Technical report 05-006

Optimal control of freeway networks with bottlenecks and static demand[*]

A. Hegyi, B. De Schutter, and J. Hellendoorn

Delft Center for Systems and Control
Delft University of Technology
Mekelweg 2, 2628 CD Delft
The Netherlands
phone: +31-15-278.24.73 (secretary)
URL: https://www.dcsc.tudelft.nl

Optimal control of freeway networks with bottlenecks and static demand

A. Hegyi, B. De Schutter, and J. Hellendoorn

Delft Center for Systems and Control
Delft University of Technology
Mekelweg 2, 2628 CD Delft, The Netherlands
email: {a.hegyi,b.deschutter,j.hellendoorn}@dcsc.tudelft.nl

Abstract

We consider optimally coordinated freeway traffic control for networks containing bottlenecks with capacity drop and hysteresis behavior. Due to the multitude of traffic jams, and the spatial and temporal relationships between control actions and traffic behavior, this problem is not as straightforward as for local control. The order in which the measures are applied may be relevant, or it may be possible that not all jams can be resolved. In that case the best possible locations of jams should be determined. We develop an approach that addresses these problems, where we use a generalized representation of flow-limiting control measures and bottlenecks. We determine whether a certain set of control measures is sufficient to improve the network performance. The approach also supplies the necessary sequence of control actions and the necessary relocation of traffic jams to achieve the network state corresponding to the best achievable performance.

1 Introduction

1.1 Traffic control in networks

In the past dynamic traffic control measures (such as ramp metering, route guidance, and dynamic speed limits) have been applied mainly locally to resolve traffic jams. However, due to the high traffic demands the spatial and temporal relationships in traffic networks have become stronger: a control measure applied at a certain place and time may also influence (positively or negatively) the traffic later and/or at more distant parts of the network. Therefore, it is necessary to take the network-wide effects of the control measures into account, which implies coordination between the measures. By network-wide coordinated control (as opposed to local control) we mean the coordination of the control measures such that the resulting traffic behavior is taken into account for the whole route of all drivers (from their origins to their destinations), not for a limited part of their route only.

Even though a network-wide coordinated approach may result in more effective traffic control, it is not guaranteed at all that it can improve the network performance for all possible congested traffic scenarios. While for a local control measure it is straightforward to determine whether it is effective (e.g., by flow measurements), for network-wide control the effects may not be so straightforward, e.g., an improved traffic flow at one location may worsen a traffic jam at another location, or even trigger a new jam.

Another difficulty in network-wide traffic control is that it may not be possible to solve all traffic jams in the network or it may be necessary to create temporarily a traffic jam somewhere else in the network in order to solve a given jam (similarly to ramp metering which locally creates a queue to solve a freeway jam). With the approach presented in this paper an answer is given to which jams need to be solved (or created) in order to achieve the highest outflow of the network.

Furthermore, there may be limitations on the traffic control measure signals, such as minimum and maximum metering rates, bounds on dynamic speed limits, or the limited rerouting effects of route guidance messages, so that these measures may not limit the flow sufficiently to solve a jam individually. However, the combination with other measures (available or to be installed) may result in a more effective control that can solve the jam. The selection of the appropriate measures will also be addressed in this paper.

1.2 Capacity drop

There are several possible causes why freeway networks do not always perform optimally. By non-optimal we mean either that while there is sufficient demand at some freeway links, they do not carry flows equal to their capacity, or that the drivers take certain routes while there is another, shorter or faster route that has enough capacity to accommodate them. The main reasons for freeway links carrying a sub-capacity flow are the capacity drop and blocking (or insufficient demand). The capacity drop (also called the two-capacity phenomenon) is the phenomenon that the outflow of a jam at a bottleneck (the so-called queue discharge rate) is lower than the capacity of the bottleneck in free-flow (the so-called pre-queue capacity). As long as the jam remains existent, the performance of a jammed freeway link will be sub-optimal. Blocking occurs when the tail of a traffic jam propagates back to a bifurcation where it blocks the traffic that has a route that does not go via the bottleneck location that caused the jam.

In this paper we focus on the capacity drop and develop an approach to determine whether the network performance can by improved by dynamic traffic control measures. For the sake of simplicity we will assume that the queues that occur at bottlenecks and control measures do not become so long that they reach bifurcations or other bottlenecks. If this assumption does not hold, a more complex model should be used.

The capacity drop may occur at several types of bottlenecks, such as on-ramps, upstream propagating jams (shock waves), off-ramps, curves, grades, tunnels and bridges. The value of the capacity drop has been estimated for some of these bottlenecks: for on-ramps [1–3] it was found to be in the range of 0–15%, and for upstream propagating jams [1, 4] around 30%. There is no consensus about the reason of the capacity drop; it may be related to traffic friction upstream the bottleneck, or to the lower acceleration of vehicles in high-density areas (the so-called slow-to-start behavior [5]).

A phenomenon related to the capacity drop is the hysteresis effect, which in this context means that in order to achieve a transition from congested to free flow, the inflow of an active bottleneck needs to be considerably lower than it can be in free flow without causing a breakdown. Hysteresis is observed at on-ramps [6], but may also occur at other types of bottlenecks. Another related concept is *metastability*, which means that under the same traffic demands both the congested and the free flow scenarios can remain existent for a long time. E.g., there may be demands possible on a main-stream freeway and on-ramp that result in a stable free-flow or in a stable jam (at the on-ramp), depending on the initial condition.

Related to this approach is the approach presented by Papageorgiou [7] where the modeled traffic network also results in linear equations and linear inequality constraints. The difference between the two approaches is that in this paper the capacity drop is explicitly modeled, and the resulting optimal control signal will take into account the hysteresis behavior of the bottlenecks. Other approaches, such as [8] do take into account the capacity drop via nonlinear traffic flow modeling, however the approach presented here is computationally more efficient, is guaranteed to be optimal,

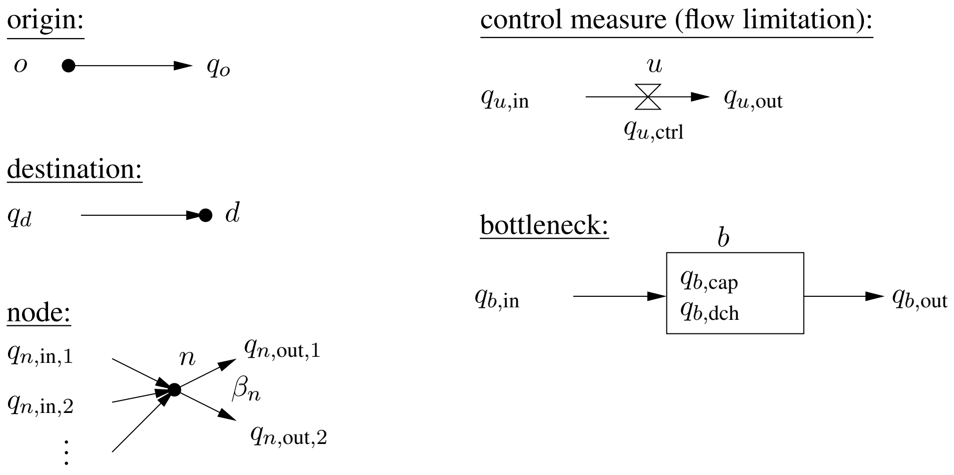


Figure 1: The network elements.

and provides insight in the reasons why certain traffic jams can or cannot be solved. This insight may be relevant to traffic operators in convincing them to apply certain control strategies, or to traffic infrastructure planners in the decisions about the placement of certain traffic control measures.

