A summary of transit models developed in Research Project TRA2008-06782-C02-02

E. Codina[1], L. Montero[1], A. Marín[2], F. López[1], D. Verastegui[3], M*ª*. Luz López[3]

[1] Universitat Politècnica de Catalunya
[2] Universidad Politécnica de Madrid
[3] Universidad de Castilla La Mancha

DR 2012/1

A summary of transit models developed in Research Project TRA2008-06782-C02-02[1]

E. Codina[†], L. Montero[†], A. Marín[‡], F. López[†],
D. Verastegui[♮], Mª. Luz López[♮]

[†] Universitat Politècnica de Catalunya

[‡] Universidad Politécnica de Madrid

[♮] Universidad de Castilla La Mancha

Abstract. This document describes the models for network design and planning in the field of urban transportation that have been developed in the course of research project TRA2008-06782-C02-02. This includes transit, rapid transit and bus networks. This project can be considered in many aspects as a follow up of previous research project TRA2005-09068-C03-02/MODAL. The network topology design is being considered in this project as well as logistic problems regarding vehicle assignment and timetabling. Also, the integration of different stages traditionally carried out in classical approaches of the modelization process (i.e. design and planning) is done via unified and compact models contributing in this way to the improvement in the consistency of the solutions. Some mathematical programming models are exposed in order to set the number of services at previously specified bus lines, which are intended to assist high demand occurring during the disruption of the Rapid Transit services or during the celebration of massive events. By means of this model two types of basic magnitudes can be determined, basically: a) the number of bus units assigned to each line and b) the number of services that should be assigned to those units. The model can be considered of the system optimum type, in the sense that the assignment of units and of services is carried out minimizing a linear combination of operation costs and total travel time of users. The model considers delays experienced by buses as a consequence of the get in/out of the passengers, queueing at stations and the delays that passengers experience waiting at the stations.
Keywords: Network design, public transportation models, congestion, mathematical programming.

Introduction

Because of huge costs in construction and exploitation of rapid transit systems a careful attention is paid to their efficiency. The underlying network design process is part of the planning process itself, which is comprised by several stages with two intertwined problems: alignment determination and stations location. Once the area has been divided into transportation zones, the classical four step model in planning is applied: trip generating, demand distribution, modal choice and trip assignment. The four step process for the selection of a network of lines in mass transit systems, allows us the identification of a set of potential corridors for rapid transit. These corridors are evaluated accordingly to a wealth of hierarchical factors by the decision taker. At a further stage the corridors are ordered accordingly to these factors and those selected are merged making up different network configuration which give rise to distinct scenarios. For on-line planning of transit lines the initial point is the analysis of traffic not taking into account congestion or everyday incidences that may occur, either planned or unexpected. Formal representation of the multimodal transportation problem is based in the concept of elastic

demand simultaneously applied to the intervening transport modes. It is a key hypothesis in these models that solutions verify a wardropian equilibrium that can be interpreted economically as an equilibrium between offer and demand. Another kind of mass transportation systems are public transportation systems on demand. These systems were introduced in the seventies in the United States under the name of "dial-a-ride" systems to incorporate within the existing urban public transport systems, an innovative service that adapts the new demand of transportation. While in Europe, the first experiments on transport demand systems, recognized as Demand Responsive Transport (DRT), did not show up until the eighties, and they were designed primarily to the neediest social groups: disabled and the old people. On the other hand, in case of malfunction of the transportation system, transportation authorities and operators have as a common practice to create temporary services as a recourse in order to help to the already congested transportation networks.

Setting properly the required services to attend transportation demand taking into account available resources in urban public transportation networks is a key aspect in order to keep their good performance as well as to ensure users confidence in public transportation as a valid alternative.

Models for the overall design of transit networks or simply for some management aspects of public transport lines which take into account demand in the design process, have an intrinsic relationship with passenger transit assignment models. Such assignment models can be classified in a first approach by two criterions: a) static or dynamic and b) frequency based or time table based. Within the classical passenger transit assignment models under the concept of strategy, the classical work in [15] must be cited. This initial model is unable to take into account congestion in public transportation systems. It has not been until very recently, that these strategy-based models have been able to reflect how effective frequencies may be altered by congestion ([6], [1], [10]). Frequency setting models have been formulated using transit assignment schemas based on strategies and time table based. Using a strategy based assignment model under a static approach, the works in [7] and in [13] must be taken into account. For the case of lines under strict time table, assignment models that must be cited are those implemented in the commercial package EMME and others very recently developed such as those in [9] and in [14]. In this paper two service setting models in [3] and in [4] are described which are able to reflect the effects of congestion under a static approach for public transportation lines intended for emergency situations or for supporting special events. In these two models the underlying passenger assignment schema is a non-strategy based user optimum and a congested shortest route choice, respectively. Also the formulation in variational inequalities in [2] for the congested transit assignment model in [6] is briefly described and some numerical results are presented for a variant of the model with sharp capacity constraints.

This document summarizes two main areas of models that are related intrinsically to the above described problems. After providing a common notation in section 1, in section 2 pure frequency setting models are described under two approaches: a) user-equilibrium and b) system equilibrium behavior of passengers. On both cases the effect of congestion on public transport lines is taken into account by means of the decrease of effective frequencies experimented by passengers. In the models presented in section 2, the assignment of passengers is not carried out following the concept os strategies already mentioned in this introduction. In section 4 the congested transit assignment in

[1] is summarized by using its formulation in variational inequalities as shown in [2] and the extension developed during this project to the case of elastic demand is described. Illustrative computational tests are shown for each of the two sections 2 and 4.

1 Notation and network model

In this section a unified notation is presented for all the models under discussion. The transit network is represented by means of a directed graph $G = (N, A)$, where N is the set of nodes and A is the set of links. The number of trips from i to d will be denoted by g_i^d. By $C \subset N$ it will be denoted the subset of nodes representing centroids or trip attraction/generation points. By $W = \{ (i, d) \in C \times C \mid g_i^d > 0 \}$ it is denoted the set of active origin-destination pairs $\omega = (i, d)$ on the network. The set of destinations in the network shall be denoted by $D = \{ d \in C \mid \exists (i, d) \in W \}$ and the set of origin nodes for a fixed destination $d \in D$ shall be denoted by $O(d) = \{ i \in C \mid (i, d) \in W \}$. For a node $i \in N$, the set of emerging links will be denoted by $E(i)$ and the set of incoming links by $I(i)$. The representation of transit lines will be in form of an expanded network, as in [15] (see figure 1 below).
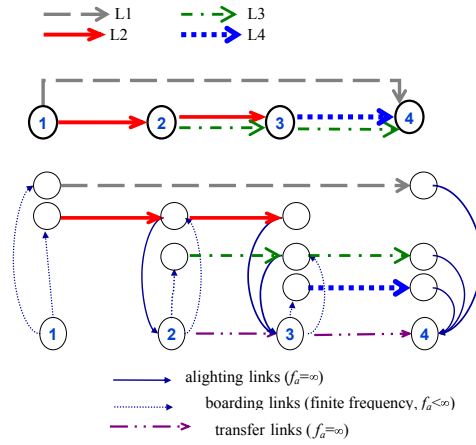

FIGURE 1. The transit expanded network assumed in the model.

By v_a^d it will be denoted the flow at link $a \in A$ with destination $d \in D$. Then the following notation will be used for the various types of vector flows and origin-destination volumes:

- $\mathrm{v}_i^d = (..., v_a^d, ...; a \in E(i)) \in \mathbb{R}_+^{|E(i)|}$, $i \in N$, $d \in D$ is the vector of flows with destination d at emerging links of node i.
- $v_i^d = \sum_{a \in E(i)} v_a^d$ is the total inflow through node $i \in N$ with destination $d \in D$.
- $\mathrm{v}^d = (..., \mathrm{v}_i^d, ...; i \in N) \in \mathbb{R}_+^{|A|}$, $d \in D$. $\mathrm{v} = (..., \mathrm{v}^d, ...; d \in D) \in \mathbb{R}_+^{|A||D|}$.
- $v = \sum_{d \in D} \mathrm{v}^d \in \mathbb{R}_+^{|A|}$. Vector of total flows on links and $v_a = \sum_{d \in D} v_a^d$, $a \in A$.
- $\mathrm{g}^d = (..., g_i^d, ...; i \in O(d)) \in \mathbb{R}_+^{|O(d)|}$, $d \in D$. $\mathrm{g} = (..., \mathrm{g}^d, ...; d \in D) \in \mathbb{R}_+^{|W|}$.

The feasibility set for the congested transit equilibrium problem can be formulated as $\text{V} = \bigotimes_{d \in D} \text{V}^d$, being each set V^d defined as:
(1)

$$\text{V}^d \triangleq \left\{ \text{v}^d \in \mathbb{R}_+^{|A|} \;\middle|\; \sum_{a \in E(i)} v_a^d - \sum_{a \in I(i)} v_a^d = g_i^d \;\; i \in N_d, \;\; \sum_{a \in I(d)} v_a^d = \sum_{i \in O(d)} g_i^d, \; v_a^d = 0, \; \forall a \in E(d) \right\}$$

The polyhedron of total link flows v is $\mathcal{V} = \{ v \in \mathbb{R}_+^{|A|} \mid v = \sum_{d \in D} \text{v}^d, \; \text{v}^d \in \text{V}^d \}$. Because of the finite capacity of vehicles, boarding of passengers may not happen at the first arriving vehicle seen by the passenger. Mean waiting times for a boarding, or inverse of effective frequencies, shall be denoted by $\sigma_a(\cdot) = 1/f_a(\cdot)$. Travel times on links are given by functions $t_a(v)$, $a \in A$ which are finite on $\mathcal{V}$. The subset of nodes for which emerging links exist with a finite efff on $\mathcal{V}$ will be denoted by $\hat{N} = \{ i \in N \mid \exists a \in E(i), \; f_a(\cdot) < +\infty \}$. The sets $\hat{N}_d = \hat{N} \setminus \{d\}$, $d \in D$ and $\hat{A} = \{ a \in A \mid \exists i \in \hat{N}, \; a \in \hat{E}(i) \}$ will be also used. For nodes $i \in N$, the subset of emerging links with finite effective frequency will be denoted by $\hat{E}(i)$. Line segments as well as pedestrian, transfer and non transit facilities shall be represented by links $a \in A$ with either constant or flow dependent travel time functions $t_a(\cdot)$ and infinite frequencies, $f_a = +\infty$. This apply also for links $a \in I(i)$, $i \in \hat{N}$, representing alighting at stops.

2 Frequency setting models

2.1 A user equilibrium based service setting model

The first model by Codina and Marín [**3**], [model SUE] below, is oriented to set the number of services when passengers have a behavior characterized by two facts: a) no recommendation or regulation is made on the assignment from passengers to lines b) at each stop they choose a transit line accordingly to a route from their origin to their destination that they consider as optimal. The design model can be stated as a bilevel programming in which the lower level is an asymmetric traffic assignment problem. Asymmetries in costs come from the fact that passenger delays at stations waiting for a bus line to arrive depend not only on passenger's flow arriving at the station to board on that line but also on the unit's occupancy of that line arriving at the station. The upper level objective function is composed by two terms. The first one evaluates the operational costs of assigning units to a line plus the operational costs of bus services. The second cost is proportional to the total time spent by all passengers. The coefficient θ can be considered as the social cost of time.

