Capacity Allocation Under Noncooperative Routing

Yannis A. Korilis, *Member, IEEE,* Aurel A. Lazar, *Fellow, IEEE,* and Ariel Orda, *Member, IEEE*

***Abstract*—The capacity allocation problem in a network that is to be shared by noncooperative users is considered. Each user decides independently upon its routing strategy so as to optimize its individual performance objective. The operating points of the network are the Nash equilibria of the underlying routing game. The network designer aims to allocate link capacities, so that the resulting Nash equilibria are efficient, according to some systemwide performance criterion. In general, the solution of such design problems is complex and at times counterintuitive, since adding link capacity might lead to degradation of user performance. For systems of parallel links, we show that such paradoxes do not occur and that the capacity allocation problem has a simple and intuitive optimal solution that coincides with the solution in the single-user case.**

***Index Terms*—Capacity allocation, Nash games, noncooperative networks, routing.**

I. Introduction

THE COMPLEXITY of high-speed, large-scale networks calls for decentralized control algorithms, where control decisions are made by each user independently, according to its own individual performance objectives.[1] Such networks are henceforth called *noncooperative*, and game theory [1], [2] provides the systematic framework to study and understand their behavior. The operating points of a noncooperative network are the *Nash equilibria* of the underlying game, that is, the points where unilateral deviation does not help any user to improve its performance.

In modern networking, game theoretic models have been employed in the context of flow control [3]–[6], routing [7]–[9], and virtual path bandwidth allocation [10]. These studies mainly investigate the structure of the Nash equilibria and provide valuable insight into the nature of networking under decentralized and noncooperative control.

The present work approaches noncooperative networking from a different viewpoint: given that the network is shared by noncooperative users, is it possible to devise a set of design rules which guarantee that the resulting Nash equilibria exhibit certain desirable properties? Network design issues are scarcely addressed under a noncooperative setting, mainly due to the complex structure—or lack thereof—of the underlying game. One exception is [11], which addresses the problem of designing the service discipline of a switch shared by users performing noncooperative flow control.

We consider the problem of optimal capacity allocation under noncooperative routing. The network is shared by a set of noncooperative users, each bifurcating its flow over the paths available in the network, in a way that optimizes its individual performance objective.[2] The noncooperative routing scenario applies to various modern networking environments. The Internet Protocol (both IPv4 and the current IPv6 Specification), for example, provides the option of source routing [16], [17] that enables the user to determine the path(s) its flow follows from source to destination. Another example is the flexible routing service as specified in the Q.1211 CCITT Recommendation for the standardized capability set of Intelligent Networks (IN CS-1) [18]. One of the goals of this service is to route calls over particular facilities based on the subscriber's routing preference list or distribution algorithm.

The network designer allocates link capacities while satisfying lower bounds specified per link and an upper bound on the total capacity of the network. A capacity allocation is sought, such that the resulting routing equilibrium exhibits the "best" performance according to some networkwide efficiency criterion. We consider several efficiency criteria, such as the "price" (marginal cost) as seen by each user, the cost of each user, or some combination of the above, for example, the average network delay. The combined capacity and routing optimization is a hard problem, even when routing is centrally controlled (single-user case) and heuristics are usually in place [12], [20]. The complexity of the problem is even more pronounced in the case of noncooperative routing. Indeed, for the design problem described above it is not even clear that the designer should deploy all the available capacity. The well-known Braess paradox indicates that, in general, addition of

Manuscript received July 21, 1995. Recommended by Associate Editor, P. Nain. This work was supported by the Office of Naval Research under Contract N00014-90-J-1289. The work of A. Orda was supported in part by the consortium for broadband communication administered by the Chief Scientist of the Israeli Ministry of Industry and Commerce. This paper was presented in part at the IEEE INFOCOM'95, Boston, MA, April 1995.

Y. A. Korilis was with the Department of Electrical Engineering, Columbia University, New York, NY, USA. He is now with Bell Laboratories, Lucent Technologies, Holmdel, NJ 07733 USA (e-mail: yannisk@dnrc.bell-labs.com).

A. A. Lazar is with the Department of Electrical Engineering and the Center for Telecommunications Research, Columbia University, NY 10027 USA.

A. Orda is with the Department of Electrical Engineering, Technion-Israel Institute of Technology, Haifa 32000, Israel. This work was in part carried out while he was visiting with the Center for Telecommunications Research at Columbia University.

Publisher Item Identifier S 0018-9286(97)02049-7.

[1] The term "user" may refer to a network user itself or, in case that the user's traffic consists of multiple connections, to individual connections that are controlled independently.

[2] Bifurcation of flows is a well-established routing mechanism that has been studied extensively in the networking literature (e.g., [12] and references therein) and has already been implemented in practical networks [12]–[14]. Bifurcated routing is often preferred to simple shortest-path methods, since the latter may result in oscillatory behavior [12], [15]. Even in cases where individual connections might not be split over different paths, optimal bifurcation can be achieved (or approximated); a "user" might represent an organization that decides on the routing of its total flow, thus it can approximate optimal bifurcated routing by assigning the various connections it controls over appropriate paths in the network.

capacity to a network may degrade the performance of each and every user [21], [22]; an example that adopts the paradox to the communication network framework considered in this paper is presented in [28].

The optimal capacity allocation problem is analyzed in detail for a simple network consisting of a common source and a common destination node interconnected by a number of parallel links, for which it is known that there exists a unique Nash equilibrium for any capacity configuration [9]. Systems of parallel links, albeit simple, represent an appropriate model for seemingly unrelated networking problems. Consider, for example, a network in which resources are preallocated to various routing paths that do not interfere. Such scenarios are common in modern networking. In broadband networks, for instance, bandwidth is separated among different virtual paths, resulting effectively in a system of parallel and noninterfering "links" between source/destination pairs. Moreover, to reduce the complexity of routing mechanisms, the network might present the users with a limited set of paths between source and destination, hiding the underlying physical topology. Another example is that of a corporation or organization that receives service from a number of different network providers. The corporation can split its total flow over the various network facilities (according to performance and cost considerations), each of which can be represented as a "link" in the parallel link model. Finally, it should be noted that routing, as a control paradigm, applies not only to the allocation of paths to messages and connections in communication networks, but in fact to any problem of splitting load among several resources, e.g., distribution of tasks among multiple processors. Consider, for example, a multimedia network with several servers that are shared by the network customers; each customer distributes its applications among the servers while competing with the other customers on the common available resources, resulting in effect with a routing game. Modeling each resource (e.g., multimedia server) as a "link," the parallel links model considered in our study fits well such scenarios.

In the single-user (parallel links) case, the optimal capacity allocation problem has a simple and intuitive solution: the best design strategy is to allot the entire additional capacity to the link with the initially highest capacity [12]. One of the main contributions of this work is to generalize this result (for the various efficiency criteria considered by the designer) to the case of noncooperative routing, independently of the number and the throughput demands of the users. While in the single-user case the proof is quite simple, in a multi-user setting it requires systematic and rather cautious analysis to establish some "order" in the complex structure of the routing game. More specifically, we decompose the problem into two subproblems: 1) the problem of adding capacity to any link and 2) the problem of transferring capacity from one link to another. These subproblems correspond to practical situations encountered in various networking environments, where the capacity of a single physical link is dynamically allocated to several logical links (e.g., virtual paths) and, therefore, provide design rules that are interesting in their own right. For the capacity addition problem, we establish that adding capacity to any link in the network improves performance. For the capacity transfer problem we show that transferring capacity toward the link with the originally highest capacity improves the performance of the network. Combining these results, we obtain the solution to the optimal capacity allocation problem.

An important practical implication of these results is that, although users make noncooperative decisions, design methodologies can be devised to improve the overall network performance. Improvements can be achieved both during the *provisioning phase*, i.e., when the network parameters are sized, and during the *run time phase*, i.e., during the actual operation of the network. The capacity allocation problem considered here aims at improvements during the provisioning phase. Strategies that improve the network performance during the run time phase are investigated in [23].

The outline of the paper is the following. In Section II we present the parallel links model and formulate the optimal capacity allocation problem. Section III explores the structure of the underlying Nash equilibria and establishes several properties that form the foundation of the subsequent analysis. In Section IV we prove that addition of capacity to a system of parallel links does not degrade performance. In Section V we show that transferring capacity toward the link with the originally highest capacity improves performance. With these results at hand, we investigate, in Section VI, the optimal strategy for adding capacity to a system of parallel links. Finally, Section VII summarizes the main results, delineates their practical implications, and discusses possible extensions.

II. Model and Problem Formulation

A. Model and Preliminaries

We consider a set $\mathcal{I} = \{1, \cdots, I\}$ of users that share a set $\mathcal{L} = \{1, \cdots, L\}$ of communication links interconnecting a common source to a common destination node. Let c_l be the capacity of link l and $C = \Sigma_{l \in \mathcal{L}}\ c_l$ be the total capacity of the system. Each user i has a throughput demand that is some process with average rate $r^i > 0$. We assume that $r^1 \geq r^2 \geq \cdots \geq r^I$. Let $R = \Sigma_{i \in \mathcal{I}}\ r^i$ denote the total demand of the users. We only consider capacity configurations $\boldsymbol{c} = (c_1, \cdots, c_L)$ that can accommodate the total user demand, that is, configurations with $C > R$.

User i ships its flow by splitting its demand r^i over the set of parallel links, according to some individual performance objective. Let f_l^i denote the expected flow that user i sends on link l. The user flow configuration $\boldsymbol{f}^i = (f_1^i, \cdots, f_L^i)$ is called a routing *strategy* of user i, and the set $F^i = \{\boldsymbol{f}^i \in \mathbb{R}^L\colon 0 \leq f_l^i \leq c_l, l \in \mathcal{L}; \Sigma_{l \in \mathcal{L}}\ f_l^i = r^i\}$ of strategies that satisfy the user's demand is called the strategy space of user i. The system flow configuration $\boldsymbol{f} = (\boldsymbol{f}^1, \cdots, \boldsymbol{f}^I)$ is called a routing *strategy profile* and takes values in the product strategy space $F = \otimes_{i \in \mathcal{I}} F^i$.

The performance objective of user i is quantified by means of a cost function $J^i(\boldsymbol{f})$. The user aims to find a strategy $\boldsymbol{f}^i \in F^i$ that minimizes its cost. This optimization problem depends on the routing decisions of the other users, described by the strategy profile $\boldsymbol{f}^{-i} = (\boldsymbol{f}^1, \cdots, \boldsymbol{f}^{i-1}, \boldsymbol{f}^{i+1}, \cdots, \boldsymbol{f}^I)$, since J^i is a function of the system flow configuration $\boldsymbol{f}$. A

Nash equilibrium of the routing game is a strategy profile from which no user finds it beneficial to unilaterally deviate. Hence, $\boldsymbol{f} \in F$ is a Nash equilibrium if

$$\boldsymbol{f}^i \in \arg \min_{\boldsymbol{g}^i \in F^i} J^i(\boldsymbol{g}^i, \boldsymbol{f}^{-i}), \qquad i \in \mathcal{I}. \tag{1}$$

The problem of existence and uniqueness of equilibria has been investigated in [9] for certain general classes of cost functions. Here, we consider cost functions that are the sum of link cost functions

$$J^i(\boldsymbol{f}) = \sum_{l \in \mathcal{L}} J_l^i(\boldsymbol{f}_l), \quad J_l^i(\boldsymbol{f}_l) = f_l^i T_l(f_l), \qquad l \in \mathcal{L} \tag{2}$$

where $\boldsymbol{f}_l = (f_l^1, \cdots, f_l^I)$, and $T_l(f_l)$ is the average delay per unit of flow on link l that depends only on the total flow $f_l = \Sigma_{i \in \mathcal{I}} f_l^i$ on that link. The average delay should be interpreted as a general *congestion cost* per unit of flow that encapsulates the dependence of the quality of a finite capacity resource on the total load f_l offered to it (see [24] for a related discussion). In the present paper, we concentrate on congestion costs of the form

$$T_l(f_l) = \begin{cases} (c_l - f_l)^{-1}, & f_l < c_l \\ \infty, & f_l \geq c_l \end{cases} \tag{3}$$

that are typical in various practical routing algorithms [25], [12].

Note that (3) describes the M/M/1 delay function. Therefore, if we assume that the delay characteristics of each link can be approximated by an M/M/1 queue, $J^i(\boldsymbol{f})/r^i$ is the average time-delay that the flow of user i experiences under strategy profile $\boldsymbol{f}$. Also, note that the stability constraint $f_l < c_l$ of link l is manifested through the definition of T_l. In particular, since the total user demand R does not exceed the total capacity C of the network, (1) and (3) guarantee that at any Nash equilibrium, we have $f_l < c_l$ for all $l \in \mathcal{L}$, and the costs of all users are finite.

Given a strategy profile $\boldsymbol{f}^{-i}$ of the other users, the cost of user i, as defined by (2) and (3), is a convex function of its strategy $\boldsymbol{f}^i$. Hence, the minimization problem in (1) has a unique solution. Since $C > R$, the Slater condition [26] is satisfied, therefore the Kuhn–Tucker optimality conditions are applicable. These conditions imply that $\boldsymbol{f}^i$ is the optimal response of user i to $\boldsymbol{f}^{-i}$ if and only if there exist (Lagrange multipliers) λ^i and $\boldsymbol{\mu}^i = (\mu_1^i, \cdots, \mu_L^i)$, such that

$$\frac{\partial J^i}{\partial f_l^i}(\boldsymbol{f}) - \lambda^i - \mu_l^i = 0, \qquad l \in \mathcal{L} \tag{4}$$

$$\sum_{l \in \mathcal{L}} f_l^i = r^i \tag{5}$$

$$\mu_l^i f_l^i = 0, \qquad l \in \mathcal{L} \tag{6}$$

$$\mu_l^i \geq 0, \quad f_l^i \geq 0, \qquad l \in \mathcal{L} \tag{7}$$

where $\boldsymbol{f} = (\boldsymbol{f}^i, \boldsymbol{f}^{-i})$. Therefore, a strategy profile $\boldsymbol{f} \in F$ is a Nash equilibrium if and only if there exist λ^i and $\boldsymbol{\mu}^i$, such that the optimality conditions (4)–(7) are satisfied for all $i \in \mathcal{I}$.

It is easy to verify that the necessary and sufficient conditions (4)–(7) are equivalent to the following:

$$\lambda^i = \frac{\partial J^i}{\partial f_l^i}(\boldsymbol{f}), \quad \text{if } f_l^i > 0, \qquad l \in \mathcal{L} \tag{8}$$

$$\lambda^i \leq \frac{\partial J^i}{\partial f_l^i}(\boldsymbol{f}), \quad \text{if } f_l^i = 0, \qquad l \in \mathcal{L} \tag{9}$$

$$\sum_{l \in \mathcal{L}} f_l^i = r^i, \quad f_l^i \geq 0, \qquad l \in \mathcal{L} \tag{10}$$

which imply that λ^i is, in fact, the marginal cost of user i at the optimality point. In accordance with the economics terminology [27], λ^i will be referred to as the *price* of user i.

For the cost function $J^i(\boldsymbol{f})$ given by (2) and (3), we have

$$\frac{\partial J^i}{\partial f_l^i}(\boldsymbol{f}) = f_l^i T_l'(f_l) + T_l(f_l) = \frac{c_l - f_l^{-i}}{(c_l - f_l)^2} \tag{11}$$

where T_l' is the derivative of T_l with respect to f_l, and $f_l^{-i} = \Sigma_{j \neq i} f_l^j$ is the total flow that all users except the ith send on link l. Note that $T_l' = T_l^2$.

In [9] it has been shown that the routing game described above has a *unique* Nash equilibrium.

At times we will concentrate on special types of users, defined in the following.

Definition 1: Users are said to be *identical* if their demands are all equal, i.e., $r^i = r^j, i, j \in \mathcal{I}$.

The Nash equilibrium of identical users is symmetrical, i.e., $f_l^i = f_l^j = f_l/I$ for all $i, j \in \mathcal{I}$ [9].

Definition 2: A user is said to be *simple* if all of its flows are routed through links (or paths) of minimal delay.

Users often route their flows according to the “simple” scheme due to practical considerations. Many typical routing algorithms send flows through shortest paths, without accounting for derivatives (T_l'), and thus bifurcating flows. The Nash equilibrium of simple users in a system of parallel links is unique with respect to the *total* link flows [9], and the corresponding necessary and sufficient conditions require the existence of some λ such that

$$\lambda = T_l, \quad \text{if } f_l > 0, \qquad l \in \mathcal{L} \tag{12}$$

$$\lambda \leq T_l, \quad \text{if } f_l = 0, \qquad l \in \mathcal{L} \tag{13}$$

$$\sum_{l \in \mathcal{L}} f_l = R, \quad f_l \geq 0, \qquad l \in \mathcal{L}. \tag{14}$$

We shall refer to the value of λ as the price of the simple users. From (12)–(14), it is easy to see that users that route according to the optimality conditions (8)–(10) become simple as their population grows to infinity and their individual demands become infinitesimally small, while their total demand remains R. This is the typical scenario in a transportation network.

Definition 3: Users are said to be *consistent* (for a given capacity configuration) if, at the Nash equilibrium, they all use the same set of links.

