Inter-Session Network Coding with Strategic Users: A Game-Theoretic Analysis of Network Coding

Amir-Hamed Mohsenian-Rad, Jianwei Huang, Vincent W.S. Wong,

Sidharth Jaggi, and Robert Schober

Abstract

A common assumption in the existing network coding literature is that the users are *cooperative* and do *not* pursue their own interests. However, this assumption can be violated in practice. In this paper, we analyze *inter-session network coding* in a wired network using game theory. We assume that the users are *selfish* and act as *strategic* players to maximize their own utility, which leads to a resource allocation *game* among users. In particular, we study network coding with strategic users for the well-known *butterfly* network topology where a bottleneck link is shared by several network coding and routing flows. We prove the existence of a Nash equilibrium for a wide range of utility functions. We also show that the number of Nash equilibria can be large (even *infinite*) for certain choices of system parameters. This is in sharp contrast to a similar game setting with traditional packet forwarding where the Nash equilibrium is always unique. We then characterize the worst-case efficiency bounds, i.e., the *Price-of-Anarchy* (PoA), compared to an *optimal* and *cooperative* network design. We show that by using a novel *discriminatory pricing* scheme which charges encoded and forwarded packets differently, we can improve the PoA in comparison with the case where a *single* pricing scheme is being used. However, regardless of the discriminatory pricing scheme being used, the PoA is still worse than for the case when network coding is not applied. This implies that, although inter-session network coding can improve performance compared to ordinary routing, it is significantly *more sensitive* to users' strategic behavior. For example, in a butterfly network where the side links have *zero* cost, the efficiency at certain Nash equilibria can be as low as 25%. If the side links have *non-zero* cost, then the efficiency at some Nash equilibria can further reduce to only 20%. These results generalize the well-known result of guaranteed 67% worst-case efficiency for traditional packet forwarding networks.

Keywords: Inter-session network coding, butterfly network, resource management, game theory, Nash equilibrium, price-of-anarchy, efficiency bound, convex optimization, network surplus maximization.

I. INTRODUCTION

Since the seminal paper by Ahlswede *et al.* [1], a rich body of work has been reported on how network coding can improve performance in both wired and wireless networks [2]–[5]. Network coding can be performed by *jointly* encoding multiple packets either from the *same* user or from *different* users. The former is called *intra-session* network coding [1], [2] while the latter is called *inter-session* network coding [3]–[5]. A common assumption in most network coding schemes in the literature is that the users are *cooperative* and do *not* pursue their own interests. However, this assumption can be violated in practice. Therefore, assuming that the users are *selfish* and *strategic*, in this paper we ask the following key questions: (a) *What is the impact of users' strategic behavior on network performance?* (b) *How does this impact change with different pricing schemes that we may potentially choose for each link?*

It is widely accepted that *pricing* is an effective approach in terms of improving the efficiency of network resource allocation, especially in *distributed* settings. In [6], Kelly *et al.* showed that if users are *price takers* (i.e., they treat network prices as *fixed*), efficient resource allocation can be achieved by properly setting *congestion prices* on each of the shared links. Recently, Johari *et al.* studied how the results on efficiency can change in both capacity-constrained [7] and capacity-unconstrained [8] networks if users are *price anticipators* who realize that the price is directly impacted by each individual user's behavior. In this case, users play a *game* with each other, and the efficiency of resource allocation is characterized by the Nash equilibrium of the game. A key performance metric is the *Price-of-Anarchy (PoA)*, which measures the *worst-case efficiency loss* at a Nash equilibrium due to users' price anticipating behavior. The PoA is equal to 1 if there is *no* efficiency loss. A smaller PoA denotes a higher efficiency loss. Other recent work on resource allocation games and the PoA include [9]–[15]. To the best of our knowledge, none of the previous works along this line study price anticipation in *network coding* systems.

The game theoretic analysis of network coding has received limited attention in the literature, e.g., in [16]–[20]. All results in [16]–[20] focus on the case of *intra-session* network coding, whereas we consider *inter-session* network coding in this paper. In [21], a game theoretic analysis for *inter-session* network coding of *unicast* flows in a single bottleneck link is considered. It is shown that in some classes of *two-user* networks, it is possible to use a rate allocation mechanism to enforce cooperation among users. In this paper, we assume that there are $N \geq 2$ users, two

of which use network coding while the rest only use routing. This helps us to better understand the interaction between network coding and routing flows. In fact, we show that the performance degrades when both network coding and routing sessions share the same link. Our results are also different from those in [21] since we consider the capacity-unconstrained case instead of the capacity-constrained case as in [21]. In fact, due to the focus on the capacity region of the network coding scheme, the work in [21] did not consider the impact of the *utility functions* of the users, the cost of the side links, price anticipation, price discrimination, and the PoA.

The key contributions of this paper are as follows:

- *New problem formulation*: We formulate the problem of maximizing the *network aggregate surplus*, i.e., the total utility of all users minus the network cost, for inter-session network coding. As far as we know, such a problem has not been studied in the literature before.

- *Innovative pricing schemes*: We consider two pricing schemes: *non-discriminatory pricing* and *discriminatory pricing*. The first one is the traditional approach in networks with routing-only users. It charges all packets with the same price. The second pricing scheme is a novel generalization of the first one where the encoded and forwarded packets are charged with different prices. We show that due to the special properties of network coding, discriminatory pricing is more reasonable in terms of reflecting the *actual load* generated by each user.

- *Characterization of Nash equilibria*: We prove that the existence of a Nash equilibrium for the formulated game is always guaranteed; however, there can be *many* (even *infinite*) Nash equilibria in the resource allocation game with inter-session network coding.