1.3 Optimality: outflow and Total Time Spent

The network optimality mentioned above is often defined mathematically as the total time that vehicles spend in the network (TTS), (see, e.g., [8, 9]). It is not difficult to show that if the traffic demand is given for a certain traffic network then the TTS is directly related to the outflow of the network. We derive a formula to compute the TTS using the inflow and the outflow of the network (see Section 3.1).

The core of the problem discussed in this paper is that if a bottleneck is active (jammed) then the inflow of the bottleneck has to be limited to a value below its outflow in order to solve the jam and return to free flow, where higher flows can be achieved than the queue discharge rate. Obviously, there must be sufficient flow-limiting control measures upstream to limit the inflow of the bottleneck. Furthermore, the applied control measure should not result in a too large flow reduction elsewhere in the network since that would adversely affect the overall flow improvement.

2 Problem description

2.1 Network elements

We model a traffic network by a directed graph that contains problem-specific elements. Each network consists of the following elements (see also Figure 1 and Table 1 for the symbols of the elements and the related variables):

- **Origins.** Origin o is an element of the set of all origins $\{\text{O}_1, \text{O}_2, \dots\}$, and provides a constant inflow to the network of q_o (veh/h).
- **Destinations.** Destinations are the sinks of traffic. The flow at destination $d \in \{\text{D}_1, \text{D}_2 \dots\}$ is denoted by q_d (veh/h).

variable	description
q_o	inflow from origin o
q_d	outflow at origin d
$q_{n,\text{in},i}$	inflow from link i to node n
$q_{n,\text{out},j}$	outflow to link j from node n
q_n	total flow through node n
$\beta_{n,j}$	fraction of traffic that leaves node n through link j (uncontrolled node)
$q_{n,\text{ctrl}}$	traffic leaving node n (controlled node, $q_{n,\text{out},1} = q_{n,\text{ctrl}}$)
$q_{n,\text{out},1,\text{min}}$	minimal outflow from node n to the first outgoing link
$q_{n,\text{out},1,\text{max}}$	maximal outflow from node n to the first outgoing link
$q_{u,\text{in}}$	inflow of measure u
$q_{u,\text{out}}$	outflow of measure u
$q_{u,\text{ctrl}}$	control input at measure u
$q_{u,\text{min,ctrl}}$	lower bound of control input at measure u
$q_{u,\text{max,ctrl}}$	upper bound of control input at measure u
$q_{b,\text{in}}$	inflow at bottleneck b
$q_{b,\text{out}}$	outflow at bottleneck b
$q_{b,\text{dch}}$	queue discharge rate of bottleneck b (bottleneck active)
$q_{b,\text{cap}}$	free-flow capacity of bottleneck b (bottleneck not active)

Table 1: Overview of the network variables and constants. All variables and constants have veh/h as unit, except for $\beta_{n,j}$, which is a dimensionless fraction.

- **Nodes with route guidance.** At nodes traffic from several incoming links may be joined and redistributed over one or two outgoing links[1]. The flows of the incoming links of node $n \in \{\mathrm{N}_1, \mathrm{N}_2, \dots\}$ are denoted by $q_{n,\text{in},i}$ (with $i \in \mathcal{I}_n$, where $\mathcal{I}_n$ denotes the set of indexes of the incoming links of node n), and the outgoing links are denoted by $q_{n,\text{out},j}$ (with $j \in \{1,2\}$). The inflows and outflows are related by:

  $$q_{n,\text{out},j} = \beta_{n,j}\ q_n\ ,$$

  where $q_n = \sum\limits_{i \in \mathcal{I}_n} q_{n,\text{in},i}$, and $\beta_{n,j}$ is the fraction of traffic that leaves node n through link j. Of course, $\beta_n \geq 0$ and $\sum_{j \in \mathcal{O}_n} \beta_{n,j} = 1$, where $\mathcal{O}_n$ is the set of indices of leaving links of node n. If there is no route guidance, a fixed turning rate $\beta_{n,j}$ is assumed. If there is route guidance at the node, we will consider $q_{n,\text{ctrl}} = q_{n,\text{out},1}\ (= \beta_{n,1}\ q_n)$ as the control variable.

  There may be bounds on the route guidance signal $q_{n,\text{out},1}$, which are expressed by $q_{n,\text{out},1,\text{min}}$ and $q_{n,\text{out},1,\text{max}}$ (which also implies corresponding bounds on $q_{n,\text{out},2}$). This leads to the following relations:

  $$q_{n,\text{out},2} = q_n - q_{n,\text{out},1},$$
  $$q_{n,\text{out},1,\text{min}} \leq q_{n,\text{out},1} \leq q_{n,\text{out},1,\text{max}}.$$

- **Flow-limiting control measures.** Traffic control measures, such as ramp metering, mainstream metering, and dynamic speed limits can be represented by a generalized control measure

[1] For simplicity we assume that there are at most two outgoing links. The extension to more outgoing links is straightforward.

$u \in \{U_1, U_2, \dots\}$ that describes the corresponding flow limitation[2].

A flow limitation can be active or inactive. The relation between the inflow, outflow, and the control input of a flow limitation is represented in Figure 2(a) and will be detailed here. If the flow limitation is active, it limits the flow and the following relations hold:

$$\begin{aligned} q_{u,\text{out}} &= q_{u,\text{ctrl}} \,, \\ q_{u,\text{ctrl}} &\leq q_{u,\text{in}} \,, \end{aligned}$$

where $q_{u,\text{ctrl}}$ is the control input of the measure, $q_{u,\text{in}}$ is the inflow and $q_{u,\text{out}}$ the outflow at the measure u. In addition there may be bounds on the control input

$$q_{u,\text{min,ctrl}} \leq q_{u,\text{ctrl}} \leq q_{u,\text{max,ctrl}} \,.$$

Note that by using the upper bounds we can represent traffic control measures that have a maximal throughput that is lower than the capacity of the road when they are switched on.

If the flow limitation is inactive, then the (out)flow is not limited by the control measure, and the following relations hold:

$$\begin{aligned} q_{u,\text{out}} &= q_{u,\text{in}} \,, \\ q_{u,\text{ctrl}} &\geq q_{u,\text{in}} \,. \end{aligned}$$

The activity status of u is denoted by χ_u, which has a value 1 if the control measure is active and 0 if it is inactive.

