Mixed Equilibrium for Multiclass Routing Games*

T. Boulogne[1,2] , E. Altman[1,3], H. Kameda[4], O. Pourtallier[1]

Abstract

We consider a network shared by non-cooperative two types of users, group users and individual users, each user of the first type has a significant impact on the load of the network, whereas a user of the second type doesn't. Both group users as well as individual users choose their routes so as to minimize their costs. We further consider the case that the group users may have side constraints. We study the concept of mixed equilibrium (mixing of Nash equilibrium and Wardrop equilibrium). We establish its existence and some conditions for its uniqueness. Then we apply the mixed equilibrium to a parallel links network and to a case of load balancing.

Keywords

Game Theory, routing, networks, Nash equilibrium, Wardrop equilibrium, mixed-equilibrium, side constraints.

I. Introduction

We consider in this paper the problem of optimal routing in networks. The entity that is routed is called a *job*. There are infinitely many jobs to ship from a source to a destination (sources, so as destinations, may be different according to the jobs). The decision maker is called a user, there exist two types of users, group users and individual users. A group user has a large amount of jobs to ship, while an individual user has only one job to ship. Each user has its own source(s) and destination(s), its own link costs functions, and its own optimization criterion. We further consider allow for side constraints, and even more generally, we consider a setting in which the space of decisions of group users is not orthogonal (see [21] for a similar setting in the case of group equilibrium only).

We group the individual users into classes, and we call also a group user a class. Then there are several classes of jobs. Each class corresponds to a large number of single jobs. In each class the routes to be taken by the jobs of that class are determined either by a decision maker that centralizes all decisions for that class, or they are done individually by each individual user. We call the first type of per-class decision making a class-centralized optimization, and the second approach a class-individual optimization.

When all class use a class-individual optimization approach then the natural optimization concept is the Wardrop equilibrium [22]. This concept was very much studied (e.g. [3], [5], [6], [9], [20] and references therein). Most of the work with this optimization approach has been done in the framework of road traffic. However, this concept has been also useful in the area of distributed computing [14], [15], and in telecommunication networks [7]. In the context of road traffic, an individual user (a "job" in our terminology) may correspond to a single driver, and the class may correspond to all the drivers of a given type of vehicle that have a given source and destination. In the context of distributed computing, a user may correspond to a single job that is sent to be processed at some computer in a computer-network. Finally, in the context of telecommunications, a single user may correspond to a single packet in networks in which the delay of each packet is minimized [7]. A generalized version of the Wardrop equilibrium which involves side constraints has been studied in [18] and references therein.

When all class use a class-centralized optimization approach then the optimization concept is the Nash equilibrium. There has been much recent interest in this framework in recent years [1], [2], [3], [11], [16], [17], [19]. In the context of road traffic, a class, or a group user, may correspond to a transportation company, or to a bus company; in both examples we may assume that the route of each vehicle is indeed determined by the company and not by the individual driver.

The concept of mixed-equilibrium (M.E.) has been introduced by Harker [8] (and further applied in [23] to a dynamic equilibrium and in [12] to a specific load balancing problem with a completely symmetrical network). Harker has established the existence of the M.E., characterized it through variational inequalities, and gave conditions for its uniqueness.

The first part of the paper consists of the mathematical model and the definition of mixed-equilibria (Sec. II), then Sec. III-IV establish the existence of equilibria under different approaches and assumptions, and Sec. V, VI derive uniqueness conditions under conditions related to strict monotonicity. This part of our paper extends Harker's model [8] in several directions: (i) A general cost function is considered, rather than the separable cost function given as the sum of link costs in [8]. This allows one to model routing games in which the performance measures are rejection probabilities of calls or loss probabilities of packets. Our general cost allows in particular different users to have different costs for the same links or the same paths, which allows to model priorities. In some cases we explicitly introduce the term of per-user "service-rate" for this purpose. (ii) We obtain existence and some uniqueness results for the case where the decisions of group users are constrained. This allows one to model side constraints,

*This work was partially supported by a research contract with France Telecom R& D No. 001B001

[1] INRIA, 2004 route des Lucioles, 06902 Sophia Antipolis, France. Tel: +33-4-92-38-77-77 Fax: +33-4-92-38-77-65
E-mail: {thomas.boulogne,eitan.altman,odile.pourtallier}@sophia.inria.fr

[2]Université Pierre et Marie Curie - Paris 6, 4 place Jussieu 75230 Paris, France.

[3]C.E.S.I.M.O., Universidad de Los Andes, Faculdad de Ingeneria, Merida, VENEZUELA

[4] Institute of Information Sciences and Electronics, University of Tsukuba, Tsukuba Science City, Ibaraki 305-8573, Japan. Tel: +81-298-53-5539 Fax: +81-298-53-5206 E-mail: kameda@is.tsukuba.ac.jp.

and to consider multiobjective problems faced by the users. For example, a group user might wish to find a strategy that minimizes its delay, and at the same time constraining its average loss probability to be below some bound.

In the second part of the paper we obtain new sets of conditions for the uniqueness of the mixed equilibrium for the case where conditions of the type of strict monotonicity (such as those that are used in [8]) do not apply. Some of the new conditions are obtained by making further assumptions on the structure of possible equilibria (Sec. VII) and others are obtained for specific topologies (Sec. VIII): the parallel link topology, and load balancing models.

II. Mixed Equilibrium (M.E.): Model and Assumptions

We consider a general network. We denote $\mathcal{M}$ the set of nodes, and $\mathcal{L} \subset \mathcal{M} \times \mathcal{M}$ the set of unidirectional links [1]. The unit entity that is routed through the network is called a *job*.

Each job j has an origin-destination pair (*O-D pair*) as well as a service rate vector, μ^j. We denote the origin, or the source by $s(j)$ and the destination by $d(j)$; $\mu^j = (\mu^j_l, l = 1, \ldots, L)$ is an L-vector (L is the number of elements of the set $\mathcal{L}$, *i.e.* $L = \#\mathcal{L}$). The interpretation of μ^j_l can be the speed at which job j is processed in link l and we assume that $\mu^i_l > 0, \ \forall i \in \mathcal{I}, \ \forall l \in \mathcal{L}$.

Each user u has a certain amount of jobs to route from a source s to a destination d, we call this amount the flow demand of the user u for the O-D pair (s, d) and we denote it by $\phi^u_{(s,d)}$.

The network is used by two types of users.

The first type of users, referred to as *group users* have to route a large amount of jobs. The choices made by each of these users have a significant impact on the load of the network, and then on the delays that any other user can expect. We denote by $\mathcal{N}$ the set of group users. Each user $i \in \mathcal{N}$ is characterized by

- one service rate vector $\mu^i = (\mu^i_l)_{l\in\mathcal{L}}$,
- a *set* of O-D pairs D^i and
- a vector of demands $\phi^i = \left(\phi^i_{(s,d)}\right)_{(s,d)\in D^i}$, $\phi^i_{(s,d)}$ denotes the rate of jobs of this class that have to be shipped from s to d.

(Note that having several sources and destinations allows in particular to handle multicast applications, in which several destinations are associated with a single source).

The second type of users, referred as *individual users* have a single job to route through the network from a given source to a given destination, with a given service rate. There are infinitely many individual users and the routing choice of a single individual user has a negligible impact on the load of the system. Individual users can be classified according to the pair source-destination and the service rate associated to their jobs. We denote $\mathcal{W}$ the set of classes of individual users. Each class i of the second type is characterized by

- one O-D pair (s^i, d^i),
- one service rate vector $\mu^i = (\mu^i_l)_{l\in\mathcal{L}}$ and
- one flow demand ϕ^i (the "number" of users belonging to class i).

Note that since all jobs of class i have the same service rate vector, we shall use μ^i to denote the service rate vector of any one of the jobs that belong to class i.

Note also that the elements of the set $\mathcal{N}$ or of the set $\mathcal{W}$ can be considered as a class of jobs characterized by a set of pair(s) of source and destination and a service rate. Nevertheless the routing decision of all the jobs of $i \in \mathcal{N}$ is taken by a single decision maker, while the routing decision of any single job of $i \in \mathcal{W}$ is taken by the individual user who is paired with it.

We denote by $\mathcal{I}$ the set of all possible classes of jobs, $\mathcal{I} = \mathcal{N} \cup \mathcal{W}$, and assume that $\mathcal{I}$ is finite.

A path p from $s \in \mathcal{M}$ to $d \in \mathcal{M}$ is a sequence of directed links that goes from s to d. For $i \in \mathcal{I}$ we denote by $\mathcal{P}^i$ the set of possible paths for class i, by $\mathcal{P}^i_{(u,v)}$ the set of possible paths for class i which go from u to v, $\mathcal{P}^i = \cup_{(u,v)\in D^i}\mathcal{P}^i_{(u,v)}$ and by $\mathcal{P}$ the set of all possible paths, $\mathcal{P} = \cup_{i\in\mathcal{I}}\mathcal{P}^i$.

In this paper we try to work as much as possible on paths (*i.e.*, the decision is what fraction of traffic of each class has to be routed over each path; this is in contrast to the more restrictive models such as [19] in which the routing decisions are how much jobs to route to each outgoing link of each node; this second type of models implicitly assumes that all sequences of directed links that lead from a source to a destination are admissible paths). Nevertheless it will be necessary, sometimes, to work on link models, *i.e.*, at each node we shall allow each class to route all the flow that it sends through that node to any of the out-going links of that node. So we introduce two notations for the flows, one in term of paths and one in term of links.

Each decision maker (a class within $\mathcal{N}$ or an individual user belonging to some class in $\mathcal{W}$) has to choose a (set of) path(s) to route its job(s). For $i \in \mathcal{I}$ and $p \in \mathcal{P}^i$ (*resp.* $l \in \mathcal{L}$), we denote by $x^i_{(p)}$(*resp.* x^i_l) the amount of jobs sent through path p (*resp.* link l) by class i. Note again that the meaning of $x^i_{(p)}$ is slightly different according to whether i belongs to the set $\mathcal{N}$ or $\mathcal{W}$. If $i \in \mathcal{N}$, then $x^i_{(p)}$ is the amount of jobs of user $i \in \mathcal{N}$ sent through p, if $i \in \mathcal{W}$, $x^i_{(p)}$ represents the amount of individual users of class $i \in \mathcal{W}$ that choose path p to ship their unique job.

Depending on the context, we will denote by x^i, the strategy of class i, either the vector $(x^i_{(1)}, \ldots, x^i_{(P^i)})^T$ of path flows, or the vector $(x^i_1, \ldots, x^i_L)^T$ of link flows, where P^i (*resp.* L) is the number of paths (*resp.* links) in the set $\mathcal{P}^i$ (*resp.* $\mathcal{L}$).

[1] A bidirectional link may be transformed into a network of unidirectional ones where some are of null cost (Appendix B), then the results presented in this paper are also valid in networks with both unidirectional and bidirectional links unless the assumptions impose that the links' cost functions are strictly increasing.

Let $\mathbf{x}$ be the flow configuration, *i.e.*, $\mathbf{x}^T$ is the vector $(x^1,\ldots,x^I)$, where $I=\#\mathcal{I}$, and $\mathcal{X}$ be the set of possible $\mathbf{x}$ (the "total" strategy set).

It will sometimes be necessary to distinguish in a routing profile $\mathbf{x}$ of $\mathcal{X}$ the part due to the group users, and the part due to the classes of individual users. We will then write $\mathbf{x}=\begin{pmatrix}\mathbf{x}^{\mathcal{N}}\\ \mathbf{x}^{\mathcal{W}}\end{pmatrix}$, where $(\mathbf{x}^{\mathcal{N}})^T=(x^1,\ldots,x^N)\in\mathcal{X}^{\mathcal{N}}$ corresponds to the choice of the group users, and
$(\mathbf{x}^{\mathcal{W}})^T=(x^1,\ldots,x^W)\in\mathcal{X}^{\mathcal{W}}$ corresponds to the choice of the classes of individual users, where $N=\#\mathcal{N}$ and $W=\#\mathcal{W}$. We assume that $\mathcal{X}=\mathcal{X}^{\mathcal{N}}\times\mathcal{X}^{\mathcal{W}}$, and that both $\mathcal{X}^{\mathcal{N}}$ and $\mathcal{X}^{\mathcal{W}}$ are convex and compact. Note that, as in [21], we do not assume that $\mathcal{X}^{\mathcal{N}}$ has a product form, and thus the policy used by some classes may restrict the policies used by other class. This is a general way of introducing constraints over the policies.

Notations:
ρ_l is the total load on link l, $\rho_l=\sum_{i\in\mathcal{I}}\rho_l^i$, where
$\rho_l^i=\frac{1}{\mu_l^i}\sum_{p\in\mathcal{P}^i}\delta_{lp}x^i_{(p)}=\frac{x_l^i}{\mu_l^i}$, where $\delta_{lp}=\begin{cases}1 & \text{if } l\in p,\\ 0 & \text{otherwise.}\end{cases}$
$\boldsymbol{\rho}$ is the utilization vector which is induced by $\mathbf{x}$, $\boldsymbol{\rho}^T=(\rho_1,\ldots,\rho_L)$,
Φ is the total vector of flow demand, $\Phi=(\phi^i)_{i\in\mathcal{I}}$.
Let "$\cdot$" be the inner product and $\nabla_i:=\frac{\partial}{\partial x^i}$.

Cost functions:

- $J^i:\mathcal{X}\to[0,\infty)$ (or $[0,\infty]$, depending on the context) is the cost function of class $i\in\mathcal{N}$.
- $F^i_{(p)}:\mathcal{X}\to[0,\infty)$ (or $[0,\infty]$, depending on the context) is the cost function of path p for each individual user of class $i\in\mathcal{W}$.

The aim of each user is to minimize its cost (according to the constraint set), *i.e.*, for $i\in\mathcal{N},\ \min_{x^i\in\mathcal{X}^i}J^i(\mathbf{x})$, and for $i\in\mathcal{W},\ \min_{p\in\mathcal{P}^i}F^i_{(p)}(\mathbf{x})$.
Let $\mathcal{P}^{i*}(\mathbf{x})$ be the set of paths for class i which have a flow strictly positive when the strategy of class i is x^i ($\forall p\in\mathcal{P}^{i*}(\mathbf{x}),\ x^i_{(p)}>0$) and let $(\mathbf{x}^{-i},y^i)$ be the flow configuration where class j $(j\neq i)$ uses strategy x^j and class i uses strategy y^i.

