HAL
open science

On the Inefficiency of Atomic Routing Games over Parallel Links

Olivier Brun, Josu Doncel

To cite this version:

Olivier Brun, Josu Doncel. On the Inefficiency of Atomic Routing Games over Parallel Links. 2023. hal-03863930v2

HAL Id: hal-03863930

https://laas.hal.science/hal-03863930v2

Preprint submitted on 4 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On the Inefficiency of Atomic Routing Games over Parallel Links

Olivier Brun[a], Josu Doncel[a]

[a] LAAS-CNRS, Toulouse, France
[b] University of the Basque Country, UPV/EHU, Leioa, Spain

October 3, 2023

Abstract

Several recent works on non-atomic routing games suggest that the performance degradation of selfish routing with respect to optimal routing is overall low and far from worst-case scenarios. In this work, we study the performance degradation induced by the lack of coordination in an atomic routing game over parallel links in which there are two types of links. The latency function of "cheap" links is of the form $c_1\phi(x)$, whereas the latency function of "expensive" links is of the form $c_2\phi(x)$, where $c_2 > c_1$. We obtain an explicit characterization of the optimal and equilibrium flow configurations, and establish sufficient conditions on the latency function $\phi(x)$ under which the worst traffic conditions occur when all users have the same traffic demand and the total traffic demand is such that "expensive" link are marginally used by selfish routing. We also obtain some partial results on the worst network configuration for the inefficiency of selfish routing. All in all, our results suggest that the worst-case scenario for the inefficiency of selfish routing corresponds to very specific traffic conditions and to highly asymmetric network configurations, and thus that the *Price of Anarchy* is probably an overly pessimistic performance measure for non-cooperative routing games, as advocated in the above-mentioned works.

1 Introduction

1.1 Motivation

In networks based on a centralized routing scheme, a central node computes the least-cost path between source and destination nodes by using some global knowledge of the network and then distributes the resulting routes to other nodes so that they can forward user traffic. The main advantage of such an approach is that it can enforce an optimal routing policy minimizing the overall cost (e.g., the overall mean packet-delay) of all users. However, it is broadly admitted that a centralized routing scheme cannot be used in large networks due to scalability, robustness and complexity reasons. An alternative approach is to let each network user selfishly decide on which path to route its traffic demand according to its own interest. Although more robust and scalable, this decentralized scheme may lead to a loss in performance as the individual optimizations performed by many interacting self-interested users does not necessarily converge to an optimal routing policy.

Noncooperative routing games provide the natural framework to study the performance degradation in the above decentralized routing scheme. These games are mathematical models of the equilibrium flow configuration resulting from the routing decisions made by selfish, uncoordinated network users. One usually distinguishes two types of such games. Atomic routing games refer to games in which there are finitely many users, and the resulting flow configuration is known as the Nash equilibrium [22]. It corresponds to a routing policy in which no user can decrease its own routing cost by deviating from its strategy unilaterally. In contrast, non-atomic routing games refer to games in which there is a continuum of users, each one controlling a negligible amount of traffic, and the resulting flow configuration corresponds to the so-called Wardrop equilibrium [31], according to which the traffic demand of each user is forwarded along a minimum-cost path. For both types of games, the equilibrium flow configuration does not always correspond to that of a globally optimal routing policy. In game theory terminology, such an optimal policy is usually referred to as a social optimum as it minimizes the sum of all user costs.

A vast body of literature has been devoted to the study of the inefficiency of selfish routing under a variety of traffic models. The most popular measure of the inefficiency of equilibria is the Price of

Anarchy (PoA) which was introduced by Koutsoupias and Papadimitriou [20] and is defined as the performance ratio between the overall cost of an optimal routing policy and that of the worst Nash/Wardrop equilibrium (that is, an equilibrium with the largest social cost). As discussed in Section 1.3, it has been shown that the PoA of some selfish routing games can be arbitrarily large. Several recent works aiming at understanding when is selfish routing bad suggest however that the PoA is an overly pessimistic measure and that non-cooperative routing achieves near-optimal performance in most realistic settings. Nevertheless, most of these works have been carried out for non-atomic routing games, which are usually much simpler to analyze thanks to the assumption of a continuum of players.

In contrast to the above mentioned works, the present paper studies the efficiency of selfish routing in atomic routing games. Owing to the complexity of the analysis, we restrict ourselves to a topology of parallel links, as introduced in the seminal paper of Orda, Rom and Shimkin [24]. Given a strictly increasing and convex function $\phi(x)$ and two cost parameters c_1 and c_2 such that $c_1 < c_2$, we further assume that there are two types of links: "cheap" links whose cost is a function of the flow on the link of the form $c_1\phi(x)$, and "expensive" links whose cost function is of the form $c_2\phi(x)$. We also assume a finite number of users, each one splitting its traffic demand over the parallel links so as to minimize its own routing cost, which is the sum of the costs incurred over all the links it uses. As the cost of a link is a non-decreasing function of the traffic flow on that link, the optimal strategy of a user depends on how the other users split their traffic demands. In this context, a Nash equilibrium is a flow configuration on the links in which no user can benefit from a unilateral deviation of its own routing strategy.

For the above atomic routing game, we assume that the number of links of each type as well as their cost parameters are fixed, and study the efficiency of selfish routing as a function of the traffic demands of users. Our goal is to understand under which traffic conditions the worst inefficiency of selfish routing is obtained for a fixed network configuration. As a measure of efficiency, we adopt the ratio of the social costs obtained at the Nash equilibrium and under a socially optimal solution. This ratio is at least one, when selfish routing is optimal, and is upper bounded by the PoA of the game. The worst value of this ratio (over all possible vectors of traffic demands) corresponds to the *Inefficiency* of the routing game, a concept introduced in [16] for load-balancing games. As opposed to the PoA, the *Inefficiency* depends on the network configuration. By calculating the worst possible value of the *Inefficiency* over all network configurations, one retrieves the PoA.

1.2 Contributions

The main contributions in this work are the following:

- For an arbitrary network configuration, we characterize the traffic conditions associated with the *Inefficiency*, i.e., the traffic conditions under which the ratio of social costs is maximum. Specifically, we establish sufficient conditions on the latency function $\phi(x)$ under which the worst traffic conditions occur when all users have the same traffic demand and when the total traffic in the network is such that "expensive" links are marginally used at Nash equilibrium.

- We show that these sufficient conditions are in particular satisfied by $\phi(x) = (1 + x)^m$ for $m \geq 2$ and $\phi(x) = e^{\nu x}$ for $\nu > 0$. These latency functions are used throughout the paper for illustration purposes. The former belongs to the class of polynomial latency functions commonly used in transportation research to model travel times in road networks [30]. The latter is reminiscent of exponential growth models used to model many physical phenomena and corresponds to situations in which the rate at which the delay over a link increases is proportional to its value, that is, $\phi'(x) = \nu\,\phi(x)$.

- We provide an explicit characterization of the optimal and equilibrium flow configurations on the links. In particular, we show that under the worst traffic conditions the ratio of the flows obtained under the decentralized and centralized schemes is maximum for the "cheap" links, whereas it is minimum for the "expensive" links. The latter result holds true under weaker conditions on the latency function $\phi(x)$, which are satisfied in particular by $\phi(x) = 1 + x^m$ and $\phi(x) = 1/(1 - x)^m$ for $m \geq 2$.

- We show that the *Inefficiency* depends only on the ratio of the number of links of each type and on the ratio $\frac{c_1}{c_2}$ of their costs (but not directly on the total number of links nor on their costs). We prove that it implies that the worst value of the *Inefficiency* is obtained when there is only one "cheap" link and the rest of the links are "expensive".

- For $\phi(x) = e^{\nu x}$, we provide a lower-bound on the PoA. Besides, we conjecture from numerical experiments that the PoA is obtained when the cost parameter of the "expensive" links c_2 is infinitely larger than that of "cheap" links c_1. Assuming that this conjecture holds, we obtain an upper-bound on the PoA. This is in contrast to the situation for $\phi(x) = (1+x)^m$ for which we observe that the *Inefficiency* is not monotone as a function of the ratio $\frac{c_1}{c_2}$.

1.3 Related work

We first review relevant works on nonatomic routing games. The analysis of the efficiency of Wardrop equilibria has a long history which dates back to 1920 and the well-known Pigou's example which shows that the outcome of a selfish routing game can be inferior to a centrally designed outcome [25]. It was shown in 1968 by Dietrich Braess that adding resources to a transportation network can sometimes hurt performance at equilibrium, a phenomemon now known as the famous Braess's paradox [5]. More recently, Roughgarden and Tardos have shown that the value of the PoA of nonatomic congestion games with affine costs is bounded above by 4/3, and that this bound is tight [27]. This shows that selfish routing is always efficient for such routing games. However, it was shown in [26] that the PoA for networks with latency functions that are polynomials with nonnegative coefficients and degree at most d is asymptotically $\Theta\left(\frac{d}{\log(d)}\right)$ as $d \to \infty$, indicating that selfish routing can be very inefficient in such games. Similarly, it was shown in [19] that the PoA of nonatomic load-balancing games over parallel servers corresponds to the number of servers (see also [1, 4]). Other relevant works on the PoA of nonatomic routing games are [28, 14, 12, 13].

On the empirical side, several recent works studied the efficiency of Wardrop equilibria in real networks and observed that the performance degradation with respect to optimal routing is overall low in spite of large theoretical values of the PoA. For instance, Monnot et al. analyze data of commuting students in Singapore and conclude that routing choices are near optimal [21] (see also [18] for a similar study). On a more theoretical side, the authors in [9, 10, 8] prove that Wardrop equilibria are efficient when the network is lightly or highly congested. The authors in [32] extend the results of [8] to a more general setting and show that selfish routing is efficient when the total traffic demand gets very large. In [11], Cominetti et al. study the efficiency of Wardrop equilibria as a function of the total traffic demand in the network. As an efficiency metric, they focus as we do on the ratio of social costs obtained under the equilibrium and optimal routing strategies. For affine link costs, they show that this ratio has a finite number of local maxima, which are achieved where the set of active links changes. In summary, all the above works suggest that the PoA is an overly pessimistic measure of the inefficiency of selfish routing and that the performance degradation is often low and far from the worst-case scenarios.

Efficiency results for atomic routing games are much scarcer, as these games are much harder to analyze. Most known results are only valid for topologies of parallel links, as introduced in [24], where the existence and unicity of the Nash equilibrium are shown for a broad class of latency functions. Ayesta et al. consider in [3] an atomic load-balancing game in which K users selfishly route their jobs to a system of S parallel M/G/1/PS servers and prove that in this case the PoA is of the order $\sqrt{K}$, independently of the number S of servers as long as $S \geq 2$. Other results on the PoA of selfish load balancing can be found in [29, 2, 15, 7] (see also [17] for a recent survey).

A closely related work to ours is presented in [16], where the authors consider an atomic load-balancing game with "fast" and "slow" servers, which are modeled as M/G/1/PS queues. The setting they consider is thus similar to the one considered here, but restricted to the latency function $\phi(x) = 1/(1-x)$ for the parallel links representing the servers. They study the ratio of social costs as a function of the total incoming traffic in the system and prove that this ratio attains its maximum when the "slow" servers are marginally used by the decentralized load-balancing scheme. Our work extends the results of [16]. Whereas the analysis in [16] heavily relies on the properties of the function $\phi(x) = 1/(1-x)$, and in particular on the fact that $\phi'(x) = \phi(x)^2$, we establish sufficient technical conditions under which a similar result holds and show that these conditions are in particular satisfied by the latency functions $\phi(x) = e^{\nu x}$ and $\phi(x) = (1+x)^m$. We emphasize however that all our results hold for any other latency function ϕ satisfying Assumptions 1-5 stated in Appendix A. Besides, we show in Section 4.1 that the worst traffic conditions for the ratio of social costs are precisely those that maximize the ratio of the equilibrium and optimal flows on the "cheap" links, a result which does not appear in [16] and hold under weaker conditions satisfied as well by other latency functions such as $\phi(x) = 1/(1-x)^m$ or $\phi(x) = 1+x^m$. Our work also complements the work in [11] which also studies the ratio of social costs as a function of the total traffic demand in general network topologies, but for a nonatomic routing game and affine cost functions on the links, whereas we consider an atomic routing game over parallel links and non-linear

cost functions on the links.

1.4 Organization of the paper

We present the mathematical model of the atomic routing game considered in this paper in Section 2. In Section 3, we establish some preliminary results and obtain the characterization of the equilibrium and optimal flow configurations as solution of convex optimization problems. Our results on the worst-case total traffic are established in Section 4. In Section 5, we investigate the worst network configuration for the inefficiency of the decentralized routing scheme, that is, we study how the PoA is obtained from the *Inefficiency*. In Section 6, we discuss some possible extensions of this work. Finally, the conclusions of this work are drawn in Section 7.

2 Problem formulation

2.1 Non-cooperative routing game

As illustrated in Figure 1, we consider an atomic splittable routing game in which K users have to send their traffic demands from a source node to a destination node. We let $\mathcal{C} = \{1, \dots, K\}$ be the set of users and denote by $\lambda_u > 0$ the traffic demand of user $u \in \mathcal{C}$. We consider a *decentralized routing scheme* in which each user freely decides how to split its traffic demand over the N parallel links joining the source node and the destination node. We shall denote by $x_{u,j}$ the quantity of traffic sent by user u on link $j = 1, 2, \dots, N$.

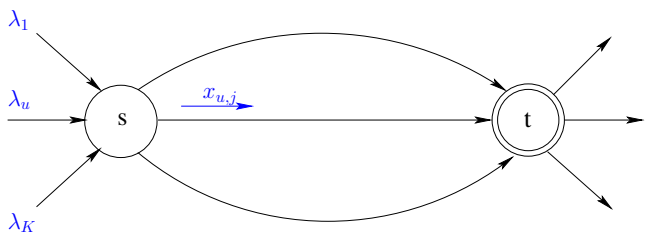


Figure 1: K users route their traffic demands over N parallel links.

The routing strategy of user u is represented by the vector $\mathbf{x}_u = (x_{u,j})_{j=1,\dots,N}$. We shall denote by $\mathcal{X}_u$ the set of feasible routing strategies for this user, that is, the set of vectors $\mathbf{x}_u \geq \mathbf{0}$ such that $\sum_j x_{u,j} = \lambda_u$. A strategy profile is then a vector $\mathbf{x} = (\mathbf{x}_u)_{u \in \mathcal{C}}$ belonging to the product strategy space $\mathcal{X} = \bigotimes_{u \in \mathcal{C}} \mathcal{X}_u$. It describes the routing strategy used by each player and represents in a way the state of the game. Given a state $\mathbf{x} \in \mathcal{X}$ of the game, we denote by $\mathbf{x}_{-u}$ the vector $(\mathbf{x}_1, \dots, \mathbf{x}_{u-1}, \mathbf{x}_{u+1}, \dots, \mathbf{x}_K)$ which describes the routing strategies of all other users than u.

We assume that the links are of two types. There are $n_1 \geq 1$ type-1 links and $n_2 = N - n_1 \geq 1$ type-2 links. In the following, we let $\mathcal{S}_1 = \{1, \dots, n_1\}$ be the set of type-1 links and $\mathcal{S}_2 = \{n_1 + 1, \dots, N\}$ be the set of type-2 links. A crucial assumption in our model is that the links have traffic-dependent cost functions. More precisely, it is assumed that the cost per unit flow on link $j \in \mathcal{S}_k$ is of the form $c_k \phi(y_j)$, where $y_j = \sum_{u \in \mathcal{C}} x_{u,j}$ represents the total traffic on the link, c_k is a cost parameter which depends on the type of the link, and ϕ is a given cost function. In this work, we shall assume that $c_1 < c_2$ and refer to type-1 links (resp. type-2 links) as "cheap" links (resp. "expensive" links).

Another key assumption is that users are self-interested agents seeking to minimize their own routing cost. More precisely, in state $\mathbf{x} = (\mathbf{x}_u, \mathbf{x}_{-u})$ of the game, user u chooses its routing strategy $\mathbf{x}_u^*$ so as to solve the following optimization problem

$$\underset{\mathbf{x}_u^* \in \mathcal{X}_u}{\text{minimize}}\ T_u(\mathbf{x}_u^*, \mathbf{x}_{-u}) = c_1 \sum_{j \in \mathcal{S}_1} x_{u,j}^* \, \phi\left(x_{u,j}^* + \sum_{i \neq u} x_{i,j} \right) + c_2 \sum_{j \in \mathcal{S}_2} x_{u,j}^* \, \phi\left(x_{u,j}^* + \sum_{i \neq u} x_{i,j} \right).$$

Note that the cost incurred by the routing agent u on a link j depends both on the amount of flow $x_{u,j}^*$ that it routes to that link, but also on the total traffic $\sum_{i \neq u} x_{i,j}$ sent by the other users on that link. Thus the optimal strategy for user u depends on the routing strategies of the other players, which gives

a non-cooperative routing game between the users. A Nash equilibrium (NE) of this game is a stable state $\mathbf{x}^{ne} \in \mathcal{X}$ of the game from which no user finds it beneficial to deviate unilaterally, that is

$$\mathbf{x}_u^{ne} \in \arg\min_{\mathbf{z}\in\mathcal{X}_u} T_u\left(\mathbf{z}, \mathbf{x}_{-u}^{ne}\right), \quad \forall u \in \mathcal{C}.$$

Throughout the paper, we shall only consider latency functions ϕ satisfying Assumptions 1-5 stated in Appendix A. These assumptions are in particular satisfied by the functions $\phi(x) = e^{\nu x}$ for $\nu > 0$ and $\phi(x) = (1+x)^m$ for $m \geq 2$ (see Remark 1 of Appendix A for other examples of latency functions satisfying these assumptions). Under Assumption 1, the existence and uniqueness of the NE follow from [24].

2.2 Inefficiency of the decentralized routing scheme

It is well-known in game theory that the outcome of a non-cooperative game between selfish players can be inferior to a centrally designed outcome (see, e.g. Chapter 17 of [23]). In our setting, it means that letting users optimize their own performances without any coordination may lead to a performance degradation with respect to a centralized routing scheme which would optimize the global performance of all users.

We adopt the point of view of a network operator who has a known *network configuration*. For the model introduced above, a network configuration is defined as a vector of parameters $\mathbf{p} = (\mathbf{n}, \mathbf{c})$, where $\mathbf{n} = (n_1, n_2)$ specifies the numbers of "cheap" and "expensive" links, whereas $\mathbf{c} = (c_1, c_2)$ gives the costs of the two types of links. The network operator does not know the traffic demands $\lambda_1, \ldots, \lambda_K$ of the users and does not control how they route their traffic demands into the network. We assume a decentralized scheme in which each user minimizes its own routing cost, without coordination with the others, as explained above. The objective is to evaluate, for the fixed network configuration $\mathbf{p}$, the performance degradation resulting from the absence of a central authority under the worst-case traffic conditions.

In order to make things more formal, we introduce below some definitions. Consider a network configuration $\mathbf{p}$ and a vector $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_K)$ of traffic demands. Let $\mathbf{x}^{ne}$ be the corresponding NE of the routing game and define $\mathbf{y}^{ne} = (y_1^{ne}, y_2^{ne}, \ldots, y_N^{ne})$, where $y_j^{ne} = \sum_{u\in\mathcal{C}} x_{u,j}^{ne}$ is the total flow on link j at the NE. The global performance (or social cost) of the *decentralized routing scheme* with K players is defined as the sum of the individual player' costs at the NE

$$D_K(\boldsymbol{\lambda}, \mathbf{p}) = \sum_{u\in\mathcal{C}} T_u(\mathbf{x}^{ne}).$$

Introducing $F_k(y) = c_k\, y\, \phi\,(y)$ for $k = 1, 2$, the social cost can be written as follows

$$\begin{aligned} D_K(\boldsymbol{\lambda}, \mathbf{p}) &= c_1 \sum_{j\in\mathcal{S}_1} \sum_{u\in\mathcal{C}} x_{u,j}^{ne}\, \phi\left(\sum_{u\in\mathcal{C}} x_{u,j}^{ne}\right) + c_2 \sum_{j\in\mathcal{S}_2} \sum_{u\in\mathcal{C}} x_{u,j}^{ne}\, \phi\left(\sum_{u\in\mathcal{C}} x_{u,j}^{ne}\right) \\ &= \sum_{j\in\mathcal{S}_1} F_1(y_j^{ne}) + \sum_{j\in\mathcal{S}_2} F_2(y_j^{ne}) \\ &= F(\mathbf{y}^{ne}), \end{aligned}$$

where $F(\mathbf{y}) = \sum_{j\in\mathcal{S}_1} F_1(y_j) + \sum_{j\in\mathcal{S}_2} F_2(y_j)$ for all $\mathbf{y} \geq \mathbf{0}$. In this paper, our objective is to compare the performance of the decentralized scheme with the optimal performance that could be achieved

$$F(\mathbf{y}^*) = \min\left\{ F(\mathbf{y}) \;:\; \mathbf{y} \geq \mathbf{0}, \sum_j y_j = \bar{\lambda} \right\}, \tag{1}$$

where $\bar{\lambda} = \sum_{u\in\mathcal{C}} \lambda_u$ is the total traffic in the system. Note that $F(\mathbf{y}^*)$ is the performance achieved by an optimal routing strategy minimizing the social cost, and that it corresponds to the cost obtained when there is a single user controlling all the traffic in the system, i.e. such that $\lambda_1 = \bar{\lambda}$. The global cost of this *centralized routing scheme* can be written as $D_1(\bar{\lambda}, \mathbf{p})$, and we thus have $F(\mathbf{y}^*) = D_1(\bar{\lambda}, \mathbf{p})$.