- **Bottlenecks.** A generalized bottleneck $b \in \{\mathrm{B}_1, \mathrm{B}_2, \dots\}$ may represent several kinds of bottlenecks, such as on-ramps, bridges, tunnels, curves, grades, shock waves[3], merges, and bifurcations. The common factor in these bottlenecks is that they have a limited capacity $q_{b,\text{cap}}$, and that there may be a capacity drop if the bottleneck is jammed. The queue discharge rate is denoted by $q_{b,\text{dch}} (\leq q_{b,\text{cap}})$, where equality holds if there is no capacity drop, but only a limited capacity.

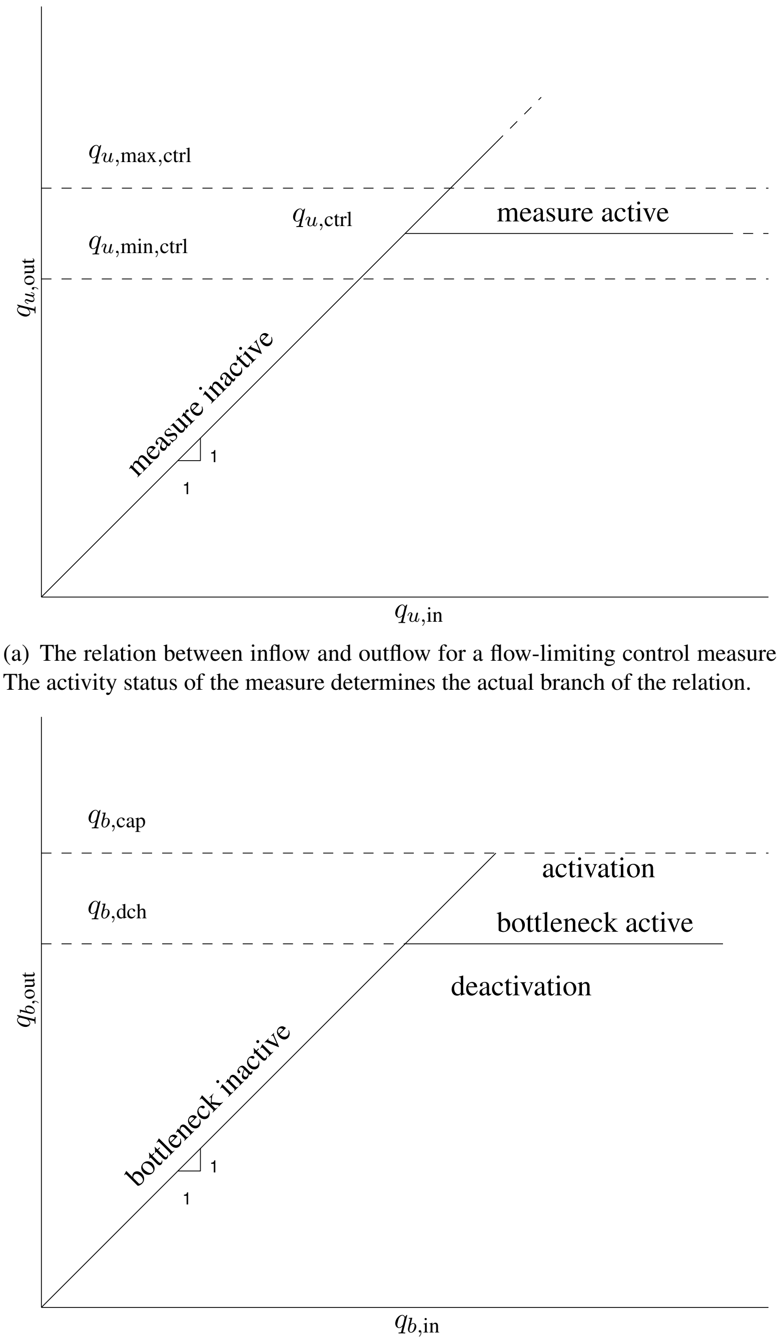
  Similarly to flow-limiting control measures a bottleneck can also be active or inactive, and the relation between the inflow and outflow depends on the activity status. This relation is depicted in Figure 2(b) and described here. The basic idea for the bottleneck modeling is that if the inflow exceeds the capacity then the bottleneck will become active (congested) and the outflow will drop to the queue discharge rate. In order to resolve the jam at the bottleneck the inflow must be limited to a value lower than the outflow (the queue discharge rate). When the jam is resolved, the bottleneck becomes inactive and the outflow may increase up to the capacity again. This switching between the active and the inactive state is equivalent to the hysteresis discussed in Section 1.2.

  Since we do not consider queuing here, we assume that both transitions occur immediately if the conditions are satisfied[4].

[2] Note that for variable speed limits, the speed limit value must be lower than the critical speed in order to limit the flow. See [10] for more information.

[3] The representation of moving shock waves is valid in our framework as long as the shock wave does not propagate upstream to other network elements.

[4] In a dynamic setting the transition from active to inactive would only occur when the queue is resolved, which takes some time depending on the net inflow and the queue length.



(a) The relation between inflow and outflow for a flow-limiting control measure. The activity status of the measure determines the actual branch of the relation.

(b) The relation between inflow and outflow for a bottleneck. The activity status of the bottleneck determines the actual branch of the relation.

Figure 2: The relations between inflow and outflow for flow-limiting control measures and bottlenecks.

If the bottleneck is active the following relationships hold:

$$q_{b,\text{out}} = q_{b,\text{dch}} \ ,$$
$$q_{b,\text{in}} \geq q_{b,\text{dch}} \ .$$

If the bottleneck is inactive the following relationships hold:

$$q_{b,\text{out}} = q_{b,\text{in}} \ ,$$
$$q_{b,\text{in}} \leq q_{b,\text{cap}} \ .$$

The activity status of b is denoted by χ_b, which has a value 1 if the bottleneck is active and 0 if it is inactive.

- **Links.** Links provide the connection between two any other elements. A link connects the outflow of the upstream element with the inflow of the downstream element. The capacity of a link is assumed to be unlimited. So, if a freeway link with limited capacity is modeled, a bottleneck element should be included.

Now we can build networks with the elements from Section 2.1 (see Figure 3 for an example).

2.2 Network properties

In this section we give some definitions that are necessary for the problem statement which will be given in Section 2.3.

- **Network activity state.** We define the network activity state $\mathcal{S}$ as the vector of the activity status of all bottlenecks and control measures

  $$\mathcal{S} = [\chi_{\text{B}_1}, \chi_{\text{B}_2}, \ldots, \chi_{\text{U}_1}, \chi_{\text{U}_2}, \ldots]^{\text{T}} \ .$$

  When the appropriate control actions are determined, the current network activity state is acquired from speed, flow and density measurements at the bottlenecks and control measures. The bottleneck activity states cannot be controlled directly, only through the available control measures, which may change the inflow of a bottleneck such that an activity state change is triggered.