Due to the structure of their Nash equilibrium, identical users are consistent. It is easy to verify that simple users are also consistent [9]. Finally, consistent users are typical of systems with heavy traffic, i.e., when R approaches C, in which case each user sends flow on all links in the network.

B. Capacity Allocation Problem

Consider a network of parallel links with initial capacity configuration $\boldsymbol{c}^0$ and total capacity $C^0 > R$. We assume that $c_1^0 \geq \cdots \geq c_L^0$. Suppose that there exists some additional capacity allowance of at most Δ, which the network designer can distribute among the network links. The aim of the designer is to come up with a capacity configuration $\boldsymbol{c}$, with $c_l \geq c_l^0$ for all links $l \in \mathcal{L}$, that results in a network with a total capacity of at most $C^0 + \Delta$ that is "efficient" at the corresponding Nash equilibrium. Without loss of generality, we can concentrate on capacity configurations $\boldsymbol{c}$ that preserve the initial link order, that is, configurations with $c_1 \geq \cdots \geq c_L$.[3] Therefore, the set of all capacity configurations that can be implemented by the designer is $\mathcal{C} = \{\boldsymbol{c} \in \mathbb{R}_+^L\colon c_1 \geq \cdots \geq c_L; c_l \geq c_l^0, l \in \mathcal{L}; \Sigma_{l \in \mathcal{L}}\, (c_l - c_l^0) \leq \Delta\}$. Each capacity configuration in $\mathcal{C}$ induces a different routing game that has a unique Nash equilibrium. Therefore, we can define a function $\mathcal{N}\colon \mathcal{C} \rightarrow F$ that assigns to each $\boldsymbol{c} \in \mathcal{C}$ the Nash equilibrium $\mathcal{N}(\boldsymbol{c})$ of its respective routing game. $\mathcal{N}$ will be referred to as the *Nash mapping*.

The designer may have different measures for characterizing the efficiency of a capacity configuration. We shall concentrate on measures that are expressed by means of either the user prices or costs. Although the user's cost is a direct measure of its level of satisfaction, prices may be a more important measure from the system's point of view since they account for the level of congestion as seen by users and are the direct indication of how each user could accommodate fluctuations in the system's state. The designer can consider various ways of combining either the prices or the costs of the users. We shall concentrate on *user* optimization, i.e., trying to reduce the price or cost of each and every user, and *overall* optimization, i.e., trying to reduce the sum of all prices or costs. The various performance measures of the designer are formally stated in the following definitions.

Definition 4: Consider two capacity configurations $\boldsymbol{c}$ and $\hat{\boldsymbol{c}}$, and let λ^i and $\hat{\lambda}^i$ (correspondingly, J^i and $\hat{J}^i$) be the price (correspondingly, cost) of user i at the respective equilibrium. Then we have the following.

1) Configuration $\hat{\boldsymbol{c}}$ is said to be the *user price (cost) efficient* relative to configuration $\boldsymbol{c}$, if $\hat{\lambda}^i \leq \lambda^i (\hat{J}^i \leq J^i)$, for all $i \in \mathcal{I}$.
2) Configuration $\hat{\boldsymbol{c}}$ is said to be the *overall price (cost) efficient* relative to configuration $\boldsymbol{c}$, if $\Sigma_{i \in \mathcal{I}}\ \hat{\lambda}^i \leq \Sigma_{i \in \mathcal{I}}\ \lambda^i (\Sigma_{i \in \mathcal{I}}\ \hat{J}^i \leq \Sigma_{i \in \mathcal{I}}\ J^i)$.

Definition 5: A capacity configuration $\boldsymbol{c}^* \in \mathcal{C}$ is called:

1) *user price (cost) optimal* in $\mathcal{C}$, if it is user price (cost) efficient relative to any $\boldsymbol{c} \in \mathcal{C}$;
2) *overall price (cost) optimal* in $\mathcal{C}$, if it is overall price (cost) efficient relative to any $\boldsymbol{c} \in \mathcal{C}$.

Obviously, user efficiency (optimality) implies overall efficiency (optimality). Price and cost efficiency (optimality), however, do not imply each other in either direction. Note also that, in general, the existence of user optima cannot be guaranteed even if overall optima do exist.

The optimal capacity allocation problem, corresponding to the various designer's performance measures, is described as follows.

[3] The properties of the Nash equilibrium in a system of parallel links with capacity configuration $\boldsymbol{c}$ depends on the actual link capacities and not on the link "labels" that are determined by the initial configuration $\boldsymbol{c}^0$. Hence, renaming the links, so that $c_1 \geq \cdots \geq c_L$, does not affect the characteristics of the resulting equilibrium.

Given a system of parallel links $\mathcal{L}$ with users $\mathcal{I}$, an initial capacity configuration $\boldsymbol{c}^0$, and an additional capacity allowance Δ, find a capacity configuration $\boldsymbol{c}^*$ that is user/overall price/cost optimal in $\mathcal{C}$.

Although the problem is formulated as allocating additional capacity to an existing network, this formulation is equivalent to the typical capacity allocation problem, where the capacity of each link has to be higher than a lower bound, e.g., due to reliability considerations. By definition, the initial capacity c_l^0 of every link l is positive, in other words, the designer can only add capacity to existing links. Nonetheless, as shown in [28], the results of the following subsections can be easily extended to the case where $c_l^0 = 0$ for some links $l \in \mathcal{L}$, that is, when the designer is also allowed to add a (finite) number of links to the network.

Solving the optimal capacity allocation problem in a network shared by noncooperative users amounts to comparing the Nash equilibria of the routing games induced by different capacity configurations in $\mathcal{C}$. Comparing the outcomes of different games is, in general, a highly complex task and requires explicit characterization of the respective equilibria. The structure of the unique Nash equilibrium of the routing game is investigated in the following section. Before we proceed, let us first summarize the main results of this study.

C. Outline of Results

1) Addition of capacity to a link results in a configuration that is:
   a) user (thus, overall) price efficient;
   b) user (thus, overall) cost efficient for consistent users;
   c) overall cost efficient when capacity is added to the link with the initially highest capacity.
2) Transferring capacity from any link to the link with the initially highest capacity results in a configuration that is:
   a) user (thus, overall) price efficient;
   b) user (thus, overall) cost efficient for consistent users and for two users.
3) The capacity configuration that results from allocating the entire additional capacity allowance Δ to the link with the initially highest capacity is:
   a) user (thus, overall) price optimal in $\mathcal{C}$;
   b) user (thus, overall) cost optimal in $\mathcal{C}$ for each of the following cases:
      - identical lower bounds on the link capacities;
      - consistent users;
      - two users.

III. Structure of the Nash Mapping

In this section we study the structure of the Nash mapping $\mathcal{N}$ that assigns to each capacity configuration $\boldsymbol{c} \in \mathcal{C}$ the Nash equilibrium $\mathcal{N}(\boldsymbol{c})$ of its respective routing game. We start by investigating the structure of the Nash equilibrium for a given capacity configuration $\boldsymbol{c}$. These results will be used to establish continuity properties of the Nash mapping.

A. Structure of the Nash Equilibrium

Consider the Nash equilibrium of the routing game in a system of parallel links with capacity configuration $\boldsymbol{c}$. A number of intuitive monotonicity properties of this equilibrium have been established in [9] and are summarized in the following.

Lemma 1: Let $\boldsymbol{f}$ be the unique Nash equilibrium of the routing game in a network of parallel links with capacity configuration $\boldsymbol{c}$. Then:

1) the expected flow of any user $i \in \mathcal{I}$ decreases in the link number, i.e., $f_1^i \geq f_2^i \geq \cdots \geq f_L^i$. In particular, for $f_l^i > 0$, we have $f_l^i = f_m^i$ if and only if $c_l = c_m$;
2) for any link $l \in \mathcal{L}$, the flows decrease in the user number, i.e., $f_l^1 \geq f_l^2 \geq \cdots \geq f_l^I$. In particular, for $f_l^i > 0$, we have $f_l^i = f_l^j$ if and only if $r^i = r^j$;
3) the residual capacity is decreasing in the link number, i.e., $c_1 - f_1 \geq c_2 - f_2 \geq \cdots \geq c_L - f_L$, or equivalently, $T_1 \leq T_2 \leq \cdots \leq T_L$. In particular, $T_l = T_m$ if and only if $c_l = c_m$;
4) for every user $i \in \mathcal{I}$, the residual capacity $c_l^i = c_l - f_l^{-i}$ seen by the user on link, l is decreasing in the link number, i.e., $c_1^i \geq c_2^i \geq \cdots \geq c_L^i$. In particular, $c_l^i = c_m^i$ if and only if $c_l = c_m$.

Let $\mathcal{L}^i$ denote the set of links that receive some flow from user i, and $\mathcal{I}_l$ denotes the set of users that send flow over link l. The first statement in Lemma 1 implies that for every user i, there exists some link L^i, such that $f_l^i > 0$ for all $l \leq L^i$, and $f_l^i = 0$ for $l > L^i$, that is, $\mathcal{L}^i = \{1, 2, \cdots, L^i\}$. Similarly, the second statement in the Lemma implies that for every link l, there exists some user I_l such that $f_l^i > 0$ for all $i \leq I_l$, and $f_l^i = 0$ for $i > I_l$, that is, $\mathcal{I}_l = \{1, 2, \cdots, I_l\}$. Moreover, $\mathcal{L}^{i+1} \subseteq \mathcal{L}^i (1 \leq i < I)$ and $\mathcal{I}_{l+1} \subseteq \mathcal{I}_l (1 \leq l < L)$.

Consider the best reply $\boldsymbol{f}^i$ of user i to a fixed strategy profile $\boldsymbol{f}^{-i}$ of the other users. This is the unique solution to the (single-user) optimal routing problem for a network of parallel links with capacity configuration $\boldsymbol{c}^i = (c_1^i, \cdots, c_L^i)$ and is determined by the Kuhn–Tucker optimality conditions (8)–(10). Note that for any link $l \in \mathcal{L}$, conditions (8) and (9) can be written as

$$\lambda^i = \frac{c_l^i}{(c_l^i - f_l^i)^2} = \frac{c_l^i}{(c_l - f_l)^2}, \quad \text{if } \; l \leq L^i \tag{15}$$

$$\lambda^i \leq \frac{c_l^i}{(c_l^i - f_l^i)^2} = \frac{1}{c_l - f_l} = \frac{1}{c_l^i}, \quad \text{if } \; l > L^i. \tag{16}$$

In the sequel, we will derive an explicit characterization of the structure of the user's equilibrium strategy $\boldsymbol{f}^i$ as a function of $\boldsymbol{c}^i$, which depends on the capacity configuration $\boldsymbol{c}$ and the strategy profile $\boldsymbol{f}^{-i}$ of the other users. To this end, let us define

$$G_l^i = \sum_{m=1}^{l-1} c_m^i - \sqrt{c_l^i} \sum_{m=1}^{l-1} \sqrt{c_m^i}, \qquad l = 2, \cdots, L \tag{17}$$

and $G_1^i = 0, G_{L+1}^i = \Sigma_{m=1}^L c_m^i = C - R^{-i}$, where $R^{-i} = \Sigma_{j \neq i} r^j$ is the total demand of all users except the ith. Note that $C - R^{-i}$ is the total residual capacity of the network as seen by user i. Since $c_l^i \geq c_{l+1}^i$, it is easy to verify that

$$G_l^i \leq G_{l+1}^i, \qquad l \in \mathcal{L} \tag{18}$$

with equality holding if and only if $c_l = c_{l+1}$. We are now ready to show the following.

Proposition 1: The Nash equilibrium $\boldsymbol{f}$ of the routing game in a system of parallel links with capacity configuration $\boldsymbol{c}$ satisfies the following relationship:

$$f_l^i = \begin{cases} c_l^i - \left(\displaystyle\sum_{m=1}^{L^i} c_m^i - r^i \right) \dfrac{\sqrt{c_l^i}}{\Sigma_{m=1}^{L^i} \sqrt{c_m^i}}, & l \leq L^i \\ 0, & l > L^i \end{cases} \tag{19}$$

where, for every user $i \in \mathcal{I}$, the threshold L^i is determined by

$$G_{L^i}^i < r^i \leq G_{L^i+1}^i. \tag{20}$$

The equilibrium price and the equilibrium cost for user i are, respectively

$$\lambda^i = \left[\frac{\Sigma_{l \in A} \sqrt{c_l^i}}{\Sigma_{l \in A} (c_l^i - f_l^i)} \right]^2, \quad \text{for any set } \; A \subseteq \mathcal{L}^i \tag{21}$$

$$= \left[\frac{\Sigma_{l=1}^{L^i} \sqrt{c_l^i}}{\Sigma_{l=1}^{L^i} c_l^i - r^i} \right]^2 \tag{22}$$

$$J^i = \lambda^i \sum_{l=1}^{L^i} (c_l - f_l) - L^i = \frac{\left[\Sigma_{l=1}^{L^i} \sqrt{c_l^i} \right]^2}{\Sigma_{l=1}^{L^i} c_l^i - r^i} - L^i. \tag{23}$$

Proof: See Appendix A. ■

Remarks:

1) The proposition implies that the information user i needs to determine its best reply $\boldsymbol{f}^i$ to any strategy profile $\boldsymbol{f}^{-i}$ of the other users in the residual capacity c_l^i seen by the user on every link $l \in \mathcal{L}$ [see (19) and (17)], and not a detailed description of $\boldsymbol{f}^{-i}$. In practice, information about the residual capacities can be acquired by measuring the link delays through an appropriate estimation technique.
2) In the special case $r^i = G_{L^i+1}^i$, (17) and (22) imply that $\lambda^i = 1/c_{L^i+1}^i$, and (16) holds tight for $l = L^i + 1$. Therefore, in this case, we can define the set of links on which the user sends flow as $\mathcal{L}^i = \{1, \cdots, L^i, L^i + 1\}$, where link $L^i + 1$ is "marginally" used with $f_{L^i+1}^i = 0$.

The structure of the Nash equilibrium of the routing game is exploited in the following section to show that the Nash mapping $\mathcal{N}$ is continuous. This fundamental property will substantially simplify the analysis of the optimal capacity allocation problem in the subsequent sections.

B. Continuity of the Nash Mapping

From Proposition 1, and especially the expressions for the equilibrium prices and costs, it is clear that the set of links over which each user sends its flow has a prominent role in the properties of the Nash equilibrium. To investigate the capacity allocation problem, we need to compare the equilibria of games that are induced by different capacity configurations in $\mathcal{C}$. If the resulting equilibria are such that the sets of links over which each user sends its flow do not coincide at both equilibria, such comparisons are extremely complex, if possible at all. In this section, we first show that the Nash mapping $\mathcal{N}$ is continuous and then explain that this result allows us to investigate the general capacity allocation problem, based solely on comparisons between capacity configurations that are such that each user sends its flow over the same links under both configurations.

Consider a fixed capacity configuration $\boldsymbol{c} \in \mathcal{C}$, and let $\boldsymbol{f}$ be its corresponding Nash equilibrium. The price λ^i of user i at this equilibrium is unique. Similarly, $\boldsymbol{\mu}^i$ is uniquely determined by (4). Therefore, there exists a unique collection of Lagrange multipliers $(\boldsymbol{\lambda}, \boldsymbol{\mu}), \boldsymbol{\lambda} = (\lambda^i)_{i \in \mathcal{I}}, \boldsymbol{\mu} = (\boldsymbol{\mu}^i)_{i \in \mathcal{I}}$ that, together with the Nash equilibrium $\boldsymbol{f}$, solve the system of necessary and sufficient conditions (8)–(10) for all $i \in \mathcal{I}$. Let us now augment the definition of the Nash mapping so that to each capacity configuration $\boldsymbol{c}$ we assign the Nash equilibrium $\boldsymbol{f}$ of the routing game and the corresponding Lagrange multipliers $(\boldsymbol{\lambda}, \boldsymbol{\mu})$, that is $\mathcal{N}: \mathcal{C} \rightarrow F \otimes \mathbb{R}^{I(L+1)}$, with $\mathcal{N}(\boldsymbol{c}) = (\boldsymbol{f}, \boldsymbol{\lambda}, \boldsymbol{\mu})$.

Theorem 1: The Nash mapping $\mathcal{N}: \mathcal{C} \rightarrow F \otimes \mathbb{R}^{I(L+1)}$ is continuous.

Proof: See Appendix A. ∎

The following corollary shows that the equilibrium costs of the users and the equilibrium link delays are also continuous functions of the capacity configuration $\boldsymbol{c}$.

Corollary 1: Let $\Phi_1: \mathcal{C} \rightarrow \mathbb{R}_+^I$ and $\Phi_2: \mathcal{C} \rightarrow \mathbb{R}_+^L$ be such that for every $\boldsymbol{c} \in \mathcal{C}, \Phi_1(\boldsymbol{c}) = (J^1, \cdots, J^I)$, where J^i is the equilibrium cost of user i under capacity configuration $\boldsymbol{c}$, and $\Phi_2(\boldsymbol{c}) = (T_1, \cdots, T_L)$, where T_l is the equilibrium delay of link l under $\boldsymbol{c}$. Then, Φ_1 and Φ_2 are continuous.