In order to compare the two routing schemes, we shall use the concept of *inefficiency* introduced in [16] in the context of server farms. The *inefficiency* of the decentralized scheme is defined as the ratio

between the performance obtained by the NE and the global optimal solution under the worst possible traffic conditions, that is

$$\textit{Inefficiency } I_K(\mathbf{p}) = \sup_{\boldsymbol{\lambda}} \frac{D_K(\boldsymbol{\lambda}, \mathbf{p})}{D_1(\bar{\lambda}, \mathbf{p})} = \sup_{\boldsymbol{\lambda}} \frac{F(\mathbf{y}^{ne})}{F(\mathbf{y}^*)}, \tag{2}$$

where the supremum is taken over all vectors $\boldsymbol{\lambda}$ of traffic demands that the network may have to accommodate and $\bar{\lambda} = \sum_{u=1}^{K} \lambda_u$. We emphasize that the inefficiency depends on the network configuration $\mathbf{p}$ but not on the traffic demands of the users. Its values are between 1 and ∞, a higher value indicating a worse performance of decentralized routing compared to centralized routing.

Another widely used measure of how the efficiency of a system degrades due to selfish behavior of its users is the so-called *Price of Anarchy* (PoA) [20], which, in our case, is related to the *inefficiency* as follows

$$PoA(K, N) = \sup_{\mathbf{p}} I_K(\mathbf{p}), \tag{3}$$

where the supremum is taken over all parameter vectors $\mathbf{p} = (\mathbf{n}, \mathbf{c})$ such that $n_1 + n_2 = N$ and $0 < c_1 < c_2$.

3 Preliminary results

We first recall in Section 3.1 some results regarding the worst-case traffic conditions for the performance of the decentralized routing scheme. We then provide simple relations for the characterization of the centralized and decentralized routing solutions in Sections 3.2 and 3.3, respectively.

3.1 Worst traffic conditions for a fixed total traffic

If the total incoming traffic $\bar{\lambda}$ is fixed, it is proven in [6] that the global cost $D_K(\boldsymbol{\lambda}, \mathbf{p})$ achieves its maximum for the symmetric game, that is

$$\sup_{\boldsymbol{\lambda}} D_K(\boldsymbol{\lambda}, \mathbf{p}) = D_K\left(\frac{\bar{\lambda}}{K}\mathbf{1}, \mathbf{p}\right), \tag{4}$$

where $\mathbf{1} = (1, 1, \ldots, 1)$, implying that

$$I_K(\mathbf{p}) = \sup_{\bar{\lambda}>0} \frac{D_K\left(\frac{\bar{\lambda}}{K}\mathbf{1}, \mathbf{p}\right)}{D_1(\bar{\lambda}, \mathbf{p})}. \tag{5}$$

As a consequence, for the calculation of the *inefficiency*, we can restrict ourselves to the symmetric game. This, coupled with the fact that in our setting the symmetric game is also a potential game, makes it more tractable for the analytic computation of the NE routing solution. More precisely, it is shown in [6] that the decentralized routing solution is the solution of a convex optimization problem, as stated in Theorem 1 below.

Theorem 1. *Let the vector* $\mathbf{y}$ *be the global optimum of the following convex optimization problem*

$$\begin{aligned} \underset{\mathbf{y}}{\textit{minimize}} \quad & \sum_{k=1}^{2} \sum_{j \in \mathcal{S}_k} F_k(y_j) + (K-1) \int_0^{y_j} c_k\, \phi(z) dz \\ \textit{s.t.} \quad & \\ & \textstyle\sum_{j=1}^{N} y_j = \bar{\lambda}, \\ & y_j \geq 0,\ j = 1, \ldots, N. \end{aligned}$$

The strategy profile $\mathbf{x}$ *such that* $x_{u,j} = \frac{y_j}{K}$ *for all* $u \in \mathcal{C}$ *and* $j = 1, \ldots, N$ *is the NE of the symmetric game.*

3.2 Characterization of the centralized routing strategy

Under the centralized routing strategy, the vector $\mathbf{y}^*$ of link flows is the optimal solution of problem (1). For a type-k link $j \in \mathcal{S}_k$, the KKT conditions then imply that $y_j^* > 0$ if and only if the marginal cost

$$F_k'(y_j^*) = c_k \left[\phi\left(y_j^*\right) + y_j^*\, \phi'\left(y_j^*\right)\right], \tag{6}$$

is minimal. Defining $\bar{\lambda}^*$ as the unique solution of $F_1'(\frac{\bar{\lambda}^*}{n_1}) = F_2'(0)$, that is,

$$c_1 \left[\phi\left(\frac{\bar{\lambda}^*}{n_1}\right) + \frac{\bar{\lambda}^*}{n_1}\, \phi'\left(\frac{\bar{\lambda}^*}{n_1}\right) \right] = c_2, \tag{7}$$

it follows from the assumption $c_1 < c_2$ that

- for $\bar{\lambda} \leq \bar{\lambda}^*$, the centralized routing strategy forwards all the traffic to the cheap links and they all receive the same amount of traffic, that is, $y_k^* = \frac{\bar{\lambda}}{n_1}$ for all $k \in \mathcal{S}_1$ and $y_j^* = 0$ for all $j \in \mathcal{S}_2$,
- for $\bar{\lambda} > \bar{\lambda}^*$, the centralized routing strategy is such that all links receive a positive amount of traffic (that is, $y_l^* > 0$ for all $l = 1, \ldots, N$), and

$$c_1 \left[\phi\left(y_k^*\right) + y_k^*\, \phi'\left(y_k^*\right)\right] = c_2 \left[\phi\left(y_j^*\right) + y_j^*\, \phi'\left(y_j^*\right)\right], \forall k \in \mathcal{S}_1,\ \forall j \in \mathcal{S}_2. \tag{8}$$

Note that in both cases two links of the same type receive exactly the same amount of traffic, that is, $y_l^* = y_m^*$ if $l, m \in \mathcal{S}_k$ for $k = 1, 2$. This implies that for $\bar{\lambda} \leq \bar{\lambda}^*$ the optimal social cost is simply $F(\mathbf{y}^*) = n_1 F_1\left(\frac{\bar{\lambda}}{n_1}\right)$, whereas for $\bar{\lambda} > \bar{\lambda}^*$, it can be written as $F(\mathbf{y}^*) = n_1 F_1\left(y_1^*\right) + n_2 F_2\left(y_N^*\right)$.

3.3 Characterization of the decentralized routing strategy

Under the decentralized routing strategy, the vector $\mathbf{y}^{ne}$ of link flows is the optimal solution of the optimization problem stated in Theorem 1. Defining $\bar{\lambda}^{ne}$ as the unique solution of

$$c_1 \left[K\phi\left(\frac{\bar{\lambda}^{ne}}{n_1}\right) + \frac{\bar{\lambda}^{ne}}{n_1}\, \phi'\left(\frac{\bar{\lambda}^{ne}}{n_1}\right) \right] = c_2 K, \tag{9}$$

it follows from the KKT conditions and the assumption $c_1 < c_2$ that

- for $\bar{\lambda} \leq \bar{\lambda}^{ne}$, the decentralized routing strategy forwards all the traffic to the cheap links only and they all receive the same amount of traffic, that is, $y_k^{ne} = \frac{\bar{\lambda}}{n_1}$ for all $k \in \mathcal{S}_1$ and $y_j^{ne} = 0$ for all $j \in \mathcal{S}_2$,
- for $\bar{\lambda} > \bar{\lambda}^{ne}$, in the decentralized routing strategy all links receive a positive amount of traffic (that is $y_l^{ne} > 0$ for all l), and

$$c_1 \left[K\phi\left(y_k^{ne}\right) + y_k^{ne}\, \phi'\left(y_k^{ne}\right)\right] = c_2 \left[K\phi\left(y_j^{ne}\right) + y_j^{ne}\, \phi'\left(y_j^{ne}\right)\right], \forall k \in \mathcal{S}_1,\ \forall j \in \mathcal{S}_2. \tag{10}$$

Note that, as in the centralized setting, links of the same type always receive the same amount of traffic. As a direct consequence, the social cost at the NE is $F(\mathbf{y}^{ne}) = n_1 F_1\left(\frac{\bar{\lambda}}{n_1}\right)$ for $\bar{\lambda} \leq \bar{\lambda}^{ne}$, whereas it can be written as $F(\mathbf{y}^{ne}) = n_1 F_1\left(y_1^{ne}\right) + n_2 F_2\left(y_N^{ne}\right)$ for $\bar{\lambda} > \bar{\lambda}^{ne}$.

It directly follows from (7) and (9) that $\bar{\lambda}^* < \bar{\lambda}^{ne}$ for $K > 1$. It means that when there are more than one user, the decentralized routing strategy uses only the "cheap" links longer than what would be optimal. Of course, when there is only one user, that is for $K = 1$, we have $\bar{\lambda}^{ne} = \bar{\lambda}^*$ and conditions (10) and (8) are equivalent in this case. In other words, the centralized routing strategy and the decentralized one coincide when there is only one user.

4 Worst-case total traffic

As discussed in Section 3.1, for a fixed total traffic $\bar{\lambda} > 0$, the worst inefficiency is obtained when all users control the same amount of traffic $\frac{\bar{\lambda}}{K}$, that is, for the symmetric game. We now study the worst-case total traffic $\bar{\lambda}$ for the ratio of social costs $D_K\left(\frac{\bar{\lambda}}{K}\mathbf{1}, \mathbf{p}\right) / D_1(\bar{\lambda}, \mathbf{p})$. As $D_K\left(\frac{\bar{\lambda}}{K}\mathbf{1}, \mathbf{p}\right) = F(\mathbf{y}^{ne})$ and $D_1(\bar{\lambda}, \mathbf{p}) = F(\mathbf{y}^*)$, we first establish some results pertaining to the comparison of equilibrium and optimal flow configurations in Section 4.1. In particular, we show that for any "cheap" link k the ratio y_k^{ne}/y_k^* reaches its maximum for $\bar{\lambda} = \bar{\lambda}^{ne}$. We then study in Section 4.2 the ratio of social costs as a function of $\bar{\lambda}$ and prove that it also achieves its maximum for $\bar{\lambda} = \bar{\lambda}^{ne}$.

4.1 Link flows under the centralized and decentralized routing strategies

Note that for $\bar{\lambda} > \bar{\lambda}^*$ (resp. $\bar{\lambda} > \bar{\lambda}^{ne}$) the link flows y_l^* (resp. y_l^{ne}) obtained under the centralized (resp. decentralized) routing scheme are implicitly defined by equation (8) (resp. (10)). We first prove in Lemma 1 below that under both routing schemes these link flows are continuous functions of $\bar{\lambda}$.

Lemma 1. *The vectors $\mathbf{y}^*$ and $\mathbf{y}^{ne}$ are continuous in $\bar{\lambda}$ over $[0, \infty)$.*

Proof. See Appendix B.1. □

Proposition 1 below prove some inequalities satisfied by the flows on cheap and expensive links, which are valid under both strategies. It is worthwhile noticing that the proof of this proposition exploits only the strict monotonicity and the convexity of the latency function ϕ.

Proposition 1. *For $K \geq 1$, it holds that*

(a) The flow on the expensive links is strictly lower than that on the cheap links, that is, $y_j^{ne} < y_k^{ne}$ for all $j \in \mathcal{S}_2$ and $k \in \mathcal{S}_1$,

(b) For any cheap link $k \in \mathcal{S}_1$ and any expensive link $j \in \mathcal{S}_2$, it holds that $c_1\phi\left(y_k^{ne}\right) < c_2\phi\left(y_j^{ne}\right)$ and $c_1 y_k^{ne}\phi'\left(y_k^{ne}\right) > c_2 y_j^{ne}\phi'\left(y_j^{ne}\right)$.

(c) For $\bar{\lambda} \geq \bar{\lambda}^{ne}$, the marginal cost of the cheap links is greater than or equal to that of the expensive ones, that is

$$F_1'(y_k^{ne}) \geq F_2'(y_j^{ne}), \quad \text{for all } k \in \mathcal{S}_1 \text{ and } j \in \mathcal{S}_2, \tag{11}$$

and the inequality is strict for $K > 1$.

Proof. See Appendix B.2. □

We emphasize that the properties stated in Proposition 1 also hold for $K = 1$, that is, for the centralized routing strategy. Hence, the centralized routing strategy forwards more traffic on the cheap links than on the expensive links, exactly as does the decentralized one. However, for $\bar{\lambda} \geq \bar{\lambda}^{ne}$, the marginal costs of cheap and expensive links are always equal under the centralized routing strategy, whereas the marginal costs of cheap links are strictly greater than those of expensive links under the decentralized routing strategy.

We now turn our attention to the comparison of the link flows obtained under both routing strategies when $\bar{\lambda}$ varies. When $\bar{\lambda}$ is in the interval $(0, \bar{\lambda}^*]$, both strategies coincide: they both forward all the traffic only on the cheap links. In the interval $(\bar{\lambda}^*, \bar{\lambda}^{ne}]$, the centralized strategy deviates a fraction of the traffic onto the expensive links, whereas the decentralized one keeps using only the cheap links. Finally, in the interval $(\bar{\lambda}^{ne}, \infty)$, both the centralized and decentralized strategies use both type of links. Proposition 2 below states our main result regarding the comparison of the equilibrium and optimal link flows when $\bar{\lambda}$ is in the latter interval. The proof relies on a technical condition which is easily shown to be satisfied by $\phi(x) = e^{\nu x}$ and $\phi(x) = (1 + x)^m$ (see Assumption 4 and Lemma 7 in Appendix A). More generally, all the results in this section are valid for any other latency function satisfying Assumptions 1, 3 and 4 in Appendix A, such as $\phi(x) = 1/(1 - x)^m$ and $\phi(x) = 1 + x^m$.

Proposition 2. *For $\bar{\lambda} > \bar{\lambda}^{ne}$, it holds that*

(a) The decentralized routing strategy forwards more (resp. less) traffic on cheap (resp. expensive) links than the centralized one does, that is

$$y_k^{ne} > y_k^* \text{ for all } k \in \mathcal{S}_1 \text{ and } y_j^{ne} < y_j^* \text{ for all } j \in \mathcal{S}_2. \tag{12}$$

(b) For any cheap link $k \in \mathcal{S}_1$, the difference $y_k^{ne} - y_k^*$ between the amount of flow forwarded on this link by the decentralized routing strategy and the centralized one decreases as $\bar{\lambda}$ increases, that is,

$$\frac{dy_k^{ne}}{d\bar{\lambda}} < \frac{dy_k^*}{d\bar{\lambda}}. \tag{13}$$

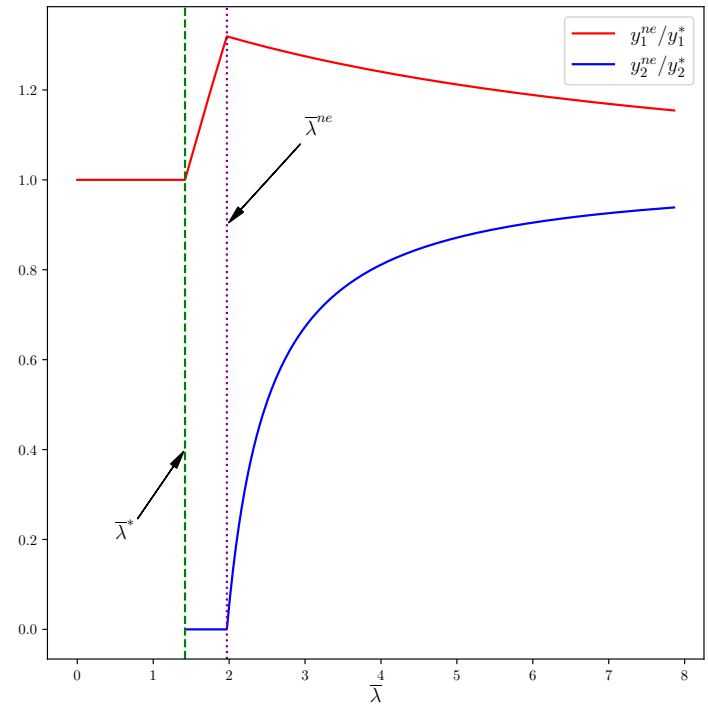


Figure 2: Evolution of the link flow ratios y_1^{ne}/y_1^* (for the cheap link) and y_2^{ne}/y_2^* (for the expensive links) as a function of $\bar{\lambda}$ for the latency function $\phi(x) = e^x$. In this example, there are $K = 5$ users, one cheap link with $c_1 = 1$ and nine expensive links with $c_2 = 10$.

(c) *For any expensive link $j \in \mathcal{S}_2$, the difference $y_j^{ne} - y_j^*$ between the amount of flow forwarded on this link by the decentralized routing strategy and the centralized one increases as $\bar{\lambda}$ increases, that is,*

$$\frac{dy_j^{ne}}{d\bar{\lambda}} > \frac{dy_j^*}{d\bar{\lambda}}. \tag{14}$$

Proof. See Appendix B.3. □

Using Proposition 1 and Proposition 2, Proposition 3 below characterizes the behaviour of the ratios y_l^{ne}/y_l^* of link flows obtained under both settings when the total traffic $\bar{\lambda}$ varies.

Proposition 3. *Let $k \in \mathcal{S}_1$ be an arbitrary cheap link and $j \in \mathcal{S}_2$ be an arbitrary expensive link and consider the ratios y_k^{ne}/y_k^* and y_j^{ne}/y_j^* as functions of the total traffic $\bar{\lambda}$. It holds that*

- *the ratio y_k^{ne}/y_k^* is strictly increasing in $\bar{\lambda}$ over the interval $\left(\bar{\lambda}^*, \bar{\lambda}^{ne}\right]$, and strictly decreasing over the interval $\left(\bar{\lambda}^{ne}, \infty\right)$,*
- *The ratio y_j^{ne}/y_j^* is 0 over the interval $\left(\bar{\lambda}^*, \bar{\lambda}^{ne}\right]$, and strictly increasing in $\bar{\lambda}$ over the interval $\left(\bar{\lambda}^{ne}, \infty\right)$.*

Proof. See Appendix B.4. □

It directly follows from Proposition 3 and Lemma 1 that for any cheap link k the maximum value of the ratio y_k^{ne}/y_k^* is obtained when $\bar{\lambda} = \bar{\lambda}^{ne}$, as formally stated in Corollary 1.

Corollary 1. *For all cheap links $k \in \mathcal{S}_1$, the ratio y_k^{ne}/y_k^* achieves its maximum for $\bar{\lambda} = \bar{\lambda}^{ne}$. At this point, the ratio y_j^{ne}/y_j^* is minimum for all expensive links $j \in \mathcal{S}_2$.*

To summarize, in the interval $(0, \bar{\lambda}^*]$, the ratio y_k^{ne}/y_k^* is constant and equal to 1 for any cheap link k. In the interval $(\bar{\lambda}^*, \bar{\lambda}^{ne}]$, the ratio y_k^{ne}/y_k^* increases as $\bar{\lambda}$ increases and it reaches its maximum for $\bar{\lambda} = \bar{\lambda}^{ne}$. From this point onwards, the ratio y_k^{ne}/y_k^* decreases with $\bar{\lambda}$. Similarly, for any expensive link

j, the ratio y_j^{ne}/y_j^* is 0 over the interval $(\bar{\lambda}^*, \bar{\lambda}^{ne}]$ (it is undefined for $\bar{\lambda} \leq \bar{\lambda}^*$) and it increases with $\bar{\lambda}$ over $(\bar{\lambda}^{ne}, \infty)$. Figure 2 illustrates this behaviour of the link flow ratios y_k^{ne}/y_k^* and y_j^{ne}/y_j^* for $\phi(x) = e^x$, $K = 5$ users and $N = 10$ parallel links. In this example, there is only one cheap link of cost $c_1 = 1$ and there are nine expensive links of cost $c_2 = 10$.

4.2 Worst-case total traffic for the ratio of social costs

We now study the ratio of social costs. It directly follows from Lemma 1 that this ratio is continuous in $\bar{\lambda}$.

Lemma 2. *As a function of $\bar{\lambda}$, the ratio $D_K\left(\frac{\bar{\lambda}}{K}\mathbf{1}, \mathbf{p}\right)/D_1(\bar{\lambda}, \mathbf{p})$ is continuous over $(0, \infty)$.*

Proof. See Appendix C.1. □

Proposition 4 below characterizes the behaviour of the ratio of social costs as $\bar{\lambda}$ varies over $(0, \infty)$.

Proposition 4. *For $K > 1$, as a function of $\bar{\lambda}$, the ratio $D_K\left(\frac{\bar{\lambda}}{K}\mathbf{1}, \mathbf{p}\right)/D_1(\bar{\lambda}, \mathbf{p})$ of the social costs obtained under the decentralized routing strategy and the centralized one is*

(a) constant and equal to 1 in the interval $(0, \bar{\lambda}^*]$,

(b) strictly increasing with $\bar{\lambda}$ in the interval $\left(\bar{\lambda}^*, \bar{\lambda}^{ne}\right]$, and,

(c) strictly decreasing with $\bar{\lambda}$ in the interval $\left(\bar{\lambda}^{ne}, \infty\right)$.

Proof. See Appendix C. □

We have shown in Section 4.1 that for any cheap link $k \in \mathcal{S}_1$, the ratio y_k^{ne}/y_k^* of the flows on this link obtained under the decentralized routing strategy and the optimal one achieves its maximum for $\bar{\lambda} = \bar{\lambda}^{ne}$. As stated in Corollary 2, the same is true for the ratio of social costs.

Corollary 2. *The ratio $D_K\left(\frac{\bar{\lambda}}{K}\mathbf{1}, \mathbf{p}\right)/D_1(\bar{\lambda}, \mathbf{p})$ achieves its maximum when $\bar{\lambda} = \bar{\lambda}^{ne}$, implying that*

$$I_K(\mathbf{p}) = \frac{n_1 F_1\left(\frac{\bar{\lambda}^{ne}}{n_1}\right)}{n_1 F_1\left(y_1^*\right) + n_2 F_2\left(y_N^*\right)}, \tag{15}$$

where y_1^* and y_N^* are the links flows over cheap and expensive links, respectively, obtained under the centralized routing strategy for $\bar{\lambda} = \bar{\lambda}^{ne}$.