Definition: $\mathbf{x}\in\mathcal{X}$ is a *Mixed Equilibrium* (M.E.) if
$\forall i\in\mathcal{N},\ \forall y^i\ s.t.\ (\mathbf{x}^{-i},y^i)\in\mathcal{X},\ J^i(\mathbf{x})\le J^i(\mathbf{x}^{-i},y^i)$ (*Nash equilibrium condition*), and
$\forall i\in\mathcal{W},\ \forall p\in\mathcal{P}^i,\ \forall p^*\in\mathcal{P}^{i*}(\mathbf{x}), F^i_{(p^*)}(\mathbf{x})\le F^i_{(p)}(\mathbf{x})$ (*Wardrop equilibrium condition*).
Remark: Wardrop equilibrium condition is equivalent to

$$F^i_{(p)}(\mathbf{x})-A^i\ge 0;\quad (F^i_{(p)}(\mathbf{x})-A^i)x^i_{(p)}=0 \tag{1}$$

$\forall i\in\mathcal{W}$ and $\forall p\in\mathcal{P}^i$, where $A^i=A^i(\mathbf{x}):=\min_{p\in\mathcal{P}^i}F^i_{(p)}(\mathbf{x})$.

III. Existence of M.E. through variational inequalities

In this section, we present a simple variational inequality method to establish the existence of M.E. in the case of no extra constraints under general conditions on the cost functions for both types of classes. More precisely, Let $n=\sum_{i\in\mathcal{N}}\#\mathcal{P}^i$, $w=\sum_{i\in\mathcal{W}}\#\mathcal{P}^i$, $n+w=\sum_{i\in\mathcal{I}}\#\mathcal{P}^i$ and define the (total) strategy set $\mathcal{X}$ as follows

$$\mathcal{X}=\{\mathbf{x}\in\mathbb{R}^{(n+w)}\mid\forall i\in\mathcal{I},\forall(u,v)\in D^i,\forall p\in\mathcal{P}^i,\ x^i_{(p)}\ge 0,\sum_{p\in\mathcal{P}^i_{(u,v)}}x^i_{(p)}=\phi^i_{(u,v)}\}$$

Assumptions:

- $(\mathcal{A}1)$ $\forall i\in\mathcal{N},\ J^i(\mathbf{x}):\mathcal{X}\to[0,\infty)$ is convex in x^i and continuously differentiable w.r.t. $x^i_{(p)},\ \forall p\in\mathcal{P}^i$
- $(\mathcal{A}2)$ $\forall i\in\mathcal{W},\forall p\in\mathcal{P}^i\quad F^i_{(p)}(\mathbf{x}):\mathcal{X}\to[0,\infty)$ is continuous.

For every $i\in\mathcal{N}$ we denote the derivative of $J^i(\mathbf{x})$ with respect to $x^i_{(p)}$ $K^i_{(p)}(\mathbf{x})$, *i.e.*, $K^i_{(p)}(\mathbf{x})=\frac{\partial}{\partial x^i_{(p)}}J^i(\mathbf{x})$

Let us now reformulate the mixed equilibrium conditions. $\mathbf{x}\in\mathcal{X}$ is a M.E. if and only if $\mathbf{x}$ satisfies

- $\exists\boldsymbol{\alpha}=\boldsymbol{\alpha}(x)$, $\boldsymbol{\alpha}=(\alpha^i_{(u,v)})_{i\in\mathcal{N},(u,v)\in D^i}$, such that $\forall i\in\mathcal{N},\ \forall(u,v)\in D^i$ and $\forall p\in\mathcal{P}^i_{(u,v)}$

$$K^i_{(p)}(\mathbf{x})-\alpha^i_{(u,v)}\ge 0;\quad (K^i_{(p)}(\mathbf{x})-\alpha^i_{(u,v)})x^i_{(p)}=0\quad\text{and}^2 \tag{2}$$

- $\forall i\in\mathcal{W}$ and $\forall p\in\mathcal{P}^i$

$$F^i_{(p)}(\mathbf{x})-A^i\ge 0;\quad (F^i_{(p)}(\mathbf{x})-A^i)x^i_{(p)}=0. \tag{3}$$

[2](2) are Kuhn-Tucker conditions where $\alpha^i_{(u,v)}$ is the Lagrange multiplier associated to the constraint $\phi^i_{(u,v)}=\sum_{p\in\mathcal{P}^i_{(u,v)}}x^i_{(p)}$.

where $A^i = A^i(\mathbf{x}) := \min_{p\in\mathcal{P}^i} F^i_{(p)}(\mathbf{x})$.

Denote by $K(\mathbf{x})$ the n-dimensional vector $K(\mathbf{x}) = (K^i_{(p)}(\mathbf{x}))_{i\in\mathcal{N},p\in\mathcal{P}^i}$,
by $F(\mathbf{x})$ the w-dimensional vector $F(\mathbf{x}) = (F^i_{(p)}(\mathbf{x}))_{i\in\mathcal{W},p\in\mathcal{P}^i}$,
by $T(\mathbf{x})$ the $(n+w)$-dimensional vector $T(\mathbf{x}) = \begin{pmatrix} K(\mathbf{x}) \\ F(\mathbf{x}) \end{pmatrix}$,
by $\mathcal{A}$ the $\left(\sum_{i\in\mathcal{N}} \# D^i + W\right)$-dimensional vector $\mathcal{A} = \begin{pmatrix} \boldsymbol{\alpha} \\ (A^i)_{i\in\mathcal{W}} \end{pmatrix}$,
and by Δ the incidence matrix (see Appendix A).

Then we have

Lemma III.1: Assume $\mathcal{A}1$-$\mathcal{A}2$. $\mathbf{x}\in\mathcal{X}$ is a M.E. if and only if $\mathbf{x}$ satisfies

$$T(\mathbf{x}) - \Delta\mathcal{A} \geq 0, \qquad (T(\mathbf{x}) - \Delta\mathcal{A})\cdot\mathbf{x} = 0, \tag{4}$$

$$\Delta^T\mathbf{x} = \Phi, \qquad \mathbf{x}\geq 0. \tag{5}$$

Proof: We have just to note that the conditions (5) are equivalent to $\mathbf{x}\in\mathcal{X}$. $\square$

Lemma III.2: Assume $\mathcal{A}1$-$\mathcal{A}2$. $\mathbf{x}\in\mathcal{X}$ is a M.E. if and only if

$$T(\mathbf{x})\cdot(\mathbf{y}-\mathbf{x}) \geq 0 \quad \forall \mathbf{y}\in\mathcal{X}. \tag{6}$$

Proof: Similar to the proof of [3, Lem. 3.2] ((6) holds if and only if $\mathbf{x}$ is solution of the linear program $\min_{\mathbf{y}} T(\mathbf{x})\cdot\mathbf{y}$, s.t. $\Delta^T\mathbf{y} = \Phi,\ \mathbf{y}\geq 0$). $\square$

Theorem III.3: Assume $\mathcal{A}1$-$\mathcal{A}2$. Then there exists a mixed equilibrium.

Proof: $\mathcal{X}\subset\mathbb{R}^n$ is a nonempty, bounded, convex set, $T:\mathcal{X}\to\mathbb{R}^n$ is a continuous mapping on $\mathcal{X}$, then there exists a solution to (6) (see [4]). $\square$

IV. Existence of M.E.: a fixed point approach

In this section we relax the assumptions on the cost functions of the group users but restrict the cost functions of the individual users. With these new assumptions we obtain the existence of the M.E. (in a setting that allows one to include extra constraints) using the following approach. It is well known that one may compute the Wardrop equilibrium by transforming the problem into an equivalent optimization problem (as if there were only one decision maker) by transforming the costs in the network, see [20] and references therein. In our setting of M.E. we shall thus (i) transform in a similar way the optimization problem faced by all individual users into an optimization problem of a new equivalent single group user by transforming the cost in a similar way as is done for the Wardrop equilibrium. (ii) Then we will be faced with a game problem of group users only, for which we shall use Rosen's existence theorem [21].

Let f_l be the cost function of the link l, $f_l : (0,\infty)\to(0,\infty]$, this function is used only for the individual users. For any individual user $i\in\mathcal{W}$, we defined the cost function of path $p\in\mathcal{P}^i$ as follows

$$F^i_{(p)}(\mathbf{x}) = \sum_{l\in\mathcal{L}} \frac{\delta_{lp}}{\mu^i_l} f_l(\rho_l) = \sum_{l\in p} \frac{1}{\mu^i_l} f_l(\rho_l)$$

Assumptions:

- $(\mathcal{A}'1)$ $\mathcal{X}$ is a nonempty, convex and compact subset of $\mathbb{R}^n$
- $(\mathcal{A}'2)$ $\mathcal{X}^{\mathcal{W}} = \{\mathbf{x}^{\mathcal{W}}\in\mathbb{R}^w \mid \forall i\in\mathcal{W}, \forall p\in\mathcal{P}^i,\ x^i_{(p)}\geq 0,\ \sum_{p\in\mathcal{P}^i} x^i_{(p)} = \phi^i\}$
- $(\mathcal{A}'3)$ $J^i(\mathbf{x}) : \mathcal{X}\to[0,\infty]$ is a continuous function of $\mathbf{x}$ and is convex in x^i,
- $(\mathcal{A}'4)$ f_l is continuous and increasing in ρ_l,
- $(\mathcal{A}'5)$ for every system flow configuration $\mathbf{x}$, if not all costs of group users are finite then at least one class, $i\in\mathcal{N}$ with infinite cost can change its own flow configuration to make its cost finite, and similarly an individual user has always a path of finite cost that it can use.

These assumptions will be imposed in the rest of the paper. They imply that the policies of the group users may be constrained, since we do not assume that $\mathcal{X}^{\mathcal{N}}$ is an orthogonal set. Thus the choice of policies by some group users may restrict the set of policies available to other group users.

Define $\tilde{f}(\mathbf{s},\mathbf{t})$ where $\mathbf{s}\in\mathcal{X}^{\mathcal{N}}, \mathbf{t}\in\mathcal{X}^{\mathcal{W}}$ by

$$\tilde{f}(\mathbf{s},\mathbf{t}) = \frac{1}{\Phi}\left[\sum_{l\in\mathcal{L}} \int_0^{\rho_l} f_l(\tau)d\tau\right],$$

where

$$\rho_l = \sum_{i=1}^{N} \frac{1}{\mu^i_l}\left(\sum_{p\in\mathcal{P}^i} \delta_{lp} s^i_{(p)}\right) + \sum_{i=N+1}^{N+W} \frac{1}{\mu^i_l}\left(\sum_{p\in\mathcal{P}^i} \delta_{lp} t^i_{(p)}\right),\ \delta_{lp} = \begin{cases} 1 & \text{if } l\in p, \\ 0 & \text{otherwise.}\end{cases}$$

Observe that $\tilde{f}$ is convex in $\mathbf{t}$ (f_l is increasing) and continuous in $\mathbf{s}$ and $\mathbf{t}$.

Introduce the convex minimization program

$$\min \tilde{f}(\mathbf{s},\mathbf{t}) \text{ with respect to } \mathbf{t} \in \mathcal{X}^{\mathcal{W}}. \tag{7}$$

Lemma IV.1: Either the convex program (7) has an optimal solution which satisfies condition (1) where $\mathbf{x} = (\mathbf{s},\mathbf{t})^T$ or $\forall \mathbf{t} \in \mathcal{X}^{\mathcal{W}},\ \tilde{f}(\mathbf{s},\mathbf{t}) = +\infty$.

Proof: We have $F^i_{(p)}(\mathbf{s},\mathbf{t}) = \frac{\partial}{\partial t^i_{(p)}}(\Phi \tilde{f}(\mathbf{s},\mathbf{t}))$. Define the Lagrangian function

$$\Lambda(\mathbf{t},\alpha) = (\Lambda_1(\mathbf{t},\alpha^1),\dots,\Lambda_W(\mathbf{t},\alpha^W))^T,$$

where $\forall i \in \mathcal{W} \quad \Lambda_i(\mathbf{t},\alpha^i) = \Phi\tilde{f}(\mathbf{s},\mathbf{t}) + \alpha^i(\phi^i - \sum_{p\in\mathcal{P}^i} t^i_{(p)})$.

Since we minimize a continuous and convex function on a convex, compact set, therefore either there exists an optimal solution or $\forall t \in \mathcal{X}^{\mathcal{W}} \quad \tilde{f}(\mathbf{s},\mathbf{t}) = +\infty$.

Next we show that an optimal solution satisfies (1). $\bar{\mathbf{t}}$ is an optimal solution , if and only if $\bar{\mathbf{t}}$ satisfies the following necessary and sufficient Kuhn-Tucker conditions: for any $i \in \mathcal{W} \quad \exists \bar{\alpha}^i \in \mathbb{R}$ (which depends on $\bar{\mathbf{t}}$) such that:

$$\nabla_i \Phi \tilde{f}(\mathbf{s},\bar{\mathbf{t}}) - \bar{\alpha}^i(1,\dots,1)^T \geq 0, \tag{8}$$

$$(\nabla_i \Phi \tilde{f}(\mathbf{s},\bar{\mathbf{t}}) - \bar{\alpha}^i(1,\dots,1)^T)^T \bar{t}^i = 0, \tag{9}$$

then $\bar{\alpha}^i = A^i$ and the result follows. $\square$

We now apply the existence theorem in [21, Thm. 1] to the convex game (in the sense of [21]) with $(N+1)$ players: the original N group users as well as the additional one who minimizes $\tilde{f}(\mathbf{x}^{\mathcal{N}},\mathbf{x}^{\mathcal{W}})$ with respect to $\mathbf{x}^{\mathcal{W}} \in \mathcal{X}^{\mathcal{W}}$.

Therefore under assumptions $\mathcal{A}'1 - \mathcal{A}'5$, there exists a mixed equilibrium.

Remark: If we wish to include constraints that involve also the individual users, such as add constraints on the links capacities (which then involves constraints on all users) then the condition of Wardrop equilibrium (all the paths used are of same cost) may not hold anymore. Nevertheless Larsson and Patriksson in [18] show that the program $\min \tilde{f}(x^{\mathcal{W}})$ with respect to the new strategy set leads to another kind of equilibrium (which they call *generalized Wardrop equilibrium*), in this case, we can also apply our *Lemma IV.1* and *Theorem 1* of [21] to obtain the existence of a "generalized mixed equilibrium".