In the formulation of [model SUE] below, $S^*(z)$ is the solution set of an asymmetric traffic model that can be stated as a variational inequality (V.I.): Find $v^* \in \mathcal{V}$ so that $T(v, z)^\top (v - v^*) \geq 0$, $\forall v \in \mathcal{V}$. This V.I., which makes up the lower level problem, is parametrized by the number of services z^ℓ assigned at each bus line $\ell \in L$. The number of services plays the role of a parameter for the links in the expanded network modeling

passenger flows on line ℓ.

$$[\text{Model SUE}] \quad \underset{n,z,v}{\text{Min}} \quad \sum_{\ell \in L}(\varsigma^\ell n^\ell + \gamma^\ell z^\ell) + \theta\, v^\top T(v,z)$$

$$s.t. \quad v \in S^*(z)$$

(2)

$$\text{A0-1} \qquad \sum_{\ell \in L} n^\ell \leq \mathsf{p}, \ \ n^\ell \geq 0, \ n^\ell \in \mathsf{Z}, \ \ \ell \in L$$

$$\text{A0-2} \qquad H^\ell n^\ell \geq z^\ell C_\ell(v,z), \ \ \ell \in L$$

$$\text{A0-3} \qquad 0 \leq z^\ell \leq \lambda^\ell \hat{f} \cdot H, \ \ z^\ell \in \mathsf{Z}, \ \lambda_\ell \in \{0,1\}, \ \ \ell \in L$$

$$\text{A0-4} \qquad z_\ell \geq \lambda^\ell\, \frac{H}{h_{max}}\,, \quad \ell \in L$$

2.2 A non-linear congested shortest path based service setting model

For the case of special services set in order to alleviate disruptions, it is difficult to impose to the passengers of a given o-d pair a splitting amongst several routes as a policy oriented to follow a system-optimum behavior. Instead it is easier to recommend a single route to be followed by all passengers of a given o-d pair. The recommended route should be optimal and should take into account congestion effects. Because of congestion, non-linearities appear and the model is similar to a non-linear shortest path choice problem and the objective function of the design model might minimize total costs. For this case, [Model SS] below was developed by Codina et al. in [4].

(3)

$$[\text{Model SS}] \quad \underset{\substack{n,z,v,\\ \tau,\lambda}}{\text{Min}} \quad \sum_{\ell \in L}(\varsigma^\ell n^\ell + \gamma^\ell z^\ell) + \theta \sum_{a \in A} v_a T_a(v,z) + \theta \sum_{\ell \in L}\sum_{b \in \Pi_\ell} \zeta_{a(\ell,b)}(v,z)$$

$$s.t. \qquad \text{constraints A0 as in [model SUE]}$$

$$\text{B0-1} \qquad v \in \mathcal{V}$$

$$\text{R0-1} \qquad \sum_{a \in \hat{E}(i)} \tau_a^\omega \leq 1, \quad \tau_a^\omega \in \{0,1\}, \ \ a \in \hat{E}(i), \ i \in N, \ \omega \in W$$

$$\text{R0-2} \qquad v_a^\omega \leq M\tau_a^\omega, \ \ a \in A \setminus A_G, \ \ \omega \in W$$

$$\text{Qb0} \qquad \sum_{\ell \in L_b} z^\ell \leq \hat{Z}_b(v,z), \ b \in \hat{N}_G$$

$$a = a(\ell,b), \quad b \in \Pi_\ell, \ \ell \in L \ \ \text{in} \ \ \text{Qp-1, Qp-2:}$$

$$\text{Qp0-1} \qquad v_a + v_{x(a)} \leq cz^\ell$$

$$\text{Qp0-2} \qquad \sum_{\ell \in L_b} \zeta_a(v,z) \leq \frac{H}{\eta_b}\, \hat{N}_b^{\mathsf{pax}}$$

It consists of the minimization of total costs, as in previous [Model SUE], but being these expressed conveniently in order to handle bulk service type queueing models for passengers at stations. The first term includes operational costs for setting and operation of services and the second plus the third one are in total the total travel time. The third term is made up by functions ζ for modeling queueing time of passengers at stations, whereas the second term includes times at links of the expanded network excluding

queueing of passengers at stations. Routing considerations appear reflected in constraints RO-1, RO-2 of the formulation, where binary decision variables τ_a^ω indicate, upon solving the problem, which of the boarding links in the expanded network, outgoing from a station, must be chosen by passengers with origin-destination pair ω. The model also includes constraints (Qb0) in order to reflect stations capacity in terms of maximum number of incoming buses per hour that the facility is able to admit taking into account the spillback of buses queueing for boarding/alighting operations and also, the maximum number of passengers that can be standing at a station, queueing for boarding (constraint Qp0-2). Constraints Qp0-1 impose a limitation in the boarding flow v_a at a boarding link a in a station b accordingly to the number of services z^ℓ of line ℓ to which the link belongs, the bus capacity c and the average number of passengers $v_{x(a)}$ *on* buses of line line ℓ arriving at station b.

[Model SS] is of the nonlinear mixed integer type and several optimization techniques are currently on essay in order to solve it. Function ζ has been determined using simulations with bulk-service queues and a convex piecewise approximation has been developed, resulting into an approximate model. A heuristic technique for obtaining suboptimal solutions has been developed showing very a good computational performance. It consists of freezing values of non-linear functions appearing in [Model SS] based on flows v and number of services z at previous iteration. In this way a mixed integer linear programming problem appears at each iteration which can be solved efficiently using CPLEX for medium size networks.

Heuristic algorithm for Model SS

(0) (a) Determine initially suitable values for the number of services and an initial value for the uncongested waiting time per passenger and service at a station, $\bar{P}_a^{(0}$. Set also temptative line cycle lengths $\bar{C}_\ell^{(0}$ for each line ; set also default bus service time at stations, $\kappa_b^{(0}$, and temptative initial values for bus waiting times at stations $w_{qb}^{0,(0}$, $b \in N_G$, so that an initial value for the maximum number of services allowable at a station, $\bar{Z}_b^{(0}$ can be evaluated using function $\tilde{Z}_b(\cdot, \cdot)$, i.e. $\bar{Z}_b^{(0} = \tilde{Z}_b(\kappa_b^{(0}, w_{qb}^{0,(0})$. Also, determine suitable link travel times $\bar{T}_a^{(0}$ accordingly.

(b) Solve model M2 for parameters $(\bar{T}^{(0}, \bar{Z}^{(0}, \bar{P}^{(0}, \bar{C}^{(0})$ in order to obtain flows and number of services $(v^{(1}, z^{(1})$. Set $\nu = 0$

At iteration $\nu + 1$:

(1) Obtain new values for packet service time $\kappa_b^{(\nu+1}$, waiting time of buses at stations, $w_{qb}^{0,(\nu+1}$, and maximum number of services allowed at each station $\bar{Z}_b^{(\nu+1}$ by an MSA step, using $\alpha_\nu = 1/(\nu + 2)$:

(4)
$$
\begin{aligned}
\kappa_b^{(\nu+1} &= \kappa_b^{(\nu} + \alpha_\nu \left(\kappa_b(v^{(\nu+1}, z^{(\nu+1}) - \kappa_b^{(\nu} \right), \\
w_{qb}^{0,(\nu+1} &= w_{qb}^{0,(\nu} + \alpha_\nu \left(w_{qb}^0(v^{(\nu+1}, z^{(\nu+1}) - w_{qb}^{0,(\nu} \right) \quad b \in N_G \\
\bar{Z}_b^{(\nu+1} &= \bar{Z}_b^{(\nu} + \alpha_\nu \left(\tilde{Z}_b(\kappa_b^{(\nu+1}, w_{qb}^{0,(\nu+1}) - \bar{Z}_b^{(\nu} \right)
\end{aligned}
$$

Update line cycles $C_\ell^{(\nu+1)} = C_\ell(v^{(\nu+1)}, z^{(\nu+1)}), \ell \in L$, uncongested waiting time per passenger and per service at stations $\bar{P}_a^{(\nu+1)} = P_a(z^{(\nu+1)})$ and link travel times $\bar{T}^{(\nu+1)}$ as follows:

(5)
$$
\begin{cases}
\bar{T}_a^{(\nu+1)} = t_a^0 + w_b^1 + \kappa_{b,\ell}(v^{(\nu+1)}, z^{(\nu+1)}) + w_{q,b'}^{0,(\nu+1)}, & \text{if } a = (j_\ell(b), j_\ell'(b')), \ \ell \in L, \ b, b' \in \Pi_\ell \\
\bar{T}_a^{(\nu+1)} = T_a(v^{(\nu+1)}, z^{(\nu+1)}) & \text{otherwise}
\end{cases}
$$

  (2) Solve approximate mixed linear integer model SS for parameters $\bar{T}^{(\nu+1)}$, $\bar{Z}^{(\nu+1)}$, $\bar{P}^{(\nu+1)}$, $\bar{C}^{\nu+1}$ and obtain flows $v^{(\nu+2)}$ and number of services $z^{(\nu+2)}$. Let $\nu \leftarrow \nu + 1$ and return to step 1.

The algorithm stops when, at a predetermined number r of consecutive iterations, the number of services assigned to bus lines do not change ($z^{\nu+1} = ... = z^{\nu+r}$) and also, during these r iterations, flows v and total delays ζ have little fluctuation ($\|v^{\nu+s+1} - v^{\nu+s}\|_2 \leq \epsilon_v$ and $\|\zeta^{\nu+s+1} - \zeta^{\nu+s}\|_2 \leq \epsilon_\zeta$, $s = 1, 2, ..., r - 1$).

