elements with a dedicated bus lane is $\Omega_i = \{(q_i^b, q_i^c) | pq_i^b \leq Q_i^b \wedge q_i^c \leq Q_i^c\}$.

Assuming the buses and cars are proportionally distributed in the traffic stream and the queuing process follows a first-in-first-out (FIFO) discipline, the transformation function for elements with mixed lanes is,

$$f_i^b(q_i^b, q_i^c) = \begin{cases} q_i^b, & pq_i^b + q_i^c \leq Q_i \\ \frac{q_i^b}{pq_i^b + q_i^c} Q_i, & pq_i^b + q_i^c > Q_i \end{cases} \quad f_i^c(q_i^b, q_i^c) = \begin{cases} q_i^c, & pq_i^b + q_i^c \leq Q_i \\ \frac{q_i^c}{pq_i^b + q_i^c} Q_i, & pq_i^b + q_i^c > Q_i \end{cases} \quad (2)$$

The potential throughput for elements with mixed lanes is $\Omega_i = \{(q_i^b, q_i^c) | pq_i^b + q_i^c \leq Q_i\}$.

Flexible sharing strategies combine the advantages of mixed lanes (the element’s capacity can be fully utilized if buses are not present) and dedicated bus lanes (the buses receive priority). Hence, some properties related to a general flexible sharing strategy could be derived. We assume that all buses receive priority in this system, then we have $f_i^b(q_i^b, q_i^c) = q_i^b$ for $pq_i^b \leq Q_i$ and $f_i^b(q_i^b, q_i^c) = \frac{Q_i}{p}$ otherwise. Note that ensuring priority for all arriving buses is relatively easy as long as the necessary infrastructure is in place, since it only entails

²Here the notation $|(q^b, q^c)|$ represents the norm of the vector. In this particular case, with any general p-norm definition, the statement is mathematically equivalent to $f_i^b(q_i^b, q_i^c) < q_i^b \vee f_i^c(q_i^b, q_i^c) < q_i^c$.

blocking cars to allow the bus to proceed ahead of them. Since capacity is not reserved in a fixed manner for either buses or cars, cars will be able to utilize the remaining capacity. This is particularly helpful when the car flow is high while the bus flow is low. However, the remaining car capacity of the element is subject to an overall reduction factor due to the added complexity of the system. Here, for simplicity, it is assumed that the reduction is a linear function of the bus input, i.e. $0 \leq f_i^c(q_i^b, q_i^c) \leq Q_i - \beta_i q_i^b$, where β_i . The linearity simplification is justified by previous research on this subject (Guler and Cassidy (2012); Guler and Menendez (2014a)). If the relationship between car discharge capacity and bus input is non-linear, a first order approximation could be obtained around the bus demand. Note that β is the only parameter required in our model to evaluate the efficiency of any flexible sharing strategy. This factor can be estimated analytically, empirically or with simulations. An example will be demonstrated in Section 5. A small β means the implementation can efficiently share the element's capacity between modes. In particular, $\beta_i = p$ would represent a perfect sharing system. However, to the authors' best knowledge, in practice a system where bus priority can be provided without loss of capacity does not exist, hence in general $\beta_i > p$. Note also that β does not indicate anything about the level of bus priority, which should be checked separately. To summarize, elements with a flexible sharing strategy have a transformation function of the form:

$$f_i^b(q_i^b, q_i^c) = \begin{cases} q_i^b, & \text{if } pq_i^b \leq Q_i \\ \frac{Q_i}{p}, & \text{if } pq_i^b > Q_i \end{cases} \quad f_i^c(q_i^b, q_i^c) = \begin{cases} q_i^c, & \text{if } q_i^c \leq Q_i - \beta_i q_i^b \\ \max(Q_i - \beta_i q_i^b, 0), & \text{if } q_i^c > Q_i - \beta_i q_i^b \end{cases} \quad (3)$$

It can then be verified that the potential throughput for an element with a flexible sharing strategy is $\Omega_i = \{(q_i^b, q_i^c) | \beta_i q_i^b + q_i^c \leq Q_i\}$ ³. In Eq.3, notice that the transformation function for buses is the same as that for dedicated lanes. However, the transformation function for cars will result in a car output somewhere in between mixed lanes and dedicated lanes. Moreover, although the transformation functions in Eq.2 and Eq.3 are different, their corresponding potential throughputs Ω_i are of very similar mathematical forms. Hence, in the rest of the paper elements with mixed lanes and flexible sharing strategies are often discussed together.

In reality, it is often possible to estimate the range of bus and car vehicle demands upstream of the arterial, either with empirical studies or demand modelling. We denote the set of all upstream bus and car vehicle demand combinations (q^b, q^c) as D , and call it the demand set. For simplicity, we can approximate the demand set with a rectangular area, i.e. $q^b \leq D^b$ and $q^c \leq D^c$. If $Q^b < D^b$ or $Q^c < D^c$ for elements with dedicated bus lanes, or if $Q < D^b + D^c$ for elements with mixed lanes, then the demand exceeds capacity. Note also that in most urban arterials the bus demand is much lower than the car demand, and often far less than the capacity of the reserved bus lane(s), i.e. $D^b \ll D^c, D^b \ll Q_i^b$. This characteristic ensures that the allowable car flow in Eq.3b is not too low, i.e. $Q_i - \beta_i q_i^b \gg 0$.

For visualization purposes, the transformation functions of Eq.1, Eq.2 and Eq.3 are graphically illustrated in Figure 1. Notice that the output for an input that falls within the potential throughput is the input point itself. For an input outside the potential throughput, the output lies on the boundary of the potential throughput.

3. Mathematical definitions for improving bi-modal arterial capacity

While improving the capacity of a single-mode arterial clearly means increasing the capacity value, improving the capacity of a bi-modal arterial could be achieved in different ways. The formulations in Section 2

³Note that points with $q_i^c = 0 \vee \frac{Q_i}{\beta} \leq q_i^b \leq \frac{Q_i}{p}$ are also fixed points of Eq.3, but in practice we do not consider such theoretical situations without any cars.

Figure 1: Illustration of the transformations of three basic elements.

allow for a mathematical definition of improvement, so the operation of different bi-modal systems can be quantitatively compared. We compare the potential throughputs of different systems by comparing the sets of vectors, instead of comparing the capacity. Since there are many different ways to compare two sets, in this section we discuss four possible definitions of what could be meant by an improvement (i.e. we define four different types of improvements),

Definition 3: If the vector transformation corresponding to one bi-modal transportation system has the potential throughput Ω and the vector transformation corresponding to another bi-modal transportation system has the potential throughput Ω' , then the former transportation system is an absolute improvement over the latter if $\Omega' \subset \Omega$.

Note that the above comparison between bi-modal systems is compatible with traditional capacity comparison for a single mode (i.e. only cars). If we have $\{q_i^c \mid q_i^c \leq Q_i^{c'e}\} \subset \{q_i^c \mid q_i^c \leq Q_i^c\}$, then we have $Q_i^{c'e} < Q_i^c$. The definition of this comparison is termed ‘absolute improvement’ since the former system is better, irrespective of the location of the arterial or the time of the day, i.e. it is an improvement irrespective of the demand. Since both the bus and car capacities are improved in this case, this is a Pareto improvement. An absolute improvement is usually difficult to achieve. However, the arterial can be improved based on a given demand set according to the definition below.

Definition 4: If the vector transformation corresponding to one bi-modal transportation system has the fixed set Ω and the vector transformation corresponding to another bi-modal transportation system has the fixed set Ω' , given the demand set D of the arterial, then the former transportation system is a demand-based improvement over the latter if $\Omega' \cap D \subset \Omega \cap D$.

It is easy to deduce that the demand-based improvement is a necessary but not sufficient condition for the absolute improvement. Hence, the name ‘demand-based’, because this is true only for the given arterial with the known demand, and might not be true at another location with a different demand pattern. Applying such improvements (e.g. by installing innovative bus priority strategies) in the city must be analyzed on a case by case basis for each arterial. Even demand-based improvements are not always achievable. Very often the improvement of one mode comes at the cost of the other mode (i.e. it is not a Pareto improvement). Mathematically, the potential throughputs of such strategies do not possess a subset relationship, and hence do not lead to demand-based improvement. Nevertheless, another possible definition for improvement could be formulated based on what portion of the demand set is covered by the potential throughput.