V. Uniqueness of M.E.: Rosen's Type Condition

Definition: Let $T(\mathbf{x}) \in \mathbb{R}^n$ be a vector, then $\sigma(\mathbf{x},r) = \sum_{i=1}^n r_i T^i(\mathbf{x})$ is DSI (Diagonally Strictly Increasing) for $\mathbf{x} \in \mathcal{X}$ and for some $r \in \mathbb{R}^n,\ r \geq 0$ if for any $\bar{\mathbf{x}}$ and $\tilde{\mathbf{x}} \in \mathcal{X}$ ($\bar{\mathbf{x}} \neq \tilde{\mathbf{x}}$) we have

$$\sum_{i=1}^n r_i \left((\bar{x}^i - \tilde{x}^i)(T^i(\bar{\mathbf{x}}) - T^i(\tilde{\mathbf{x}}))\right) > 0,$$

$$\text{or equivalently, } (\bar{\mathbf{x}} - \tilde{\mathbf{x}})\cdot\gamma(\tilde{\mathbf{x}},r) + (\tilde{\mathbf{x}} - \bar{\mathbf{x}})\cdot\gamma(\bar{\mathbf{x}},r) < 0, \tag{10}$$

$$\text{where } \gamma^T(\mathbf{x},r) = (r_1T^1(\mathbf{x}), r_2T^2(\mathbf{x}),\dots,r_nT^n(\mathbf{x})).$$

The notion of DSI comes from the DSC (Diagonal Strict Convexity) of [21], in fact J.B. Rosen introduces the Diagonal Strict Concavity for a maximization problem, when we talk about a minimization problem we have to reverse the inequality in order to obtain convexity. The DSC is a condition on the derivatives ($\nabla_i T_i$), that we cannot apply in our case to the cost functions of individual users, that's why we introduce the DSI. Note that $\sigma(\mathbf{x},r) = \sum_{i=1}^n r_i \nabla_i T^i(\mathbf{x})$ DSI is equivalent to $s(\mathbf{x},r) = \sum_{i=1}^n r_i T^i(\mathbf{x})$ DSC.

In the previous section we considered a general convex, compact set $\mathcal{X}^{\mathcal{N}}$. In this section we need that $\mathcal{X}$ be orthogonal, then we restrict to sets that can be described as follows. Let $\mathcal{X} = \mathcal{X}^1 \times \mathcal{X}^2 \times \ldots \times \mathcal{X}^I$, where for any $i \in \mathcal{I}$, $\mathcal{X}^i$ is a bounded, closed and convex set defined by

- $\forall i \in \mathcal{N},\ \mathcal{X}^i = \{x^i | g^i(x^i) \leq 0\}$, where ${g^i}^T(x^i) = (g^i_1(x^i),\dots,g^i_{c_i}(x^i)),\ g^i_j(x^i),\ j = 1,\dots,c_i$, is a convex function of x^i, continuously differentiable and c_i (for $i \in \mathcal{N}$) is a constant.
- $\forall i \in \mathcal{W} \quad \mathcal{X}^i = \{x^i | -x^i_{(p)} \leq 0,\ \sum_{p\in\mathcal{P}^i} x^i_{(p)} = \phi^i\}$,

Then $\mathcal{X}$ is an orthogonal constraint set, which is convex.

Remark: g^i may represent (for $i \in \mathcal{N}$) the positivity constraints, the flow conservation constraints and some "extra" constraints. For $i \in \mathcal{W}$ positivity and demand constraints are explicitly described.

We introduce the following assumptions:

Assumptions:

- ($\mathcal{B}1$) There exists an interior point in the set of constraints which are not linear.
- ($\mathcal{B}2$) Wherever finite, J^i is continuously differentiable in x^i (which imposes that J^i is continuous in x^i).

- ($\mathcal{B}3$) J^i depends on $\mathbf{x}$ only through x^i and $\boldsymbol{\rho}$.

Notations:
$(\nabla J)^T(\mathbf{x}) = (\nabla_1 J^1(\mathbf{x}), \nabla_2 J^2(\mathbf{x}), \ldots, \nabla_N J^N(\mathbf{x}))$,
$f^T(\mathbf{x}) = (f_1(\rho_1), f_2(\rho_2), \ldots, f_L(\rho_L))$,
$\tilde{T} = \begin{pmatrix} \nabla J \\ f \end{pmatrix}$.

Let $\mathbf{y}$ be the function from $\mathcal{X}$ to $\mathcal{X}^{\mathcal{N}} \times \mathbb{R}^L$, defined by $\mathbf{y}(\mathbf{x}) = \mathbf{y} = (y^1, \ldots, y^N, y^{\mathcal{W}})^T$, where $\forall i \in \mathcal{N}, y^i = x^i$ and $y^{\mathcal{W}} = (y_1^{\mathcal{W}}, \ldots, y_L^{\mathcal{W}})^T$, and where $\forall l \in \mathcal{L}, y_l^{\mathcal{W}} = \rho_l^{\mathcal{W}} = \sum_{i \in \mathcal{W}} \frac{x_l^i}{\mu_l^i}$.
With some abuse of notation, we shall write for $\mathbf{y} = \mathbf{y}(\mathbf{x})$

$$J^i(\mathbf{y}) = J^i(\mathbf{y}(\mathbf{x})) = J^i(\mathbf{x}) \text{ and } f(\mathbf{y}) = f(\mathbf{y}(\mathbf{x})) = f(\mathbf{x}).$$

That J^i depends on $\mathbf{x}$ only through $\mathbf{y}(\mathbf{x})$ follows from ($\mathcal{B}3$) and that f depends on $\mathbf{x}$ only through $\mathbf{y}(\mathbf{x})$ follows from the fact that f depends on $\mathbf{x}$ only through ρ. Let

$$\sigma(\mathbf{y}, r) = \sum_{i \in \mathcal{N} \cup \mathcal{W}} r_i \tilde{T}^i(\mathbf{y}),$$

where $\forall i \in \mathcal{N}$ $\tilde{T}^i(\mathbf{y}) = J^i(\mathbf{y})$ and $\tilde{T}^{\mathcal{W}}(\mathbf{y}) = f(\mathbf{y})$, and let

$$\gamma^T(\mathbf{y}, r) = (r_1 \tilde{T}^1(\mathbf{y}), r_2 \tilde{T}^2(\mathbf{y}), \ldots, r_N \tilde{T}^N(\mathbf{y}), r_{\mathcal{W}} \tilde{T}^{\mathcal{W}}(\mathbf{y})).$$

Note again that for $i \in \{1, \ldots, N\} \cup \{\mathcal{W}\}$, $\tilde{T}^i(\mathbf{x}) = \tilde{T}^i(\mathbf{y})$.

Theorem V.1: If $\sigma(\mathbf{y}, r)$ is DSI for some $r > 0$, then all mixed equilibria $\mathbf{x}$ have the same utilization on links and moreover $\mathbf{x}^{\mathcal{N}}$ is unique.

Proof: Let $\bar{\mathbf{x}}$ and $\tilde{\mathbf{x}}$ be two M.E. Then we have for $\mathbf{x} = \bar{\mathbf{x}}$ and for $\mathbf{x} = \tilde{\mathbf{x}}$, $\forall i \in \mathcal{N}$, by the necessity of the Kuhn-Tucker conditions,

$$g^i(x^i) \leq 0 \text{ and } \exists \alpha^i \geq 0 \text{ such that} \tag{11}$$

$$\alpha^i \cdot g^i(x^i) = 0 \quad \left(\alpha^i = (\alpha_1^i, \ldots, \alpha_{c_i}^i)^T\right), \tag{12}$$

$$\nabla_i J^i(\mathbf{x}) + \sum_{j=1}^{c_i} \alpha_j^i \nabla_i g_j^i(x^i) = 0 \tag{13}$$

and $\forall i \in \mathcal{W}$ according to our *Lemma IV.1*, we have that

$$\text{for } p \in \mathcal{P}^i, \quad (F_{(p)}^i(\mathbf{x}) - A^i) x_{(p)}^i = 0 \tag{14}$$

which can be restated as

$$x^i \geq 0,\ g^i(x^i) \leq 0 \text{ and } \exists \beta^i \geq 0 \text{ such that} \tag{15}$$

$$\beta^i \cdot x^i = 0 \quad \left(\beta^i = (\beta_{(1)}^i, \ldots, \beta_{(P^i)}^i)^T\right) \tag{16}$$

$$F^i(\mathbf{x}) - \beta^i - A^i(1, \ldots, 1)^T = 0, \tag{17}$$

where $F^{i^T}(\mathbf{x}) = (F_{(1)}^i(\mathbf{x}), \ldots, F_{(P^i)}^i(\mathbf{x}))$. Note that $\boldsymbol{\beta} = (\beta^i,\ i \in \mathcal{W})$ depends on $\mathbf{x}$.

We multiply (13) (*resp.* (17)) by $r_i(\tilde{x}^i - \bar{x}^i)^T$ (*resp.* $r_{\mathcal{W}}(\tilde{x}^i - \bar{x}^i)^T$) for $\bar{\mathbf{x}}$ and by $r_i(\bar{x}^i - \tilde{x}^i)^T$ (*resp.* $r_{\mathcal{W}}(\bar{x}^i - \tilde{x}^i)^T$) for $\tilde{\mathbf{x}}$, and sum on i. This gives

$$d_{\mathcal{N}} + d_{\mathcal{W}} + \delta_{\mathcal{N}} + \delta_{\mathcal{W}} = 0, \tag{18}$$

where

$$d_{\mathcal{N}} = \sum_{i \in \mathcal{N}} r_i \left((\bar{x}^i - \tilde{x}^i) \cdot \tilde{T}^i(\tilde{\mathbf{x}}) + (\tilde{x}^i - \bar{x}^i) \cdot \tilde{T}^i(\bar{\mathbf{x}}) \right), \tag{19}$$

$$d_{\mathcal{W}} = r_{\mathcal{W}} \sum_{i \in \mathcal{W}} \sum_{p \in \mathcal{P}^i} (\tilde{x}_{(p)}^i - \bar{x}_{(p)}^i)(F_{(p)}^i(\bar{\mathbf{x}}) - F_{(p)}^i(\tilde{\mathbf{x}})), \tag{20}$$

$$\delta_{\mathcal{N}} = \sum_{i \in \mathcal{N}} r_i B_i, \quad \delta_{\mathcal{W}} = \sum_{i \in \mathcal{W}} r_{\mathcal{W}} B_i' + \sum_{i \in \mathcal{W}} r_{\mathcal{W}} C_i, \tag{21}$$

where

$$B_i = \sum_{j=1}^{c_i} \left\{ \bar{\alpha}_j^i (\tilde{x}^i - \bar{x}^i) \cdot \nabla_i g_j^i(\bar{x}^i) + \tilde{\alpha}_j^i (\bar{x}^i - \tilde{x}^i) \cdot \nabla_i g_j^i(\tilde{x}^i) \right\},$$

$$B_i' = \tilde{\beta}^i \cdot (\tilde{x}^i - \bar{x}^i) - \bar{\beta}^i \cdot (\tilde{x}^i - \bar{x}^i)$$

and

$$C_i = (\bar{A}^i - \tilde{A}^i) \sum_{p \in \mathcal{P}^i} (\bar{x}^i_{(p)} - \tilde{x}^i_{(p)}).$$

Since $\forall i \in \mathcal{N}$, $\forall j \in \{1, \dots, c_i\} \quad g_j^i$ is convex in x^i, therefore

$$(\tilde{x}^i - \bar{x}^i) \cdot \nabla_i g_j^i(\bar{x}^i) \leq g_j^i(\tilde{x}^i) - g_j^i(\bar{x}^i)$$

and

$$(\bar{x}^i - \tilde{x}^i) \cdot \nabla_i g_j^i(\tilde{x}^i) \leq g_j^i(\bar{x}^i) - g_j^i(\tilde{x}^i).$$

Moreover $\forall i \in \mathcal{N}$, $\forall j \in \{1, \dots, c_i\} \quad \alpha_j^i \geq 0$. Hence

$$\begin{aligned} B_i &\leq \sum_{j=1}^{c_i} \left[\bar{\alpha}_j^i \left(g_j^i(\tilde{x}^i) - g_j^i(\bar{x}^i) \right) + \tilde{\alpha}_j^i \left(g_j^i(\bar{x}^i) - g_j^i(\tilde{x}^i) \right) \right] \\ &\leq \left(\bar{\alpha}^i \cdot g^i(\tilde{x}^i) + \tilde{\alpha}^i \cdot g^i(\bar{x}^i) \right). \end{aligned}$$

The last inequality is due to (12) and (16). By (11) we have $B_i \leq 0$. By (15) and (16) we have also that $B_i' \leq 0$. Since

$$\forall i \in \mathcal{W}, \ \sum_{p \in \mathcal{P}^i} \tilde{x}^i_{(p)} = \sum_{p \in \mathcal{P}^i} \bar{x}^i_{(p)} = \phi^i,$$

therefore $C_i = 0$. Further,

$$\begin{aligned} d_{\mathcal{W}} &= r_{\mathcal{W}} \sum_{i \in \mathcal{W}} \sum_{p \in \mathcal{P}^i} \sum_{l \in \mathcal{L}} (\tilde{x}^i_{(p)} - \bar{x}^i_{(p)}) \delta_{lp} \frac{1}{\mu_l^i} (f_l(\bar{\rho}_l) - f_l(\tilde{\rho}_l)) \\ &= r_{\mathcal{W}} \sum_{l \in \mathcal{L}} (f_l(\bar{\rho}_l) - f_l(\tilde{\rho}_l)) (\tilde{y}_l^{\mathcal{W}} - \bar{y}_l^{\mathcal{W}}). \end{aligned}$$

Then

$$d_{\mathcal{N}} + d_{\mathcal{W}} = (\bar{\mathbf{y}} - \tilde{\mathbf{y}}) \cdot \gamma(\tilde{\mathbf{y}}, r) + (\tilde{\mathbf{y}} - \bar{\mathbf{y}}) \cdot \gamma(\bar{\mathbf{y}}, r)$$

Hence $\delta_{\mathcal{N}} \leq 0$, $\delta_{\mathcal{W}} \leq 0$ and it follows from (10) that $d_{\mathcal{N}} + d_{\mathcal{W}} < 0$ if $\bar{\mathbf{y}} \neq \tilde{\mathbf{y}}$. But this contradicts (18), so $\bar{\mathbf{y}} = \tilde{\mathbf{y}}$, *i.e.*, $\bar{\mathbf{x}}^{\mathcal{N}} = \tilde{\mathbf{x}}^{\mathcal{N}}$ and $\forall l \in \mathcal{L} \quad \bar{\rho}_l = \tilde{\rho}_l$. □

A. Sufficient Condition for DSI

Let $\mathcal{Y}$ be the set of $\mathbf{y}$ which correspond to a $\mathbf{x} \in \mathcal{X}$. Let the $(N+1) \times (N+1)$ matrix $\Gamma(\mathbf{y}, r)$ be the Jacobian of $\gamma(\mathbf{y}, r))$ for fixed $r > 0$. That is the jth column of $\Gamma(\mathbf{y}, r)$ is $\frac{\partial \gamma(\mathbf{y}, r)}{\partial y^j}, \ j \in \{1, \dots, N\} \cup \{\mathcal{W}\}$, where $\gamma(\mathbf{y}, r)$ is defined by (10). Then the condition given in [21, Thm. 6] holds for our definition of DSI, *i.e.*

Theorem V.2: A sufficient condition that $\sigma(\mathbf{y}, r)$ be diagonally strictly increasing for $\mathbf{y} \in \mathcal{Y}$ and fixed $r > 0$ is that the symmetric matrix $[\Gamma(\mathbf{y}, r) + \Gamma^T(\mathbf{y}, r)]$ be positive definite for $\mathbf{y} \in \mathcal{Y}$.