3 Computational tests for frequency setting models

3.1 Tests for SUE models

[model SUE] was solved by means of the simulated annealing algorithm on the expanded transit network of figure 4.1 and with a passenger's demand given in table 4. Figure 2.1 shows the evolution of the objective function for 2000 iterations of S.A. algorithm with low temperature. Execution time on a HP laptop with 2Gb took 1h15min for 2000 iterations problems. In the computational experiences, the V.I., once the number of services were set, was solved using a diagonalization algorithm using a maximum of 500 iterations for each run of the diagonalization algorithm. A technique for reducing the number of iterations of this algorithm was used resulting in 25% savings in CPU time. As it can be seen from the figure, good objective function values for model SUE above were reached at a much earlier iteration than the 2000-th one. Runs with high temperature provided much worse computational results requiring almost all the 2000 iterations in order to reach very similar objective function values.
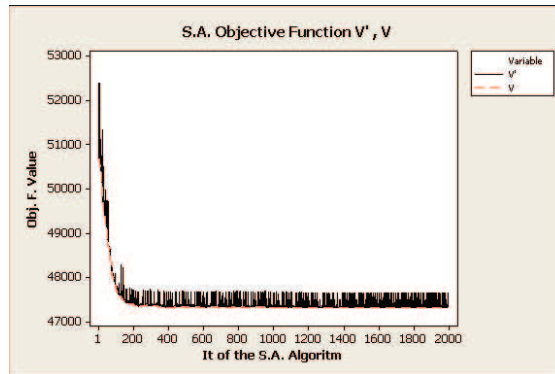

FIGURE 2. Evolution of objective function in model SUE using the simulated annealing algorithm.
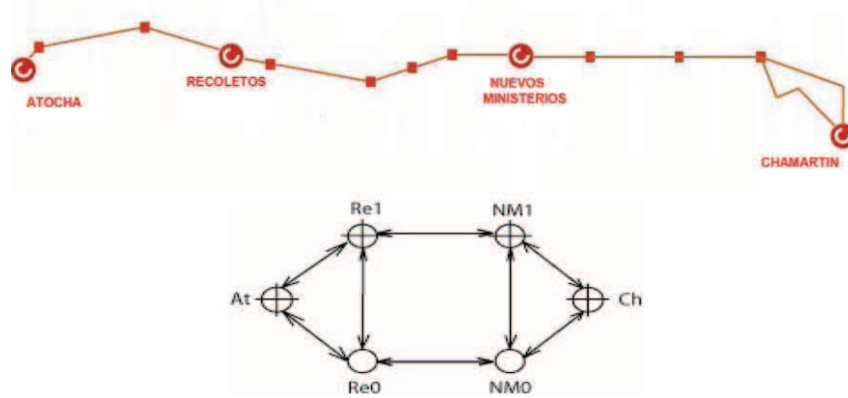
Figure 3. (Top) Schematic representation of Paseo de la Castellana's layout in Madrid where the bus-bridging system must operate with representation of the four disrupted railway stations. (Bottom) Representation of the subgraph (N_G, A_G) for movements of passengers outside the bus-bridging lines. Bus stops at Recoletos and Nuevos Ministerios have been mirrored on both sides of the Paseo de la Castellana. Disrupted railway stations appear as nodes marked with a cross, which in this model are considered also as possible bus stops.

3.2 Tests for the non-linear congested shortest path based model

In order to illustrate the model's performance and the proposed heuristic method to solve it, two test cases have been used.

The first one is a set of 20 candidate bus lines operating on 6 bus stops for a bridging system to assist disruption of a railway's corridor in Madrid, comprised by 4 railway stations. The expanded bus network, as defined in section 1, consists of 118 nodes and 240 links, 4 centroids which correspond with the physical location of disrupted railway stations and 12 origin-destination pairs. The origin destination trip table, with more than 37000 passengers in a three hours peak period, for Madrid's corridor appears in table 4 and was provided by railway authorities. The second test network is a set of 48 candidate bus lines operating on 17 bus stops for a bus-bridging system assisting a disruption of the line 1 metro's network in Barcelona (from stations Plaça d'Espanya to Clot). Figure 4 shows a schematic representation of the streets where the bus-bridging system must operate and on which the expanded bus network can be defined. This expanded bus network consists 310 nodes and 640 links, 10 centroids (corresponding to the disrupted
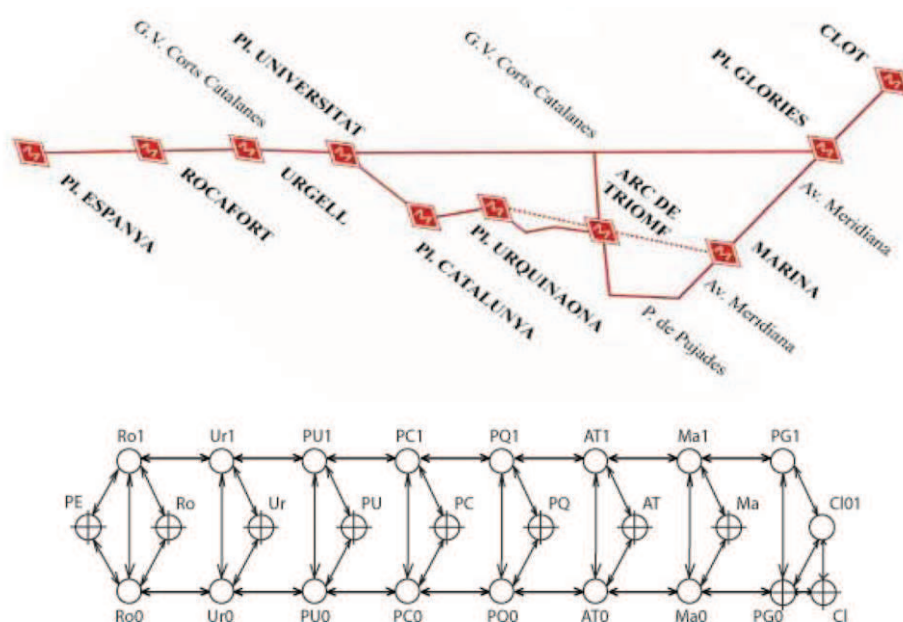
FIGURE 4. (Top) Schematic representation of main arterials in Barcelona where the bus-bridging system must operate for a disruption of metro line 1. (Bottom) Representation of the subgraph (N_G, A_G) for movements of passengers outside the bus-bridging lines. Excluding Plaça d'Espanya and Clot all other bus stations have been mirrored on both sides of the arterials. Disrupted metro stations appear as nodes marked with a cross.

metro stations) and 88 origin-destination pairs. The origin destination trip table for Barcelona's line 1 has been estimated with a total of 37.992 trips also during a period of three hours. Figures 3 and 4 show on the upper side a schema of the streets on which the bus-bridging systems will operate in both test cases. Figure 3 depicts a schematic representation of Paseo de la Castellana's layout in Madrid, where the bus-bridging system is intended to operate and figure 4 shows a schema of main arterials where the bus-bridging system may operate in Barcelona (Gran Via de les Corts Catalanes, Av. Meridiana, as well as some minor streets). Both expanded bus networks, contain a set of links A_G for movements of passengers carried out outside the bus network (access from disrupted stations to bus-stops, transfers between bus lines, portions of the trip carried out by foot). This set of links A_G appear depicted at the bottom in figures 3 and 4.

In both cases, because of the linear structure of the disrupted transportation network, the set of candidate lines that was considered for inclusion in the bus-bridging system was comprised by bus lines which can be classified under different topologies. On the one hand one may consider bus lines with only two terminal bus stops and bus lines with more than two bus stops. On the other hand one may consider lines as symmetric or asymmetric. Symmetric lines visit the same stops in both directions, whereas asymmetric lines may skip some or all of the stops (excluding terminals) in one of the directions. Figure 5 shows these types of lines using links $a \notin A_G$ of the transit expanded network. Figure 5(a) shows an example of a symmetric line defined on a bidirectional arterial visiting stops T1 P1, T2, P1' and again T1, being stop P1' the mirror of stop P1 located on the other side of the arterial more or less in front of stop P1. Physically, passengers need to cross the arterial if they want to go from stop P1 to stop P1' by means of links which are not represented in figure 5. Terminal stops T1 and T2 will be located also on a given side of the arterial. Terminal bus stops are not mirrored in the lines considered for the two test cases. Examples of asymmetrical lines are also given in figure 5(b) and (c). Another class of asymmetrical bus lines also considered in the bus-bridging systems of the two test cases, are those which terminal bus stops are only for boarding or for alighting, i.e. no passengers are transported on unit's return to the starting terminal. Taking into account the previous types of bus lines, 20 of them where selected for the Madrid's test case and 48 the Barcelona's test case. In both test networks, bus units with a capacity for 100 passengers has been assumed. In both cases the time period under consideration has been a morning peak period of $H = 180$ minutes.

The runs shown in this section for both test cases assume that exploitation and setting costs of the bus-bridging system are nearly zero, so actually the objective under minimization is the total travel time of passengers.

TABLE 1. O-D Trip matrix (station-to-station) during period $H = 180$ minutes for Madrid's railway corridor. Last row and column are average rates for arrivals and departures per minute at stations.

	At	Re	NM	Ch	Total Or.	pax/min
At	0	2.011	22.097	368	24.476	$135, 98$
Re	170	0	3.066	230	3.466	$19, 25$
NM	4.386	150	0	170	4.706	$26, 14$
Ch	2.504	150	2.438	0	5.092	$28, 28$
Total Dest.	7.060	2.311	27.601	768	**37.740**	—
pax/min	$39, 22$	$12, 84$	$153, 34$	$4, 26$	—	—

The following tables 2 and 3 show the main results. The heuristic algorithm has been implemented in AMPL, using CPLEX 11.0 as solver. Runs shown in the previously mentioned tables have been performed on a laptop 1.2GHz, 2Gb RAM. In order to see how the increase in the level of demand congests the bus network and affects the performance of the heuristic algorithm, several runs have been performed taking the original origin-destination trip matrix shown in table 4, with each cell multiplied by a common factor η. Thus $\eta = 1, 3$ implies that all cells of the matrix have been increased by 30% and $\eta = 1$ implies that the original trip table is being used. Column Dem shows the
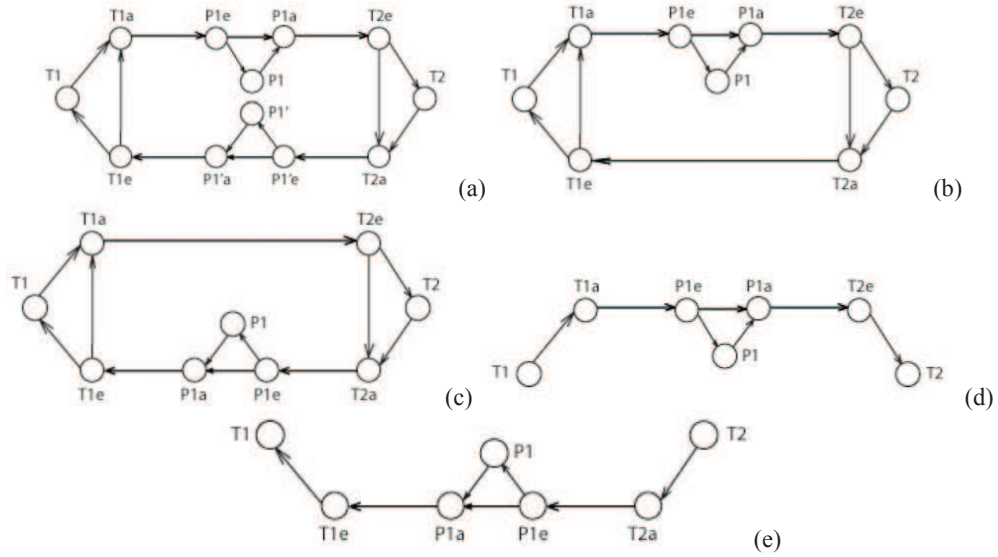
FIGURE 5. Lines with more than two stops. (a) symmetric lines, (b) asymmetric lines for one direction loading passengers on the return, (c) asymmetric lines for the opposite direction loading passengers on the return, (d) asymmetric lines on one direction without loading passengers at the return and (e) asymmetric lines on the opposite direction without loading passengers at the return.

total resulting trips. Column UnCov reports the percentage of the trips that have not used at all the auxiliary bus system. Column Tcpu reports the total number of elapsed seconds for the run and column #iter the total number of iterations required by the heuristic algorithm described in section 2.2 to converge. Column AvT reports the average trip time in minutes for the mass of passengers and column $\bar{v}$ reports their average speed. This speed is calculated as the total time required to go from origin to destination divided by the shortest distance that travelers would use by walking. Column %Walk reports the percentage of the total time used by passengers for reaching their corresponding bus stops, transfers between bus stops or simply because their journey is made completely by walking from origin to destination due to congestion of the lines that the model assigns a positive number of services. Finally, columns f^* and f^0 report final objective function

values reached by the heuristic and initial objective function value at starting solution respectively.