Proof: From (2) and (3), one can see that the link delays and the user cost functions are continuous at every $(\boldsymbol{c}, \boldsymbol{f})$, as long as the stability condition $f_l < c_l$ for all links $l \in \mathcal{L}$ is satisfied. Suppose now that $\boldsymbol{f}$ is the Nash equilibrium under capacity configuration $\boldsymbol{c}$. Then, as explained in Section II-A, the stability condition is satisfied, and since $\boldsymbol{f}$ is a continuous function of $\boldsymbol{c}$ (Theorem 1) the result follows. ∎

As explained in Section I, we will investigate the optimal capacity allocation problem by decomposing it into two subproblems, namely the problem of adding capacity to any link and the problem of transferring capacity from one link to another. In the rest of this section we explain that the continuity properties of the Nash mapping allow us to analyze these problems under the assumption that each user sends its flow over the same set of links before and after the capacity addition/transfer. We will concentrate on the case of capacity transfer; the analysis can be readily adopted to the problem of capacity addition.

Consider two capacity configurations $\boldsymbol{c}, \hat{\boldsymbol{c}} \in \mathcal{C}$, such that $\hat{\boldsymbol{c}}$ results from $\boldsymbol{c}$ by shifting an amount of capacity Δ_q from some link q to a link l, i.e., $\hat{\boldsymbol{c}} = \boldsymbol{c} + \Delta_q(\boldsymbol{e}_l - \boldsymbol{e}_q)$.[4] For every $\delta \in [0, \Delta_q]$, let $\boldsymbol{c}(\delta) = \boldsymbol{c} + \delta(\boldsymbol{e}_l - \boldsymbol{e}_q)$ be the capacity configuration that results from $\boldsymbol{c}$ by a transfer of capacity δ from link q to link l. All quantities of interest, e.g., the equilibrium prices and costs, can be treated as functions of δ. Let h be such a quantity, and suppose that we aim at showing that $h(0) \leq h(\Delta_q)$, i.e., that h is higher under configuration $\hat{\boldsymbol{c}} = \boldsymbol{c}(\Delta_q)$ than under $\boldsymbol{c} = \boldsymbol{c}(0)$. To achieve this goal, it suffices to show that h is a nondecreasing function of $\delta \in [0, \Delta_q]$.

The set of links that receive flow from user i is determined by (20). From (17), one can see that G_l^i is a continuous function of the capacity configuration $\boldsymbol{c}$ and the equilibrium strategies of the other users $\boldsymbol{f}^{-i}$. The continuity of the Nash mapping, then, implies that G_l^i is a continuous function of $\delta \in [0, \Delta_q]$. Let $A_l^i = \{\delta \in [0, \Delta_q]: G_l^i(\delta) \leq r^i \leq G_{l+1}^i(\delta)\}$ denote the set of $\delta \in [0, \Delta_q]$ for which user i sends flow on links $\{1, \cdots, l\}$ under configuration $\boldsymbol{c}(\delta)$.[5] Continuity of $G_l^i, l \in \mathcal{L}$ implies that A_l^i is a closed set [29]. Define $A_{l_1, \cdots, l_I} = \cap_{i \in \mathcal{I}} A_{l_i}^i$, which is also a closed set. If $\delta_1, \delta_2 \in A_{l_1, \cdots, l_I}$, for some $(l_1, \cdots, l_I)$, then each user sends its flow over the same set of links under configurations $\boldsymbol{c}(\delta_1)$ and $\boldsymbol{c}(\delta_2)$. Note that $\cup_{i \in \mathcal{I}} \cup_{l_i \in \mathcal{L}} A_{l_1, \cdots, l_I} = [0, \Delta_q]$.

Theorem 5 in Appendix A implies that, to prove that h is nondecreasing in $[0, \Delta_q]$, it suffices to establish this property in every set $A_{l_1, \cdots, l_I} \subset [0, \Delta_q]$, i.e., to show that for every $\delta_1, \delta_2 \in A_{l_1, \cdots, l_I}, \delta_1 < \delta_2$ implies that $h(\delta_1) \leq h(\delta_2)$. In other words, all comparisons between $\boldsymbol{c}$ and $\hat{\boldsymbol{c}} = \boldsymbol{c} + \Delta_q(\boldsymbol{e}_l - \boldsymbol{e}_q)$ can be carried based on the assumption that each user sends its flow over the same set of links under both configurations.

IV. Capacity Addition

As previously mentioned, in [28] we present an example that adopts the Braess paradox to the communication network framework considered in this paper. That example demonstrates that addition of capacity to a network may increase user prices and/or costs. In this section we investigate the problem of adding capacity to systems of parallel links and show that, under various conditions, this paradoxical behavior cannot occur in this setting.

A capacity configuration $\hat{\boldsymbol{c}}$ is called an *augmentation* of configuration $\boldsymbol{c}$, if $\hat{c}_l \geq c_l$ for all l and $\Sigma_l \hat{c}_l > \Sigma_l c_l$. Throughout this section we shall compare the Nash equilibrium of a capacity configuration $\boldsymbol{c}$ to that of some augmentation $\hat{\boldsymbol{c}}$. "Hat" values will refer to configuration $\hat{\boldsymbol{c}}$, while "nonhat" values refer to $\boldsymbol{c}$. For example, $\hat{\lambda}^i$ and λ^i are the prices of user i under $\hat{\boldsymbol{c}}$ and $\boldsymbol{c}$, respectively.

A. Price Efficiency

The following proposition shows that an addition of capacity is always price efficient.

Proposition 2: If a capacity configuration $\hat{\boldsymbol{c}}$ is an augmentation of configuration $\boldsymbol{c}$, then $\hat{\boldsymbol{c}}$ is user price efficient relative to $\boldsymbol{c}$, i.e., $\hat{\lambda}^i \leq \lambda^i$, for all $i \in \mathcal{I}$. Moreover, the equilibrium delay of each link l is lower (not higher) under configuration $\hat{\boldsymbol{c}}$, i.e., $\hat{T}_l \leq T_l$, for all $l \in \mathcal{L}$.

Proof: Assume by contradiction that the set $\mathcal{T}^+ = \{l \in \mathcal{L}: \hat{T}_l > T_l\}$ is nonempty. Since the flow in each $l \in \mathcal{T}^+$ is higher under configuration $\hat{\boldsymbol{c}}$, there must be a user i and links $l \in \mathcal{T}^+$ and $n \notin \mathcal{T}^+$ such that $\hat{f}_l^i > f_l^i$ and $\hat{f}_n^i < f_n^i$. Since $\hat{f}_l^i > f_l^i \geq 0$, the optimality conditions (8)–(10) imply that $\hat{f}_n^i \hat{T}_n' + \hat{T}_n \geq \hat{\lambda}^i = \hat{f}_l^i \hat{T}_l' + \hat{T}_l$. Similarly, since $f_n^i > \hat{f}_n^i \geq 0$, we have $f_n^i T_n' + T_n = \lambda^i \leq f_l^i T_l' + T_l$. Combining these

[4] e_l is the vector in $\mathbb{R}^L$ with the lth component equal to one and all other components equal to zero.

[5] If $G_l^i(\delta) = r^i$, user i marginally uses link l under $\boldsymbol{c}(\delta)$; see the second remark following Proposition 1.

inequalities with $\hat{T}_n \le T_n$ and $\hat{f}_n^i < f_n^i$, we get $\hat{f}_l^i \hat{T}_l' + \hat{T}_l \le f_l^i T_l' + T_l$. Since $\hat{f}_l^i > f_l^i$ and $\hat{T}_l > T_l$, this is a contradiction. Hence, the set $\mathcal{T}^+$ is empty, i.e., $\hat{T}_l \le T_l$ for all links l.

Since the demand of each user i is r^i in both capacity configurations, the user cannot increase its flow on each and every link, and there must be a link l for which $f_l^i > 0$ and $\hat{f}_l^i \le f_l^i$. Therefore, $\hat{\lambda}^i \le \hat{f}_l^i \hat{T}_l' + \hat{T}_l \le f_l^i T_l' + T_l = \lambda^i$, thus concluding the proof. ∎

B. Cost Efficiency

The following proposition shows that if capacity is added exclusively to link 1, the resulting configuration is overall cost efficient.

Proposition 3: Let $\boldsymbol{c}$ and $\hat{\boldsymbol{c}}$ be two capacity configurations such that $\hat{c}_l = c_l$ for all $l > 1$, and $\hat{c}_1 > c_1$. Then $\hat{\boldsymbol{c}}$ is overall cost efficient relative to $\boldsymbol{c}$.

Proof: From Proposition 2 we have that $\hat{T}_l \le T_l$ for all links l. This means that for all $l > 1$, we have $\hat{f}_l \le f_l$. Therefore, when "moving" from $\boldsymbol{c}$ to $\hat{\boldsymbol{c}}$, we observe flow being "transferred" from all links $l > 1$ to link 1. Since $\hat{T}_l \le T_l$ for all links l, the flow that remains (under configuration $\hat{\boldsymbol{c}}$) in a link $l > 1$, experiences a delay which is not higher than the previous one. Moreover, since $\hat{T}_1 \le T_1 \le T_l$, we also conclude that the flow that moved to link 1 experiences a delay which is not higher than the previous one, and the result follows. ∎

The following two propositions, whose proofs appear in Appendix B, establish user cost efficiency of capacity addition in some special cases of interest.

Proposition 4: Let $\boldsymbol{c}$ and $\hat{\boldsymbol{c}}$ be two capacity configurations such that $\hat{\boldsymbol{c}}$ is an augmentation of $\boldsymbol{c}$. Assume that users are consistent under both $\hat{\boldsymbol{c}}$ and $\boldsymbol{c}$. Then $\hat{\boldsymbol{c}}$ is user cost efficient relative to $\boldsymbol{c}$, that is, $\hat{J}^i \le J^i$, for all $i \in \mathcal{I}$.

The above result applies, in particular, both to identical users and to simple users, since they belong to the class of consistent users under all capacity configurations.

Proposition 5: Let $\boldsymbol{c}$ and $\hat{\boldsymbol{c}}$ be two capacity configurations such that $\hat{c}_l = c_l$ for all $l > 1$, and $\hat{c}_1 > c_1$. Then, for $I = 2, \hat{\boldsymbol{c}}$ is user cost efficient relative to $\boldsymbol{c}$.

Remark: We have a proof that an augmented capacity configuration is user cost efficient also for the (dual) case of two links ($L = 2$) and any number of users.

V. Capacity Transfer

In this section we investigate the problem of transferring capacity from one link to another. Specifically, we establish that transferring capacity from any link (while observing the lower bound) to link 1 improves performance according the various efficiency criteria defined in Section II-B. Except for being a stage toward the solution of the optimal allocation problem, this result provides an interesting design rule per se. In broadband networks, for example, capacity is routinely released upon the completion of a session. The network management may redistribute the excess capacity among virtual paths or circuits, thus facing the problem considered in this section.

Consider two capacity configurations $\boldsymbol{c}$ and $\hat{\boldsymbol{c}}$ in $\mathcal{C}$, such that $\hat{\boldsymbol{c}} = \boldsymbol{c} + \Delta_q(\boldsymbol{e}_1 - \boldsymbol{e}_q)$ is derived from $\boldsymbol{c}$ by a transfer of capacity $\Delta_q (0 < \Delta_q \le c_q - c_q^0)$ from some link q, with $c_q < c_1$, to link 1. Note that the total capacity of the system remains the same after the capacity transfer, i.e., $\Sigma_{l \in \mathcal{L}} \hat{c}_l = \Sigma_{l \in \mathcal{L}} c_l = C$. As before, "hat" values will refer to configuration $\hat{\boldsymbol{c}}$, while "nonhat" values refer to the initial configuration $\boldsymbol{c}$.

A. Price Efficiency

We begin by showing that capacity configuration $\hat{\boldsymbol{c}}$ is user price efficient relative to $\boldsymbol{c}$. The comparison of configurations $\hat{\boldsymbol{c}}$ and $\boldsymbol{c}$ is carried out in a series of lemmas. Lemma 2 examines the effect of the transfer of capacity from link q to link 1 on the equilibrium delays of these two links. Lemmas 3 and 4 show that the transfer of capacity affects the prices of the users and the equilibrium delays of the links in $\mathcal{L} \setminus \{1, q\}$ in an "ordered" way, in a sense that will be explained in the lemmas. Finally, user price efficiency of $\hat{\boldsymbol{c}}$ relative to $\boldsymbol{c}$ will be established in Theorem 2.

Let $\mathcal{T}^+$ (respectively, $\mathcal{T}^-$) denote the set of links in $\mathcal{L} \setminus \{1, q\}$ whose equilibrium delay is higher (respectively, not higher) under $\hat{\boldsymbol{c}}$, that is, $\mathcal{T}^+ = \{l \in \mathcal{L} \setminus \{1, q\}\colon \hat{T}_l > T_l\}$ and $\mathcal{T}^- = \{l \in \mathcal{L} \setminus \{1, q\}\colon \hat{T}_l \le T_l\}$. Consider now any link $l \in \mathcal{L} \setminus \{1, q.\}$ Since $c_l = \hat{c}_l$, we have that $l \in \mathcal{T}^+$ if and only if $\hat{f}_l > f_l$, while $l \in \mathcal{T}^-$ if and only if $\hat{f}_l \le f_l$.

The following lemma shows that the transfer of capacity from link q to link 1 decreases the equilibrium delay on link 1, while it increases the delay on link q.

Lemma 2: Consider two capacity configurations $\boldsymbol{c}, \hat{\boldsymbol{c}} \in \mathcal{C}$ with $\hat{\boldsymbol{c}} = \boldsymbol{c} + \Delta_q(\boldsymbol{e}_1 - \boldsymbol{e}_q)$. Then, $\hat{T}_1 \le T_1$ and $\hat{T}_q > T_q$.

Proof: See Appendix C. ∎

In the sequel, we present two lemmas that will play a key role in the proof of price efficiency of $\hat{\boldsymbol{c}}$ relative to $\boldsymbol{c}$. Both refer to the case where the transfer of capacity Δ_q from link q to link 1 is such that each user sends its flow over the same set of links under $\boldsymbol{c}$ and $\hat{\boldsymbol{c}}$, i.e., $\hat{\mathcal{L}}^i = \mathcal{L}^i$ for all $i \in \mathcal{I}$, and $\hat{\mathcal{I}}_l = \mathcal{I}_l$ for all $l \in \mathcal{L}$. The first lemma asserts that the transfer affects the prices of all users that send flow to link q in the same way, that is, either $\hat{\lambda}^i > \lambda^i$ for all $i \in \mathcal{I}_q$ or $\hat{\lambda}^i \le \lambda^i$ for all $i \in \mathcal{I}_q$. Similarly, either all links with capacity lower than link q increase their equilibrium delays with $\hat{\boldsymbol{c}}$, or else all of them decrease their equilibrium delays. The proofs of both lemmas are presented in Appendix C.

Lemma 3: Consider two capacity configurations $\boldsymbol{c}, \hat{\boldsymbol{c}} \in \mathcal{C}$ with $\hat{\boldsymbol{c}} = \boldsymbol{c} + \Delta_q(\boldsymbol{e}_1 - \boldsymbol{e}_q)$, where Δ_q is such that $\hat{\mathcal{L}}^i = \mathcal{L}^i$ for all $i \in \mathcal{I}$, and assume that $\mathcal{I}_q \ne \emptyset$. Then, either:

1) $\hat{\lambda}^i > \lambda^i$ for all $i \in \mathcal{I}_q$, and $\hat{T}_l > T_l$ for all links l with $q < l \le L^1$;
2) $\hat{\lambda}^i \le \lambda^i$ for all $i \in \mathcal{I}_q$, and $\hat{T}_l \le T_l$, for all links l with $q < l \le L^1$.

In the following lemma we show that if the delay of some link l in $\{2, \cdots, q-1\}$ is higher under configuration $\hat{\boldsymbol{c}}$, then the same is true for all links in $\{l+1, \cdots, q-1\}$.

Lemma 4: Consider two capacity configurations $\boldsymbol{c}, \hat{\boldsymbol{c}} \in \mathcal{C}$ with $\hat{\boldsymbol{c}} = \boldsymbol{c} + \Delta_q(\boldsymbol{e}_1 - \boldsymbol{e}_q)$, where Δ_q is such that $\hat{\mathcal{L}}^i = \mathcal{L}^i$ for all $i \in \mathcal{I}$. For any link $l < q-1$, if $l \in \mathcal{T}^+$, then $n \in \mathcal{T}^+$ for all links $n \in \{l+1, \cdots, q-1\}$.

An immediate consequence of the lemma is that there exists a link $l_0, 1 \le l_0 < q$ such that $\hat{T}_l \le T_l$ for all $l \in \{1, \cdots, l_0\}$, and $\hat{T}_l > T_l$ for all $l \in \{l_0 + 1, \cdots, q\}$.

We are now ready to prove that $\hat{\boldsymbol{c}}$ is user price efficient compared to $\boldsymbol{c}$. The proof is given in the following theorem, which asserts also that the equilibrium delays on all links except link q are lower under configuration $\hat{\boldsymbol{c}}$, i.e., that the set $\mathcal{T}^+$ is empty.