The corollary of this theorem given in [19, Cor 3.1] holds as well; define

$$\Gamma_l(y_l, r) = \left\{ r_i \frac{\partial \tilde{T}_l^i}{\partial y_l^j} \right\}_{i,j \{1, \dots, N\} \cup \{\mathcal{W}\}},$$

where for $i \in \mathcal{N}, \ \tilde{T}_l^i = \frac{\partial J_l^i}{\partial y_l^i}$ and $\tilde{T}_l^{\mathcal{W}} = f_l$.

Corollary V.3: Assume that for some positive $r \in \mathbb{R}^{N+1}$, the symmetric matrix $\left(\Gamma_l(y_l, r) + \Gamma_l^T(y_l, r) \right)$ is positive definite for every possible y_l and $l \in \mathcal{L}$. Then all mixed equilibria $\mathbf{x}$ have the same utilization on links and moreover $\mathbf{x}^{\mathcal{N}}$ is unique.

Proof: It may easily be seen that, up to re-indexing of rows and columns, $\Gamma(\mathbf{y}, r)$ equals diag$\{\Gamma_l(y_l, r), l \in \mathcal{L}\}$, and the required conclusion follows from Theorem V.2 and V.1. □

Example: Linear Costs. To illustrate Theorem V.1, we consider the following cost structure, for which uniqueness has already been obtained in [8] using an alternative approach. Define the cost functions as follows

- $\forall l \in \mathcal{L} \quad f_l(x_l) = p_l x_l + q_l$, where $p_l > 0$ and $q_l \geq 0$.
- $\forall i \in \mathcal{N},\ \forall l \in \mathcal{L} \quad J_l^i(x_l^i, x_l) = x_l^i . f_l(x_l)$.
- $\forall i \in \mathcal{W},\ \forall p \in \mathcal{P}^i \quad F_{(p)}^i(\mathbf{x}) = \sum_{l \in p} f_l(x_l)$.

Indeed in such a case we have

$$\Gamma_l(y_l, 1) = (\underline{1} \cdot \underline{1}^T + Q)p_l,$$

where $\underline{1}$ denotes the $(N+1)$-dimensional vector with entries all 1 and Q is the diagonal matrix with 1 everywhere on its diagonal, except at position $(N+1, N+1)$ where it is a 0.
For any $l \in \mathcal{L},\ \Gamma_l(y_l, 1)$ is a symmetric positive definite matrix.

Note that this network is a special case of equal service rates. One may easily find examples of networks with linear costs but different service rates where one cannot satisfy the hypothesis of *Corollary V.3*. In the next section we deal with such cases and we obtain a result on uniqueness for the links utilization.

VI. Uniqueness of M.E.: Linear Costs

We next obtain uniqueness of the utilization of some of the links in general networks with linear costs allowing for prioritizations through different service rates (thus extending the uniqueness results of [8]).
Assumptions:

- $(\mathcal{C}1)$ We define the cost function of user $i \in \mathcal{N}$ as follows:

$$J^i(\mathbf{x}) = \sum_{p \in \mathcal{P}} x_{(p)}^i \sum_{l \in \mathcal{L}} \delta_{lp} \frac{1}{\mu_l^i} f_l(\rho_l),$$

- $(\mathcal{C}2)$ f_l is linear and increasing.

We make the following assumption on the links:

- $(\mathcal{D})$ The set $\mathcal{L}$ is composed of two disjoint sets of links:

(i) $\mathcal{L}_{\mathcal{I}}$, for which $f_l(\rho_l)$ are strictly increasing,
(ii) $\mathcal{L}_{\mathcal{C}}$, for which $f_l(\rho_l) = f_l$ are constant (independent of ρ_l).
Remark: The f_l's are the same for all users (group and individual users).
We have

$$K_{(p)}^i(\mathbf{x}) := \frac{\partial J^i(\mathbf{x})}{\partial x_{(p)}^i} = \sum_{l \in L} \delta_{lp} \frac{1}{\mu_l^i} \left(f_l(\rho_l) + \frac{x_l^i}{\mu_l^i} \frac{\partial f_l(\rho_l)}{\partial \rho_l} \right) \text{ where } T(\mathbf{x}) = \begin{pmatrix} K(\mathbf{x}) \\ F(\mathbf{x}) \end{pmatrix}.$$

Lemma VI.1: Assume $\mathcal{C}1, \mathcal{C}2$ and $\mathcal{D}$. For arbitrary $\mathbf{x}$ and $\tilde{\mathbf{x}}$ ($\mathbf{x} \neq \tilde{\mathbf{x}}$), if $T(\mathbf{x})$ are finite or $T(\tilde{\mathbf{x}})$ are finite then

$$(\mathbf{x} - \tilde{\mathbf{x}}) \cdot [T(\mathbf{x}) - T(\tilde{\mathbf{x}})] > 0 \text{ if } \exists l \in \mathcal{L}_{\mathcal{I}} \text{ such that } \rho_l \neq \tilde{\rho}_l. \tag{22}$$

Proof: Assume that $\exists l \in \mathcal{L}_{\mathcal{I}}$ such that $\rho_l \neq \tilde{\rho}_l$. Note that $\nabla f_l := \frac{\partial f_l(\rho_l)}{\partial \rho_l} = \frac{\partial f_l(\tilde{\rho}_l)}{\partial \tilde{\rho}_l} \geq 0$ due to Assumption $(\mathcal{C}2)$. Then

$$\begin{aligned}(\mathbf{x} - \tilde{\mathbf{x}}) \cdot [T(\mathbf{x}) - T(\tilde{\mathbf{x}})] = &\left[\sum_{i \in \mathcal{N}} \sum_{p \in \mathcal{P}^i} \left(x_{(p)}^i - \tilde{x}_{(p)}^i \right) \left(K_{(p)}^i(\mathbf{x}) - K_{(p)}^i(\tilde{\mathbf{x}}) \right) \right] \\ &+ \left[\sum_{i \in \mathcal{W}} \sum_{p \in \mathcal{P}^i} \left(x_{(p)}^i - \tilde{x}_{(p)}^i \right) \left(F_{(p)}^i(\mathbf{x}) - F_{(p)}^i(\tilde{\mathbf{x}}) \right) \right],\end{aligned}$$

where $\forall i \in \mathcal{N}$,

$$\sum_{p \in \mathcal{P}^i} \left(x_{(p)}^i - \tilde{x}_{(p)}^i \right) \left(K_{(p)}^i(\mathbf{x}) - K_{(p)}^i(\tilde{\mathbf{x}}) \right)$$

$$= \sum_{p \in \mathcal{P}^i} \sum_{l \in \mathcal{L}} \left(x_{(p)}^i - \tilde{x}_{(p)}^i \right) \times \delta_{lp} \frac{1}{\mu_l^i} \left[\left(f_l(\rho_l) + \frac{x_l^i}{\mu_l^i} \frac{\partial f_l(\rho_l)}{\partial \rho_l} \right) - \left(f_l(\tilde{\rho}_l) + \frac{\tilde{x}_l^i}{\mu_l^i} \frac{\partial f_l(\tilde{\rho}_l)}{\partial \tilde{\rho}_l} \right) \right]$$

$$= \sum_{l \in \mathcal{L}} (\rho_l^i - \tilde{\rho}_l^i)(f_l(\rho_l) - f_l(\tilde{\rho}_l)) + \sum_{l \in \mathcal{L}} (\rho_l^i - \tilde{\rho}_l^i)^2 \nabla f_l,$$

and $\forall i \in \mathcal{W}$

$$\sum_{p\in\mathcal{P}^i}\left(x^i_{(p)}-\tilde{x}^i_{(p)}\right)\left(F^i_{(p)}(\mathbf{x})-F^i_{(p)}(\tilde{\mathbf{x}})\right)$$

$$=\sum_{p\in\mathcal{P}^i}\sum_{l\in\mathcal{L}}\left(x^i_{(p)}-\tilde{x}^i_{(p)}\right)\times\frac{\delta_{lp}}{\mu^i_l}(f_l(\rho_l)-f_l(\tilde{\rho}_l))$$

$$=\sum_{l\in\mathcal{L}}(\rho^i_l-\tilde{\rho}^i_l)(f_l(\rho_l)-f_l(\tilde{\rho}_l)).$$

Moreover we know that $\forall l \in \mathcal{L}_{\mathcal{C}},\ f_l(\rho_l)-f_l(\tilde{\rho}_l)=0$. Then

$$(\mathbf{x}-\tilde{\mathbf{x}})\cdot[T(\mathbf{x})-T(\tilde{\mathbf{x}})]=\sum_{l\in\mathcal{L}_{\mathcal{I}}}(f_l(\rho_l)-f_l(\tilde{\rho}_l))(\rho_l-\tilde{\rho}_l)+\sum_{i\in\mathcal{N}}\sum_{l\in\mathcal{L}}(\rho^i_l-\tilde{\rho}^i_l)^2\nabla f_l>0. \tag{23}$$

The last inequality is due to the existence of a link $l\in\mathcal{L}_{\mathcal{I}}$ such that $\rho_l\neq\tilde{\rho}_l$. Therefore we have the relation (22). □

Theorem VI.2: Assume $\mathcal{C}1$-$\mathcal{C}3$ and $\mathcal{D}$. Then all mixed equilibria have the same utilization on links $l\in\mathcal{L}_{\mathcal{I}}$.

Proof: Let $\bar{\mathbf{x}},\ \tilde{\mathbf{x}}$ be two mixed equilibria. Then according to *Lemma III.2* we have

$$T(\bar{\mathbf{x}})\cdot(\tilde{\mathbf{x}}-\bar{\mathbf{x}})\geq 0,\ \text{and}\ T(\tilde{\mathbf{x}})\cdot(\bar{\mathbf{x}}-\tilde{\mathbf{x}})\geq 0.$$

Then we obtain that

$$(\tilde{\mathbf{x}}-\bar{\mathbf{x}})\cdot(T(\tilde{\mathbf{x}})-T(\bar{\mathbf{x}}))\leq 0.$$

We can apply *Lemma VI.1*, since if the assumptions imposed are

- $\mathcal{A}1$-$\mathcal{A}2$, then $T(\mathbf{x})$ is finite for all $\mathbf{x}\in\mathcal{X}$,
- $\mathcal{A}'1$-$\mathcal{A}'5$, then $T(\mathbf{x})$ is finite if $\mathbf{x}$ is a mixed equilibrium, due to ($\mathcal{A}'5$). It implies that $\forall l\in\mathcal{L}_{\mathcal{I}}\quad \bar{\rho}_l=\tilde{\rho}_l$. □

VII. Uniqueness of M.E.: Positive Flows

The first Theorem of this Section shows under quite general conditions that if the global load on some links are the same under two equilibria, then also the flows of each user on these links are the same for the group users. Under more restrictive conditions, the second theorem in the section then, which extends [19, Thm. 3.3], establishes conditions for the uniqueness of the global load at equilibrium.

Assumptions:

- ($\mathcal{D}'$) The set $\mathcal{L}$ is composed of two disjoint sets of links:

(i) $\mathcal{L}_{\mathcal{I}}$, for which $h_l(\rho_l)$ are strictly increasing,
(ii) $\mathcal{L}_{\mathcal{C}}$, for which $h_l(\rho_l)=h_l$ are constant (independent of ρ_l).

- ($\mathcal{E}1$) All the individual users are grouped in a unique class, denoted $\mathcal{W}$, then we have $\mathcal{I}=\{1,\dots,N\}\cup\{\mathcal{W}\}$.
- ($\mathcal{E}2$) The service rate μ^i_l can be represented as $a^i\mu_l$, and $0<\mu^i_l$ is finite for all $i\in\mathcal{I}$, and $l\in\mathcal{L}$.
- ($\mathcal{E}3$) At each node, each class may re-route all the flow that it sends through that node to any of the out-going links of that node.
- ($\mathcal{E}4$) $J^i(\mathbf{x})=\sum_{l\in\mathcal{L}}J^i_l(x^i_l,\rho_l)$, where
- ($\mathcal{E}5$) J^i_l the cost function on link l for user i satisfies

$$J^i_l(x^i_l,\rho_l)=\frac{x^i_l}{\mu^i_l}h_l(\rho_l)=\rho^i_l h_l(\rho_l),\ \text{where}$$

- ($\mathcal{E}6$) h_l is continuous and increasing and J^i_l is continuously differentiable wherever finite.
- ($\mathcal{E}7$) ϕ^i_v is the amount of traffic of class i that enters the network at node $v\in\mathcal{M}$, if this quantity is negative this means that traffic of class i leaves node v at an amount of $|\phi^i_v|$. We assume that $\sum_{v\in\mathcal{M}}\phi^i_v=0$.

Note that in this section the cost functions on links for the group users, h_l, may be different of those used by individual users, f_l; while in the previous section we required that both types of users have the same ones, *i.e.*, $h_l=f_l$.

For each node v, we denote by $In(v)$ the set of its in-going links, and by $Out(v)$ the set of its out-going links. For each node v we have the following demand-conservation constraint

$$\sum_{l\in Out(v)}x^i_l=\sum_{l\in In(v)}x^i_l+\phi^i_v.$$

In order to minimize cost functions, we introduce the Lagrangian function

$$\Lambda^i(\mathbf{x},\alpha^i) = \sum_{l\in\mathcal{L}} \rho_l^i h_l(\rho_l) - \sum_{v\in\mathcal{M}} \alpha_v^i \left[\sum_{l\in Out(v)} x_l^i - \sum_{l\in In(v)} x_l^i - \phi_v^i \right],$$

where $\alpha^i = (\alpha_1^i, \alpha_2^i, \ldots, \alpha_M^i)^T$ is the vector of Lagrange multipliers for class i. Then for $\mathbf{x}$ to be a mixed equilibrium the following conditions on group users are necessary
for any $i \in \mathcal{N}$ there exists $\alpha^i = \alpha^i(x)$ such that for any $l \in \mathcal{L}$

$$\frac{\partial\Lambda^i(\mathbf{x},\alpha^i)}{\partial x_l^i} \geq 0, \quad \left(\frac{\partial\Lambda^i(\mathbf{x},\alpha^i)}{\partial x_l^i}\right) x_l^i = 0, \tag{24}$$

$$x_l^i \geq 0, \sum_{l\in Out(v)} x_l^i = \sum_{l\in In(v)} x_l^i + \phi_v^i.$$

We have

$$K_l^i(x_l^i,\rho_l) := \frac{\partial \rho_l^i h_l(\rho_l)}{\partial x_l^i} = \frac{1}{\mu_l^i}\left(\rho_l^i \frac{\partial h_l(\rho_l)}{\partial \rho_l} + h_l(\rho_l)\right).$$

With $l = (u,v)$, conditions (24) can be rewritten as

$$K_l^i(x_l^i,\rho_l) - \alpha_u^i + \alpha_v^i \geq 0; \quad \left(K_l^i(x_l^i,\rho_l) - \alpha_u^i + \alpha_v^i\right) x_l^i = 0, \tag{25}$$

Remark: For $l \in \mathcal{L}_\mathcal{I},\ K_l^i(x_l^i,\rho_l)$ is strictly increasing in both of its arguments.