- *Calculation of the PoA in a butterfly network with zero-cost side links*: We show that, among the two aforementioned pricing schemes, a properly chosen discriminatory pricing leads to a better PoA compared with the non-discriminatory approach. We also show that the PoA is always smaller (i.e., worse) compared with the case without network coding. In particular, at certain Nash equilibria, the PoA can be as low as 25%, which is less than the well-known result of guaranteed 67% worst-case efficiency in [8] for packet forwarding networks.

- *Calculation of the PoA in a butterfly network with non-zero-cost side links*: We further show that if the side links in the studied butterfly network topology have non-zero cost, then the PoA can further reduce to only 20%. This occurs due to the fact that in this case *none* of the users have an incentive to participate in network coding. This implies that if the users have strategic behavior, then it is important to design mechanisms to encourage users to perform

network coding; otherwise, we cannot benefit from the advantages of network coding.

The key results of this paper together with a comparison with the related state-of-the-art results *without* considering network coding in [8] are summarized in Table I.

The rest of the paper is organized as follows. In Section II, we review some recent results on resource allocation games with routing. In Section III, we extend those results to the case when two users can jointly perform inter-session network coding in a butterfly network where the side links have zero cost. We further extend our analysis to the case where the side links have non-zero cost in Section IV. Conclusions and future work are discussed in Section V. The key notations we used in this paper are summarized in Table II.

II. BACKGROUND: RESOURCE ALLOCATION GAME WITH ROUTING FLOWS

In this section, we consider a resource allocation game in which multiple end-to-end users compete to send their packets through a single shared link as in Fig. 1. By construction, no inter-session network coding is performed in this case. This is a well-known problem which has been widely studied in [6]–[11]. Here, we summarize the key results in [8], which present the proper terminology and serve as a benchmark for our later discussions.

In Fig. 1, a set of users $\mathcal{N} = \{1, \ldots, N\}$ *shares* the bottleneck link (i, j) between nodes i and j. All packets that arrive at node i are simply *forwarded* to node j through link (i, j). For each user $n \in \mathcal{N}$, we denote the transmitter and receiver nodes by s_n and t_n, respectively. Let x_n denote the transmission rate of user $n \in \mathcal{N}$. We assume that each user $n \in \mathcal{N}$ has a *utility function* U_n, representing its degree of satisfaction based on its achievable data rate x_n. On the other hand, the shared link has a *cost function* C, which depends on the total rate (i.e., $\sum_{n \in \mathcal{N}} x_n$). As in [8]–[10], we make the following common assumptions throughout this paper:

Assumption 1 (Users' Utility Functions): For each user $n \in \mathcal{N}$, the utility function $U_n(x_n)$ is *concave*, *non-negative*, *increasing*, and *differentiable*.

Assumption 2 (Link Cost and Price Functions): There exists a *differentiable*, *convex*, and *non-decreasing* function $p(q)$ over $q \geq 0$, with $p(0) \geq 0$ and $p(q) \to \infty$ as $q \to \infty$, such that for each $q \geq 0$, the cost is modeled as $C(q) = \int_0^q p(z) \, dz$. Here, $C(q)$ is *convex* and *non-decreasing*. In particular, we assume that there exists $a > 0$ such that $p(q) = aq$ and $C(q) = \int_0^q a \, z \, dz = \frac{a}{2} q^2$. That is, the cost function $C(q)$ is *quadratic* and the price function $p(q)$ is *linear*. Notice that

linear price functions are the *only* price functions that satisfy the four well-known axioms of *rescaling*, *consistency*, *additivity*, and *positivity* for cost-sharing systems[1] [9], [22].

Assumption 1 is often used to model applications with *elastic* traffic, e.g., for remote file transfer using the file transfer protocol (FTP) [6]. Examples of utility functions which satisfy Assumption 1 include the well-known class of α-fair utility functions with $\alpha \in (0,1)$ [23]. Assumption 2 is also a common assumption in the network resource management literature (cf. [9], [14], [24]). In practice, the cost function C may reflect the *actual cost* (e.g., in dollars) of transmitting units of data over link (i,j) or simply the *delay* that the packets experience over link (i,j). The more the aggregate data on the link, the higher is the average queueing delay.

Let $\boldsymbol{x} = (x_1, \ldots, x_N)$. Given *complete* knowledge and *centralized* control of the network in Fig. 1, an efficient rate allocation can be characterized as a solution of the following problem:

Problem 1 (Surplus Maximization with Routing):

$$\underset{\boldsymbol{x}}{\text{maximize}} \quad \sum_{n=1}^{N} U_n(x_n) - C\left(\sum_{n=1}^{N} x_n\right)$$

$$\text{subject to} \quad x_n \geq 0, \quad n = 1, \ldots, N.$$

The objective function in Problem 1 is the *network aggregate surplus* [25]. Network aggregate surplus maximization is a common network design objective (cf. [9], [10], [14], [24]). Clearly, Problem 1 is a *convex* optimization problem. In general, since the utility functions are local to the users and are *not* known to each link, efficient resource allocation can be achieved via *pricing*. Given the rate vector $\boldsymbol{x}$ from the users, the shared link (i,j) can set a *single* price

$$\mu(\boldsymbol{x}) = p\left(\sum_{n=1}^{N} x_n\right) \tag{1}$$

for each unit of data rate it carries. Each user $n \in \mathcal{N}$ then pays $x_n \mu(\boldsymbol{x})$ for its data rate x_n.