Theorem 2: Consider two capacity configurations $\boldsymbol{c}, \hat{\boldsymbol{c}} \in \mathcal{C}$ with $\hat{\boldsymbol{c}} = \boldsymbol{c} + \Delta_q(\boldsymbol{e}_1 - \boldsymbol{e}_q), 0 < \Delta_q \leq c_q - c_q^0$. Then:

1) configuration $\hat{\boldsymbol{c}}$ is user (thus, overall) price efficient relative to $\boldsymbol{c}$, i.e., $\hat{\lambda}^i \leq \lambda^i$ for all $i \in \mathcal{I}$;
2) $\hat{T}_l \leq T_l$ for all $l \in \mathcal{L} \setminus \{q\}$, and $\hat{T}_q > T_q$.

Proof: $\hat{T}_1 \leq T_1$ and $\hat{T}_q > T_q$ have been established in Lemma 2, thus we only have to prove the remaining statements in the theorem. Let us first establish these claims under the assumption that each user sends its flow over the same set of links under $\boldsymbol{c}$ and $\hat{\boldsymbol{c}}$, i.e., that $\hat{\mathcal{L}}^i = \mathcal{L}^i$ for all $i \in \mathcal{I}$, and then generalize them to the case where $\hat{\mathcal{L}}^i \neq \mathcal{L}^i$ for some user i.

Assume that $\hat{\mathcal{L}}^i = \mathcal{L}^i$ for all $i \in \mathcal{I}$. If no user sends flow to link q, that is, if $L^1 < q$, transferring capacity from link q to link 1 is equivalent to adding capacity to the system of parallel links$\{1, \cdots, L^1\}$,and the result is immediate from Proposition 2. Thus, we only need to consider the case $\mathcal{I}_q \neq \emptyset$, that is, $L^1 \geq q$. Without loss of generality, we assume that user 1 sends flow on all links in the network, i.e., that $L^1 = L$.

Let us first show that $\hat{\lambda}^1 \leq \lambda^1$. Assume by contradiction that $\hat{\lambda}^1 > \lambda^1$. Then, from Lemma 3, we have $\hat{T}_l > T_l$, for all links $l \in \{q+1, \cdots, L\}$. As already explained, Lemma 4 implies that there exists some link $l_0, 1 \leq l_0 < q$, such that $\hat{T}_l \leq T_l$ for all $l \in \{1, \cdots, l_0\}$, and $\hat{T}_l > T_l$ for all $l \in \{l_0+1, \cdots, q\}$. Define $y_l = |(\hat{c}_l - \hat{f}_l) - (c_l - f_l)|, l \in \mathcal{L}$, and note that

$$\sum_{l=1}^{l_0} y_l = \sum_{l=l_0+1}^{L} y_l$$

since

$$\sum_{l \in \mathcal{L}} (\hat{c}_l - \hat{f}_l) = \sum_{l \in \mathcal{L}} (c_l - f_l) = C - R.$$

Recalling that $c_l - f_l \geq c_{l+1} - f_{l+1}$, we have

$$\begin{aligned}
\sum_{l=1}^{L} (\hat{c}_l - \hat{f}_l)^2 &- \sum_{l=1}^{L} (c_l - f_l)^2 \\
&= \sum_{l=1}^{L} y_l^2 + 2\sum_{l=1}^{l_0} (c_l - f_l) y_l \\
&\quad - 2\sum_{l=l_0+1}^{L} (c_l - f_l) y_l \\
&\geq \sum_{l=1}^{L} y_l^2 + 2(c_{l_0} - f_{l_0}) \sum_{l=1}^{l_0} y_l \\
&\quad - 2(c_{l_0+1} - f_{l_0+1}) \sum_{l=l_0+1}^{L} y_l \\
&= \sum_{l=1}^{L} y_l^2 + 2[(c_{l_0} - f_{l_0}) \\
&\quad - (c_{l_0+1} - f_{l_0+1})] \sum_{l=1}^{l_0} y_l > 0. \qquad (24)
\end{aligned}$$

Using the expression for λ^1 given by (39) in Appendix C and (24), we have

$$\hat{\lambda}^1 = \frac{\Sigma_{l \in \mathcal{L}^1} c_l - R^{-1}}{\Sigma_{l \in \mathcal{L}^1} (\hat{c}_l - \hat{f}_l)^2} < \frac{\Sigma_{1 \in \mathcal{L}^1} c_l - R^{-1}}{\Sigma_{1 \in \mathcal{L}^1} (c_l - f_l)^2} = \lambda^1$$

since $\mathcal{L}^1 = \mathcal{L}$. But this contradicts the assumption $\hat{\lambda}^1 > \lambda^1$. Therefore, $\hat{\lambda}^1 \leq \lambda^1$. Lemma 3, then, implies that $\hat{T}_l \leq T_l$ for all $l > q$, and $\hat{\lambda}^i \leq \lambda^i$ for all $i \in \mathcal{I}_q$. Thus

$$\sum_{i \in \mathcal{I}_q} \hat{\lambda}^i \leq \sum_{i \in \mathcal{I}_q} \lambda^i.$$

As explained in the remark following the proof of Lemma 4 in Appendix C, this implies that $\hat{T}_{q-1} \leq T_{q-1}$. Applying Lemma 4 inductively for $l = q-2, q-3, \cdots, 2$, it follows that $\hat{T}_l \leq T_l$, for every link l in $\{2, \cdots, q-1\}$. Therefore, all links in $\mathcal{L}^1 \setminus \{1, q\}$ belong to $\mathcal{T}^-$ and $\mathcal{T}^+ = \emptyset$. This concludes the proof of the second statement in the lemma in the case $\hat{\mathcal{L}}^i = \mathcal{L}^i, i \in \mathcal{I}$.

We now proceed to show that $\hat{\lambda}^i \leq \lambda^i$ for all users $i \in \mathcal{I}$. Assume by contradiction that there exists some user j, such that $\hat{\lambda}^j > \lambda^j$. Then, $j \in \mathcal{I} \setminus \mathcal{I}_q$, since $\hat{\lambda}^i \leq \lambda^i$ for all $i \in \mathcal{I}_q$. Therefore, $\hat{T}_l \leq T_l$ for all $l \in \mathcal{L}^j$. Since $f_l^j T_l' + T_l = \lambda^j < \hat{\lambda}^j = \hat{f}_l^j \hat{T}_l' + \hat{T}_l$, this implies that $\hat{f}_l^j > f_l^j$ for all $l \in \mathcal{L}^j$. Thus

$$r^j = \sum_{l \in \mathcal{L}^j} \hat{f}_l^j > \sum_{l \in \mathcal{L}^j} f_l^j = r^j$$

which is a contradiction. Therefore, for all $i \in \mathcal{I}$, we have $\hat{\lambda}^i \leq \lambda^i$, and this completes the proof of the theorem in the case $\hat{\mathcal{L}}^i = \mathcal{L}^i, i \in \mathcal{I}$.

Let us now consider the case where the transfer of capacity Δ_q from link q to link 1 forces some users to change the set of links over which they send their flow. As in Section III-B, let $A_{l_1, \cdots, l_I}$ denote the set of capacity transfers $\delta \in [0, \Delta_q]$ from link q to link 1, that are such that the set of links over which user $i \in \mathcal{I}$ sends its flow is $\{1, \cdots, l_i\}$. Note that the previous analysis establishes that for any capacity transfers δ_1, δ_2, if $\delta_1, \delta_2 \in A_{l_1, \cdots, l_I}$—for some $(l_1, \cdots, l_I)$—and $\delta_1 < \delta_2$, then: 1) $\lambda^i(\delta_1) \geq \lambda^i(\delta_2)$ for all $i \in \mathcal{I}$; and 2) $T_l(\delta_1) \geq T_l(\delta_2)$ for all $l \in \mathcal{L} \setminus \{q\}$. As explained in Section III-B, this implies that: 1) $\lambda^i = \lambda^i(0) \geq \lambda^i(\Delta_q) = \hat{\lambda}^i, i \in \mathcal{I}$; and 2) $T_l = T_l(0) \geq T_l(\Delta_q) = \hat{T}_l, l \in \mathcal{L} \setminus \{q\}$, by virtue of Theorem 5. ■

B. Cost Efficiency

Let us now proceed to show that $\hat{\boldsymbol{c}}$ is user cost efficient in some special cases of interest. We start by showing the following corollary of Theorem 2.

Corollary 2: Consider two capacity configurations $\boldsymbol{c}, \hat{\boldsymbol{c}} \in \mathcal{C}$ with $\hat{\boldsymbol{c}} = \boldsymbol{c} + \Delta_q(\boldsymbol{e}_1 - \boldsymbol{e}_q)$, where Δ_q is such that $\hat{\mathcal{L}}^i = \mathcal{L}^i$ for all $i \in \mathcal{I}$. Then, the equilibrium costs of all users $i \in \mathcal{I}_q$ are lower under $\hat{\boldsymbol{c}}$, i.e., $\hat{J}^i \leq J^i$ for all $i \in \mathcal{I}_q$.

Proof: From Theorem 2, we have $\hat{c}_l - \hat{f}_l \geq c_l - f_l$, for all links $l > q$. In view of

$$\sum_{l \in \mathcal{L}} (\hat{c}_l - \hat{f}_l) = \sum_{l \in \mathcal{L}} (c_l - f_l) = C - R$$

this implies that

$$\sum_{l=1}^{m}(\hat{c}_l - \hat{f}_l) \le \sum_{l=1}^{m}(c_l - f_l), \qquad m = q, \cdots, L. \tag{25}$$

Using the expression for J^i given in (23), we get

$$\begin{aligned} J^i - \hat{J}^i &= \lambda^i \sum_{l=1}^{L^i}(c_l - f_l) - \hat{\lambda}^i \sum_{l=1}^{L^i}(\hat{c}_l - \hat{f}_l) \\ &\ge \hat{\lambda}^i \sum_{l=1}^{L^i}[(c_l - f_l) - (\hat{c}_l - \hat{f}_l)] \ge 0, \quad i \in \mathcal{I}_q \end{aligned}$$

where the first inequality follows from $\hat{\lambda}^i \le \lambda^i$ (Theorem 2) and the second from (25), since $L^i \ge q$ for all $i \in \mathcal{I}_q$. ■

In the following proposition we establish that $\hat{\boldsymbol{c}}$ is user cost efficient in the case of users that are consistent under both $\boldsymbol{c}$ and $\hat{\boldsymbol{c}}$, which includes the cases of identical and simple users.

Proposition 6: Consider a system of parallel links shared by I users, consistent at the capacity configurations $\boldsymbol{c}$ and $\hat{\boldsymbol{c}}$, where $\hat{\boldsymbol{c}} = \boldsymbol{c} + \Delta_q(\boldsymbol{e}_1 - \boldsymbol{e}_q)$. Then $\hat{\boldsymbol{c}}$ is user (thus, overall) cost efficient compared to $\boldsymbol{c}$.

Proof: For consistent users we have $\mathcal{L}^i = \mathcal{L}^j$, for all $i, j \in \mathcal{I}$. As in the proof of Theorem 2, it suffices to establish cost efficiency in the case where $\hat{\mathcal{L}}^i = \mathcal{L}^i, i \in \mathcal{I}$. If $q \notin \mathcal{L}^i$, transferring capacity from link q to link 1 is equivalent to adding capacity to the system, and the result is immediate from Proposition 4. If, on the other hand, $q \in \mathcal{L}^i$, the result follows from Corollary 2. ■

The following proposition asserts that $\hat{\boldsymbol{c}}$ is user cost efficient in the special case of two users.

Proposition 7: Consider a system of parallel links shared by two users. The capacity configuration $\hat{\boldsymbol{c}} = \boldsymbol{c} + \Delta_q(\boldsymbol{e}_1 - \boldsymbol{e}_q)$ is user (thus, overall) cost efficient compared to $\boldsymbol{c}$.

Proof: It suffices to establish cost efficiency under the assumption that $\hat{\mathcal{L}}^i = \mathcal{L}^i, i = 1, 2$. The result is immediate from Proposition 5, if no user sends flow on link q. If both users send flow on link q, the result is also immediate from Corollary 2. Hence, we only have to consider the case $L^2 < q \le L^1$. Then, from Corollary 2, we have $\hat{J}^1 \le J^1$. Following the same proof as in Lemma 7 in Appendix B, one can show that there must be a user j such that $\hat{c}_l^j \ge c_l^j$ for all links $l \in \mathcal{L} \setminus \{q\}$. If $j = 2$, then the residual capacity seen by user 2 at all links in $\mathcal{L}^2$ is higher under $\hat{\boldsymbol{c}}$; therefore, $\hat{J}^2 \le J^2$. Suppose now that $j = 1$. For any link $l \in \mathcal{L}^2 \setminus \{1\}$, inequality $\hat{c}_l^1 \ge c_l^1$ implies that $\hat{f}_l^2 \le f_l^2$, since $\hat{c}_l = c_l$. Therefore, user 2 decreases its flow on all links in $\mathcal{L}^2 \setminus \{1\}$. Since $\hat{T}_l \le T_l$ for all links $l \in \mathcal{L}^2$ and the delay at link 1 is always minimal among all links, we conclude that $\hat{J}^2 \le J^2$.

VI. Optimal Capacity Allocation

We now proceed to investigate the optimal capacity allocation problem for a system of parallel links, according to the various performance measures defined in Section II-B. The main results of this section, namely Theorems 3 and 4, assert that the capacity configuration $\boldsymbol{c}^* = \boldsymbol{c}^0 + \Delta\boldsymbol{e}_1$, that results from allocating the entire additional capacity to the link with the initially highest capacity is 1) user price optimal in $\mathcal{C}$ and 2) user cost optimal in $\mathcal{C}$ if the lower bounds on the link capacities are equal for all links. Furthermore, $\boldsymbol{c}^*$ will be shown to be cost optimal for a number of special cases of interest, even if the lower bounds on the link capacities are not identical.

A. Price Optimality

The following theorem establishes price optimality of capacity configuration $\boldsymbol{c}^*$.

Theorem 3: Consider a system of parallel links with initial capacity configuration $\boldsymbol{c}^0$, shared by I noncooperative users, and an additional capacity allowance Δ. The capacity configuration $\boldsymbol{c}^* = \boldsymbol{c}^0 + \Delta\boldsymbol{e}_1$, that results from allocating the entire additional capacity to the link with the initially highest capacity, is user (thus, overall) price optimal in $\mathcal{C}$.

Proof: Let $\mathcal{D}$ denote the set of capacity configurations that can be implemented by the designer by allocating an additional capacity of exactly $\delta, 0 \le \delta \le \Delta$, to a system of parallel links with initial configuration $\boldsymbol{c}^0$. For every δ, define $\boldsymbol{c}^*(\delta) = \boldsymbol{c}^0 + \delta\boldsymbol{e}_1$. Theorem 2 implies that $\boldsymbol{c}^*(\delta)$ is user price optimal in $\mathcal{D}$. To see this, consider any $\boldsymbol{c} \in \mathcal{D}$. From Theorem 2, the capacity configuration $\boldsymbol{c} + (c_L - c_L^0)(\boldsymbol{e}_1 - \boldsymbol{e}_L)$, that is obtained by reducing the capacity of link L to its lower bound and adding the excess capacity $(c_L - c_L^0)$ to link 1, is user price efficient compared to $\boldsymbol{c}$. Proceeding inductively, for every $m > 1$, the configuration $\boldsymbol{c} + \Sigma_{l=m}^{L}\,(c_l - c_l^0)(\boldsymbol{e}_1 - \boldsymbol{e}_l)$ is user price efficient compared to $\boldsymbol{c} + \Sigma_{l=m+1}^{L}\,(c_l - c_l^0)(\boldsymbol{e}_1 - \boldsymbol{e}_l)$. Hence, $\boldsymbol{c}^0 + \delta\boldsymbol{e}_1 = \boldsymbol{c} + \Sigma_{l=2}^{L}\,(c_l - c_l^0)(\boldsymbol{e}_1 - \boldsymbol{e}_l)$ is user price efficient with respect to $\boldsymbol{c}$, that is, $\boldsymbol{c}^*(\delta)$ is user price optimal in $\mathcal{D}$. From Proposition 2, $\boldsymbol{c}^* = \boldsymbol{c}^*(\Delta)$ is user price efficient with respect to any $\boldsymbol{c}^*(\delta)$ with $0 \le \delta < \Delta$. Therefore, $\boldsymbol{c}^*$ is user price optimal in $\mathcal{C}$. ■

B. Cost Optimality

The following theorem shows that configuration $\boldsymbol{c}^*$ is user cost optimal if the lower bounds on the link capacities are equal for all links.

Theorem 4: Consider a system of parallel links with initial capacity configuration $\boldsymbol{c}^0$, shared by I users, and an additional capacity allowance Δ. If $c_l^0 = c_m^0$, for all $l, m \in \mathcal{L}$, then the capacity configuration $\boldsymbol{c}^* = \boldsymbol{c}^0 + \Delta\boldsymbol{e}_1$, that results from allocating the entire additional capacity to link 1, is user (thus, overall) cost optimal in $\mathcal{C}$.

Proof: As in the proof of Theorem 3, let $\mathcal{D}$ be the set of capacity configurations that can be implemented by the designer by allocating an additional capacity of exactly δ to a system of parallel links with initial configuration $\boldsymbol{c}^0$. Let us first show that $\boldsymbol{c}^*(\delta) = \boldsymbol{c}^0 + \delta\boldsymbol{e}_1$ is user cost efficient relative to any $\boldsymbol{c} \in \mathcal{D}$.