Definition: $\mathcal{L}_0(\mathbf{x})$ is the set of links l such that for any $i \in \mathcal{I},\ x_l^i > 0$.

Theorem VII.1: Assume $\mathcal{D}'$ and $\mathcal{E}3$-$\mathcal{E}7$. If for all links $l \in \mathcal{L}_\mathcal{I}$ we have $\rho_l = \tilde{\rho}_l$ then $\forall l \in \mathcal{L}_\mathcal{I},\ \forall i \in \mathcal{N},\ x_l^i = \tilde{x}_l^i$.

Proof: Suppose that there exists a link $\bar{l} \in \mathcal{L}_\mathcal{I}$ and a group user $i \in \mathcal{N}$ such that $\tilde{x}_{\bar{l}}^i > x_{\bar{l}}^i$. We construct a directed network $G'(\mathcal{M}',\mathcal{L}')$, whose set of nodes is identical to the one of our original network, *i.e.*, $\mathcal{M}' = \mathcal{M}$ and the set of links $\mathcal{L}'$ is constructed as follows

- for each link $l = (u,v) \in \mathcal{L}$ such that $\tilde{x}_l^i > x_l^i$ we have a link $l' = (u,v) \in \mathcal{L}'$ which we assign a flow value $x_{l'} = \tilde{x}_l^i - x_l^i > 0$,
- for each link $l = (u,v) \in \mathcal{L}$ such that $\tilde{x}_l^i < x_l^i$ we have a link $l' = (v,u) \in \mathcal{L}'$ which we assign a flow value $x_{l'} = x_l^i - \tilde{x}_l^i > 0$.

In other words, we redirect links according to the relation between $\tilde{x}_l^i$ and x_l^i. Remark that $\forall l' \in \mathcal{L}',\ x_{l'} > 0$ and $\mathcal{L}' \neq \emptyset$ since for the link $\bar{l},\ \tilde{x}_{\bar{l}}^i > x_{\bar{l}}^i$. Then obviously the values $x_{l'}$ constitute a positive, directed flow in the network. Since at each node $w \in \mathcal{M}'$ we must have

$$\sum_{l\in Out(w)} x_l^i - \sum_{l\in In(w)} x_l^i = \phi_w^i = \sum_{l\in Out(w)} \tilde{x}_l^i - \sum_{l\in In(w)} \tilde{x}_l^i,$$

this flow has no sources (it is a circulation). Then there exists a cycle C of links in G' such that $x_{l'} > 0$ for all $l' \in C$ and such that $\bar{l} \in C$.
Consider now a link $l' = (u,v) \in C$. Therefore either $\tilde{x}_{uv}^i > x_{uv}^i$ or $x_{vu}^i > \tilde{x}_{vu}^i$. Then in the first case we have

$$\tilde{\alpha}_u^i - \tilde{\alpha}_v^i = K_{uv}^i(\tilde{x}_{uv}^i,\tilde{\rho}_{uv}) > K_{uv}^i(x_{uv}^i,\rho_{uv}) \geq \alpha_u^i - \alpha_v^i, \tag{26}$$

if $l \in \mathcal{L}_\mathcal{I}$, where the first equality and the last inequality follow from Kuhn-Tucker conditions and the first inequality follows from the hypothesis $\tilde{x}_{uv}^i > x_{uv}^i$ and the fact that $\tilde{\rho}_{uv} = \rho_{uv}$. Furthermore, if $l \in \mathcal{L}_\mathcal{C}$ then

$$\tilde{\alpha}_u^i - \tilde{\alpha}_v^i = K_{uv}^i(\tilde{x}_{uv}^i,\tilde{\rho}_{uv}) = K_{uv}^i(x_{uv}^i,\rho_{uv}) \geq \alpha_u^i - \alpha_v^i. \tag{27}$$

In the second case we have

$$\alpha_v^i - \alpha_u^i = K_{vu}^i(x_{vu}^i,\rho_{vu}) > K_{vu}^i(\tilde{x}_{vu}^i,\tilde{\rho}_{vu}) \geq \tilde{\alpha}_v^i - \tilde{\alpha}_u^i, \text{ if } l \in \mathcal{L}_\mathcal{I}, \tag{28}$$

$$\alpha_v^i - \alpha_u^i = K_{vu}^i(x_{vu}^i,\rho_{vu}) = K_{vu}^i(\tilde{x}_{vu}^i,\tilde{\rho}_{vu}) \geq \tilde{\alpha}_v^i - \tilde{\alpha}_u^i, \text{ if } l \in \mathcal{L}_\mathcal{C}. \tag{29}$$

Note that the results of equations (26) and (28) are in fact identical as well as those of (27) and (29).
Denote $\delta\alpha_w = \tilde{\alpha}_w^i - \alpha_w^i$ (for all $w \in \mathcal{M}$). From (26)-(29) we conclude that for a link $l' = (u,v) \in \mathcal{L}'$ which comes from a link $l \in \mathcal{L}_\mathcal{I}$ in the original network

$$x_{l'} > 0 \text{ implies that } \delta\alpha_u > \delta\alpha_v.$$

and for a link $l' = (u,v) \in \mathcal{L}'$ which comes from a link $l \in \mathcal{L}_\mathcal{C}$ in the original network

$$x_{l'} > 0 \text{ implies that } \delta\alpha_u \geq \delta\alpha_v.$$

This means that along the cycle C we would have a monotonically increasing sequence of $\delta\alpha$'s where a step is with a strict increase (due to the existence of the link $\bar{l}$), then $\delta\alpha_u > \delta\alpha_u$, which is a contradiction.

We conclude that $\forall l \in \mathcal{L}_\mathcal{I},\ \forall i \in \mathcal{N},\ x_l^i = \tilde{x}_l^i$. □

Theorem VII.2: Assume all users have the same source and destination. Assume $\mathcal{D}'$ and $\mathcal{E}1 - \mathcal{E}7$. Let $\mathbf{x}$ and $\tilde{\mathbf{x}}$ be two mixed equilibria. Assume that $\forall i \in \mathcal{I},\ \forall l \notin \mathcal{L}_0(\mathbf{x})\quad x_l^i = 0$ and $\forall i \in \mathcal{I},\ \forall l \notin \mathcal{L}_0(\tilde{\mathbf{x}}),\ \tilde{x}_l^i = 0$. Then $\forall l \in \mathcal{L}_\mathcal{I},\ \rho_l = \tilde{\rho}_l$ and moreover $\forall l \in \mathcal{L}_\mathcal{I},\ \forall i \in \mathcal{N},\ x_l^i = \tilde{x}_l^i$.

it, then the assumption cannot be satisfied. *Proof:* Denote $\alpha_u = \sum_{i\in\mathcal{N}} a^i\alpha_u^i$ (where a^i is defined in ($\mathcal{E}2$)) and

$$S_l(\rho_l) = \rho_l \frac{\partial h_l(\rho_l)}{\partial \rho_l} + N h_l(\rho_l).$$

$S_l(\rho_l)$ are finite (due to our assumption ($\mathcal{A}'5$)).

Let $\boldsymbol{\alpha}$ (*resp.* $\tilde{\boldsymbol{\alpha}}$) be the vector of Lagrange multipliers associated to $\mathbf{x}$ (*resp.* $\tilde{\mathbf{x}}$).

Since $\rho_l^{\mathcal{W}} \geq 0$, we have

$$S_l(\rho_l) \geq \sum_{i\in\mathcal{N}} \left(\rho_l^i \frac{\partial h_l(\rho_l)}{\partial \rho_l} + h_l(\rho_l) \right).$$

(25) implies that

$$\frac{1}{\mu_{uv}} S_{uv}(\rho_{uv}) \geq \alpha_u - \alpha_v, \tag{30}$$

with equality for $(u,v) \in \mathcal{L}_0(\mathbf{x})$. A similar relation holds for $\tilde{\mathbf{x}}$.

We obtain for

$$\begin{aligned} 0 \quad &\leq \sum_{(u,v)\in\mathcal{L}} (\rho_{uv} - \tilde{\rho}_{uv})(S_{uv}(\rho_{uv}) - S_{uv}(\tilde{\rho}_{uv})) \\ &\leq \sum_{(u,v)\in\mathcal{L}} \mu_{uv}(\rho_{uv} - \tilde{\rho}_{uv})((\alpha_u - \tilde{\alpha}_u) - (\alpha_v - \tilde{\alpha}_v)) = 0. \end{aligned} \tag{31}$$

The first inequality follows from the monotonicity and the convexity of $h_l(\rho_l)$ for $l \in \mathcal{L}_\mathcal{I}$. The second inequality holds in fact for each pair u, v (and not just for the sum). Indeed, for $(u,v) \in \mathcal{L}_0(\mathbf{x}) \cap \mathcal{L}_0(\tilde{\mathbf{x}})$ this relation holds with equality due to (30). This is also the case for $(u,v) \notin \mathcal{L}_0(\mathbf{x}) \cup \mathcal{L}_0(\tilde{\mathbf{x}})$, since in that case $\rho_{uv} = \tilde{\rho}_{uv} = 0$. Consider next the case $(u,v) \in \mathcal{L}_0(\mathbf{x})$,
$(u,v) \notin \mathcal{L}_0(\tilde{\mathbf{x}})$. Then we have

$$\begin{aligned} (\rho_{uv} - \tilde{\rho}_{uv})(S_{uv}(\rho_{uv}) - S_{uv}(\tilde{\rho}_{uv})) \quad &= \quad \rho_{uv}(S_{uv}(\rho_{uv}) - S_{uv}(\tilde{\rho}_{uv})) \\ &\leq \quad \mu_{uv}\rho_{uv}((\alpha_u - \tilde{\alpha}_u) - (\alpha_v - \tilde{\alpha}_v)). \end{aligned}$$

A symmetric argument establishes the case $(u,v) \notin \mathcal{L}_0(\mathbf{x}),\ (u,v) \in \mathcal{L}_0(\tilde{\mathbf{x}})$. We finally establish the last equality in (31).

$$\sum_{(u,v)\in\mathcal{L}} \mu_{uv}(\rho_{uv} - \tilde{\rho}_{uv})((\alpha_u - \tilde{\alpha}_u) - (\alpha_v - \tilde{\alpha}_v))$$

$$= \sum_{r\in\mathcal{M}} (\alpha_r - \tilde{\alpha}_r) \sum_{s\in\mathcal{M}(r,s)\in\mathcal{L}} (\rho_{rs} - \tilde{\rho}_{rs})\mu_{rs} - \sum_{r\in\mathcal{M}} (\alpha_r - \tilde{\alpha}_r) \sum_{s\in\mathcal{M},(s,r)\in\mathcal{L}} (\rho_{sr} - \tilde{\rho}_{sr})\mu_{sr}$$

$$\begin{aligned} &= \sum_{r\in\mathcal{M}} (\alpha_r - \tilde{\alpha}_r) \left(\sum_{l\in Out(r)} (\rho_l - \tilde{\rho}_l)\mu_l - \sum_{l\in In(r)} (\rho_l - \tilde{\rho}_l)\mu_l \right) \\ &= \sum_{i\in\mathcal{I}} \frac{1}{a^i} \left[\sum_{r\in\mathcal{M}} (\alpha_r - \tilde{\alpha}_r) \left(\sum_{l\in Out(r)} (x_l^i - \tilde{x}_l^i) - \sum_{l\in In(r)} (x_l^i - \tilde{x}_l^i) \right) \right] \\ &= \sum_{i\in\mathcal{I}} \frac{1}{a^i} \left[\sum_{r\in\mathcal{M}} (\alpha_r - \tilde{\alpha}_r) \left(\sum_{l\in Out(r)} x_l^i - \sum_{l\in In(r)} x_l^i - \left(\sum_{l\in Out(r)} \tilde{x}_l^i - \sum_{l\in In(r)} \tilde{x}_l^i \right) \right) \right] = 0. \end{aligned}$$

Since for all $l \in \mathcal{L}$ (resp. $\in \mathcal{L}_{\mathcal{I}}$), S_l is increasing (resp. strictly increasing), we conclude from (31) that $\rho_l = \tilde{\rho}_l$ for all links in $\mathcal{L}_{\mathcal{I}}$. The first part of the theorem is established.

From Theorem VII.1 we conclude that $\forall l \in \mathcal{L},\ \forall i \in \mathcal{N} \quad x_l^i = \tilde{x}_l^i$. Thus the theorem is established. $\square$

Remark The above Theorem extends [19, Thm 3.3]. The latter first establishes, under a more restrictive setting, the uniqueness of global link flows. Then it proceeds to conclude the uniqueness of the actual flows by hinting at an argument different than Theorem VII.1, taken from the proof of [19, Thm 2.1], which deals with the case of parallel links. We have not been able to reconstruct that argument, as it uses the fact that the sum of link flows of each user between two nodes does not depend on the equilibrium; this indeed is trivially true in the case of a parallel link topology, but one still needs to show that this extends to general topology. Our Theorem VII.1 of course implies this.

VIII. Uniqueness of M.E. for specific topologies

We establish below the uniqueness of M.E. in networks with specific topologies: a network of parallel links, and two load balancing models from [11]. The uniqueness of the Nash equilibrium ($\mathcal{W} = \emptyset$) for these models in the equal service rate case has been established in [19, thm 2.1], [11, Thm 5.1] and [10]. For the load balancing networks, uniqueness and characterization of the M.E. has been derived in [12] for the case of a completely symmetric network. Introduce the following

Assumptions:

- $(\mathcal{F}1)$ $J_l^i : [0,\infty)^2 \to [0,\infty]$, $J_l^i(x_l^i, \rho_l)$ is a continuous function, convex in x_l^i.
- $(\mathcal{F}2)$ Wherever finite, J_l^i is continuously differentiable in x_l^i. We denote $K_l^i := \partial J_l^i / \partial x_l^i$.
- $(\mathcal{F}3)$ K_l^i depends of two arguments x_l^i and ρ_l and is strictly increasing in both of them.