Next, we analyze how the users determine their rates based on the price set by link (i,j). First, assume that the users are *price takers*, i.e., they do *not* anticipate the effect of a change of their rates on the resulting price. In that case, each user $n \in \mathcal{N}$ selects its rate x_n to maximize

---

[1]The first axiom, i.e., *rescaling*, requires that the prices should be *independent* of the *units* of measurement. The second axiom, i.e., *consistency*, requires that two users having the same effect on the cost should face the same price. The third axiom, i.e., *additivity*, requires that if the cost function can be decomposed, then so should the price function. Finally, the fourth axiom, i.e., *positivity*, requires that if the cost is positive, then the price should be at least non-negative. Notice that the last axiom reflects a notion of fairness toward the service provider. [22].

its *own* surplus, i.e., utility minus payment, by solving the following *local* problem [6]:

$$\max_{x_n \geq 0} \ (U_n(x_n) - x_n \mu) \qquad \Rightarrow \qquad x_n = U_n'^{-1}(\mu), \tag{2}$$

where $U_n'^{-1}$ denotes the inverse of the derivative of utility function U_n and price μ is as in (1). From the first fundamental theorem of welfare economics, if each user $n \in \mathcal{N}$ selects its rate as in (2), then the network aggregate surplus is maximized at equilibrium [25, p. 326].

Next, we consider *price anticipating* users: each user can anticipate the effect of its selected data rate on the resulting price. In this case, each user $n \in \mathcal{N}$ no longer selects its rate as in (2). Instead, it *strategically* selects x_n to maximize its surplus given the knowledge that the price $\mu(\boldsymbol{x})$ is set according to (1) and is *not* fixed. Clearly, the decision made by user n also depends on the rates selected by other users, leading to a *resource allocation game* among all users:

Game 1 (Resource Allocation Game Among Routing Flows):

- *Players*: Users in set $\mathcal{N}$.

- *Strategies*: Transmission rates $\boldsymbol{x}$ for all users.

- *Payoffs*: $P_n(x_n; \boldsymbol{x}_{-n})$ for each user $n \in \mathcal{N}$, where

$$P_n(x_n; \boldsymbol{x}_{-n}) = U_n(x_n) - x_n \, p\left(\sum_{n=1}^{N} x_n\right),$$

and $\boldsymbol{x}_{-n}$ denotes the vector of selected data rates for all users *other than* user n.

In Game 1, each user $n \in \mathcal{N}$ selects its rate $x_n \geq 0$ to *maximize* its payoff $P_n(x_n; \boldsymbol{x}_{-n})$. A *Nash equilibrium* of Game 1 is defined as a non-negative rate vector $\boldsymbol{x}^* = (x_1^*, \ldots, x_N^*)$ such that for each user $n \in \mathcal{N}$, we have $P_n(x_n^*; \boldsymbol{x}_{-n}^*) \geq P_n(\bar{x}_n; \boldsymbol{x}_{-n}^*)$ for *any* $\bar{x}_n \geq 0$. In a Nash equilibrium $\boldsymbol{x}^*$, no user $n \in \mathcal{N}$ can increase its payoff by *unilaterally* changing its strategy x_n.

Definition 1: Let $\boldsymbol{x}^S = (x_1^S, \ldots, x_N^S)$ be an optimal solution for Problem 1 and $\boldsymbol{x}^*$ be a Nash equilibrium for Game 1 for the *same* choice of system parameters. The *efficiency* at Nash equilibrium $\boldsymbol{x}^*$ is defined as the *ratio* of the network aggregate surplus at $\boldsymbol{x}^*$ to the network aggregate surplus at $\boldsymbol{x}^S$:

$$\frac{\sum_{n=1}^{N} U_n\left(x_n^*\right) - C\left(\sum_{n=1}^{N} x_n^*\right)}{\sum_{n=1}^{N} U_n\left(x_n^S\right) - C\left(\sum_{n=1}^{N} x_n^S\right)}. \tag{3}$$

Definition 2: The *price-of-anarchy* PoA (Game 1, Problem 1) is defined as the *worst-case* efficiency of a Nash equilibrium of Game 1 among *all* possible selections of system parameters (i.e., number of users, utility, cost, and price functions) as long as Assumptions 1 and 2 hold.

The following key result is based on [8, Theorem 3]:

Theorem 1: Suppose Assumptions 1 and 2 hold. (a) Game 1 always has a *unique* Nash equilibrium. (b) We have

$$\text{PoA}\,(\text{Game 1}, \text{Problem 1}) = \frac{2}{3}. \tag{4}$$

From Theorem 1, for *any* choice of parameters, the network aggregate surplus at a Nash equilibrium of Game 1 is guaranteed to be *at least* $\frac{2}{3} \approx 67\%$ of the *optimal* network aggregate surplus. Notice that the PoA indicates how bad the network performance can become due to strategic behavior of the users. In the rest of this paper, we generalize Theorem 1 to the case where some of the users can perform inter-session network coding. We show that such a generalization is non-trivial and the results are drastically different from those in Theorem 1 in several aspects.

III. RESOURCE ALLOCATION GAME WITH INTER-SESSION NETWORK CODING AND ROUTING FLOWS: THE CASE WITH ZERO COSTS FOR SIDE LINKS

In this section, we reformulate Problem 1 and Game 1 in a network scenario where a bottleneck link is shared not only by routing flows, but also by some inter-session network coding flows. We then extend the results in Theorem 1 according to a *new* network resource allocation game. We show that the new game setting may have *multiple* Nash equilibria. In addition, the 67% efficiency bound in Theorem 1 is no longer guaranteed. In fact, although the efficiency loss is still *bounded*, the performance at some Nash equilibria is only 25% of the optimal performance.

A. Problem Formulation

Consider the modified network model in Fig. 2. The network topology in this figure is called a *butterfly* network in the network coding literature [5], [26]. The network in Fig. 2 is similar to the one in Fig. 1, except that it includes two direct *side* links (s_1, t_N) and (s_N, t_1). In this scenario, the source node of user 1 is located closer to the destination node of user N than to its own destination node (and vice versa). Thus, users 1 and N can perform *inter-session* network coding. In this setting, we can distinguish two different types of users in the system:

- *Network Coding Users*: Users 1 and N, who *can* perform inter-session network coding.
- *Routing Users*: Users $2, \ldots, N-1$, who *cannot* perform inter-session network coding.