Definition 5: If the vector transformation corresponding to one bi-modal transportation system has the fixed set Ω and the vector transformation corresponding to another bi-modal transportation system has the fixed set Ω' , given the demand set D of the arterial, then the former transportation system is a coverage-based improvement over the latter if $A(\Omega' \cap D) < A(\Omega \cap D)$ ⁴.

⁴Here the notation $A(\Omega)$ represents the Lebesgue measure of the set Ω . Graphically, it is the area covered by the set on the (q^b, q^c) plane.

It is easy to deduce that the coverage-based improvement is a necessary but not sufficient condition for the demand-based improvement (and hence also for the absolute improvement). It should be noted that the covered area does not differentiate whether it includes regions with higher bus flows or regions with higher car flows. Therefore, this is not a Pareto improvement because the improvement of either bus or car capacity comes at the cost of the other mode. Moreover, certain regions of demand could be of a higher weight. However, since the weight of different modes is subjective, the rest of the paper will focus on absolute improvements and demand-based improvements⁵.

The three types of improvements above are graphically illustrated in Figure 2 for comparison. In the figures, Ω represents the fixed set for an arterial with car capacity Q^c and bus capacity Q^b ; Ω' represents the fixed set for an arterial with car capacity $Q^{c'}$ and bus capacity $Q^{b'}$.

Figure 2: Absolute improvement, demand-based improvement, and coverage-based improvement.

According to the above definitions, in order to compare different bi-modal transportation systems, mathematically we have to compare the fixed sets of the corresponding output functions. This could be tedious because the functional forms of Eq.2 and Eq.3 are complicated. Additionally, it is tedious to derive an explicit expression for the multivariate composite function corresponding to the arterial and its fixed set, especially when there are more than two elements. Therefore, we establish here two theorems to simplify the comparison. The proofs of the theorems are included in Appendixes A and B, respectively.

Theorem 1: Suppose Ω is the fixed set for an arterial with mixed lanes or flexible sharing strategies. If a certain demand combination $(q_0^b, q_0^c) \in \Omega$, then for any other fixed set Ω' such that $\Omega' \subset \{(q^b, q^c) \mid q^b \leq q_0^b \wedge q^c \leq q_0^c\}$, we have $\Omega' \subset \Omega$.

This theorem simplifies the comparison between fixed sets with various forms. Notice that among dedicated bus lanes, mixed lanes, and flexible sharing strategies, the mathematical form of fixed sets is the simplest for dedicated bus lanes. When considering a facility (with fixed set Ω), Theorem 1 enables us to use a hypothetical facility with dedicated bus lanes as a proxy to compare with another facility (with fixed set Ω'). The hypothetical facility has a capacity share of (q_0^b, q_0^c) , which belongs to the potential throughput of the considered facility (i.e. $(q_0^b, q_0^c) \in \Omega$). If this hypothetical capacity share results in a capacity improvement, then the considered facility is also an improvement (i.e. $\Omega' \subset \Omega$).

Theorem 2: If Ω_i are the fixed sets of the corresponding transformation f_i of element i , and Ω is the fixed set of the transformation $f = f_{|I|} \dots f_3 f_2 f_1$ of the arterial, then $\Omega = \Omega_1 \cap \Omega_2 \cap \dots \cap \Omega_{|I|}$.

This means that if a given input is within the capacity of all individual elements of the bi-modal system, then it is within the capacity of the system. This is true because all our transformations f_i are contractive. Note that, for a single-mode system with only cars, Theorem 2 could be interpreted as the usual statement that the capacity of the arterial is the capacity of the most restrictive element (when there is no spillback

⁵With an appropriate weight for the different modes, the reader could also formulate a weighted-coverage definition for improvement.

within the system). For a bi-modal arterial, Theorem 2 asks for separate analysis of the transformation of each individual element to simplify the comparison. An implicit assumption of Theorem 2 is that the traffic signals along the arterial are properly coordinated such that upstream intersections do not create unutilized green times for downstream intersections. In particular, the signal coordination must not be so poor as to cause unnecessary spillbacks.

To illustrate Theorem 2, we first examine the case with dedicated bus lanes, and then the case with mixed lanes or flexible sharing strategies. Index the elements with dedicated bus lanes with $j, j \in J, J \subseteq I$. Their fixed sets are of the form $\{(q^b, q^c) \mid pq^b \leq Q_j^b \wedge q^c \leq Q_j^c\}$. Graphically, each fixed set forms a rectangle with the coordinate axes. The intersection of all these fixed sets is $\{(q^b, q^c) \mid pq^b \leq \min_{j \in J} Q_j^b \wedge q^c \leq \min_{j \in J} Q_j^c\}$, which also forms a rectangle with the coordinate axes, with a boundary given by the lower envelope of all the individual rectangles superimposed. Index the elements with mixed lanes or flexible sharing strategy with $k, k \in K, K \subseteq I$. Their fixed sets are of the form $\{(q^b, q^c) \mid \beta_k q^b + q^c \leq Q_k\}$, where $\beta_k = p$ for mixed lanes. Graphically, each fixed set forms a triangle with the coordinate axes. The intersection of all these fixed sets is $\{(q^b, q^c) \mid \beta_k q^b + q^c \leq Q_k, \forall k \in K\}$, which forms a convex polygon with the coordinate axes, with a boundary given by the lower envelope of all the individual triangles superimposed. Overall, note that the constraints are linear for all elements discussed in this paper, and the resulting fixed sets are convex. This can be used to formulate optimization problems on bi-modal arterials as a linear programming problem. An example is illustrate in Section 4.2.

Theorems 1 and 2 may still seem rather abstract. To illustrate how Theorems 1 and 2 could be used in practice to quickly gauge and compare arterial capacities, we discuss here an example of the arterial depicted in Figure 3. The arterial has in total 8 elements, including 4 links and 4 nodes. The two upstream elements (1-2) have two lanes each, and all the downstream elements (3-8) have three lanes each. A bus lane (the shaded lane in Figure 3) runs throughout the arterial. Assume all the signals have a cycle length of 60s, and the green time duration for elements 2, 4, 6 and 8 are respectively 20s, 30s, 20s, and 15s. Further assume each lane has a capacity of λ . Simple calculations give the theoretical bus and car capacities for each element as indicated in Figure 3. Hence, as discussed above, according to Theorem 2 the potential throughput of this arterial is $\Omega' = \Omega_1 \cap \Omega_2 \cap \dots \cap \Omega_8 = \{(q^b, q^c) \mid pq^b \leq \min q_i^b \wedge q^c \leq \min q_i^c\} = \{(q^b, q^c) \mid pq^b \leq \frac{1}{4}\lambda \wedge q^c \leq \frac{1}{3}\lambda\}$. The most restrictive element for buses is element 8 and for cars is element 2. Let's assume now that a flexible sharing strategy replaces the bus lane in element 2 and can achieve a capacity share of $\frac{1}{4}\lambda$ for buses and $\frac{3}{8}\lambda$ for cars, i.e. $(\frac{1}{4}\lambda, \frac{3}{8}\lambda) \in \Omega$, where Ω is the fixed set of the new arterial. This is possible because the total capacity $\frac{5}{8}\lambda$ is less than the capacity $\frac{2}{3}\lambda$ of the element. We would like to know if the new configuration is an improvement, without deriving the explicit expression for Ω , which could be tedious due to the flexible sharing element. To do this, we suppose a hypothetical dedicated bus lane system with the capacity share $(Q^b = \frac{1}{4}\lambda, Q^c = \frac{3}{8}\lambda)$ is implemented on element 2. Then it is very easy to find the potential throughput of this hypothetical system, i.e. $\{(q^b, q^c) \mid pq^b \leq \frac{1}{4}\lambda \wedge q^c \leq \frac{3}{8}\lambda\}$. Note that Ω' is a subset of this hypothetical fixed set, then from Theorem 1 we have $\Omega' \subset \Omega$, and the flexible sharing strategy will improve the arterial capacity.

Figure 3: An example of an arterial.

4. Improving the operation of a bi-modal arterial

4.1. Systematic evaluation of an arterial for installing flexible sharing strategies

While it may be possible to implement flexible sharing strategies at every element of the arterial to increase capacity, this might not be efficient or cost-effective because similar results might be achieved by focusing only on a few ‘critical bottlenecks’. It is important then to provide the practitioners with the right tools to systematically evaluate any given bi-modal arterial. The following three questions are likely most relevant: i) Will installing flexible sharing strategies improve the capacity of the arterial? ii) If yes, where should flexible sharing strategies be installed? iii) What are the requirements for the flexible sharing strategies to be installed?