In both tables runs marked with 'a' correspond to solutions of the model M0 where a single route is imposed to passengers of a given origin destination pair on the expanded transit network, as stated in constraints `RO-1` and `RO-2`. Results show that, as the level of demand increases these routes become more and more congested, making that for the total number of trips of a given origin destination pair travel by walking may become attractive. This happens specially with those origin destination pairs with a large number of trips. For runs marked with 'b', passengers of a given origin destination pair may ride a bus at different bus stops, but at any bus stop that they reach they wait for a determined bus line without following a behavior based on strategies. In this last case the performance of the bus-bridging system is slightly enhanced.

The model is able to detect bottlenecks of the system, which in all of the runs shown appear in the capacity of bus stops for allocating bus queues, which impose by means of constraints `Qb0` in model M0, a limitation in the input flow of buses that the bus stop may admit. It must be also remarked that bus service time has not appeared as a limitation for the input flow of buses that a bus stop may admit. In all runs, the number of passengers that a bus station may allocate queuing for his/her bus to arrive has shown to be sufficient.

In the auxiliary bus system for Barcelona's metro (line 1), runs 3 and 6 (table 2) performed with and extra 30% of demand and bus stops at Plaça d'Espanya (PE) and Arc de Triomf (AT1) were at capacity. By increasing the space allowed for buses to queue at these stop (2 extra buses were allowed to queue), run 7, marked with 'c' shows that the performance of the whole system is enhanced. It must be noticed that the level of congestion has a serious impact on the performance of the heuristic algorithm on this test case.

In the auxiliary bus system for Madrid's corridor, forcing a unique route for all passengers of an origin destination pair or allowing them to choose more than one bus stop affects the performance of the solutions, because of the big flow Atocha (At) to Nuevos Ministerios (NM) which can not be allocated by a single bus line during the period of 3 hours for $\eta = 1, 0$ and $\eta = 1, 3$.

4 Congested transit assignment models

Strategy based transit assignment models used in modeling passenger flows in regular lines of urban public transportation do not reflect congestion effects until very recently. Because of that frequency setting or service setting models which take into account congestion when passengers follow strategies have not yet been developed. A classical uncongested model is that of Spiess [15], which can be formulated as the following linear program:

$$
\begin{aligned}
\text{Min}_{v,w} \quad & \sum_{d \in D} \sum_{a \in A} t_a v_a^d + \sum_{d \in D} \sum_{i \in \hat{N}_d} w_i^d \\
\text{[PL]}(r,t) \quad s.t: \quad & v_a^d \leq r_a w_i^d, \quad \mid \theta_a^d, \quad a \in \hat{E}(i), \ i \in \hat{N}, \ d \in D \\
& v \in V
\end{aligned}
$$

(6)

TABLE 2. Results for the auxiliary bus system for Barcelona's metro line 1 (a) when all passengers for an O-D pair are forced to follow a single route (b) when passengers of an O-D pair are allowed to board on the auxiliary system at different bus stops, riding there at a unique recommended bus line

#		η	Dem.	UnCov.	Tcpu	#iter	AvT	%Walk	$\bar{v}$	f^*	$f^{(0}$
			trips		(s)		(min)		(km/h)	$(\$)$	$(\$)$
1	a	$0,5$	18.996	29%	929	140	27	56%	$6,9$	41.552	74.213
2	a	$1,0$	37.992	52%	1.958	151	$34,9$	76%	$5,5$	107.290	135.380
3	a	$1,3$	49.130	67%	3.366	251	$37,3$	$83,4\%$	$5,1$	149.190	175.230
4	b	$0,5$	18.996	$22,7\%$	797	150	27	$56,2\%$	$6,9$	41.546	74.213
5	b	$1,0$	37.992	$46,5\%$	1.144	180	$34,8$	$73,4\%$	$5,4$	106.840	135.380
6	b	$1,3$	49.130	65%	681	140	$37,1$	$82,4\%$	$5,1$	148.420	180.280
7	c	$1,3$	49.130	$51,5\%$	183	31	$35,2$	$76,5\%$	$7,2$	140.880	180.680

TABLE 3. Results for the auxiliary bus system for Madrid's railway corridor (a) when all passengers for an O-D pair are forced to follow a single route (b) when passengers of an O-D pair are allowed to board on the auxiliary system at different bus stops, riding there at a unique recommended bus line

#		η	Dem.	UnCov.	Tcpu	#iter	AvT	%Walk	$\bar{v}$	f^*	$f^{(0}$
			trips		(s)		(min)		(km/h)	$(\$)$	$(\$)$
1	a	$0,5$	18.870	$14,3\%$	$\sim 6,8$	51	$25,3$	40%	12	38.741	70.521
2	a	$1,0$	37.740	$72,8\%$	$\sim 8,5$	37	$54,3$	$92,1\%$	$7,1$	181.140	181.230
3	a	$1,3$	49.062	$72,8\%$	$\sim 8,6$	42	$59,5$	$91,8\%$	$5,5$	236.270	238.250
1	b	$0,5$	18.870	$14,3\%$	$\sim 5,5$	14	25	40%	12	38.741	70.521
2	b	$1,0$	37.740	$14,3\%$	~ 9	21	$41,7$	$63,8\%$	$7,8$	127.350	168.520
3	b	$1,3$	49.062	$14,3\%$	$\sim 9,5$	23	$46,3$	$71,7\%$	$5,5$	184.170	239.360

Based on the results of Cominetti and Correa in [**6**], Cepeda *et al.* in [**1**] prove that their strategy based congested network equilibrium transit notion is equivalent to the minimization of the following nonconvex, nondifferentiable gap function $\tilde{G}_{\text{CCF}}(v)$

$$(7) \qquad \tilde{G}_{\text{CCF}}(v) = \sum_{d \in D} \left[\sum_{a \in A} v_a^d t_a(v) + \sum_{i \in N_d} \underset{a \in E(i)}{\text{Max}} \left\{ \frac{v_a^d}{f_a(v)} \right\} - \sum_{i \in N_d} g_i^d \tilde{\tau}_i^d(v) \right]$$

over the feasible set of destination flow vectors V, i.e. solutions of the congested transit equilibrium model are also global minima of the problem $\text{Min}_{v \in V} \tilde{G}_{\text{CCF}}(v)$. In [**2**] it is proved that solving this problem is equivalent to the following variational inequality (VI):

$$(8) \quad (\text{VI}) \quad \left[\begin{array}{l} \text{Find } (v, \zeta) \in V \times S \text{ so that, } \forall (u, z) \in V \times S: \\[2mm] \displaystyle\sum_{d \in D} \sum_{i \in N_d} \left(\sum_{a \in E(i)} (t_a(v) + s_a(v)\zeta_a^d)(u_a^d - v_a^d) - y_a^d(v)(z_a^d - \zeta_a^d) \right) \geq 0 \end{array} \right]$$

where $y_a^d(\text{v}) = s_a(v)\text{v}_a^d$, $a \in A$ and functions $s_a(v)$ are defined as $s_a(v) = \sigma_a(v)$, if $a \in \hat{E}(i)$ and $s_a(v) \equiv 0$, if $a \in E(i) \setminus \hat{E}(i)$. $S = \bigotimes_{d \in D} \bigotimes_{i \in \hat{N}_d} S_i^d$ and $S_i^d = \left\{ \alpha \in \Re_+^{|\hat{E}(i)|} \mid \sum_{a \in \hat{E}(i)} \alpha_a = 1 \right\}$ associated to node $i \in \hat{N}$. In [2] previous results are also extended to the case of sharp capacity constraints on bus lines either explicitly or implicitly imposed by effective frequency functions $\sigma_a(v)$ and the MSA algorithm developed for the congested strategy based transit assignment problem in [1] can be easily adapted for this case. Next subsection shows some computational results for this case.

4.1 Some computational results for the capacitated transit assignment problem

The transit network for this example is made up of eight transit lines and its expanded transit network is shown in figure 4.1. Effective frequency functions for boarding links are of the type $f_a(v) = 0.2(1 - \rho_a^2(v))$ and $\rho_a(v) = v_a/(c - v_{m(a)})$. Capacity c at boarding links is 9600 passengers for a period of 3 hours. Link travel times are given in [2]. Boarding links (i, j) are those whose i-node is either 1, 2, 3 or 4. Demands in passengers for a 3 hours period are shown in table 4 below. This matrix has been uniformly augmented by a factor τ in order to conduct computational experiments

	1	2	3	4	Total Or.	pax/min
1	0	2011	22097	368	24476	135.98
2	170	0	3066	230	3466	19.25
3	4386	150	0	170	4706	26.14
3	2504	150	2438	0	5092	28.28
Total Dest.	7060	2311	27601	768	**37740**	–
pax/min	39.22	12.84	153.34	4.26	–	–

TABLE 4.   O-D Trip table for a period of 180 minutes. Last row and column are average arrival and departure rates of passengers at bus stops.

For this test network, the MSA algorithm in [1] with implicit capacity constraints behaves well for $\tau = 1.0, 1.2$. For $\tau = 1.3$, iterates violating capacity constraints appear during the run of the algorithm, although it finally converges to the solution. For $\tau = 1.6, 2.0$ and larger values, it shows unstable behaviour because capacity infeasible iterates appear too often. Results comparing algorithm in [1] (implicit capacity constraints) and the algorithm using explicit capacity constraints are shown in table 5. Also the self-regulated MSA step in [11] has been used for this example with $\tau = 2$ for the MSA algorithm in [1] and for the algorithm using explicit capacities.
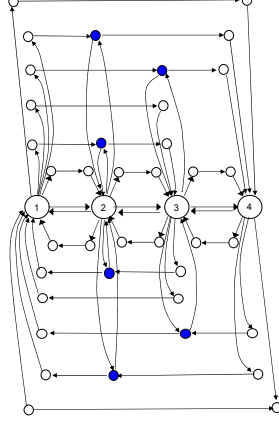
FIGURE 6. Expanded network model for the example.

