Node modeling for congested urban road networks

Saif Eddin Jabari

Division of Engineering, New York University Abu Dhabi, Saadiyat Island, P.O. Box 129188, Abu Dhabi, U.A.E.

Abstract

First-order network flow models are coupled systems of differential equations which describe the build-up and dissipation of congestion along network road segments, known as link models. Models describing flows across network junctions, referred to as node models, play the role of the coupling between the link models and are responsible for capturing the propagation of traffic dynamics through the network. Node models are typically stated as optimization problems, so that the coupling between the link dynamics is not known explicitly. This renders network flow models analytically intractable. This paper examines the properties of node models for urban networks. Solutions to node models that are free of traffic holding, referred to as holding-free solutions, are formally defined and it is shown that flow maximization is only a sufficient condition for holding-free solutions. A simple greedy algorithm is shown to produce holding-free solutions while also respecting the invariance principle. Staging movements through nodes in a manner that prevents conflicting flows from proceeding through the nodes simultaneously is shown to simplify the node models considerably and promote unique solutions. The staging also models intersection capacities in a more realistic way by preventing unrealistically large flows when there is ample supply in the downstream and preventing artificial blocking when some of the downstream supplies are restricted.

Keywords: urban networks, kinematic waves, node models, traffic holding, flow maximization, invariance principle, simultaneous movements and conflicts, signalized intersections

E-mail: sej7@nyu.edu.

URL http://engineering.nyu.edu/people/saif-eddin-jabari.

NOTICE: this is the author’s version of a work that was accepted for publication in Transportation Research Part B. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Transportation Research Part B, vol.91, (2016) DOI: 10.1016/j.trb.2016.06.001. http://dx.doi.org/10.1016/j.trb.2016.06.001
1. Introduction

In congested urban networks, we now know that combating a traffic problem at one location might best be done by adjusting the controls at another location. But how to do so requires an understanding of how congestion propagates in a network. Over the past six decades, numerous modeling approaches have been developed that can successfully describe the spatio-temporal evolution of congestion along individual road segments (Lighthill and Whitham, 1955; Richards, 1956; Prigogine and Andrews, 1960; Paveri-Fontana, 1975; Kerner and Konháuser, 1994; Zhang, 1998; Treiber et al., 1999; Aw and Rascle, 2000). Most of these theories were developed for roads of infinite length (or closed circuits) and draw their inspiration from models that describe fluid flow or gas dynamics.

It was not until the mid 1990’s that traffic flow theories of road segments with boundaries began to emerge (Daganzo, 1995; Lebacque, 1996; Holden and Risebro, 1995). Today, we have good theories for describing/computing flows through simple junctions that connect one incoming road segment to two (or more) outgoing road segments, diverge nodes, and those that connect one outgoing segment to two (or more) incoming segments, merge nodes (Daganzo, 1995; Ni and Leonard, 2005; Jin, 2010; Jin and Zhang, 2013; Jin, 2015). These node models, combined with suitable models of flow in individual links, referred to as link models, can be used to effectively model/compute traffic dynamics in expressway networks. In such networks, the link models play a major role in representing traffic conditions since the network nodes tend to be located at fairly long distances from one another. The situation is quite different for urban/city networks, where the road segments are short and the nodes can have multiple inbound links and multiple outbound links. In such networks, the nodes play a more critical role in describing the dynamics in the network.

When modeling flow rates across link boundaries, certain conditions need to be satisfied to ensure that the boundary flows make sense physically. For example, flows into (out of) the upstream (downstream) boundary of a link should not exceed the supply (demand) at the boundary. Treatments in which boundary flows are appended as source terms to continuity equations without imposing these conditions can lead to ill-posed problems. In general, the conditions on the boundary flows are imposed in such a way as to ensure realism in addition to well posedness of the mathematical model. Some of the literature on general node models have focused on the propagation of kinematic waves through nodes (Holden and Risebro, 1995; Coclite et al., 2005; Garavello and Piccoli, 2006; Jin, 2012). Due to the finite speeds at which traffic information propagates through network links, it is reasonable to investigate the node models independently of the link dynamics as in (Lebacque, 2005; Tampère et al., 2011; Flötteröd and Rohde, 2011; Smits et al., 2015). In general, all of these models satisfy supply and demand restrictions at link boundaries and mass-balance requirements on flows across the nodes. Most of them also consider restrictions on the distribution of demands to different outgoing links so as to reflect path/destination desires of drivers.

In many papers, flow maximization is imposed as an entropy condition aimed at restricting node model solution spaces. Flow maximization states that the total flows
through the node at any time instant should be maximized. The principle is motivated by the physical requirement that flows through a network node should not involve traffic holding, which is the undesirable situation in which vehicles do not proceed through the node despite availability of supply in the downstream. In this paper, it is demonstrated that traffic holding and flow maximization are not one and the same except in special cases, namely, merge and diverge nodes. For general intersections, flow maximized solutions are only sufficient conditions for holding-free solutions.

One of the themes in the node modeling literature is the derivation of formulas for calculating node flows which satisfy the aforementioned restrictions, e.g., (Bliemer, 2007; Jin and Zhang, 2004). Such closed form expressions for node flows are desirable: the link dynamics are represented by systems of differential equations, the node flows represent a coupling between the differential equations. Hence, closed form solutions not only facilitate the analysis of the network dynamics, but can also help to dramatically reduce the computational effort for dynamic network loading tools. There exist many efficient techniques in the literature for computing the forward dynamics, e.g., (Gentile, 2010; Raadsen et al., 2015). However, a challenge that persists today is the analytical intractibility of the network models. For example, using the most efficient loading tools, the effect of a perturbation to the inputs can, in many circumstances, only be investigated by re-solving the entire problem. As a result, their complexity can be prohibitive for real-time applications such as signal control, traffic assignment, and real-time estimation. An exception to this can be found in recent work that applies the link transmission model for link dynamics; the reader is referred to (Himpe et al., 2016) and references therein.

Lebacque and Khoshyaran (2005) demonstrated that some solution approaches can result in unrealistic dynamics on the links, namely, the possibility of waves propagating in the wrong direction. To overcome this behavior, they proposed a principle, referred to as the invariance principle, which states that when restricted by the demands (supplies), the node flow solutions should be invariant to increases in the supplies (demands). Unlike the other conditions mentioned above, the undesirable behavior that results from violations of the invariance principle is dynamical in nature. Hence, invariance is a property of node model solutions and the algorithms used to compute them. Violations of the invariance principle tend to arise when there are multiple solutions and the algorithm picks out a “bad” solution. The more recent solution techniques in the literature tend to honor the invariance principle; some do so implicitly (Tampère et al., 2011; Flötteröd and Rohde, 2011; Corthout et al., 2012) and some have employed other mechanisms to ensure that waves propagate in the correct direction (Jin, 2012).

Despite all of the restrictions that have been imposed on node models, the majority of (realistic) models and their associated algorithms can still produce different solutions (Corthout et al., 2012; Smits et al., 2015) to the same problem. These differences can result in dramatically different network dynamics under the same initial and boundary conditions. This non-uniqueness can be attributed to the node flow problems being insufficiently constrained to produce unique solutions for any set of boundary conditions. Different approaches in the literature incorporate different physical insights in order to pick out one of the possibly many solutions that all honor the general requirements mentioned.
above. The physical insight incorporated in this paper to overcome this non-uniqueness issue is that conflicting movements should not be allowed to proceed through the node simultaneously. This couples with knowledge of the signal phasing scheme results in trivial movement priorities. This, in turn, promotes unique solutions.

The rest of this paper is organized as follows: Section 2 gives an overview of first-order link models and summarizes the principles that node models enforce. Section 3 formulates the demand, supply, and demand distribution conditions and Section 4 formally defines holding-free solutions and their relation to flow maximizing solutions. Section 5 presents the invariance principle and an invariant greedy algorithm that produces holding-free solutions. The non-simultaneity of conflicting movements is discussed in Section 6. Section 7 gives simple numerical examples and Section 8 concludes the paper.

2. Background

A network flow model consists of a system of (partial) differential equations (namely, conservation laws), which are coupled at the road segment (link) boundaries. We give below a brief background on link models with boundaries and node models, which constitute the boundary couplings. Note that the types of networks that are primarily considered in this paper are dense urban grid networks.

2.1. Link models

Let $G = (L, N)$ be a directed graph representing a road network with $L$ being the set of link indices and $N$ the set of node indices. Scalar conservation laws describe the traffic flows along the interiors of the network links and the node models provide the boundary conditions in the network. A schematic is shown in Fig. 1.

![Link-node network schematic](image)

Fig. 1: Link-node network schematic

For each $i \in L$ with upstream boundary position $l_i$ and downstream boundary position $u_i$, we denote the traffic density at time $t \in \mathbb{R}_+$ and location $x$ by $\rho_i(x, t)$. With slight abuse
of notation, the restriction of \( x \) to the interval \([l_i, u_i]\) is to be understood implicitly. Denote by \( Q_i \) the equilibrium flow density relation pertaining to link \( i \) and let \( \Delta_i \) and \( \Sigma_i \) denote the associated demand and supply functions, respectively. These relations are depicted in Fig. 2. One may also consider modified demand functions as proposed, for instance, in Monamy et al. (2012); Srivastava et al. (2015) in order to capture slow starts and the resulting capacity drops that are observed; for such cases, the demand function is denoted by \( \Delta_{\text{mod},i} \). See Fig. 3 for an example of modified demand functions.

![Fig. 2: Equilibrium relations (a) fundamental relation, (b) demand function, (c) supply function](image1)

![Fig. 3: Modified demand function](image2)

In this paper, it will be assumed that individual links are homogeneous road segments, so that \( Q_i \) does not depend on \( x \). Where the link characteristics do vary (e.g., a lane drop or change in speed limit), a node can be placed to separate links with different traffic characteristics. This is reasonable since in real-world networks changes in traffic characteristics that result in different equilibrium relations occur over distances that are large enough to warrant separate links.

Various approaches appear in the literature for modeling (and computing) first-order traffic dynamics. The different types of models are, in essence, variants of the non-linear hyperbolic conservation laws first proposed in (Lighthill and Whitham, 1955; Richards, 1956), which are achieved by means of coordinate transformation. The latter variants include Hamilton-Jacobi equations applied to traffic flow (Daganzo, 2005; Claudel and Bayen, 2010; Friesz et al., 2013) and Lagrangian coordinate transformations (Leclercq et al., 2010).