The first question could be answered by graphically drawing the potential throughput of the given arterial. Here we examine the fixed set for the transformation function of the arterial when the arterial has all three types of elements. Since $I = J \cup K$, according to Theorem 2 the fixed set for the arterial is $\{(q^b, q^c) \mid pq^b \leq \min_{j \in J} Q_j^b \wedge q^c \leq \min_{j \in J} Q_j^c\} \cap \{(q^b, q^c) \mid \beta_k q^b + q^c \leq Q_k, \forall k \in K\}$. There are three possible outcomes of this intersection as graphically illustrated in Figure 4. Again, note that the overall constraints are linear and the arterial fixed set forms a convex polygon graphically. Figure 4a illustrates the case where the capacity is limited by the elements with mixed lanes or flexible sharing strategies. Therefore, further installing flexible sharing strategies on elements with dedicated bus lane will not help in this case (hereafter named a Type A arterial). Although making the existing sharing strategies more efficient might marginally improve the arterial capacity, substantial improvement could only be achieved by expanding the capacity of some bottleneck elements (for example by adding more lanes or extending the effective green time duration of a traffic light). Figure 4b illustrates the case where the capacity is limited by the elements with dedicated bus lanes (hereafter named a Type B arterial). Note that Type B arterials include the case where there are dedicated bus lanes throughout the arterial. For a Type B arterial, removing the capacity restriction due to the bottleneck element(s) will definitely improve the arterial capacity. Figure 4c illustrates the case where the capacity is limited by both the elements with mixed lanes or flexible sharing strategies, and the elements with dedicated bus lanes (hereafter named a Type C arterial). For a Type C arterial, removing the capacity restriction due to the bottleneck element(s) might improve the arterial capacity.

In summary, installing flexible sharing strategies might improve capacity of Type B and Type C arterials. Next, we establish two mathematical conditions to select the elements to install flexible sharing strategies for Type B and Type C arterials, respectively. This analytical procedure helps to answer question ii) and iii) stated in the beginning of the section. The proofs of the conditions are included in the Appendixes C and D, respectively.

Figure 4: Graphical tool to classify a given bi-modal arterial.

Condition 1: For a Type B arterial, the capacity of the arterial can be improved if the following condition

is NOT satisfied:

$$\exists j^* \in J, \text{ such that } Q_{j^*}^b = \min_{j \in J} \{Q_j^b\} \wedge Q_{j^*}^c = \min_{j \in J} \{Q_j^c\} \quad (4)$$

This means that for the optimal operation of the arterial, there must exist at least one element which is both the bottleneck for buses and the bottleneck for cars. Note that a sufficient but not necessary condition for Eq.4 is $\operatorname{argmin}_{j \in J} \{Q_j^c\} = \operatorname{argmin}_{j \in J} \{Q_j^b\}$, which means that all the car bottlenecks are also the bus bottlenecks.

This is, however, not necessary because there could be multiple bottlenecks of one mode, with only some being also the bottleneck of the other mode. Condition 1 also leads to the following corollary to improve the capacity.

Corollary 1: When Eq.4 is not met for a Type B arterial, the capacity of the arterial could be improved by redistributing the bus and car capacities at the car bottleneck $j \in \operatorname{argmin}_{j \in J} \{Q_j^c\}$, if the flexible sharing

$$\text{strategies to be implemented satisfy } \beta_j \leq \frac{Q_j - \min_{j \in J} \{Q_j^c\}}{\min_{j \in J} \{Q_j^b\}}.$$

This means that if none of the car bottlenecks is a bus bottleneck, then increasing its car discharge capacity while decreasing its bus discharge capacity could improve the operation of the arterial. It also gives the requirement for the flexible sharing strategy to be installed. To further illustrate Condition 1 and Corollary 1, consider again the example in Figure 3. Hence, the bus bottleneck is element 8, while the car bottleneck is element 2. Therefore, Condition 1 is not satisfied. From Corollary 1, the capacity of the arterial could be improved by redistributing the bus and car capacities at element 2 with some flexible sharing strategy, for example as discussed previously in Section 3 with a capacity share of $\frac{1}{4}\lambda$ for buses and $\frac{3}{8}\lambda$ for cars. In this case, element 2 becomes both a bus bottleneck and a car bottleneck, hence Condition 1 is now satisfied. A similar approach could be to implement flexible sharing strategies on element 8 to increase the bus capacity while reducing the car capacity. In this way, element 8 could also become both a bus bottleneck and a car bottleneck. However, since car queues are the problem in most urban arterials, the former approach is more likely in practice (recall the assumption that $D^b \ll D^c$).

For a Type B arterial which satisfies Condition 1, there exists an element (usually a signalized intersection) which is both the bottleneck for buses and the bottleneck for cars. The next step is to estimate the maximum bus demand D^b and car demand D^c , and hence the demand set D . To further improve the given Type B arterial, or to improve a Type C arterial, we notice that for any given bus flow, the car flow could either be limited by the elements $j \in J$, $f^c(q^b, q^c) \leq \min_{j \in J} \{Q_j^c\}$, or the elements $k \in K$, $f^c(q^b, q^c) \leq \min_{k \in K} \{Q_k - \beta_k q^b\}$.

We therefore check the following condition.

Condition 2: For a Type B arterial which satisfies condition 1 or a Type C arterial, the capacity of the arterial can be improved if the following condition is satisfied:

$$\min_{j \in J} \{Q_j^c\} < \min \{ \min_{k \in K} \{Q_k - \beta_k D^b\}, D^c \} \quad (5)$$

When the above condition is satisfied, the minimum car capacity on the dedicated bus lane elements is less than the flow allowable on the mixed lane elements. This means that on the dedicated bus lane segments capacity might be over-reserved for buses. Hence, for the optimal operation of the arterial, the bottleneck of the arterial (which is both the bottleneck for buses and cars) must not over-reserve capacity for buses and waste capacity for cars (assuming car demand is much higher than bus demand). The bottleneck car capacity should be increased as much as possible (up to the highest car demand) without reducing the arterial bus capacity. It is important for the bus capacity to be kept intact, because more people can be moved

by bus within the limited road space. Condition 2 also leads to the following corollary to improve the capacity.

Corollary 2: When Eq.5 is not met for a Type B arterial (which satisfies Eq.4) or a Type C arterial, the capacity of the arterial could be improved by redistributing the bus and car capacities at the car bottleneck $j \in \operatorname{argmin}_{j \in J} \{Q_j^c\}$, if the flexible sharing strategies to be implemented satisfy $\beta_j \leq \frac{Q_j - \min_{j \in J} \{Q_j^c\}}{D^b}$.

This means that if the capacity allocation between buses and cars at the bottleneck is not optimal, flexible sharing strategies should be installed at these bottleneck elements to achieve a better capacity allocation, while still serving the entire bus demand. It also gives the requirement for the flexible sharing strategy to be installed. To further illustrate Condition 2 and Corollary 2, we continue with the example in Figure 3. Assume we know the bus demand on this arterial is low, for example at $\frac{1}{100}\lambda$ (this is approximately 20 buses/hour), and the car demand is high, for example at λ . After a flexible sharing strategy has been installed on element 2 according to Corollary 1, we notice that the car demand λ still exceeds the arterial car capacity $\frac{3}{8}\lambda$ while the bus demand (which is much smaller than $\frac{1}{10}\lambda$) is lower than the arterial bus capacity $\frac{1}{4}\lambda$. This means much of the precious discharging capacity of the bottleneck intersection is wasted. The capacity of the arterial could be improved if we redistribute the bus and car capacities at element 2 and element 8. For example, we can implement a better flexible sharing strategy which could achieve a capacity share of $\frac{1}{10}\lambda$ for buses and $\frac{8}{15}\lambda$ for cars on both elements 2 and 8 (this is possible because the total capacity $\frac{19}{30}\lambda$ is less than the capacity of both elements). This results in a demand-based improvement of the arterial capacity.