Let X_1 and X_N denote packets sent from source nodes s_1 and s_N, respectively. Node i can encode packets X_1 and X_N jointly, and then send out the resulting encoded packet, denoted by

$X_1 \oplus X_N$, towards node j (and from there towards t_1 and t_N). Given the *remedy* data X_1 from the side link (s_1, t_N) and the *remedy* data X_N from the side link (s_N, t_1), nodes t_N and t_1 can *decode* the encoded packets that they receive. In fact, in this setting, nodes t_1 and t_N can decode both X_1 and X_N. Clearly, the benefit of network coding is to reduce the traffic load on link (i, j) (thus reducing the cost) while achieving the *same* data rates compared to the case that no network coding is performed. It is worth mentioning that although the network coding scenario in Fig. 2 is simple, it is the building block for more general network coding scenarios. For example, in [2], [3], [27] the network is modeled as a *superposition* of several butterfly networks. Thus, understanding this model is the key to understand more general networks. Further to Assumptions 1 and 2, in this section, we also assume that:

Assumption 3 (Zero Cost for Side Links): The two side links (s_1, t_N) and (s_N, t_1) in Fig. 2 always have *zero* cost and impose *zero* prices.

For example, if the link cost is used to model the link *delay* and the side links (s_1, t_N) and (s_N, t_1) have a higher capacity than the shared link (i, j), then the costs of the side links can be neglected. The case where the side links have *non-zero* cost is studied in Section IV.

For the network in Fig. 2, the network aggregate surplus maximization problem becomes:

Problem 2 (Surplus Maximization with Network Coding and Zero-Cost Side Links):

$$\underset{\boldsymbol{x}}{\text{maximize}} \quad \sum_{n=1}^{N} U_n(x_n) - C\left(\sum_{n=2}^{N-1} x_n + \max(x_1, x_N) \right)$$

$$\text{subject to} \quad x_n \geq 0, \quad n = 1, \dots, N.$$

Comparing Problems 1 and 2, we can see that the cost term $C(\sum_{n=1}^{N} x_n)$ in Problem 1 is now replaced by a new cost term $C(\sum_{n=2}^{N-1} x_n + \max(x_1, x_N))$. The intuition behind the objective function in Problem 2 is as follows. Since x_1 and x_N are selected *independently* by users 1 and N, in general, we may have $x_1 \neq x_N$. Thus, regardless of the choice of an *efficient* network coding scheme, the intermediate node i can perform network coding *only* at rate $\min(x_1, x_N)$. Those packets which are *not* encoded (e.g., at rate $x_1 - \min(x_1, x_N)$ if $x_1 \geq x_N$, and at rate $x_N - \min(x_1, x_N)$ if $x_1 \leq x_N$) are simply *forwarded*, leading to an aggregate rate of $\sum_{n=2}^{N-1} x_n + \max(x_1, x_N)$ on link (i, j). Note that if $x_1 = x_N$, then *all* packets from users 1 and N are jointly encoded. In fact, this is the case at optimality as the following result suggests:

Theorem 2: Let $\boldsymbol{x}^S = (x_1^S, \dots, x_N^S)$ be an *optimal* solution for Problem 2. We have $x_1^S = x_N^S$.

The proof of Theorem 2 is given in Appendix A. From Theorem 2, users 1 and N should have *equal* rates at optimality. Notice that since Problem 2 is a convex optimization problem, it can be solved in a centralized fashion using convex programming techniques [28]. Distributed resource allocation can also be done via pricing as explained next.

Following the same pricing scheme as in Section II, the shared link may apply a *single* price for *all* (i.e., coded and routed) packets:

$$\mu(\boldsymbol{x}) = p \left(\sum_{n=2}^{N-1} x_n + \max(x_1, x_N) \right). \tag{5}$$

Each user n pays $x_n \mu(\boldsymbol{x})$. However, this leads to *double charging* for *encoded* packets. Note that each encoded packet includes the data from *both* users 1 and N. Thus, the *single* pricing model in (5) leads to more payment from the users than the actual link cost. This can be avoided by *price discrimination*, i.e., charging the routed and network-coded packets with *different* prices.

Let $\mu(\boldsymbol{x})$ in (5) denote the price to be charged for routed packets. Under the discriminatory pricing scheme, we define another price $\delta(\boldsymbol{x})$ for network coded packets. In general, we have

$$\delta(\boldsymbol{x}) = \beta \, \mu(\boldsymbol{x}), \tag{6}$$

where $0 < \beta \leq 1$ is a pricing parameter. If $\beta = 1$, then there is only a single price. If $\beta < 1$, then the encoded packets are charged less than the routed packets as they carry more information compared to routing packets of the same size. In this paper, we focus on the case of $\beta = \frac{1}{2}$. This is the only choice of β that avoids *over-* or *under*-charging with two network coding flows. Based on the this pricing scheme, user 1 pays $\min(x_1, x_N)\delta(\boldsymbol{x}) + (x_1 - \min(x_1, x_N)) \, \mu(\boldsymbol{x})$. That is, it pays for transmission of its encoded packets at a price of $\delta(\boldsymbol{x})$ and for transmission of its forwarded (not coded) packets at a price of $\mu(\boldsymbol{x})$. From (6), the *total* payment by user 1 is

$$(x_1 - (1 - \beta) \min(x_1, x_N)) \, \mu(\boldsymbol{x}). \tag{7}$$

A similar payment model applies to user N. Notice that each user $n = 2, \ldots, N-1$ pays $x_n \mu(\boldsymbol{x})$.