- **Feasibility.** We call a network state $\mathcal{S}$ *feasible* for given demands q_o ($o \in \mathcal{O}$, where $\mathcal{O}$ is the set of all origins) if there exists a set of flows (inflows and outflows) for all nodes, control measures, bottlenecks, and destinations and a set of control inputs $q_{u,\text{ctrl}}$ ($u \in \mathcal{U}$, where $\mathcal{U}$ is the set of all flow-limiting control measures) and $q_{n,\text{ctrl}}$ ($n \in \mathcal{N}$, where $\mathcal{N}$ is the set of all controlled nodes) that satisfy the relations given in Section 2.1.

  It is not difficult to show that for a given set of demands there exists always at least one feasible activity state. The proof is given by assuming any control input, and computing the flows from the origins to the destinations and assuming activity states that satisfy the inequalities corresponding to the inflow of the bottleneck or control measure. Since for any inflow value of a bottleneck or control measure there exists at least one valid activity state, there exist a network activity state that is feasible.

  Furthermore, since for certain inflows two (both active and inactive) activity states may satisfy the related equalities and inequalities, for a certain set of demands and control inputs more than one activity state may be feasible. E.g., a bottleneck may be both active or inactive if the inflow is between the capacity and the queue discharge rate. By this property metastability is be represented in networks.

- **Reachability.** If the network is in a feasible activity state, then other activity states may be reached by varying the control inputs. If the control inputs are changed, then the activity status of a control measure or a bottleneck may change. We say that activity state $\mathcal{S}_2$ is *reachable* from activity state $\mathcal{S}_1$ if one or more inequality conditions related to the activity state $\mathcal{S}_1$ can be violated by changing the control inputs such that the conditions related to state $\mathcal{S}_2$ and the control input bounds are satisfied, under the same demands. Furthermore, if state $\mathcal{S}_3$ is reachable from activity state $\mathcal{S}_2$ and state $\mathcal{S}_2$ from state $\mathcal{S}_1$ then state $\mathcal{S}_3$ is also said to be reachable from state $\mathcal{S}_1$, i.e., the reachability property is transitive. If state $\mathcal{S}_2$ is reachable from state $\mathcal{S}_1$ without intermediate activity states, we say that $\mathcal{S}_2$ is *directly reachable* from $\mathcal{S}_1$.

  Note that it is necessary to require that at least one inequality condition is *violated* since only in that case the transition will be triggered. It may be possible that both states $\mathcal{S}_1$ and $\mathcal{S}_2$ are feasible, but the control measures are not powerful enough to trigger a transition. E.g., a traffic jam at a bridge that has capacity drop: when the demand is between the queue discharge rate and the capacity both activity states (jammed and free flow) are feasible. If there is no control measure upstream the bridge, the transition obviously cannot be made and none of the activity states is reachable from the other one.

  Furthermore, reachability is not a symmetric relation, since due to the hysteresis it may be possible that the available control measures can trigger a bottleneck to become active, but that they are not powerful enough to resolve the jam at the bottleneck. It may also be possible that the control measures can solve a jam, but due to the low demands the jam cannot be recreated.

The outflow of a network is uniquely defined for any control input and corresponding network activity state . The purpose of the control measures is to maximize the network outflow. This can be achieved by changing the control inputs which may lead to another network activity state, or may lead to different outflow within the actual activity state. The goal of the control is to reach the activity state and find the control inputs within that activity state that leads to the highest possible total outflow of the network.

2.3 Problem statement

The control problem can now be formulated as follows:

Given constant traffic demands $q_o, o \in \mathcal{O}$ at the origins, the control inputs $q_{u,\text{ctrl}}, u \in \mathcal{U}$ and $q_{n,\text{ctrl}}, n \in \mathcal{N}$, and given a corresponding feasible network state $\mathcal{S}$

- find the state $\mathcal{S}^*$ that is reachable from state $\mathcal{S}$ and control inputs $q^*_{u,\text{ctrl}}$ and $q^*_{n,\text{ctrl}}$ that maximize the network outflow $\sum\limits_{d\in\mathcal{D}} q_d$, and therefore minimize the TTS,
- find all the possible activity state sequences and corresponding control inputs that lead from state $\mathcal{S}$ and control inputs $q_{u,\text{ctrl}}$ and $q_{n,\text{ctrl}}$ to state $\mathcal{S}^*$ and $q^*_{u,\text{ctrl}}$ and $q^*_{n,\text{ctrl}}$.

3 Approach

In this section we present an approach to solve the stated problem. First, we show that in our case maximizing the outflow is equivalent to minimizing the TTS. Next, we discuss the procedure for finding the control inputs that maximize the outflow for a given activity state. Furthermore, we discuss the algorithms for finding all reachable activity states, and for finding all possible activity state sequences that lead to the optimal state. We illustrate the approach with an example.

3.1 Relation between inflow, outflow and TTS

It is well-known that there is a direct relationship between the TTS and the inflow and outflow of a traffic network (see, e.g., [11]). Here we give the relationship for static inflows and outflows for the case the inflow of the network exceeds the outflow. Denote the number of vehicles in a network by the discrete time variable $n(k)$, where k is the time index. The total time t_{TTS} that all vehicles spend in the network over a period $0, \ldots, K-1$ is given by

$$t_{\text{TTS}} = T \sum_{k=0}^{K-1} n(k) \,, \tag{1}$$

where T is the sampling time. If the total inflow of the network is denoted by q_{in} and the total outflow by q_{out} then the evolution of the number of vehicles in the network over time is given by

$$n(k) = n(k-1) + T(q_{\text{in}} - q_{\text{out}}) \,,$$

or in a more useful form,

$$n(k) = n(0) + Tk(q_{\text{in}} - q_{\text{out}}) \,. \tag{2}$$

For these equations we assume that the network does not become completely empty, i.e., $n(k) > 0$. Combining equation (1) with equation (2) gives

$$t_{\text{TTS}} = TKn(0) + T^2 \frac{K(K-1)}{2} (q_{\text{in}} - q_{\text{out}}) \,.$$

This expression can be used to determine the (relative) improvement of the TTS when the outflow of a network is increased by means of control measures.

From this relationship it is clear that if the outflow is increased then the TTS will decrease given the same initial condition and traffic demand. So, there is a one to one relation between the outflow and the TTS. In the rest of the paper we will focus on outflow maximization instead of the equivalent TTS minimization, since this makes the solution of the control problem easier.

3.2 Solving for one activity state

In order to find the optimal control inputs for a given activity state $\mathcal{S}$, we first derive the equations for the total outflow and for the constraints that contain only the control variables and constants. We distinguish between *constants*, *control variables*, and *destination flows*. Constants are the independent quantities in the network that serve as a flow source for the connected downstream variables, such as the origin flows q_o, and the outflows of the *active* bottlenecks $q_{b,\text{dch}}$. The control variables are the control values $q_{u,\text{ctrl}}$ of flow limitations and the outflow $q_{n,\text{ctrl}}$ of nodes with route guidance. Destination flows are the flows of all destinations.