In summary, the capacity of the arterial could be improved by identifying the critical bottlenecks and redistributing the bus and car capacities there. This redistribution, however, could not be achieved with the traditional method of implementing dedicated bus lanes. This is because adding or taking away one dedicated bus lane changes the capacity distribution by one full lane, while the two proposed principles above require much finer tuning of the capacity distribution. Therefore, flexible sharing strategies are advantageous and could be implemented to make absolute or demand-based improvement of the arterial capacity.

4.2. Optimization of person-throughput

Another common concern in bi-modal arterials is its person-throughput, which can represent the balance between car throughput and bus throughput. One way to look at person-throughput, is to compare the maximum person-throughput of two arterial configurations. Recall from Figure 4 that for all types of bi-modal arterials, the overall constraints on the potential throughput are linear, and the resulting fixed set forms a convex polygon. Denote f^p as the person throughput, we have $f^p(q^b, q^c) = X^b f^b(q^b, q^c) + X^c f^c(q^b, q^c)$, where X^b and X^c are the passenger occupancies of buses and cars, respectively. To find the optimal person-throughput, we need to solve the linear program $\max f^p(q^b, q^c)$ subject to $(q^b, q^c) \in \Omega$. This could easily be done graphically for a given arterial with a known potential throughput, as illustrated in Figure 5 below. The optimal person throughput of two arterials can also be directly compared graphically.

Another aspect of person-throughput is to look at over-saturated situations. To analyze this, we first look at the person-throughput of a single element. The person throughput from an element with bus lanes, mixed lanes, and a flexible sharing strategy could be easily compared with Eq.1, Eq.2, and Eq.3. For an element i with n_i lanes and total capacity Q_i , comparing using mixed lanes, implementing n_i^b bus lanes, or implementing a flexible sharing strategy with reduction factor β_i , the flexible sharing strategy results in the

Figure 5: Comparing optimal person-throughput of two bi-modal arterials: in this case arterial 2 with potential throughput 2 has a higher person-throughput for the given passenger occupancy levels.

highest person throughput when,

$$\begin{cases} q^b < \frac{n_i Q_i}{n_i^b \beta_i} \\ q^c > \frac{X_i^b - X^c}{X^b - \beta X^c} Q_i - q^b \\ X^b > \beta X^c \end{cases} \quad (6)$$

This means the bus flow cannot be too high to avoid blocking the discharge of cars, and the car flow cannot be too low for the flexible sharing strategy to help discharge more cars. Lastly, the bus occupancy must be high enough to justify its priority over cars. Although the above three conditions are easy to obtain by comparing the output functions of the elements, this does not provide much insight into the nature of the problem. We therefore provide below a qualitative explanation.

Figure 6: Transformation of input demands with three different strategies.

As stated before in Section 2, in most urban arterials we have $D^b \ll D^c$. Under this assumption, the demand set D in relation to different possible fixed sets could be divided into four areas. The division and outputs for three different strategies are illustrated in Figure 6. For demands within area 1, the input will be within capacity of all three considered systems. Hence, the person throughput of all three systems are the same. For demands within area 2, the input exceeds the capacity when bus lanes are implemented, but is still within the capacity if mixed lanes or a flexible sharing strategy is implemented. Hence, the latter two

result in the same person throughput, which is higher than when bus lanes are implemented. For demands within area 3, the input exceeds the capacity when bus lanes or a flexible sharing strategy is implemented, but is still within the capacity if mixed lanes are implemented. In this case, implementing mixed lanes will result in the highest person throughput. Lastly, for demands within area 4, the capacity is exceeded with whichever of the three systems is implemented. In this case, the mixed lanes system results in the lowest bus flow output while the bus lanes system results in the lowest car flow output. The system with a flexible sharing strategy achieves the same bus throughput as with bus lanes, while maintaining a high level of car throughput. Consequently, the system with a flexible sharing strategy will result in the highest person throughput when the passenger occupancy of buses is high enough. To summarize, the only situation when implementing a flexible sharing strategy on this element does not result in the highest person throughput, is when the demand falls within area 3 and a small part of area 4. With most demands, the flexible sharing strategy results in either the highest person throughput, or the same person throughput as with some other implementations.

Eq.6 is satisfied in most cases with our assumption that the bus flow is low while the car flow is high. Therefore, we see that implementing flexible sharing strategies at the critical bottlenecks will not only improve the capacity, but for most over-saturated demands it will also improve the person throughput.

4.3. Analysis of spillbacks

Since each link element has a limited physical length and storage capacity, congestion at the bottleneck element could eventually spill back to upstream elements (Daganzo (2007); Liu and Chang (2011)). A queue spillback would require a large demand sustained over a period of time. If a high demand is expected for long enough, our framework can easily predict whether or not a spillback would occur since the capacity of each element is known. Unfortunately, for temporarily large demands our framework could not predict whether or not a spillback would occur since the temporal and spatial dimensions of queues are not considered as part of this framework. However, a queueing analysis which takes into consideration the capacity and demand for each element could be conducted on elements to determine whether spillbacks would be expected.

An important consideration of flexible sharing strategies or dedicated bus lanes is that a dedicated bus lane would need to be introduced on an element where it did not exist before (i.e., mixed lanes would need to be changed to dedicated lanes). This would create a bottleneck due to sorting of vehicles. At this bottleneck, all upstream lanes are accessible by cars, but some lanes are reserved for buses downstream. For cars, clearly this is a merge bottleneck due to lane reduction. For buses, although this location is unlikely a bottleneck (as $D^b \ll Q^b$), they will still need to travel through the car queue to access the downstream bus lane. Hence, for buses this can be modelled as a diverge bottleneck (Newell (1999)), where buses in the queue are not constrained by their own downstream discharge capacity, but are delayed from reaching their destination due to car queues. We adopt below a similar approach to that from Newell (1999). The analysis below discusses a transition area between element $i - 1$ and element i .

We start from the time $t = 0$ when the arrival of the vehicles first exceeds the capacity of the bottleneck and a queue starts. Let $A^b(t), A^c(t)$ be the cumulative number of buses and cars, respectively, which would have entered element i by time t if there were no queue (i.e. the virtual arrival accumulation, hereafter termed the cumulative arrival). Let $H^b(t), H^c(t)$ be the cumulative number of buses and cars, respectively, which actually enter element i by time t (i.e. the actual cumulative departure, hereafter termed the cumulative departure). The total cumulative arrival and cumulative departure of all vehicles are, respectively, $A(t) = A^b(t) + A^c(t)$ and $H(t) = H^b(t) + H^c(t)$.

We consider here steady states with stable demands. Assume for a time period $[0, t_0]$, the car demand exceeds the capacity of the downstream car lane, $\frac{dA^c}{dt} = q_i^c > Q_i^c (0 \leq t \leq t_0)$. Afterwards, the car demand drops to $\frac{dA^c}{dt} = q_i^c < Q_i^c (t > t_0)$. Assume the bus demand, on the other hand, remains below capacity as scheduled, $\frac{dA^b}{dt} = q_i^b < Q_i^b (\forall t)$. We use the geometry of cumulative arrival and departure curves to determine how

long it takes for this excess accumulation to dissipate. The solution is drawn in Figure 7, and the analysis is provided below.

Figure 7: Cumulative arrival and departure curves.

To approach this problem, assume all vehicles queue in a FIFO system. Then, cars and buses will be discharged from the bottleneck element according to their proportions in the queue. Here we conservatively assume that both cars and buses merge into the car lane and then the buses are sorted into the bus lane. In this case, the total discharge rate of the bottleneck element would be Q_i^c . Although the total discharge rate is constant, the proportion of buses and cars discharged depends on their arrival rates. Evidently, up to time t_0 , the discharge proportion is the same as the arrival proportion. Hence we have, $\frac{dH^b}{dt} = \frac{q_i^b}{pq_i^b + q_i^c} Q_i^c$ and $\frac{dH^c}{dt} = \frac{q_i^c}{pq_i^b + q_i^c} Q_i^c$ for $0 \leq t \leq t_0$.

The car arrival rate drops at time t_0 such that the total discharge rate exceeds the total arrival rate and the queue starts to dissipate. However, the proportion between bus and car discharges remains the same because the bottleneck is still discharging vehicles which have arrived before t_0 (i.e., discharge proportion is not the same as arrival proportion). For a FIFO system, the discharge proportion will not change until all vehicles which arrived before t_0 have been discharged. Denote this time as t_1 . On the cumulative curves ($t_1 - t_0$) is the horizontal distance between the curves $A^c(t)$ and $H^c(t)$ at the height of $A^c(t_0) = q_i^c t_0$. It can be calculated from the geometry that $t_1 = \frac{q_i^c}{Q_i^c} t_0$. Note that during the period $[t_0, t_1]$, the excessive accumulation of buses continues to increase because the bus arrival rate still exceeds the bus discharge rate.