While the node modeling principles presented below should apply regardless of the type of first-order link model used, for the sake of concreteness, we will consider non-
linear hyperbolic conservation laws as models of link dynamics:

$$\frac{\partial \rho_i}{\partial t} + \frac{\partial Q_i(\rho_i)}{\partial x} = 0, \ i \in \mathcal{L}, \ x \in (l_i, u_i),$$ (1)

where the dependence on \((x, t)\) is omitted for notation clarity. Equations (1) represent the traffic dynamics within the interiors of the network links, i.e., \(x \in (l_i, u_i)\). Let \(q_i(l_i, t)\) and \(q_i(u_i, t)\) denote prescribed flows at the boundaries of link \(i\) at time \(t\). Then the traffic densities at the link boundaries ensuring well posedness are given, for each \(i \in \mathcal{L}\), by\(^2\):

$$\rho_i(l_i, t+) = \begin{cases} \Delta_i^{-1}(q_i(l_i, t)), & q_i(l_i, t) \leq \Sigma_i(\rho_i(l_i, t)) \\ \rho_i(l_i+, t), & q_i(l_i, t) > \Sigma_i(\rho_i(l_i, t)) \end{cases}$$ (2)

and

$$\rho_i(u_i, t+) = \begin{cases} \Sigma_i^{-1}(q_i(u_i, t)), & q_i(u_i, t) \leq \Delta_{mod,i}(\rho_i(u_i, t)) \\ \rho_i(u_i-, t), & q_i(u_i, t) > \Delta_{mod,i}(\rho_i(u_i, t)) \end{cases}$$ (3)

Given the prescribed boundary flows, \(q_i(l_i, t)\) and \(q_i(u_i, t)\), these conditions ensure that the boundary traffic densities reflect the correct traffic regime (e.g., free-flow or congested) in accordance with whether demands and supplies are satisfied. For further details on these relations, we refer to (Lebacque, 2005; Lebacque and Khoshyaran, 2005) and references therein.

2.2. Modeling node flows (node models)

Node models will play the role of determining the prescribed boundary flows, \(q_i(l_i, t)\) and \(q_i(u_i, t)\) for each \(i \in \mathcal{L}\) in such a way as to capture the interactions between the inbound and outbound link demands and supplies and ensure that the boundary flows are physically reasonable. The conditions enforced by the node models considered in this paper are summarized as follows:

1. Supply and demand constraints: flow rates into a node through inbound links should not exceed the local demands at the downstream boundaries of those links and flow rates through the node into its outbound links should not exceed the local supplies at the upstream boundaries of those links.

2. Mass balance (or conservation of node flows): the total flow out of the inbound links should equal the total flow entering the outbound links.

3. Demand distribution constraints: the distribution of demand at each inbound link to the outbound links should be in accordance with splits representing the path/destination desires of the drivers.

\(^2\)When the demand function is the modified version, \(\Delta_{mod,i}\), and \(q_i(l_i, t) \leq \Sigma_i(\rho_i(l_i, t))\), the upstream boundary density at time instant \(t+\) is subcritical. We therefore, retain the unmodified version of the demand function, \(\Delta_i\) for calculating \(\rho_i(l_i, t+)\)
4. **No traffic holding**: when there is available supply in drivers’ desired destination links, drivers will use it.

5. **The invariance principle**: when a prescribed flow is demand (supply) constrained, it should be invariant to increases in supply (demand).

6. **Non-simultaneity of conflicting flows**: conflicting movements do not proceed through an intersection node simultaneously.  

Denote by $I_n \subset L$ and $O_n \subset L$ the sets of indices of inbound and outbound links to node $n \in \mathcal{N}$, respectively. **First-order network flow model** consists of $|L|$ link models, the boundary conditions (2)-(3), and prescribed boundary flows $\{q_i(l_i,t)\}_{i \in I_n} \cup \{q_i(u_i,t)\}_{i \in O_n}$. The prescribed flows should respect conditions 1-6 above for each node $n \in \mathcal{N}$ and each time step $t \in \mathbb{R}_+$. The term “first-order” is a characteristic of the link models; in this paper, we restrict our attention to Lighthill-Whitham-Richards (LWR) link models.

### 3. Supply, demand, and distribution restrictions

Since the wave speeds are finite and the conditions are to be honored for each time instant, node dynamics can be studied for each node and each time instant separately. Consequently, we may simplify notation and minimize clutter. We will omit the subscript $n$ and the dependence on a fixed time $t$ where no confusion will arise. We will also denote by $q_i \equiv q_i(u_i,t)$ and $q_o \equiv q_o(l_o,t)$ the prescribed inbound and outbound flows for $i \in I$ and $o \in O$, respectively. Similarly, $\delta_i \equiv \Delta_i(\rho_i(u_i,t))$ and $\sigma_o \equiv \Sigma_o(\rho_o(l_o,t))$.

The supply and demand conditions (along with non-negativity) are given by

\[ 0 \leq q_i \leq \delta_i \text{ for each } i \in I \]
\[ 0 \leq q_o \leq \sigma_o \text{ for each } o \in O \]  

(4)

and the mass balance condition is written as

\[ \sum_{i \in I} q_i - \sum_{o \in O} q_o = 0. \]  

(5)

\[ \sum_{o \in O} f_{i,o} = 1. \]  

(6)

Then condition 3 is given by

\[ q_o = \sum_{i \in I} f_{i,o} q_i, \text{ for each } o \in O. \]  

(7)

Notice that, from (7) and since $q_i \geq 0$ for each incoming link $i$ and $f_{i,o} \geq 0$ for every $(i,o)$ pair, $q_o \geq 0$ is implicitly ensured for each $o$. Next, summing both sides of (7) over $o$,  

\[ \sum_{o \in O} q_o = \sum_{i \in I} \sum_{o \in O} f_{i,o} q_i \leq \sum_{i \in I} \sum_{o \in O} \delta_i f_{i,o} = \sum_{i \in I} \delta_i, \]  

which shows that the total supply is bounded by the total demand.

---

3We note that this last condition only applies in the case of interrupted flow junctions and is not needed when modeling expressway junctions.
we have that
\[
\sum_{o \in O} q_o = \sum_{o \in O} \sum_{i \in I} f_{i,o} q_i = \sum_{i \in I} \sum_{o \in O} f_{i,o} q_i = \sum_{i \in I} q_i \left( \sum_{o \in O} f_{i,o} \right) = \sum_{i \in I} q_i. \tag{8}
\]

Hence, mass balance is also implicitly ensured and conditions 1-3 are honored provided that the following inequalities hold
\[
0 \leq q_i \leq \delta_i \text{ for each } i \in I
\]
\[
\sum_{i \in I} f_{i,o} q_i \leq \sigma_o \text{ for each } o \in O. \tag{9}
\]

In the following sections, we discuss the impact of incorporating each of the conditions 4-6 incrementally.

4. Traffic holding and flow maximization

Traffic holding has a clear physical meaning and should be prevented by any realistic node model. Flow maximization on the other hand has no clear physical/behavioral interpretation, as noted by Smits et al. (2015). In this section, it will be demonstrated that flow maximization is a sufficient, but not necessary, condition for holding-free node flows. That is, for a general network node, there exist holding-free solutions which are not flow maximizing. We will also discuss special cases where holding-free solutions are necessarily flow maximizing as well.

4.1. Holding-free solutions

We first formally define holding-free node flows and then show that they arise as solutions to a complementarity problem.

**Definition 1** (Holding-free solution). A holding-free solution (HFS) is defined as a set of non-negative flows that satisfy conditions 1-3 and where no unsatisfied demands can proceed to their desired outbound links without violating at least one supply constraint.

**Theorem 1** (HFS complementarity condition). Let \( q \equiv [q_{i_1} \cdots q_{i_{|\mathcal{I}|}} \, q_{o_1} \cdots q_{o_{|\mathcal{O}|}}]^T \) be a flow vector, where \( \mathcal{I} = \{i_1, \cdots, i_{|\mathcal{I}|} \} \) and \( \mathcal{O} = \{o_1, \cdots, o_{|\mathcal{O}|} \} \).

(i) If \( q \) is a HFS, then
\[
(\delta_i - q_i) \prod_{o \in \mathcal{O} : f_{i,o} > 0} (\sigma_o - q_o) = 0, \quad i \in \mathcal{I}. \tag{10}
\]

(ii) Conversely, if \( q \) satisfies (9) and (10), then \( q \) is a HFS.
PROOF. (i) A demand \( i \) is satisfied if and only if \( q_i = \delta_i \) so that this case is trivial. Let \( I' \subset I \) denote the subset of indices corresponding to flows with unsatisfied demands, that is, for each \( i \in I' \), \( q_i < \delta_i \). Then, since \( q \) is a HFS, for each unsatisfied demand \( i \), there exists at least one \( o \in O \) such that \( f_{i,o} > 0 \) and for which \( q_o = \sigma_o \). Otherwise, it would be possible for unsatisfied demands along link \( i \) to proceed to their desired outbound link. Hence (10) holds whenever \( q \) is a HFS.

(ii) Since \( q \) satisfies (9), conditions 1-3 are also satisfied. For each \( i \) with \( q_i = \delta_i \), the holding free property holds trivially since in this case, there are no unsatisfied demands.

Let \( I' \) be as above, then for all \( i \in I' \), we have that

\[
\prod_{o \in O: f_{i,o} > 0} (\sigma_o - q_o) = 0, \quad (11)
\]

so that there exists at least one \( o' \in O \) with \( f_{i',o'} > 0 \) and \( q_{o'} = \sigma_{o'} \). From (7), and since \( q \) satisfies conditions 1-3 by assumption, we have that

\[
\sigma_{o'} = q_{o'} = \sum_{i \in I} f_{i,o'} q_i, \quad (12)
\]

Hence, any increase in \( q_i \) results in a violation of the supply constraint associated with outbound link \( o' \). It follows that \( q \) is a HFS. \( \square \)

As a consequence of Theorem 1 it follows that a HFS is any set of flows that satisfy the following conditions:

\[
(\delta_i - q_i) \prod_{o \in O: f_{i,o} > 0} (\sigma_o - \sum_{i' \in I} f_{i',o} q_{i'}) = 0, \quad i \in I
\]

\[
\sum_{i \in I} f_{i,o} q_i \leq \sigma_o, \quad o \in O
\]

\[
0 \leq q_i \leq \delta_i, \quad i \in I. \quad (13)
\]

Many of the node flow solution algorithms in the literature produce holding-free solutions, but some do not. For example, node flow algorithms that split the supply in accordance with the ratios of capacities may not produce holding-free solutions. Denote the capacity of link \( i \) by \( q_{i,\max} = \max_{\rho_i} Q_i(\rho_i) = \max_{\rho_i} \Delta_i(\rho_i) = \max_{\rho_i} \Sigma_i(\rho_i) \) and consider the simple merge shown in Fig. 4.