We denote the constants by the vector $q_{\text{const}} = [q_o, q_{b,\text{dch}}]^{\text{T}}$, which includes the origin flows ($o \in \mathcal{O}$) and the outflows of active bottlenecks ($b \in \{b|\chi_b = 1\}$), the control variables by the vector $q_{\text{ctrl}} = [q_{u,\text{ctrl}}, q_{n,\text{ctrl}}]^{\text{T}}$ which includes the outflows of flow limitations ($u \in \mathcal{U}$) and the outflow of nodes n for which there is route guidance present ($n \in \mathcal{N}$), the destination variables by the vector $q_{\text{dest}} = [q_d]^{\text{T}}$.

By inspection, i.e., by tracking all possible paths from all origins, all outputs of active bottlenecks, and all outputs of active control measures to all destinations, the destination flows can be written as a linear combination of constants and control variables:

$$C_1 q_{\text{ctrl}} + C_2 q_{\text{const}} = q_{\text{dest}} \,, \tag{3}$$

where C_1 and C_2 are appropriate matrices. Similarly, by tracking all possible paths from all origins, all outputs of active bottlenecks, and all outputs of active control measures to all (active and inactive) bottlenecks and all (active and inactive) control measures, all inequality conditions that are related to the network activity state can be written in terms of control variables and constants. Furthermore, the constraints regarding the control input bounds are already in terms of control inputs. Consequently, all constraints can be written as a linear combination of constants and control variables:

$$A_1 q_{\text{ctrl}} + A_2 q_{\text{const}} \geq q_{\text{bounds}} \,, \tag{4}$$

where A_1 and A_2 are matrices and the vector q_{bounds} contains appropriate values, such as the control input bounds, the bottleneck capacities, and queue discharge rates.

Noting that $C_2 q_{\text{const}}$ and $A_2 q_{\text{const}}$ are constant vectors, and that we are interested in maximizing the total outflow of the network, the problem of finding the control inputs that maximize the outflow for the actual activity state, can be written as

$$\max_{q_{\text{ctrl}}} c q_{\text{ctrl}} \,, \text{ subject to} \tag{5}$$

$$A_{\text{activity}} q_{\text{ctrl}} \geq b_1 \,, \tag{6}$$

$$A_{\text{bounds}} q_{\text{ctrl}} \geq b_2 \,, \tag{7}$$

where vector c contains the sum of the rows of C_1, and where the matrix A_1 is split into the matrix A_{activity} and into the matrix A_{bounds} related to the bounds on the control inputs, i.e., $A_1 = \begin{bmatrix} A_{\text{activity}} \\ A_{\text{bounds}} \end{bmatrix}$, and vector $b = q_{\text{bounds}} - A_2 q_{\text{const}}$, which is also split accordingly into b_1 and b_2. This problem is a standard form of a linear programming (LP) problem, which can be solved by existing standard techniques, such as the simplex method or the interior point method (see [12–14] for more information on solving LP problems). In general these kind of LP problems can have three different type of solutions:

1. The problem may be infeasible. In this case the solution set is empty.
2. The solution may tend to infinity.
3. The solution may be (a set of) real-valued points.

In our case:

1. There will always a feasible solution, since we consider only feasible network activity states.
2. The solution will never tend to infinity since the total outflow cannot exceed the total inflow.
3. There will be real-valued solution that we denote by $q^*_{\mathcal{S},\text{ctrl}}$, where $\mathcal{S}$ is the current activity state. If the solution is a set, we pick an arbitrary element.

3.3 Activity state transitions

Each row in equation (6) represents one of the inequalities presented in Section (2.1) related to the network activity state (the state of the control measures and the state of the bottlenecks), and equation (7) represents the bounds on the control inputs. An activity state transition occurs when the control vector is such that one or more inequalities in equation (6) are violated, but the inequalities of equation (7) are satisfied. In other words, if changing the greater or equal sign ($\geq$) to a less sign ($<$) for one or more rows related to bottleneck states, results in a feasible problem, and if there is a control input such that equality holds for these rows, then the transition can be made.

Activity state transitions related to the activity state change of control measures can always be made (as long as the bounds are satisfied) since any control input will result in a feasible activity state.

By testing the feasibility for all possible LP problems that can be obtained by changing the greater or equal signs, all reachable activity states from the current activity state can be found. However, in practice often a pre-selection of states can be made based on heuristics and insight into the traffic system. This will also be illustrated by the example in Section 3.5.

3.4 Finding all reachable activity states and the corresponding paths

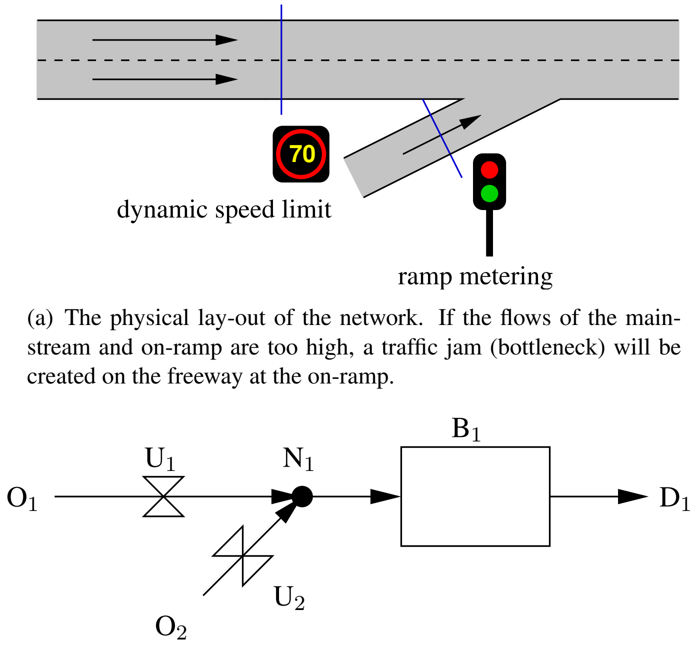
Now that we can find all directly reachable activity states from a given activity state, we are ready to formulate the algorithm to find all reachable activity states with any number of activity state transitions, without visiting the same activity state twice.

With the following algorithm we will find new reachable states by examining the states that are reachable from the currently known reachable states. If a new reachable state is found, the states reachable from this new state are examined, and so on.