Proof. The proof directly follows from Lemma 2 and Proposition 4. □

Figure 3 illustrates the evolution of the ratio of social costs as the total traffic $\bar{\lambda}$ in the system varies for the latency functions $\phi(x) = e^x$, $\phi(x) = (1+x)^3$ and $\phi(x) = (1+x)^4$. The setting is the same as in Figure 2, that is, there are $K = 5$ users, one cheap link with $c_1 = 1$ and nine expensive links with $c_2 = 10$. For $\phi(x) = e^x$, the ratio is constant before $\bar{\lambda}^* = 1.42$, it increases from $\bar{\lambda}^*$ up to $\bar{\lambda}^{ne} = 1.97$ where it reaches a maximum value of 1.21, and then decreases with $\bar{\lambda}$. Similarly, for $\phi(x) = (1+x)^4$, the ratio of social costs is constant before $\bar{\lambda}^* = 0.45$. From this point onwards it increases up to $\bar{\lambda}^{ne} = 0.66$ where it reaches a maximum value of 1.18, and then decreases with $\bar{\lambda}$.

In summary, we have shown that, given a fixed network configuration $\mathbf{p}$, the worst inefficiency of the decentralized routing scheme is obtained when all users control the same amount of traffic and when the total traffic in the system is $\bar{\lambda} = \bar{\lambda}^{ne}$. This corresponds to the value of the total traffic for which the decentralized routing scheme starts using the expensive links. When $K > 1$, this value is strictly greater than $\bar{\lambda}^*$, which means that selfish users send all their traffic demands on the cheap links longer than what would be globally optimal. Although this result is proven here only for latency functions satisfying Assumptions 1-5, we note that a similar result was proven in [16] for the M/M/1 latency function, that is, for $\phi(x) = 1/(1-x)$. As discussed in Section 6, numerical experiments suggest that Corollary 2 probably holds for a much broader class of latency functions.

We would like to remark that the result of Corollary 2 is consistent with the work of [11], in which the authors show that, for nonatomic routing games and affine costs, the local maxima of the ratio of social costs are obtained when the total traffic is such that a new set of links is used. Another direct consequence of Corollary 2 is that, given a fixed network configuration, the worst inefficiency of the

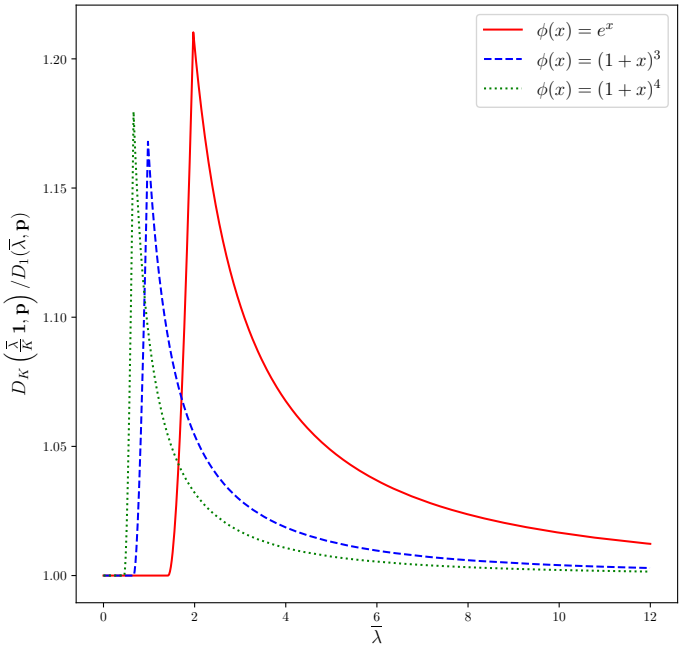


Figure 3: Evolution of the ratio $D_K\left(\frac{\bar{\lambda}}{K}\mathbf{1}, \mathbf{p}\right)/D_1(\bar{\lambda}, \mathbf{p})$ as $\bar{\lambda}$ increases for the latency functions $\phi(x) = e^x$, $\phi(x) = (1+x)^3$ and $\phi(x) = (1+x)^4$. The setting is the same as in Figure 2.

decentralized routing scheme is obtained neither for $\bar{\lambda} \to 0$ nor for $\bar{\lambda} \to \infty$, as could be expected. Interestingly, this result is consistent with the work of [8] where the authors show that the PoA of nonatomic routing games is not achieved when the traffic is very small or very large. Nevertheless, for $\phi(x) = (1+x)^m$, the value of $\bar{\lambda}^{ne}$ can be made arbitrarily small, as shown in Lemma 3 below.

Lemma 3. *For $\phi(x) = (1+x)^m$, $\bar{\lambda}^{ne}$ decreases with m and $\bar{\lambda}^{ne} \to 0$ as $m \to \infty$.*

Proof. See Appendix C.3. □

5 From Inefficiency to Price of Anarchy

In Section 4, we have characterized the worst traffic conditions for the inefficiency $I_K(\mathbf{p})$ of the decentralized routing scheme, assuming a fixed network configuration $\mathbf{p}$. In this section, we assume that the worst traffic conditions are met and investigate the worst network configuration for the inefficiency of the decentralized routing scheme. In other words, we study the PoA for the routing game over parallel links, where in (3), the PoA was defined as the supremum over all the network configurations $\mathbf{p}$ of $I_K(\mathbf{p})$. We first show that the *Inefficiency* depends only on the ratios n_1/n_2 and c_1/c_2.

Proposition 5. *The Inefficiency $I_K(\mathbf{p})$ depends on the parameters $\mathbf{p} = (\mathbf{n}, \mathbf{c})$ only through the ratios $\alpha = \frac{n_1}{n_2}$ and $\gamma = \frac{c_1}{c_2}$.*

Proof. See Appendix D. □

Note in particular that the *Inefficiency* depends on the total number of links only through the values that the ratio $\alpha = \frac{n_1}{n_2}$ can take. As a consequence of Proposition 5, we shall write $I_K(\alpha, \gamma)$ instead of $I_K(\mathbf{n}, \mathbf{c})$ to denote the *Inefficiency* in the following. We study below how the *Inefficiency* varies with α and γ.

We first study how $I_K(\alpha, \gamma)$ varies with α for a fixed value of γ. Assuming that $K = 5$ and $N = 100$, Figure 4 shows $I_K(\alpha, \gamma)$ as a function of n_1/N for $\gamma = 0.5$, $\gamma = 0.2$ and $\gamma = 0.1$ when $\phi(x) = (1+x)^3$. We observe that the inefficiency of the decentralized routing scheme seems to decrease as the proportion of cheap links increases. A similar behavior was observed for $\phi(x) = e^x$. As $n_1/N = \alpha/(1+\alpha)$ is an increasing function of α, this suggests that the inefficiency decreases with the ratio α of the numbers of cheap and expensive links. This is formally proven in Proposition 6 below.

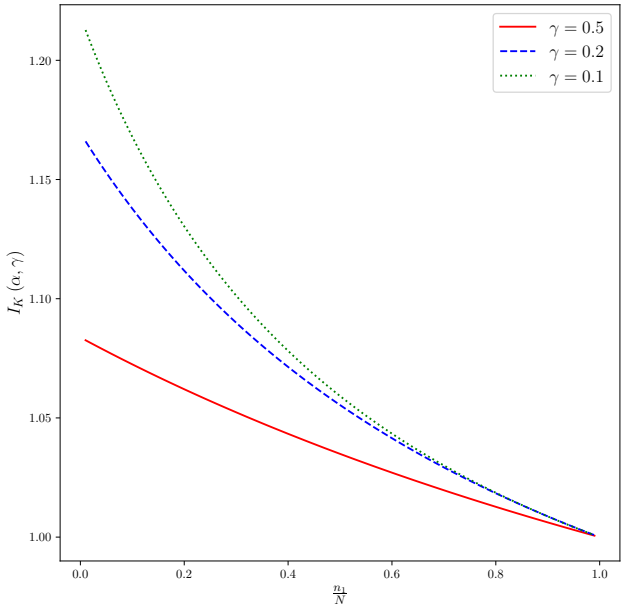


Figure 4: Inefficiency $I_K(\alpha, \gamma)$ as a function of $\frac{n_1}{N}$ for different values of γ and for $\phi(x) = (1+x)^3$.

Proposition 6. *The Inefficiency $I_K(\alpha, \gamma)$ is strictly decreasing with α.*

Proof. See Appendix D. □

An immediate consequence is the following corollary.

Corollary 3. *The PoA is obtained when there is only one "cheap" link and $N-1$ " expensive" links, that is,*

$$PoA(K, N) = \sup_{\alpha,\gamma} I_K(\alpha, \gamma) = \sup_{\gamma} I_K\left(\frac{1}{N-1}, \gamma\right). \tag{16}$$

In the following, we shall therefore assume that $\alpha = 1/(N-1)$ and study how the *Inefficiency* varies as a function of γ. Assuming that $\phi(x) = e^x$, Figure 5a shows the values obtained for $I_K\left(\frac{1}{N-1}, \gamma\right)$ as γ varies from 0 to 1 in scenarios with $K = 2$, $K = 3$ and $K = 5$ users and $N = 10$ parallel links. We observe that for all values of K the *Inefficiency* is strictly decreasing with γ, which, according to (16), implies that the PoA is obtained when γ tends to zero. As a result, we conjecture that for $\phi(x) = e^{\nu x}$,

$$PoA(K, N) = \lim_{\gamma \to 0} I_K\left(\frac{1}{N-1}, \gamma\right).$$

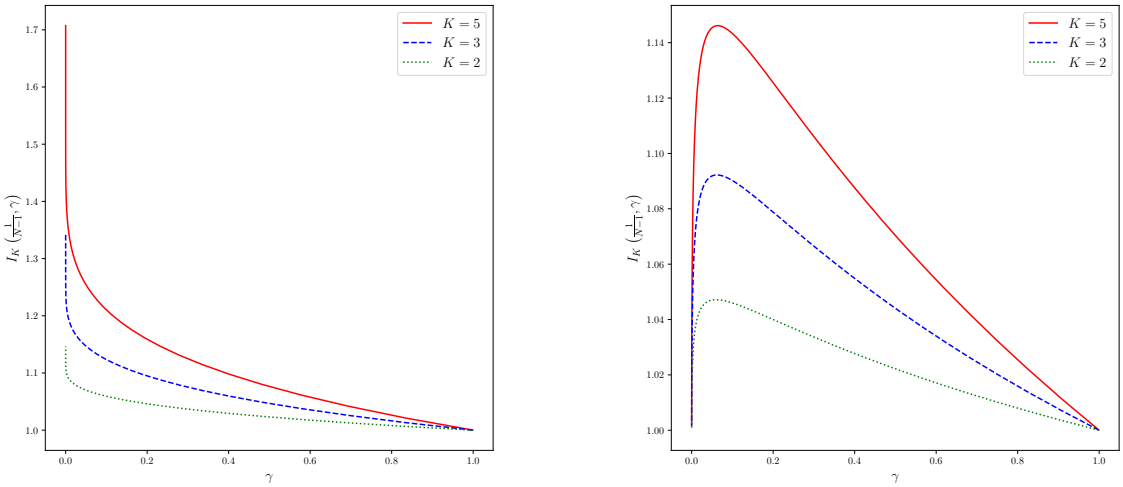
Besides, we get the following bounds on the performance degradation.

Lemma 4. *For $\phi(x) = e^{\nu x}$ and $N \geq 2$, it holds that*

$$K^{\frac{N-1-\nu}{N-1}} \frac{N-1}{N(1+\log K)-1} \leq \lim_{\gamma \to 0} I_K\left(\frac{1}{N-1}, \gamma\right) \leq K,$$

from which it follows that $K/(1+\log(K)) \leq \lim_{N\to\infty} PoA(K, N)$.

Proof. See Appendix D. □



(a) $\phi(x) = e^x$.

(b) $\phi(x) = (1+x)^2$.

Figure 5: Inefficiency $I_K(\frac{1}{N-1}, \gamma)$ as a function of γ for different values of K and for (a) $\phi(x) = e^x$ and (b) $\phi(x) = (1+x)^2$.

In words, we conjecture that for $\phi(x) = e^{\nu x}$ the worst inefficiency is achieved when the cost of the "expensive" links is infinitely larger than the cost of the "cheap" link. If this conjecture holds, this implies that the PoA is upper bounded by K. Furthermore, regardless this conjecture is true or not, the above result implies that the PoA for $\phi(x) = e^{\nu x}$ is, at least, $K^{\frac{N-1-\nu}{N-1}}(N-1)/[N(1+\log K)-1]$. For $\nu = 1$, $N = 10$ and $K = 5$, it yields 1.499, which is to be compared to the value 1.7 obtained in Figure 5a. Moreover, we can conclude that the PoA for $\phi(x) = e^{\nu x}$ is unbounded in nonatomic routing games as $K/(1+\log(K))$ tends to ∞ when $K \to \infty$. Surprisingly, the monotonicity property shown in Figure 5a for $\phi(x) = e^{\nu x}$ does not seem to hold for $\phi(x) = (1+x)^m$. As illustrated for $\phi(x) = (1+x)^2$ in Figure 5b, in which we also assume that $N = 10$, the inefficiency $I_K(\frac{1}{N-1}, \gamma)$ obtained for different values of K is not monotone as a function of γ. A similar behaviour was observed for other values of m. Unfortunately, we were not able to characterize the value of γ yielding the worst inefficiency. Therefore, the precise value of γ for which the PoA is achieved when $\phi(x) = (1+x)^m$ remains an open question.

6 Extensions of this work

Our main result is that the worst inefficiency of the decentralized routing scheme is obtained when the traffic demands of all users are $\bar{\lambda}^{ne}/K$. The key ingredient to prove this result is Proposition 4, which characterizes how the ratio of equilibrium and optimal social costs varies with the total traffic demand $\bar{\lambda}$. Proposition 4 has been established under sufficient conditions on the latency function $\phi(x)$, assuming that there are only two types of links with the same latency function $\phi(x)$. We discuss below several interesting extensions of this work.

6.1 Generalization to other latency functions

As already mentioned, it was proven in [16] that for $\phi(x) = 1/(1-x)$ the ratio of social costs varies with $\bar{\lambda}$ exactly as stated in Proposition 4 for latency functions satisfying Assumptions 1-5 such as $\phi(x) = e^{\nu x}$ and $\phi(x) = (1+x)^m$. Our numeric experiments suggest however that Proposition 4 holds for a much broader class of latency functions.

For instance, in Figure 6 we plot the evolution of the ratio of social costs as $\bar{\lambda}$ increases when $\phi(x) = 1 + x^m$ for different values of m. A similar behavior of the ratio of social costs is obtained for $\phi(x) = 1/(1-x)^m$. For both types of latency function, we observe that the ratio of social costs varies with $\bar{\lambda}$ as stated in Proposition 4, and that the worst inefficiency is obtained when $\bar{\lambda} = \bar{\lambda}^{ne}$.

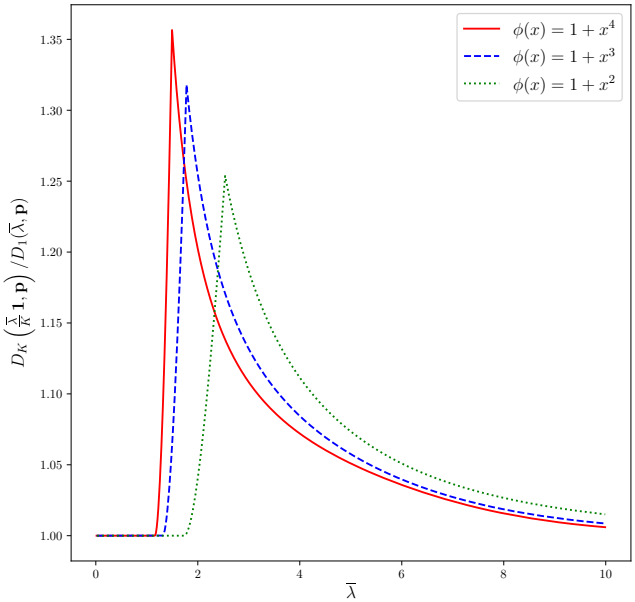


Figure 6: Evolution of the ratio of social costs as $\bar{\lambda}$ increases for the latency function $\phi(x) = 1 + x^m$. In this example, there are $K = 5$ users, one cheap link with $c_1 = 1$ and nine expensive links with $c_2 = 10$.

Unfortunately, the arguments used to prove Proposition 4 do not readily apply to these latency functions, as we now briefly explain:

- For $\phi(x) = 1 + x^m$, with $m \geq 1$, it is straightforward to show that Lemma 8 holds, and thus that the ratio of social costs is strictly decreasing with $\bar{\lambda}$ over the interval $\left(\bar{\lambda}^{ne}, +\infty\right)$. However, the argument used to prove that this ratio is strictly increasing with $\bar{\lambda}$ over $\left(\bar{\lambda}^*, \bar{\lambda}^{ne}\right)$ does not apply because the latency function $\phi(x) = 1 + x^m$ does not satisfy Assumption 2 in Appendix A. Therefore, a different argument should be used to prove that the ratio of social costs increases over $\left(\bar{\lambda}^*, \bar{\lambda}^{ne}\right)$.
- For $\phi = \frac{1}{(1-x)^m}$ with $m \geq 2$, the same approach as for $\phi(x) = e^{\nu x}$ and $\phi(x) = (1+x)^m$ can be used to show that the ratio of social costs is strictly increasing with $\bar{\lambda}$ over $\left(\bar{\lambda}^*, \bar{\lambda}^{ne}\right)$. Unfortunately, our numerical experiments suggest that Assumption 5 in Appendix A is not met by $\phi(x) = 1/(1-x)^m$, implying that a different approach should be used to prove that the ratio of social costs is strictly decreasing with $\bar{\lambda}$ for $\bar{\lambda} > \bar{\lambda}^{ne}$.

6.2 Extension to more than two types of links

Our results on the inefficiency of selfish routing have been established assuming that there are only two types of links which differ by their cost parameters c_1 and $c_2 > c_1$. Numeric experiments suggest however that similar results hold for more than two types of links.

In Figure 7 we plot the ratio of social costs obtained in the symmetric game for $\phi(x) = e^x$ and $\phi(x) = (1 + x)^2$ as a function of the total traffic demand $\bar{\lambda}$ when there are 4 links, each of a different type. It can be observed that as $\bar{\lambda}$ increases the ratio goes through peaks and valleys, and finally converges towards 1 as $\bar{\lambda} \to \infty$. The peaks correspond to values of the total traffic demand $\bar{\lambda}$ at which the decentralized routing scheme starts using one new link (these values are shown with dotted vertical lines in Figure 7). A similar behavior of the ratio of social costs was observed in [16] for the latency function $\phi(x) = 1/(1 - x)$, and in nonatomic routing games with affine costs [11].

The analysis is nevertheless much more complex for more than two types of links as one needs to compare multiple local maxima to determine the *Inefficiency*. The extension of our results to more than two types of links is therefore left for future work.

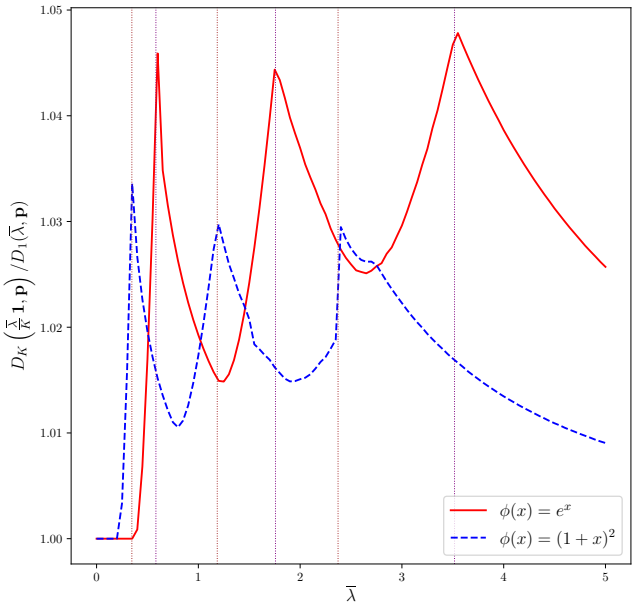


Figure 7: Evolution of the ratio of social costs as a function of the total traffic demand for $\phi(x) = e^x$ and $\phi(x) = (1+x)^2$ when there are $K = 5$ users and $N = 4$ links, each of a different type. The cost parameters of the links are as follows: $c_1 = 1$, $c_2 = 2$, $c_3 = 4$ and $c_4 = 8$.

6.3 Extension to heterogeneous latency functions

It was assumed throughout this paper that the two types of links share the same latency function $\phi(x)$ and differ only through their cost parameters c_1 and c_2. A natural extension of this work would be to investigate the inefficiency of selfish routing for heterogeneous latency functions, that is, when the cost function of type-i links is $c_i\,\phi_i(x)$ (e.g., $c_1 e^x$ for type-1 links and $c_2(1+x)^m$ for type-2 links). It is known that even in this case the Nash Equilibrium exists and is unique under mild assumptions on the latency functions $\phi_i(x)$ [24]. It is not clear however whether the global cost $D_K(\boldsymbol{\lambda}, \mathbf{p})$ achieves its maximum for the symmetric game $\boldsymbol{\lambda} = \left(\frac{\bar{\lambda}}{K}, \ldots, \frac{\bar{\lambda}}{K}\right)$ in this case. The proof of this result in [6] relies on a monotonic property regarding the order of preference of links as seen by each user (see Proposition 2 and Lemma 1 in [6]). Proving that this monotonic property is still valid in the case of heterogeneous latency functions is highly non-trivial. As a consequence, it is not clear whether the worst inefficiency of the decentralized routing scheme is obtained for a symmetric game, which is crucial to characterize the decentralized routing strategy as the solution of a convex optimization problem (see Theorem 1). This extension is therefore also left for future work.