Notation: In order to simplify the reading, let $f_l^i(\rho_l) = \frac{1}{\mu_l^i} f_l(\rho_l)$.

A. Parallel Links

In a network with parallel links, all the users have the same origin and the same destination, and moreover each link is a path and vice-versa. Then we have $F_{(p)}^i(\mathbf{x}) = \sum_{l \in p} f_l^i(\rho_l) = f_l^i(\rho_l)$. Even for such a simple network, and even if we took equal service rates, the conditions in Harker [8] or the DSI condition are typically not satisfied; indeed, it is shown in [19] that these type of conditions do not hold in the special case of two links, two group users (with no individual users), and link costs that are of the type of an M/M/1 queue, except for very low traffic demands.

Lemma VIII.1: In a network of parallel links where the cost function of each user satisfies $(\mathcal{A}'1) - (\mathcal{A}'5)$ and $(\mathcal{F}1) - (\mathcal{F}3)$, all mixed equilibria $\mathbf{x}$ have the same utilization on links and moreover $\mathbf{x}^{\mathcal{N}}$ is unique.

Proof: We recall that $\rho_l^i = \frac{x_l^i}{\mu_l^i}$ and $\tilde{\rho}_l^i = \frac{\tilde{x}_l^i}{\mu_l^i}$.

Let $\mathbf{x}$ and $\tilde{\mathbf{x}} \in \mathcal{X}$ be two mixed equilibria. Then $\mathbf{x}$ and $\tilde{\mathbf{x}}$ satisfy the following conditions:
for $i \in \mathcal{W},\ \forall l \in \mathcal{L}$

$$f_l^i(\rho_l) - A^i \geq 0; \quad (f_l^i(\rho_l) - A^i) x_l^i = 0, \tag{32}$$

$$f_l^i(\tilde{\rho}_l) - \tilde{A}^i \geq 0; \quad (f_l^i(\tilde{\rho}_l) - \tilde{A}^i) \tilde{x}_l^i = 0, \tag{33}$$

$\exists\ \boldsymbol{\alpha},\ \tilde{\boldsymbol{\alpha}}$ such that $\forall i \in \mathcal{N},\ \forall l \in \mathcal{L}$

$$K_l^i(x_l^i, \rho_l) - \alpha^i \geq 0; \quad (K_l^i(x_l^i, \rho_l) - \alpha^i) x_l^i = 0, \tag{34}$$

$$K_l^i(\tilde{x}_l^i, \tilde{\rho}_l) - \tilde{\alpha}^i \geq 0; \quad (K_l^i(\tilde{x}_l^i, \tilde{\rho}_l) - \tilde{\alpha}^i) \tilde{x}_l^i = 0, \tag{35}$$

where $\boldsymbol{\alpha} = (\alpha^1, \ldots, \alpha^N)^T$ (*resp.* $\tilde{\boldsymbol{\alpha}} = (\tilde{\alpha}^1, \ldots, \tilde{\alpha}^N)^T$) is the vector of Lagrange multipliers associated to $\mathbf{x}$ (*resp.* $\tilde{\mathbf{x}}$).

The first step is to establish that $\rho_l = \tilde{\rho}_l,\ \forall l \in \mathcal{L}$. To this end, we use the relations of the proof of [19, Thm. 2.1], *i.e* we prove that for each $l \in \mathcal{L}$ and $i \in \mathcal{N}$, the following relations hold

$$\{\tilde{\alpha}^i \leq \alpha^i,\ \tilde{\rho}_l \geq \rho_l\} \text{ implies that } \tilde{x}_l^i \leq x_l^i, \tag{36}$$

$$\{\tilde{\alpha}^i \geq \alpha^i,\ \tilde{\rho}_l \leq \rho_l\} \text{ implies that } \tilde{x}_l^i \geq x_l^i. \tag{37}$$

We shall only prove (36), since (37) is symmetric. Assume that $\tilde{\alpha}^i \leq \alpha^i$ and $\tilde{\rho}_l \geq \rho_l$ for some $l \in \mathcal{L}$ and some $i \in \mathcal{N}$. Note that (36) holds trivially if $\tilde{x}_l^i = 0$. Otherwise, if $\tilde{x}_l^i > 0$, then (34)-(35) together with our assumption imply that

$$K_l^i(\tilde{x}_l^i, \tilde{\rho}_l) = \tilde{\alpha}^i \leq \alpha^i \leq (K_l^i(x_l^i, \rho_l) \leq (K_l^i(x_l^i, \tilde{\rho}_l), \tag{38}$$

where the last inequality follows from the monotonicity of K_l^i in its second argument. Now, since K_l^i is non-decreasing in its first argument, this implies that $\tilde{x}_l^i \leq x_l^i$, and (36) is established.

Furthermore we have $l \in \mathcal{L}$ and $i \in \mathcal{W}$

$$\{\tilde{A}^i \leq A^i,\ \tilde{\rho}_l > \rho_l\} \text{ (or equivalently } \{\tilde{A}^i < A^i,\ \tilde{\rho}_l \geq \rho_l\}) \text{ implies that } \tilde{x}_l^i = 0, \tag{39}$$

$$\{\tilde{A}^i > A^i,\ \tilde{\rho}_l \leq \rho_l\} \text{ (or equivalently } \{\tilde{A}^i \geq A^i,\ \tilde{\rho}_l < \rho_l\}) \text{ implies that } x_l^i = 0. \tag{40}$$

Actually suppose $\tilde{x}_l^i > 0$, hence by (32)-(33) we obtain:

$$f_l(\tilde{\rho}_l) = \tilde{A}^i \leq A^i \leq f_l(\rho_l) \text{ (or } f_l(\tilde{\rho}_l) = \tilde{A}^i < A^i \leq f_l(\rho_l)),$$

which is a contradiction with our assumption on f_l. Therefore $\tilde{x}_l^i = 0$ and a fortiori $\tilde{x}_l^i \leq x_l^i$, (40) is symmetric.

Let $\mathcal{L}_1 = \{l :\ \tilde{\rho}_l > \rho_l\}$ and $\mathcal{I}_a = \{i :\ \tilde{\alpha}^i > \alpha^i \text{ or } \tilde{A}^i > A^i\}$,
$\mathcal{L}_2 = \mathcal{L} - \mathcal{L}_1 = \{l :\ \tilde{\rho}_l \leq \rho_l\}$. Assume that $\mathcal{L}_1$ is not empty. Since $\sum_l \tilde{x}_l^i = \sum_l x_l^i = \phi^i$, therefore (37) and (40) imply that for any $i \in \mathcal{I}_a$

$$\sum_{l \in \mathcal{L}_1} \tilde{x}_l^i = \phi^i - \sum_{l \in \mathcal{L}_2} \tilde{x}_l^i \leq \phi^i - \sum_{l \in \mathcal{L}_2} x_l^i = \sum_{l \in \mathcal{L}_1} x_l^i.$$

Since (36) and (39) imply that $\tilde{x}_l^i \leq x_l^i$ for $l \in \mathcal{L}_1$ and $i \notin \mathcal{I}_a$, therefore

$$\sum_{l \in \mathcal{L}_1} \tilde{\rho}_l = \sum_{l \in \mathcal{L}_1} \sum_{i \in \mathcal{I}} \frac{\tilde{x}_l^i}{\mu_l^i} \leq \sum_{l \in \mathcal{L}_1} \sum_{i \in \mathcal{I}} \frac{x_l^i}{\mu_l^i} = \sum_{l \in \mathcal{L}_1} \rho_l.$$

This inequality contradicts our non-emptiness assumption on $\mathcal{L}_1$, and then $\mathcal{L}_1 = \emptyset$. By symmetry it follows that the set $\{l :\ \tilde{\rho}_l < \rho_l\}$ is also empty. Thus, it has been established that

$$\tilde{\rho}_l = \rho_l,\ \forall l \in \mathcal{L}, \tag{41}$$

i.e., all mixed equilibria have the same utilization of links.

So it is now sufficient to establish that $\tilde{\alpha}^i = \alpha^i$ in order to prove the theorem.
(36) may be strengthened as follows

$$\{\tilde{\alpha}^i < \alpha^i,\ \tilde{\rho}_l = \rho_l\} \text{ implies that either } \tilde{x}_l^i < x_l^i \text{ or } \tilde{x}_l^i = x_l^i = 0. \tag{42}$$

Indeed, if $\tilde{x}_l^i = 0$ then the implication is trivial. Otherwise, if $\tilde{x}_l^i > 0$, it follows that

$$K_l^i(\tilde{x}_l^i, \tilde{\rho}_l) = \tilde{\alpha}^i < \alpha^i \leq K_l^i(x_l^i, \rho_l),$$

so that $\tilde{x}_l^i < x_l^i$ as required.
Assume that $\tilde{\alpha}^i < \alpha^i$ for some $i \in \mathcal{N}$. Since $\sum_{l \in \mathcal{L}} \tilde{x}_l^i = \phi_i > 0$, then $\exists l \in \mathcal{L}$ such that $\tilde{x}_l^i > 0$ and (42) implies that

$$\sum_{l \in \mathcal{L}} x_l^i > \sum_{l \in \mathcal{L}} \tilde{x}_l^i = \phi_i,$$

which contradicts the demand constraint for user i. Hence $\tilde{\alpha}^i < \alpha^i$ does not hold for any user $i \in \mathcal{N}$, we obtain by a symmetric argument that $\tilde{\alpha}^i > \alpha^i$ does not hold as well.
Finally we have for $i \in \mathcal{N},\ \tilde{\alpha}^i = \alpha^i$.
Combined with (41), this implies by (36) and (37) that $\tilde{x}_l^i = x_l^i$ for all $l \in \mathcal{L},\ i \in \mathcal{N}$ and the lemma is proved. □

B. Load Balancing with unidirectional links

We consider a model consisting of two processors, a and b, and a two-way communication lines, c and c', between them; the two processors can be seen as two links, a and b, between two different sources, u and v, and a unique destination.

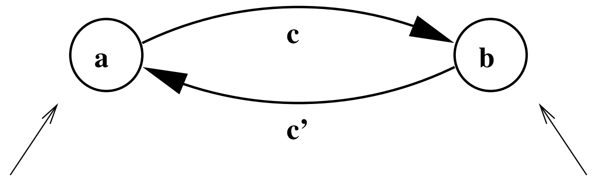

Fig. 1. Two processors and two unidirectional links between them

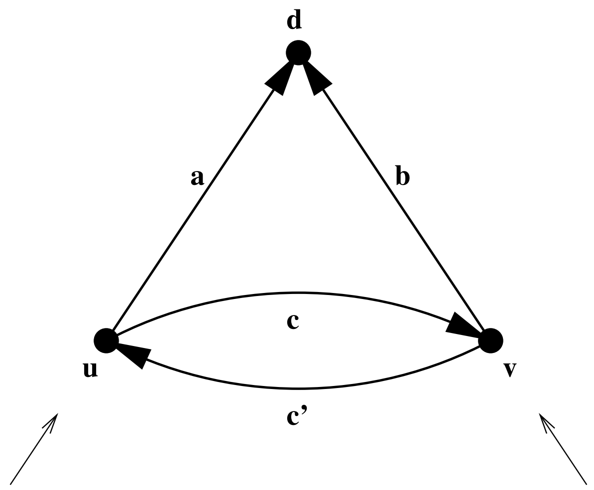

Fig. 2. The corresponding network

This network can be reformulated in a network of parallel links ([10]), then we can apply *Lemma VIII.1* and obtain the same uniqueness than previously for this model of load balancing. This result can be extended to a model of n processors with a two-way communication lines between each couple of processors (there are exactly $2(n-1)^2$ lines).

C. Load Balancing with a Communication Bus

We now consider a model made up of two processors, a and b, and a communication bus, c, between them; the two processors can be seen as two links, a and b, between two different sources, u and v, and a unique destination, d.

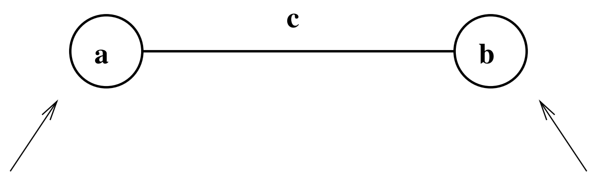

Fig. 3. Two processors and a communication bus

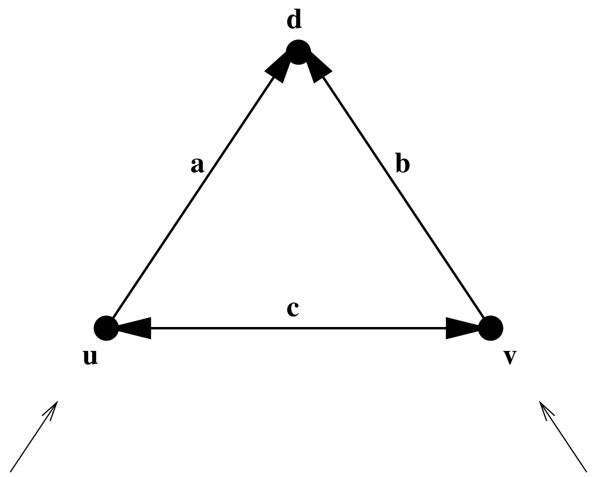

Fig. 4. The corresponding network

Since a bidirectional link can be transformed in a network of unidirectional links (Appendix B) mixed equilibria in networks with unidirectional and bidirectional links exist.
Notation: We denote for $w = u, v$ by i_w the part of class i whose origin is w, *i.e.*, $s(i_w) = w$ and by ϕ^i_w the initial flow demand of user i at node w. In this model there exist four paths, $\mathcal{P} = \{(a),\ (ca),\ (b),\ (cb)\}$.

Each class $i \in \mathcal{N}$ is faced with the minimization program

$$\min_{x^i} J^i(\mathbf{x}) = \sum_{p \in \mathcal{P}^i} J^i_{(p)}(x^i_{(p)}, \boldsymbol{\rho}), \quad s.t.$$

$$x^i_{(a)} + x^i_{(cb)} = \phi^i_u, \quad x^i_{(a)} \geq 0,\ x^i_{(cb)} \geq 0,$$

$$x^i_{(b)} + x^i_{(ca)} = \phi^i_v, \quad x^i_{(b)} \geq 0,\ x^i_{(ca)} \geq 0.$$

Note that we allow $\phi^i_w = 0$, if the class $i \in \mathcal{N}$ has only one O-D pair, then $\forall i \in \mathcal{N},\ \mathcal{P}^i = \mathcal{P}$.