τ	#iter	lowest $\check{g}A$	$\check{g}B$	lowest $\check{g}B$
1.6	1000	—	8.85E-04	8.36E-05
1.6	4000	0.011	—	—
1.6	50000	—	2.43E-04	—
2.0	4000	0.008	0.0233	1.95E-05
2.0	50000	1.95E-03	8.30E-04	7.68E-04
2.0	$50000^{(*)}$	1.25E-03	2.15E-05	1.91E-05

TABLE 5. Comparison between algorithm with (B) and without (A) explicit capacity constraints. $\check{g}$ = relative gap. (*) use of self-regulated MSA step in [**11**].

5 Approximating models with constant frequencies

Let us consider

$$(9) \qquad w_i^d(\mathrm{v}_i^d) = \underset{a \in \hat{E}(i)}{\mathsf{Max}} \left\{ \frac{v_a^d}{f_a(v)} \right\}, \; i \in \hat{N}_d, \, d \in D$$

When frequencies $f_a(\cdot) = r_a$, $a \in \hat{E}(i)$, $i \in \hat{N}$ are flow independent and constant and travel costs $t_a(v)$ at links $a \in A$ have a diagonal and positive semi-definite jacobian , or equivalently $t_a(v) = t_a(v_a)$, it will be shown that models developed in Spiess (1984) and in Spiess and Florian (1989) are reproduced. If frequencies are constant, then $x_i^d(\mathrm{v}_i^d) = (..., v_a^d/r_a, ...; \; a \in \hat{E}(i))$ and the subgradient of w_i^d are:

$$(10)$$

$$\partial w_i^d(\mathrm{v}_i^d) = \left\{ (..., \zeta_a^d/r_a, ...; \; a \in \hat{E}(i)) \in \mathbb{R}^{|\hat{E}(i)|} \mid \sum_{a \in \hat{E}_*^d(i,\mathrm{v})} \zeta_a^d = 1, \; \zeta_a^d = 0, \; a \notin \hat{E}_*^d(i,\mathrm{v}) \right\}$$

V.I. (`VIcond`) (8) can be also stated as:

(11) $v \in V$

(12) $t_a(v_a) + (\partial w_i^d(v_i^d))_a = \lambda_i^d - \lambda_j^d + \xi_a^d, \quad d \in D, i \in \hat{N}_d, \ a = (i,j) \in \hat{E}(i),$

(13) $t_a(v_a) = \lambda_i^d - \lambda_j^d + \xi_a^d, \quad d \in D, \ i \in \hat{N}_d, \ a = (i,j) \in E(i) \setminus \hat{E}(i),$

(14) $v_a^d \geq 0, \ \ \xi_a^d \geq 0, \ \ v_a^d \xi_a^d = 0,$

which are, in fact, first order conditions of the following optimization problem for the semicongested transit assignment problem in Spiess and Florian (1989):

(15)
$$\operatorname{Min}_{v,w} \ \sum_{a \in A} \int_0^{v_a} t_a(\alpha) \, d\alpha + \sum_{d \in D} \sum_{i \in \hat{N}} w_i^d$$
$$s.t: \quad v_a = \sum_{d \in D} v_a^d, \ a \in A$$
$$v_a^d \leq r_a w_i^d, \ a \in \hat{E}(i), \ i \in \hat{N}, \ d \in D$$
$$v \in V$$

Finally problem (15) reduces to the transit assignment model in Spiess (1984) for $t_a(v_a) = t_a = \mathtt{ctant}$ which will be designated by $[\mathtt{PL}](r,t)$:

(16) $[\mathtt{PL}](r,t)$
$$\operatorname{Min}_{v,w} \ \sum_{d \in D} \sum_{a \in A} t_a v_a^d + \sum_{d \in D} \sum_{i \in \hat{N}} w_i^d$$
$$s.t: \quad v_a^d \leq r_a w_i^d, \ a \in \hat{E}(i), \ i \in \hat{N}, \ d \in D$$
$$v \in V$$

Proposition 5.1 *Assume now that $t_a > 0$, $\forall a \in A$ in problem (16). Let $v^* = (...,v^{*,d},...; d \in D)$ be a per destination flow vector, solution of problem (16). Then, each of the vectors $v^{*,d}$ can be expressed as a convex combination of vertexes $\hat{v}_\nu^d \in V$, so that for each of them there exists a collection of reverse trees of acyclic paths on the expanded network with common destination (root) d. If by W_d it is denoted the subset of O-D pairs with common destination $d \in D$, then by loading a path tree with O-D flows g_ω, $\omega \in W_d$, a vertex $\hat{v}_\nu^d$ in the collection of vertexes for destination $d \in D$ is obtained.*

   Proof: Problem (16) can be rewritten as:

(17) $\operatorname{Min}_{w \geq 0} \ \ U(w) + \sum_{d \in D} \sum_{i \in \hat{N}} w_i^d$

where function $U(w)$ is defined for $w \geq 0$ as:

(18)
$$U(w) \overset{\triangle}{=} \ \operatorname{Min}_v \ \sum_{d \in D} \sum_{a \in A} t_a v_a^d$$
$$s.t: \quad v_a^d \leq r_a w_i^d, \ a \in \hat{E}(i), \ i \in \hat{N}, \ d \in D$$
$$v \in V$$

By lagrangian duality, for any $w \geq 0$, there exists $\gamma_i^d \geq 0$ so that solutions of previous problem (18) are also solutions of

$$(19) \qquad \mathsf{Min}_{v \in V} \quad \sum_{d \in D} \sum_{a \in A} t_a v_a^d - \sum_{d \in D} \sum_{i \in \hat{N}} \gamma_i^d (r_a w_i^d - v_a^d)$$

and that there is null duality gap. Now it suffices to note that previous problem (19) has the structure of a separable multidestination flow problem with positive arc costs, where each of the subproblems per destination $d \in D$ has as optimal basis those corresponding with reverse shortest path trees rooted at destination $d \in D$. $\square$

Let now consider Γ_ω, the set of acyclic paths joining O-D pair $\omega \in W$ on the expanded transit network and let H_ω the polytope of path flows for acyclic paths joining O-D pair $\omega \in W$:

$$(20) \qquad H_\omega \triangleq \left\{ \mathrm{h} \in \mathrm{I\!R}_+^{|\Gamma_\omega|} \ \middle| \ \sum_{r \in \Gamma_\omega} h_r^\omega = g_\omega \right\}$$

Correspondingly let H_d, $d \in D$ and H be defined as:

$$(21) \qquad H_d \triangleq \bigotimes_{\omega \in W_d} H_\omega, \ \ d \in D$$

$$(22) \qquad H \triangleq \bigotimes_{\omega \in W} H_\omega$$

The following corollary is a straightforward consequence of previous proposition 5.1.

Corollary 5.2 *Let* $\mathrm{v}^* = (..., \mathrm{v}^{d,*}, ...; \ d \in D)$ *be a per-destination flow vector, solution of problem (16) and let W_d be the set of O-D pairs with destination $d \in D$. Then each of the vectors $\mathrm{v}^{d,*}$ can be decomposed in terms of flows $\mathrm{h}^d \in H_d$ on acyclic paths (generally non unique decompostion). In other words, if $\Delta^d = (\delta_a^r)$ is a link-path incidence matrix for paths with common destination $d \in D$ on the expanded transit network, then $\exists \mathrm{h}^{d,*} \in H_d$ so that $\mathrm{v}^{d,*} = \Delta^d \mathrm{h}^{d,*}$, or explicitly*

$$(23) \qquad v_a^{d,*} = \sum_{\omega \in W_d} \sum_{r \in \Gamma_\omega} \delta_a^r h_r^{\omega,*}, \ a \in A, \ d \in D$$

Corollary 5.3 *Consider now a solution $\mathrm{v}^* \in \mathrm{V}^*$ of variational inequality (VI) (8). Then results in corollary 5.2 also hold for such $\mathrm{v}^* \in \mathrm{V}^*$.*

Proof: It is clear that, as link flows v^* are a solution of $\mathrm{PL}[f(v^*), t(v^*)]$ and the following equivalence must hold:

$$(24) \qquad v^* \in \mathsf{Sol}(\,[\mathrm{PL}](f(v^*), t(v^*))\,) \ \Leftrightarrow \ G_{\mathrm{CCF}}(\mathrm{v}^*) = 0$$

being $v^* = \sum_{d \in D} \mathrm{v}^{d,*}$ the total link flows as usual. This equivalence implies that previous corollary 5.2 must also hold for any $\mathrm{v}^* \in \mathrm{V}^*$. $\square$

5.1 A smoothing approximation to C3F transit equilibrium model

A natural approximation for functions $w_i^d(\cdot)$ is by replacing them by the smoothing approximations $\phi(x,t)$. The relevant properties of these functions are summarized in Appendix 7. In the context of this paper, let $\psi_i^d(x_i^d(\mathrm{v}),z)$ be defined for some parameter $z > 0$ at node $i \in \hat{N}_d$ and destination $d \in D$ as:

$$(25) \qquad \psi_i^d(x_i^d(\mathrm{v}),z) = \frac{1}{z} \log \left(\sum_{a \in \hat{E}(i)} \exp\left(z\, v_a^d \sigma_a(v)\right) \right)$$

Now, variables ζ_a^d in (8) can be approximated by

$$(26) \qquad \zeta_a^d \approx \bar{\zeta}_a^d = (\nabla_x \psi_i^d(x_i^d(\mathrm{v}),z))_a = \frac{\exp(z\, v_a^d \sigma_a(v))}{\sum_{a' \in \hat{E}(i)} \exp(z\, v_{a'}^d \sigma_{a'}(v))} \;,\; a \in \hat{E}(i),\; i \in \hat{N}$$

Let $\tilde{\mathrm{T}}^d(\mathrm{v},z) = (...\tilde{\Psi}_a^d(\mathrm{v},z)...;a \in A)$ where $\tilde{\Psi}_a^d$ are defined as

$$(27) \quad \tilde{\Psi}_a^d(\mathrm{v},z) = \begin{cases} t_a(v) + \sigma_a(v)(\nabla_x \psi_i^d(x_i^d(\mathrm{v}),z))_a, & a \in \hat{E}(i) \\ t_a(v) & a \in E(i) \setminus \hat{E}(i) \end{cases} \quad d \in D,\; i \in N_d$$

By means of the smoothing functions ψ_i^d, the variational inequality (VI) in (8) can be approximated by the following one

$$(28) \qquad (\texttt{VIaprox}) \quad \begin{aligned} 0 &\in \tilde{\mathrm{T}}^d(\mathrm{v},z) + \mathsf{N}_{\mathrm{V}^d}(\mathrm{v}^d), & d \in D \\ \mathrm{v}^d &\in \mathrm{V}^d, & d \in D \end{aligned}$$

When $z \to \infty$ solutions v_z^* of previous V.I. (28) lie in V^*, the solution set of (VI) in (8).