Let us denote the initial activity state by $\mathcal{S}_1$ and define $\mathcal{T}_{\text{done}}$ as the set of activity states for which directly reachable activity states already have been determined, and the set $\mathcal{T}_{\text{todo}}$ as the set of states for which directly reachable states not determined yet. Furthermore, we will use the set $\mathcal{T}_{\text{help}}$ as to temporarily store a set of states, and $\mathcal{S}_{\text{help}}$ to temporarily store a state.

```
begin
    1. Initialize $\mathcal{T}_{\text{done}} = \emptyset, \mathcal{T}_{\text{todo}} = \emptyset, \mathcal{T}_{\text{help}} = \emptyset$
    2. Add $\mathcal{S}_1$ to $\mathcal{T}_{\text{todo}}$.
    3. Stop if $\mathcal{T}_{\text{todo}}$ is empty, otherwise set $\mathcal{S}$ equal to any element in $\mathcal{T}_{\text{todo}}$, and compose the
       corresponding LP problem.
    4. Find all states that are directly reachable from $\mathcal{S}$ as described in Section 3.3. Store
       these states in $\mathcal{T}_{\text{help}}$.
    5. For all elements $\mathcal{S}_{\text{help}}$ in $\mathcal{T}_{\text{help}}$ do:
       5.1 If $\mathcal{S}_{\text{help}} \in \mathcal{T}_{\text{done}}$, discard this element, otherwise add it to $\mathcal{T}_{\text{todo}}$.
    6. Go to step 3.
end
```

Following this algorithm $\mathcal{T}_{\text{done}}$ will contain all reachable activity states, and all possible state transitions within these states will be explored. Once we know all reachable activity states we can determine the maximum outflow by solving the LP problems corresponding to these states. By comparison the activity state $\mathcal{S}^*$ and the control inputs $q^*_{u,\text{ctrl}}$ and $q^*_{n,\text{ctrl}}$ that result in the overall maximum outflow can be found, and the first item of the problem statement is solved.



(a) The physical lay-out of the network. If the flows of the mainstream and on-ramp are too high, a traffic jam (bottleneck) will be created on the freeway at the on-ramp.

(b) The network in terms of the defined network elements.

Figure 3: Example network with ramp metering and dynamic speed limits. Dynamic speed limits can complement ramp metering when ramp metering alone is insufficient to resolve the on-ramp jam.

To answer the second item we used the following rule: all paths that lead from $\mathcal{S}_1$ to $\mathcal{S}_2$ can be enumerated by combining all paths from states $\mathcal{S}_1$ to states that are directly reachable from $\mathcal{S}_1$ with all paths that go from the directly reachable states to state $\mathcal{S}_2$. By applying this rule recursively, all paths from $\mathcal{S}_1$ to $\mathcal{S}^*$ can be found.

What rests, is the selection of a state sequence that is favorable to achieve the state transition from $\mathcal{S}_1$ to $\mathcal{S}^*$. The most reasonable choice would be to select the state sequence that leads to the least delay for the traffic. This question will be addressed in future research, since this delay cannot be determined without the explicit representation of queues, because the queue dynamics will determine how long it takes before a congested bottleneck can change state to uncongested.

The main conclusion here is that we have developed an approach to determine and reach the activity state and control inputs that result in the highest network outflow, given the current state of the traffic network, the demands, the available control measures and the network topology.

3.5 Example

In this section we give a simple example to clarify the approach presented in the previous section. The example network is a freeway stretch with a controlled on-ramp. We will show that ramp metering may not be able to resolve a jam occurring on the freeway at the on-ramp, and that the combination of dynamic speed limits and ramp metering may be more effective in these situations.

The network for this example is shown in Figure 3. This example is also related to the publications [4, 15] where we show by numerical optimization that dynamic speed limits can complement ramp metering. Here we show the same by the approach developed in the previous sections.

In the network of Figure 3 O_1 and D_1 are connected by a 2-lane freeway, and the link from O_2

represents a single-lane on-ramp. The bottleneck B_1 represents the bottleneck on the freeway caused by the on-ramp. The on-ramp is metered by U_2 and there is a dynamic speed limit U_1 on the freeway. The parameters of this problem are chosen as:

$q_{\mathrm{O}_1} = 3500\,\mathrm{veh/h}$,
$q_{\mathrm{O}_2} = 600\,\mathrm{veh/h}$,
$q_{\mathrm{U}_1,\mathrm{max,ctrl}} = 4200\,\mathrm{veh/h}$,
$q_{\mathrm{U}_1,\mathrm{min,ctrl}} = 3300\,\mathrm{veh/h}$,
$q_{\mathrm{U}_2,\mathrm{max,ctrl}} = 2000\,\mathrm{veh/h}$,
$q_{\mathrm{U}_2,\mathrm{min,ctrl}} = 300\,\mathrm{veh/h}$,
$q_{\mathrm{B}_1,\mathrm{cap}} = 4200\,\mathrm{veh/h}$,
$q_{\mathrm{B}_1,\mathrm{dch}} = 3800\,\mathrm{veh/h}$.

We assume that in the initial activity state the bottleneck and the ramp metering are active and the speed limit is inactive, i.e., $\mathcal{S}_1 = [\chi_{\mathrm{B}_1}, \chi_{\mathrm{U}_1}, \chi_{\mathrm{U}_2}]^{\mathrm{T}} = [1, 0, 1]^{\mathrm{T}}$.

Solution

The corresponding LP problem written in the form of equations (3) and (4) reads:

$$\max_{q_{\mathrm{ctrl}}} \; q_{\mathrm{D}_1}, \text{ subject to}$$

$$\begin{bmatrix} 0 & 0 \end{bmatrix} \begin{bmatrix} q_{\mathrm{U}_1,\mathrm{ctrl}} \\ q_{\mathrm{U}_2,\mathrm{ctrl}} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} q_{\mathrm{O}_1} \\ q_{\mathrm{O}_2} \\ q_{\mathrm{B}_1,\mathrm{dch}} \end{bmatrix} = \begin{bmatrix} q_{\mathrm{D}_1} \end{bmatrix} \;, \tag{8}$$

$$\begin{bmatrix} 0 & 1 \\ 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} q_{\mathrm{U}_1,\mathrm{ctrl}} \\ q_{\mathrm{U}_2,\mathrm{ctrl}} \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} q_{\mathrm{O}_1} \\ q_{\mathrm{O}_2} \\ q_{\mathrm{B}_1,\mathrm{dch}} \end{bmatrix} \geq \begin{bmatrix} q_{\mathrm{B}_1,\mathrm{dch}} \\ -q_{\mathrm{O}_2} \\ q_{\mathrm{O}_1} \end{bmatrix} \;, \tag{9}$$

$$\begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} q_{\mathrm{U}_1,\mathrm{ctrl}} \\ q_{\mathrm{U}_2,\mathrm{ctrl}} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} q_{\mathrm{O}_1} \\ q_{\mathrm{O}_2} \\ q_{\mathrm{B}_1,\mathrm{dch}} \end{bmatrix} \geq \begin{bmatrix} q_{\mathrm{U}_1,\mathrm{min}} \\ -q_{\mathrm{U}_1,\mathrm{max}} \\ q_{\mathrm{U}_2,\mathrm{min}} \\ -q_{\mathrm{U}_2,\mathrm{max}} \end{bmatrix} \;, \tag{10}$$

where the first row of the system of inequalities (9) is related to the active state of the bottleneck, the second row to the active state of the ramp metering, the third row to the inactive state of the speed limit, and the system of inequalities (10) to the bounds on the control values.