Suppose \( \delta_1 = 2100 \text{ veh/hr}, \delta_2 = 1400 \text{ veh/hr}, \sigma = 3000 \text{ veh/hr}, q_{1,\max} = q_{2,\max} = 2200 \text{ veh/hr}, \) and \( f_1 = f_2 = 1.4 \).\(^4\) Consider the following solution formula:

\[
q_i = \min \left\{ \delta_i, \sigma \frac{q_{\max}}{\sum_{I} f_{i'} q_{i',\max}} \right\}, \quad i \in I. \quad (14)
\]

\(^4\)This example was adopted from (Lebacque and Khoshyaran, 2005).
Using (14), we get $q_1 = 1500$ veh/hr, $q_2 = 1400$ veh/hr, and $q = 2900$ veh/hr. The left-hand side of the complementarity equation for $i = 1$ is then $(\delta_1 - q_1)(\sigma - q) = (2100 - 1500)(3000 - 2900) > 0$, so that (14) does not produce holding-free solutions.

As another example, consider the following formula, which was first proposed by Jin and Zhang (2004):

$$\hat{q}_i = \min \left\{ 1, \min \left\{ \frac{\sigma_{o'}}{\sum_{i' \in I} f_{i',o'} \delta_{i'}} \right\} \right\}, \; i \in I. \tag{15}$$

It can be demonstrated that this formula does not produce holding-free solutions: Suppose

$$\min_{o' \in O} \left\{ \frac{\sigma_{o'}}{\sum_{i' \in I} f_{i',o'} \delta_{i'}} \right\} < 1. \tag{16}$$

Then, for all $i \in I$, we have that $\hat{q}_i < \delta_i$. Now suppose there exists a pair $i' \in I$ and $o' \in O$ such that $f_{i',o'} = 1$, i.e., $f_{i,o'} = 0$ for all $i \neq i'$; that is, we have an outbound link $o'$ that can only be reached by a single inbound link $i'$ and all flow on $i'$ is destined to $o'$. Also, assume that $\sigma_{o'} \geq \delta_{i'}$. Then, using (15), we have that $\hat{q}_{i'} < \delta_{i'}$ and $\sum_{i \in I} f_{i,o'} \hat{q}_i < \sigma_{o'}$. Hence, (15) cannot guarantee holding-free solutions.

4.2. Flow maximizing solutions

A flow maximizing scheme is one that honors conditions 1-3 while also maximizing the total inbound flows to the node (hence, also the total outbound flows by mass-balance). A flow maximizing solution is one that solves the following linear program:

$$\max \{ q_i \}_{i \in I} \sum_{i \in I} q_i$$

s.t. $\sum_{i \in I} f_{i,o} q_i \leq \sigma_o$, $o \in O$

$$0 \leq q_i \leq \delta_i, \; i \in I. \tag{17}$$

Theorem 2 below formally establishes the intuitive, albeit subtle, fact that any flow maximizing solution is also holding-free.

**Theorem 2** (Sufficiency of flow maximizing solutions). Let $q^* \equiv [q_1^* \cdots q_{|I|}^*]^{\top}$ solve (17) to optimality. Then, $q^*$ is a HFS.
PROOF. Assume $q^*$ is not a HFS. Then, there exists at least one $i \in \mathcal{I}$ for which
\[
(\delta_i - q^*_i) \prod_{o \in \mathcal{O}: \ f_{i,o} > 0} (\sigma_o - \sum_{i' \in \mathcal{I}} f_{i',o} q^*_{i'}) > 0.
\] (18)

For any $i$ satisfying (18), define
\[
\varepsilon(i) \equiv \min \left\{ (\delta_i - q^*_i), \min_{o \in \mathcal{O}} \left\{ \frac{1}{f_{i,o}} (\sigma_o - \sum_{i' \in \mathcal{I}} f_{i',o} q^*_{i'}) \right\} \right\},
\] (19)

where if $f_{i,o} = 0$ for some $o \in \mathcal{O}$, then $\frac{1}{f_{i,o}} (\sigma_o - \sum_{i' \in \mathcal{I}} f_{i',o} q^*_{i'}) = \infty$, which is always larger than $\delta_i - q^*_i$ so that (19) makes sense.

It follows from (18) that $\varepsilon(i) > 0$. Let $\hat{q} \equiv q^* + \varepsilon(i) e_i$, where $e_i$ is a standard basis vector of size $|\mathcal{I}|$. It is easy to see that $\hat{q}$ is a feasible solution to (17) with an objective value that is strictly larger than $q^*$. Hence, $q^*$ is not an optimal solution and we have a contradiction.

Flow maximization is not a necessary condition for ensuring holding-free solutions. This non-necessity can be easily demonstrated by means of a counter-example. Consider the following solution formula:
\[
\hat{q}_i = \delta_i \min \left\{ 1, \min_{o' \in \mathcal{O}: \ f_{i,o'} > 0} \left\{ \frac{\sigma_{o'}}{\sum_{i' \in \mathcal{I}} f_{i',o'} \delta_{i'}} \right\} \right\}, \ i \in \mathcal{I}.
\] (20)

This formula modifies (15) by taking the minimum inside over outbound links with strictly positive splits as opposed to positive splits. Tampère et al. (2011) attribute this formula to Bliemer (2007).

It can be shown that the vector $\hat{q}$ with elements given by (20) is a HFS (see Appendix A); this is in contrast to solutions obtained using (15) as demonstrated above. However, it cannot be guaranteed that $\hat{q}$ solves (17) to optimality; that is, $\hat{q}$ is not necessarily flow maximizing. In fact, in most cases it is not.

To demonstrate that (20) cannot guarantee flow maximizing solutions, consider a general intersection with inbound links $\mathcal{I}$ and outbound links $\mathcal{O}$. Assume, without any loss of generality, that $\delta_i > 0$ for each $i$, $\sigma_o > 0$ for each $o$, and $f_{i,o} > 0$ for each $(i,o)$ pair. Also, assume that the optimal solution to (17) for this node is unique\(^5\). In this case, the flow maximizing solution lies at an extreme point of the feasible region (a basic feasible solution). This is guaranteed, for instance, whenever none of the hyperplanes corresponding to the left hand sides of the supply constraints are parallel to the objective function. Clearly, if (20) produces non-basic solutions, points that lie on the interior of the feasible

\(^5\)After all, it only takes a single counter-example to establish the result and, more often than not, for a general network node the optimal solution to (17) is indeed unique.
region of (17), then such solutions cannot be flow maximizing (i.e., they cannot solve (17) to optimality). This is indeed the case, especially in situations when not all demands can be satisfied by the available supplies. This corresponds to the situation when

\[
\min \left\{ 1, \min_{\sigma_o' \in O, f_{i,o'} > 0} \left\{ \frac{\sigma_o'}{\sum_{i' \in I} f_{i',o'} \delta_{i'}} \right\} \right\} < 1.
\] (22)

In this case, we have that \(0 < \hat{q}_i < \delta_i\) for each \(i \in \mathcal{I}\). This means that we have at least \(2|\mathcal{I}|\) basic variables in the solution space of (17): the \(\hat{q}_i\) variables themselves and the slack variables associated with the demand constraints. Hence, if \(|\mathcal{I}| > |\mathcal{O}|\), then \(\hat{q}\) is not a basic feasible solution (extreme point). This serves as one example of a HFS that is not flow maximizing.

Moreover, if \(|\mathcal{I}| < |\mathcal{O}|\), then the minimum over \(o' \in \mathcal{O}\) needs to be attained for at least \(|\mathcal{O}| - |\mathcal{I}|\) outbound links. To be precise, let \(\mathcal{O}' \subset \mathcal{O}\) be such that (i) for all \(o' \in \mathcal{O}'\)

\[
\frac{\sigma_o'}{\sum_{i' \in I} f_{i',o'} \delta_{i'}} = \min_{o \in \mathcal{O}} \left\{ \frac{\sigma_o}{\sum_{i' \in I} f_{i,o'} \delta_{i'}} \right\}
\] (23)

and (ii) \(|\mathcal{O}'| = |\mathcal{O}| - |\mathcal{I}|\). It is only when both of these (very restrictive) conditions are satisfied that \(\hat{q}\) corresponds to a basic feasible solution to (17), i.e., an extreme point. Note that this still does not guarantee optimality. However, it does exhibit the many ways in which a HFS is not necessarily flow maximizing. In Subsection 7.1, we give an example demonstrating that \(\hat{q}\) is not flow maximizing, despite being an extreme point of the feasible region.

In Appendix B, we show that (13) is equivalent to the incremental node model (INM) of Flötteröd and Rohde (2011). The complementarity formulation (13) has the advantage of offering simple conditions, which are easy to analyze. As a result of this equivalence, we also have that the approach proposed by Tampère et al. (2011) also produces holding-free solutions. This is also the case for the algorithm proposed by Gibb (2011).

**Remark 1.** Flow maximization and holding free solutions are commonly confused in the literature and it is important to stress the difference for reasons that go beyond node models. If it were necessary, one implication is that solution formulas such as (20) would serve as closed form formulas for any linear programming problem of the form (17). For instance, generalizations with non-negative cost coefficients have the same form. Also,

---

6Clearly, if

\[
\min \left\{ 1, \min_{\sigma_o' \in O, f_{i,o'} > 0} \left\{ \frac{\sigma_o'}{\sum_{i' \in I} f_{i',o'} \delta_{i'}} \right\} \right\} = 1
\] (21)

then \(\hat{q}\) solves (17) to optimality, since in this case \(\hat{q}_i = \delta_i\) for all \(i\).