Starting from $\boldsymbol{c}$, we construct a process of capacity transfers from links $l \in \mathcal{L} \setminus \{1\}$ to link 1 such that at each step the resulting capacity configuration is user cost efficient relative to the previous one, and the configuration at the final step coincides with $\boldsymbol{c}^*(\delta)$. More specifically, at the first step we reduce the capacity of link 2 to c_3 and transfer the excess capacity $c_2 - c_3$ to link 1. At the second step, the capacity of both links 2 and 3 is reduced to c_4 and the excess capacity $2(c_3 - c_4)$ is added to link 1. Proceeding this way, at step $L-1$, the capacity of all links in $\mathcal{L}\setminus\{1\}$ is equal to c_L. At the final step, the capacity of all these links is reduced from c_L to c_L^0 and the excess capacity is transferred to link 1. The final capacity configuration coincides with $\boldsymbol{c}^*(\delta)$.

In order to prove that at each step of this process the resulting capacity configuration is user cost efficient relative to the previous one, it suffices to show that for any capacity configuration $\boldsymbol{c}\in\mathcal{D}$, with $c_2=c_3=\cdots=c_n$ for some $n\in\mathcal{L}$, the configuration $\hat{\boldsymbol{c}}=\boldsymbol{c}+\Delta_n\Sigma_{l=2}^n\ (\boldsymbol{e}_1-\boldsymbol{e}_l)$ that results from *simultaneously* reducing the capacity of each link in $\{2,\cdots,n\}$ by $\Delta_n\le c_n-c_{n+1}$ and transferring the excess capacity $(n-1)\Delta_n$ to link 1, is user cost efficient relative to $\boldsymbol{c}$. As explained in Section III-B, we only need to establish this result under the assumption that Δ_n is such that $\hat{\mathcal{L}}^i=\mathcal{L}^i$ for all $i\in\mathcal{I}$; by virtue of Theorem 5, the result extends for any $\Delta_n\in[0,c_n-c_{n+1}]$.

Consider two configurations $\boldsymbol{c}$ and $\hat{\boldsymbol{c}}$ in $\mathcal{D}$, as described above, such that $\hat{\mathcal{L}}^i=\mathcal{L}^i$ for all $i\in\mathcal{I}$. Then, one can show that $\hat{T}_1\le T_1$, while $\hat{T}_l>T_l$ for all $l\in\{2,\cdots,n\}$, following precisely the proof of Lemma 2 with $\{2,\cdots,n\}$ playing the role of q in that proof. Furthermore, note that the proof of Lemma 3 is based on the structure of $\boldsymbol{c}$ and $\hat{\boldsymbol{c}}$—as defined in that lemma—only to the extent that there exists some link q such that $\Sigma_{l=1}^q\ \hat{c}_l=\Sigma_{l=1}^q\ c_l$ and $\hat{c}_l=c_l$ for all $l>q$. Therefore, the results in that lemma—with q replaced by n—readily apply to configurations $\boldsymbol{c}$ and $\hat{\boldsymbol{c}}$ of the form considered here. Based on these results, one can adopt the proof of Theorem 2 to show that: 1) $\hat{\boldsymbol{c}}$ is user price efficient relative to $\boldsymbol{c}$; and 2) $\hat{T}_l\le T_l$ for all links $l\in\mathcal{L}\setminus\{2,\cdots,n\}$. Then, following precisely the proof of Corollary 2, one can show that $\hat{J}^i\le J^i$, for all users $i\in\mathcal{I}_n$. Since $c_2=\cdots=c_n$, we have $\mathcal{I}_2=\cdots=\mathcal{I}_n$, i.e., the users that do not send flow on link n send their entire flow on link 1. For any such user i we have $\hat{J}^i=r^i\hat{T}_1\le r^iT_1=J^i$. Therefore, $\hat{\boldsymbol{c}}$ is user cost efficient with respect to $\boldsymbol{c}$.

As already explained, user cost efficiency of $\hat{\boldsymbol{c}}$ relative to $\boldsymbol{c}$ implies that $\boldsymbol{c}^*(\delta)$ is user cost optimal in $\mathcal{D}$. Thus, in order to show user cost optimality of $\boldsymbol{c}^*=\boldsymbol{c}^*(\Delta)$ in $\mathcal{C}$, it suffices to show that for any $\delta_1,\delta_2\in[0,\Delta]$, if $\delta_1<\delta_2$, then $\boldsymbol{c}^*(\delta_2)$ is user cost efficient relative to$\boldsymbol{c}^*(\delta_1)$. By virtue of Theorem 5, we only need to establish this result under the assumption that each user sends its flow over the same links under both configurations. Note that in both configurations the capacity of all links in $\mathcal{L}\setminus\{1\}$ is equal to their common lower bound, while the capacity of link 1 is higher in $\boldsymbol{c}^*(\delta_2)$. Since, by Proposition 2, the delay of each link is lower under $\boldsymbol{c}^*(\delta_2)$, the cost of each user that sends flow only to link 1 is lower under this configuration, i.e., $J^i(\delta_2)\le J^i(\delta_1)$, for all $i\in\mathcal{I}_1\setminus\mathcal{I}_2$. Since $\mathcal{I}=\mathcal{I}_1\cup\mathcal{I}_2$, it remains to be shown that the same inequality holds for all $i\in\mathcal{I}_2$.

Note that the users that send flow only to link 1 simply occupy a fixed capacity of that link at both $\boldsymbol{c}^*(\delta_1)$ and $\boldsymbol{c}^*(\delta_2)$. Therefore, in order to compare the performance of the rest of the users, we can consider two configurations, $\tilde{c}^*(\delta_1)$ and $\tilde{c}^*(\delta_2)$, derived from $\boldsymbol{c}^*(\delta_1)$ and $\boldsymbol{c}^*(\delta_2)$, respectively, by reducing the capacity of link 1 by $\Sigma_{i\in\mathcal{I}_1\setminus\mathcal{I}_2}\ r^i$, and neglect the users in $i\in\mathcal{I}_1\setminus\mathcal{I}_2$. All users in $\mathcal{I}_2$ are consistent under both $\tilde{c}^*(\delta_1)$ and $\tilde{c}^*(\delta_2)$, since they send flow on all links. Proposition 4, then, implies that $J^i(\delta_2)\le J^i(\delta_1)$, for all $i\in\mathcal{I}_2$, and this concludes the proof. ■

Remark: The results in Theorems 3 and 4 rely on the assumption that the network designer adds capacity to an existing system of parallel links, i.e., that the initial capacity of every link c_l^0 is nonzero, as has been noted in Section II-B. In [28] we show that all the optimality results of this section apply also to the case where $c_l^0=0$, for some links $l\in\mathcal{L}$, that is, when the designer is allowed to add a finite number of links.

The following proposition indicates that the user price optimal capacity configuration $\boldsymbol{c}^*$ is also user cost optimal in some special cases of interest, even if the lower bounds on the link capacities are not identical.

Proposition 8: Consider a system of parallel links with initial capacity configuration $\boldsymbol{c}^0$, shared by I users, and an additional capacity allowance Δ. The capacity configuration $\boldsymbol{c}^*=\boldsymbol{c}^0+\Delta\boldsymbol{e}_1$, that results from allocating the entire additional capacity to the link with the initially highest capacity, is user (thus, overall) cost optimal in $\mathcal{C}$ if 1) users are consistent at all capacity configurations in $\mathcal{C}$; or 2) there are two users ($I=2$).

Proof: Part 1) follows from Propositions 4 and 6. Part 2) follows from Propositions 5 and 7. ■

VII. Conclusions

We investigated the optimal capacity allocation problem in a network where users noncooperatively implement their optimal routing strategies. The problem was formulated as allocating an additional capacity allowance to an existing network. This formulation is equivalent to a standard capacity allocation problem, for which a lower bound is specified on the capacity of each link, e.g., due to reliability considerations.

For a system of parallel links we established the efficiency of two elementary capacity provisioning operations: capacity *addition* to any network link, and capacity *transfer* to the link with the originally largest capacity. Given these results, we showed that the capacity allocation problem has a simple and intuitive solution: the optimal allocation assigns the additional capacity exclusively to the link with the initially highest capacity. This solution coincides with the optimal capacity allocation when routing is centrally controlled.

In this study we concentrated on cost functions that are based on M/M/1 delays. As previously mentioned, these link delay functions, $T_l(f_l)$, should be interpreted as a general congestion cost per unit of flow, that encapsulates the dependence of the quality of service provide by a finite capacity resource on the total load f_l offered to it. Functions of such form have been used to express this dependence in various practical routing schemes [25], [12]. Their suitability as "generic" cost functions can be observed also in the noncooperative framework, where the routing equilibrium, corresponding to the M/M/1 cost function, exhibits properties that one would expect in practice. We also note that our results readily apply to other classes of cost functions, such as queuing delays of M/D/1 systems. Furthermore, some of our results, e.g., efficiency of capacity addition to a system of parallel links, apply to a broader class of cost functions and do not depend on the specific structure of M/M/1 delays.

An extension of this study for general network topologies appears in [28].

APPENDIX A
PROOFS OF RESULTS IN SECTION III

Proof of Proposition 1: Let us concentrate on the equilibrium strategy $\boldsymbol{f}^i$ of user i. As explained earlier, Lemma 1 implies that there exists some $L^i \leq L$, such that

$$f_l^i > 0 \quad \text{if} \quad l \leq L^i, \quad f_l^i = 0 \quad \text{if} \quad l > L^i \tag{26}$$

i.e., $\mathcal{L}^i = \{1, \cdots, L^i\}$. Equation (21) is immediate by summing $\sqrt{\lambda^i}(c_l^i - f_l^i) = \sqrt{c_l^i}$ [see (15)] over all links $l \in A$. Taking $A = \mathcal{L}^i$ and recalling that $\Sigma_{l \in \mathcal{L}^i} f_l^i = r^i$, (22) follows.

Equation (22), together with (16) for $l = L^i + 1$, implies that

$$\frac{\Sigma_{l=1}^{L^i}\sqrt{c_l^i}}{\Sigma_{l=1}^{L^i} c_l^i - r^i} = \sqrt{\lambda^i} \leq \frac{1}{\sqrt{c_{L^i+1}^i}} \Rightarrow r^i \leq G_{L^i+1}^i. \tag{27}$$

Similarly, taking $A = \{1, \cdots, L^i - 1\}$ in (21) and $l = L^i$ in (15), we have

$$\frac{\Sigma_{l=1}^{L^i-1}\sqrt{c_l^i}}{\Sigma_{l=1}^{L^i-1}(c_l^i - f_l^i)} = \sqrt{\lambda^i} = \frac{\sqrt{c_{L^i}^i}}{c_{L^i}^i - f_{L^i}^i} > \frac{1}{\sqrt{c_{L^i}^i}}$$

which implies $G_{L^i}^i < \Sigma_{l=1}^{L^i-1} f_l^i < r^i$. Hence, (20) is a necessary condition for (26). Using (18), it is easy to see that it is also sufficient. From (15) and (22), we get $f_l^i = c_l^i - \sqrt{c_l^i/\lambda^i}$, and (19) follows, using the expression for λ^i given in (27).

From (15), we note that

$$\frac{f_l^i}{c_l - f_l} = \lambda^i(c_l - f_l) - 1, \qquad l = 1, \cdots, L^i. \tag{28}$$

Therefore, by summing over all links $l \in \mathcal{L}^i$, we obtain

$$J^i = \sum_{l=1}^{L^i} \frac{f_l^i}{c_l^i - f_l^i} = \lambda^i \sum_{l=1}^{L^i}(c_l - f_l) - L^i$$

and (23) follows using (27). ∎

We proceed to present the proof of Theorem 1, which asserts that the Nash mapping $\mathcal{N}\colon \mathcal{C} \to F \otimes \mathbb{R}^{I(L+1)}$, as defined in Section III-B, is continuous. The proof is based on the following lemma from [6].

Lemma 5: Let X and Y be subsets of two (finite dimensional) Euclidean spaces, where Y is a compact set. A function $h\colon X \to Y$ is continuous if and only if its graph $\mathcal{G}h = \{(x, y) \in X \otimes Y \colon y = h(x)\}$ is a closed subset of $X \otimes Y$, i.e., if for any convergent sequence $\{(x(n), y(n)), n \geq 0\}$ in $\mathcal{G}h$, we have $\lim_{n\to\infty}(x(n), y(n)) = (x, y) \in \mathcal{G}h$.

The following lemma shows that the Nash mapping takes values in a compact subset of $F \otimes \mathbb{R}^{I(L+1)}$.

Lemma 6: There exists a compact subset $\Theta \subseteq F \otimes \mathbb{R}^{I(L+1)}$, such that $\mathcal{N}(\boldsymbol{c}) \in \Theta$, for all $\boldsymbol{c} \in \mathcal{C}$.

Proof: Let $\boldsymbol{c} \in \mathcal{C}$ and $\mathcal{N}(\boldsymbol{c}) = (\boldsymbol{f}, \boldsymbol{\lambda}, \boldsymbol{\mu})$. Then, $0 \leq f_l^i \leq r^i$, for all $l \in \mathcal{L}$ and $i \in \mathcal{I}$. Turning our attention to λ^i, note that (15) and (16) imply that for all $l \in \mathcal{L}$, we have

$$\sqrt{\lambda^i}(c_l - f_l) \leq \sqrt{c_l - f_l^{-i}} \leq \sqrt{c_l}.$$

Summing over all $l \in \mathcal{L}$, we get

$$0 \leq \lambda^i \leq \left[\frac{\sum_{l \in \mathcal{L}} \sqrt{c_l}}{\left(\sum_{l \in \mathcal{L}} c_l - R\right)}\right]^2$$

for all $i \in \mathcal{I}$. For every $\boldsymbol{c} \in \mathcal{C}$, we have

$$R < C^0 \leq \sum_{l \in \mathcal{L}} c_l \leq C^0 + \Delta$$

and since

$$\sum_{l \in \mathcal{L}} \sqrt{c_l} \leq L\sqrt{C^0 + \Delta}$$

we get

$$0 \leq \lambda^i \leq \left[\frac{L\sqrt{C^0 + \Delta}}{C^0 - R}\right]^2 \equiv \overline{\lambda} < \infty, \qquad i \in \mathcal{I}. \tag{29}$$

Let us now consider μ_l^i. By (7), $\mu_l^i \geq 0$, for all $l \in \mathcal{L}$ and $i \in \mathcal{I}$. If $f_l^i > 0$, then $\mu_l^i = 0$, by (6). Hence, we need only consider the case $f_l^i = 0$. Then $f_l = f_l^{-i}$, and (4) gives $\mu_l^i = (c_l - f_l)^{-1} - \lambda^i \leq (c_l - f_l)^{-1}$. If $f_l = 0$, then $\mu_l^i \leq 1/c_l \leq 1/c_L^0 < \infty$. If, on the other hand, $f_l > 0$, there exists some user $j \in \mathcal{I}$ with $f_l^j > 0$, and (15) together with $f_l^{-j} = f_l + f_l^j$ gives $(c_l - f_l)^{-1} = \lambda^j - f_l^j(c_l - f_l)^{-2} \leq \lambda^j \leq \overline{\lambda}$. Therefore, defining $\overline{\mu} = \max\{\overline{\lambda}, 1/c_L^0\} < \infty$, we have $0 \leq \mu_l^i \leq \overline{\mu}$. Hence, taking

$$\Theta = \bigotimes_{i \in \mathcal{I}} [0, r^i]^L \otimes ([0, \overline{\lambda}] \otimes [0, \overline{\mu}]^L)^I$$

we have $\mathcal{N}(\boldsymbol{c}) \in \Theta$, and the result follows. ∎

Proof of Theorem 1: From Lemma 5, it suffices to show that $\mathcal{N}$ has a closed graph. To this end, let us consider convergent sequences $\boldsymbol{c}(n) \to \boldsymbol{c}$ in $\mathcal{C}$ and $(\boldsymbol{f}(n), \boldsymbol{\lambda}(n), \boldsymbol{\mu}(n)) \to (\boldsymbol{f}, \boldsymbol{\lambda}, \boldsymbol{\mu})$ in Θ, such that $\mathcal{N}(\boldsymbol{c}(n)) = (\boldsymbol{f}(n), \boldsymbol{\lambda}(n), \boldsymbol{\mu}(n)), n \geq 0$. We, then, have to show that $\mathcal{N}(\boldsymbol{c}) = (\boldsymbol{f}, \boldsymbol{\lambda}, \boldsymbol{\mu})$.

By virtue of the optimality conditions (4)–(7), for all $i \in \mathcal{I}$ and all $n \geq 0$, we have

$$\frac{c_l(n) - f_l^{-i}(n)}{(c_l(n) - f_l(n))^2} - \lambda^i(n) - \mu_l^i(n) = 0, \qquad l \in \mathcal{L} \tag{30}$$

$$\sum_{l \in \mathcal{L}} f_l^i(n) = r^i \tag{31}$$

$$\mu_l^i(n) f_l^i(n) = 0, \qquad l \in \mathcal{L} \tag{32}$$

$$\mu_l^i(n) \geq 0, \quad f_l^i(n) \geq 0, \qquad l \in \mathcal{L}. \tag{33}$$

Taking the limit as $n \to \infty$ in (30)–(33), for all $i \in \mathcal{I}$, we get precisely the necessary and sufficient conditions for $\mathcal{N}(\boldsymbol{c}) = (\boldsymbol{f}, \boldsymbol{\lambda}, \boldsymbol{\mu})$. Hence, $\mathcal{G}\mathcal{N}$ is a closed set, and continuity of $\mathcal{N}$ follows from Lemma 5.