After t_1 , the discharge proportion between buses and cars changes to the new arrival proportion that started at t_0 . Then, the discharge rates of buses and cars both exceed their respective arrival rates, and the excess accumulation dissipates at time T ,

$$T = \frac{q_i^c - q_i'^c}{Q_i^c - q_i'^c - q_i^b} t_0 \quad (7)$$

Therefore, for each transient period of high car demand, t_0 , the temporary excess accumulation upstream of element i can fully dissipate if there is a long enough period, T . Moreover, if the maximum accumulation (after t_0) could be stored within the link, the temporary excess accumulation will not block the upstream intersection. In this case, there is no spillback and our analytical framework is applicable. Otherwise, if

there are frequent high car demands arriving at the transition area, or the link does not have enough storage capacity upstream of the transition area, the queue cannot dissipate in time and will eventually propagate to the upstream intersection. This results in detrimental spillbacks. To avoid this, Eq.7, along with the maximum queue length, should be used to ensure there will be no spillbacks when implementing flexible sharing strategies or dedicated bus lanes.

5. A case study using a pre-signal as a flexible sharing strategy

On an urban arterial, signalized intersections are often the arterial bottlenecks. This section examines how a previously proposed bus-priority strategy for signalized intersections, namely the pre-signal, could work as a flexible sharing strategy. We use the pre-signal as an example in this section to illustrate how flexible sharing of capacity between modes could be achieved, and how the operation of the arterial could be improved.

A pre-signal could be installed when there is a dedicated bus lane upstream of the signalized intersection. The pre-signal terminates the dedicated bus lane immediately upstream of the intersection i and stops cars at this location. Therefore, when the pre-signal is red, buses can jump the car queue and hence bus priority is provided. Meanwhile, cars can use all the lanes at the main signal to discharge, so that the capacity of the intersection is allocated in a flexible manner. A typical configuration of a pre-signal is illustrated in Figure 8. Note that the downstream element ($i + 1$) could either be immediately converted back to an element with dedicated bus lane (by installing a transition area upstream of this element), or could continue to be an element with mixed lanes. Different operating strategies of pre-signals have been proposed and implemented (for example, in UK, Denmark, and Switzerland). This paper assumes a previously proposed operating strategy with exact formulas given in (Guler and Menendez (2014a)). Four main principles behind this strategy are listed below.

Figure 8: Configuration of a typical pre-signal.

1. The pre-signal has the same cycle length as the main signal. It does not affect the operation of the main signal.
2. The pre-signal turns red before the main signal such that the last discharged car from the pre-signal must pass the main signal. Therefore, the pre-signal will not stop any car for an additional cycle.
3. When a bus is not present, the pre-signal is red for the maximum possible duration such that it does not become an additional bottleneck.
4. When a bus is detected, the pre-signal is red for a duration R_b seconds in advance of the bus arrival at the pre-signal, so that the space upstream of the main signal is cleared.

Next, we provide some quantitative analysis of pre-signals in terms of our definition of flexible sharing strategies. Suppose the element i is a signalized intersection. The element $(i - 1)$ is the link immediately upstream of the intersection i . Note that originally when the bus lanes are implemented, both elements have the same number of bus lanes ($n_i^b = n_{i-1}^b$) and car lanes ($n_i^c = n_{i-1}^c$). The signal has a cycle length of C_i seconds and effective green duration G_i seconds, and we have $Q_i = \frac{G_i}{C_i} Q_{i-1}$. Evidently, the fixed set of this intersection is $\{(q_i^b, q_i^c) | pq_i^b \leq (1 - \alpha_i) Q_i \wedge q_i^c \leq \alpha_i Q_i\}$, where $\alpha_i = \frac{n_i^c}{n_i^b + n_i^c}$.

We discuss here the properties of a properly timed pre-signal (according to the four principles above), which is to be installed immediately upstream of intersection i . From principle 1, its cycle length should also be C_i seconds. When a bus is not detected, the pre-signal has a fixed green duration of G_{pi} seconds. Principle 3 gives that $G_{pi} = \frac{G_i}{\alpha_i}$, which ensures that the pre-signal can discharge as many cars as the main signal, i.e. the overall constraint on car flow remains $q^c \leq Q_i$. This, however, will be reduced in the presence of buses. Here we assume there is at most one bus arriving at the intersection during every signal cycle. This is a reasonable assumption about bus flow for most urban arterials. Under this assumption, the bus either arrives when the pre-signal is green, resulting in an additional R_b seconds of red pre-signal duration; or arrives when the pre-signal is red, resulting in no additional red. Hence, the total time of additional red pre-signal triggered by buses is on average $\frac{G_{pi}}{C_i} q^b R_b$ per hour (bus flow in vehicle/hour)⁶. The capacity loss per hour due to the additional red of pre-signal is $\frac{G_{pi}}{C_i} q^b R_b Q_{i-1} = \frac{R_b Q_i}{\alpha_i} q^b$. Therefore, the constraint on car flow is $q^c \leq Q_i - \beta_i q^b$, where $\beta_i = \frac{R_b Q_i}{\alpha_i}$. This is exactly as defined in Eq.3, hence the pre-signal works as a flexible sharing strategy. Note that in this case, the linearity assumption is theoretically exact. Compared to the output function of the dedicated bus lane, installing the pre-signal results in a higher car discharge capacity for bus flow $q^b \leq \frac{1-\alpha_i}{\beta_i} Q_i$. Note that in general we have $D^b \leq \frac{1-\alpha_i}{\beta_i} Q_i$, so installing a pre-signal improves the capacity based on most given demands. In particular, during periods when a bus is not present, the car discharge capacity when a pre-signal is installed is $\frac{1}{\alpha_i}$ times of that before.

Evidently, the theoretical calculation above is not able to account for many practical factors, such as bounded acceleration, reaction time, and lane changes. To verify the theoretical calculations and to illustrate some typical values of the coefficients, a micro-simulation model for an isolated intersection is built for an example case using VISSIM. A micro-simulation for the arterial will be analyzed subsequently in this section. Both simulations are calibrated and validated based on empirical data from a pre-signal in Zurich. More details about the calibration and validation can be found in (He et al. (2014)). The layout of the isolated intersection in the simulation model is exactly as illustrated in Figure 8, with one bus lane and one car lane. Table 1 summarizes the variables used in this simulation model (assumed or calculated). With these values, the theoretical linear factor for this pre-signal can be calculated using the analytical expression derived above. This results in $\beta_i=7.8$.

Table 1: Variables used in the simulation model.

n_i^c	1
n_i^b	1
α_i	0.5
Q_{i-1}	4200veh/h
C_i	60s
G_i	20s
Q_i	1400veh/h
G_{pi}	40s
R_b	10s

In the simulation model, a maximum bus demand of 60buses/hour is assumed (i.e. $D_b=60$ buses/hour). This is reasonable because in most urban arterials, the bus headway is larger than 1 minute. 7 levels of bus flow between 0 and 60buses/hour are tested. The car capacity (i.e. maximum car flow) at the intersection is obtained from the simulation for each level of bus flow. Each result is averaged over 10 simulation runs (each with a different random seed) and summarized in Figure 9. The car capacities obtained in the simulation

⁶It is possible that buses arrive near the boundary of the red/green phases of the pre-signal, resulting in less than R_b seconds of red pre-signal. However, the reader could easily verify that the results remain the same.

are fitted using a least squares method. The linear fit is very good ($R^2=0.98$). The linear factor obtained from simulation (8.1) is also close to the theoretically calculated value (7.8). This confirms both our linearity assumption and the theoretical expression of the linear factor for our proposed pre-signals. The linear factor of the simulation results is slightly higher than the theoretical value, likely because in reality more capacity is lost due to driver behaviours and physical vehicle constraints. Overall, from Figure 9 we see that installing a pre-signal enables flexible sharing of the capacity between buses and cars. In particular, during periods when buses are not present, the car capacity is doubled compared to implementing a dedicated bus lane. Therefore, pre-signal improves the operation of this intersection.

Figure 9: Car capacity vs. bus flow when a pre-signal is installed upstream of an isolated intersection.