5.2 Properties of the smoothing approximation with constant frequencies

From previous conditions (11) to (14) it is clear that model (15) can be rewritten as

$$(29) \qquad \begin{aligned} \mathsf{Min}_{\mathrm{v}} &\quad \sum_{a \in A} \int_0^{v_a(\mathrm{v})} t_a(r)dr + \sum_{d \in D}\sum_{i \in \hat{N}} w_i^d(\mathrm{v}_i^d) \\ s.t: &\quad \mathrm{v} \in \mathrm{V} \end{aligned}$$

where $v_a(\mathrm{v}) = \sum_{d \in D} v_a^d$ and functions $w_i^d(\cdot)$ are defined with constant frequencies $f_a(v) = f_a$. The objective function of previous program (29) is convex but non-differentiable. If, for the case of constant frequencies, we define

$$(30) \qquad \bar{\psi}_i^d(\mathrm{v}_i^d,z,r_i) = \frac{1}{z} \log \left(\sum_{a \in \hat{E}(i)} \exp\left(z\frac{v_a^d}{r_a}\right) \right),\; i \in \hat{N},\; d \in D$$

then a smooth approximation for problem (29) using functions $\bar{\psi}_i^d(x_i^d(\mathrm{v}),z)$ results in

$$\text{Min }_{v} \quad \sum_{a \in A} \int_0^{v_a(v)} t_a(r)dr + \sum_{d \in D} \sum_{i \in \hat{N}} \bar{\psi}_i^d(v_i^d, z, r_i)$$

(31)

$$s.t : \quad v \in V$$

where by $r_i = (...,r_a,...; a \in \hat{E}(i))$ it is denoted a vector of constant frequencies at node $i \in \hat{N}$.

Lemma 5.4 *Assume that boarding links $a \in \hat{E}(i)$ at stops $i \in \hat{N}$ have constant boarding times, $t_a = $ ctant. Then solutions $v^* \in V^*$ of model (31) are such that per-destination flows at boarding links are unique, i.e. $(v_i^{d,*})_a = \{v_a^{d,*}\}$ $d \in D$, $i \in \hat{N}_d$, $a \in \hat{E}(i)$.*

Proof: Because objective function of program (31) is convex and differentiable for $z > 0$, its gradient must be constant on the relative interior of the solution set (see, Mangasarian (1988), lemma 1.a) and the set of equations

(32) $\quad t_a + \sigma_a(\nabla_x \psi_i^d(x_i^d(v), z))_a = (\text{ctant})_{a,d}, \quad \forall v_i^d \in (\text{int}V^*)_{i,d}, \; d \in D, \; i \in \hat{N}_d, \; a \in \hat{E}(i)$

define unique per-destination flows $v_a^{d,*}$ at boarding links $a \in \hat{E}(i)$ of stops $i \in \hat{N}$ as ψ_i^d is a strictly convex function $\square$

Uniqueness of solutions in total flows on the remaining links depend on characteristics of the travel time functions t_a of that links. Assume now that the following representation given in the left side of figure 7 below is adopted for the transit network. As per-destination flows $v_a^{d,*}$ are uniquely defined by problem (31), then it will suffice to consider the original transit network on which boarding links have been suppressed and flows $v_a^{d,*}$ are injected on the boarding node of the corresponding transit line as shown in the right side of the following figure 7.
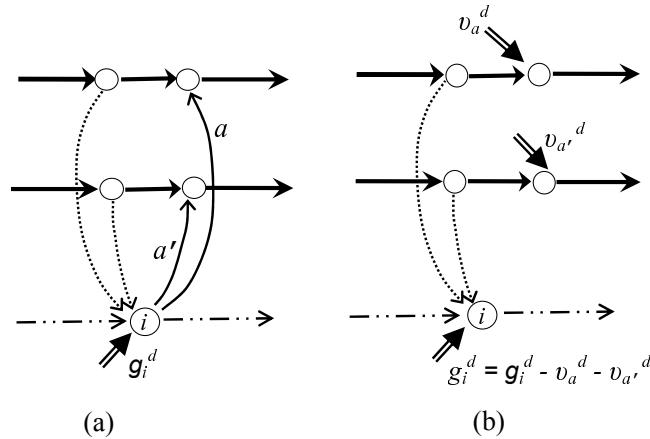

FIGURE 7.   (a) A convenient representation of the transit network and (b) an equivalent network after freezing boarding flows at a solution of model (31)

Formally consider now a traffic assignment network model $G = (N, A)$ based on the structure of the expanded transit network $G = (N, A)$, with the following set of origin-destination pairs W and with the following origin-destination trip matrix $\mathbf{g}$, being all them defined as follows

$$(33) \qquad A = A \setminus \left(\bigcup_{i \in \hat{N}} \hat{E}(i) \right),$$

$$(34) \qquad W = W \bigcup \left(\bigcup_{d \in D} \left(\bigcup_{i \in \hat{N}} \left\{ (j_a, d) \mid (i, j_a) \in \hat{E}(i) \right\} \right) \right)$$

$$(35) \qquad \mathbf{g}_p^d = \begin{cases} g_i^d & \text{if } (i, d) \in W \ \& \ i \notin \hat{N} \\ g_i^d - \sum_{a \in \hat{E}(i)} v_a^{d,*} & \text{if } (i, d) \in W \ \& \ i \in \hat{N} \qquad (p, d) \in W \\ v_a^{d,*} & \text{if } p = j_a, \ a = (i, j_a) \in \hat{E}(i), \ i \in \hat{N} \end{cases}$$

If Γ_ω, $\omega = (p, d) \in W$ is the set of paths joining origin p with destination d, let $\tilde{H}$ denote the polyhedron of feasible path flows h_r on paths $r \in \Gamma_\omega$

$$(36) \qquad \tilde{H} = \left\{ h_r \geq 0 \mid \sum_{r \in \Gamma_\omega} h_r = \mathbf{g}_p^d, \ (p, d) \in W \right\}$$

and let $\tilde{V}$ be the polyhedron of total flows on links

$$(37) \qquad \tilde{V} = \left\{ v_a, a \in A \mid v_a = \sum_{\omega \in W} \sum_{\substack{r \in \Gamma_\omega \\ r \ni a}} h_r \right\}$$

If travel time functions $t_a(\cdot)$ are adopted for the links in A, with previous definitions (33) to (37), the following fixed demand traffic assignment problem must be considered

$$(38) \qquad \begin{aligned} & \text{Min}_v \ \sum_{a \in A} \int_0^{v_a(\mathbf{v})} t_a(r) dr \\ & s.t : \quad v \in \tilde{V} \end{aligned}$$

Assumption 5.5 *Travel time functions $t_a(\cdot)$ at links $a \in A$ are such that the fixed demand traffic assignment problem (38) has uniqueness of solutions in total link flows.*

Sufficient conditions for uniqueness of solutions in total link flows in traffic assignment problems can be found, for instance in Patriksson (1994), theorem 2.5, page 43.

Lemma 5.6 *Assume, as in previous lemma 5.4, that boarding links $a \in \hat{E}(i)$ at stops $i \in \hat{N}$ have constant boarding times, $t_a = \mathsf{ctant}$ and that link travel time functions $t_a(\cdot)$ for the remaining links verify previous assumption 5.5. Then total link flows $v^* \in \mathrm{V}^*$ of model (31) are unique.*

Proof: Because of previous lemma 5.4, problem (31) has uniqueness of solutions in per-destination flows at boarding links. The total flows at remaining links of the transit network will be then given by solving previous fixed demand traffic assignment problem (38). As a result problem (31) has uniqueness of solutions in total link flows. □

5.3 Elastic demand formulations from hierarchical logit models

Assume a transportation network composed by a set N_T of transportation modes and that users choose them by means of their perception of its utility. Assume that the mean utility of all modes but the one corresponding to transit are known and constant. Let $\tilde{u}_p^\omega$ the mean utility for transportation mode $p \in N_T$ for the origin-destination pair $\omega \in W$

Let $T = (N_T, A_T)$ be a tree where each leaf node represents a pure mode and each non leaf node represents a composed mode. If $E_T(p)$ denotes the emerging nodes of a node p in the tree, then $L \triangleq \{ p \in N_T \mid E_T(p) = \emptyset \}$ be the set of leaf nodes and $C \triangleq \{ p \in N_T \mid E_T(p) \neq \emptyset \}$ the set of non leaf nodes. Root node will be assigned label R and node corresponding to transit mode will be denoted by the label B.

Assume that the total demand (all modes) is fixed and known and given by g_{R}^ω, $\omega \in W$. Let $n_t = |N_T|$. and that node B is a pure transportation mode corresponding to a leaf in the tree T.

For a mode corresponding to a node $p \in N_T$ and destination $d \in D$, let g_p^d denote the vector of o-d trip flows, $g_p^d = (..., g_p^\omega, ...; \omega = (i,d),\ i \in O(d)) \in \Re^{|O(d)|}$. O-D trip flows per mode g_p^ω, accordingly to the hierarchical tree T will be subject to the following relationships

$$(39) \qquad \sum_{p' \in E_T(p)} g_{p'}^\omega = g_p^\omega, \mid \beta_p^\omega, \ \ \omega \in W, \ \ p \in C$$

where β_p^ω are the corresponding dual variables. Consider also the vector $g^d = (..., g_p^d, ...; p \in N_T) \in \Re^{|O(d)|+n_t}$ and the vector $g = (..., g^d, ...; d \in D)$.

Let G be feasibility set for g which can be decomposed as $G = \otimes_{d \in D} G^d$, where each of the G^d is defined as:

$$(40) \qquad G^d = \left\{ g^d \in \Re_+^{|O(d)|+n_t} \mid \text{relationships (39) are verified} \right\}$$

The set Ω_G^d of feasible transit flows per destination on links and feasible trip o-d flows per mode will be given now by

(41)

$$\Omega_G^d = \left\{ (v^d, g^d) \in \Re_+^{|A|+|N_d|+n_t} \ \middle| \ g^d \in G^d, \ \sum_{a \in E(i)} v_a^d - \sum_{a \in I(i)} v_a^d - g_{\mathsf{B}}^{i,d} = 0, \right. $$
$$\left. \sum_{a \in I(d)} v_a^d - \sum_{i \in O(d)} g_{\mathsf{B}}^{i,d} = 0, \ \ v_a^d = 0, \forall a \in E(d),\ i \in N_d \right\}, \quad d \in D$$

and $\Omega_G = \otimes_{d \in D} \Omega_G^d$. Consider now the function $L(g) = (..., L_p^\omega(g), ...; \omega \in W, p \in N_T)$ whose components $L_p^\omega(g)$ are defined as:

$$
(42) \qquad L_{p'}^{\omega}(\mathrm{g}) = \begin{cases} \dfrac{-\eta_p}{\vartheta} \, \log\!\left(\dfrac{g_{p'}^{\omega}}{g_p^{\omega}}\right) + \tilde{u}_{p'}^{\omega}, & \text{if } p' \in L \setminus \{\mathsf{B}\} \\[2ex] \dfrac{-\eta_p}{\vartheta} \, \log\!\left(\dfrac{g_{p'}^{\omega}}{g_p^{\omega}}\right), & \text{if } p' \in C \cup \{\mathsf{B}\} \end{cases} , \ p' \in E_T(p), \ p \in C
$$

where $\eta_{\mathsf{R}} \stackrel{\triangle}{=} 1$.