7 Conclusions and future work

For the specific atomic routing game considered in this paper, it was shown that the worst traffic conditions occur when all users have the same traffic demand and the total traffic demand is such that "expensive links" are marginally used by selfish routing. Moreover, if these worst traffic conditions are met, the worst inefficiency of the selfish routing scheme is obtained when the number of "expensive" links is infinitely larger than the number of "cheap" links and under a very specific condition on the ratio c_1/c_2 (which we conjecture to be $c_1/c_2 \to 0$ for $\phi(x) = e^{\nu x}$). The worst-case scenarios for the inefficiency of selfish routing therefore corresponds to very specific traffic conditions and to highly asymmetric network configurations, which explain why the PoA is probably an overly pessimistic performance measure, as advocated in many recent works on non-atomic routing games.

As future work, we plan, as discussed in Section 6, to investigate the extension of our results to other latency functions, to more than two types of links and to heterogeneous latency functions for the different

types of links.

References

[1] E. Altman, U. Ayesta, and B. J. Prabhu. Load balancing in processor sharing systems. *Telecommunication Systems*, 47:35–48, 2011.

[2] J. Anselmi and B. Gaujal. Optimal routing in parallel, non-observable queues and the price of anarchy revisited. In *22nd International Teletraffic Congress (ITC)*, Amsterdam, 2010.

[3] U. Ayesta, O. Brun, and B. J. Prabhu. Price of anarchy in non-cooperative load balancing games. *Performance Evaluation*, 68(12):1312–1332, 2011.

[4] C. H. Bell and S. Stidham. Individual versus social optimization in the allocation of customers to alternative servers. *Management Science*, 29:831–839, 1983.

[5] D. Braess. Über ein paradoxon aus der verkehrsplanung. *Unternehmensforschung*, 12(1):258–268, 1968.

[6] O. Brun and B. Prabhu. Worst-case analysis of non-cooperative load balancing. *Annals of Operations Research*, 239(2):471–495, Apr 2016.

[7] H. L. Chen, J. Marden, and A. Wierman. The effect of local scheduling in load balancing designs. In *Proceedings of IEEE Infocom*, 2009.

[8] R. Colini-Baldeschi, R. Cominetti, P. Mertikopoulos, and M. Scarsini. When is selfish routing bad? the price of anarchy in light and heavy traffic. *Operations Research*, 68(2):411–434, 2020.

[9] R. Colini-Baldeschi, R. Cominetti, and M. Scarsini. On the price of anarchy of highly congested nonatomic network games. In *International Symposium on Algorithmic Game Theory*, pages 117–128. Springer, 2016.

[10] R. Colini-Baldeschi, R. Cominetti, and M. Scarsini. Price of anarchy for highly congested routing games in parallel networks. *Theory of Computing Systems*, 63(1):90–113, 2019.

[11] R. Cominetti, V. Dose, and M. Scarsini. The price of anarchy in routing games as a function of the demand. *Mathematical Programming*, pages 1–28, 2021.

[12] J. R. Correa, A. S. Schulz, and N. E. Stier-Moses. Selfish routing in capacitated networks. *Mathematics of Operations Research*, 29(4):961–976, 2004.

[13] J. R. Correa, A. S. Schulz, and N. E. Stier-Moses. Fast, fair, and efficient flows in networks. *Operations Research*, 55(2):215–225, 2007.

[14] J. R. Correa, A. S. Schulz, and N. E. Stier-Moses. A geometric approach to the price of anarchy in nonatomic congestion games. *Games and Economic Behavior*, 64(2):457–469, 2008.

[15] A. Czumaj, P. Krysta, and B. Vocking. Selfish traffic allocation for server farms. In *Proceedings of STOC*, 2002.

[16] J. Doncel, U. Ayesta, O. Brun, and B. Prabhu. Is price of anarchy the right measure for load-balancing games? *ACM Transactions on Internet Technology*, 14(2–3), Sep. 2014.

[17] S. Ghosh and R. Hassin. Inefficiency in stochastic queueing systems with strategic customers. *European Journal of Operational Research*, 295(1):1–11, 2021.

[18] Y. H, G. MT, and J. H. Price of anarchy in transportation networks: efficiency and optimality control. *Phys Review Letter*, 101(12), September 2008.

[19] M. Haviv and T. Roughgarden. The price of anarchy in an exponential multi-server. *Operations Research Letters*, 35:421–426, 2007.

[20] E. Koutsoupias and C. H. Papadimitriou. Worst-case equilibria. In *STACS 1999*, 1999.

[21] B. Monnot, F. Benita, and G. Piliouras. Routing games in the wild: Efficiency, equilibration and regret. In *International Conference on Web and Internet Economics*, pages 340–353. Springer, 2017.

[22] J. Nash. Non-cooperative games. *Annals of mathematics*, pages 286–295, 1951.

[23] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, editors. *Algorithmic Game Theory*. Cambridge University Press, 2007.

[24] A. Orda, R. Rom, and N. Shimkin. Competitive routing in multi-user communication networks. *IEEE/ACM Transactions on Networking*, 1:510–521, October 1993.

[25] A. Pigou. *The economics of welfare*. Macmillan, London, 1920.

[26] T. Roughgarden. The price of anarchy is independent of the network topology. In *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, pages 428–437, 2002.

[27] T. Roughgarden and É. Tardos. How bad is selfish routing? *Journal of the ACM (JACM)*, 49(2):236–259, 2002.

[28] T. Roughgarden and É. Tardos. Bounding the inefficiency of equilibria in nonatomic congestion games. *Games and economic behavior*, 47(2):389–403, 2004.

[29] S. Suri, C. D. Tóth, and Y. Zhou. Selfish load balancing and atomic congestion games. In *Proceedings of the Sixteenth Annual ACM Symposium on Parallelism in Algorithms and Architectures*, SPAA '04, pages 188–195, New York, NY, USA, 2004. Association for Computing Machinery.

[30] United States. Bureau of Public Roads. *Traffic Assignment Manual for Application with a Large, High Speed Computer*, volume 2. U.S. Department of Commerce, Bureau of Public Roads, Office of Planning, Urban Planning Division, 1964.

[31] J. G. Wardrop. Road paper. some theoretical aspects of road traffic research. *Proceedings of the institution of civil engineers*, 1(3):325–362, 1952.

[32] Z. Wu, R. H. Möhring, Y. Chen, and D. Xu. Selfishness need not be bad. *Operations Research*, 69(2):410–435, 2021.

A Some Properties of Latency Functions

As mentioned before, in this work we consider the following latency functions: $\phi(x) = e^{\nu x}$ for $\nu > 0$ and $\phi(x) = (1 + x)^m$ for $m \geq 2$. However, in order to keep the discussion as general as possible and ease the proofs for other latency functions, we shall regroup here all the technical assumptions on the latency function ϕ that are required for our results to hold, and systematically indicate in which proof an assumption is used. We also prove in this appendix that all the assumptions hold for the latency functions $\phi(x) = e^{\nu x}$ and $\phi(x) = (1+x)^m$. We emphasize however that all the results proven in Appendix B, C and D hold for any other latency function ϕ satisfying Assumption 1-5 below. Table 1 summarizes which assumption is required for which proof and shows whether or not an assumption is satisfied by one of the latency functions considered in the present paper.

Remark 1. *One can also easily check that Assumptions A1-A3 are also true for $\phi(x) = (1 + x^m)e^{\nu x}$ and $\phi(x) = (1 + x)^m e^{\nu x}$, where $\nu > 0$ and $m \geq 2$. Our numerical experiments suggest that these latency functions also verify Assumptions A4 and A5.*

Our basic assumption on the latency function ϕ is formally stated in Assumption 1 below.

Assumption 1. *The latency function ϕ is a continuously differentiable, strictly increasing and convex function over the interval $[0, +\infty)$ which verifies that $\phi(0) = 1$ and $\lim_{x\to+\infty} \phi(x) = +\infty$. In addition, its second derivative $\phi''()$ exists at all points in the interval $[0, +\infty)$.*

As discussed in [6], any latency function ϕ for which Assumption 1 is satisfied is such that the conditions given in [24] for the existence of a unique NE for the routing game are satisfied. It is clear that these conditions are satisfied both for $\phi(x) = e^{\nu x}$ and $\phi(x) = (1 + x)^m$.

To prove assertion (b) of Proposition 4, that is, that the ratio of social costs is strictly increasing over $(\bar{\lambda}^*, \bar{\lambda}^{ne}]$, we will need the following assumption.

Assumption	Required for	$\phi(x) = e^{\nu x}$	$\phi(x) = (1+x)^m$	$\phi(x) = \frac{1}{(1-x)^m}$	$\phi(x) = 1 + x^m$
Assumption 1	Existence of a unique NE and Theorem 1	✓	✓	✓	✓
Assumption 2	Assertion (b) of Proposition 4	✓	✓	✓	✗
Assumption 3	Proposition 3	✓	✓	✓	✓
Assumption 4	Proposition 2	✓	✓	✓	✓
Assumption 5	Proposition 4	✓	✓	✗	✓

Table 1: This table shows which assumptions are satisfied or not by some classic latency functions and indicates where they are used.

Assumption 2. *The function $g : [0, +\infty) \to [0, +\infty)$ defined by*

$$g(y) = \frac{F_1(y)}{F_1'(y)} = \frac{F_2(y)}{F_2'(y)} = \frac{y\,\phi(y)}{\phi(y) + y\phi'(y)} \tag{17}$$

is strictly increasing and strictly concave on $[0, +\infty)$.

As stated in Lemma 5 below, Assumption 2 is satisfied for $\phi(x) = e^{\nu x}$ and $\phi(x) = (1+x)^m$.

Lemma 5. *For $\phi(x) = e^{\nu x}$ and $\phi(x) = (1+x)^m$, Assumption 2 holds.*

Proof. It is enough to check that in both cases we have $g'(y) > 0$ and $g''(y) < 0$ for all $y \in [0, \infty)$. For $\phi(x) = e^{\nu x}$, we have $g(y) = \frac{y}{1+\nu\, y}$, and therefore $g'(y) = \frac{1}{(1+\nu y)^2} > 0$ and $g''(y) = -\frac{2\nu}{(1+\nu y)^3} < 0$ for all $y \in [0, \infty)$. For $\phi(x) = (1+x)^m$, we obtain $g(y) = \frac{y+y^2}{1+(m+1)y}$, which yields $g'(y) = \frac{1+2y+(m+1)y^2}{(1+(m+1)y)^2} > 0$ and $g''(y) = -\frac{2m}{(1+(m+1)y)^3} < 0$ for all $y \in [0, \infty)$. □

The third assumption that we shall require is stated below.

Assumption 3. *The function $f : [0, +\infty) \to [0, +\infty)$ defined by*

$$f(y) = \frac{F_1''(y)}{F_1'(y)} y = \frac{F_2''(y)}{F_2'(y)} y = \frac{2\phi'(y) + y\phi''(y)}{\phi(y) + y\phi'(y)}\, y \tag{18}$$

is strictly increasing over $[0, +\infty)$.

Assumption 3 is required in the proof of Lemma 22 to show that the ratio y_k^{ne}/y_k^* increases as $\bar{\lambda}$ increases in the interval $(\bar{\lambda}^*, \bar{\lambda}^{ne}]$ for all cheap links k. Lemma 22 is in turn used to prove Proposition 3. Lemma 6 below establishes that Assumption 3 is satisfied for $\phi(x) = e^{\nu x}$ and $\phi(x) = (1+x)^m$.

Lemma 6. *For $\phi(x) = e^{\nu x}$ and $\phi(x) = (1+x)^m$, Assumption 3 holds.*

Proof. For $\phi(x) = e^{\nu x}$, we have $f(y) = \nu y + \frac{\nu y}{1+\nu y}$ from which it follows that $f'(y) = \nu + \frac{\nu}{(1+\nu y)^2} > 0$. For $\phi(x) = (1+x)^m$, $f(y) = \frac{my}{1+y}\left(1 + \frac{1}{1+(m+1)y}\right)$, from which we obtain

$$f'(y) = \frac{2m}{(1+y)^2(1+(m+1)y)^2}\left[1 + (m+1)y + \frac{m(m+1)}{2}y^2\right] > 0.$$

□

In order to prove that $\frac{dy_k^{ne}}{d\bar{\lambda}} < \frac{dy_k^*}{d\bar{\lambda}}$ for $\bar{\lambda} \geq \bar{\lambda}^{ne}$ and any cheap link $k \in \mathcal{S}_1$ (see Proposition 2), we will need Assumption 4 below.

Assumption 4. *Let*

$$H(x, y) = x\phi'(x) - \phi(x)\left(1 + y\frac{\phi''(y)}{\phi'(y)}\right), \tag{19}$$

for $x, y \in [0, +\infty)$. Then, it is assumed that

- *the function* $A(x,y) = [c_1 H(x,y) - c_2 H(y,y)]/\phi'(x)$ *is decreasing in* x *over* $\left[y, \phi^{-1}(\frac{c_2}{c_1}\phi(y))\right)$ *for* y *fixed, and increasing in* y *over* $\left[\phi^{-1}(\frac{c_1}{c_2}\phi(x)), x\right)$ *for* x *fixed, and at least one of the monotonicities is strict;*
- *the function* $B(x,y) = [c_1 H(x,x) - c_2 H(y,x)]/\phi'(y)$ *is increasing in* x *over* $\left[y, \phi^{-1}(\frac{c_2}{c_1}\phi(y))\right)$ *for* y *fixed, and decreasing in* y *over* $\left[\phi^{-1}(\frac{c_1}{c_2}\phi(x)), x\right)$ *for* x *fixed, and at least one of the monotonicities is strict.*

Lemma 7 proves that this assumption is satisfied by the latency functions $\phi(x) = e^{\nu x}$ and $\phi(x) = (1+x)^m$.

Lemma 7. *For* $\phi(x) = e^{\nu x}$ *and* $\phi(x) = (1+x)^m$*, Assumption 4 holds.*

Proof. We first give the proof for $\phi(x) = e^{\nu x}$. Define $a_1 = e^{\nu x}$, $a_2 = e^{\nu y}$ and $b = \frac{a_1}{a_2}$. We have

$$H(x,y) = a_1 \log(a_1) - a_1(1 + \log(a_2)) = a_1(\log(b) - 1), \tag{20}$$
$$H(y,y) = a_2 \log(a_2) - a_2(1 + \log(a_2)) = -a_2, \tag{21}$$

and by symmetry, it follows that $H(x,x) = -a_1$ and $H(y,x) = -a_2(\log(b)+1)$. We thus obtain $\nu A(x,y) = c_1(\log(b) - 1) + c_2/b$, which is strictly decreasing in b over the interval $[1, \frac{c_2}{c_1})$. As $\phi()$ is a strictly increasing function, this implies that $A(x,y)$ is strictly decreasing in x over $\left[y, \log\left(\frac{c_2}{c_1}e^y\right)\right)$ for y fixed, and strictly increasing in y over $\left[\log\left(\frac{c_1}{c_2}e^x\right), x\right)$ for x fixed. Hence, the statement of Lemma 7 pertaining to $A(x,y)$ is satisfied for $\phi(x) = e^{\nu x}$. Similarly, we obtain $\nu B(x,y) = -c_1 b + c_2(\log(b)+1)$, which is clearly strictly increasing in b over the interval $[1, \frac{c_2}{c_1})$. It follows that $B(x,y)$ is strictly increasing in x over $\left[y, \log\left(\frac{c_2}{c_1}e^y\right)\right)$ for y fixed, and strictly decreasing in y over $\left[\log\left(\frac{c_1}{c_2}e^x\right), x\right)$ for x fixed, as claimed.

The proof for $\phi(x) = (1+x)^m$ is similar. In this case, we have

$$A(x,y) = -\frac{1}{m}\left[c_1(m - (m-1)b^{1/m}) - c_2 b^{-(1-1/m)}\right],$$
$$B(x,y) = -\frac{1}{m}\left[c_1 b^{1-1/m} - c_2(m - (m-1)b^{-1/m})\right],$$

where $b = \phi(x)/\phi(y)$. It is easy to check that $A(x,y)$ (resp. $B(x,y)$) is strictly decreasing (resp. increasing) in b over the interval $[1, \frac{c_2}{c_1})$, from which the result follows. □

Finally, in order to prove Proposition 4, we will need one last assumption. Before introducing this assumption, we define some additional notations. Given a fixed value of $\bar{\lambda} \geq \bar{\lambda}^{ne}$, let y_1^* (resp. y_N^*) be the flow on an arbitrary "cheap" (resp."expensive") link under the centralized routing strategy. Let us define the vector-valued function $\mathbf{y}(\Delta) = \left(y_1^* + \Delta, y_N^* - \frac{n_1}{n_2}\Delta\right)$ for $\Delta \in \left[0, \frac{n_2}{n_1}y_N^*\right]$, and

$$Q(\mathbf{y}) = \frac{n_1 F_1(y_1) + n_2 F_2(y_N)}{\delta F_1'(y_1) + (1-\delta)F_2'(y_N)}, \tag{22}$$

where $\delta = n_1 \frac{dy_1^*}{d\bar{\lambda}}$. With a slight abuse of notation, we write $Q(\Delta)$ for $Q(\mathbf{y}(\Delta))$. Note that $\mathbf{y}(0) = \mathbf{y}^*$ and that there exists $\Delta^{ne} \in \left(0, \frac{n_2}{n_1}y_N^*\right)$ such that $\mathbf{y}(\Delta^{ne}) = \mathbf{y}^{ne}$. Our last assumption is then as follows.

Assumption 5. *The latency function* ϕ *is such that* $Q(\Delta) > Q(0)$ *for all* $\Delta \in \left(0, \frac{n_2}{n_1}y_N^*\right]$.

Lemma 8 below shows that Assumption 5 is satisfied for both $\phi(x) = e^{\nu x}$ and $\phi(x) = (1+x)^m$. This lemma implies in particular that $Q(0) < Q(\Delta^{ne})$. The latter inequality is used in the proof of Proposition 4, which can be found in Appendix C.2.

Lemma 8. *For* $\phi(x) = e^{\nu x}$ *and* $\phi(x) = (1+x)^m$*, Assumption 5 holds.*

Proof. See Appendix A.1 for the proof for $\phi(x) = e^{\nu x}$, and Appendix A.2 for the proof for $\phi(x) = (1+x)^m$. □

A.1 Proof of Lemma 8 for $\phi(x) = e^{\nu x}$

Define $a_k = 1/(1+\nu y_l^*)$ for $l \in \mathcal{S}_k$ and $k = 1, 2$. Note that as $y_l^* = y_m^*$ for any links $l, m \in \mathcal{S}_k$, the value of a_k does not depend on the link $l \in \mathcal{S}_k$ considered. Moreover, it follows from property (a) in Proposition 1 that $1 > a_2 > a_1 > 0$. We first establish the expressions of δ and $Q(\Delta)$ in terms of a_1 and a_2 in Lemma 9.

Lemma 9. *It holds that*

$$\delta = \frac{n_1(1+a_2)}{n_2(1+a_1)+n_1(1+a_2)} \quad \text{and } 1-\delta = \frac{n_2(1+a_1)}{n_2(1+a_1)+n_1(1+a_2)}, \tag{23}$$

and

$$Q(\Delta) = \frac{1}{\nu}\,\frac{n_1 e^{\nu\Delta}\left(\bar{a}_1+\nu a_1\Delta\right)+n_2 e^{-\nu\omega\Delta}\left(\bar{a}_2-\nu a_2\omega\Delta\right)}{\delta e^{\nu\Delta}\left(1+\nu a_1\Delta\right)+\bar{\delta}\,e^{-\nu\omega\Delta}\left(1-\nu a_2\omega\Delta\right)}. \tag{24}$$

where $\bar{a}_k = 1 - a_k$ for $k = 1, 2$ and $\bar{\delta} = 1 - \delta$.

Proof. We have

$$F_k'(y) = c_k\left[\phi(y) + y\phi'(y)\right] = c_k(1+\nu y)e^{\nu y} \tag{25}$$

for $k = 1, 2$. Similarly, for any link $l \in \mathcal{S}_k$ we have

$$\begin{aligned} c_k\left[2\phi'(y_l^*) + y_l^*\phi''(y_l^*)\right] &= c_k\nu e^{\nu y_l^*}\left(2+\nu y_l^*\right) \\ &= c_k(1+\nu y_l^*)e^{\nu y_l^*}\,\nu\frac{2+\nu y_l^*}{1+\nu y_l^*} \\ &= F_k'(y_l^*)\nu(1+a_k), \end{aligned} \tag{26}$$

where $a_k = 1/(1+y_l^*)$. Note that the value of a_k does not depend on the link $l \in \mathcal{S}_k$ considered. Moreover, it follows from property (a) in Proposition 1 that $a_2 > a_1$. We first use (26) to establish the expression of δ. We have from (52)

$$\begin{aligned} \delta &= n_1\frac{\partial y_1^*}{\partial\bar{\lambda}}, \\ &= \frac{n_1 c_2\left[2\phi'\left(y_N^*\right) + y_N^*\,\phi''\left(y_N^*\right)\right]}{n_2 c_1\left[2\phi'\left(y_1^*\right) + y_1^*\,\phi''\left(y_1^*\right)\right] + n_1 c_2\left[2\phi'\left(y_N^*\right) + y_N^*\,\phi''\left(y_N^*\right)\right]}, \\ &= \frac{n_1 F_2'(y_N^*)\nu(1+a_2)}{n_2 F_1'(y_1^*)\nu(1+a_1) + n_1 F_2'(y_N^*)\nu(1+a_2)}, \end{aligned}$$

and with the optimality condition $F_1'(y_1^*) = F_2'(y_N^*)$, it yields (23).