Next, we test the effects of implementing pre-signals on the arterial level. The arterial in Figure 3 is implemented in a micro-simulation model in VISSIM. The length of each link (i.e. elements 1,3,5,7) is assumed to be 180m⁷, and the length of each intersection is assumed to be 15m. Signal timings are as described in Figure 3 with a green wave coordination⁸. We compare the following three strategies.

1. Dedicated bus lane: There is a dedicated bus lane throughout the whole arterial (as the original case depicted in Figure 3).
2. Pre-signals upstream of critical intersections: Pre-signals are implemented upstream of elements 2 and 8 (as suggested in Section 3). The pre-signal is installed 80m upstream of elements 2, and 60m upstream of element 8. Cars must change back to the car lanes within 100m downstream of elements 2 and 8.
3. Intermittent bus lane: The bus lane is terminated 80m upstream of element 2 and 60m upstream of element 8, but no pre-signal is installed. Cars must change back to the car lanes within 100m downstream of elements 2 and 8.
4. Pre-signals upstream of all intersections: Pre-signals are implemented upstream of elements 2, 4, 6 and 8. Each with an appropriate distance upstream of the intersection. Cars must change back to the car lanes within 100m downstream of each intersection.

Bus demand is assumed to be 60buses/hour. Car demands between 0.1λ and λ are tested ($\lambda = 2100\text{veh/h}$). In each simulation run, the car and bus outputs are measured for a 10-minute period starting from when the first car leaves element 8. For each demand, the outputs are averaged over 5 simulation runs, each

⁷Due to car storage capacity, the pre-signal is generally not applicable on links shorter than 150m.

⁸The green wave coordination is used because the signals on the arterial are likely coordinated before pre-signals are installed. However, tests show that the results do not depend on the coordination.

with a different random seed. The total person output is first tested with an assumption of bus to car passenger occupancy ratio $X^b/X^c = 50$. Then, the sensitivity of the system to passenger occupancy ratios between 20 and 100 is tested to find the domains of application for pre-signals. The results when pre-signals are only installed upstream of the critical intersections, and when pre-signals are installed upstream of all intersections, are exactly the same. This is as expected from our theoretical framework. Hence only results from the former case are presented. Moreover, if mixed lanes are implemented throughout the entire arterial, the queues will still be only upstream of element 2 and element 8. Therefore, the vehicle outputs will be exactly the same as our intermittent bus lane case (although this could increase bus delay). All the results are summarized in Figure 10.

From Figure 10a we see that implementing pre-signals as a flexible sharing strategy increases the arterial car output compared to a dedicated bus lane. Meanwhile, the bus output is the same as that of a dedicated bus lane, hence the same level of bus priority is provided. For the case with $X^b/X^c = 50$, at low car demands, the person output is the same for all the strategies. At very high car demands (e.g. during rush hours when car demand is greater than 0.8λ), implementing pre-signals as a flexible sharing strategy results in the largest person output for the arterial. When the car demand is between 0.4λ and 0.8λ , it is best to implement intermittent bus lane. However, its person output is very unstable and sensitive to the car demand. A drastic drop in person output is observed when the car demand exceeds 0.6λ . In contrast, the person output is very stable over a wide range of high demands when pre-signals are implemented. Since either the pre-signal strategy or the intermittent bus lane strategy results in the highest arterial person output at higher car demands, their domains of application are calculated for different passenger occupancy ratios and summarized in Figure 10b. We see that the system is not very sensitive to the passenger occupancy ratio, and implementing the pre-signals at the critical locations always improve the operation of the arterial in terms of person output at higher car demands. The upper bound for an intermittent bus lane is almost the same as our previous theoretical results. However, the lower bound is a little lower, and hence pre-signals perform worse than an intermittent bus lane for a slightly larger range of low car demands. This is because the simulated car outputs when pre-signals are installed are a bit lower than those predicted with out theoretical calculations when the traffic is congested. In other words, the capacity sharing achieved by pre-signals is less efficient than theoretically expected. This agrees with our previous result that the β factors calculated from simulation are a bit lower than the theoretical results.

Figure 10: Comparison of three different strategies for the arterial example.

6. Conclusions

This paper develops a capacity analysis framework to analyze bi-modal facilities. In a bi-modal facility, different modes can discharge utilizing separate capacities, or with a total capacity. We quantify the performance of a bi-modal facility with its set of potential throughputs. In particular, our framework enables us to analytically evaluate flexible sharing strategies on bi-modal arterials. These strategies provide priority to buses without allocating road space capacity between modes in a fixed manner. We build a parsimonious model with only a single parameter which defines the reduction in car capacity to describe the efficiency of any flexible sharing strategy. This factor could be estimated analytically, empirically, or with simulations. Based on our framework, graphical tools and mathematical principles to systematically and quantitatively evaluate any given bi-modal arterial are developed. The tools presented in this paper enable practitioners to decide if installing flexible sharing strategies will improve the capacity of the arterial, and if so where to implement them. We state two principles to efficiently install these flexible sharing strategies along a given arterial at some ‘critical bottlenecks’, and provide the requirements for these flexible sharing strategies.

A specific example of flexible sharing strategy, the pre-signal, is illustrated both analytically and through simulations. Previously, pre-signals had only been studied at the isolated intersection level in the literature. Our framework enables a first analysis for its implementation on the arterial level. Results show that arterials with properly installed pre-signals perform better during rush hours, compared to implementing an intermittent bus lane or a dedicated bus lane. The bus priority could be provided without significant damage to car operations. This makes it more feasible for different stakeholders to agree on the implementation of bus priority strategies, which could lead to a more sustainable transportation system in the long run, with more passengers using the bus mode.

It is worth noting that, although the pre-signal is analyzed as a specific example of flexible sharing strategies, the formulation and conclusions are applicable to any general flexible sharing strategy applied to any kind of bottlenecks along the arterial (for example, tunnels, bridges, construction sites, etc.). Moreover, since our framework is applicable to any general queuing system, the conclusions are applicable in other contexts beyond transportation. It could be any serial queuing system with two or more types of objects sharing the capacity. The key characteristic is that one type of objects have higher priority, but come with relatively low volume. An example could be priority versus normal mail in the post system. Our conclusion implies that it might not be optimal to devote an entire unit of work-force exclusively to priority mails, especially in some critical logistic steps. Rather, it could be more beneficial to share the workforce and the processing capacity. Overall, the framework developed in this paper could be readily extended to study other queuing systems of similar nature.

Acknowledgement

This research has been partially funded by the Swiss National Science Foundation (SNSF) under the project name SignBus, contract 205121_150180.

Appendix A. Proof of Theorem 1

Theorem 1: Suppose Ω is the fixed set for an arterial with mixed lanes or flexible sharing strategies. If a certain demand combination $(q_0^b, q_0^c) \in \Omega$, then for any other fixed set Ω' such that $\Omega' \subset \{(q^b, q^c) \mid q^b \leq q_0^b \wedge q^c \leq q_0^c\}$, we have $\Omega' \subset \Omega$.

Proof. To prove $\Omega' \subset \Omega$, we proceed to show that $\{(q^b, q^c) \mid q^b \leq q_0^b \wedge q^c \leq q_0^c\} \subset \Omega$.

Index the elements with dedicated bus lanes with $j, j \in J$. Index the elements with mixed lanes or flexible

sharing strategy with $k, k \in K$. We proceed to show that $(q^b, q^c) \in \Omega_j, \forall j \in J$ and $(q^b, q^c) \in \Omega_k, \forall k \in K$. To prove $(q^b, q^c) \in \Omega_j, \forall j \in J$, notice we have $(q_0^b, q_0^c) \in \Omega$ which means $pq_0^b \leq \min_{j \in J} Q_j^b \wedge q_0^c \leq \min_{j \in J} Q_j^c, \forall j \in J$.

Hence, we have $pq^b \leq pq_0^b \leq \min_{j \in J} Q_j^b \wedge q^c \leq q_0^c \leq \min_{j \in J} Q_j^c, \forall j \in J$.