If now $\mathrm{L}^d(\mathrm{g}) = (\ ..., L_p^{\omega}(\mathrm{g}), ... \ ; \ \omega = (i,d) \in W\ )$, $d \in D$, the following V.I. (43) defines an elastic transit assignment model where the elasticity is originated implicitly by the hierarchical modal choice expressed by the tree structure T

$$
(43) \qquad (\texttt{VIcond-H}) \qquad \left[\begin{array}{l} \mathsf{Find}\ (\mathrm{v},\mathrm{g},\zeta) \in \Omega_G \times S \ \ \mathsf{so\ that,} \\[2ex] 0 \in \begin{pmatrix} \mathrm{T}^d(v,\zeta^d) \\ \mathrm{L}^d(\mathrm{g}) \end{pmatrix} + \mathsf{N}_{\Omega_G^d}(\mathrm{v}^d, \mathrm{g}^d), \ d \in D \\[2ex] 0 \in -x_i^d(\mathrm{v}) + \mathsf{N}_{S_i^d}(\zeta_i^d), \qquad d \in D, \ i \in \hat{N}_d \end{array} \right]
$$

If we write out components of the previous V.I. in variational condition form corresponding to variables g_p^{ω}

(44)

$$
g_{p'}^{\omega} : \begin{cases} \dfrac{\eta_p}{\vartheta} \, \log\!\left(\dfrac{g_{p'}^{\omega}}{g_p^{\omega}}\right) + \tilde{u}_{p'}^{\omega} \ = \beta_p^{\omega} + \mu_{p'}^{\omega}, & \text{if } p' \in L \setminus \mathsf{B}, \ \{p\} = I_T(p') \\[2ex] \dfrac{\eta_p}{\vartheta} \, \log\!\left(\dfrac{g_{\mathsf{B}}^{\omega}}{g_p^{\omega}}\right) \qquad\ \ = \beta_p^{\omega} + \mu_{\mathsf{B}}^{\omega} - \lambda_i^d, & \text{if } p' = \mathsf{B}, \ \{p\} = I_T(\mathsf{B}) \\[2ex] \dfrac{\eta_p}{\vartheta} \, \log\!\left(\dfrac{g_{p'}^{\omega}}{g_p^{\omega}}\right) \qquad\ \ = \beta_p^{\omega} + \mu_{p'}^{\omega} - \beta_{p'}^{\omega}, & \text{if } p' \in C, \ \{p\} = I_T(p') \end{cases} , \ \begin{array}{l} p' \in N_T \\ \omega \in W \end{array}
$$

Lemma 5.7 *Solutions of V.I. (43) are such that* $\mathrm{g} > 0$ *and reproduce a hierarchical logit modal split model.*

Proof: Consider $p' \in E(\mathsf{R})$, $p \neq \mathsf{B}$, then because of (44)

$$
(45) \qquad\qquad\qquad g_{p'}^{\omega} = \exp\left(\vartheta(\beta_{\mathsf{R}}^{\omega} + \mu_{p'}^{\omega} - \tilde{u}_{p'}^{\omega})\right)
$$

Because a solution of V.I. (43) verifies Mangasarian-Fromovitz Constraint Qualification (`MFCQ`) then the Kuhn-Tucker set is bounded (see, for instance Proposition 3.2.1, page 254 in [**8**]) and this implies that $g_{p'}^{\omega} > 0$ and consequently $\mu_{p'}^{\omega} = 0$. If now $\vartheta_p \stackrel{\triangle}{=} \vartheta/\eta_p$, $p \in C$, then for $p' \neq \mathsf{B}$

$$
(46) \qquad\qquad g_{p'}^{\omega} = \exp\left(\vartheta_p(\beta_p^{\omega} - \tilde{u}_{p'}^{\omega})\right), \ p' \in L \cap E_T(p), \ \omega \in W
$$

$$
(47) \qquad\qquad g_{p'}^{\omega} = \exp\left(\vartheta_p(\beta_p^{\omega} - \beta_{p'}^{\omega})\right), \ p' \in C \cap E_T(p), \ \omega \in W
$$

Because of (39) a log-sum relationship applies for node $p \in C$, such that $\mathsf{B} \notin E_T(p)$:

$$
(48) \qquad \exp\left(-\vartheta_p \beta_p^{\omega}\right) = \sum_{p' \in C \cap E_T(p)} \exp\left(-\vartheta_p \beta_{p'}^{\omega}\right) + \sum_{p' \in L \cap E_T(p)} \exp\left(-\vartheta_p \tilde{u}_{p'}^{\omega}\right)
$$

and for $p \in C$, $\mathsf{B} \notin E_T(p)$:

(49)
$$\frac{g_{p'}^{\omega}}{g_p^{\omega}} = \frac{\exp\left(-\vartheta_p \beta_{p'}^{\omega}\right)}{\exp\left(-\vartheta_p \beta_p^{\omega}\right)} = \frac{\exp\left(-\vartheta_p \beta_{p'}^{\omega}\right)}{\sum_{p' \in C \cap E_T(p)} \exp\left(-\vartheta_p \beta_{p'}^{\omega}\right) + \sum_{p' \in L \cap E_T(p)} \exp\left(-\vartheta_p \tilde{u}_{p'}^{\omega}\right)} \ ,$$

whereas for node B, the corresponding term $\beta_{\mathsf{B}}^{\omega}$ is the o-d travel time λ_i^d, $\omega = (i, d) \in W$, as stated in (44)

(50)
$$\frac{g_{\mathsf{B}}^{\omega}}{g_p^{\omega}} = \frac{\exp\left(-\vartheta_p \beta_{\mathsf{B}}^{\omega}\right)}{\exp\left(-\vartheta_p \beta_p^{\omega}\right)} \ , \ \{p\} = I_T(\mathsf{B}) \qquad \square$$

Lemma 5.8 *The jacobian* $\left(\dfrac{\partial \mathrm{L}^{\omega}}{\partial \mathrm{g}^{\omega}}\right)$ *is positive definite.*

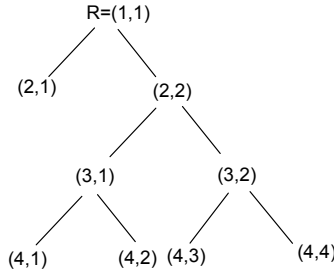*Proof:* Without loss of generality it will suffice to examine the example shown in figure 8 below:

FIGURE 8. A sample hierarchical tree.

(51)
$$\left(\frac{\partial L^{\omega}}{\partial \mathrm{g}^{\omega}}\right) = \frac{1}{\vartheta}
\begin{pmatrix}
\frac{1}{g^{2,1}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{g^{2,2}} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{-\eta_{2,2}}{g^{2,2}} & \frac{\eta_{2,2}}{g^{3,1}} & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{-\eta_{2,2}}{g^{2,2}} & 0 & \frac{\eta_{2,2}}{g^{3,2}} & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{-\eta_{3,1}}{g^{3,1}} & 0 & \frac{\eta_{3,1}}{g^{4,1}} & 0 & 0 & 0 \\
0 & 0 & \frac{-\eta_{3,1}}{g^{3,1}} & 0 & 0 & \frac{\eta_{3,1}}{g^{4,2}} & 0 & 0 \\
0 & 0 & 0 & \frac{-\eta_{3,2}}{g^{3,2}} & 0 & 0 & \frac{\eta_{3,2}}{g^{4,3}} & 0 \\
0 & 0 & 0 & \frac{-\eta_{3,2}}{g^{3,2}} & 0 & 0 & 0 & \frac{\eta_{3,2}}{g^{4,4}}
\end{pmatrix}$$

where the superscript ω has been omitted for clarity. Because of lema 5.7 all $g^{(m,n)} > 0$, thus implying that $g^{(m,n)} < g^{(m-1,n')}$ and always the previous matrix is diagonally dominant and thus definite positive:

$$\frac{\eta_{m-1,n'}}{g^{(m,n)}} \; > \; \left| \; \frac{-\eta_{m-1,n'}}{g^{(m-1,n')}} \; \right| \quad \square$$

6 Conclusions

Service setting models for public transportation lines in congested situations have been presented under two different passenger transit assignment approaches. The first model assumes that passengers make a choice accordingly to a user equilibrium principle following no recommendation but without assuming possible strategies. The second model assumes that passengers follow a recommendation based on a shortest congested route. Both models have been formulated as nonlinear mixed integer programming problems. The first one by means of simulated annealing and the second one by means of an *ad hoc* developed heuristic. Also, the congested transit assignment model based on strategies developed in [**6**] and its formulation in V.I. in [**2**] has been briefly introduced. Computational results using an MSA algorithm have been presented on a small test network with strict capacities. The adaptation of the model to the elastic demand case is is also described.

References

[1] M. Cepeda, R. Cominetti and M. Florian, *A frequency-based assignment model for congested transit networks with strict capacity constraints : characterization and computation of equilibria*, Transportation Research B **40** (2006) 437-459 .

[2] E. Codina, *A variational inequality reformulation of a congested transit assignment model by Cominetti, Correa, Cepeda and Florian. Transportation Science.* To appear. (2012)

[3] E. Codina, A. Marín, Link restoration for rapid transit network incidences. Preprints of the XXIII EURO Conference, Bonn, Germany (2009)

[4] E. Codina, A. Marín, F. López, Bus complementary line design for disruptions in rapid transit networks. Preprints of the World Conference on Transportation Research, Lisbon, Portugal (2010)

[5] E. Codina, A. Marín, F. López, A model for sertting services on auxiliary bus lines under congestion. To appear in TOP (2012)

[6] R. Cominetti and J. Correa, *Common-lines and passenger assignment in congested transit networks*, Transportation Science 35 (**3**) (2001) 250-267.

[7] I. Constantin and M. Florian (1995) Optimizing frequencies in a transit network : a non-linear bi-level programming approach. Int. Trans. Opl. Res. Vol. 2 p 149-164.

[8] F. Facchinei and J-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems. Vol I, II*, Springer Series in Operations Research (2002)

[9] T. Haupt, K. Nökel and U. Reiter (2009) *Schedule based dynamic transit assignment.* In Schedule-based modeling of transportaion networks, N.H.M. Wilson, A. Nuzzolo Ed. Springer.

[10] F. Kurauchi, M.G.H. Bell, J-D. Schmöcker (2003) *Capacity constrained transit assignment with common lines. Journal of Mathematical modelling and algorithms* 2, 309-327.

[11] H.X. Liu, X. He and B. He, *Method of successive weighted averages and self-regulated averaging schemes for solving stochastic user equilibrium problems. Networks and Spatial Economics* **9** (2009) 485-503.

[12] F. López, E. Codina, A. Marín, Integrating network design and line planning in rapid transit systems. In review process. Journal of Public Transport.

[13] Noriega Y. y M. Florian (2003) *L'optimisation des fréquences d'un réseau de transport en commun moyennement congestionné. INFOR* vol 41, no **2**. 129-153.

[14] Papola N., F. Filippi, G.Gentile and L.Meschini (2009) *Schedule based transit assignment: new dynamic equilibrium model with vehicle capacity constraints.* In Schedule-based modeling of transportation networks, N.H.M. Wilson, A. Nuzzolo Ed. Springer.