We are now ready to define a resource allocation game for the network setting in Fig. 2, when users can anticipate prices μ and δ according to (5) and (6), respectively:

Game 2: (Resource Allocation Game with Inter-session Network Coding and Routing Flows where the Side Links Have Zero Costs and Zero Prices)

- *Players*: Users in set $\mathcal{N}$.
- *Strategies*: Transmission rates $\boldsymbol{x}$ for all users.

- *Payoffs*: $Q_n(x_n; \boldsymbol{x}_{-n})$ for each user $n \in \mathcal{N}$. The network coding users 1 and N have

$$Q_1(x_1; \boldsymbol{x}_{-1}) = U_1(x_1) - (x_1 - (1-\beta)\min(x_1, x_N)) p\left(\sum_{r=2}^{N-1} x_r + \max(x_1, x_N)\right),; \quad (8)$$

$$Q_N(x_N; \boldsymbol{x}_{-N}) = U_N(x_N) - (x_N - (1-\beta)\min(x_1, x_N)) p\left(\sum_{r=2}^{N-1} x_r + \max(x_1, x_N)\right), \quad (9)$$

and each routing user $n \in \mathcal{N}\backslash\{1, N\}$ has

$$Q_n(x_n; \boldsymbol{x}_{-n}) = U_n(x_n) - x_n p\left(\sum_{r=2}^{N-1} x_r + \max(x_1, x_N)\right).$$

Comparing Games 1 and 2, we can see that Game 2 introduces significantly more complex payoff functions. In the rest of this section, we answer the following questions:

1) Does Game 2 always (i.e., for *any* choice of system parameters) have a Nash equilibrium?

2) If a Nash equilibrium exists for Game 2, is it always unique?

3) What is the *worst-case* efficiency (i.e., the PoA) at a Nash equilibrium of Game 2?

B. Existence and Non-uniqueness of Nash Equilibria

A Nash equilibrium of Game 2 with both routing and inter-session network coding flows can be defined as a data rate selection vector $\boldsymbol{x}^* \succeq \boldsymbol{0}$, where the inequality is coordinate-wise, such that for all users $n \in \mathcal{N}$, we have $Q_n(x_n^*; \boldsymbol{x}_{-n}^*) \geq Q_n(\bar{x}_n; \boldsymbol{x}_{-n}^*)$ for any choice of $\bar{x}_n \geq 0$.

Theorem 3: There exists *at least* one Nash equilibrium in Game 2.

The proof of Theorem 3 is given in Appendix B. The key to prove Theorem 3 is to apply *Rosen's existence theorem for concave N-person games* [29, Theorem 1]. In this regard, we show that for *all* users $n \in \mathcal{N}$, the payoff function $Q_n(x_n; \boldsymbol{x}_{-n})$ is a *concave* function with respect to x_n, even though Q_1 and Q_N are *not* differentiable due to the $\max$ and $\min$ functions.

From Theorem 3, the existence of Nash equilibria for the resource allocation game is still guaranteed when network coding is applied. However, as we will see in Section III-C, there can multiple Nash equilibria in this case. This can change the results on efficiency loss and the PoA.

C. Users' Best Responses

The strategic behavior of users can be modeled based on their *best responses*. In this regard, each user $n \in \mathcal{N}$ selects its data rate as x_n^B to *maximize* its own payoff Q_n, given $\boldsymbol{x}_{-n}$:

$$x_n^B(\boldsymbol{x}_{-n}) = \arg\max_{x_n \geq 0} Q_n(x_n; \boldsymbol{x}_{-n}), \quad \forall n \in \mathcal{N}. \quad (10)$$

Since the problem in (10) is *convex*, we can readily show the following for *routing* users.

Proposition 1: For each routing user $n \in \mathcal{N}\backslash\{1, N\}$, the best response $x_n^B(\boldsymbol{x}_{-n})$ is obtained as the value of x_n which satisfies the following equation (bounded below by 0):

$$U_n'(x_n) - a\left(\sum_{r=2, r\neq n}^{N-1} x_r + \max(x_1, x_N)\right) - 2ax_n = 0. \tag{11}$$

Recall that the *linear* pricing parameter a is defined in Assumption 2. The proof of Proposition 1 is given in Appendix C. The key idea is to take the derivative of the payoff $Q_n(x_n; \boldsymbol{x}_{-n})$ with respect to x_n and solve the resulted Karush-Kuhn-Tucker (KKT) optimality condition [28].

Obtaining the best response functions for network coding users 1 and N is more complex, mostly due to the non-differentiability of the payoff functions $Q_1(x_1; \boldsymbol{x}_{-1})$ and $Q_N(x_N; \boldsymbol{x}_{-N})$. In fact, network coding user 1 should *separately* examine two scenarios:

(a) Selecting its strategy (data rate) x_1 to be *greater* than or equal to x_N:

$$\tilde{x}_1^B(\boldsymbol{x}_{-1}) = \arg\max_{x_1 \geq x_N} \ U_1(x_1) - (x_1 - (1-\beta)x_N) \ a \left(\sum_{n=2}^{N-1} x_n + x_1\right). \tag{12}$$

(b) Selecting its strategy (data rate) x_1 to be *less* than or equal to x_N:

$$\hat{x}_1^B(\boldsymbol{x}_{-1}) = \arg\max_{0 \leq x_1 \leq x_N} \ U_1(x_1) - \beta x_1 \ a \left(\sum_{n=2}^{N-1} x_n + x_N\right). \tag{13}$$

The intuition behind the objective functions in (12) and (13) is as follows. In (12), since the strategy of user 1, i.e., its data rate x_1, is *lower bounded* by x_N, we have: $\min(x_1, x_N) = x_N$ and $\max(x_1, x_N) = x_1$. Thus, the payoff function $Q_1(x_1; \boldsymbol{x}_{-1})$ in (8) reduces to the objective function in (12). On the other hand, in (13), since the data rate x_1 is *upper bounded* by x_N, we have: $\min(x_1, x_N) = x_1$, $\max(x_1, x_N) = x_N$, and $x_1 - (1-\beta)\min(x_1, x_N) = x_1 - (1-\beta)x_1 = \beta x_1$. Thus, the payoff function $Q_1(x_1; \boldsymbol{x}_{-1})$ in (8) reduces to the objective function in (13).