If we assume a control vector of $q_{\mathrm{ctrl}} = [q_{\mathrm{U}_1,\mathrm{ctrl}}, q_{\mathrm{U}_2,\mathrm{ctrl}}]^{\mathrm{T}} = [4200, 500]^{\mathrm{T}}$ then the LP is feasible.

Other states may be reached by flipping the $\geq$ sign for one or more rows of the system of inequalities (9) and checking whether the problem is feasible. If the inequality of the first row is flipped it will always result in an infeasible problem, since row 1 of inequalities (10) will not be satisfied. So, only the second or the third inequality may flipped or both, which will result in changing controller U_2 from active to inactive (state $\mathcal{S}_2 = [1, 0, 0]^{\mathrm{T}}$), changing controller U_1 from inactive to active (state $\mathcal{S}_3 = [1, 1, 1]^{\mathrm{T}}$), or both (state $\mathcal{S}_4 = [1, 1, 0]^{\mathrm{T}}$). For all the states $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3, \mathcal{S}_4$ the bottleneck is active, so the maximum achievable outflow will be $q^*_{\mathcal{S}_{\{1,2,3,4\}},\mathrm{D}_1} = q_{\mathrm{B}_1,\mathrm{dch}} = 3800\,\mathrm{veh/h}$.

It can be verified that the states $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3$, and $\mathcal{S}_4$ are reachable from each other, and no other states can be reached from states $\mathcal{S}_1, \mathcal{S}_2$, and $\mathcal{S}_4$, only from state $\mathcal{S}_3$.

For this reason, we now continue with state $\mathcal{S}_3 = [1, 1, 1]^{\mathrm{T}}$, where controller U_1 is also active. The corresponding LP reads:

$$\max_{q_{\mathrm{ctrl}}} \; q_{\mathrm{D}_1}, \text{ subject to}$$

$$\begin{bmatrix}0 & 0\end{bmatrix}\begin{bmatrix}q_{\mathrm{U}_1,\mathrm{ctrl}}\\ q_{\mathrm{U}_2,\mathrm{ctrl}}\end{bmatrix}+\begin{bmatrix}0 & 0 & 1\end{bmatrix}\begin{bmatrix}q_{\mathrm{O}_1}\\ q_{\mathrm{O}_2}\\ q_{\mathrm{B}_1,\mathrm{dch}}\end{bmatrix}=\begin{bmatrix}q_{\mathrm{D}_1}\end{bmatrix}\ , \tag{11}$$

$$\begin{bmatrix}1 & 1\\ 0 & -1\\ -1 & 0\end{bmatrix}\begin{bmatrix}q_{\mathrm{U}_1,\mathrm{ctrl}}\\ q_{\mathrm{U}_2,\mathrm{ctrl}}\end{bmatrix}+\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}q_{\mathrm{O}_1}\\ q_{\mathrm{O}_2}\\ q_{\mathrm{B}_1,\mathrm{dch}}\end{bmatrix}\geq\begin{bmatrix}q_{\mathrm{B}_1,\mathrm{dch}}\\ -q_{\mathrm{O}_2}\\ -q_{\mathrm{O}_1}\end{bmatrix}\ . \tag{12}$$

$$\begin{bmatrix}1 & 0\\ -1 & 0\\ 0 & 1\\ 0 & -1\end{bmatrix}\begin{bmatrix}q_{\mathrm{U}_1,\mathrm{ctrl}}\\ q_{\mathrm{U}_2,\mathrm{ctrl}}\end{bmatrix}+\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}q_{\mathrm{O}_1}\\ q_{\mathrm{O}_2}\\ q_{\mathrm{B}_1,\mathrm{dch}}\end{bmatrix}\geq\begin{bmatrix}q_{\mathrm{U}_1,\min}\\ -q_{\mathrm{U}_1,\max}\\ q_{\mathrm{U}_2,\min}\\ -q_{\mathrm{U}_2,\max}\end{bmatrix}\ . \tag{13}$$

The only changes are in the first and the third row of equation (12). These changes are related to the fact that the inflow of the bottleneck is not determined by q_{O_1} anymore, but by $q_{\mathrm{U}_1,\mathrm{ctrl}}$ (since the measure is active), and to the fact that the measure U_1 is active. Also in this state, the flow q_{D_1} is determined by the bottleneck queue discharge rate $q_{\mathrm{D}_1,\mathrm{dch}}$. The reachable states from this state include the already examined states $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_4$, and include some new activity states, since flipping the inequality of the first row now results in a feasible LP. It can be verified that the states $\mathcal{S}_5 = [0,1,1]^{\mathrm{T}}, \mathcal{S}_6 = [0,0,1]^{\mathrm{T}}$, and $\mathcal{S}_7 = [0,1,0]^{\mathrm{T}}$ are reachable. These states correspond to an inactive bottleneck, which means that the combination of a speed limit and ramp metering can resolve the jam at the bottleneck. These states are not reachable from the states where only ramp metering is active ($\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_4$). So, ramp metering only is not powerful enough to resolve the jam at bottleneck B_1 *for the given traffic demand.*

We continue here with the optimization of state $\mathcal{S}_5$, the other states could be optimized similarly. The LP for state $\mathcal{S}_5$ reads:

$$\max_{q_{\mathrm{ctrl}}}\ q_{\mathrm{D}_1},\ \text{subject to}$$

$$\begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}q_{\mathrm{U}_1,\mathrm{ctrl}}\\ q_{\mathrm{U}_2,\mathrm{ctrl}}\end{bmatrix}+\begin{bmatrix}0 & 0\end{bmatrix}\begin{bmatrix}q_{\mathrm{O}_1}\\ q_{\mathrm{O}_2}\end{bmatrix}=\begin{bmatrix}q_{\mathrm{D}_1}\end{bmatrix}\ , \tag{14}$$

$$\begin{bmatrix}-1 & -1\\ 0 & -1\\ -1 & 0\end{bmatrix}\begin{bmatrix}q_{\mathrm{U}_1,\mathrm{ctrl}}\\ q_{\mathrm{U}_2,\mathrm{ctrl}}\end{bmatrix}+\begin{bmatrix}0 & 0\\ 0 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}q_{\mathrm{O}_1}\\ q_{\mathrm{O}_2}\end{bmatrix}\geq\begin{bmatrix}-q_{\mathrm{B}_1,\mathrm{cap}}\\ -q_{\mathrm{O}_2}\\ -q_{\mathrm{O}_1}\end{bmatrix}\ . \tag{15}$$

$$\begin{bmatrix}1 & 0\\ -1 & 0\\ 0 & 1\\ 0 & -1\end{bmatrix}\begin{bmatrix}q_{\mathrm{U}_1,\mathrm{ctrl}}\\ q_{\mathrm{U}_2,\mathrm{ctrl}}\end{bmatrix}+\begin{bmatrix}0 & 0\\ 0 & 0\\ 0 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}q_{\mathrm{O}_1}\\ q_{\mathrm{O}_2}\end{bmatrix}\geq\begin{bmatrix}q_{\mathrm{U}_1,\min}\\ -q_{\mathrm{U}_1,\max}\\ q_{\mathrm{U}_2,\min}\\ -q_{\mathrm{U}_2,\max}\end{bmatrix}\ . \tag{16}$$

The changes here are now in equation (14) where the outflow does not depend on the queue discharge rate of the bottleneck anymore, but on the control vector q_{ctrl}, and in the first row of equation (15). which contains the condition for the inactive bottleneck. Furthermore, the constant $q_{\mathrm{B}_1,\mathrm{dch}}$ has disappeared from q_{const} since the bottleneck is not active anymore.