To prove $(q^b, q^c) \in \Omega_k, \forall k \in K$, denote the linear factors for the flexible sharing strategies as β_k (which equals p in the case of mixed lanes). Notice we have $(q_0^b, q_0^c) \in \Omega$ which means $\beta_k q_0^b + q_0^c \leq Q_k, \forall k \in K$. Hence, we have $\beta_k q^b + q^c \leq \beta_k q_0^b + q_0^c \leq Q_k, \forall k \in K$. $\square$

Appendix B. Proof of Theorem 2

Theorem 2: If Ω_i are the fixed sets of the corresponding transformation f_i of element i , and Ω is the fixed set of the transformation $f = f_{|I|} \dots f_3 f_2 f_1$ of the arterial, then $\Omega = \Omega_1 \cap \Omega_2 \cap \dots \cap \Omega_{|I|}$.

Proof. To prove $\Omega = \Omega_1 \cap \Omega_2 \cap \dots \cap \Omega_{|I|}$, we proceed to show $\Omega_1 \cap \Omega_2 \cap \dots \cap \Omega_{|I|} \subseteq \Omega$ and $\Omega \subseteq \Omega_1 \cap \Omega_2 \cap \dots \cap \Omega_{|I|}$. For the first direction, we proceed to show that $\forall (q^b, q^c) \in \Omega_1 \cap \Omega_2 \cap \dots \cap \Omega_{|I|}$, we have $(q^b, q^c) \in \Omega$.

Note that $(q^b, q^c) \in \Omega_1 \cap \Omega_2 \cap \dots \cap \Omega_{|I|}$ means $(q^b, q^c) \in \Omega_i, \forall i \in I$.

Hence by the definition of fixed sets, we have $f_i(q^b, q^c) = (q^b, q^c), \forall i \in I$. Hence, we can calculate the composite function step by step, $f_{|I|} \dots f_2 f_1(q^b, q^c) = f_{|I|} \dots f_2(q^b, q^c) = \dots = f_{|I|}(q^b, q^c) = (q^b, q^c)$. Hence, (q^b, q^c) is also a fixed point of the composite function, i.e. $(q^b, q^c) \in \Omega$.

For the second direction we would like to show that $\forall (q^b, q^c) \in \Omega$, we have $(q^b, q^c) \in \Omega_1 \cap \Omega_2 \cap \dots \cap \Omega_{|I|}$. This is not easy, so we proceed to prove its negation is false. Note that its negation is $\exists (q^b, q^c) \in \Omega$, s.t. $(q^b, q^c) \notin \Omega_1 \cap \Omega_2 \cap \dots \cap \Omega_{|I|}$. This means $\exists i \in I$, s.t. $(q^b, q^c) \notin \Omega_i$.

By the definition of fixed sets, this means $\exists i \in I$, s.t. $f_i(q^b, q^c) \neq (q^b, q^c)$. Hence, either $f_i^b(q^b, q^c) < q^b$ or $f_i^c(q^b, q^c) < q^c$. Without loss of generality, assume $f_i^b(q^b, q^c) < q^b$. Then $f_{|I|}^b \dots f_2^b f_1^b(q^b, q^c) \leq f_{|I|}^b \dots f_i^b(q^b, q^c) < f_{|I|}^b \dots f_{i+1}^b(q^b, q^c) \leq q^b$. This is a contradiction because $(q^b, q^c) \in \Omega$. $\square$

Appendix C. Proof of Condition 1

Condition 1: For a Type B arterial, the capacity of the arterial could be improved if the following condition is NOT satisfied:

$$\exists j^* \in J, \text{ such that } Q_{j^*}^b = \min_{j \in J} \{Q_j^b\} \wedge Q_{j^*}^c = \min_{j \in J} \{Q_j^c\} \quad (\text{C.1})$$

Proof. If this condition is not satisfied, denote $J^b = \operatorname{argmin}_{j \in J} \{Q_j^b\}$ the set of all bus bottlenecks, and $J^c = \operatorname{argmin}_{j \in J} \{Q_j^c\}$ the set of all car bottlenecks. Among these, we pick one bus bottleneck element $j_1 \in J^b$ and one car bottleneck element $j_2 \in J^c$. Since no element is a bottleneck for both buses and cars, we have $j_1 \neq j_2$, and then $Q_{j_2}^c < Q_{j_1}^c$ and $Q_{j_1}^b < Q_{j_2}^b$. The fixed set of the arterial is $\Omega = \{(q^b, q^c) | pq^b \leq Q_{j_1}^b \wedge q^c \leq Q_{j_2}^c\}$. To improve the capacity of the arterial, replace all elements $j \in J^c$ with flexible sharing strategies with β_j which satisfy $\beta_j \leq \frac{Q_j - Q_{j_2}^c}{Q_{j_1}^b}$. This requirement will ensure that all the newly implemented flexible sharing strategies could achieve the capacity share of $q^b = Q_{j_1}^b, q^c = \min_{j \in J^c} \{(Q_j - \beta_j Q_{j_1}^b)\}$ where $q^c \geq Q_{j_2}^c$. Hence this capacity share could be achieved by the improved arterial. Denote the fixed set of the improved arterial

as Ω' .

$\forall (q^b, q^c) \in \Omega$, we have $pq^b \leq Q_{j_1}^b$ and $q^c \leq Q_{j_2}^c \leq \min_{j \in J^c} (Q_j - \beta_j Q_{j_1}^b)$.

Then, $(q^b, q^c) \in \left\{ (q^b, q^c) | pq^b \leq Q_{j_1}^b \wedge q^c \leq \min_{j \in J^c} (Q_j - \beta_j Q_{j_1}^b) \right\}$, and hence $\Omega \subset \left\{ (q^b, q^c) | pq^b \leq Q_{j_1}^b \wedge q^c \leq \min_{j \in J^c} (Q_j - \beta_j Q_{j_1}^b) \right\}$.

From Theorem 1 we have $\Omega \subset \Omega'$. Hence, the capacity is improved. $\square$

Appendix D. Proof of Condition 2

Condition 2: For a Type B arterial which satisfies condition 1 or a Type C arterial, the capacity of the arterial could be improved if the following condition is satisfied.

$$\min_{j \in J} \{Q_j^c\} < \min\{\min_{k \in K} \{Q_k - \beta_k D^b\}, D^c\} \quad (\text{D.1})$$

Proof. To prove this, we will show that when this condition is satisfied, we could implement flexible sharing strategies which result in a demand-based improvement. Denote the potential throughput of the arterial as Ω . First, we will show that $\Omega \cap D = \{(q^b, q^c) | (q^b \leq D^b, q^c \leq \min_{j \in J} \{Q_j^c\})\}$. To prove this we will show

$\Omega \cap D \subseteq \{(q^b, q^c) | (q^b \leq D^b, q^c \leq \min_{j \in J} \{Q_j^c\})\}$ and $\{(q^b, q^c) | (q^b \leq D^b, q^c \leq \min_{j \in J} \{Q_j^c\})\} \subseteq \Omega \cap D$.

For the first direction, if $(q^b, q^c) \in \Omega \cap D$, from Theorem 2 we have $(q^b, q^c) \in \Omega_i, \forall i \in I$. In particular $(q^b, q^c) \in \Omega_j, \forall j \in J$, hence $q^c \leq \min_{j \in J} \{Q_j^c\}$. Also $(q^b, q^c) \in D$, hence $q^b \leq D^b$. Combining the above two restrictions on q^b and q^c , we have $(q^b, q^c) \in \{(q^b, q^c) | (q^b \leq D^b, q^c \leq \min_{j \in J} \{Q_j^c\})\}$.

For the second direction, if $(q^b, q^c) \in \{(q^b, q^c) | (q^b \leq D^b, q^c \leq \min_{j \in J} \{Q_j^c\})\}$, then we have $q^c \leq \min_{j \in J} \{Q_j^c\} < \min\{\min_{k \in K} \{Q_k - \beta_k D^b\}, D^c\} \leq D^c$. Also, $q^b \leq D^b$, hence $(q^b, q^c) \in D$. From the assumption of low bus demand D^b , we have $q^b \leq D^b \leq \min_{j \in J} \{Q_j^b\}$. Also, $q^c \leq \min_{j \in J} \{Q_j^c\}$, hence $(q^b, q^c) \in \Omega_j, \forall j \in J$. We also have $q^c \leq \min_{j \in J} \{Q_j^c\} < \min\{\min_{k \in K} \{Q_k - \beta_k D^b\}, D^c\} \leq \min_{k \in K} \{Q_k - \beta_k D^b\}$. Hence $(q^b, q^c) \in \Omega_k, \forall k \in K$. Then $(q^b, q^c) \in \Omega_i, \forall i \in I$. Then $(q^b, q^c) \in \Omega \cap D$. This completes the first part of the proof.