7Note that we no longer need to state the \(f_{i,o'} > 0\) condition in the minimum since we are assuming that this holds for all \(i,o\) pairs.
problems with left-hand side coefficients that add up to any constant (column-wise) other than unity still have the same structure and would be solvable using a formula like (20) or the greedy algorithm presented in Section 5. Alas, this is not the case.

4.3. Diverge and merge nodes

In the special cases of diverge and merge nodes, HFS and flow maximization are equivalent. We have already established in Theorem 2 that flow maximization implies the HFS property. To demonstrate that they are equivalent, first consider the case of a diverge node, i.e., $|\mathcal{I}| = 1$. Let $\hat{q}$ denote the HFS. In this case (13) becomes

\[
(\delta - \hat{q}) \prod_{o \in \mathcal{O}; f_o > 0} (\sigma_o - f_o \hat{q}) = 0
\]

\[
\hat{q} \leq \frac{\sigma_o}{f_o}, \ o \in \mathcal{O}
\]

\[
0 \leq \hat{q} \leq \delta.
\] (24)

From the equality condition, we have that either $\hat{q} = \delta$ or $\hat{q} = \frac{\sigma_o}{f_o}$ for at least one $o$, or both. Since $\hat{q}$ also satisfies the supply and demand constraints, we have that

\[
\hat{q} = \min \left\{ \delta, \min_{o \in \mathcal{O}; f_o > 0} \frac{\sigma_o}{f_o} \right\}.
\] (25)

It can be easily shown that (25) also solves (17), which in this case is given by

\[
\begin{align*}
\max q \\
\text{s.t. } q & \leq \frac{\sigma_o}{f_o} \text{ for } o \in \mathcal{O} \\
0 & \leq q \leq \delta.
\end{align*}
\] (26)

In the case of a merge node, i.e, $|\mathcal{O}| = 1$, (13) becomes

\[
(\delta_i - \hat{q}_i)(\sigma - \sum_{i' \in \mathcal{I}} \hat{q}_{i'}) = 0, \ i \in \mathcal{I}
\]

\[
\sum_{i \in \mathcal{I}} \hat{q}_i \leq \sigma
\]

\[
0 \leq \hat{q}_i \leq \delta_i, \ i \in \mathcal{I}.
\] (27)

From the equalities, we have that either $\sum_{i \in \mathcal{I}} \hat{q}_i = \sigma$ or $\hat{q}_i = \delta_i$ for all $i$. If $\sigma > \sum_{i \in \mathcal{I}} \delta_i$, the former case is ruled out since this would result in a violation of at least one demand constraint. Therefore,

\[
\sum_{i \in \mathcal{I}} \hat{q}_i = \min \left\{ \sigma, \sum_{i \in \mathcal{I}} \delta_i \right\}.
\] (28)

If the minimum is $\sum_{i \in \mathcal{I}} \delta_i$, then $\hat{q}_i = \delta_i$ for all $i$ is a unique solution to (27). On the other hand, if the minimum is $\sigma$, then any solution that satisfies $\sum_{i \in \mathcal{I}} \hat{q}_i = \sigma$ and $0 \leq \hat{q}_i \leq \delta_i$ for all $i$ is a HFS.
A flow maximizing solution, denoted $q^*$, for the case of a merge node is one that solves

$$\max_{\{q_i\}_{i\in\mathcal{I}}} \sum_{i\in\mathcal{I}} q_i$$

s.t. $\sum_{i\in\mathcal{I}} q_i \leq \sigma$

$$0 \leq q_i \leq \delta_i, \ i \in \mathcal{I}. \quad (29)$$

Clearly, (28) holds for $q^*$ and in the case where the minimum is $\sigma$, we also have that any solution that satisfies $\sum_{i\in\mathcal{I}} q_i^* = \sigma$ and $0 \leq q_i^* \leq \delta_i$ is flow maximizing. Similarly, in case the minimum is $\sum_{i\in\mathcal{I}} \delta_i$, the solution is uniquely given by $q_i^* = \delta_i$ for all $i$.

The discussion above can be formally summarized by the following corollary.

**Corollary 1** (Necessity of flow maximizing solutions). Suppose that either $|\mathcal{I}| = 1$, $|\mathcal{O}| = 1$, or both. Then a set of node flows is a HFS if and only if it is flow maximizing.

**Remark 2.** Any algorithm that produces holding free solutions will produce the same (unique) solution in the case of a diverge node. This immediately follows from the fact that (25) is a unique HFS.

The analysis above highlights the fact that flow maximizing solutions (hence holding-free solutions) are not unique in some cases, e.g., the case of a simple merge node. Additional conditions are required in order to pick out a unique solution to the problem. In the case of a merge node, there exists no unique HFS and existing solution techniques may produce different solutions. Consider the example in Fig. 4. Using the algorithm in (Ni and Leonard, 2005) we get the solution $q_1 = 1600$ veh/hr, $q_2 = 1400$ veh/hr. Using the approach in (Daganzo, 1995), we have an entire family of solutions $q_1 = \text{mid}\{2100,1600,3000p_1\}$ and $q_2 = \text{mid}\{1400,900,3000p_2\}$, where $p_1, p_2$ are priorities such that $p_1 + p_2 = 1$. In other words, one requires (additional) knowledge regarding the priorities in order for the solutions to be unique in this case. In practice, such information is not readily available for the case of general urban intersections. In Section 5, we present a greedy algorithm that produces holding-free solutions, which are unique given a binary priority scheme. In Section 6, we demonstrate how knowledge of signal phasing schemes results in unique priorities.

5. The invariance principle

Unlike conditions 1-4, the invariance principle (IP) pertains to the dynamics. Generally speaking, it is a condition that relates to flows at different time instants. Consequently, strictly speaking, static node models such as the majority of the models cited above cannot be extended to accomodate the IP. In other words, let $q(t)$ be a node flow solution for the boundary conditions at time $t$ that violates the IP. The undesirable consequences of violating the IP come into being at some time $t + \epsilon$, where $\epsilon > 0$ can be very small.
To illustrate the possible consequences of violating the IP, suppose a node flow solution at time $t$ is such that $q_i(u_i, t) < \Delta_i(\rho_i(u_i, t))$. From (3), we see in this case that $\rho_i(u_i, t^+) = \Sigma_i^{-1}(q_i(u_i, t))$ is supercritical. Consequently, the demand at the downstream end of link $i$ at time $t^+$ is given by $\Delta_i(\rho_i(u_i, t^+)) = q_i^{\text{max}}$. In other words, since some of the demand in the downstream of link $i$ is not satisfied (according to the node flow solution prescribed), congestion sets in and, all else held constant, this should trigger a shockwave that propagates in the upstream direction along link $i$. These dynamics arise as a consequence of the prescribed node flow solution. Now, if under this new system state at time $t^+$, the node flow solution approach used can produce $q_i(u_i, t^+) > q_i(u_i, t)$, this can trigger a downstream moving shockwave along link $i$. (See the example in (Lebacque and Khoshyaran, 2005) for a numerical illustration.) Note that this occurs on a very small time scale. In other words, an IP violating node flow solution can produce nonsensical wave propagation behavior, such as an upstream moving shockwave followed instantaneously by a downstream moving shockwave.

To avoid the type of behavior described above, the IP states that node flows should be invariant to increases in supply when constrained by demand and vice versa. For instance, suppose $q = [q_1 \cdots q_{|I|}]^\top$ solves either (13) or (17) and suppose, for some $i$, that $q_i = \delta_i$, then increasing $\sigma_o$ for any $o$ for which $f_{i,o} > 0$ should have no impact on $q_i$. Note that the invariance property pertains to the solution technique and not the solution itself. That is, given a node flow solution, it is not possible to determine whether the solution is invariant or not. Solution formulas that do satisfy the IP will be referred to as invariant holding-free solution (I-HFS) algorithms. These are defined formally below.

**Definition 2** (Invariant holding-free solutions). Let $A(\delta, \sigma)$ be an algorithm that calculates holding-free solutions and let $q$ be a HFS calculated using $A$. Define $\delta \equiv [\delta_1 \cdots \delta_{|I|}]^\top$ and $\sigma \equiv [\sigma_1 \cdots \sigma_{|O|}]^\top$, where $\delta_i = q_i^{\text{max}}$ for each $i \in I$ such that $q_i < \delta_i$, $\sigma_o = q_o^{\text{max}}$ for each $o \in O$ such that $q_o < \sigma_o$, and $\delta_i = \sigma_o = \sigma_o$ otherwise. If $q = A(\delta, \sigma) = A(\delta, \sigma)$, then $A$ is an invariant holding-free solution (I-HFS) algorithm.

One example of a solution approach that is not an I-HFS scheme is the formula (20). To demonstrate this, suppose

$$\min_{o' \in O} \left\{ \frac{\sigma_{o'}}{\sum_{i' \in I} f_{i', o'} \delta_{i'}} \right\} < 1,$$

(30)

then $q_i < \delta_i$ for all $i$ and any increase in any of the $\delta_i$’s will impact the solution, since the solution still depends on the $\delta_i$’s. Examples of I-HFS algorithms, for general nodes, include those proposed by Tampère et al. (2011); Flötteröd and Rohde (2011) and (Bliemer et al., 2014, Appendix A).

**Remark 3.** Whenever the optimal solution to (17) is unique (when the objective function is not parallel to any of the left hand sides of the supply constraints), any algorithm that can produce optimal linear programming solutions is invariant. That is, the case should be clear since the demands and supplies that would change ($\delta$ and $\sigma$ in **Definition 2**).
correspond to inactive constraints at the optimal solution. Increasing the right hand sides for these constraints has no impact on the solution.

Algorithm 1 below runs $|\mathcal{I}|$ iterations and in each iteration it performs $|\mathcal{O}|$ updates. This gives an upper bound complexity of $O(|\mathcal{I}| \cdot |\mathcal{O}|)$, which is reasonable considering the small sizes of $\mathcal{I}$ and $\mathcal{O}$ in real-world settings. It is quite possible to terminate the algorithm early, thereby reducing computation times even further. For instance, if after iteration $k$, the residual supplies $\sigma^k_o = 0$ for all $o$, the algorithm may be terminated and $q_j = 0$ for all $j > k$. In this paper, we will not pursue algorithmic complexity issues any further as this is not the main focus of the paper.