We now derive an expression for the function

$$Q(\Delta) = \frac{n_1 F_1(y_1^*+\Delta_1) + n_2 F_2(y_N^*+\Delta_2)}{\delta F_1'(y_1^*+\Delta_1) + (1-\delta)F_2'(y_N^*+\Delta_2)},$$

where $\Delta_1 = \Delta$, $\Delta_2 = -\omega\Delta$ and $\omega = n_1/n_2$. For the terms in the numerator of $Q(\Delta)$, it follows from (25) that for $k = 1, 2$ and $l \in \mathcal{S}_k$

$$\begin{aligned} F_k(y_l^*+\Delta_k) &= c_k(y_l^*+\Delta_k)e^{\nu(y_l^*+\Delta_k)} \\ &= e^{\nu\Delta_k}F_k'(y_l^*)\frac{y_l^*+\Delta_k}{1+\nu y_l^*} \\ &= e^{\nu\Delta_k}F_k'(y_l^*)\frac{1}{\nu}\left[1-a_k+\nu a_k\Delta_k\right]. \end{aligned} \tag{27}$$

Similarly, for the terms in the denominator of $Q(\Delta)$, we obtain

$$
\begin{aligned}
F_k'(y_l^* + \Delta_k) &= c_k(1 + \nu(y_l^* + \Delta_k))e^{\nu(y_l^* + \Delta_k)} \\
&= e^{\nu\Delta_k} F_k'(y_l^*) \frac{1 + \nu y_l^* + \nu\Delta_k}{1 + \nu y_l^*} \\
&= e^{\nu\Delta_k} F_k'(y_l^*) \left[1 + \nu a_k \Delta_k\right]. \qquad (28)
\end{aligned}
$$

Combining (27) and (28) and using the optimality condition $F_1'(y_1^*) = F_2'(y_N^*)$, we obtain (24). □

Interestingly, we note from (23) that $a_2 > a_1$ implies that $\delta = n_1 \frac{\partial y_1^*}{\partial \lambda} > 1 - \delta = n_2 \frac{\partial y_N^*}{\partial \lambda}$.

We shall use Lemma 9 to derive a sufficient condition for $Q(\Delta) > Q(0)$ to hold true for all $\Delta \in \left(0, \frac{n_2}{n_1} y_N^*\right]$. To this end, we will need the following result.

Lemma 10. *Let $\omega = \frac{n_1}{n_2}$. For any strictly positive numbers m_1, m_2 and p such that $m_2 - m_1 = \nu(1+\omega)p$, it holds that*

$$
e^{\nu\Delta}\left[m_1\Delta + p\right] > e^{-\nu\omega\Delta}\left[m_2\Delta + p\right], \qquad (29)
$$

for any $\Delta > 0$.

Proof. The proof is based on the well-known inequality $e^x \geq x + 1$. It yields $e^{\nu(1+\omega)\Delta} \geq \nu(1+\omega)\Delta + 1$, and it is therefore enough to show that

$$
\nu(1+\omega)\Delta > \frac{m_2\Delta + p}{m_1\Delta + p} - 1 = \frac{(m_2 - m_1)\Delta}{m_1\Delta + p} = \frac{\nu(1+\omega)p\Delta}{m_1\Delta + p}, \qquad (30)
$$

which is equivalent to $m_1\Delta > 0$. As $m_1 > 0$, the latter inequality clearly holds true. □

We are now in position to prove Lemma 8 for $\phi(x) = e^{\nu x}$.

Proof of Lemma 8 for $\phi(x) = e^{\nu x}$. With Lemma 9, we have

$$
Q(0) < Q(\Delta) \iff e^{\nu\Delta}\left[n_1\bar{a}_1 - Q(0)\delta + (n_1 - Q(0)\delta)\nu a_1\Delta\right] > e^{-\nu\omega\Delta}\left[Q(0)\bar{\delta} - n_2\bar{a}_2 + (n_2 - Q(0)\bar{\delta})\nu a_2\omega\Delta\right],
$$

Let $p = n_1\bar{a}_1 - Q(0)\delta$. Note from (24) that $Q(0) = n_1\bar{a}_1 + n_2\bar{a}_2$. It follows that

$$
Q(0)\bar{\delta} - n_2\bar{a}_2 = Q(0) - Q(0)\delta - n_2\bar{a}_2 = n_1\bar{a}_1 - Q(0)\delta = p.
$$

Letting $m_1 = (n_1 - Q(0)\delta)\nu a_1$ and $m_2 = (n_2 - Q(0)\bar{\delta})\nu a_2\omega$, it yields

$$
Q(0) < Q(\Delta) \iff e^{\nu\Delta}\left[m_1\Delta + p\right] > e^{-\nu\omega\Delta}\left[m_2\Delta + p\right]. \qquad (31)
$$

With $Q(0) = n_1\bar{a}_1 + n_2\bar{a}_2$ and (23), we obtain after some algebra

$$
p = n_1 n_2 \frac{a_2^2 - a_1^2}{n_2(1 + a_1) + n_1(1 + a_2)} > 0.
$$

Similarly, we have

$$
\begin{aligned}
m_1 &= \nu\, n_1 a_1 \frac{a_2(n_1 a_1 + n_2 a_2) + a_1(n_1 + n_2)}{n_2(1 + a_1) + n_1(1 + a_2)}, \\
m_2 &= \nu\, n_1 a_2 \frac{a_1(n_1 a_1 + n_2 a_2) + a_2(n_1 + n_2)}{n_2(1 + a_1) + n_1(1 + a_2)},
\end{aligned}
$$

from which we conclude that $m_1 > 0$, $m_2 > 0$ and $m_2 - m_1 = \nu(1+\omega)p$. The assumptions of Lemma 10 are met and we can thus conclude that inequality (31) is satisfied for all $\Delta > 0$. We thus get that $Q(0) < Q(\Delta)$ for all $\Delta \in \left(0, \frac{1}{\omega} y_N^*\right]$, as claimed.

□

A.2 Proof of Lemma 8 for $\phi(x) = (1+x)^m$

We proceed as in the case $\phi(x) = e^x$. Define $a_k = \frac{1}{1+y_l^*}$ and $b_k = \frac{1}{1+(m+1)y_l^*}$ for $k = 1, 2$ and $l \in \mathcal{S}_k$. Note that the values of a_k and b_k do not depend on the chosen link $l \in \mathcal{S}_k$. Moreover, note that $y_1^* > y_N^*$ implies that $a_1 < a_2$ and $b_1 < b_2$.

We first establish the expressions of δ and $Q(\Delta)$ in terms of a_1, a_2, b_1 and b_2 in Lemma 11.

Lemma 11. *It holds that*

$$\delta = \frac{\omega a_2(1+b_2)}{a_1(1+b_1) + \omega a_2(1+b_2)}, \tag{32}$$

$$1-\delta = \frac{a_1(1+b_1)}{a_1(1+b_1) + \omega a_2(1+b_2)}, \tag{33}$$

where $\omega = \frac{n_1}{n_2}$ *as in Appendix A.1, and*

$$Q(\Delta) = \frac{n_2}{m+1} \frac{\frac{\omega}{a_1}\left[1 - b_1 + (m+1)b_1\Delta\right](1+a_1\Delta)^m + \frac{1}{a_2}\left[1 - b_2 - (m+1)b_2\omega\Delta\right](1-a_2\omega\Delta)^m}{\delta\left[1+(m+1)b_1\Delta\right](1+a_1\Delta)^{m+1} + (1-\delta)\left[1-(m+1)b_2\omega\Delta\right](1-a_2\omega\Delta)^{m+1}}, \tag{34}$$

Proof. For $\phi(x) = (1+x)^m$, we have $\phi'(x) = m(1+x)^{m-1}$ and $\phi''(x) = m(m-1)(1+x)^{m-2}$, from which we obtain

$$F_k'(y) = c_k\left[\phi(y) + y\phi'(y)\right] = c_k(1+y)^{m-1}\left[1+(m+1)y\right], \quad k = 1, 2. \tag{35}$$

Besides, we also have

$$\begin{aligned} 2\phi'(y) + y\phi''(y) &= m(1+y)^{m-2}\left[2+(m+1)y\right] \\ &= \frac{m}{1+y}(1+y)^{m-1}\left[1+(m+1)y\right]\left[1 + \frac{1}{1+(m+1)y}\right], \end{aligned}$$

which yields with (35)

$$c_k\left[2\phi'(y_l^*) + y_l^*\phi''(y_l^*)\right] = ma_kF_k'(y_l^*)(1+b_k), \quad l \in \mathcal{S}_k,\ k = 1, 2. \tag{36}$$

We now use (36) to establish the expression of δ. We have from (52)

$$\begin{aligned} \delta &= n_1\frac{\partial y_1^*}{\partial \overline{\lambda}} \\ &= \frac{n_1c_2\left[2\phi'\left(y_N^*\right) + y_N^*\,\phi''\left(y_N^*\right)\right]}{n_2c_1\left[2\phi'\left(y_1^*\right) + y_1^*\,\phi''\left(y_1^*\right)\right] + n_1c_2\left[2\phi'\left(y_N^*\right) + y_N^*\,\phi''\left(y_N^*\right)\right]} \\ &= \frac{n_1ma_2F_2'(y_N^*)(1+b_2)}{n_2ma_1F_1'(y_1^*)(1+b_1) + n_1ma_2F_2'(y_N^*)(1+b_2)}, \end{aligned}$$

and with the optimality condition $F_1'(y_1^*) = F_2'(y_N^*)$, it yields (32) and (33).

We now derive an expression for the function

$$Q(\Delta) = \frac{n_1F_1(y_1^* + \Delta_1) + n_2F_2(y_N^* + \Delta_2)}{\delta_1F_1'(y_1 + \Delta_1) + \delta_2F_2'(y_N^* + \Delta_2)},$$

where $\delta_1 = \delta$, $\delta_2 = 1-\delta$, $\Delta_1 = \Delta$, $\Delta_2 = -\omega\Delta$. Regarding the terms in the numerator of $Q(\Delta)$, we observe that

$$\begin{aligned} F_k(y_l^* + \Delta_k) &= \frac{c_k}{m+1}\left[1+(m+1)y_l^* + (m+1)\Delta_k - 1\right](1+y_l^*)^m\left(1 + \frac{\Delta_k}{1+y_l^*}\right)^m \\ &= \frac{1}{(m+1)a_k}F_k'(y_l^*)\left[1 - b_k + (m+1)b_k\Delta_k\right](1+a_k\Delta_k)^m. \end{aligned} \tag{37}$$

For the terms in the denominator of $Q(\Delta)$, we have

$$\begin{aligned} F_k'(y_l^* + \Delta_k) &= c_k(1+y_l^*)^{m-1}\,(1+a_k\Delta_k)^m\,[1+(m+1)y_l^*]\,[1+(m+1)b_k\Delta_k] \\ &= F_k'(y_l^*)\,[1+(m+1)b_k\Delta_k]\,(1+a_k\Delta_k)^{m-1}\,. \end{aligned} \tag{38}$$

Combining (37) and (38), and using $F_1'(y_1^*) = F_2'(y_N^*)$, we obtain

$$Q(\Delta) = \frac{1}{m+1}\frac{\sum_{k=1}^{2}\frac{n_k}{a_k}\,[1-b_k+(m+1)b_k\Delta_k]\,(1+a_k\Delta_k)^m}{\sum_{k=1}^{2}\delta_k\,[1+(m+1)b_k\Delta_k]\,(1+a_k\Delta_k)^{m-1}}, \tag{39}$$

which corresponds to (34).

□

We now exploit the expression of $Q(\Delta)$ given in Lemma 11 to obtain an equivalent condition for $Q(0) < Q(\Delta)$ to be true for all $\Delta > 0$. Note from (34) that

$$Q(0) = \frac{n_2}{m+1}\left\{\omega\frac{1-b_1}{a_1} + \frac{1-b_2}{a_2}\right\}, \tag{40}$$

and define $Q_0(0) = \frac{m+1}{n_2}Q(0) = \omega\frac{1-b_1}{a_1} + \frac{1-b_2}{a_2}$. We also define the following quantities:

- $p_1 = \delta Q_0(0) - \frac{\omega(1-b_1)}{a_1}$, $p_2 = b_1(m+1)\delta Q_0(0) - \omega(1-b_1) - \frac{\omega(m+1)b_1}{a_1}$ and $p_3 = -\omega(m+1)b_1$,
- $q_1 = Q_0(0)(1-\delta) - \frac{1-b_2}{a_2}$, $q_2 = -Q_0(0)(1-\delta)(m+1)b_2\omega + (1-b_2)\omega + \frac{(m+1)b_2\omega}{a_2}$ and $q_3 = -(m+1)b_2\omega^2$.

The equivalent condition is stated in Lemma 12 below.

Lemma 12. *The condition $Q(0) < Q(\Delta)$ is equivalent to*

$$\left(\frac{1+a_1\Delta}{1-a_2\omega\Delta}\right)^{m-1}(p_1 + p_2\Delta + p_3\Delta^2) + (q_1 + q_2\Delta + q_3\Delta^2) < 0. \tag{41}$$

Proof. With (34) and (40), the condition $Q(0) < Q(\Delta)$ is equivalent to

$$\begin{aligned} (1+a_1\Delta)^{m-1}\left\{Q_0(0)\delta[1+(m+1)b_1\Delta] - \frac{\omega}{a_1}[1-b_1+(m+1)b_1\Delta](1+a_1\Delta)\right\} &+ \\ (1-a_2\omega\Delta)^{m-1}\left\{Q_0(0)(1-\delta)[1-(m+1)b_2\omega\Delta] - \frac{1}{a_2}[1-b_2-(m+1)b_2\omega\Delta](1-a_2\omega\Delta)\right\} &< 0 \end{aligned}$$

Observe that

$$Q_0(0)\delta[1+(m+1)b_1\Delta] - \frac{\omega}{a_1}[1-b_1+(m+1)b_1\Delta](1+a_1\Delta) = p_1 + p_2\Delta + p_3\Delta^2,$$

and

$$Q_0(0)(1-\delta)[1-(m+1)b_2\omega\Delta] - \frac{1}{a_2}[1-b_2-(m+1)b_2\omega\Delta](1-a_2\omega\Delta) = q_1 + q_2\Delta + q_3\Delta^2,$$

which yields the equivalent condition (41).

□

We shall use the equivalent condition provided in Lemma 12 to derive a sufficient condition for $Q(0) < Q(\Delta)$ to be true for all $\Delta > 0$. To this end, we will need the following technical result.

Lemma 13. *Let p_1, p_2 and q_2 be as defined above. It holds that*

$$(m-1)(a_1 + a_2\omega)\,p_1 + p_2 + q_2 = 0. \tag{42}$$

Proof. First, we note from (32) and (40) that

$$\begin{aligned}
p_1 &= \delta Q_0(0) - \frac{\omega(1-b_1)}{a_1} \\
&= \frac{\omega a_2(1+b_2)}{a_1(1+b_1)+\omega a_2(1+b_2)}\left(\frac{\omega(1-b_1)}{a_1}+\frac{1-b_2}{a_2}\right) - \frac{\omega(1-b_1)}{a_1} \\
&= \frac{\omega(b_1^2-b_2^2)}{a_1(1+b_1)+\omega a_2(1+b_2)}.
\end{aligned}$$

Thus, we aim to show that

$$\omega(b_1^2-b_2^2)(m-1)(a_1+a_2\omega)+(p_2+q_2)(a_1(1+b_1)+\omega a_2(1+b_2)) = 0. \tag{43}$$

From the definition of p_2, and exploiting again (32) and (40), we obtain

$$\begin{aligned}
&p_2(a_1(1+b_1)+\omega a_2(1+b_2)) \\
&= \left(b_1(m+1)\delta Q_0(0) - \omega(1-b_1) - \frac{\omega(m+1)b_1}{a_1}\right)(a_1(1+b_1)+\omega a_2(1+b_2)) \\
&= \left(\frac{b_1(m+1)\omega a_2(1+b_2)}{a_1(1+b_1)+\omega a_2(1+b_2)}\left(\frac{\omega(1-b_1)}{a_1}+\frac{1-b_2}{a_2}\right) - \omega(1-b_1) - \frac{\omega(m+1)b_1}{a_1}\right) \\
&\quad (a_1(1+b_1)+\omega a_2(1+b_2)) \\
&= b_1(m+1)\omega a_2(1+b_2)\left(\frac{\omega(1-b_1)}{a_1}+\frac{1-b_2}{a_2}\right) - \omega(1-b_1)(a_1(1+b_1)+\omega a_2(1+b_2)) \\
&\quad - \frac{\omega(m+1)b_1}{a_1}(a_1(1+b_1)+\omega a_2(1+b_2)) \\
&= \frac{b_1a_2(m+1)\omega(1+b_2)\omega(1-b_1)}{a_1} + b_1(m+1)\omega(1-b_2^2) - a_1\omega(1-b_1^2) \\
&\quad - a_2\omega^2(1-b_1)(1+b_2) - \omega(m+1)b_1(1+b_1) - \frac{a_2\omega^2(m+1)b_1(1+b_2)}{a_1}.
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
q_2(a_1(1+b_1)+\omega a_2(1+b_2)) &= \left(-Q_0(0)(1-\delta)(m+1)b_2\omega + (1-b_2)\omega + \frac{(m+1)b_2\omega}{a_2}\right) \\
&\quad (a_1(1+b_1)+\omega a_2(1+b_2)) \\
&= -\left(\frac{\omega(1-b_1)}{a_1}+\frac{1-b_2}{a_2}\right)a_1(1+b_1)(m+1)b_2\omega \\
&\quad + (1-b_2)\omega(a_1(1+b_1)+\omega a_2(1+b_2)) \\
&\quad + \frac{(m+1)b_2\omega}{a_2}(a_1(1+b_1)+\omega a_2(1+b_2)) \\
&= -\omega(1-b_1^2)(m+1)b_2\omega - \frac{a_1b_2(m+1)(1-b_2)(1+b_1)\omega}{a_2} \\
&\quad + (1-b_2)\omega a_1(1+b_1) + (1-b_2)\omega^2 a_2(1+b_2) \\
&\quad + \frac{(m+1)b_2\omega a_1(1+b_1)}{a_2} + (m+1)b_2\omega^2(1+b_2)
\end{aligned}$$

We also note that

$$\omega(b_1^2-b_2^2)(m-1)(a_1+a_2\omega) = \omega(1-b_2^2)(m-1)(a_1+a_2\omega) - \omega(1-b_1^2)(m-1)(a_1+a_2\omega)$$

Hence, replacing the above expressions in (43) and moving all the negative terms to the other side, we

derive the following equivalent equality:

$$\begin{aligned}
&(a_1+a_2\omega)\omega(1-b_1^2)(m-1)+a_1\omega(1-b_1^2)+a_2\omega^2(1-b_1)(1+b_2)+\omega b_1(1+b_1)(m+1)\\
&+\omega^2 b_2(1-b_1^2)(m+1)+\frac{a_2\omega^2(m+1)b_1(1+b_2)}{a_1}+\frac{a_1b_2(m+1)(1-b_2)(1+b_1)\omega}{a_2}\\
=&(a_1+a_2\omega)\omega(1-b_2^2)(m-1)+a_2\omega^2(1-b_2^2)+a_1\omega(1+b_1)(1-b_2)+\omega^2b_2(1+b_2)(m+1)\\
&+\omega b_1(1-b_2^2)(m+1)+\frac{a_1b_2\omega(m+1)(1+b_1)}{a_2}+\frac{a_2b_1(m+1)(1+b_2)(1-b_1)\omega^2}{a_1}
\end{aligned}$$

Dividing by ω, we obtain

$$\begin{aligned}
&(a_1+a_2\omega)(1-b_1^2)(m-1)+a_1(1-b_1^2)+a_2\omega(1-b_1)(1+b_2)+b_1(1+b_1)(m+1)\\
&+\omega b_2(1-b_1^2)(m+1)+\frac{a_2\omega b_1(m+1)(1+b_2)}{a_1}+\frac{a_1b_2(m+1)(1+b_1)(1-b_2)}{a_2}\\
=&(a_1+a_2\omega)(1-b_2^2)(m-1)+a_2\omega(1-b_2^2)+a_1(1+b_1)(1-b_2)+\omega b_2(1+b_2)(m+1)\\
&+b_1(1-b_2^2)(m+1)+\frac{a_1b_2(m+1)(1+b_1)}{a_2}+\frac{a_2\omega b_1(m+1)(1+b_2)(1-b_1)}{a_1},
\end{aligned}$$

which yields

$$\begin{aligned}
&(a_1+a_2\omega)(1-b_1^2)(m-1)+a_1(1-b_1^2)+a_2\omega(1-b_1)(1+b_2)+b_1(1+b_1)(m+1)\\
&+\omega b_2(1-b_1^2)(m+1)-\frac{a_1b_2^2(1+b_1)(m+1)}{a_2}\\
=&(a_1+a_2\omega)(1-b_2^2)(m-1)+a_2\omega(1-b_2^2)+a_1(1+b_1)(1-b_2)+\omega b_2(1+b_2)(m+1)\\
&+b_1(1-b_2^2)(m+1)-\frac{a_2\omega b_1^2(1+b_2)(m+1)}{a_1},
\end{aligned}$$

or equivalently

$$\begin{aligned}
&(a_1+a_2\omega)(1-b_1^2)(m-1)+a_1(1-b_1^2)+a_2\omega(1-b_1)(1+b_2)+b_1(1+b_1)(m+1)\\
&+\omega b_2(1-b_1^2)(m+1)+\frac{a_2\omega b_1^2(1+b_2)(m+1)}{a_1}\\
=&(a_1+a_2\omega)(1-b_2^2)(m-1)+a_2\omega(1-b_2^2)+a_1(1+b_1)(1-b_2)+\omega b_2(1+b_2)(m+1)\\
&+b_1(1-b_2^2)(m+1)+\frac{a_1b_2^2(1+b_1)(m+1)}{a_2}.
\end{aligned}$$

Multiplying on both sides of the above equality by a_1a_2, we obtain

$$\begin{aligned}
&a_1a_2(a_1+a_2\omega)(1-b_1^2)(m-1)+a_1^2a_2(1-b_1^2)+a_1a_2^2\omega(1-b_1)(1+b_2)+a_1a_2b_1(1+b_1)(m+1)\\
&+a_1a_2\omega b_2(1-b_1^2)(m+1)+a_2^2\omega b_1^2(1+b_2)(m+1)\\
=&a_1a_2(a_1+a_2\omega)(1-b_2^2)(m-1)+a_1a_2^2\omega(1-b_2^2)+a_1^2a_2(1+b_1)(1-b_2)+a_1a_2\omega b_2(1+b_2)(m+1)\\
&+a_1a_2b_1(1-b_2^2)(m+1)+a_1^2b_2^2(1+b_1)(m+1),
\end{aligned}$$

that is

$$\begin{aligned}
&a_1a_2(a_1+a_2\omega)(1-b_1^2)(m-1)+a_1^2a_2(1-b_1^2)+a_1a_2^2\omega(1-b_1)(1+b_2)+a_1a_2b_1^2(m+1)\\
&-a_1a_2\omega b_1^2b_2(m+1)+a_2^2\omega b_1^2(1+b_2)(m+1)\\
=&a_1a_2(a_1+a_2\omega)(1-b_2^2)(m-1)+a_1a_2^2\omega(1-b_2^2)+a_1^2a_2(1+b_1)(1-b_2)+a_1a_2\omega b_2^2(m+1)\\
&-a_1a_2b_1b_2^2(m+1)+a_1^2b_2^2(1+b_1)(m+1).
\end{aligned}$$