The solution of this LP problem will result in multiple solutions for which the outflow reaches the capacity of the bottleneck, e.g., $q_{\mathrm{U}_1,\mathrm{ctrl}} = 3700$ veh/h and $q_{\mathrm{U}_2,\mathrm{ctrl}} = 500$ veh/h. It is obvious that the other states ($\mathcal{S}_6$ and $\mathcal{S}_7$) cannot result in a better performance, where $q^{*}_{\mathcal{S}_{\{5,6,7\}},\mathrm{D}_1} = q_{\mathrm{B}_1,\mathrm{cap}} =$ 4200 veh/h.

Note also that state $\mathcal{S}_8 = [0,0,0]^{\mathrm{T}}$ is infeasible, since for this state the inflow of the bottleneck equals 4100 veh/h (sum of all demands) which is higher than its capacity, which would lead to an active bottleneck, not to an inactive one as expressed by the state.

We may conclude that the optimal state $\mathcal{S}^* = \mathcal{S}_5$ is found, and also the corresponding control values and the state sequence ($\mathcal{S}_1 \to \mathcal{S}_3 \to \mathcal{S}_5$) is identified.

4 Conclusions and future research

We have developed an approach that finds the best achievable network activity state (the best achievable combination of active and inactive bottlenecks) including the corresponding control signals, and active and inactive control measures assuming constant origin demands and no blocking in the network. If the optimal network state is not directly reachable from the initial activity state, the approach also provides all possible state transition sequences that lead to the optimal state. The physical interpretation of this result is that this method finds the sequence of necessary control measures that solve all traffic jams – or if that is not possible – to reach the best possible relocation of traffic jams.

The most important network element in this approach is the bottleneck, with which the capacity drop, hysteresis effects, and metastability can be reproduced. The main goal of this approach was to minimize the total time spent by the vehicles, – which was shown to be equivalent to maximizing the outflow of the network,– while to coping with these effects.

Since most traffic control measures can only produce a limited range of flows, upper and lower bounds were incorporated in the approach. These bounds are useful to determine whether a certain combination of control measures is powerful enough to solve a certain traffic jam. In this sense this approach is also suitable to determine whether the addition of an extra control measure has the potential to improve the network performance.

Most of the topics for future research are related to the relaxation of the above assumptions:

- **Inclusion of queues.** Explicit modeling of the dynamics of the queues occurring at bottlenecks and control measures will enable us to compute the TTS corresponding to a certain activity state sequence, to model blocking effects and to include queue length constraints.
- **Representation of upstream propagating shock waves.** Similarly to vertical queues, upstream propagating shock waves may also influence the behavior of other network elements (such as control measures and bottlenecks).
- **Extension to dynamic demands.** Dynamic demands may also trigger activity state changes which may interfere with the state transitions triggered by the control measures.
- **Investigation of more efficient algorithms.** The algorithms presented in this paper for finding the optimal state are not optimized for efficiency yet. The computation time to solve the LP problems is not expected to become a problem since its size increases linearly with the number of network elements. However, finding the activity state sequence leading to the optimal activity state may grow exponentially in complexity, since the number of states increases exponentially with the number of bottlenecks and control measures. In the future more effective search algorithms could be used to limit the number of cases for which an LP has to be solved.

Acknowledgments

Research supported by the project “Multi-agent control of large-scale hybrid systems”(DWV.6188) of the Dutch Technology Foundation STW, by the TU Delft spearhead program “Transport Research Centre Delft: Towards Reliable Mobility”, and by the BSIK project “Transition Towards Sustainable Mobility (TRANSUMO)”.

References

[1] B. S. Kerner. Empirical features of congested patterns at highway bottlenecks. In *Proceedings of the 81st Annual Meeting of the Transportation Research Board*, Washington, D.C., 2002.

[2] M. Treiber, A. Hennecke, and D. Helbing. Congested traffic states in empirical observations and microscopic simulations. *Physical Review E*, 62(2):1805–1824, August 2000.

[3] F.L. Hall and K. Agyemang-Duah. Freeway capacity drop and the definition of capacity. *Transportation Research Record*, (1320):91–98, 1991.

[4] A. Hegyi. *Model Predictive Control for Integrating Traffic Control Measures*. Ph.D. thesis, TRAIL thesis series T2004/2, Delft University of Technology, Delft, The Netherlands, February 2004. ISBN 90-5584-053-X.

[5] R. Barlovic, L. Santen, A. Schadschneider, and M. Schreckenberg. Metastable states in cellular automata for traffic flow. *The European Physical Journal B*, 5(3):793–800, 1998.

[6] B. Kerner and H. Rehborn. Experimental properties of phase transitions in traffic flow. *Physical Review Letters*, 79(20):4030–4033, November 1997.

[7] M. Papageorgiou. An integrated control approach for traffic corridors. *Transportation Research Part C*, 3(1):19–30, 1995.

[8] A. Kotsialos, M. Papageorgiou, M. Mangeas, and H. Haj-Salem. Coordinated and integrated control of motorway networks via non-linear optimal control. *Transportation Research Part C*, 10:65–84, 2002.

[9] A. Alessandri, A. Di Febbraro, A. Ferrara, and E. Punta. Optimal control of freeways via speed signalling and ramp metering. *Control Engineering Practice*, 6:771–780, 1998.

[10] A. Hegyi, B. De Schutter, and J. Hellendoorn. Optimal coordination of variable speed limits to suppress shock waves. *Transportation Research Record*, (1852):167–174, 2004.

[11] M. Papageorgiou, J.M. Blosseville, and H. Hadj-Salem. La fluidification des rocades de l'Ile de France: Un projet d'importance. Technical report, Dynamic Systems and Simulation Laboratory, Technical University of Crete, Chania, Greece, 1998. Internal Report No. 1998-17.

[12] A. Schrijver. *Theory of Linear and Integer Programming*. John Wiley & Sons, Chichester, UK, 1986.

[13] P.M. Pardalos and M.G.C. Resende, editors. *Handbook of Applied Optimization*. Oxford University Press, Oxford, UK, 2002.

[14] Y. Nesterov and A. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM, Philadelphia, Pennsylvania, 1994.

[15] A. Hegyi, B. De Schutter, and J. Hellendoorn. Model predictive control for optimal coordination of ramp metering and variable speed limit control. *Transportation Research Part C*, 2004. Accepted for publication.