**Algorithm 1: An I-HFS algorithm**

1: Initialize:
2: $\mathcal{P} \equiv (p_1, \ldots, p_{|\mathcal{I}|})$ ← a permutation of $\mathcal{I}$
3: $k \leftarrow 1$
4: $\sigma^k_o \leftarrow \sigma_o$ for each $o \in \mathcal{O}$
5: Iterate:
6: for $k \leq |\mathcal{I}|$ do
7: $q_{p_k} \leftarrow \min \left\{ \delta_{p_k}, \min_{o' \in \mathcal{O}, f_{p_k,o'} > 0} \frac{\sigma^k_{o'}}{f_{p_k,o'}} \right\}$
8: $\sigma^{k+1}_o \leftarrow \sigma^k_o - f_{k,o}q_{p_k}$ for each $o \in \mathcal{O}$
9: $k \leftarrow k + 1$
10: end for

We will next turn to formally establishing that Algorithm 1 is indeed an I-HFS algorithm.

**Theorem 3.** Algorithm 1 is an I-HFS algorithm for any permutation, $\mathcal{P}$, of $\mathcal{I}$.

**Proof.** We first show that the algorithm produces holding-free solutions: That $0 \leq q_{p_k} \leq \delta_{p_k}$ is trivial and since the set $\mathcal{P} = (p_1, \ldots, p_{|\mathcal{I}|})^8$ is in one-to-one correspondence with $\mathcal{I}$, we have that all demand constraints are satisfied. Now suppose for some $p_k$ that

$$\min_{o' \in \mathcal{O}, f_{p_k,o'} > 0} \frac{\sigma^k_{o'}}{f_{p_k,o'}} < \delta_{p_k}$$

and let $o$ be the index of any outbound link for which the minimum is not attained. That is,

$$\frac{\sigma^k_o}{f_{p_k,o}} > \min_{o' \in \mathcal{O}, f_{p_k,o'} > 0} \frac{\sigma^k_{o'}}{f_{p_k,o'}}.$$  

---

8The use of round brackets as opposed to the curly brackets used to list the elements of $\mathcal{I}$ is to emphasize that the ordering of the elements in $\mathcal{P}$ is important.
$\sigma^k_o$ can be interpreted as the residual supply of outbound link $o$ after the demands on links $p_1, \ldots, p_k$ have been assigned. Noting that (32) holds only if $\sigma^k_o > 0$ and $f_{p_k,o} > 0$, we have that the residual supply of $o$ after assigning the demands of link $p_k$ is

$$\sigma^{k+1}_o = \sigma^k_o - f_{p_k,o} \min_{o' \in O} \frac{\sigma^k_o}{f_{p_k,o'}} > \sigma^k_o - f_{p_k,o} \frac{\sigma^k_o}{f_{p_k,o}} = 0.$$ (33)

Now let $o$ be the index of an outbound link for which the minimum is attained. Note that this implies that $f_{p_k,o} > 0$. If $\sigma^k_o = 0$, then $q_{p_k} = 0$; in this case all of the supply of link $o$ has been exhausted by inbound links $p_j$ where $j < k$ and no additional demand can utilize link $o$. If, on the other hand, $\sigma^k_o > 0$, then

$$\sigma^{k+1}_o = \sigma^k_o - f_{p_k,o} q_{p_k} = \sigma^k_o - f_{p_k,o} \frac{\sigma^k_o}{f_{p_k,o}} = 0$$ (34)

since $f_{p_k,o} > 0$. In this case, there is no supply along $o$ that can be utilized by any inbound link $p_j$ with $j > k$. Hence, Algorithm 1 produces solutions that satisfy the supply constraints.

As a consequence, since (31) implies that there exists $o \in O$ for which either $\sigma^k_o = 0$ or $\sigma^{k+1}_o = 0$ with $f_{p_k,o} > 0$, we have that (31) also implies that $\sigma^k_o - q_k = 0$. On the other hand, if

$$\min_{o' \in O} \frac{\sigma^k_{o'}}{f_{p_k,o'}} \geq \delta_{p_k}$$ (35)

we have that $q_{p_k} - \delta_{p_k} = 0$. Hence, Algorithm 1 produces holding-free solutions. That the algorithm produces invariant solutions is straightforward: for any $p_k$, suppose (31) holds, then

$$\min \left\{ q_{p_k} \max_{o'} \frac{\sigma^k_{o'}}{f_{p_k,o'}}, \min_{o' \in O} \frac{\sigma^k_{o'}}{f_{p_k,o'}} \right\} = \min \left\{ \delta_{p_k} \max_{o'} \frac{\sigma^k_{o'}}{f_{p_k,o'}}, \min_{o' \in O} \frac{\sigma^k_{o'}}{f_{p_k,o'}} \right\} = q_{p_k}$$ (36)

since $\delta_{p_k} \leq q_{p_k}$. Hence the solution is invariant under increases in demand when restricted by the supplies. On the other hand, assume the solution is demand constrained and let $O' \subset O$ denote the set of outbound links with non-zero residual supplies. That is, for each $o \in O'$, we have that $\sigma^{|I|+1}_o > 0$. This only holds when (32) holds for $k = |I|$, and hence for all $k$ as well, by induction. Consequently, increasing $\sigma_o$ to $q_{o}^\text{max}$ for all $o \in O'$ does not impact the solution since (32) continues to hold and the right hand side remains unchanged.

Finally, since $P$ is arbitrary, these properties hold for any permutation of $I$. \hfill \square

Theorem 3 states that any permutation can be used and Algorithm 1 remains invariant. However, changing the permutation can change the solution dramatically. This is much like how different priorities effect the solution of any node model that requires knowledge
of priorities, e.g., Daganzo (1995). Different permutations can be interpreted as different (binary) priority rules. It is possible to explicitly relate the (binary) permutation to a priority rule. For instance, consider a merge and a (binary) permutation that gives priority to links with higher capacity. Take the algorithm proposed in (Ni and Leonard, 2005) and replace the fairness ratio calculated in step 1 of their algorithm with an indicator that link $i$ has the greatest capacity amongst inbound links under consideration (i.e., let $r_i$ denote the fair share of link $i$, then under a binary rule, $r_i = 1$ if link $i$ has the greatest capacity and $r_i = 0$, otherwise).

To illustrate the differences obtained using different permutations and how they relate to different priority schemes, consider the example given at the end of Subsection 4.1 for the network depicted in Fig. 4. We have two possible permutations in this case $P_1 = (1,2)$, i.e., $p_1 = 1, p_2 = 2$, and $P_2 = (2,1)$, i.e., $p_1 = 2, p_2 = 1$. For $P_1$ the algorithm proceeds as follows:

- $k = 1, \sigma_0^1 = \sigma_0 = 3000$:
  - $q_{p_1} = q_1 = \min\{2100, 3000\} = 2100$,
  - $\sigma_0^2 = \sigma_0^1 - q_{p_1} = 3000 - 2100 = 900$.

- $k = 2, \sigma_0^2 = 900$:
  - $q_{p_2} = q_2 = \min\{1400, 900\} = 900$,
  - $\sigma_0^3 = \sigma_0^2 - q_{p_2} = 900 - 900 = 0$.

It can be easily seen that the solution $q_1 = 2100 \text{ veh/hr}, q_2 = 900 \text{ veh/hr}$ is a HFS. It is also easy to demonstrate that increasing $\delta_2 = 1400 \text{ veh/hr}$ to $q_2^{\text{max}} = 2200 \text{ veh/hr}$ would not impact the solution. Hence, it is invariant. This same solution would be obtained using the formula given in (Daganzo, 1995) and the algorithm given in (Flötteröd and Rohde, 2011) for any priority $\frac{q_1}{q_2} \geq \frac{0.7}{0.3}$. Such a priority rule can arise in a broad set of situations in practice.

Now consider the permutation $P_2$. We get:

- $k = 1, \sigma_0^1 = \sigma_0 = 3000$:
  - $q_{p_1} = q_2 = \min\{1400, 3000\} = 1400$,
  - $\sigma_0^2 = \sigma_0^1 - q_{p_1} = 3000 - 2100 = 1600$.

- $k = 2, \sigma_0^2 = 1600$:
  - $q_{p_2} = q_1 = \min\{2100, 1600\} = 1600$,
  - $\sigma_0^3 = \sigma_0^2 - q_{p_2} = 1600 - 1600 = 0$. 

18
Like the solution obtained with $\mathcal{P}_1$, it can be easily seen that the solution $q_1 = 1600$ veh/hr, $q_2 = 1400$ veh/hr is a HFS. Furthermore, increasing $\delta_1 = 2100$ veh/hr to $q_1^{\text{max}} = 2200$ veh/hr would have no impact on the solution. This same solution would be obtained using the formula given in (Daganzo, 1995) and the algorithm given in (Flötteröd and Rohde, 2011) for any priority \( \frac{n_1}{n_2} \leq \frac{0.53}{0.47} \). It is also the solution that would be obtained using the algorithm provided by (Ni and Leonard, 2005); i.e., a capacity based splitting scheme. Such a priority rule can arise in a broad set of situations in practice. Also note that, consistent with Corollary 1, both permutations produce solutions that are flow maximizing.

6. Non-simultaneity of conflicting flows

The node models (13) and (17) allow for conflicting movements to proceed through the node simultaneously, which is unrealistic. In reality, at any time instant, only a subset of non-conflicting movements proceed through the intersection simultaneously. This results in a natural decomposition of the movements at the intersection into “phases” of protected and permitted intersection movements. The time scales over which a single vehicle in each movement of a phase completes its maneuver is long enough to warrant considering only subsets of movements at each time instant. In this paper we will only consider the case of signalized intersections, so that the time scales over which non-conflicting movements complete their maneuvers correspond to signal phases.

In addition to allowing vehicles to cross paths, thereby overestimating the capacity of the intersection when there are no downstream supply restrictions, another physically adverse consequence of simultaneously allowing conflicting flows through a general intersection is the blocking effect. Suppose $\sigma_o = 0$ for some outbound link $o$. Then in both (13) and (17), we have that $q_i = 0$ for all $i$ with $f_{i,o} > 0$. That is, in this case vehicles not destined to $o$ are always blocked by those that are. This is only reasonable in very specific circumstances (e.g., a pure merge nodes with spill-over from the mainline) but not general network intersections.