We move the negative terms on the LHS to the RHS and the negative terms of the RHS to the LHS to obtain

$$\begin{aligned}
&a_1a_2(a_1+a_2\omega)(1-b_1^2)(m-1)+a_1^2a_2(1-b_1^2)+a_1a_2^2\omega(1-b_1)(1+b_2)+a_1a_2b_1^2(m+1)\\
&+a_1a_2b_1b_2^2(m+1)+a_2^2\omega b_1^2(1+b_2)(m+1)\\
=&a_1a_2(a_1+a_2\omega)(1-b_2^2)(m-1)+a_1a_2^2\omega(1-b_2^2)+a_1^2a_2(1+b_1)(1-b_2)+a_1a_2\omega b_2^2(m+1)\\
&+a_1a_2\omega b_1^2b_2(m+1)+a_1^2b_2^2(1+b_1)(m+1),
\end{aligned}$$

which after simplification yields

$$\begin{aligned}
&a_1a_2(a_1+a_2\omega)b_2^2(m-1)+a_1^2a_2(1+b_1)b_2+a_1a_2^2\omega(1+b_2)b_2+a_1a_2b_1^2(m+1)\\
&+a_1a_2b_1b_2^2(m+1)+a_2^2\omega b_1^2(1+b_2)(m+1)\\
=&a_1a_2(a_1+a_2\omega)b_1^2(m-1)+a_1^2a_2(1+b_1)b_1+a_1a_2^2\omega(1+b_2)b_1+a_1a_2\omega b_2^2(m+1)\\
&+a_1a_2\omega b_1^2b_2(m+1)+a_1^2b_2^2(1+b_1)(m+1). \qquad (44)
\end{aligned}$$

We prove below that the terms in factor of ω in (44) are equal and the other terms are equal as well. For the terms in factor of ω, we prove that

$$\begin{aligned}
a_1a_2^2b_2^2(m-1)+a_1a_2^2(1+b_2)b_2+a_2^2b_1^2(1+b_2)(m+1)=&a_1a_2^2b_1^2(m-1)+a_1a_2^2(1+b_2)b_1\\
&+a_1a_2b_2^2(m+1)+a_1a_2b_1^2b_2(m+1), \qquad (45)
\end{aligned}$$

whereas for the other terms we prove that

$$\begin{aligned}
&a_1^2a_2b_2^2(m-1)+a_1^2a_2(1+b_1)b_2+a_1a_2b_1^2(m+1)+a_1a_2b_1b_2^2(m+1)\\
&\qquad=a_1^2a_2b_1^2(m-1)+a_1^2a_2(1+b_1)b_1+a_1^2b_2^2(1+b_1)(m+1). \qquad (46)
\end{aligned}$$

We first focus on (45). Dividing by a_2, we obtain

$$\begin{aligned}
&a_1a_2b_2^2(m-1)+a_1a_2(1+b_2)b_2+a_2b_1^2(1+b_2)(m+1)\\
&\qquad=a_1a_2b_1^2(m-1)+a_1a_2(1+b_2)b_1+a_1b_2^2(m+1)+a_1b_1^2b_2(m+1),
\end{aligned}$$

which yields after some algebra

$$a_1a_2(b_2-b_1)(mb_2+(m-1)b_1+1)+(m+1)(a_2b_1^2-a_1b_2^2+b_1^2b_2(a_2-a_1))=0.$$

Let $c_k=m+1-ma_k$ for $k=1,2$. Hence, $b_k=a_k/c_k$ and the previous equality can be written as follows

$$a_1a_2\left(\frac{a_2}{c_2}-\frac{a_1}{c_1}\right)\left(m\frac{a_2}{c_2}+(m-1)\frac{a_1}{c_1}+1\right)+(m+1)\left(a_2\left(\frac{a_1}{c_1}\right)^2-a_1\left(\frac{a_2}{c_2}\right)^2+\left(\frac{a_1}{c_1}\right)^2\frac{a_2}{c_2}(a_2-a_1)\right)=0.$$

Dividing by a_1a_2, it yields

$$\left(\frac{a_2}{c_2}-\frac{a_1}{c_1}\right)\left(m\frac{a_2}{c_2}+(m-1)\frac{a_1}{c_1}+1\right)+(m+1)\left(\frac{a_1}{c_1^2}-\frac{a_2}{c_2^2}+\frac{a_1a_2}{c_1^2c_2}-\frac{a_1^2}{c_1^2c_2}\right)=0.$$

We note that $\frac{a_2}{c_2}-\frac{a_1}{c_1}=\frac{(m+1)(a_2-a_1)}{c_1c_2}$ and thus

$$\frac{a_2-a_1}{c_1c_2}\left(m\frac{a_2}{c_2}+(m-1)\frac{a_1}{c_1}+1\right)+(m+1)\left(\frac{a_1}{c_1^2}-\frac{a_2}{c_2^2}+\frac{a_1a_2}{c_1^2c_2}-\frac{a_1^2}{c_1^2c_2}\right)=0 \iff$$

$$m\frac{a_2(a_2-a_1)}{c_1c_2^2}+(m-1)\frac{a_1(a_2-a_1)}{c_1^2c_2}+\frac{a_2-a_1}{c_1c_2}+\frac{a_1}{c_1^2}-\frac{a_2}{c_2^2}+\frac{a_1a_2}{c_1^2c_2}-\frac{a_1^2}{c_1^2c_2}=0 \iff$$

$$m\frac{a_2(a_2-a_1)}{c_1c_2^2}+m\frac{a_1(a_2-a_1)}{c_1^2c_2}+\frac{a_2-a_1}{c_1c_2}+\frac{a_1}{c_1^2}-\frac{a_2}{c_2^2}=0. \qquad (47)$$

Multiplying (47) by $c_1^2c_2$, we obtain

$$mc_1a_2(a_2-a_1)+mc_2a_1(a_2-a_1)+c_1c_2(a_2-a_1)+=c_1^2a_2-c_2^2a_1 \iff$$

$$m(a_2-a_1)(c_1a_2+c_2a_1)=c_1^2a_2-c_2^2a_1-c_1c_2(a_2-a_1) \iff$$

$$m(a_2-a_1)(c_1a_2+c_2a_1)=(c_1-c_2)(c_1a_2+c_2a_1) \iff$$

$$m(a_2 - a_1) = (c_1 - c_2),$$

and since the last expression holds, we have shown that (45) holds. We now focus on (46). Dividing by a_1, we obtain

$$\begin{aligned} a_1a_2b_2^2(m-1) + a_1a_2(1+b_1)b_2 + a_2b_1^2(m+1) + a_2b_1b_2^2(m+1) \\ = a_1a_2b_1^2(m-1) + a_1a_2(1+b_1)b_1 + a_1b_2^2(1+b_1)(m+1). \end{aligned}$$

After some algebra, it yields

$$a_1a_2(b_2 - b_1)((m-1)b_2 + mb_1 + 1) + (m+1)(a_2b_1^2 - a_1b_2^2 + b_1b_2^2(a_2 - a_1)) = 0.$$

Using that $c_k = m+1-ma_k$ and $b_k = a_k/c_k$ for $k = 1, 2$, the previous equality can be written as follows

$$a_1a_2\left(\frac{a_2}{c_2} - \frac{a_1}{c_1}\right)\left((m-1)\frac{a_2}{c_2} + m\frac{a_1}{c_1} + 1\right) + (m+1)\left(a_2\left(\frac{a_1}{c_1}\right)^2 - a_1\left(\frac{a_2}{c_2}\right)^2 + \frac{a_1}{c_1}\left(\frac{a_2}{c_2}\right)^2(a_2 - a_1)\right) = 0,$$

which upon division by a_1a_2 yields

$$\left(\frac{a_2}{c_2} - \frac{a_1}{c_1}\right)\left((m-1)\frac{a_2}{c_2} + m\frac{a_1}{c_1} + 1\right) + (m+1)\left(\frac{a_1}{c_1^2} - \frac{a_2}{c_2^2} + \frac{a_2}{c_1c_2^2}(a_2 - a_1)\right) = 0.$$

Using that $\frac{a_2}{c_2} - \frac{a_1}{c_1} = \frac{(m+1)(a_2-a_1)}{c_1c_2}$ and simplifying, we get

$$\frac{a_2 - a_1}{c_1c_2}\left(m\frac{a_2}{c_2} + m\frac{a_1}{c_1} + 1\right) + \left(\frac{a_1}{c_1^2} - \frac{a_2}{c_2^2}\right) = 0,$$

which coincides with (47). As the latter equality has been shown to hold above, (46) also holds true, which concludes the proof. □

We are now in position to prove Lemma 8 for $\phi(x) = (1+x)^m$.

Proof of Lemma 8 for $\phi(x) = (1+x)^m$. We use (41) to obtain a sufficient condition for $Q(0) < Q(\Delta)$ to be true, and then prove that this condition is satisfied. We first show that $p_1 < 0$ as follows

$$\begin{aligned} \delta Q_0(0) - \frac{\omega(1-b_1)}{a_1} < 0 &\iff \delta\left(\frac{\omega(1-b_1)}{a_1} + \frac{1-b_2}{a_2}\right) - \frac{\omega(1-b_1)}{a_1} < 0 \\ &\iff \delta\frac{\omega(1-b_1)}{a_1} + \delta\frac{1-b_2}{a_2} - \frac{\omega(1-b_1)}{a_1} < 0 \\ &\iff (\delta-1)\frac{\omega(1-b_1)}{a_1} + \delta\frac{1-b_2}{a_2} < 0 \\ &\iff \frac{-a_1(1+b_1)\omega(1-b_1)}{a_1} + \frac{\omega a_2(1+b_2)(1-b_2)}{a_2} < 0 \\ &\iff -(1+b_1)(1-b_1) + (1+b_2)(1-b_2) < 0 \\ &\iff b_1^2 - b_2^2 < 0, \end{aligned}$$

and since $b_1 < b_2$ the last expression is true and we have shown that $p_1 < 0$. Likewise, we can show that p_2 is negative. In fact, after some algebra, we get that

$$b_1(m+1)\delta Q_0(0) - \omega(1-b_1) - \frac{\omega(m+1)b_1}{a_1} = \left(\frac{\omega(-b_1 - b_2^2)}{a_1(1+b_1) + \omega a_2(1+b_2)} - \frac{b_1\delta\omega}{a1}\right)b_1(m+1) - \omega(1-b_1),$$

and the RHS of the above expression is clearly negative.

Since $p_3 < 0$ and from the above reasoning, we know that that $p_1 + \Delta p_2 + p_3\Delta^2 < 0$. As a result, using that $(1+x)^m \geq 1 + xm$ for $x > 0$, it follows that a sufficient condition for (41) to hold is

$$\left(1 + \Delta(m-1)\frac{a_1 + a_2\omega}{1 - a_2\omega\Delta}\right)[p_3\Delta^2 + p_2\Delta + p_1] + q_3\Delta^2 + q_2\Delta + q_1 < 0. \tag{48}$$

Let us note that

$$p_1 + q_1 = \delta Q_0(0) - \frac{\omega(1-b_1)}{a_1} + \bar{\delta} Q_0(0) - \frac{(1-b_2)}{a_2} = Q_0(0) - \frac{\omega(1-b_1)}{a_1} - \frac{(1-b_2)}{a_2} = 0,$$

that is, $q_1 = -p_1$. Thus, we need to show that

$$\left(1 + \Delta(m-1)\frac{a_1 + a_2\omega}{1 - a_2\omega\Delta}\right)[p_3\Delta^2 + p_2\Delta] + q_3\Delta^2 + q_2\Delta + p_1\Delta(m-1)\frac{a_1 + a_2\omega}{1 - a_2\omega\Delta} < 0,$$

for all $\Delta > 0$. Dividing by Δ, we get

$$\left(1 + \Delta(m-1)\frac{a_1 + a_2\omega}{1 - a_2\omega\Delta}\right)[p_3\Delta + p_2] + q_3\Delta + q_2 + p_1(m-1)\frac{a_1 + a_2\omega}{1 - a_2\omega\Delta} < 0. \tag{49}$$

Since $p_3 < 0$ and $p_2 < 0$, we know that

$$\Delta(m-1)\frac{a_1 + a_2\omega}{1 - a_2\omega\Delta}[p_3\Delta + p_2] < 0.$$

Besides, it follows from $p_1 < 0$ that

$$p_1(m-1)\frac{a_1 + a_2\omega}{1 - a_2\omega\Delta} \leq p_1(m-1)(a_1 + a_2\omega).$$

Therefore, a sufficient condition for (49) to be true is

$$p_3\Delta + p_2 + q_3\Delta + q_2 + p_1(m-1)(a_1 + a_2\omega) < 0$$

that is

$$(p_3 + q_3)\Delta + (p_2 + q_2 + (m-1)(a_1 + a_2\omega)p_1) < 0,$$

for all $\Delta > 0$. With Lemma 13, we know that $p_2 + q_2 + (m-1)(a_1 + a_2\omega)p_1 = 0$, and the above sufficient condition reduces to $(p_3 + q_3)\Delta < 0$. Note from their definitions that $p3 < 0$ and $q_3 < 0$. The sufficient condition is therefore satisfied for all $\Delta > 0$, which proves that $Q(0) < Q(\Delta)$ for all $\Delta > 0$. □

B Proof of Results in Section 4.1

B.1 Proof of Lemma 1

Proof of Lemma 1. We only give the proof for the centralized routing scheme as the proof for $\mathbf{y}^{ne}$ follows similar arguments. As $y_k^* = y_1^*$ for all cheap links $k \in \mathcal{S}_1$ and $y_j^* = y_N^*$ for all expensive links $j \in \mathcal{S}_2$, it is enough to show the continuity of y_1^* and y_N^*. Moreover, as $y_N^* = (\bar{\lambda} - n_1 y_1^*)/n_2$, it is enough to show that y_1^* is continuous in $\bar{\lambda}$.

For $\bar{\lambda} < \bar{\lambda}^*$, we know that $y_1^* = \bar{\lambda}$, from which the continuity of y_1^* over $[0, \bar{\lambda}^*)$ follows.

For $\bar{\lambda} \geq \bar{\lambda}^*$, y_1^* is defined in (8) as the solution of $J(\bar{\lambda}, y_1^*) = 0$, where

$$J(\bar{\lambda}, y_1^*) = F_1'(y_1^*) - F_2'\left(\frac{\bar{\lambda} - n_1 y_1^*}{n_2}\right).$$

As J is of class C^1 and $\frac{\partial J}{\partial y_1^*}(\bar{\lambda}, y_1^*) > 0$ at all points $(\bar{\lambda}, y_1^*)$ such that $J(\bar{\lambda}, y_1^*) = 0$, it follows from the *Implicit Function Theorem* that y_1^* is a continuously differentiable function of $\bar{\lambda}$ in the neighborhood of all such points. Hence, y_1^* is a continuous function of $\bar{\lambda}$ over $(\bar{\lambda}^*, \infty)$.

It remains to show that y_1^* is a continuous function of $\bar{\lambda}$ at point $\bar{\lambda} = \bar{\lambda}^*$, that is, that $y_1^* \to \frac{\bar{\lambda}^*}{n_1}$ as $\bar{\lambda} \to \bar{\lambda}^*$. As the result is obvious when $\bar{\lambda} \to \bar{\lambda}^*$ from below, we focus on the case when $\bar{\lambda} \to \bar{\lambda}^*$ from above. We want to show that for any $\epsilon > 0$, there exists $\theta > 0$ such that $\frac{\bar{\lambda}^*}{n_1} \geq y_1^* > \frac{\bar{\lambda}^*}{n_1} - \epsilon$ for all $\bar{\lambda} \in (\bar{\lambda}^*, \bar{\lambda}^* + \theta)$. Assume on the contrary that there exists $\epsilon > 0$ such that for all $\theta > 0$ we can find $\bar{\lambda}$ in the interval $(\bar{\lambda}^*, \bar{\lambda}^* + \theta)$ such that the solution y_1^* of $J(\bar{\lambda}, y_1^*) = 0$ satisfies $y_1^* \leq \frac{\bar{\lambda}^*}{n_1} - \epsilon$. Note that it also implies that $y_N^* > \frac{n_1}{n_2}\epsilon$. As the functions $F_1'(y)$ and $F_2'(y)$ are strictly increasing in y, we then have $F_1'(y_1^*) < F_1'(\bar{\lambda}^*/n_1)$ and $F_2'(y_N^*) > c_2$, from which we obtain $J(\bar{\lambda}, y_1^*) < J(\bar{\lambda}^*, \bar{\lambda}^*/n_1) = 0$, where the last equality follows from the definition of $\bar{\lambda}^*$ in (7). As this is a contradiction with our assumption $J(\bar{\lambda}, y_1^*) = 0$, we conclude that, as a function of $\bar{\lambda}$, y_1^* is continuous at point $\bar{\lambda} = \bar{\lambda}^*$, and therefore over $[0, \infty)$. □

B.2 Proof of Proposition 1

Since all cheap links receive the same amount of traffic, as do all expensive links, we just need to prove Proposition 1 for an arbitrary cheap link, say link 1, and an arbitrary expensive link, say link N. We prove a series of technical lemmata, from which the proof of Proposition 1 directly follows.

Lemma 14. *It holds that $y_N^{ne} < y_1^{ne}$.*

Proof. With (10), $c_2 > c_1$ implies that $K\phi(y_1^{ne}) + y_1^{ne}\phi'(y_1^{ne}) > K\phi(y_N^{ne}) + y_N^{ne}\phi'(y_N^{ne})$. As the functions $\phi()$ and $\phi'()$ are increasing, this implies that $y_N^{ne} < y_1^{ne}$. □

Lemma 15. *It holds that*

$$c_1\phi\left(y_1^{ne}\right) < c_2\phi\left(y_N^{ne}\right) \iff c_1 y_1^{ne}\phi'\left(y_1^{ne}\right) > c_2 y_N^{ne}\phi'\left(y_N^{ne}\right) \tag{50}$$

Proof. It follows from (10) that

$$c_2\phi\left(y_N^{ne}\right) - c_1\phi\left(y_1^{ne}\right) = \frac{1}{K}\left\{c_1 y_1^{ne}\phi'\left(y_1^{ne}\right) - c_2 y_N^{ne}\phi'\left(y_N^{ne}\right)\right\},$$

from which the result follows. □

Lemma 16. *It holds that $c_1\phi\left(y_1^{ne}\right) < c_2\phi\left(y_N^{ne}\right)$.*

Proof. Assume on the contrary that $c_1\phi\left(y_1^{ne}\right) \geq c_2\phi\left(y_N^{ne}\right)$. With Lemma 15, it yields $c_1 y_1^{ne}\phi'\left(y_1^{ne}\right) \leq c_2 y_N^{ne}\phi'\left(y_N^{ne}\right)$. We therefore have

$$y_1^{ne}\frac{\phi'\left(y_1^{ne}\right)}{\phi\left(y_1^{ne}\right)} \leq y_N^{ne}\frac{\phi'\left(y_N^{ne}\right)}{\phi\left(y_N^{ne}\right)}. \tag{51}$$

As the function $w \to w\,\frac{\phi'(w)}{\phi(w)}$ is strictly increasing on $[0,\infty)$ for any strictly increasing and convex latency function $\phi(x)$ (see Lemma 11 in [6]), (51) implies that $y_N^{ne} \geq y_1^{ne}$. As $\phi()$ is increasing, we thus have $c_2\phi\left(y_N^{ne}\right) \leq c_1\phi\left(y_1^{ne}\right) \leq c_1\phi\left(y_N^{ne}\right)$, from which we conclude that $c_2 \leq c_1$. As it is assumed that $c_2 > c_1$, this is a contradiction. □

Lemma 17. *It holds that $F_1'(y_1^{ne}) \geq F_2'(y_N^{ne})$, and the inequality is strict for $K > 1$.*

Proof. From (10), we obtain $F_1'(y_1^{ne}) - F_2'(y_N^{ne}) = (K-1)\left[c_2\phi\left(y_N^{ne}\right) - c_1\phi\left(y_1^{ne}\right)\right]$, and the result follows from Lemma 16. □

We are now in position to prove Proposition 1.

Proof of Proposition 1. The proof of property (a) directly follows from Lemma 14. The proof of property (b) follows from Lemma 16 and Lemma 15. The inequality on the marginal costs stated in property (c) is proven in Lemma 17. □

B.3 Proof of Proposition 2

We assume throughout this section that $\bar{\lambda} > \bar{\lambda}^{ne}$. As for the proof of Proposition 1, it is enough to prove the results for an arbitrary cheap link and an arbitrary expensive links. We choose links 1 and N. We first establish some technical results which are required to prove Proposition 2. We first prove in Lemma 18 below that for $K > 1$, the decentralized routing strategy forwards more (resp. less) traffic on cheap (resp. low) links than the centralized one does.