[15] H. Spiess and M. Florian, *Optimal strategies: a new assignment model for transit networks*, Transportation Research Part B 23 (**2**) (1989) 83-102

7 Appendix

7.1 Mangasarian-Fromovitz constraint qualification

Assume a (generally non-convex) set $X \subset \mathbb{R}^n$ with a finite representation as:

$$(52) \qquad X = \{\, x \in \mathbb{R}^n \,|\, h(x) = 0,\ g(x) \geq 0 \,\}$$

where $h : \mathbb{R}^n \to \mathbb{R}^\ell$, $g : \mathbb{R}^n \to \mathbb{R}^m$, are continuously differentiable functions. By definition, the Mangasarian Fromovitz Constraint Qualification, abbreviated as `MFCQ`, holds at a vector $x \in X$ if

  (1) the gradients $\nabla h_i(x),\ i = 1, ..., \ell$ are linearly independent, and
  (2) there exists a vector $u \in \mathbb{R}^n$ such that

$$(53) \qquad \begin{aligned} \nabla h_i(x)^\top u &= 0, \quad 1 \leq i \leq \ell \\ \nabla g_j(x)^\top u &> 0, \quad j \in I(x) \end{aligned}$$

being $I(x) \overset{\Delta}{=} \{\, 1 \leq k \leq m \mid g_k(x) = 0 \,\}$ the set of active indexes at x.

Any point in a simplex $S = \{\, x \in \mathbb{R}^n_+ \mid \sum_{k=1}^n x_k = 1 \,\}$ trivially satisfies `MFCQ`.

A relevant characteristic of a V.I. with continuous operator defined on a set X whose points verify `MFCQ` is that multipliers for the corresponding constraints are bounded (see Facchinei and Pang (2002), proposition 3.2.1 (b)). Let $M(x)$ the polyhedral set of multipliers for a V.I. $0 \in F(x) + \mathsf{N}_X(x)$ defined for a point $x \in X$ as follows

$$(54)$$

$$M(x) \overset{\Delta}{=} \{\, (u,v) \in \mathbb{R}^{\ell+m} \mid F(x) = \sum_{i=1}^{\ell} u_i \nabla h_i(x) + \sum_{j=1}^{m} v_j \nabla g_j(x),\ g(x)^\top v = 0,\ v_j \geq 0 \,\}$$

If F is continuous on X, and `MFCQ` hold at $x \in X$ then $M(x)$ is a bounded set.

7.2 Properties of H

Let $H_n(\cdot)$, be the function on $\mathbb{R}^n$, given by

$$(55) \qquad H_n(x) = \underset{1 \leq \ell \leq n}{\mathsf{Max}} \{\, x_\ell \,\}$$

Property 7.1 *Let $I(x) = \{\, 1 \leq \ell \leq n \mid x_\ell = H_n(x) \,\}$. Then, Clarke's subgradient of H_n at $x \in \mathbb{R}^n_+$ is*

$$(56) \qquad \partial H(x) = \left\{\, \alpha \in \mathbb{R}^n_+ \mid \sum_{\ell \in I(x)} \alpha_\ell = 1,\ \alpha_\ell = 0 \text{ if } \ell \notin I(x) \,\right\},\ x \in \mathbb{R}^n_+$$

Property 7.2 *The function H is homogeneous of degree 1, i.e. it verifies the differential inclusion:*

$$(57) \qquad \sum_{\ell=1}^{n} x_\ell \zeta_\ell = H(x),\ \ \zeta \in \partial H(x)$$

Proof: (of properties 7.1 and 7.2) Function $H(x)$ can be expressed as the optimal value function of the linear program

$$
\begin{aligned}
(58) \qquad H(x) \quad &= \mathsf{Min}_{w,s} \quad w \\
&\quad s.t. : \quad w - s_\ell = x_\ell \ \mid \zeta_\ell, \ \ \ell = 1, 2, ...n \\
(\mathtt{P}) \quad & \qquad\qquad s_\ell \geq 0
\end{aligned}
$$

with dual (D) given by:

$$
\begin{aligned}
(59) \qquad H(x) \quad &= \mathsf{Max}_{\zeta} \quad \sum_{\ell=1}^{n} x_\ell \zeta_\ell \\
(\mathtt{D}) \quad s.t. : \quad & \sum_{\ell=1}^{n} \zeta_\ell = 1 \\
& \quad \zeta_\ell \geq 0, \ \ \ell = 1, 2, ...n
\end{aligned}
$$

This result clearly states that H is the dual support function of the simplex S_n in $\mathbb{R}^n$ and that ∂H is the set defined in (56). See, for instance Corollary 8.25 in Rockafellar and Wets (1998). Because of complementary slackness it must be verified that $(w^* - x_\ell^*)\zeta_\ell^* = 0$ or equivalently:

$$
(60) \qquad\qquad (H(x) - x_\ell)\zeta_\ell = 0, \ \ell = 1, 2, ..., n
$$

which is precisely (57) as $\sum_\ell \zeta = 1$. $\square$

For simplicity this section will be restricted to $x \in \mathbb{R}_+^n$. Typically, two smoothing family of functions to H are:

$$
(61) \qquad\qquad \phi(x, t) = \left(\sum_{\ell=1}^{n} x_\ell^{1/t} \right)^t, \ t > 0
$$

$$
(62) \qquad \psi(x, t) = \log \phi(\exp_c(x), t) = t \log \left(\sum_{\ell=1}^{n} \exp\left(x_\ell/t\right) \right), \ t > 0
$$

here $\exp_c(x)$, $x \in \mathbb{R}$ denotes the componentwise exponential function, i.e.: $\exp_c(x) = (\exp(x_1), ..., \exp(x_n))$. This notation will be extended to any scalar function $h : \mathbb{R} \to \mathbb{R}$: i.e. $h_c(x) = (h(x_1), ..., h(x_n))^\top$.

Functions $\phi(\cdot, t)$ and $\psi(\cdot, t)$ are strictly convex in x for $t > 0$ and their gradients have the following formulas:

$$
(63) \qquad\qquad (\nabla_x \phi(x, t))_\ell = \left(\frac{x_\ell}{\phi(x, t)} \right)^{\frac{1}{t} - 1}, \ \ \ell = 1, 2, ...n
$$

$$
(64) \qquad (\nabla_x \psi(x, t))_\ell = \frac{exp\left(x_\ell/t\right)}{\sum_{j=1}^{n} \exp\left(x_j/t\right)} = \left(\frac{\exp\left(x_\ell\right)}{\exp(\psi(x, t))} \right)^{\frac{1}{t}}, \ \ \ell = 1, 2, ...n
$$

Property 7.3 *Let $\{x^{(k)}\} \to x$ and let $\{t^{(k)}\} \to 0+$, then both $\psi(x^{(k)}, t^{(k)}) \to H(x)$ and $\phi(x^{(k)}, t^{(k)}) \to H(x)$*

Property 7.4 *(See lemma 2.1,* vii *in Peng (1999)), p. 296) If $\{x^{(k)}\}$ and $\{t^{(k)}\}$ are convergent sequences, $\{x^{(k)}\} \to \bar{x}$ and $\{t^{(k)}\} \to 0+$, then also for both approximations:*

(65)
$$\ell im_{k\to\infty} d(\ \nabla_x \psi(x^{(k)}, t^{(k)}), \partial H(\bar{x})\ ) = 0$$

(66)
$$\ell im_{k\to\infty} d(\ \nabla_x \phi(x^{(k)}, t^{(k)}), \partial H(\bar{x})\ ) = 0$$

Proof: for (65) see Peng (1999), p. 296. For (66) just notice that as e^x is an increasing function, then $I(exp_c(\cdot)) \equiv I(\cdot)$ and this implies that $\partial H(exp_c(x)) = \partial H(x)$. Now let the sequence $\{z^{(k}\}$ be defined as $z^{(k} = log_c(x^{(k)})$ and let $z = log_c(x)$; then

$$d(\ \nabla_x \phi(z^{(k}, t^{(k)}), \partial H(z)\ ) = d(\ \nabla_x \phi(z^{(k}, t^{(k)}), \partial H(exp_c(z))\ ) = d(\ \nabla_x \psi(x^{(k)}, t^{(k)}), \partial H(x)\ ) \to 0. \ \square$$

Property 7.5 *Let $x \in \mathbb{R}^n_+$. For any $\eta \in \partial H(x)$, $\exists \{(x_k, t_k)\} \overset{t \geq 0}{\to} (x, 0)$, $x_k \geq 0$, so that $\lim_{k\to\infty} \nabla_x \psi(x_k, t_k) = \eta$. The same property applies for function ϕ.*

Property 7.6 *Complementary slackness condition (60), or equivalently (57), is approximately verified by $\phi(x, t)$ at $x \in \mathbb{R}^n_+$ and $t > 0$:*

Proof:
$$\mathtt{CS}_\ell(x, t) \overset{\Delta}{=} (\phi(x, t) - x_\ell)(\nabla_x \phi(x, t))_\ell = (\phi(x, t) - x_\ell) \left(\frac{x_\ell}{\phi(x, t)} \right)^{\frac{1}{t} - 1} =$$
(67)
$$= x_\ell \left(1 - \frac{x_\ell}{\phi(x, t)} \right) \left(\frac{x_\ell}{\phi(x, t)} \right)^{\frac{1}{t} - 2} \leq x_\ell \left(1 - \frac{x_\ell}{\phi(x, t)} \right) \left(\frac{x_\ell}{\phi(x, t)} \right)$$

Clearly, if $\ell \in I(x)$ then $\ell im_{t\to 0+} \phi(x, t) = x_\ell$ and if $\ell \notin I(x)$, then

$$\ell im_{t\to 0+} \left(\frac{x_\ell}{\phi(x, t)} \right)^{\frac{1}{t} - 2} = 0$$

Thus $\ell im_{t\to 0+} \mathtt{CS}_\ell(x, t) = 0$, $1 \leq \ell \leq n \ \square$

Property 7.7 *Complementary slackness condition (60), or equivalently (57), is approximately verified by $\psi(y, t)$ at $y \in \mathbb{R}^n_+$ and $t > 0$:*

Proof: Using $\mathtt{CS}$ for ϕ defined in (67),

(68)
$$\mathtt{CS}_\ell(\ exp_c(y), t) \overset{\Delta}{=} (\exp(\psi(y, t)) - \exp(y_\ell)) \left(\frac{\exp(y_\ell)}{\exp(\psi(y, t))} \right)^{\frac{1}{t} - 1}, \quad 1 \leq \ell \leq n$$

and this proofs that the following $\mathtt{CS}$' will be approximately zero at $t \to 0+$.

(69)
$$\mathtt{CS'}_\ell(y, t) \overset{\Delta}{=} (\psi(y, t) - y_\ell) \left(\frac{\exp(y_\ell)}{\exp(\psi(y, t))} \right)^{\frac{1}{t}} \overset{t \to 0+}{\longrightarrow} 0, \quad 1 \leq \ell \leq n \ \square$$