For the case of a signalized intersection, the phasing scheme and the phase lengths are dictated by the timing plans. For an unsignalized intersection, the phasing is self-organized by the drivers. In this case, the phasing schemes may not be well defined and could vary with time and driver aggressiveness, levels of compliance with the conventional rule, and so forth. These details are delicate and difficult to handle in a deterministic context; we leave the subject of phasing at unsignalized intersections to future research. In either case, in incident-free traffic, conflicting movements never proceed simultaneously. In this paper, we will assume that the phasing dynamics are known a priori\(^9\). An example of a typical phasing diagram is given in Fig. 5. In this example, the signal cycle consists of five phases. During Phase $p$, only the movements that are active during that phase are allowed. For example, during Phase 2, for all $t \in [t_1, t_2)$ west-

---

\(^9\)In recent years such knowledge has been made possible by intersection monitoring technologies such as the SMART-Signal system (Ma, 2008).
bound through, westbound right, eastbound through, and eastbound right are allowed. All other movements are not.

Fig. 5: Example phasing diagram for a signal cycle. Phase $p$ is active during the time interval $[t_{p-1}, t_p)$

A reasonable phasing scheme will have the following properties:

1. Merging and crossing conflicts between movements, whenever they are allowed occur between a pair of movements: a protected movement and a permitted movement.

2. Movements that form a diverging conflict shall not block each other.

There are many ways to stage the movements at a node in ways that respect point 1 above. Phasing schemes at signalized intersections typically do. To give some examples, consider the two-way four-leg intersection depicted in Fig. 6 and assume the movements across the intersection consist of through, left, and right turns for all inbound flows (i.e., we have 12 movements).

Fig. 6: Example two-way four leg intersection

Four phases are depicted in Fig. 7. Phase (a) consists of both protected movements (solid colors) and permitted movements (shaded colors). Phases (a), (b), and (c) all respect point 1. Phase (d) does not: it allows two crossing conflicts and two merging conflicts between protected movements to proceed simultaneously, the two left-turns with the opposing through movements and with the opposing right turns, respectively.

In the case of crossing conflicts, such as the permitted left and the protected through in Fig. 7(a), capturing the priority scheme may require the introduction of an *internal node*. 

20
This is needed since crossing conflicts do not compete for supply at the outbound links. When crossing conflicts are not allowed, this is not an issue.

The advantage of staging the movements is that the node models become very easy to solve for each phase (as long as the phasing scheme prohibits conflicts). This should be easy to see: Since conflicts do not take place between protected movements, we do not have any merging conflicts. This, combined with conflicts taking place between pairs of permitted and protected movements, makes determining movement priority levels a trivial matter. In essence, the node model is decomposed into a set of diverge node models, which can be solved explicitly.

While it is easy to see how the node models can be solved, how to translate these solutions into boundary conditions for the links becomes more subtle. There are two issues than need to be addressed in this context: First, the splits, \( f_{i,o} \) are no longer constant over the course of a “cycle”; some fractions of the demands get depleted while others do not during the course of a phase. Second, the dynamics of movements that are part of a phase at time \( t \) are different from those that are not. For instance, consider the dual left phase in Fig. 7(b); during this phase the through movement demands should be at capacity even if those portions of demand are small, while the left turn demands should only be at capacity if there are not sufficient supplies in the downstream even if those portions of demand are large (but below capacity).

These subtleties dictate a multi-commodity approach to dealing with the boundary flows. Due to the staging, the usual treatment that hinges on all “commodities” having the same speed is not applicable. Instead, it will be assumed that each movement has it’s own fundamental relation, \( Q_{(i,o)} \), but that this decomposition only applies to the inbound flows. That is, decomposing the traffic densities at the downstream boundaries of the inbound links (by outbound movement): \( \rho_i(u_i,t) = \sum_{o \in O} \rho_{(i,o)}(u_i,t) \), the following holds\(^{10}\)

\[^{10}\]The dependence on \( (u_i,t) \) has been suppressed in (37). It should be clear that this only applies to the inbound links. i.e., \( Q_o(\rho_o) \neq \sum_{i \in I} Q_{(i,o)}(\rho_{(i,o)}) \) is allowed.
\[ Q_i(\rho_t) = \sum_{o \in O} Q_{(i,o)}(\rho_{(i,o)}). \] (37)

The movement-specific fundamental relations give rise to movement-specific demand functions, \( \Delta_{(i,o)}(\rho_{(i,o)}) \). This will allow for modeling different traffic characteristics amongst the different movements emanating from each link \( i \in \mathcal{L} \) during each phase. The boundary movement-specific densities are given by
\[
\rho_{(i,o)}(u_{i,t}+) = \begin{cases} 
\Sigma_{(i,o)}^{-1}(q_{(i,o)}(u_{i,t})) & \text{if } q_{(i,o)}(u_{i,t}) < \Delta_{(i,o)}(\rho_{(i,o)}(u_{i,t})), \\
\rho_{(i,o)}(u_{i,t}) & \text{if } q_{(i,o)}(u_{i,t}) \geq \Delta_{(i,o)}(\rho_{(i,o)}(u_{i,t})),
\end{cases}
\] (38)

where \( q_{(i,o)} \) are the prescribed individual movement flows. In essence, each movement-specific flow is treated as an independent inbound flow. This, coupled with simultaneous flows in a phase not competing for the same supplies, has the advantage that blocking does not take place. The traffic densities at the upstream boundaries of the outbound links are calculated in accordance with (2) after aggregating the appropriate node flows: \( q_o = \sum_{i \in \mathcal{I}} q_{(i,o)}, \) where \( q_{i,o} = 0 \) for \( (i,o) \) movements that are not part of the active phase.

Let \( \mathcal{M}_n(t) \subseteq \mathcal{I} \times \mathcal{O} \) denote the set of movements across node \( n \) that are active at time \( t \). The set \( \mathcal{M}_n(t) \) can be partitioned into two disjoint sets of movements, a set of permitted movements \( \mathcal{M}_n^{pe}(t) \) and a set of protected movements \( \mathcal{M}_n^{pr}(t) \) with \( \mathcal{M}_n(t) = \mathcal{M}_n^{pe}(t) \sqcup \mathcal{M}_n^{pr}(t) \). A phase can consist of protected movements only, but not only of permitted movements: \( \mathcal{M}_n^{pe}(t) = \emptyset \) is allowed, while \( \mathcal{M}_n^{pr}(t) \neq \emptyset \). The movements in a phase are assigned one of two priority levels depending on whether they are protected or permitted. Algorithm 1 then simplifies to a two-step algorithm: all protected movement flows are calculated in parallel (since they do not interact) and then whatever remaining supply is allocated to the permitted movements. The solution is given uniquely by the explicit solution formula\(^{11}\):

\[
q_{(i,o)} = \begin{cases} 
\min \left\{ \delta_{(i,o)}, \sigma_0 \right\} & \text{if } (i,o) \in \mathcal{M}_n^{pr}(t) \\
\min \left\{ \delta_{(i,o)}, \sigma_0 - \Sigma_{(i',o) \in \mathcal{M}_n^{pr}(t)} \min \left\{ \delta_{(i',o)}, \sigma_0 \right\} \right\} & \text{if } (i,o) \in \mathcal{M}_n^{pe}(t).
\end{cases}
\] (39)

Since (39) is a specialization of Algorithm 1, it is an I-HFS formula. To demonstrate this, let \( \mathcal{I}_n^a(t) \subseteq \mathcal{I}_n \) denote the set of inbound links to node \( n \) which are active at time \( t \). That is, for each \( i \in \mathcal{I}_n^a(t) \), there exists at least one active movement \( (i,o) \in \mathcal{M}_n(t) \) at time \( t \). Similarly, let \( \mathcal{O}_n^a(t) \subseteq \mathcal{O}_n \) denote the set of active outbound links. The staged flows problem is then written as
\[
(\delta_i - q_i) \prod_{o \in \mathcal{O}_n^a(t) : \delta_{(i,o)} \geq 0} (\sigma_0 - q_o) = 0, \quad i \in \mathcal{I}_n^a(t)
\] 0 \leq q_o \leq \sigma_0, \quad o \in \mathcal{O}_n^a(t)
\] 0 \leq q_i \leq \delta_i, \quad i \in \mathcal{I}_n^a(t). \] (40)

\(^{11}\)Note that since merging conflicts between protected movements is not allowed, the sum in the second part of (39) is over a single \((i',o)\).
First note that for each \( o \in O^a_n(t) \), \( q_o = \sum_{i \in I^a_n(t)} q_{(i,o)} \). Since conflicting movements are not allowed, this sum is over at most two movements, one protected and one permitted. Denote these respectively by \((i_1,o)\) and \((i_2,o)\). From (39), we have that

\[
q_o = q_{(i_1,o)} + q_{(i_2,o)} \leq q_{(i_1,o)} + (\sigma_o - q_{(i_1,o)}) = \sigma_o.
\]  

(41)

Similarly, \( q_i = \sum_{o \in O^a_n(t)} q_{(i,o)} \leq \sum_{o \in O^a_n(t)} \delta_{(i,o)} = \delta_i \). Hence, (39) satisfies the local supply and local demand constraints. Next, if for any \((i,o) \in M^p_n(t)\), \( \min\{\delta_{(i,o)}, \sigma_o\} = \sigma_o \), then \( q_o = \sigma_o \). If, on the other hand, \( \min\{\delta_{(i,o)}, \sigma_o\} = \delta_{(i,o)} \) for all \( o \), then \( q_i = \sum_{o \in O^a_n(t)} q_{(i,o)} = \sum_{o \in O^a_n(t)} \delta_{(i,o)} = \delta_i \). Hence, the complementarity condition also holds. This proves that (39) only produces holding free solutions. That it is invariant is trivial. Hence, it is a I-HFS solution formula. Moreover, since the movements in a phase do not interact, (39) is also flow maximizing (in accordance with Corollary 1).