Given $\tilde{x}_1^B(\boldsymbol{x}_{-1})$ and $\hat{x}_1^B(\boldsymbol{x}_{-1})$, if $Q_1(\tilde{x}_1^B(\boldsymbol{x}_{-1}); \boldsymbol{x}_{-1}) \geq Q_1(\hat{x}_1^B(\boldsymbol{x}_{-1}); \boldsymbol{x}_{-1})$, then user 1 selects its best response $x_1^B(\boldsymbol{x}_{-1}) = \tilde{x}_1^B(\boldsymbol{x}_{-1})$; otherwise, it selects $x_1^B(\boldsymbol{x}_{-1}) = \hat{x}_1^B(\boldsymbol{x}_{-1})$. The best response for user N is obtained similarly. We can show the following for network coding users.

Proposition 2: For network coding user 1, the data rate $\tilde{x}_1^B(\boldsymbol{x}_{-1})$ is obtained as the value of x_1 that satisfies the following equation (bounded below by x_N)

$$U_1'(x_1) - a\left(\sum_{n=2}^{N-1} x_n + x_1\right) + a(1-\beta)x_N - ax_1 = 0. \tag{14}$$

Furthermore, if the utility function $U_1(x_1)$ is *non-linear*, then $\hat{x}_1^B(\boldsymbol{x}_{-1})$ is obtained as the value

of x_1 that satisfies the following equation (bounded between 0 and x_N)

$$U_1'(x_1) - \beta a \left(\sum_{n=2}^{N-1} x_n + x_N \right) = 0. \tag{15}$$

When the utility function $U_1(x_1)$ is *linear* (i.e., $U_1'(x_1)$ is a *constant* for all $x_1 \geq 0$), we have $\hat{x}_1^B(\boldsymbol{x}_{-1}) = x_N$, if $U_1'(x_1) > \beta a(\sum_{n=2}^{N-1} x_n + x_N)$; and $\hat{x}_1^B(\boldsymbol{x}_{-1}) = 0$, if $U_1'(x_1) < \beta a(\sum_{n=2}^{N-1} x_n + x_N)$. In this case, if $U_1'(x_1) = \beta a(\sum_{n=2}^{N-1} x_n + x_N)$, then $\hat{x}_1^B(\boldsymbol{x}_{-1})$ is *any* value between 0 and x_N.

The proof of Proposition 2 is given in Appendix D. The key idea is to solve the KKT optimality conditions [28] for the optimization problems in (12) and (13) for user 1 (and N). We can see that the best responses in Propositions 1 and 2 only depend on the *first derivatives* of the utility functions. We will use this key observation to characterize the Nash equilibrium in Section III-D.

D. Nash Equilibrium and Price-of-Anarchy

Given the best response functions in Section III-C, we are now ready to characterize the Nash equilibria. Let $\mathcal{X}^*$ denote the set of *all* Nash equilibria of Game 2. Recall that set $\mathcal{X}^*$ has *at least* one member as shown in Theorem 3. By definition, for any Nash equilibrium $\boldsymbol{x}^* \in \mathcal{X}^*$, given $\boldsymbol{x}_{-n}^*$, the best response for user $n \in \mathcal{N}$ is the same as its strategy at Nash equilibrium [30]. That is, $x_n^B(\boldsymbol{x}_{-n}^*) = x_n^*$ for all $n \in \mathcal{N}$. Thus, all Nash equilibria of Game 2 can be obtained using the best response equations in Propositions 1 and 2. Recall from Section III-C that the best responses in (13), (15), (15) only depend on the *first derivatives* of the utility functions. Therefore, for each Nash equilibrium $\boldsymbol{x}^* \in \mathcal{X}^*$, if we define the following *linear* utility functions:

$$\bar{U}_n(x_n) = U_n'(x_n^*) \, x_n, \qquad \forall n \in \mathcal{N}, \tag{16}$$

then $\boldsymbol{x}^*$ continues to be a Nash equilibrium for a *new* game with *new* utilities $\bar{U}_1(x_1), \ldots, \bar{U}_N(x_N)$. In fact, $\boldsymbol{x}^*$ is a Nash equilibrium for the *family* of games with utility functions $U_1(x_1), \ldots, U_N(x_N)$ having their first derivatives equal to $U_1'(x_1^*), \ldots, U_N'(x_N^*)$ at Nash equilibrium, respectively.

Theorem 4: For each Nash equilibrium $\boldsymbol{x}^* \in \mathcal{X}^*$ of Game 2 and any optimal solution $\boldsymbol{x}^S$ of Problem 2, the following inequality holds:

$$\frac{\sum_{n=1}^{N} U_n(x_n^*) - C\left(\sum_{n=2}^{N-1} x_n^* + \max(x_1^*, x_N^*) \right)}{\sum_{n=1}^{N} U_n(x_n^S) - C\left(\sum_{n=2}^{N-1} x_n^S + \max(x_1^S, x_N^S) \right)} \geq \frac{\sum_{n=1}^{N} \bar{U}_n(x_n^*) - C\left(\sum_{n=2}^{N-1} x_n^* + \max(x_1^*, x_N^*) \right)}{\max_{\tilde{q} \geq 0} \left[\sigma \, \tilde{q} - C(\tilde{q}) \right]}, \tag{17}$$

where $\sigma = \max \left\{ U_2'(x_2^*), \ldots, U_{N-1}'(x_{N-1}^*), U_1'(x_1^*) + U_N'(x_N^*) \right\}$.