For $l = a, b$, we have $J^i_{(l)}(x^i_{(l)}, \boldsymbol{\rho}) = J^i_l(x^i_{(l)}, \rho_l)$
and $J^i_{(cl)}(x^i_{(cl)}, \boldsymbol{\rho}) = J^i_l(x^i_{(cl)}, \rho_l) + J^i_c(x^i_{(cl)}, \rho_c)$.

We define for $p \in \mathcal{P}$

$$K^i_{(p)}(x^i_{(p)}, \boldsymbol{\rho}) := \frac{\partial J^i_{(p)}(x^i_{(p)}, \boldsymbol{\rho})}{\partial x^i_{(p)}}.$$

We have for $l = a, b$

$$K^i_{(l)}(x^i_{(l)}, \boldsymbol{\rho}) = K^i_l(x^i_{(l)}, \rho_l) \tag{43}$$

and

$$K^i_{(cl)}(x^i_{(cl)}, \boldsymbol{\rho}) = K^i_l(x^i_{(cl)}, \rho_l) + K^i_c(x^i_{(cl)}, \rho_c), \tag{44}$$

where for $p = l, cl$

$$K^i_l(x^i_{(p)}, \rho_l) = \frac{\partial J^i_l(x^i_{(p)}, \rho_l)}{\partial x^i_{(p)}},$$

and

$$K^i_c(x^i_{(cl)}, \rho_c) = \frac{\partial J^i_c(x^i_{(cl)}, \rho_c)}{\partial x^i_{(cl)}}.$$

Then for $\mathbf{x}$ to be a mixed equilibrium we have the following necessary Kuhn-Tucker conditions:
There exists $\boldsymbol{\alpha}$ such that $\forall i \in \mathcal{N}$

$$K^i_{(p)}(x^i_{(p)}, \boldsymbol{\rho}) - \alpha^i_w \geq 0; \quad (K^i_{(p)}(x^i_{(p)}, \boldsymbol{\rho}) - \alpha^i_w) x^i_{(p)} = 0, \tag{45}$$

where $p = a, ca$ if $w = u$ and $p = b, cb$ if $w = v$.

Lemma VIII.2: Assume that the service rate of each user does not depend on the links. Then in a network with two processors and a communication bus between them and where the cost function of each user satisfies $(\mathcal{A}'1) - (\mathcal{A}'5)$ and $(\mathcal{F}1) - (\mathcal{F}3)$, all mixed equilibria $\mathbf{x}$ have the same utilization on links.
Proof: See Appendix C

IX. Conclusion

We have focused in this paper the M.E. concept introduced in [8] and studied it under more general assumptions on the costs and for more general setting of optimization (which allows one to use constraints). M.E. involves groups that contain a continuum of users, where some of the groups have a single decision maker for the whole group and others have a decision maker per user. We further established uniqueness of the mixed equilibrium by either restricting to specific topologies or making some extra assumptions on the equilibrium flows.

A future research direction would be to add also extra constraints on the individual users (see [18]). We have not included these constraints here (except for a Remark in the end of Section IV) since in their presence, the Wardrop principles need not hold anymore. For example, consider a network of two parallel links having both load independent costs, and in which there is a capacity constraint on a link with the lowest cost. If the latter link cannot accommodate all the flow then we would expect the outcome of individual optimization to yield the full utilization of that link and partial utilization of the other one. Hence the costs of different links (paths) that carry positive flow are not the same, thus violating Wardrop principle.

APPENDIX

A. Constraints in General Networks

In a general network the configuration flows which are feasible are the $\mathbf{x}$ which satisfy

$$\Delta^T \mathbf{x} = \Phi \ \& \ \mathbf{x} \geq 0, \text{ where } \mathbf{x} = \begin{pmatrix} \mathbf{x}^{\mathcal{N}} \\ \mathbf{x}^{\mathcal{W}} \end{pmatrix}.$$

In order to simplify the reading let $d^i = \# D^i$.

$$\mathcal{T} = \begin{pmatrix} \tau_{11}^{11} & \tau_{12}^{11} & \cdots & \tau_{1d^1}^{11} & \tau_{11}^{12} & \cdots & \tau_{1d^N}^{1N} \\ \tau_{21}^{11} & \tau_{22}^{11} & \cdots & \tau_{2d^1}^{11} & \tau_{21}^{12} & \cdots & \tau_{2d^N}^{1N} \\ \vdots & \vdots & \ddots & \vdots & & & \\ \tau_{P^1 1}^{11} & \tau_{P^1 2}^{11} & \cdots & \tau_{P^1 d^1}^{11} & \tau_{P^1 1}^{12} & \cdots & \tau_{P^1 d^N}^{1N} \\ \tau_{11}^{21} & \tau_{12}^{21} & \cdots & \tau_{1d^1}^{21} & \tau_{11}^{22} & \cdots & \tau_{1d^N}^{2N} \\ \vdots & \vdots & \ddots & \vdots & & & \\ \tau_{(P^N-1)1}^{N1} & \tau_{(P^N-1)2}^{N1} & \cdots & \tau_{(P^N-1)d^1}^{N1} & \tau_{(P^N-1)1}^{N2} & \cdots & \tau_{(P^N-1)d^N}^{NN} \\ \tau_{P^N 1}^{N1} & \tau_{P^N 2}^{N1} & \cdots & \tau_{P^N d^1}^{N1} & \tau_{P^N 1}^{N2} & \cdots & \tau_{P^N d^N}^{NN} \end{pmatrix}$$

whose element (q, r) is τ_{pd}^{ij}, where $i, j \in \mathcal{N}$, $d \in D^j$, $p \in \mathcal{P}^i$, $q = k + \sum_{k=1}^{i-1} P^k$ and $r = d + \sum_{s=1}^{j-1} d^s$ and $\tau_{dk}^{ij} = \begin{cases} 1 & \text{if } i = j \text{ and } k \in d, \\ 0 & \text{otherwise.} \end{cases}$

$$\Theta = \begin{pmatrix} \theta_1^{11} & \theta_1^{12} & \cdots & \theta_1^{1W} \\ \theta_2^{11} & \theta_2^{12} & \cdots & \theta_2^{1W} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{P^1}^{11} & \theta_{P^1}^{12} & \cdots & \theta_{P^1}^{1W} \\ \theta_1^{21} & \theta_1^{22} & \cdots & \theta_1^{2W} \\ \theta_2^{21} & \theta_2^{22} & \cdots & \theta_2^{2W} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{P^W-1}^{W1} & \theta_{P^W-1}^{W2} & \cdots & \theta_{P^W-1}^{WW} \\ \theta_{P^W}^{W1} & \theta_{P^W}^{W2} & \cdots & \theta_{P^W}^{WW} \end{pmatrix}$$

whose element (q, r) is θ_k^{ij}, where $i, j \in \mathcal{W}$, $k \in \mathcal{P}^i$, $q = k + \sum_{s=1}^{i-1} P^s$ and $r = j$ and $\theta_k^{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$
and

$$\Delta = \begin{pmatrix} \mathcal{T} & 0 \\ 0 & \Theta \end{pmatrix}$$

Remark: Here for all $i \quad x^i$ is expressed in term of paths, *i.e.*,
$\forall i \in \mathcal{N} \cup \mathcal{W} \quad x^i = (x^i_{(1)}, \ldots, x^i_{(P^i)})^T$.

B. Relation between a Bidirectional Link and a Network of Unidirectional Ones

A bidirectional link may always be expressed as an equivalent network of unidirectional ones. Indeed, consider a bidirectional link l between nodes u and v. where the cost function of this link is $f_l^i(x_l)$, and where x_l is the aggregate flow through link l. Then we can transform this link into the network of unidirectional links $(l1, l2, l3, l4, l5)$ in Fig. 5 with the cost functions: $f_{l1}^i(x_{l1}) = f_{l2}^i(x_{l2}) = f_{l4}^i(x_{l4}) = f_{l5}^i(x_{l5}) = 0$ and $f_{l3}^i(x_{l3}) = f_l^i(x_{l3})$. These two subnetworks are not equivalent, since in the second one a user i can go from u to u (of from v to v) with the cost $f^i(x_{l3})$ which is not possible in the first one. Neverthelesss, they become equivalent if we add a constraint in the second network excluding cycles. This does not affect the equilibrium, since at equilibrium the paths $(l1, l3, l4)$ and $(l2, l3, l5)$ will not be used (as long as costs are nonnegative, of course).

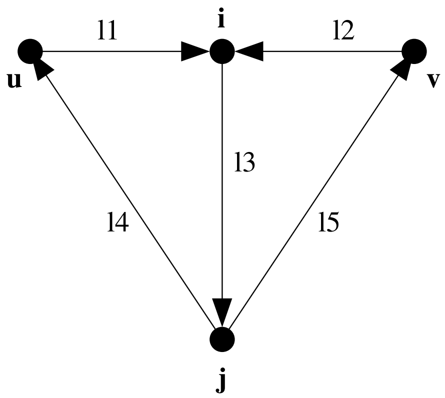

Fig. 5. the corresponding network

C. Proof of Lemma VIII.2

Let $\mathbf{x}$ and $\tilde{\mathbf{x}} \in \mathcal{X}$ be two mixed equilibria.
Assume that $\tilde{\rho}_c \geq \rho_c$. Let $k \in \{a, b\}$ such that $\tilde{\rho}_k \geq \rho_k$ (which is equivalent to $\tilde{\rho}_l \leq \rho_l,\ l \neq k$). Let $\nu(l')$ be the source associated with l', *i.e.*, $l' = (\nu(l'), d)$, $l' \in \{l, k\}$, where d is the unique destination.
We recall that each class $i \in \mathcal{W}$ has only one O-D pair, then in this model, either $s(i) = u$ or $s(i) = v$; define $x^i_{(a)} = x^i_{(cb)} = 0$ if $s(i) = v$ and $x^i_{(b)} = x^i_{(ca)} = 0$ if $s(i) = u$.

First step: We first prove that for all $i \in \mathcal{W}$,

$$\tilde{x}^i_{(l)} \geq x^i_{(l)} \text{ or } \tilde{\rho}_l = \rho_l. \tag{46}$$

It is trivial if $s(i) = \nu(k)$, then we assume that $s(i) = \nu(l)$, we are faced with the two following cases

1) $\tilde{A}^i \geq A^i$
Then either $x^i_{(l)} = 0$ and $\tilde{x}^i_{(l)} \geq x^i_{(l)}$ or $x^i_{(l)} > 0$ and $\tilde{\rho}_l = \rho_l$.
The first implication is trivial, so we have only to check the second implication.
Let $x^i_{(l)} > 0$ then we have

$$F^i_{(l)}(\tilde{\mathbf{x}}) = f^i_l(\tilde{\rho}_l) \geq \tilde{A}^i \geq A^i = F^i_{(l)}(\mathbf{x}) = f^i_l(\rho_l).$$

Therefore, since f^i_l is strictly increasing in ρ_l, we have $\tilde{\rho}_l \geq \rho_l$ and finally $\tilde{\rho}_l = \rho_l$.

2) $\tilde{A}^i \leq A^i$
Then either $\tilde{x}^i_{(ck)} = 0$ and $\tilde{x}^i_{(l)} \geq x^i_{(l)}$ or $\tilde{x}^i_{(ck)} > 0$ and $\tilde{\rho}_l = \rho_l$.
As above we have only to check the second implication, the first one being trivial ($\tilde{x}^i_{(ck)} \leq x^i_{(ck)}$ is the same as $\tilde{x}^i_{(l)} \geq x^i_{(l)}$). Let $\tilde{x}^i_{(cl)} > 0$ then we have

$$F^i_{(ck)}(\tilde{\mathbf{x}}) = f^i_k(\tilde{\rho}_k) + f^i_c(\tilde{\rho}_c) = \tilde{A}^i \leq A^i \leq F^i_{(ck)}(\mathbf{x}) = f^i_k(\rho_k) + f^i_c(\rho_c),$$

which implies, since $\tilde{\rho}_k \geq \rho_k$, that $\rho_c \geq \tilde{\rho}_c$ which contradicts our assumption $\tilde{\rho}_c \geq \rho_c$, unless we have equality for c and for k. Then (46) is established.
Second step: We next prove a statement in the spirit of the first step, but for $i \in \mathcal{N}$:

$$\tilde{x}^i_{(l)} \geq x^i_{(l)} \text{ or equivalently } \tilde{x}^i_{(ck)} \leq x^i_{(ck)}. \tag{47}$$

First remark that the equivalence follows from the constraint on the sum of the flows. (47) holds trivially if $x^i_{(l)} = 0$, and so we have to check only the case $x^i_{(l)} > 0$. To do so, fix some $i \in \mathcal{N}$ and consider the following two subcases. Assume that
(a) $\tilde{\alpha}^i_{\nu(l)} \geq \alpha^i_{\nu(l)}$. Hence

$$K^i_{(l)}(\tilde{x}^i_{(l)}, \tilde{\boldsymbol{\rho}}) = K^i_l(\tilde{x}^i_{(l)}, \tilde{\rho}_l) \geq \tilde{\alpha}^i_{\nu(l)} \geq \alpha^i_{\nu(l)} = K^i_{(l)}(x^i_{(l)}, \boldsymbol{\rho}) = K^i_l(x^i_{(l)}, \rho_l) \geq K^i_l(x^i_{(l)}, \tilde{\rho}_l).$$

The first and last equalities follow from the definition of $K^i_{(l)}$. The other equality as well as the first inequality follow from the Kuhn Tucker conditions, whereas the last inequality follows from the monotonicity assumption ($\mathcal{F}3$). Using again ($\mathcal{F}3$), this time for the first argument, we conclude from the fact $K^i_l(\tilde{x}^i_{(l)}, \tilde{\rho}_l) \geq K^i_l(x^i_{(l)}, \tilde{\rho}_l)$ that (47) holds. Thus we try instead of (a):

(b) $\tilde{\alpha}^i_{\nu(l)} \leq \alpha^i_{\nu(l)}$. (47) holds trivially if $\tilde{x}^i_{(ck)} = 0$. So it remains to check the case $\tilde{x}^i_{(ck)} > 0$. Recall that $K^i_{(ck)}(x^i_{(ck)}, \boldsymbol{\rho}) = K^i_c(x^i_{(ck)}, \rho_c) + K^i_k(x^i_{(ck)}, \rho_k)$. We then have for $i \in \mathcal{N}$

$$K^i_c(x^i_{(ck)}, \rho_c) + K^i_k(x^i_{(ck)}, \rho_k) \geq \alpha^i_{\nu(l)} \geq \tilde{\alpha}^i_{\nu(l)} = K^i_c(\tilde{x}^i_{(ck)}, \tilde{\rho}_c) + K^i_k(\tilde{x}^i_{(ck)}, \tilde{\rho}_k)$$

$$\geq K^i_c(\tilde{x}^i_{(ck)}, \rho_c) + K^i_k(\tilde{x}^i_{(ck)}, \rho_k).$$

Here, the first inequality and the equality follow from the Kuhn Tucker conditions, whereas the last inequality follows from ($\mathcal{F}3$). Using again ($\mathcal{F}3$), we conclude that (47) holds in case (b) as well.
Then (47) is established.