In taking the limit in (30), we have assumed that $\lim_n(c_l(n) - f_l(n)) = c_l - f_l > 0$ for all $l \in \mathcal{L}$. Let us now justify this assumption. Suppose by contradiction that there exists some link $l \in \mathcal{L}$ such that $f_l = c_l > 0$. Then, there exists some user $j \in \mathcal{I}$, such that $f_l^j > 0$ and some $n_0 \in \mathbb{N}$, such that $f_l^j(n) > 0$, for all $n \geq n_0$. Then

$$\frac{\lambda^j(n) = (c_l(n) - f_l(n) + f_l^j(n))}{(c_l(n) - f_l(n))^2}$$

for all $n \geq n_0$, and taking the limit as $n \to \infty, \lambda^j = \lim_n \lambda^j(n) = \infty$. But this is a contradiction to $\lambda^i \leq \overline{\lambda} < \infty$, for all $i \in \mathcal{I}$. Thus, $\lim_n (c_l(n) - f_l(n)) = c_l - f_l > 0$. ■

In the rest of this Appendix, we prove a monotonicity result for real functions, which is used in Section III-B to establish that the capacity allocation problem can be investigated based on comparisons of capacity configurations that are such that each user sends its flow over the same set of links under both configurations. More specifically, we will show the following.

Theorem 5: Let h: $X \to \mathbb{R}$, where $X \subset \mathbb{R}$ is a closed interval. Consider a family $\mathcal{A} = \{A_1, \cdots, A_n\}$ of closed subsets of X, such that 1) $\cup_{i=1}^n A_i = X$; and 2) for every $A_i \in \mathcal{A}$, we have: $x, y \in A_i$ and $x < y \Rightarrow h(x) \leq h(y)$. Then h is nondecreasing in X.

Proof: Assume, by contradiction, that there exist $x, y \in X$, such that $x < y$ and $h(x) > h(y)$. Then, there exist $A_i \neq A_j$ in $\mathcal{A}$, such that $x \in A_i$ and $y \in A_j$. Let

$$z_1 = \min\{t \in A_j : t \geq x\}. \tag{34}$$

By definition, $z_1, y \in A_j$, and $z_1 \leq y$, thus $h(z_1) \leq h(y)$. If $z_1 = x$, this contradicts the assumption $h(x) > h(y)$. Therefore $z_1 > x$, and (34) implies that z_1 is the minimum point in the boundary $A_j \cap A_j^c$ of A_j that is greater than x. Since $z_1 \in A_j^c \cap X$, there exists $j_1 \in \{1, \cdots, n\} \setminus \{j\}$, such that $z_1 \in A_{j_1}$. Note that $j_1 \neq i$, since $j_1 = i$ would imply $x, z_1 \in A_i$, and thus $h(x) \leq h(z) \leq h(y)$.

Now define $z_2 = \min\{t \in A_{j_1} : t \geq x\}$, for which we have $z_2 \leq z_1$ and $h(z_2) \leq h(z_1) \leq h(y)$. Similarly to z_1, z_2 is the minimum point in the boundary of A_{j_1} that is greater than x, and there exists $j_2 \in \{1, \cdots, n\} \setminus \{j_1\}$, such that $z_2 \in A_{j_2}$. As in the case of z_1, it must be $j_2 \neq i$. If $z_2 < z_1$, then $j_2 \neq j$, since $j_2 = j$ would contradict the definition of z_1 in (34). If, on the other hand, $z_2 = z_1$, then j_2 can be chosen, such that $j_2 \neq j$. Indeed, if the claim is not true, then for every $k \neq j, j_1$, we have $z_2 \notin A_k$, and there is an $\varepsilon > 0$, such that $[z_2 - \varepsilon, z_2] \cap A_k = \emptyset$. Furthermore, by the definition of z_2, there are no points of A_j or A_{j_1} in $[z_2 - \varepsilon, z_2)$. Thus, $[z_2 - \varepsilon, z_2) \not\subset X$, which implies that z_2 is the left endpoint of interval X. But this is a contradiction, since, as already explained, $z_2 > x$.

Proceeding this way, we can construct a sequence $z_1, z_2, \cdots, z_{n-1}$, such that for all k

$$\begin{aligned} z_k &= \min\{t \in A_{j_{k-1}} : t \geq x\} \in A_{j_k} \\ j_k &\in \{1, \cdots, n\} \setminus \{j, j_1, \cdots, j_{k-1}\}. \end{aligned} \tag{35}$$

Since $z_k, z_{k-1} \in A_{j_{k-1}}$ and $z_k \leq z_{k-1}$, we have $h(z_k) \leq h(z_{k-1})$. Therefore, $h(z_k) \leq h(y)$, for all $k \in \{1, \cdots, n-1\}$. Furthermore, for all $k \in \{1, \cdots, n-1\}$, it must be $j_k \neq i$, since $j_k = i$ would give $h(x) \leq h(z_k) \leq h(y)$. From (35) we have $j_{n-1} \in \{1, \cdots, n\} \setminus \{j, j_1, \cdots, j_{n-2}\}$, and since $i \notin \{j, j_1, \cdots, j_{n-2}\}$, this implies that $j_{n-1} = i$, which is a contradiction. Therefore, it must be $h(x) \leq h(y)$, and the result follows. ■

Appendix B
Proofs of Results in Section IV

Let us start by proving a technical result that will be used in the proofs presented in this Appendix. The result applies to capacity configurations $\boldsymbol{c}$ and $\hat{\boldsymbol{c}}$, such that each user sends its flow over the same links under both configurations, i.e., $\hat{\mathcal{L}}^i = \mathcal{L}^i$ for all $i \in \mathcal{I}$ and $\hat{\mathcal{I}}_l = \mathcal{I}_l$ for all $l \in \mathcal{L}$.

Lemma 7: Consider two capacity configurations $\boldsymbol{c}$ and $\hat{\boldsymbol{c}}$ such that $\hat{\boldsymbol{c}}$ is an augmentation of $\boldsymbol{c}$, and $\hat{\mathcal{L}}^i = \mathcal{L}^i, i \in \mathcal{I}$. Let j be a user such that $\hat{\lambda}^j/\lambda^j \geq \hat{\lambda}^i/\lambda^i$ for all $i \in \mathcal{I}$. Then, the residual capacity $\hat{c}_l^j \geq c_l^j$ for all $l \in \mathcal{L}$, and $\hat{J}^j \leq J^j$.

Proof: Assume that $\hat{c}_m^j < c_m^j$ for some link $m \in \mathcal{L}^j$. Using (15), we have

$$\begin{aligned} \frac{\hat{c}_m^i}{c_m^i} &= \frac{\hat{\lambda}^i(\hat{c}_m - \hat{f}_m)^2}{\lambda^i(c_m - f_m)^2} \leq \frac{\hat{\lambda}^j(\hat{c}_m - \hat{f}_m)^2}{\lambda^j(c_m - f_m)^2} \\ &= \frac{\hat{c}_m^j}{c_m^j} < 1, \qquad i \in \mathcal{I}_m. \end{aligned} \tag{36}$$

Note that

$$\begin{aligned} &\sum_{i \in \mathcal{I}_m} (\hat{c}_m^i - c_m^i) \\ &\quad = \sum_{i \in \mathcal{I}_m} [(\hat{c}_m - \hat{f}_m + \hat{f}_m^i) - (c_m - f_m + f_m^i)] \\ &\quad = (I_m - 1)[(\hat{c}_m - \hat{f}_m) - (c_m - f_m)] \\ &\qquad + (\hat{c}_m - c_m) \end{aligned}$$

and since (by assumption) $\hat{c}_m \geq c_m$ and (by Proposition 2) $\hat{c}_m - \hat{f}_m \geq c_m - f_m$, we have

$$\sum_{i \in \mathcal{I}_m} \hat{c}_m^i \geq \sum_{i \in \mathcal{I}_m} c_m^i$$

that is, (36) is a contradiction. Therefore, $\hat{c}_l^j \geq c_l^j$ for all $l \in \mathcal{L}^j$. By virtue of Proposition 2, the same is true for all $l \in \mathcal{L} \setminus \mathcal{L}^j$, since $\hat{f}_l^j = f_l^j = 0$ for all such links. Hence, the residual capacity seen by user j at every link is higher under configuration $\hat{\boldsymbol{c}}$ and, therefore, $\hat{J}^j \leq J^j$. ■

Proof of Proposition 4: First we prove the proposition under the assumption that $\boldsymbol{c}$ and $\hat{\boldsymbol{c}}$ are such that $\hat{\mathcal{L}}^i = \mathcal{L}^i$ for all users $i \in \mathcal{I}$ and then generalize the result for any augmentation $\hat{\boldsymbol{c}}$ of $\boldsymbol{c}$. Let user j be as in Lemma 7 (note that such a user always exists). Since the users are consistent at $\boldsymbol{c}$ and $\hat{\boldsymbol{c}}$, using (23), $\hat{\lambda}^j/\lambda^j \geq \hat{\lambda}^i/\lambda^i$ gives

$$\frac{\hat{J}^j + L^i}{J^j + L^i} \geq \frac{\hat{J}^i + L^i}{J^i + L^i}, \quad i \in \mathcal{I}.$$

Since $\hat{J}^j \leq J^j$ (Lemma 7), this implies that $\hat{J}^i \leq J^i$. Hence the proposition holds in the case where $\hat{\mathcal{L}}^i = \mathcal{L}^i$, for all $i \in \mathcal{I}$.

Let us now generalize the result for any augmentation $\hat{\boldsymbol{c}}$ of $\boldsymbol{c}$. It suffices to establish this result for augmentations of the form $\hat{\boldsymbol{c}} = \boldsymbol{c} + \Delta_l \boldsymbol{e}_l$, for any link l. As in Section III-B, let $A_{l_1, \cdots, l_I}$ denote the set of capacity additions $\delta \in [0, \Delta_l]$ to link l such that the set of links over which user $i \in \mathcal{I}$ sends its flow is $\{1, \cdots, l_i\}$. Note that the previous analysis shows that for any δ_1, δ_2, if $\delta_1, \delta_2 \in A_{l_1,\dots,l_I}$—for some $(l_1, \cdots, l_I)$—and $\delta_1 < \delta_2$, then $J^i(\delta_1) \geq J^i(\delta_2)$ for all $i \in \mathcal{I}$. As explained in Section III-B, this implies that $J^i = J^i(0) \geq J^i(\Delta_l) = \hat{J}^i, i \in \mathcal{I}$, by virtue of Theorem 5.

Proof of Proposition 5: As in the proof of Proposition 4, it suffices to establish the result under the assumption that each user sends its flow over the same links under both configurations, i.e., $\hat{\mathcal{L}}^i = \mathcal{L}^i$ for all $i \in \mathcal{I}$.

Let user j be as in Lemma 7. Suppose first that $j = 2$. Then $\hat{J}^2 \leq J^2$. Summing (28) over all $l \in \mathcal{L}^2$, it is easy to see that $\hat{\lambda}^2/\lambda^2 \geq \hat{\lambda}^1/\lambda^1$ gives

$$\frac{\Sigma_{l\in\mathcal{L}^2} \hat{J}_l^1 + L^2}{\Sigma_{l\in\mathcal{L}^2} J_l^1 + L^2} \leq \frac{\Sigma_{l\in\mathcal{L}^2} \hat{J}_l^2 + L^2}{\Sigma_{l\in\mathcal{L}^2} J_l^2 + L^2}$$

and since $\hat{J}^2 \leq J^2$, we have

$$\sum_{l\in\mathcal{L}^2} \hat{J}_l^1 \leq \sum_{l\in\mathcal{L}^2} J_l^1. \tag{37}$$

Consider now a link $l \in \mathcal{L}^1 \setminus \mathcal{L}^2$. Since $\hat{T}_l \leq T_l$ (Proposition 2) and $\hat{c}_l = c_l$, we have $\hat{f}_l = \hat{f}_l^1 \leq f_l^1 = f_l$. Thus, $\hat{J}_l^1 \leq J_l^1$, for all $l \in \mathcal{L}^1 \setminus \mathcal{L}^2$, and using (37) we conclude that $\hat{J}^1 \leq J^1$.

Suppose now that $j = 1$. According to Lemma 7, $\hat{J}^1 \leq J^1$ and $\hat{c}_l^1 \geq c_l^1, l \in \mathcal{L}$. Since $\hat{c}_l = c_l$ for any link $l \in \mathcal{L}^2 \setminus \{1\}$, the latter inequality implies that $\hat{f}_l^2 \leq f_l^2$. This means that when moving from configuration $\boldsymbol{c}$ to $\hat{\boldsymbol{c}}$, user 2 moves flow only into link 1. Since $\hat{T}_l \leq T_l$ for all links l and the delay at link 1 is always minimal among all links, we conclude that $\hat{J}^2 \leq J^2$. ■

APPENDIX C
PROOFS OF RESULTS IN SECTION V

Let us start by proving the following technical result.

Lemma 8: Consider two capacity configurations $\boldsymbol{c}, \hat{\boldsymbol{c}} \in \mathcal{C}$. For any user $i \in \mathcal{I}$:

1) if $\hat{\lambda}^i > \lambda^i$, then $\hat{f}_l^i \geq f_l^i$ for all links l such that $\hat{T}_l \leq T_l$;
2) if $\hat{\lambda}^i \leq \lambda^i$, then $\hat{f}_l^i \leq f_l^i$ for all links l such that $\hat{T}_l > T_l$;
3) there cannot be two links $m, n \in \mathcal{L}$, such that $\hat{T}_n > T_n, \hat{T}_m \leq T_m, \hat{f}_n^i > f_n^i$ and $\hat{f}_m^i < f_m^i$.

Proof: Assume that $\hat{\lambda}^i > \lambda^i$ and that there is a link l such that $\hat{T}_l \leq T_l$ and $0 \leq \hat{f}_l^i < f_l^i$. Then, using the optimality conditions (8), (9), we have

$$\hat{\lambda}^i \leq \hat{f}_l^i \hat{T}_l' + \hat{T}_l \leq f_l^i T_l' + T_l = \lambda^i$$

which contradicts $\hat{\lambda}^i > \lambda^i$. The proof of part 2) is symmetric. For part 3), assume that there are such links $n, m \in \mathcal{L}$. Since $f_m^i > \hat{f}_m^i \geq 0$, using (8), (9), we get

$$\lambda^i = f_m^i T_m' + T_m \geq \hat{f}_m^i \hat{T}_m' + T_m \geq \hat{\lambda}^i.$$

Then part 2) implies that $\hat{f}_n^i \leq f_n^i$, i.e., a contradiction. ■

Proof of Lemma 2: Assume by contradiction that $\hat{T}_1 > T_1$. We have to consider two cases.

Case 1—$\hat{T}_q > T_q$: For any link $l \in \mathcal{T}^+ \cup \{1, q\}$, we have $\hat{T}_l > T_l$. Thus

$$\sum_{l\in\mathcal{T}^+\cup\{1,q\}} (\hat{c}_l - \hat{f}_l) < \sum_{l\in\mathcal{T}^+\cup\{1,q\}} (c_l - f_l)$$

and

$$\sum_{l\in\mathcal{T}^+\cup\{1,q\}} \hat{f}_l > \sum_{l\in\mathcal{T}^+\cup\{1,q\}} f_l$$

since $\hat{c}_1 + \hat{c}_q = c_1 + c_q$ and $\hat{c}_n = c_n$, for all $n \in \mathcal{T}^+$. This implies that there must be a user j whose total flow in $\mathcal{T}^+ \cup \{1, q\}$ is higher under $\hat{\boldsymbol{c}}$, that is

$$\sum_{l\in\mathcal{T}^+\cup\{1,q\}} \hat{f}_l^j > \sum_{l\in\mathcal{T}^+\cup\{1,q\}} f_l^j$$

and thus

$$\sum_{l\in\mathcal{T}^-} \hat{f}_l^j < \sum_{l\in\mathcal{T}^-} f_l^j.$$

Therefore, there must be links $n \in \mathcal{T}^+ \cup \{1, q\}$ and $m \in \mathcal{T}^-$, such that $\hat{f}_n^j > f_n^j$ and $\hat{f}_m^j < f_m^j$. Since $\hat{T}_n > T_n$ and $\hat{T}_m \leq T_m$, this is a contradiction to part 3) of Lemma 8.