Second, we improve the capacity of the arterial by replacing all elements $j \in J^c$, $J^c = \text{argmin}\{Q_j^c\}$, with flexible sharing strategies with β_j which satisfy $\beta_j \leq \frac{Q_j - \min_{j \in J} \{Q_j^c\}}{D^b}$. This requirement will ensure that all the newly implemented flexible sharing strategies could achieve the capacity share of $q^b = D^b, q^c = \min_{j \in J^c} \{Q_j - \beta_j D^b\} > \min_{j \in J} \{Q_j^c\}$. Hence this capacity share could be achieved by the improved arterial. Denote the new fixed set of the improved arterial as Ω' .

Third, we show that that $\Omega \cap D \subset \{(q^b, q^c) | (q^b \leq D^b, q^c \leq \min_{j \in J^c} \{Q_j - \beta_j D^b\})\}$. If $(q^b, q^c) \in \Omega \cap D$, then $(q^b, q^c) \in D$, and we have $q^b \leq D^b$. Also, $q^c \leq \min_{j \in J} \{Q_j^c\} < \min_{j \in J^c} \{Q_j - \beta_j D^b\}$. Hence $(q^b, q^c) \in \{(q^b, q^c) | (q^b \leq D^b, q^c \leq \min_{j \in J^c} \{Q_j - \beta_j D^b\})\}$.

Therefore, from Theorem 1 we have $\Omega \cap D \subset \Omega' \cap D$. $\square$

References

Agrawal, A. W., Goldman, T., Hannaford, N., 2013. Shared-use bus priority lanes on city streets: approaches to access and enforcement. *Journal of Public Transportation* 16 (4), 2.

- Allsop, R. E., 1972. Estimating the traffic capacity of a signalized road junction. *Transportation Research* 6 (3), 245–255.
- Arnet, K., Guler, S. I., Menendez, M., nov 2015. Effects of Multimodal Operations on Urban Roadways. *Transportation Research Record: Journal of the Transportation Research Board* 2533, 1–7.
- Baker, R. J., Collura, J., Dale, J. J., Head, L., Hemily, B., Ivanovic, M., Jarzab, J. T., McCormick, D., Obenberger, J., SMITH, L., Stoppenhagen, G. R., 2002. An Overview of Transit Signal Priority.
- Boyac, B., Geroliminis, N., 2011. Estimation of the Network Capacity for Multimodal Urban Systems. *Procedia - Social and Behavioral Sciences* 16, 803–813.
- Chiabaut, N., 2015. Evaluation of a multimodal urban arterial: The passenger macroscopic fundamental diagram. *Transportation Research Part B: Methodological* 81, 410–420.
- Chiabaut, N., Xie, X., Leclercq, L., 2013. Road Capacity and Travel Times with Bus Lanes and Intermittent Priority Activation. *Journal of the Transportation Research Board* 2315 (1), 182–190.
- Chiabaut, N., Xie, X., Leclercq, L., 2014. Performance analysis for different designs of a multimodal urban arterial. <http://dx.doi.org/10.1080/21680566.2014.939245>.
- Christofa, E., Ampountolas, K., Skabardonis, A., 2016. Arterial traffic signal optimization: A person-based approach. *Transportation Research Part C: Emerging Technologies* 66, 27–47.
- Daganzo, C. F., 2007. Urban gridlock: Macroscopic modeling and mitigation approaches. *Transportation Research Part B: Methodological* 41 (1), 49–62.
- Daganzo, C. F., Geroliminis, N., 2008. An analytical approximation for the macroscopic fundamental diagram of urban traffic. *Transportation Research Part B: Methodological* 42 (9), 771–781.
- Eichler, M., Daganzo, C. F., 2006. Bus lanes with intermittent priority: Strategy formulae and an evaluation. *Transportation Research Part B: Methodological* 40 (9), 731–744.
- Geroliminis, N., Daganzo, C. F., 2008. Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B: Methodological* 42 (9), 759–770.
- Geroliminis, N., Sun, J., 2011. Properties of a well-defined macroscopic fundamental diagram for urban traffic. *Transportation Research Part B: Methodological* 45 (3), 605–617.
- Geroliminis, N., Zheng, N., Ampountolas, K., 2014. A three-dimensional macroscopic fundamental diagram for mixed bi-modal urban networks. *Transportation Research Part C: Emerging Technologies* 42, 168–181.
- Greenshields, B., 1935. A study of traffic capacity. *Highway Research Board Proceedings* 14, 448–477.
- Guler, S. I., Cassidy, M. J., 2012. Strategies for sharing bottleneck capacity among buses and cars. *Transportation Research Part B: Methodological* 46 (10), 1334–1345.
- Guler, S. I., Gayah, V. V., Menendez, M., 2016. Bus priority at signalized intersections with single-lane approaches: A novel pre-signal strategy. *Transportation Research Part C: Emerging Technologies* 63, 51–70.
- Guler, S. I., Menendez, M., 2014a. Analytical formulation and empirical evaluation of pre-signals for bus priority. *Transportation Research Part B: Methodological* 64, 41–53.
- Guler, S. I., Menendez, M., 2014b. Evaluation of Presignals at Oversaturated Signalized Intersections. <http://dx.doi.org/10.3141/2418-02>.
- He, H., Guler, S. I., Menendez, M., 2014. The effects of pre-signals at an isolated intersection: simulation results. In: *Swiss Transport Research Conference (STRC)*. Ascona, Switzerland.

- He, H., Guler, S. I., Menendez, M., 2016. Adaptive control algorithm to provide bus priority with a pre-signal. *Transportation Research Part C: Emerging Technologies* 64, 28–44.
- He, H., Menendez, M., Guler, I., 2017. Analytical evaluation of flexible sharing strategies on multi-modal arterials. *Transportation Research Procedia* 23, 980–999.
- Lighthill, M. J., Whitham, G. B., may 1955. On Kinematic Waves. II. A Theory of Traffic Flow on Long Crowded Roads. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 229 (1178), 317–345.
- Liu, Y., Chang, G.-L., 2011. An arterial signal optimization model for intersections experiencing queue spillback and lane blockage. *Transportation Research Part C: Emerging Technologies* 19 (1), 130–144.
- Loder, A., Ambühl, L., Menendez, M., Axhausen, K. W., 2017. Empirics of multi-modal traffic networks—using the 3d macroscopic fundamental diagram. *Transportation Research Part C: Emerging Technologies* 82, 88–101.
- Mahmassani, H. S., Saberi, M., Zockaie, A., 2013. Urban network gridlock: Theory, characteristics, and dynamics. *Transportation Research Part C: Emerging Technologies* 36, 480–497.
- Mahmassani, H. S., Williams, J. C., Herman, R., 1984. Investigation of network-level traffic flow relationships: some simulation results. *Transportation Research Record: Journal of the Transportation Research Board* 971 (971), 121–130.
- Mahmassani, H. S., Williams, J. C., Herman, R., 1987. Performance of urban traffic networks. *Proceedings of the 10th International Symposium on Transportation and Traffic Theory*, 1–20.
- Newell, G., 1999. Delays caused by a queue at a freeway exit ramp. *Transportation Research Part B: Methodological* 33 (5), 337–350.
- Nowlin, L., Fitzpatrick, K., 1997. Performance of Queue Jumper Lanes. *American Society of Civil Engineers*.
- Smith, H. R., Hemily, B., Ivanovic, M., 2005. *Transit Signal Priority (TSP): A Planning and Implementation Handbook*.
- Transportation Research Board, 1997. *Special Report 209: Highway Capacity Manual*, 3rd Ed.
- Truong, L. T., Sarvi, M., Currie, G., 2015. Exploring Multiplier Effects Generated by Bus Lane Combinations. *Transportation Research Record: Journal of the Transportation Research Board*.
- Wu, J., Hounsell, N., 1998. Bus priority using pre-signals. *Transportation Research Part A: Policy and Practice* 32 (8), 563–583.
- Wu, X., Liu, H. X., Geroliminis, N., 2011. An empirical analysis on the arterial fundamental diagram. *Transportation Research Part B: Methodological* 45 (1), 255–266.
- Zheng, N., Geroliminis, N., 2013. On the distribution of urban road space for multimodal congested networks. *Transportation Research Part B: Methodological* 57, 326–341.