### 7. Numerical examples

#### 7.1. Four-leg intersection example

Consider the two-way four-leg intersection depicted in Fig. 8 adopted from (Tampère et al., 2011). For this node, \( I = \{1,2,3,4\} \) and \( O = \{5,6,7,8\} \). The demands and supplies at some time instant \( t \), shown in Fig. 8, are all in units of vehicles per hour. The demand distribution (at time instant \( t \)) is given in Table 1 and the capacities are \( q^\text{max}_1 = q^\text{max}_5 = 1000 \) veh/hr, \( q^\text{max}_2 = q^\text{max}_6 = 2000 \) veh/hr, \( q^\text{max}_3 = q^\text{max}_7 = 1000 \) veh/hr, and \( q^\text{max}_4 = q^\text{max}_8 = 2000 \) veh/hr.

![Fig. 8: Example two-way four-leg intersection with supplies and demands at time t](image)

From Table 1, it is easy to see that none of the supply constraints in (17) are parallel to the objective function. Therefore, the maximum flow solution to the problem is unique. This also implies that the maximum flow solution is invariant. To see this, we first use the simplex algorithm to solve (17) to optimality. The optimal solution is \( q^*_1(t) = 0 \) veh/hr, \( q^*_2(t) = 1750 \) veh/hr, \( q^*_3(t) = 800 \) veh/hr, and \( q^*_4(t) = 1567.19 \) veh/hr.
The objective value at optimality is $\sum_{i \in I} q_i^*(t) = 4117.19$ veh/hr. First note that the objective value is greater than that reported in (Tampère et al., 2011, Table 8), so their algorithm is clearly not flow maximizing. Furthermore, (Tampère et al., 2011, Table 2) provides the solution obtained using the formula (20), which is also not flow maximizing: $\sum_{i \in I} \hat{q}_i(t) = 4000 < \sum_{i \in I} q_i^*(t) = 4117.19$. Similarly, applying the algorithm given in (Gibb, 2011) results in the following solution $q(t) = [385.99 \ 1576 \ 676.81 \ 1376.60]$ $^\top$, which has a total flow of $4015.4 < 4117.19$.

Under the solution $q^*(t)$, the inbound links with unsatisfied demands immediately enter a congested regime at the link boundaries. The demand at downstream boundaries of link 1 changes to $\delta_1(t+) = \Delta_1(\Sigma_1^{-1}(q_1^*(t))) = 1000$ veh/hr and the demand at the downstream boundary of link 4 changes to $\delta_4(t+) = \Delta_4(\Sigma_4^{-1}(q_4^*(t))) = 2000$ veh/hr (in accordance with (3)). This change occurs instantaneously. Solving the max flow problem again (using the simplex algorithm) with these new right hand sides, the solution obtained remains unchanged as expected, that is, we have that $q^*(t+) = q^*(t)$. Thus, the solution is invariant.

We now apply Algorithm 1. Consider the permutation $\mathcal{P} = (1,2,3,4)$. The steps of the algorithm for the traffic state at time $t$ are listed below.

- $k = 1, \sigma^1 = \sigma = [1000 \ 2000 \ 1000 \ 2000]^\top$:
  - $q_1 = \min\{500,\infty,20000,3333.3,3333.3\} = 500$,
  - $\sigma^2 = \sigma^1 - 500 \cdot [0.1 \ 0.3 \ 0.6]^\top$.

- $k = 2, \sigma^2 = [1000 \ 1950 \ 850 \ 1700]^\top$:
  - $q_2 = \min\{2000,200000,\infty,5666.7,2125\} = 2000$,
  - $\sigma^3 = \sigma^2 - 2000 \cdot [0.05 \ 0.15 \ 0.8]^\top$.

- $k = 3, \sigma^3 = [900 \ 1950 \ 550 \ 100]^\top$:
  - $q_3 = \min\{800,7200,15600,\infty,133.3\} = 133.3$,
  - $\sigma^4 = \sigma^3 - 133.3 \cdot [0.125 \ 0.125 \ 0.75]^\top$.

- $k = 4, \sigma^4 = [883.3 \ 1933.3 \ 550 \ 0]^\top$:
  - $q_4 = \min\{1700,15016.7,4108.3,1168.8,\infty\} = 1168.8$. 

<table>
<thead>
<tr>
<th>$f_{i,o}$</th>
<th>$o = 5$</th>
<th>$o = 6$</th>
<th>$o = 7$</th>
<th>$o = 8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i = 1$</td>
<td>0</td>
<td>0.1</td>
<td>0.3</td>
<td>0.6</td>
</tr>
<tr>
<td>$i = 2$</td>
<td>0.05</td>
<td>0</td>
<td>0.15</td>
<td>0.8</td>
</tr>
<tr>
<td>$i = 3$</td>
<td>0.125</td>
<td>0.125</td>
<td>0</td>
<td>0.75</td>
</tr>
<tr>
<td>$i = 4$</td>
<td>1/17</td>
<td>8/17</td>
<td>8/17</td>
<td>0</td>
</tr>
</tbody>
</table>
\(- \sigma^5 = \sigma^4 - 1168.8 \cdot [1/17\ 8/17\ 8/17\ 0]^{\top} = [814.6\ 1383.3\ 0\ 0]^{\top}.\)

It is easy to check that the solution \(q(t) = [500\ 2000\ 133.3\ 1168.8]^{\top}\) is a HFS. Notice too that \(q(t)\) is an extreme point solution (i.e., it is basic feasible), but it is not flow maximizing. Under this solution, we have that \(\delta_3(t+) = \Delta_3(\Sigma_3^{-1}(q(t))) = 1000\) veh/hr and \(\delta_4(t+) = \Delta_4(\Sigma_4^{-1}(q_4(t))) = 2000\) veh/hr. The steps of Algorithm 1 with the new supplies at time \(t+\) proceed as follows (\(k = 1\) and \(k = 2\) are the same as above and are omitted below; the new data are highlighted in red):

- \(k = 3, \sigma^3 = [900\ 1950\ 550\ 100]^{\top}:\)
  - \(q_3 = \min\{1000, 7200, 15600, \infty, 133.3\} = 133.3,\)
  - \(\sigma^4 = \sigma^3 - 133.3 \cdot [0.125\ 0.125\ 0\ 0.75]^{\top}.\)

- \(k = 4, \sigma^4 = [883.3\ 1933.3\ 550\ 0]^{\top}:\)
  - \(q_4 = \min\{2000, 15016.7, 4108.3, 1168.8, \infty\} = 1168.8,\)
  - \(\sigma^5 = \sigma^4 - 1168.8 \cdot [1/17\ 8/17\ 8/17\ 0]^{\top} = [814.6\ 1383.3\ 0\ 0]^{\top}.\)

We find that \(q(t+)=q(t),\) demonstrating that Algorithm 1 is an I-HFS algorithm.

### 7.2. Two-way major and one-way minor

The example we present next was adopted from (Flötteröd and Rohde, 2011). A schematic of the node along with the traffic state at time \(t\) are depicted in Fig. 9. The east-to-west direction is a minor one-way road and the north-south direction is a two-way major road. For this node, \(I = \{i_E, i_N, i_S\}\) and \(O = \{o_N, o_W, o_S\}\) (the chosen labeling is meant to resemble that in the original paper). The demand distribution fractions are given in Table 2. In this case, the maximum flow problem (17) has multiple solutions:

The two extreme points \(q^{*,(1)}(t) = [0\ 600\ 200]^{\top}\) and \(q^{*,(2)}(t) = [0\ 200\ 600]^{\top}\) and any
Table 2: Demand distribution for the example in Fig. 9

<table>
<thead>
<tr>
<th>$f_{i,0}$</th>
<th>$o_N$</th>
<th>$o_W$</th>
<th>$o_S$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i_E$</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>$i_N$</td>
<td>0</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>$i_S$</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
</tr>
</tbody>
</table>

linear combination of these two solutions is optimal. All of these solutions have an optimal objective value of $\sum_{i \in I} q^*_i(t) = 800$ veh/hr. The solution to the same problem given in (Flötteröd and Rohde, 2011) using their incremental node model algorithm is $q_{i_E}(t) = 16.7$ veh/hr, $q_{i_N}(t) = 600$ veh/hr, and $q_{i_S}(t) = 166.7$ veh/hr. The objective value for this solution is $\sum_{i \in I} q_i(t) = 783.3$ veh/hr, i.e, their algorithm produces suboptimal solutions.

More importantly, it is easy to see that $q^*_{(1)}(t)$ leads to dramatically different dynamics than does $q^*_{(2)}(t)$ at time $t^+$. The former leads to an upstream-moving shockwave along $i_S$ and uncongested conditions along $i_N$, the latter leads to exactly the opposite behavior. This demonstrates the important role solution uniqueness has to play in whether the solution is invariant or not. Interestingly, Algorithm 1 produces the solution $q^*_{(1)}(t)$ when using the permutation $(i_N, i_S, i_E)$ and the solution $q^*_{(1)}(t)$ using the permutation $(i_S, i_N, i_E)$, both of which favor the major approaches over the minor approach.

7.3. A traffic signal

Consider the single intersection network in Fig. 10. Assume free-flow conditions at time $t_0 = 0$ seconds. Assume there are no outflow restrictions at the downstream ends of outbound links 5, 6, 7, and 8. The fundamental relation, $Q$, is the one proposed by Del Castillo and Benitez (1995):

$$Q(\rho) = v_f \rho \left(1 - \frac{|v|}{v_f} \left(1 - \frac{\rho_{jam}}{\rho}\right)\right).$$

(42)
All single-lane links (1, 3, 5, and 7) have a free-flow speed of $v_f = 50$ km/hr, a jam density of $\rho_{jam} = 150$ veh/km, and a backward wave speed of $w = -23$ km/hr. For all two-lane links (2, 4, 6, and 8), these are given as $v_f = 50$ km/hr, $\rho_{jam} = 300$ veh/km, and $w = -23$ km/hr. The demand functions at the downstream ends of the inbound links are depicted in Fig. 11 for the single-lane; for the two-lane links, the relation is simply scaled up to twice the size (i.e., double the capacity and double the jam and critical densities). The simulation time horizon consists of 8 signal cycles, each of length 120 seconds; that is, the horizon is 960 seconds. Each of the links is 1400 meters long and the signal phasing diagram is show in Fig. 12.