Case 2—$\hat{T}_q \leq T_q$: In this case $\hat{f}_q < f_q$, in view of $\hat{c}_q < c_q$. Since $\hat{f}_l \leq f_l$ for all $l \in \mathcal{T}^-$, this implies that

$$\sum_{l\in\mathcal{T}^-\cup\{q\}} \hat{f}_l < \sum_{l\in\mathcal{T}^-\cup\{q\}} f_l.$$

Thus there must be a user j such that

$$\sum_{l\in\mathcal{T}^-\cup\{q\}} \hat{f}_l^j < \sum_{l\in\mathcal{T}^-\cup\{q\}} f_l^j$$

and

$$\sum_{l\in\mathcal{T}^+\cup\{1\}} \hat{f}_l^j > \sum_{l\in\mathcal{T}^+\cup\{1\}} f_l^j.$$

This implies that there must be links $m \in \mathcal{T}^- \cup \{q\}$ and $n \in \mathcal{T}^+ \cup \{1\}$, such that $\hat{f}_m^j < f_m^j$ and $\hat{f}_n^j > f_m^j$. Since $\hat{T}_n > T_n$ and $\hat{T}_m \leq T_m$, this is a contradiction to part 3) of Lemma 8. Therefore, the delay on link 1 cannot be higher under capacity configuration $\hat{\boldsymbol{c}}$, that is, $\hat{T}_1 \leq T_1$. Let us now proceed to show the second part of the lemma, i.e., that $\hat{T}_q > T_q$.

Suppose that $\hat{T}_q \leq T_q$. Let us first show that this implies $\mathcal{T}^+ = \emptyset$. Assume by contradiction that $\mathcal{T}^+$ is nonempty. Since the total flow sent over links in $\mathcal{T}^+$ is higher under $\hat{\boldsymbol{c}}$, there must be a user j, such that

$$\sum_{l\in\mathcal{T}^+} \hat{f}_l^j > \sum_{l\in\mathcal{T}^+} f_l^j$$

and thus

$$\sum_{l\in\mathcal{T}^-\cup\{1,q\}} \hat{f}_l^j < \sum_{l\in\mathcal{T}^-\cup\{1,q\}} f_l^j.$$

Therefore, there exist links $n \in \mathcal{T}^+$ and $m \in \mathcal{T}^- \cup \{1, q\}$, such that $\hat{f}_n^j > f_n^j$ and $\hat{f}_m^j < f_m^j$. Since $\hat{T}_n > T_n$ and $\hat{T}_m \leq T_m$, this contradicts part 3) of Lemma 8. Therefore, $\mathcal{T}^+ = \emptyset$, and $\hat{c}_l - \hat{f}_l \geq c_l - f_l$, for all $l \in \mathcal{L}$. Recalling that

$$\sum_{l \in \mathcal{L}} (\hat{c}_l - \hat{f}_l) = \sum_{l \in \mathcal{L}} (c_l - f_l) = C - R$$

the last inequality must hold as an equality for all $l \in \mathcal{L}$, or equivalently $\hat{T}_l = T_l$, for all $l \in \mathcal{L}$. Note that this implies that no user can change its flow on any link in the network. To see this, assume that there exists a user j such that $\hat{f}_n^j > f_n^j \geq 0$ and $0 \leq \hat{f}_m^j < f_m^j$, for some links $m, n \in \mathcal{L}$. Then

$$\begin{aligned} \hat{\lambda}^j &\leq \hat{f}_m^j \hat{T}_m' + \hat{T}_m < f_m^j T_m' + T_m \\ &= \lambda^j \leq f_n^j T_n' + T_n < \hat{f}_n^j \hat{T}_n' + \hat{T}_n = \hat{\lambda}^j \end{aligned}$$

i.e., a contradiction. Thus, no user modifies its flow configuration, and $\hat{f}_l = f_l$, for all $l \in \mathcal{L}$. But this contradicts $\hat{T}_q \leq T_q$, since $\hat{c}_q < c_q$ and $\hat{f}_q = f_q$, and the result follows.

Proof of Lemma 3: We start by deriving an alternative expression for the user prices. Writing (15) as $\lambda^i (c_l - f_l)^2 = c_l - f_l^{-i}$ and summing over any set of links $A \subseteq \mathcal{L}^i$ that receive some flow from user i, we get

$$\lambda^i = \frac{\Sigma_{l \in A} (c_l - f_l^{-i})}{\Sigma_{l \in A} (c_l - f_l)^2} = \frac{\Sigma_{l \in A} c_l - R^{-i} + \Sigma_{l \in \mathcal{L}^1 \setminus A} f_l^{-i}}{\Sigma_{l \in A} (c_l - f_l)^2} \tag{38}$$

since

$$\sum_{l \in A} f_l^{-i} = R^{-i} - \sum_{l \in \mathcal{L} \setminus A} f_l^{-i}$$

and $\mathcal{L} \setminus A$ can be replaced by $\mathcal{L}^1 \setminus A$ because no user sends flow over links in $\mathcal{L} \setminus \mathcal{L}^1$. Taking $A = \mathcal{L}^i$ in (38), we have

$$\lambda^i = \frac{\Sigma_{l \in \mathcal{L}^i} c_l - R^{-i} + \displaystyle\sum_{l \in \mathcal{L}^1 \setminus \mathcal{L}^i} f_l}{\displaystyle\sum_{l \in \mathcal{L}^i} (c_l - f_l)^2} \tag{39}$$

since, for all $l \in \mathcal{L}^1 \setminus \mathcal{L}^i$, we have $f_l^i = 0$; therefore, $f_l^{-i} = f_l$.

From (8) and (11), the sum of the user prices over link $l \in \mathcal{L}$ is given by

$$\sum_{i \in \mathcal{I}_l} \lambda^i = f_l T_l' + I_l T_l.$$

Therefore, for any link $l \in \mathcal{L} \setminus \{1, q\}$, the following statements are equivalent:

$$l \in \mathcal{T}^+ \Leftrightarrow \hat{f}_l > f_l \Leftrightarrow \sum_{i \in \mathcal{I}_l} \hat{\lambda}^i > \sum_{i \in \mathcal{I}_l} \lambda^i \tag{40}$$

$$l \in \mathcal{T}^- \Leftrightarrow \hat{f}_l \leq f_l \Leftrightarrow \sum_{i \in \mathcal{I}_l} \hat{\lambda}^i \leq \sum_{i \in \mathcal{I}_l} \lambda^i. \tag{41}$$

We now proceed with the proof of the lemma. Let us first assume that $\hat{\lambda}^1 > \lambda^1$, and prove that $\hat{\lambda}^i > \lambda^i$ for all $i \in \mathcal{I}_q$ and $\hat{T}_l > T_l$ for all $l > q$. By assumption $\mathcal{I}_q \neq \emptyset$. If $I_q = 1$, user 1 is the only user sending flow on links $l > q$ and the result is immediate from $\hat{\lambda}^1 > \lambda^1$ and (40). Therefore, we concentrate on the case $I_q \geq 2$. Since $\hat{\lambda}^1 > \lambda^1$, (39) for $i = 1$ gives

$$\frac{\Sigma_{l \in \mathcal{L}^1} \hat{c}_l - R^{-1}}{\Sigma_{l \in \mathcal{L}^1} (\hat{c}_l - \hat{f}_l)^2} > \frac{\Sigma_{l \in \mathcal{L}^1} c_l - R^{-1}}{\Sigma_{l \in \mathcal{L}^1} (c_l - f_l)^2}.$$

Since

$$\sum_{l=1}^m \hat{c}_l = \sum_{l=1}^m c_l$$

for all $m \geq q$, and $q \in \mathcal{L}^1$, this implies

$$\sum_{l \in \mathcal{L}^1} (\hat{c}_l - \hat{f}_l)^2 < \sum_{l \in \mathcal{L}^1} (c_l - f_l)^2. \tag{42}$$

Let us first prove that $\hat{\lambda}^2 > \lambda^2$. If $\mathcal{L}^2 = \mathcal{L}^1$, the result is immediate from (39) and (42). Therefore, we have to consider only the case $\mathcal{L}^1 \setminus \mathcal{L}^2 \neq \emptyset$. Note that user 1 is the only user that sends flow on any link in $\mathcal{L}^1 \setminus \mathcal{L}^2$. Moreover, $q \notin \mathcal{L}^1 \setminus \mathcal{L}^2$, since $q \in \mathcal{L}^2$. Therefore, $\hat{\lambda}^1 > \lambda^1$ and (40) imply that

$$\hat{f}_l = \hat{f}_l^1 > f_l^1 = f_l, \qquad l \in \mathcal{L}^1 \setminus \mathcal{L}^2. \tag{43}$$

From (38) with $A = \mathcal{L}^2$ and $i = 1$, we have

$$\frac{\displaystyle\sum_{l \in \mathcal{L}^2} c_l - R^{-1}}{\displaystyle\sum_{l \in \mathcal{L}^2} (\hat{c}_l - \hat{f}_l)^2} = \hat{\lambda}^1 > \lambda^1 = \frac{\displaystyle\sum_{l \in \mathcal{L}^2} c_l - R^{-1}}{\displaystyle\sum_{l \in \mathcal{L}^2} (c_l - f_l)^2}$$

and thus

$$\sum_{l \in \mathcal{L}^2} (\hat{c}_l - \hat{f}_l)^2 < \sum_{1 \in \mathcal{L}^2} (c_l - f_l)^2.$$

Therefore, (39) and (43) give

$$\begin{aligned} \hat{\lambda}^2 &= \frac{\displaystyle\sum_{l \in \mathcal{L}^2} c_l - R^{-2} + \sum_{l \in \mathcal{L}^1 \setminus \mathcal{L}^2} \hat{f}_l}{\displaystyle\sum_{l \in \mathcal{L}^2} (\hat{c}_l - \hat{f}_l)^2} \\ &> \frac{\displaystyle\sum_{l \in \mathcal{L}^2} c_l - R^{-2} + \sum_{l \in \mathcal{L}^1 \setminus \mathcal{L}^2} f_l}{\displaystyle\sum_{l \in \mathcal{L}^2} (c_l - f_l)^2} = \lambda^2 \end{aligned}$$

which completes the proof for $i = 2$.

Proceeding inductively, let us assume that $\hat{\lambda}^i > \lambda^i$ for all $i \leq k < I_q$ and show that the same holds for $i = k + 1$. If $\mathcal{L}^{k+1} = \mathcal{L}^1$, the proof of $\hat{\lambda}^{k+1} > \lambda^{k+1}$ is immediate from (39) and (42). Thus, we only have to consider the case $\mathcal{L}^1 \setminus \mathcal{L}^{k+1} \neq \emptyset$. Let $\mathcal{I}_0$ denote the set of users that send flow on some link in $\mathcal{L}^1 \setminus \mathcal{L}^{k+1}$, that is, $\mathcal{I}_0 = \cup_{\mathcal{L}^1 \setminus \mathcal{L}^{k+1}} \mathcal{I}_l$. Note that user $k + 1$ does not send flow on any link in $\mathcal{L}^1 \setminus \mathcal{L}^{k+1}$. By Lemma 1, the same is true for all users $i > k + 1$. Thus, $\mathcal{I}_0 \subseteq \{1, \cdots, k\}$, and by the inductive hypothesis we have $\hat{\lambda}^i > \lambda^i$ for all $i \in \mathcal{I}_0$. Since $\mathcal{I}_l \subseteq \mathcal{I}_0$ for all $l \in \mathcal{L}^1 \setminus \mathcal{L}^{k+1}$, this implies

$$\sum_{i \in \mathcal{I}_l} \hat{\lambda}^i > \sum_{i \in \mathcal{I}_l} \lambda^i, \qquad l \in \mathcal{L}^1 \setminus \mathcal{L}^{k+1}$$

and since $q \notin \mathcal{L}^1 \setminus \mathcal{L}^{k+1}$, (40) gives

$$\hat{f}_l > f_l, \qquad l \in \mathcal{L}^1 \setminus \mathcal{L}^{k+1}. \tag{44}$$

For any user $i \in \mathcal{I}_0$, we have $k + 1 > i$ and thus $\mathcal{L}^{k+1} \subseteq \mathcal{L}^i$. Therefore, any user that sends flow on some link in $\mathcal{L}^1 \setminus \mathcal{L}^{k+1}$, also sends flow on all links in $\mathcal{L}^{k+1}$. Hence, taking $A=\mathcal{L}^{k+1}$ in (38), we get

$$\lambda^i = \frac{\sum_{l\in\mathcal{L}^{k+1}} c_l - R^{-i} + \sum_{l\in\mathcal{L}^1\setminus\mathcal{L}^{k+1}} (f_l - f_l^i)}{\sum_{l\in\mathcal{L}^{k+1}} (c_l - f_l)^2}, \qquad i\in\mathcal{I}_0. \tag{45}$$

From (39), we have

$$\begin{aligned}\lambda^{k+1} &= \frac{\sum_{l\in\mathcal{L}^{k+1}} c_l - R^{-(k+1)} + \sum_{l\in\mathcal{L}^1\setminus\mathcal{L}^{k+1}} f_l}{\sum_{l\in\mathcal{L}^{k+1}} (c_l - f_l)^2}\\ &= \frac{\sum_{l\in\mathcal{L}^{k+1}} c_l - R^{-i} + \sum_{l\in\mathcal{L}^1\setminus\mathcal{L}^{k+1}} (f_l - f_l^i)}{\sum_{l\in\mathcal{L}^{k+1}} (c_l - f_l)^2}\\ &\quad - \frac{r^i - r^{k+1} - \sum_{l\in\mathcal{L}^1\setminus\mathcal{L}^{k+1}} f_l^i}{\sum_{l\in\mathcal{L}^{k+1}} (c_l - f_l)^2}, \qquad i\in\mathcal{I}_0\end{aligned} \tag{46}$$

where we have used $R^{-(k+1)} = R^{-i} + r^i - r^{k+1}$. Equations (45) and (46) give

$$\lambda^{k+1} = \lambda^i - \frac{\sum_{l\in\mathcal{L}^{k+1}} f_l^i - r^{k+1}}{\sum_{l\in\mathcal{L}^{k+1}} ,(c_l - f_l)^2}, \qquad i\in\mathcal{I}_0. \tag{47}$$

Let us assume (by contradiction) that $\hat{\lambda}^{k+1} \le \lambda^{k+1}$. Since $q\in\mathcal{L}^{k+1}$, we have

$$\sum_{l\in\mathcal{L}^{k+1}} \hat{c}_l = \sum_{l\in\mathcal{L}^{k+1}} c_l.$$

Using (46), we get

$$\frac{\sum_{l\in\mathcal{L}^{k+1}} c_l - R^{-(k+1)} + \sum_{l\in\mathcal{L}^1\setminus\mathcal{L}^{k+1}} \hat{f}_l}{\sum_{l\in\mathcal{L}^{k+1}} (\hat{c}_l - \hat{f}_l)^2} \le \frac{\sum_{l\in\mathcal{L}^{k+1}} c_l - R^{-(k+1)} + \sum_{l\in\mathcal{L}^1\setminus\mathcal{L}^{k+1}} f_l}{\sum_{l\in\mathcal{L}^{k+1}} (c_l - f_l)^2}$$

and since, from (44)

$$\sum_{l\in\mathcal{L}^1\setminus\mathcal{L}^{k+1}} \hat{f}_l > \sum_{l\in\mathcal{L}^1\setminus\mathcal{L}^{k+1}} f_l$$

we have

$$\sum_{l\in\mathcal{L}^{k+1}} (\hat{c}_l - \hat{f}_l)^2 > \sum_{l\in\mathcal{L}^{k+1}} (c_l - f_l)^2. \tag{48}$$

Using (47) and $\hat{\lambda}^{k+1}\le\lambda^{k+1}$ we have, for all $i\in\mathcal{I}_0$

$$\frac{\sum_{l\in\mathcal{L}^{k+1}} \hat{f}_l^i - r^{k+1}}{\sum_{l\in\mathcal{L}^{k+1}} (\hat{c}_l - \hat{f}_l)^2} - \frac{\sum_{l\in\mathcal{L}^{k+1}} f_l^i - r^{k+1}}{\sum_{l\in\mathcal{L}^{k+1}} (c_l - f_l)^2} \ge \hat{\lambda}^i - \lambda^i > 0$$

since $\hat{\lambda}^i > \lambda^i$, for all $i\in\mathcal{I}_0$. In view of (48), this implies, for all $i\in\mathcal{I}_0$

$$\sum_{l\in\mathcal{L}^{k+1}} \hat{f}_l^i > \sum_{l\in\mathcal{L}^{k+1}} f_l^i$$

and thus

$$\sum_{l\in\mathcal{L}^1\setminus\mathcal{L}^{k+1}} \hat{f}_l^i < \sum_{l\in\mathcal{L}^1\setminus\mathcal{L}^{k+1}} f_l^i.$$

By the definition of $\mathcal{I}_0$, summing the last inequality over all $i\in\mathcal{I}_0$, we get

$$\sum_{l\in\mathcal{L}^1\setminus\mathcal{L}^{k+1}} \hat{f}_l < \sum_{l\in\mathcal{L}^1\setminus\mathcal{L}^{k+1}} f_l$$

which stands in contradiction with (44). Hence $\hat{\lambda}^{k+1} > \lambda^{k+1}$. Thus, by induction, we have $\hat{\lambda}^i > \lambda^i$, for all $i\in\mathcal{I}_q$. Finally, for any link $l>q$, inequality $\hat{\lambda}^i > \lambda^i$ for all $i\in\mathcal{I}_l$, together with (40) implies $\hat{T}_l > T_l$. This completes the proof for case $\hat{\lambda}^1 > \lambda^1$. It remains to be shown that if $\hat{\lambda}^1 \le \lambda^1$, then $\hat{\lambda}^i \le \lambda^i$ for all $i\in\mathcal{I}_q$ and $\hat{T}_l \le T_q$ for all $l>q$. The proof is symmetric. ■