Lemma 18. *For $K > 1$, it holds that $y_1^{ne} > y_1^*$ and $y_N^{ne} < y_N^*$.*

Proof. Assume that $y_1^{ne} \leq y_1^*$. As the function $F_1'()$ is increasing, this implies that $F_1'(y_1^{ne}) \leq F_1'(y_1^*)$. Besides, $y_N^{ne} = (\bar{\lambda} - n_1 y_1^{ne})/n_2 \geq (\bar{\lambda} - n_1 y_1^*)/n_2 = y_N^*$ implies that $F_2'(y_N^{ne}) \geq F_2'(y_N^*)$. As we know that $F_2'(y_N^*) = F_1'(y_1^*)$, it yields $F_1'(y_1^{ne}) \leq F_2'(y_N^{ne})$, which is a contradiction with statement (c) of Proposition 1 for $K > 1$. It follows that $y_1^{ne} > y_1^*$, and thus that $y_N^{ne} = (\bar{\lambda} - n_1 y_1^{ne})/n_2 < (\bar{\lambda} - n_1 y_1^*)/n_2 = y_N^*$. □

Lemma 19 below establishes the expression of $\frac{\partial y_1^{ne}}{\partial \bar{\lambda}}$, which is required to prove statements (b) and (c) of Proposition 2.

Lemma 19. *For $K \geq 1$ and all $\bar{\lambda} > \bar{\lambda}^{ne}$, we have*

$$\frac{\partial y_1^{ne}}{\partial \bar{\lambda}} = \frac{c_2\left[(K+1)\phi'\left(y_N^{ne}\right) + y_N^{ne}\,\phi''\left(y_N^{ne}\right)\right]}{n_2 c_1\left[(K+1)\phi'\left(y_1^{ne}\right) + y_1^{ne}\,\phi''\left(y_1^{ne}\right)\right] + n_1 c_2\left[(K+1)\phi'\left(y_N^{ne}\right) + y_N^{ne}\,\phi''\left(y_N^{ne}\right)\right]}. \tag{52}$$

Proof. Taking the derivative of (10) with respect to $\bar{\lambda}$, we obtain

$$c_1\left[(K+1)\phi'\left(y_1^{ne}\right) + y_1^{ne}\,\phi''\left(y_1^{ne}\right)\right]\frac{\partial y_1^{ne}}{\partial \bar{\lambda}} = c_2\left[(K+1)\phi'\left(y_N^{ne}\right) + y_N^{ne}\,\phi''\left(y_N^{ne}\right)\right]\frac{\partial y_N^{ne}}{\partial \bar{\lambda}}. \tag{53}$$

The equality $n_1 y_1^{ne} + n_2 y_N^{ne} = \bar{\lambda}$ implies that $\frac{\partial y_N^{ne}}{\partial \bar{\lambda}} = \frac{1}{n_2}\left(1 - n_1 \frac{\partial y_1^{ne}}{\partial \bar{\lambda}}\right)$. Combining the latter result with (53), the desired result follows after some algebra. □

We prove in Lemma 20 a property that we will use in Lemma 21.

Lemma 20. *For $K \geq 1$ and all $\bar{\lambda} > \bar{\lambda}^{ne}$, it holds that*

$$K = \frac{c_1 y_1^{ne}\phi'(y_1^{ne}) - c_2 y_N^{ne}\phi'(y_N^{ne})}{c_2\phi(y_N^{ne}) - c_1\phi(y_1^{ne})}. \tag{54}$$

Proof. From (10), we obtain

$$K\left[c_2\phi(y_N^{ne}) - c_1\phi(y_1^{ne})\right] = c_1 y_1^{ne}\phi'(y_1^{ne}) - c_2 y_N^{ne}\phi'(y_N^{ne}).$$

From statement (b) of Proposition 1 we know that $c_2\phi(y_N^{ne}) - c_1\phi(y_1^{ne}) > 0$, from which (54) follows. □

We now prove in Lemma 21 below that as $\bar{\lambda}$ increases in the interval $\left(\bar{\lambda}^{ne}, \infty\right)$, the difference $y_1^{ne} - y_1^*$ between the amount of flow forwarded on cheap link 1 by the decentralized routing strategy and the centralized one decreases.

Lemma 21. *For $K > 1$ and all $\bar{\lambda} > \bar{\lambda}^{ne}$, $\frac{dy_1^{ne}}{d\lambda} < \frac{dy_1^*}{d\lambda}$ and $\frac{dy_N^{ne}}{d\lambda} > \frac{dy_N^*}{d\lambda}$.*

Proof. Let $\psi(K, y) = (K+1)\phi'(y) + y\phi''(y)$. With Lemma 19, the condition $\frac{dy_1^{ne}}{d\lambda} < \frac{dy_1^*}{d\lambda}$ can be written as

$$\frac{c_2\psi(K, y_N^{ne})}{n_2 c_1\psi(K, y_1^{ne}) + n_1 c_2\psi(K, y_N^{ne})} < \frac{c_2\psi(1, y_N^*)}{n_2 c_1\psi(1, y_1^*) + n_1 c_2\psi(1, y_N^*)},$$

which is equivalent to

$$\frac{\psi(K, y_N^{ne})}{\psi(K, y_1^{ne})} < \frac{\psi(1, y_N^*)}{\psi(1, y_1^*)},$$

that is,

$$\frac{(K+1)\phi'(y_N^{ne}) + y_N^{ne}\phi''(y_N^{ne})}{(K+1)\phi'(y_1^{ne}) + y_1^{ne}\phi''(y_1^{ne})} < \frac{2\phi'(y_N^*) + y_N^*\phi''(y_N^*)}{2\phi'(y_1^*) + y_1^*\phi''(y_1^*)}. \tag{55}$$

Let $\mathbf{y}^{ne} = (y_1^{ne}, y_N^{ne})$ and $\mathbf{y}^* = (y_1^*, y_N^*)$. With Lemma 20, (55) can equivalently be written as $R(\mathbf{y}^{ne}) < R(\mathbf{y}^*)$, where the function $R(\mathbf{y})$ is defined as follows

$$R(\mathbf{y}) = \frac{\left(c_1 y_1\phi'(y_1) - c_2 y_N\phi'(y_N)\right)\phi'(y_N) + \left(\phi'(y_N) + y_N\phi''(y_N)\right)\left(c_2\phi(y_N^{ne}) - c_1\phi(y_1^{ne})\right)}{\left(c_1 y_1\phi'(y_1) - c_2 y_N\phi'(y_N)\right)\phi'(y_1) + \left(\phi'(y_1) + y_1\phi''(y_1)\right)\left(c_2\phi(y_N^{ne}) - c_1\phi(y_1^{ne})\right)}. \tag{56}$$

Consider now the vector-valued function $\mathbf{y}(t) = (y_1(t), y_N(t))$ defined by $y_1(t) = (1+t)y_1^*$ and $y_N(t) = y_N^* - t\frac{n_1}{n_2}y_1^*$ over the interval $\left[0, \frac{n_2 y_N^*}{n_1 y_1^*}\right]$. Observe that $n_1 y_1(t) + n_2 y_N(t) = \bar{\lambda}$ for all t in this interval and that $\mathbf{y}(0) = \mathbf{y}^*$. Moreover, defining $t^{ne} = \frac{n_2(y_N^* - y_N^{ne})}{n_1 y_1^*}$, we note from Lemma 18 that $t^{ne} > 0$ and $t^{ne} < \frac{n_2 y_N^*}{n_1 y_1^*}$, and we obtain that $\mathbf{y}(t^{ne}) = \mathbf{y}^{ne}$. We also note that as $y_1(t)$ increases with t and $y_N(t)$ decreases with t, $y_1(0) = y_1^* > y_N^* = y_N(0)$ implies that $y_1(t) > y_N(t)$ for all $t \leq t^{ne}$. Similarly, we know

that for $t \leq t^{ne}$ we have $y_1(t) < y_1^{ne}$ and $y_N(t) > y_N^{ne}$. Since ϕ is strictly increasing, it follows from statement (b) of Proposition 1 that

$$c_1\phi(y_1(t)) < c_1\phi(y_1^{ne}) < c_2\phi(y_N^{ne}) < c_2\phi(y_N(t)),$$

so that it holds that

$$y_N(t) < y_1(t) < \phi^{-1}\left(\frac{c_2}{c_1}\phi(y_N(t))\right) \quad \text{and } \phi^{-1}\left(\frac{c_1}{c_2}\phi(y_1(t))\right) < y_N(t) < y_1(t), \tag{57}$$

for all t in the interval $[0, t^{ne}]$.

In order to show that $R(\mathbf{y}^{ne}) < R(\mathbf{y}^*)$, and hence that inequality (55) holds, we just need to show that $R(\mathbf{y}(t))$ is a decreasing function of t. To this end, observe that $R(\mathbf{y})$ can be written as follows

$$\begin{aligned} R(\mathbf{y}) &= \frac{\phi'(y_N)}{\phi'(y_1)} \frac{c_1\left[y_1\phi'(y_1) - \phi(y_1)\left(1 + y_N\frac{\phi''(y_N)}{\phi'(y_N)}\right)\right] - c_2\left[y_N\phi'(y_N) - \phi(y_N)\left(1 + y_N\frac{\phi''(y_N)}{\phi'(y_N)}\right)\right]}{c_1\left[y_1\phi'(y_1) - \phi(y_1)\left(1 + y_1\frac{\phi''(y_1)}{\phi'(y_1)}\right)\right] - c_2\left[y_N\phi'(y_N) - \phi(y_N)\left(1 + y_1\frac{\phi''(y_1)}{\phi'(y_1)}\right)\right]} && (58) \\ &= \frac{\left[c_1H(y_1, y_N) - c_2H(y_N, y_N)\right]/\phi'(y_1)}{\left[c_1H(y_1, y_1) - c_2H(y_N, y_1)\right]/\phi'(y_N)}, && (59) \\ &= \frac{A(\mathbf{y})}{B(\mathbf{y})}, && (60) \end{aligned}$$

where the functions $H(x, y)$, $A(x, y)$ and $B(x, y)$ have been introduced in Assumption 4. We shall prove that $A(\mathbf{y}(t))$ is a strictly positive and decreasing function of t over $[0, t^{ne}]$, and that $B(\mathbf{y}(t))$ is a strictly positive and increasing function of t over $[0, t^{ne}]$, implying that $R(\mathbf{y}(t))$ decreases with t over $[0, t^{ne}]$.

We first prove that $A(\mathbf{y}(t)) > 0$ and $B(\mathbf{y}(t)) > 0$ for all $t \leq t^{ne}$. To this end, we first consider (56) and show that both the numerator and the denominator are strictly positive over $[0, t^{ne}]$. Indeed, the function $c_2\phi(y_N(t)) - c_1\phi(y_1(t))$ is strictly decreasing with t, which implies that $c_2\phi(y_N(t)) - c_1\phi(y_1(t)) \geq c_2\phi(y_N^{ne}) - c_1\phi(y_1^{ne})$ for all $t \leq t^{ne}$. With statement (b) of Proposition 11 we can thus conclude that $c_2\phi(y_N(t)) - c_1\phi(y_1(t)) > 0$ for all $t \leq t^{ne}$. Similarly, $c_1y_1(t)\phi'(y_1(t)) - c_2y_N(t)\phi'(y_N(t))$ is a strictly increasing function of t, and thus $c_1y_1(t)\phi'(y_1(t)) - c_2y_N(t)\phi'(y_N(t)) \geq c_1y_1^*\phi'(y_1^*) - c_2y_N^*\phi'(y_N^*)$ for all $t \leq t^{ne}$. With statement (b) of Proposition 11, it yields $c_1y_1(t)\phi'(y_1(t)) - c_2y_N(t)\phi'(y_N(t)) > 0$ for all $t \leq t^{ne}$. As all the other terms in the numerator and the denominator of (56) are strictly positive, this shows that both the numerator and the denominator of (56) are strictly positive. Considering the rewriting of $R(\boldsymbol{\rho})$ in (58)-(60), this clearly implies that $A(\mathbf{y}(t)) > 0$ and $B(\mathbf{y}(t)) > 0$ over $[0, t^{ne}]$.

We now show that $A(\mathbf{y}(t))$ is strictly decreasing with t over $[0, t^{ne}]$. We have

$$\frac{d}{dt}A(\mathbf{y}(t)) = y_1^*\left[\frac{\partial A}{\partial y_1}(\mathbf{y}(t)) - \frac{n_1}{n_2}\frac{\partial A}{\partial y_N}(\mathbf{y}(t))\right]. \tag{61}$$

It follows from (57) and Assumption 4 that $\frac{\partial A}{\partial y_1}(\mathbf{y}(t)) \leq 0$ and $\frac{\partial A}{\partial y_N}(\mathbf{y}(t)) \geq 0$, and that at least one of the inequalities is strict. With (61), it yields $\frac{d}{dt}A(\mathbf{y}(t)) < 0$, as claimed. Similarly, we have

$$\frac{d}{dt}B(\mathbf{y}(t)) = y_1^*\left[\frac{\partial B}{\partial y_1}(\mathbf{y}(t)) - \frac{n_1}{n_2}\frac{\partial B}{\partial y_N}(\mathbf{y}(t))\right]. \tag{62}$$

Since it follows from (57) and Assumption 4 that $\frac{\partial B}{\partial y_1}(\mathbf{y}(t)) \geq 0$ and $\frac{\partial B}{\partial y_N}(\mathbf{y}(t)) \leq 0$ with at least one of these two inequalities being strict, this proves that $B(\mathbf{y}(t))$ is strictly increasing with t over $[0, t^{ne}]$. To conclude, we just use (60) to obtain

$$\frac{d}{dt}R(\mathbf{y}(t)) = \frac{1}{B^2(\mathbf{y}(t))}\left\{\frac{d}{dt}A(\mathbf{y}(t))B(\mathbf{y}(t)) - A(\mathbf{y}(t))\frac{d}{dt}B(\mathbf{y}(t))\right\} < 0, \tag{63}$$

for all $t \leq t^{ne}$, from which it follows that $R(\mathbf{y}^{ne}) = R(\mathbf{y}(t^{ne})) < R(\mathbf{y}(0)) = R(\mathbf{y}^*)$. This implies that inequality (55) holds, and thus that $\frac{dy_1^{ne}}{d\lambda} < \frac{dy_1^*}{d\lambda}$ as claimed.

We now show that $\frac{dy_N^{ne}}{d\lambda} > \frac{dy_N^*}{d\lambda}$. From $n_1\frac{dy_1^{ne}}{d\lambda} + n_2\frac{dy_N^{ne}}{d\lambda} = n_1\frac{dy_1^*}{d\lambda} + n_2\frac{dy_N^*}{d\lambda}$, we obtain

$$n_2\left(\frac{dy_N^{ne}}{d\lambda} - \frac{dy_N^*}{d\lambda}\right) = n_1\left(\frac{dy_1^*}{d\lambda} - \frac{dy_1^{ne}}{d\lambda}\right) > 0, \tag{64}$$

which shows that $\frac{dy_N^{ne}}{d\lambda} > \frac{dy_N^*}{d\lambda}$. □

Proof of Proposition 2. The proof of property (a) directly follows from Lemma 18, whereas properties (b) and (c) follow from Lemma 21. □

B.4 Proof of Proposition 3

We first consider the behaviour of the link flow ratios when $\bar{\lambda}$ varies over the interval $\left(\bar{\lambda}^*, \bar{\lambda}^{ne}\right]$. Lemma 22 below proves that the ratio is increasing for cheap links, and equal to 0 for expensive links. It is worthwhile noticing that the proof is valid for all latency functions ϕ such that the function $f(y)$ defined in (18) is strictly increasing.

Lemma 22. *For $\bar{\lambda} \in \left(\bar{\lambda}^*, \bar{\lambda}^{ne}\right]$, the ratio y_j^{ne}/y_j^* is 0 for all expensive links $j \in \mathcal{S}_2$. In this interval, the ratio y_k^{ne}/y_k^* increases as $\bar{\lambda}$ increases for all cheap links $k \in \mathcal{S}_1$.*

Proof. In the interval $\left(\bar{\lambda}^*, \bar{\lambda}^{ne}\right]$, the decentralized strategy forwards all the traffic on the cheap links, from which it follows that $y_j^{ne} = 0$ for all $j \in \mathcal{S}_2$ in this interval, which proves the first statement of the lemma. Consider now an arbitrary cheap link, say link 1, and an arbitrary expensive link, say link N. Under the centralized routing strategy, the flows on these links are such that $F_1'(y_1^*) = F_2'(y_N^*)$ for all $\bar{\lambda}$ in $\left[\bar{\lambda}^*, \bar{\lambda}^{ne}\right)$. Hence, $F_1''(y_1^*)\frac{dy_1^*}{d\bar{\lambda}} = F_2''(y_N^*)\frac{dy_N^*}{d\bar{\lambda}}$, and we obtain

$$\frac{F_1''(y_1^*)}{F_1'(y_1^*)}\frac{dy_1^*}{d\bar{\lambda}} = \frac{F_2''(y_N^*)}{F_2'(y_N^*)}\frac{dy_N^*}{d\bar{\lambda}} \tag{65}$$

Note that it implies that both $\frac{dy_1^*}{d\bar{\lambda}}$ and $\frac{dy_N^*}{d\bar{\lambda}}$ are positive. The flow conservation constraint implies that $n_1 y_1^* + n_2 y_N^* = n_1 y_1^{ne} = \bar{\lambda}$. Dividing by y_1^* and taking the derivative with respect to $\bar{\lambda}$ and using (65), we obtain

$$\begin{aligned}
\frac{d}{d\bar{\lambda}}\left(\frac{y_1^{ne}}{y_1^*}\right) &= \frac{n_2}{n_1}\frac{d}{d\bar{\lambda}}\left(\frac{y_N^*}{y_1^*}\right), \\
&= \frac{n_2}{n_1(y_1^*)^2}\left[y_1^*\frac{dy_N^*}{d\bar{\lambda}} - y_N^*\frac{dy_1^*}{d\bar{\lambda}}\right], \\
&= \frac{n_2}{n_1(y_1^*)^2}\frac{F_2'(y_N^*)}{F_2''(y_N^*)}\left[y_1^*\frac{F_1''(y_1^*)}{F_1'(y_1^*)} - y_N^*\frac{F_2''(y_N^*)}{F_2'(y_N^*)}\right]\frac{dy_1^*}{d\bar{\lambda}}, \\
&= \frac{n_2}{n_1(y_1^*)^2}\frac{F_2'(y_N^*)}{F_2''(y_N^*)}\left[f(y_1^*) - f(y_N^*)\right]\frac{dy_1^*}{d\bar{\lambda}},
\end{aligned}$$

where the function $f(y)$ is defined in (18). With Assumption 3 and statement (a) in Proposition 1, we have $f(y_1^*) - f(y_N^*) > 0$, which proves that $\frac{d}{d\bar{\lambda}}\left(\frac{y_1^{ne}}{y_1^*}\right) > 0$. □

We now prove Proposition 3.

Proof of Proposition 3. For $\bar{\lambda}$ varying over the interval $\left(\bar{\lambda}^*, \bar{\lambda}^{ne}\right]$, the proof directly follows from Lemma 22. For $\bar{\lambda}$ varying over the interval $\left(\bar{\lambda}^{ne}, \infty\right)$, Proposition 2 implies that the ratio y_k^{ne}/y_k^* decreases as $\bar{\lambda}$ increases for all cheap links $k \in \mathcal{S}_1$. Indeed, we have

$$\frac{d}{d\bar{\lambda}}\left(\frac{y_k^{ne}}{y_k^*}\right) = \frac{1}{(y_k^*)^2}\left[\frac{dy_k^{ne}}{d\bar{\lambda}}y_k^* - y_k^{ne}\frac{dy_k^*}{d\bar{\lambda}}\right] < \frac{1}{(y_k^*)^2}\left[y_k^* - y_k^{ne}\right]\frac{dy_k^{ne}}{d\bar{\lambda}} < 0, \tag{66}$$

where the first (resp. second) inequality results from assertion (b) (resp. (a)) of Proposition 2. A similar argument shows that the ratio y_j^{ne}/y_j^* increases as $\bar{\lambda}$ increases for all expensive links $k \in \mathcal{S}_2$. □

C Proof of Results in Section 4.2

In this appendix, we prove Lemma 2 and Lemma 3, as well as Proposition 4.