The splits $\{f_{i,o}\}$ are given in Table 1. The prescribed inflows, $q_i(l_i, t)$, at the upstream boundaries of the inbound links ($i = 1, 2, 3,$ and 4) are given, respectively, by 500, 2000, 800, and 1700 veh/hr, for all $t$. These inflow rates (coupled with the phasing scheme in Fig. 12) result in queues that will slowly build up over the course of the eight cycles along the inbound links as depicted in Fig. 13. The node model used in Fig. 13 is (39). In this example, since the network boundary conditions are assumed not to change, we have steady state conditions at the downstream boundaries of the incoming links. Ignoring the phasing would result in node flows that overestimate the node’s capacity. For example, solving (17) would result in throughputs that satisfy all the demands and, therefore, fail to capture the true (cyclic) nature of queue build-up along the inbound links.
The overestimation of the capacity by a model that does not consider phasing is not a surprising finding. In fact, one can argue that suppressing the demands (or the supplies) in a way that captures the effect of a traffic signal on average can produce reasonable results (depending on application). However, one cannot avoid the blocking effect using such tricks. For the same example above, assume heavy congestion along link 5 at time $t = 0$ due to an incident at the downstream end of link 5 (i.e., $q_5(u_5,t) = 0$ for all $t$ in the study period). Then a node model that ignores the signal phasing, e.g., (17), would severely underestimate the node’s capacity. In this case, jammed traffic conditions would build up along links 2, 3, and 4 in an unrealistic way: traffic not destined to link 5 would still be blocked by traffic that is during the entire study period.

On the other hand, using (39), the effect of the blockage is captured in a more realistic manner. Link 5 is a small link and only a small portion of incoming traffic is destined to link 5. The effect of the blockage should be a commensurate increase in congestion. This is indeed the case. The dynamics using (39) are depicted in Fig. 14. Comparing the dynamics in Fig. 13 and Fig. 14, it can be seen that congestion reaches the downstream ends of link 2, 3, and 4 earlier when link 5 is blocked. Link 1 remains unaffected since no traffic along link 1 is destined to link 5.

8. Conclusion

In congested urban networks, traffic dynamics across the nodes play a prominent role in describing the network dynamics. Doing so accurately is needed to capture the rela-
Fig. 14: Traffic density dynamics along inbound links when link 5 is blocked; (a) link 1, (b) link 2, (c) link 3, (d) link 4

tionships between traffic conditions in adjacent network segments. These relationships play a critical role in a variety of applications: from estimation of traffic conditions where data may not be available to network sensor placement problems to sampling vehicle trajectories for optimal network coverage with as little data processing as possible.

In this paper, holding free solutions to node flow problems were formally defined and it was shown that flow maximization, which is a common way of preventing traffic holding, is sufficient but not a necessary condition for preventing traffic holding. This means that a max flow solution may not correctly reflect the true traffic conditions. A complementarity condition ensuring holding free node solutions was given and while the resulting model is non-linear, it was demonstrated that a simple greedy algorithm can solve it while also honoring the invariance principle.

This paper also investigated the impact of preventing conflicting movements from proceeding through the node simultaneously. This practical consideration was shown to produce very simple node flow solutions and promote solution uniqueness. The approach requires knowledge of the phasing scheme at the node; for signalized intersections, it is possible to measure the phasing scheme (in the case of actuated and adapted signals) or simply know the phasing scheme in advance (for fixed timing signals). This might be a lot more challenging for unsignalized intersections. As a topic for future research, the sensitivity of the node flows to uncertainties about the phasing schemes can be analyzed.
Acknowledgements

The author wishes to thank two anonymous referees of the paper. Their valuable comments helped in improving the quality of the manuscript.

Appendix A. HFS property of \((20)\)

It follows immediately from \((20)\) that \(0 \leq \hat{q}_i \leq \delta_i\). For each \(i \in \mathcal{I}\) define \(\mathcal{O}_1(i) \equiv \{o \in \mathcal{O} : f_{i,o} > 0\}\) and \(\mathcal{O}_2(i) \equiv \{o \in \mathcal{O} : f_{i,o} = 0\}\). Notice that for any \(i\), \(\mathcal{O}_1(i) \cup \mathcal{O}_2(i)\). Similarly, for each \(o \in \mathcal{O}\) define \(\mathcal{I}_1(o) \equiv \{i \in \mathcal{I} : f_{i,o} > 0\}\) and \(\mathcal{I}_2(o) \equiv \{i \in \mathcal{I} : f_{i,o} = 0\}\). Then, for any \(o \in \mathcal{O}\)

\[
\sum_{i \in \mathcal{I}} f_{i,o} \hat{q}_i = \sum_{i \in \mathcal{I}_1(o)} f_{i,o} \hat{q}_i \leq \sum_{i \in \mathcal{I}_1(o)} f_{i,o} \delta_i \min \left\{1, \min_{o' \in \mathcal{O}_1(i)} \left\{\frac{\sigma_{o'}}{\sum_{i' \in \mathcal{I}} f_{i',o'} \delta_{i'}}\right\}\right\}. \tag{Appendix A.1}
\]

But since \(i \in \mathcal{I}_1(o)\) for any \(o \in \mathcal{O}\) implies that \(o \in \mathcal{O}_1(i)\) we have that

\[
\sum_{i \in \mathcal{I}_1(o)} f_{i,o} \delta_i \min \left\{1, \min_{o' \in \mathcal{O}_1(i)} \left\{\frac{\sigma_{o'}}{\sum_{i' \in \mathcal{I}} f_{i',o'} \delta_{i'}}\right\}\right\} \leq \sum_{i \in \mathcal{I}_1(o)} f_{i,o} \delta_i \frac{\sigma_0}{\sum_{i' \in \mathcal{I}} f_{i',o} \delta_{i'}}, \tag{Appendix A.2}
\]

which implies that

\[
\sum_{i \in \mathcal{I}} f_{i,o} \hat{q}_i \leq \sigma_0 \tag{Appendix A.3}
\]

and hence, \((20)\) satisfies the inequalities in \((13)\).

For any \(i \in \mathcal{I}\), when

\[
\min \left\{1, \min_{o' \in \mathcal{O}_1(i)} \left\{\frac{\sigma_{o'}}{\sum_{i' \in \mathcal{I}} f_{i',o'} \delta_{i'}}\right\}\right\} = 1 \tag{Appendix A.4}
\]

the equality in \((13)\) is clearly satisfied. In this case, \(\hat{q}_i = \delta_i\). On the other hand, when

\[
\min_{o' \in \mathcal{O}_1(i)} \left\{\frac{\sigma_{o'}}{\sum_{i' \in \mathcal{I}} f_{i',o'} \delta_{i'}}\right\} < 1 \tag{Appendix A.5}
\]

there is an \(o \in \mathcal{O}\) for which \(f_{i,o} > 0\) and

\[
\hat{q}_i = \delta_i \frac{\sigma_0}{\sum_{i' \in \mathcal{I}} f_{i',o} \delta_{i'}}, \tag{Appendix A.6}
\]

so that

\[
\sum_{i \in \mathcal{I}} f_{i,o} \hat{q}_i = \sum_{i \in \mathcal{I}} f_{i,o} \delta_i \frac{\sigma_0}{\sum_{i' \in \mathcal{I}} f_{i',o} \delta_{i'}} = \sigma_0. \tag{Appendix A.7}
\]

Hence, for each \(i \in \mathcal{I}\) either \(\hat{q}_i = \delta_i\) or \(\sum_{i' \in \mathcal{I}} f_{i',o} \hat{q}_i = \sigma_0\) for some \(o\) and the equality in \((13)\) is satisfied in general. Thus, \((20)\) is a HFS.
Appendix B. Consistency of HFS with the INM of Flötteröd and Rohde

The approach proposed by Flötteröd and Rohde (2011) is one that starts with zero flows and incrementally increases the flows until no flow can proceed through the node without violating either a supply or a demand constraint. The termination condition of their algorithm is when a set, \(D\), (defined below) is empty.

Since the supply and demand restrictions are explicitly given in (13), it suffices to show that Flötteröd and Rohde’s set \(D\) is empty if and only if (10) is satisfied. Consider the vector of inbound and outbound flows \(q = [q_i \cdots q_{i|I|} q_o \cdots q_{o|O|}]^T\) and define the set valued function \(D(q) \equiv D_{\text{in}}(q) \cup D_{\text{out}}(q)\), where

\[
D_{\text{in}}(q) \equiv \{i \in I : q_i < \delta_i \text{ and } q_o < \sigma_o \text{ for all } o \in O \text{ with } f_{i,o} > 0\} \quad (\text{Appendix B.1})
\]

and

\[
D_{\text{out}}(q) \equiv \{o \in O : q_o < \sigma_o \text{ and there exists } i \in D_{\text{in}}(q) \text{ with } f_{i,o'} > 0\}. \quad (\text{Appendix B.2})
\]

Clearly, \(D_{\text{in}}(q) = \emptyset\) implies that \(D_{\text{out}}(q) = \emptyset\), so that \(D(q) = \emptyset\) if and only if \(D_{\text{in}}(q) = \emptyset\). For each \(i\), if (10) holds, then either \(q_i = \delta_i\) or \(q_o = \sigma_o\) for at least one \(o \in O\) with \(f_{i,o} > 0\). Hence, (10) implies that \(D_{\text{in}}(q) = \emptyset\).

Conversely, assume \(D_{\text{in}}(q) = \emptyset\). To show that this implies that \(q\) is a HFS, we will demonstrate that if the contrary is assumed, we arrive at a contradiction: Assume, for some \(i \in I\), that

\[
(\delta_i - q_i) \prod_{o \in O : f_{i,o} > 0} (\sigma_o - q_o) > 0. \quad (\text{Appendix B.3})
\]

Then, \(q_i < \delta_i\) and \(q_o < \sigma_o\) for all \(o \in O\) with \(f_{i,o} > 0\). Hence \(i \in D_{\text{in}}(q) \neq \emptyset\) and we have a contradiction. Thus, \(D(q) = \emptyset\) if and only if (10) holds.

References


Himpe, W., Corthout, R., Tampère, M., 2016. An efficient iterative link transmission model. Transportation Research Part B.


32


Raadsen, M., Bliemer, M., Bell, M., 2015. An efficient and exact event-based algorithm for solving simplified first order dynamic network loading problems in continuous time. Transportation Research Part B.