Proof of Lemma 4: It suffices to show that for any link $l<q-1$, if $l\in\mathcal{T}^+$, then $l+1\in\mathcal{T}^+$. Assume by contradiction that there exists a link $l<q-1$ such that $l\in\mathcal{T}^+$ and $l+1\in\mathcal{T}^-$. Then (40) and (41) give

$$\hat{f}_{l+1} \le f_{l+1} \quad\text{and}\quad \sum_{i\in\mathcal{I}_{l+1}} \hat{\lambda}^i \le \sum_{i\in\mathcal{I}_{l+1}} \lambda^i \tag{49}$$

$$\hat{f}_l > f_l \quad\text{and}\quad \sum_{i\in\mathcal{I}_l} \hat{\lambda}^i > \sum_{i\in\mathcal{I}_l} \lambda^i. \tag{50}$$

If $\mathcal{I}_{l+1}=\mathcal{I}_l$, (49) and (50) lead to a contradiction. Thus, we need to consider only the case $\mathcal{I}_l\setminus\mathcal{I}_{l+1}\ne\emptyset$. Note that this is the set of users that send flow on link l and do not send flow on link $l+1$. For any such user $i, L^i = l$. Summing (8) for link l over all $i\in\mathcal{I}_{l+1}\subset\mathcal{I}_l$ and using (11), we get

$$\begin{aligned}\sum_{i\in\mathcal{I}_{l+1}} \hat{f}_l^i \hat{T}_l' + I_{l+1}\hat{T}_l &= \sum_{i\in\mathcal{I}_{l+1}} \hat{\lambda}^i \le \sum_{i\in\mathcal{I}_{l+1}} \lambda^i\\ &= \sum_{i\in\mathcal{I}_{l+1}} f_l^i T_l' + I_{l+1}T_l\end{aligned}$$

and since $l\in\mathcal{T}^+$, this implies

$$\sum_{i\in\mathcal{I}_{l+1}} \hat{f}_l^i < \sum_{i\in\mathcal{I}_{l+1}} f_l^i. \tag{51}$$

Recalling that $\hat{f}_l > f_l$, (51) gives

$$\begin{aligned}\sum_{i\in\mathcal{I}_l\setminus\mathcal{I}_{l+1}} \hat{f}_l^i &= \hat{f}_l - \sum_{i\in\mathcal{I}_{l+1}} \hat{f}_l^i > f_l - \sum_{i\in\mathcal{I}_{l+1}} f_l^i\\ &= \sum_{i\in\mathcal{I}_l\setminus\mathcal{I}_{l+1}} f_l^i\end{aligned}$$

and since for any $i \in \mathcal{I}_l \setminus \mathcal{I}_{l+1}$ we have

$$r^i = \sum_{m=1}^{l} f_m^i = \sum_{m=1}^{l} \hat{f}_m^i$$

this implies

$$\sum_{m<l} \sum_{i\in\mathcal{I}_l\setminus\mathcal{I}_{l+1}} \hat{f}_m^i < \sum_{m<l} \sum_{i\in\mathcal{I}_l\setminus\mathcal{I}_{l+1}} f_m^i. \tag{52}$$

Subtracting (49) from (50), we obtain

$$\sum_{i\in\mathcal{I}_l\setminus\mathcal{I}_{l+1}} \hat{\lambda}^i > \sum_{i\in\mathcal{I}_l\setminus\mathcal{I}_{l+1}} \lambda^i.$$

Summing (8) over all $i \in \mathcal{I}_l \setminus \mathcal{I}_{l+1}$, for any link $m \leq l$, it is easy to see that

$$\begin{aligned} &\sum_{i\in\mathcal{I}_l\setminus\mathcal{I}_{l+1}} \hat{f}_m^i \hat{T}_m' + (I_l - I_{l+1})\hat{T}_m \\ &\quad > \sum_{i\in\mathcal{I}_l\setminus\mathcal{I}_{l+1}} f_m^i T_m' + (I_l - I_{l+1})T_m. \end{aligned} \tag{53}$$

Consider any link $m<l$ such that $m \in \mathcal{T}^- \cup \{1\}$. Since $\hat{T}_m \leq T_m$, (53) gives

$$\sum_{i\in\mathcal{I}_l\setminus\mathcal{I}_{l+1}} \hat{f}_m^i > \sum_{i\in\mathcal{I}_l\setminus\mathcal{I}_{l+1}} f_m^i.$$

Hence, from (52), we have

$$\sum_{\substack{m\in\mathcal{T}^+\\ m<l}} \sum_{i\in\mathcal{I}_l\setminus\mathcal{I}_{l+1}} \hat{f}_m^i < \sum_{\substack{m\in\mathcal{T}^+\\ m<l}} \sum_{i\in\mathcal{I}_l\setminus\mathcal{I}_{l+1}} f_m^i. \tag{54}$$

Using a similar argument as in the proof of (51), one can see that (49) implies

$$\sum_{i\in\mathcal{I}_{l+1}} \hat{f}_m^i < \sum_{i\in\mathcal{I}_{l+1}} f_m^i, \quad \mathcal{T}^+ n\{m<l\}.$$

Summing this inequality over all $m \in \mathcal{T}^+$ and $m<l$, and adding it to (54), we have

$$\sum_{\substack{m\in\mathcal{T}^+\\ m<l}} \sum_{i\in\mathcal{I}_l} \hat{f}_m^i < \sum_{\substack{m\in\mathcal{T}^+\\ m<l}} \sum_{i\in\mathcal{I}_l} f_m^i. \tag{55}$$

For any link $m \in \mathcal{T}^+, \hat{f}_m > f_m$. Therefore, the total flow sent through the set of links $\mathcal{T}^+ \cap \{1, \cdots, l-1\}$ is larger under configuration $\hat{\boldsymbol{c}}$, and (55) implies

$$\sum_{\substack{m\in\mathcal{T}^+\\ m<l}} \sum_{i\in\mathcal{I}\setminus\mathcal{I}_l} \hat{f}_m^i > \sum_{\substack{m\in\mathcal{T}^+\\ m<l}} \sum_{i\in\mathcal{I}\setminus\mathcal{I}_l} f_m^i.$$

Therefore, there exists some user $j \in \mathcal{I} \setminus \mathcal{I}_l$, such that

$$\sum_{\substack{m\in\mathcal{T}^+\\ m<l}} \hat{f}_m^j > \sum_{\substack{m\in\mathcal{T}^+\\ m<l}} f_m^j \Rightarrow \sum_{m\in\mathcal{T}^+} \hat{f}_m^j > \sum_{m\in\mathcal{T}^+} f_m^j \tag{56}$$

since $j \in \mathcal{I} \setminus \mathcal{I}_l$, i.e., user j does not send any flow to links $m \geq l$. Note that (56) implies that there exists some link $m' \in \mathcal{T}^+$, such that $\hat{f}_{m'}^j > f_{m'}^j$, and thus $\hat{\lambda}^j > \lambda^j$. Moreover, from Lemma 8, we have $\hat{f}_m^j \geq f_m^j$, for all $m \in \mathcal{T}^-$. Since $\hat{f}_q^j = f_q^j = 0$, the last inequality together with (56) imply

$$\hat{f}_1^j - f_1^j = \sum_{m\in\mathcal{T}^+\cup\mathcal{T}^-} f_m^j - \sum_{m\in\mathcal{T}^+\cup\mathcal{T}^-} \hat{f}_m^j < 0$$

which together with $\hat{T}_1 \leq T_1$ (Lemma 2) implies that $\hat{\lambda}^j < \lambda^j$. But this is a contradiction to $\hat{\lambda}^j > \lambda^j$. Hence, it must be $l+1 \in \mathcal{T}^+$. ■

Remark: In the proof of the lemma above, we assumed that there exists a link $l<q-1$, such that $l \in \mathcal{T}^+$ and $l+1 \in \mathcal{T}^-$, and arrived at a contradiction. The only implication of the assumption $l+1 \in \mathcal{T}^-$ that was used was

$$\sum_{i\in\mathcal{I}_{l+1}} \hat{\lambda}^i \leq \sum_{i\in\mathcal{I}_{l+1}} \lambda^i.$$

Thus, the same proof can be used to show that if $q-1 \in \mathcal{T}^+$, then

$$\sum_{i\in\mathcal{I}_q} \hat{\lambda}^i > \sum_{i\in\mathcal{I}_q} \lambda^i.$$

ACKNOWLEDGMENT

The authors would like to thank R. Rom and N. Shimkin for helpful discussions.

REFERENCES

[1] R. B. Myerson, *Game Theory: Analysis of Conflict*. Cambridge, MA: Harvard Univ. Press, 1991.
[2] D. Fudenberg and J. Tirole, *Game Theory*. Cambridge, MA: MIT Press, 1992.
[3] M.-T. T. Hsiao and A. A. Lazar, "Optimal decentralized flow control of Markovian queueing networks with multiple controllers, " *Performance Evaluation*, vol. 13, no. 3, pp. 181–204, 1991.
[4] Z. Zhang and C. Douligeris, "Convergence of synchronous and asynchronous greedy algorithms in a multiclass telecommunications environment, " *IEEE Trans. Commun.*, vol. 40, pp. 1277–1281, Aug. 1992.
[5] E. Altman, "Flow control using the theory of zero-sum Markov games," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 814–818, Apr. 1994.
[6] Y. A. Korilis and A. A. Lazar, "On the existence of equilibria in noncooperative optimal flow control, " *J. ACM*, vol. 42, pp. 584–613, May 1995.
[7] A. A. Economides and J. A. Silvester, "Multi-objective routing in integrated services networks: A game theory approach, " in *Proc. INFOCOM'91*, pp. 1220–1225.
[8] E. Altman and N. Shimkin, "Worst-case and Nash routing policies in parallel queues with uncertain service allocations," IMA Preprint Series 1120, Institute for Mathematics and Applications, Univ. Minnesota, Minneapolis, 1993.
[9] A. Orda, R. Rom, and N. Shimkin, "Competitive routing in multiuser communication networks, " *IEEE/ACM Trans. Networking*, vol. 1, pp. 510–521, Oct. 1993.
[10] A. A. Lazar, A. Orda, and D. E. Pendarakis, "Virtual path bandwidth allocation in multi-user networks," in *Proc. IEEE INFOCOM'95*, Boston, MA, pp. 312–320.
[11] S. J. Shenker, "Making greed work in networks: A game-theoretic analysis of switch service disciplines, " *IEEE/ACM Trans. Networking*, vol. 3, pp. 819–831, Dec. 1995.
[12] D. Bertsekas and R. Gallager, *Data Networks*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1992.
[13] P. A. Humblet and S. R. Soloway, "Algorithms for data communication networks—Part 1," Codex Corp., Tech. Rep., 1986.
[14] P. A. Humblet, S. R. Soloway, and B. Steinka, "Algorithms for data communication networks—Part 2," Codex Corp., Tech. Rep., 1986.
[15] A. Kershenbaum, *Telecommunications Network Design Algorithms*. New York: McGraw-Hill, 1993.
[16] Information Sciences Institute, "Internet Protocol," Univ. Southern California, Marina del Rey, CA, Tech. Rep. RFC 791, Sept. 1981.

[17] S. Deering and R. Hinden, "Internet protocol version 6 specification," Internet Draft, IETF, Mar. 1995.
[18] J. J. Garrahan, P. A. Russo, K. Kitami, and R. Kung, "Intelligent network overview, " *IEEE Commun. Mag.*, vol. 31, pp. 30–36, Mar. 1993.
[19] P. A. Russo, K. Bechard, E. Brooks, R. L. Corn, R. Gove, W. L. Honig, and J. Young, "IN rollout in the United States," *IEEE Commun. Mag.*, vol. 31, pp. 56–63, Mar. 1993.
[20] M. Gerla and L. Kleinrock, "On the topological design of computer communication networks," *IEEE Trans. Commun.*, vol. COM-25, no. 1, pp. 48–60, 1977.
[21] W. I. Zangwill and C. B. Garcia, *Pathways to Solutions, Fixed Points and Equilibria*. Englewood Cliffs, NJ: Prentice-Hall, 1981.
[22] J. E. Cohen and F. P. Kelly, "A paradox of congestion in a queuing network," *J. Appl. Probability*, vol. 27, pp. 730–734, 1990.
[23] Y. A. Korilis, A. A. Lazar, and A. Orda, "Achieving network optima using Stackelberg routing strategies," in *IEEE/ACM Trans. Networking*, Feb. 1997. Available at URL http://www.ctr.columbia.edu/∼john/stackel.html.
[24] J. K. MacKie-Mason and H. R. Varian, "Pricing cognestible network resources, " *IEEE J. Select. Areas Commun.*, vol. 13, pp. 1141–1149, Sept. 1995.
[25] M. Schwartz, *Telecommunication Networks: Protocols, Modeling and Analysis*. Reading, MA: Addison-Wesley, 1987.
[26] A. Ben-Israel, A. Ben-Tal, and S. Zlobec, *Optimality in Nonlinear Programming: A Feasible Directions Approach*. New York: Wiley, 1981.
[27] W. Nicholson, *Intermediate Microeconomics and its Applications*, 6th ed. New York: Dryden, 1994.
[28] Y. A. Korilis, A. A. Lazar, and A. Orda, "Capacity allocation under noncooperative routing," Center Telecommun. Res., Columbia Univ., New York, NY, Tech. Rep. 372-94-19, 1994. Available at URL http://www.ctr.columbia.edu/∼john/capacity.html.
[29] H. L. Royden, *Real Analysis*, 3rd ed. New York: Macmillan, 1988.

Yannis A. Korilis (S'89–M'95) was born in Thessaloniki, Greece, in 1966. He received the Diploma from the National Technical University of Athens, Greece, in 1989, and the M.S., M.Phil., and Ph.D. degrees from Columbia University, New York, in 1990, 1992, and 1995, respectively, all in electrical engineering.

During the summer of 1993, he worked as Senior Staff Technologist at Bellcore, Morristown, NJ. Since 1995, he has been a Member of the Technical Staff at Bell Laboratories, Lucent Technologies, Holmdel, NJ. During the spring of 1996, he held an adjunct Assistant Professor position at Columbia University, New York, NY. His research interests include control and management of multimedia networks, pricing of multimedia services, economic and game theoretic models for networking, and distributed algorithms.

Dr. Korilis has received various academic excellence awards from the National Scholarship Foundation and the Technical Chamber of Greece. He has been a winner in the Interschool Fellowship Competition at Columbia University for four consecutive years. He has also received the E. I. Jury Award for excellence in the area of Systems, Communications and Signal Processing. He is a member of the ACM and INFORMS.

Aurel A. Lazar (S'77–M'80–SM'90–F'93) is a Professor of Electrical Engineering at Columbia University. His research interests span both theoretical and experimental studies of telecommunication networks and multimedia systems. The theoretical research he conducted during the 1980's pertains to the modeling, analysis, and control of broadband networks. He formulated optimal flow and admission control problems and, by building upon the theory of point processes, derived control laws for Markovian queueing network models in single control as well as game theoretic settings. He was the chief architect of two experimental networks, generically called MAGNET. This work introduced traffic classes with explicit quality of service constraints to broadband switching and led to the concepts of schedulable, admissible load and contract regions in real-time control of broadband networks. In the early 1990's his research efforts shifted to the foundations of the control, management and telemedia architecture of future multimedia networks. His involvement with gigabit networking research lead to the first fully operational service management system on ATM-based broadband networks. The system was implemented on top of AT&T's XUNET III gigabit platform spanning the continental United States. His management and control research pioneered the application of virtual reality to the management of ATM-based broadband networks. His current research in broadband networking with quality of service guarantees focusses on modeling of video streams and analyzing their multiplexing behavior, with emphasis on multiple time scales and subexponentiality. He is also leading investigations into multimedia networking architectures supporting interoperable exchange mechanisms for interactive and on demand multimedia applications with quality of service requirements. The main focus of this work is on building an open signalling infrastructure that enables the rapid creation, deployment, and management of multimedia services. Also, he was instrumental in establishing the OPENSIG international working group with the goal of exploring network programability and next-generation signalling technology.

Ariel Orda (S'84–M'92) was born in Argentina in 1961. He received the B.Sc. (summa cum laude), M.Sc., and D.Sc. degrees in electrical engineering from the Technion-Israel Institute of Technology, Haifa, Israel, in 1983, 1985, and 1991, respectively.

From 1983 to 1985, he was a Teaching Assistant at the Technion. From 1985 to 1990, he served in the IDF as a Research Engineer. From 1990 to 1991, he was a Teaching Instructor at the Technion. In 1991 he joined the Faculty of Electrical Engineering at the Technion, as a Lecturer. During the academic year 1993–1994, he was a Visiting Scientist with the Center for Telecommunications Research, Columbia University, New York, NY. Since 1994, he has been a Senior Lecturer at the Technion. During 1991–1996, he held consulting positions in AT&T Bell Laboratories and in the Israeli industry. During the summers of 1995 and 1996, he was an Academic Visitor at IBM T. J. Watson Research Center. His current research interests include the control and management of broadband networks, the application of game theory to computer networking, QoS routing, and on-line network algorithms.

In 1991, Dr. Orda received the award of the Chief Scientist in the Ministry of Communication and a Gutwirth Award for Outstanding Distinction. In 1993, he received the research award of the Association of Computer and Electronic Industries in Israel.