Third Step: We shall next prove that our two mixed equilibria have the same utilization of links.
(I) If there exists some $i \in \mathcal{W}$ such that $\tilde{x}^i_{(ck)} > 0$ and $\tilde{A}^i \leq A^i$, it has already been proved (see case 2 in the first step of the proof of our Lemma).
(II) If there exists some $i \in \mathcal{W}$ such that $x^i_{(l)} > 0$ and $\tilde{A}^i \geq A^i$, then $\tilde{\rho}_k = \rho_k$ and $\tilde{\rho}_l = \rho_l$ (This is the first case of the first step of the proof of the Lemma). It then remains to show that $\tilde{\rho}_c = \rho_c$.

Suppose that $\tilde{\rho}_c > \rho_c$, thus there exists $i_0 \in \mathcal{N} \cup \mathcal{W}$ such that $\tilde{x}^{i_0}_c > x^{i_0}_c$. (i) if $i_0 \in \mathcal{N}$, then $\tilde{x}^{i_0}_{(ck)} + \tilde{x}^{i_0}_{(cl)} > x^{i_0}_{(ck)} + x^{i_0}_{(cl)}$, which implies, due to (47), that $\tilde{x}^{i_0}_{(cl)} > x^{i_0}_{(cl)}$ and $x^{i_0}_{(k)} > \tilde{x}^{i_0}_{(k)}$.
We obtain

$$K^{i_0}_k(x^{i_0}_{(k)}, \rho_k) = \alpha^{i_0}_{\nu(k)} \leq K^{i_0}_c(x^{i_0}_{(cl)}, \rho_c) + K^{i_0}_l(x^{i_0}_{(cl)}, \rho_l)$$

$$< K^{i_0}_c(\tilde{x}^{i_0}_{(cl)}, \tilde{\rho}_c) + K^{i_0}_l(\tilde{x}^{i_0}_{(cl)}, \tilde{\rho}_l) = \tilde{\alpha}^{i_0}_{\nu(k)} \leq K^{i_0}_k(\tilde{x}^{i_0}_{(k)}, \tilde{\rho}_k),$$

where the strict inequality follows from ($\mathcal{F}3$) and the equality and the others inequality follow from the Kuhn Tucker conditions. Using again ($\mathcal{F}3$), we conclude that $x^{i_0}_{(k)} < \tilde{x}^{i_0}_{(k)}$, which is a contradiction, then $\tilde{\rho}_c = \rho_c$.
(ii) if $i_0 \in \mathcal{W}$, then we shall show that this implies that

$$\text{there exists } i_1 \in \mathcal{W} \text{ such that } s(i_1) = \nu(l) \text{ and } \tilde{x}^{i_1}_{(ck)} > x^{i_1}_{(ck)}. \tag{48}$$

Indeed, if $i_0 \in \mathcal{W}$ and $s(i_0) = \nu(k)$, we have $\tilde{x}^{i_0}_{(cl)} > x^{i_0}_{(cl)}$ but since $\tilde{\rho}_k = \rho_k$, then

– either there exists $j \in \mathcal{W}$ such that $\tilde{x}^j_{(ck)} > x^j_{(ck)}$, which implies $s(j) = \nu(l)$, and hence (48) is established,

– or there exists $j \in \mathcal{N}$ such that $\tilde{x}^j_{(ck)} + \tilde{x}^j_{(k)} > x^j_{(ck)} + x^j_{(k)}$. In this case (47) implies that $\tilde{x}^j_{(k)} > x^j_{(k)}$ or equivalently $\tilde{x}^j_{(cl)} < x^j_{(cl)}$. Then there must exist at least one other class j such that $\tilde{x}^{i_0}_c > x^{i_0}_c$ (so that $\tilde{\rho}_c > \rho_c$).

Then we can conclude from (i) that there exists $i_1 \in \mathcal{W}$ such that $s(i_1) = \nu(l)$ and $\tilde{x}^{i_1}_{(ck)} > x^{i_1}_{(ck)}$. If $\tilde{A}^{i_0} \leq A^{i_0}$, we are faced with the second case of the first step and we have seen that $\tilde{\rho}_c = \rho_c$. Then assume that $\tilde{A}^{i_0} \geq A^{i_0}$, therefore $x^{i_0}_{(l)} > 0$ ($x^{i_0}_{(l)} > \tilde{x}^{i_0}_{(l)}$) and we have

$$\min_{p \in P^{i_0}} F^{i_0}_{(p)}(\mathbf{x}) = F^{i_0}_{(l)}(\mathbf{x}) = F^{i_0}_{(l)}(\tilde{\mathbf{x}}) \leq F^{i_0}_{(ck)}(\mathbf{x}) < F^{i_0}_{(ck)}(\tilde{\mathbf{x}}),$$

where the first equality is due to $x^{i_0}_{(l)} > 0$, the second one to $\tilde{\rho}_l = \rho_l$ and the last inequality to $\tilde{\rho}_k = \rho_k$ and our first assumption $\tilde{\rho}_c > \rho_c$.
Hence $\tilde{x}^{i_0}_{(ck)} = 0$, which is a contradiction.

(III) If $\forall i \in \mathcal{W},\ \tilde{x}^i_{(l)} \geq x^i_{(l)}$, then according to (47), we conclude that

$$\sum_{i \in \mathcal{I}} \tilde{\rho}^{i_{\nu(l)}}_k \leq \sum_{i \in \mathcal{I}} \rho^{i_{\nu(l)}}_k. \tag{49}$$

Combining this with $\tilde{\rho}_k \geq \rho_k$, we obtain that $\sum_{i \in \mathcal{I}} \tilde{\rho}^{i_{\nu(k)}}_k \geq \sum_{i \in \mathcal{I}} \rho^{i_{\nu(k)}}_k$.

However, since for $i \in \mathcal{N} \cup \mathcal{W},\ \tilde{x}^i_{(l)} + \tilde{x}^i_{(ck)} = \phi^i_{\nu(l)}$ and $\tilde{x}^i_{(k)} + \tilde{x}^i_{(cl)} = \phi^i_{\nu(k)}$ (note that if $i \in \mathcal{W}$ and $s(i) = \nu(l)$ (*resp.* $s(i) = \nu(k)$), then $\phi^i_{\nu(k)} = 0$ (*resp.* $\phi^i_{\nu(l)} = 0$)) which is equivalent to $\tilde{\rho}^{i_{\nu(l)}}_k + \tilde{\rho}^{i_{\nu(l)}}_l = \frac{\phi^i_{\nu(l)}}{\mu^i}$ and $\tilde{\rho}^{i_{\nu(k)}}_k + \tilde{\rho}^{i_{\nu(k)}}_l = \frac{\phi^i_{\nu(k)}}{\mu^i}$ by our assumption on μ^i. It follows that

$$\sum_{i \in \mathcal{I}} \tilde{\rho}^{i_{\nu(k)}}_l \leq \sum_{i \in \mathcal{I}} \rho^{i_{\nu(k)}}_l. \tag{50}$$

But this, combined with (49), implies that $\tilde{\rho}_c \leq \rho_c$ and then $\tilde{\rho}_c = \rho_c$ (since $x^i_c = x^{i_{\nu(k)}}_l + x^{i_{\nu(l)}}_k$).
Suppose now that $\tilde{\rho}_l < \rho_l$ then due to (47) and our assumption $\forall i \in \mathcal{W}_l\ \tilde{x}^i_{(l)} \geq x^i_{(l)}$, it follows that $\exists i \in \mathcal{I}$ such that $\tilde{x}^i_{(cl)} < x^i_{(cl)}$,

i.e., $\tilde{x}^i_{(k)} > x^i_{(k)}$.
If $i \in \mathcal{N}$, then we have

$$K^i_k(x^i_{(k)}, \rho_k) < K^i_k(\tilde{x}^i_{(k)}, \tilde{\rho}_k) = \tilde{\alpha}^i_{\nu(k)} \leq K^i_l(\tilde{x}^i_{(cl)}, \tilde{\rho}_l) + K^i_c(\tilde{x}^i_{(cl)}, \tilde{\rho}_c)$$
$$< K^i_l(x^i_{(cl)}, \rho_l) + K^i_c(x^i_{(cl)}, \rho_c) = \alpha^i_{\nu(k)} \leq K^i_k(x^i_{(k)}, \rho_k),$$

where the first and the third inequalities are due to the strict increase in each argument of the K's and the equalities to $\tilde{x}^i_{(k)} > 0$, $x^i_{(cl)} > 0$ and (43). But this is impossible.
If $i \in \mathcal{W}$, therefore we have

$$F^i_{(cl)}(\tilde{\mathbf{x}}) < F^i_{(cl)}(\mathbf{x}) = \min_{p \in P^i} F^i_{(p)}(\mathbf{x}) \leq F^i_{(k)}(\mathbf{x}) < F^i_{(k)}(\tilde{\mathbf{x}}),$$

where the first equality is due to $x^i_{(cl)} > 0$, the first inequality to $\tilde{\rho}_l < \rho_l$ and $\tilde{\rho}_c = \rho_c$ and the last inequality to $\tilde{\rho}_k > \rho_k$. This implies that $\tilde{x}^i_{(k)} = 0$ and this is a contradiction.

Then $\forall i \in \mathcal{I} \quad \tilde{x}^i_{(cl)} \geq x^i_{(cl)}$ and finally $\tilde{\rho}_l = \rho_l$. Hence the lemma is established. □

REFERENCES

[1] E. Altman, T. Başar, T. Jiménez and N. Shimkin, “Competitive routing in networks with polynomial cost”, to appear in *IEEE Trans. on Automatic Control*, 2001.

[2] E. Altman, T. Başar, T. Jiménez and N. Shimkin, ”Routing into two parallel links: game-theoretic distributed algorithms”, to appear in the Special Issue of *Journal of Parallel and Distributed Computing* on ”Routing in Computer and Communication Networks”, 2001.

[3] E. Altman and H. Kameda, “Uniqueness of solutions for optimal static routing in multi-class open networks”, submitted, Research report ISE-TR-98-156, Institute of Information Sciences and Electronics University of Tsukuba, Japan.

[4] H. Brézis, “équations et inéquatios non linéaires dans les espaces vectoriels en dualité”, *Annales de l'Institut Fourier*, vol 18, pp.115-175, 1968.

[5] S. Dafermos and A. Nagurney, “On some traffic equilibrium theory paradoxes”, *Transportation Research B*, **18**, pp. 101-110, 1984.

[6] S. Dafermos and F. T. Sparrow, “The traffic assignment problem for a general network”, *Journal of Research of the National Bureau of Standards-B. Mathematical Sciences* **73B**, No. 2, pp. 91-118, 1969.

[7] P. Gupta and P. R. Kumar, ”A system and traffic dependent adaptive routing algorithm for Ad Hoc networks”, *Proceedings of the 36th IEEE Conference on Decision and Control*, San-Diego, USA, Dec. 1997.

[8] P. Harker, “Multiple equilibrium behaviors on networks”, *Transportation Research B*, **22**, pp. 39-46, 1988.

[9] A. Haurie and P. Marcotte, “On the relationship between Nash-Cournot and Wardrop Equilibria”, *Networks*, **15**, pp. 295-308, 1985.

[10] H. Kameda, E. Altman and T. Kozawa, “Braess-like Paradoxes of Nash Equilibria for Load Balancing in Distributed Computer Systems”, ISE technical Report ISE-TR-98-157, Institute of Information Sciences and Electronics, University of Tsukuba, Japan, 1998.

[11] H. Kameda, E. Altman and T. Kozawa, “A case where a paradox like Braess's occurs in the Nash equilibrium but does not occur in the Wardrop equilibrium - a situation of load balancing in distributed computer systems”, Proceedings of IEEE CDC'99 , Phoenix, Arizona, USA, Dec., 1999.

[12] H. Kameda, E. Altman, O. Pourtallier, “Analytic study of distributed optima coexisting within a network of homogeneous computers”, Research report ISE-TR-00-174, Institute of Information Sciences and Electronics, University of Tsukuba, Japan, 2000. Shorter version in *Proceedings of the 9th International Symposium on Dynamic Games and Applications*, Adelaide, Australia, Dec. 18-21, 2000.

[13] H. Kameda, E. Altman, O. Pourtallier, J. Li, Y. Hosokawa, “Paradoxes in Performance Optimization of Distributed Systems”, Proceedings of *SSGRR 2000 Computer and ebusiness conference*, l'Aquila, Italy, Jul 31 - Aug 6, 2000.

[14] H. Kameda, T. Kozawa and J. Li, “Anomalous Relations among Various Performance Objectives in Distributed Computer Systems,” *Proc. 1st World Congress on Systems Simulation*, IEEE, 1997, pp.459-465.

[15] H. Kameda, J. Li, C. Kim and Y. Zhang, *Optimal Load Balancing in Distributed Computer Systems*, Springer, 1997.

[16] Y. A. Korilis, A. A. Lazar and A. Orda, “Architecting noncooperative networks”, *IEEE J. on Selected Areas in Communications*, **13**, No. 7, sept 1995, pp. 1241-1251.

[17] R. J. La and V. Anantharam, “Optimal routing control: game theoretic approach”, *Proc. of the 36th IEEE Conference on Decision and Control*, San Diego, California, Dec. 1997.

[18] T Larsson and M. Patriksson, “Side Constrained Traffic Equilibrium Models - Analysis, Computation and Applications”, *Transportation Research*, **33B**, pp 233-264, 1999.

[19] A. Orda, R. Rom and N. Shimkin, “Competitive Routing in Multiuser Communication Networks”, *IEEE/ACM Trans. on Networking 1*, 1993, pp.614-627.

[20] M. Patriksson, *The Traffic Assignment Problem: Models and Methods* VSP BV, P.O. Box 346, 3700 AH Zeist, The Netherlands, 1994.

[21] J. B. Rosen, “Existence and Uniqueness of Equilibrium Points for Concave N-person Games”, *Econometrica 33*, pp.153-163, 1965.

[22] J.G. Wardrop, “Some Theoretical Aspects of Road Traffic Research”, *Proc. Inst. Civ. Eng.* Part 2, 1, 325–378, 1952.

[23] B. W. Wie, “A differential game approach to the dynamic mixed behavior traffic network equilibrium problem”, *European Journal of Operational Research*, 83:117–136, 1995.