C.1 Proof of Lemma 2

Proof of Lemma 2. Note that $D_1(\bar{\lambda}, \mathbf{p}) > 0$ for all $\bar{\lambda} > 0$. As the ratio of two continuous functions is continuous at all points where the denominator does not vanishes, it is enough to show that $D_K\left(\frac{\bar{\lambda}}{K}\mathbf{1}, \mathbf{p}\right)$ and $D_1(\bar{\lambda}, \mathbf{p})$ are continuous functions of $\bar{\lambda}$ over $(0, \infty)$. As $D_1(\bar{\lambda}, \mathbf{p}) = F(\mathbf{y}^*)$ and F is a continuous function of the vector $\mathbf{y}^*$ of link flows obtained under the centralized routing scheme, the continuity of $D_1(\bar{\lambda}, \mathbf{p})$ follows from the continuity of $\mathbf{y}^*$ as a function of $\bar{\lambda}$ over $(0, \infty)$, which is proven in Lemma 1. Similarly, the continuity of $D_K\left(\frac{\bar{\lambda}}{K}\mathbf{1}, \mathbf{p}\right) = F(\mathbf{y}^{ne})$ follows from the continuity of $\mathbf{y}^{ne}$ as a function of $\bar{\lambda}$ over $(0, \infty)$, which is also proven in Lemma 1. □

C.2 Proof of Proposition 4

Proof of Proposition 4. We first prove assertion (a). It readily follows from $\bar{\lambda}^* < \bar{\lambda}^{ne}$ that for $\bar{\lambda} \leq \bar{\lambda}^*$

$$\frac{D_K\left(\frac{\bar{\lambda}}{K}\mathbf{1}, \mathbf{p}\right)}{D_1(\bar{\lambda}, \mathbf{p})} = \frac{F(\mathbf{y}^{ne})}{F(\mathbf{y}^*)} = \frac{n_1 F_1(\frac{\bar{\lambda}}{n_1})}{n_1 F_1(\frac{\bar{\lambda}}{n_1})} = 1.$$

We now consider assertion (b) and assume that $\bar{\lambda} \in \left(\bar{\lambda}^*, \bar{\lambda}^{ne}\right]$. Since for $\bar{\lambda} \leq \bar{\lambda}^{ne}$ we have $\mathbf{y}^{ne} = \left(\frac{\bar{\lambda}}{n_1}, \ldots, \frac{\bar{\lambda}}{n_1}, 0, \ldots, 0\right)$, the ratio $D_K\left(\frac{\bar{\lambda}}{K}\mathbf{1}, \mathbf{p}\right)/D_1(\bar{\lambda}, \mathbf{p})$ is strictly increasing in $\bar{\lambda}$ if and only if

$$F_1'\left(\frac{\bar{\lambda}}{n_1}\right) F(\mathbf{y}^*) > \left(n_1 F_1'(y_1^*)\frac{dy_1^*}{d\bar{\lambda}} + n_2 F_2'(y_N^*)\frac{dy_N^*}{d\bar{\lambda}}\right) n_1 F_1(\frac{\bar{\lambda}}{n_1}). \tag{67}$$

The constraint $n_1 y_1^* + n_2 y_N^* = \bar{\lambda}$ implies that $n_1 \frac{dy_1^*}{d\bar{\lambda}} + n_2 \frac{dy_N^*}{d\bar{\lambda}} = 1$. Moreover, $\bar{\lambda} > \bar{\lambda}^*$ implies that $F_1'(y_1^*) = F_2'(y_N^*)$. It yields $n_1 F_1'(y_1^*)\frac{dy_1^*}{d\bar{\lambda}} + n_2 F_2'(y_N^*)\frac{dy_N^*}{d\bar{\lambda}} = F_1'(y_1^*)$. Therefore, the ratio $D_K\left(\frac{\bar{\lambda}}{K}\mathbf{1}, \mathbf{p}\right)/D_1(\bar{\lambda}, \mathbf{p})$ is strictly increasing in $\bar{\lambda}$ if and only if

$$F_1'\left(\frac{\bar{\lambda}}{n_1}\right)[n_1 F_1(y_1^*) + n_2 F_2(y_N^*)] > n_1 F_1'(y_1^*) F_1\left(\frac{\bar{\lambda}}{n_1}\right),$$

which can equivalently be written as

$$n_1 \frac{F_1(y_1^*)}{F_1'(y_1^*)} + n_2 \frac{F_2(y_N^*)}{F_2'(y_N^*)} > n_1 \frac{F_1(\frac{\bar{\lambda}}{n_1})}{F_1'(\frac{\bar{\lambda}}{n_1})},$$

where we have used the equality $F_1'(y_1^*) = F_2'(y_N^*)$. Observing that $\frac{F_1(y_1^*)}{F_1'(y_1^*)} = g(y_1^*)$ and $\frac{F_2(y_N^*)}{F_2'(y_N^*)} = g(y_N^*)$, where the function $g()$ is defined by (17) in Appendix A, it is therefore enough to show that

$$n_1 g(y_1^*) + n_2 g(y_N^*) > n_1 g\left(\frac{\bar{\lambda}}{n_1}\right). \tag{68}$$

To this end, let us consider the function $h(t) = n_1 g\left(\frac{t\bar{\lambda}}{n_1}\right) + n_2 g\left(\frac{(1-t)\bar{\lambda}}{n_2}\right)$. We have

$$h'(t) = \bar{\lambda} \times \left(g'\left(\frac{t\bar{\lambda}}{n_1}\right) - g'\left(\frac{(1-t)\bar{\lambda}}{n_2}\right)\right),$$

which implies that $h'(t) < 0$ for $t \in \left[\frac{n_1}{n_1+n_2}, 1\right]$ because g is strictly concave for $\phi(x) = e^x$ and $\phi(x) = (1+x)^m$ (see Assumption 2 in Appendix A). We know from Proposition 1 that $y_1^* > y_N^*$. Together with $n_1 y_1^* + n_2 y_N^* = \bar{\lambda}$, it implies that $\frac{y_1^*}{\bar{\lambda}} > \frac{1}{n_1+n_2}$. Taking $t = \frac{n_1 y_1^*}{\bar{\lambda}} < 1$, we hence obtain that $h(1) < h(t)$, which proves (68). We thus conclude that the ratio $D_K\left(\frac{\bar{\lambda}}{K}\mathbf{1}, \mathbf{p}\right)/D_1(\bar{\lambda}, \mathbf{p})$ is strictly increasing in $\bar{\lambda}$ over the interval $\left(\bar{\lambda}^*, \bar{\lambda}^{ne}\right]$.

Finally, we focus on assertion (c) assuming that $\bar{\lambda} > \bar{\lambda}^{ne}$. The ratio $D_K\left(\frac{\bar{\lambda}}{K}\mathbf{1}, \mathbf{p}\right)/D_1(\bar{\lambda}, \mathbf{p})$ is strictly decreasing in $\bar{\lambda}$ if and only if

$$\left\{n_1 F_1'(y_1^{ne})\frac{dy_1^{ne}}{d\bar{\lambda}} + n_2 F_2'(y_N^{ne})\frac{dy_N^{ne}}{d\bar{\lambda}}\right\} F(\mathbf{y}^*) < \left\{n_1 F_1'(y_1^*)\frac{dy_1^*}{d\bar{\lambda}} + n_2 F_2'(y_N*)\frac{dy_N^*}{d\bar{\lambda}}\right\} F(\mathbf{y}^{ne}). \tag{69}$$

From Proposition 2, we know that $\frac{dy_1^{ne}}{d\bar{\lambda}} < \frac{dy_1^*}{d\bar{\lambda}}$. Let $\delta = n_1 \frac{dy_1^*}{d\bar{\lambda}}$. We have

$$\begin{aligned} n_1 F_1'(y_1^{ne}) \frac{dy_1^{ne}}{d\bar{\lambda}} + n_2 F_2'(y_N^{ne}) \frac{dy_N^{ne}}{d\bar{\lambda}} &= n_1 F_1'(y_1^{ne}) \frac{dy_1^{ne}}{d\bar{\lambda}} + n_2 F_2'(y_N^{ne}) \frac{1}{n_2} \left(1 - n_1 \frac{dy_1^{ne}}{d\bar{\lambda}}\right) \\ &= n_1 \left(F_1'(y_1^{ne}) - F_2'(y_N^{ne})\right) \frac{dy_1^{ne}}{d\bar{\lambda}} + F_2'(y_N^{ne}) \\ &< \delta \left(F_1'(y_1^{ne}) - F_2'(y_N^{ne})\right) + F_2'(y_N^{ne}) \\ &= \delta F_1'(y_1^{ne}) + (1-\delta) F_2'(y_N^{ne}), \end{aligned}$$

where we have used the inequality $F_1'(y_1^{ne}) > F_2'(y_N^{ne})$ which is proven in Proposition 1. As a consequence, a sufficient condition for (69) to hold is that

$$\{\delta F_1'(y_1^{ne}) + (1-\delta) F_2'(y_N^{ne})\}\ F(\mathbf{y}^*) < \{\delta F_1'(y_1^*) + (1-\delta) F_2'(y_N^*)\}\ F(\mathbf{y}^{ne}), \tag{70}$$

which can equivalently be written as follows

$$\frac{F(\mathbf{y}^*)}{\delta F_1'(y_1^*) + (1-\delta) F_2'(y_N^*)} < \frac{F(\mathbf{y}^{ne})}{\delta F_1'(y_1^{ne}) + (1-\delta) F_2'(y_N^{ne})}, \tag{71}$$

that is $Q(0) < Q(\Delta^{ne})$. Assumption 5 completes the proof. □

C.3 Proof of Lemma 3

Proof of Lemma 3. For $\phi(x) = (1+x)^m$, (9) yields

$$\left(1 + \frac{\bar{\lambda}^{ne}}{n_1}\right)^m \left[1 + \frac{m}{K} \frac{\bar{\lambda}^{ne}/n_1}{1 + \bar{\lambda}^{ne}/n_1}\right] = \frac{c_2}{c_1}. \tag{72}$$

Differentiating on both sides with respect to m, we obtain after some algebra

$$\frac{\partial}{\partial m}\left(\frac{\bar{\lambda}^{ne}}{n_1}\right) \frac{m}{1+\bar{\lambda}^{ne}/n_1} \left[1 + \frac{1 + m\bar{\lambda}^{ne}/n_1}{K(1+\bar{\lambda}^{ne}/n_1)}\right] = -\log\left(1 + \frac{\bar{\lambda}^{ne}}{n_1}\right) \left[1 + \frac{m}{K} \frac{\bar{\lambda}^{ne}/n_1}{1+\bar{\lambda}^{ne}/n_1}\right] - \frac{1}{K} \frac{\bar{\lambda}^{ne}/n_1}{1+\bar{\lambda}^{ne}/n_1}.$$

As the RHS of the above expression is negative and $\frac{\partial}{\partial m}\left(\frac{\bar{\lambda}^{ne}}{n_1}\right)$ is multiplied by a positive term, we conclude that $\bar{\lambda}^{ne}/n_1$ decreases with m. Since n_1 is a constant, this implies that $\bar{\lambda}^{ne}$ decreases with m. Assuming that there exists $\epsilon > 0$ such that $\bar{\lambda}^{ne}/n_1 > \epsilon$ for some m as large as we want, we get a contradiction with (72). We thus conclude that $\bar{\lambda}^{ne} \to 0$ as $m \to \infty$. □

D Proof of Results in Section 5

We give below the proofs of Proposition 5, Proposition 6 and Lemma 4.

Proof of Proposition 5. Assume that $\bar{\lambda} = \bar{\lambda}^{ne}$. It follows from (15) that

$$I_K(\mathbf{p}) = \frac{n_1 c_1 y_1^{ne} \phi\left(y_1^{ne}\right)}{n_1 c_1 y_1^* \phi\left(y_1^*\right) + n_2 c_2 y_N^* \phi\left(y_N^*\right)}, \tag{73}$$

$$= \frac{y_1^{ne} \phi\left(y_1^{ne}\right)}{y_1^* \phi\left(y_1^*\right) + \frac{1}{\alpha\gamma} y_N^* \phi\left(y_N^*\right)}. \tag{74}$$

Note that the constraint $n_1 y_1^* + n_2 y_N^* = \bar{\lambda}^{ne}$ implies that $y_1^* + \frac{y_N^*}{\alpha} = y_1^{ne}$. It yields

$$I_K(\mathbf{p}) = \frac{y_1^{ne} \phi\left(y_1^{ne}\right)}{y_1^* \phi\left(y_1^*\right) + \frac{1}{\gamma}(y_1^{ne} - y_1^*) \phi\left(\alpha(y_1^{ne} - y_1^*)\right)}. \tag{75}$$

As a consequence, to show that the inefficiency $I_K(\mathbf{p})$ depends only on the ratios α and γ, it is enough to show that y_1^{ne} and y_1^* depend only on these ratios. From (9), we get that

$$\phi\left(y_1^{ne}\right)+\frac{1}{K}y_1^{ne}\phi'\left(y_1^{ne}\right)=\frac{1}{\gamma}, \tag{76}$$

which shows that y_1^{ne} is a function of γ only. Similarly, (8) yields

$$\begin{aligned}\phi\left(y_1^*\right)+y_1^*\,\phi'\left(y_1^*\right) &= \frac{1}{\gamma}\left[\phi\left(y_N^*\right)+y_N^*\,\phi'\left(y_N^*\right)\right], &(77)\\ &= \frac{1}{\gamma}\left[\phi\left(\alpha(y_1^{ne}-y_1^*)\right)+\alpha(y_1^{ne}-y_1^*)\,\phi'\left(\alpha(y_1^{ne}-y_1^*)\right)\right], &(78)\end{aligned}$$

which shows that y_1^* is a function of the ratios α and γ. The desired result follows. □

It follows from Proposition 5 that the *Inefficiency* depends only on the ratios α and γ. As a consequence, in the proof of Proposition 6 which is given below, we shall write $I_K(\alpha,\gamma)$ to denote the *Inefficiency*.

Proof of Proposition 6. We first note from (76) that y_1^{ne} does not depend on α, which implies that the numerator of $I_K(\alpha,\gamma)$ in (74) is constant when α varies. As a consequence, it is enough to prove that the denominator of $I_K(\alpha,\gamma)$ is increasing with α, that is, that

$$D=y_1^*\phi\left(y_1^*\right)+\frac{1}{\gamma}(y_1^{ne}-y_1^*)\phi\left(\alpha(y_1^{ne}-y_1^*)\right),$$

increases when α increases. Taking the derivative with respect to α and using the identity $y_N^*=\alpha(y_1^{ne}-y_1^*)$ to simplify notations, we obtain

$$\begin{aligned}\frac{\partial D}{\partial\alpha} &= \left[\phi\left(y_1^*\right)+y_1^*\,\phi'\left(y_1^*\right)\right]\frac{\partial y_1^*}{\partial\alpha}+\frac{1}{\gamma}\Bigg[-\phi\left(\alpha(y_1^{ne}-y_1^*)\right)\frac{\partial y_1^*}{\partial\alpha}+\\ &\quad (y_1^{ne}-y_1^*)\phi'\left(\alpha(y_1^{ne}-y_1^*)\right)\left((y_1^{ne}-y_1^*)-\alpha\frac{\partial y_1^*}{\partial\alpha}\right)\Bigg],\\ &= \frac{1}{\gamma}(y_1^{ne}-y_1^*)^2\phi'\left(y_N^*\right)+\\ &\quad \left[\left(\phi\left(y_1^*\right)+y_1^*\,\phi'\left(y_1^*\right)\right)-\frac{1}{\gamma}\left(\phi\left(y_N^*\right)+y_N^*\,\phi'\left(y_N^*\right)\right)\right]\frac{\partial y_1^*}{\partial\alpha},\\ &= \frac{1}{\gamma}(y_1^{ne}-y_1^*)^2\phi'\left(y_N^*\right),\end{aligned}$$

where the last equality follows from (77). Since ϕ is strictly increasing, it follows that $\frac{\partial D}{\partial\alpha}>0$, implying that $I_K(\alpha,\gamma)$ is strictly decreasing with α, as claimed. □

Proof of Lemma 4. Assume that $n_1=1$ and $n_2=N-1$. For $\phi(x)=e^{\nu x}$, we know that $\bar\lambda^*$ is defined by $\gamma(1+\nu\bar\lambda^*)e^{\nu\bar\lambda^*}=1$. This implies that $\bar\lambda^*\to\infty$ as $\gamma\to 0$. As a consequence,

$$\lim_{\gamma\to 0}\gamma\nu\bar\lambda^*e^{\nu\bar\lambda^*}=\lim_{\gamma\to 0}\gamma(1+\nu\bar\lambda^*)e^{\nu\bar\lambda^*}\frac{\nu\bar\lambda^*}{1+\nu\bar\lambda^*}=\lim_{\gamma\to 0}\frac{\nu\bar\lambda^*}{1+\nu\bar\lambda^*}=1.$$

It implies that for any $\epsilon>0$ and for γ sufficiently close to 0, we have $(1-\epsilon)/\gamma<\nu\bar\lambda^*e^{\nu\bar\lambda^*}<(1+\epsilon)/\gamma$, that is, $W_0\left((1-\epsilon)/\gamma\right)<\nu\bar\lambda^*<W_0\left((1+\epsilon)/\gamma\right)$, where W_0 is the principal branch of the Lambert W function. As it is known that $W_0(x)=\log(x)-\log(\log(x))+o(1)$ for large values of x, we obtain that

$$\log\left(\frac{1-\epsilon}{\gamma}\right)-\log\left(\log\left(\frac{1-\epsilon}{\gamma}\right)\right)+o(1)<\nu\bar\lambda^*<\log\left(\frac{1+\epsilon}{\gamma}\right)-\log\left(\log\left(\frac{1+\epsilon}{\gamma}\right)\right)+o(1),$$

for any $\epsilon>0$ and for γ sufficiently close to 0.

Similarly, it follows from $\bar\lambda^{ne}>\bar\lambda^*$ that $\bar\lambda^{ne}\to\infty$ as $\gamma\to 0$. Since $\bar\lambda^{ne}$ is defined by $\gamma(1+\frac{\nu\bar\lambda^{ne}}{K})e^{\nu\bar\lambda^{ne}}=1$, we have

$$\lim_{\gamma\to 0}\gamma\frac{\nu\bar\lambda^{ne}}{K}e^{\nu\bar\lambda^{ne}}=\lim_{\gamma\to 0}\gamma(1+\frac{\nu\bar\lambda^{ne}}{K})e^{\bar\lambda^{ne}}\frac{\nu\bar\lambda^{ne}/K}{1+\nu\bar\lambda^{ne}/K}=\lim_{\gamma\to 0}\frac{\nu\bar\lambda^{ne}}{K+\nu\bar\lambda^{ne}}=1.$$

and we can show with a similar reasoning that

$$\log\left(K\frac{1-\epsilon}{\gamma}\right)-\log\left(\log\left(K\frac{1-\epsilon}{\gamma}\right)\right)+o(1)<\nu\bar{\lambda}^{ne}<\log\left(K\frac{1+\epsilon}{\gamma}\right)-\log\left(\log\left(K\frac{1+\epsilon}{\gamma}\right)\right)+o(1),$$

for any $\epsilon>0$ and for γ sufficiently close to 0. With the above bounds on $\bar{\lambda}^*$ and $\bar{\lambda}^{ne}$, we get

$$\nu(\bar{\lambda}^{ne}-\bar{\lambda}^*)<\log K+\log\left(\frac{1+\epsilon}{1-\epsilon}\right)+\log\left(\frac{\log\left(\frac{1-\epsilon}{\gamma}\right)}{\log K+\log\left(\frac{1+\epsilon}{\gamma}\right)}\right),$$

and

$$\nu(\bar{\lambda}^{ne}-\bar{\lambda}^*)>\log K+\log\left(\frac{1-\epsilon}{1+\epsilon}\right)+\log\left(\frac{\log\left(\frac{1+\epsilon}{\gamma}\right)}{\log K+\log\left(\frac{1-\epsilon}{\gamma}\right)}\right),$$

for any $\epsilon>0$ and for γ sufficiently close to 0. We thus conclude that

$$\lim_{\gamma\to 0}\left(\bar{\lambda}^{ne}-\bar{\lambda}^*\right)=\frac{1}{\nu}\log(K). \tag{79}$$

Consider now the ratio of social costs at $\bar{\lambda}=\bar{\lambda}^{ne}$, which is by definition

$$I_K\left(\frac{1}{N-1},\gamma\right)=\frac{F_1\left(\bar{\lambda}^{ne}\right)}{F_1\left(y_1^*\right)+(N-1)F_2\left(y_2^*\right)}=\frac{\gamma\bar{\lambda}^{ne}e^{\nu\bar{\lambda}^{ne}}}{\gamma y_1^*e^{\nu y_1^*}+(N-1)y_2^*e^{\nu y_2^*}}, \tag{80}$$

where y_1^* and y_2^* are such that $y_1^*+(N-1)y_2^*=\bar{\lambda}^{ne}$ and $\gamma(1+\nu y_1^*)e^{\nu y_1^*}=(1+\nu y_2^*)e^{\nu y_2^*}$. It then follows from $\lim_{\gamma\to 0}\gamma\bar{\lambda}^{ne}e^{\bar{\lambda}^{ne}}=K/\nu$ that

$$\lim_{\gamma\to 0}I_K\left(\frac{1}{N-1},\gamma\right)=\frac{K}{\lim_{\gamma\to 0}\left\{\gamma\nu y_1^*e^{\nu y_1^*}+(N-1)\nu y_2^*e^{\nu y_2^*}\right\}}. \tag{81}$$

Note that

$$\begin{aligned}\gamma\nu y_1^*e^{\nu y_1^*}+(N-1)\nu y_2^*e^{\nu y_2^*}&=\gamma(1+\nu y_1^*)e^{\nu y_1^*}\frac{\nu y_1^*}{1+\nu y_1^*}+(N-1)\nu y_2^*e^{\nu y_2^*},\\&=(1+\nu y_2^*)e^{\nu y_2^*}\frac{\nu y_1^*}{1+\nu y_1^*}+(N-1)\nu y_2^*e^{\nu y_2^*},\end{aligned}$$

and as $y_1^*>\bar{\lambda}^*$ implies that $y_1^*\to\infty$ as $\gamma\to 0$, we deduce that

$$\lim_{\gamma\to 0}\left\{\gamma\nu y_1^*e^{y_1^*}+(N-1)\nu y_2^*e^{y_2^*}\right\}=\lim_{\gamma\to 0}\left\{(1+N\nu y_2^*)e^{\nu y_2^*}\right\}.$$

With (81), it yields

$$\lim_{\gamma\to 0}I_K\left(\frac{1}{N-1},\gamma\right)=\frac{K}{\lim_{\gamma\to 0}\left\{(1+N\nu y_2^*)e^{\nu y_2^*}\right\}}. \tag{82}$$

We know that $y_2^*>0$ and that $y_1^*>\bar{\lambda}^*$. With $y_2^*=\left(\bar{\lambda}^{ne}-y_1^*\right)/(N-1)$, it implies that

$$0<y_2^*<\frac{\bar{\lambda}^{ne}-\bar{\lambda}^*}{N-1},$$

for any value of $\gamma<1$, and therefore that

$$1<(1+N\nu y_2^*)e^{\nu y_2^*}<\left[1+N\nu\frac{\bar{\lambda}^{ne}-\bar{\lambda}^*}{N-1}\right]\exp\left(\nu\frac{\bar{\lambda}^{ne}-\bar{\lambda}^*}{N-1}\right). \tag{83}$$

With (79) it yields

$$1 \leq \lim_{\gamma \to 0} (1 + N\nu y_2^*) e^{\nu y_2^*} \leq K^{\frac{\nu}{N-1}} \left[1 + \frac{N}{N-1} \log K \right]. \tag{84}$$

Combining (84) and (82), we get

$$K^{\frac{N-1-\nu}{N-1}} \frac{N-1}{N(1+\log K)-1} \leq \lim_{\gamma \to 0} I_K \left(\frac{1}{N-1}, \gamma \right) \leq K. \tag{85}$$

For $N \to \infty$, it yields $K/(1+\log(K)) \leq \lim_{N\to\infty} PoA(K,N)$. □