The proof of Theorem 4 is given in Appendix E. Notice that $\max_{\tilde{q} \geq 0} [\, \sigma\, \tilde{q} - C(\tilde{q}) \,]$ denotes the optimal objective value of Problem 2 for the special case of having *linear* utility functions (see Appendix E). Therefore, for the inequality in (17), the right hand side denotes the efficiency for *linear* utility functions while the left hand side denotes the efficiency for *any* utility functions, assuming that the rest of the system parameters (i.e., number of users, cost function, and price functions) are the same. This leads us to the following helpful theorem.

Theorem 5: The *worst-case* efficiency at a Nash equilibrium of Game 2 with respect to the optimal solution of Problem 2 occurs when the utility functions are *linear* for all users. That is,

$$U_n(x_n) = \gamma_n x_n, \qquad \forall n \in \mathcal{N}, \tag{18}$$

where utility parameter $\gamma_n > 0$ for all users $n \in \mathcal{N}$.

From Theorem 5, to obtain the PoA for Game 2 for *arbitrary* choices of utility functions (as long as they satisfy Assumption 1), it is *enough* to only analyze the case when all utility functions are *linear*. This key observation can make our analysis significantly more tractable. Notice that for the case of linear utilities, we have $U'_n(x_n) = \gamma_n$ for all $n \in \mathcal{N}$. As a result, the best responses for all users can be obtained in *closed form* using Propositions 1 and 2.

Next, we obtain the exact value(s) of the Nash equilibrium(s) and PoA for Game 2.

Theorem 6: Suppose Assumptions 1, 2, and 3 hold. Also assume that the utility functions are *linear*. Consider the case where $N \geq 2$ and let $\boldsymbol{x}^*$ denote the Nash equilibrium for Game 2. Without loss of generality, assume that $\gamma_1 \geq \gamma_N$. For notational simplicity, we also define

$$q^* = \sum_{n=2}^{N-1} x_n^*. \tag{19}$$

(a) If $\gamma_N \leq \gamma_1 \leq \left(1 + \frac{1}{\beta}\right)\gamma_N - \beta a q^*$, then

$$\max\left\{0, \frac{\gamma_1 - a q^*}{a(1+\beta)}\right\} \leq x_1^* = x_N^* \leq \max\left\{0, \frac{\gamma_N - \beta a q^*}{\beta a}\right\}. \tag{20}$$

(b) If $\left(1 + \frac{1}{\beta}\right)\gamma_N - \beta a q^* \leq \gamma_1 \leq \frac{2}{\beta}\gamma_N - a q^*$, then

$$x_1^* = \frac{\gamma_N}{\beta a} - q^*, \qquad x_N^* = \frac{\frac{2}{\beta}\gamma_N - \gamma_1}{a(1-\beta)} - \frac{q^*}{1-\beta}. \tag{21}$$

(c) If $\gamma_1 \geq \frac{2}{\beta}\gamma_N - a q^*$, then

$$x_1^* = \max\left\{0, \frac{\gamma_1}{2a} - \frac{q^*}{2}\right\}, \qquad x_N^* = 0. \tag{22}$$

(d) For any choice of system parameters in (a)-(c), the routing users have the following rates[2]

$$x_n^* = \begin{cases} 0, & \text{if } \gamma_n \leq a(q^* + x_1^*), \\ \frac{\gamma_n}{a} - q^* - x_1^*, & \text{otherwise}, \end{cases} \qquad n = 2, \ldots, N - 1. \qquad (23)$$

The proof of Theorem 6 is given in Appendix F. From Theorem 6(a), if the slopes of the linear utility functions for users 1 and N (i.e., γ_1 and γ_N) are identical or close, then at Nash equilibrium, users 1 and N choose to have the same data rates and there are *infinite* Nash equilibria. Theorem 6(b) and 6(c) show that if γ_1 and γ_N are *not* close, then users 1 and N will choose *different* rates at the Nash equilibrium. Comparing this with the results in Theorem 2, we shall expect a drastic efficiency loss, especially if $\gamma_1 \geq \frac{2}{\beta}\gamma_N - aq^*$ as it results in $x_N^* = 0$. We also notice that the Nash equilibrium directly depends on the value of the pricing parameter β.

To study the properties of Nash equilibria of Game 2, we consider two different cases:

1) Two Users Case: Consider the butterfly network in Fig. 2 and assume that $N = 2$. In this case, the network includes two network coding users and *no* routing users. We can obtain the Nash equilibria using Theorem 6 by setting $q^* = 0$. The Nash equilibria when $\beta = 1$ and $\beta = \frac{1}{2}$ are shown in Fig. 3(a) and (b), respectively. We can see that the data rates at the Nash equilibria are always less than the optimal rates, except when $\gamma_1 = \gamma_2$ and $\beta = \frac{1}{2}$. In addition, in many cases (i.e., $\gamma_1 \geq 2\gamma_2$ for $\beta = 1$ and $\gamma_1 \geq 3\gamma_2$ for $\beta = \frac{1}{2}$) we have $x_2^* < x_1^*$. This leads to further deviation from the optimal performance. Recall from Theorem 2 that at optimality, the data rates of users 1 and N should be *equal*. We can show the following in the two-user case:

Theorem 7: In a network as in Fig. 2 with $N = 2$, under the *single* pricing scheme ($\beta = 1$),

$$\text{PoA (Game 2, Problem 2)} = \frac{1}{3}, \qquad (24)$$

and under the *discriminatory* pricing scheme with $\beta = \frac{1}{2}$,

$$\text{PoA (Game 2, Problem 2)} = \frac{12}{25}. \qquad (25)$$

The proof of Theorem 7 is given in Appendix G. Here, PoA (Game 2, Problem 2) denotes the *lowest* (i.e., *worst-case*) ratio of the network aggregate surplus at a Nash equilibrium of Game

---

[2]Notice that for each routing user $n \in \mathcal{N}\backslash\{1, N\}$, the strategy at Nash equilibrium, i.e., data rate x_n^*, only depends on q^* and x_1^*, but *not* x_N^*. In fact, since we have assumed that $\gamma_N \leq \gamma_1$, we indeed have $x_N^* \leq x_1^*$, as shown in (20)-(22). Therefore, $\max(x_1^*, x_N^*) = x_1^*$ and neither the cost function nor the price functions for the shared link (i, j) depend on data rate x_N^*.

2 to the network aggregate surplus at the optimal solution of Problem 2. Theorem 7 extends the results on efficiency bounds for routing flows in Theorem 1 to the case where two inter-session network coding users share a link. We can see that even for this simple scenario, the efficiency bound in Theorem 1 *cannot* be guaranteed anymore. From Theorem 7, inter-session network coding with no price discrimination can reduce the PoA from 0.67 down to $\frac{1}{3} \approx 0.33$. On the other hand, even if we use price discrimination by setting $\beta = \frac{1}{2}$, i.e., users 1 and N split the price of encoded packets, the PoA improves only to $\frac{12}{25} = 0.48$. This implies that inter-session network coding is significantly *more sensitive* to strategic users. Thus, unlike the case of routing networks, a simple pricing scheme (even with price discrimination) may not be sufficient to encourage cooperation in inter-session network coding systems.

It is worth mentioning that the above results do *not* imply superiority of routing versus network coding. In fact, we can verify that at *any* Nash equilibrium of Game 2, the network surplus is higher than or equal to the network surplus at the Nash equilibrium of Game 1 for the *same* parameters. In other words, non-cooperative network coding results in an *absolute* performance which is *no worse* than the *absolute* performance of non-cooperative routing. However, the *relative* performance in non-cooperative network coding compared to optimal cooperative network coding is worse than the relative performance in the routing-only case.

Numerical results on efficiency of the Nash equilibrium of Game 2 for 200 *randomly* generated scenarios with different choices of system parameters in the two-user case are shown in Fig. 4. In particular, in each scenario, the utility functions of the users are chosen to be α-fair (cf. [23]) with a randomly selected utility parameter $\alpha \in (0, 1)$. We can see that by using price discrimination with parameter $\beta = \frac{1}{2}$, the guaranteed worst-case efficiency bound (i.e., the PoA) improves from 0.33 to 0.48. For the rest of this paper, we focus on the case with $\beta = \frac{1}{2}$. That is, the network coding users *split* the charge of transmitting their jointly encoded packets.

2) General Case: Next, consider the case where the topology is as in Fig. 2 and there are $N > 2$ users in the network. The presence of both network coding and routing users makes the analysis more complex. To see this, consider the network in Fig. 2 and assume that $N = 3$, $a = 1$, $\beta = \frac{1}{2}$, $\gamma_1 \geq \gamma_3$, $\gamma_3 = 1$, and $\gamma_2 = 3$. In this case, users 1 and 3 are the network coding users and user 2 is a routing user. From Theorem 6, the Nash equilibria are obtained as shown in Fig. 5. Comparing the results in Fig. 5 with those in Fig. 3, we can see that adding an extra routing user forces the network coding users 1 and 3 to *reduce* their data rates at Nash equilibrium.

However, there still exist multiple (infinite) Nash equilibria when γ_1 and γ_3 are close. It is easy to numerically verify that in this scenario, the *worst-case* efficiency at Nash equilibrium of Game 2 is 46.5%. Comparing this with the results in Theorem 7, we can expect that adding routing users will further reduce the PoA. In fact, we can show the next theorem in a *general* case:

Theorem 8: Consider a network coding system as in Fig. 2 and assume that $N \geq 2$.

(a) If the price discrimination parameter $\beta = \frac{1}{2}$, we have

$$\text{PoA (Game 2, Problem 2)} = \frac{1}{4}. \tag{26}$$

(b) The worst-case efficiency occurs when $N \to \infty$.

The proof of Theorem 8 is given in Appendix H. Comparing the results in Theorem 8 with those in Theorems 1 and 7, we can see that a resource allocation game with *both* network coding and routing users has a *worse* PoA compared to the *routing only* and *network coding only* cases.

IV. RESOURCE ALLOCATION GAME WITH INTER-SESSION NETWORK CODING AND ROUTING FLOWS: THE CASE WITH NON-ZERO COSTS FOR SIDE LINKS

In Section III, we considered a network coding scenario in a butterfly network where the side links have *zero* cost as stated in Assumption 3. In this section, we study the case where the side links have *non-zero* cost. We show that the results will be noticeably different. In particular, the network coding users are no longer interested in participating in network coding. This can further reduce the efficiency to as low as only 20% of the optimal network coding performance.

A. Problem Formulation

Consider the network in Fig. 6. In this figure, the side link (s_1, t_N) has price p_1 and cost C_1 while the side link (s_N, t_1) has price p_N and cost C_N. Suppose that Assumption 2 also holds for the price and cost functions of both side links. In addition, we make the following assumption.

Assumption 4 (Non-Zero Cost for Side Links): The side links (s_1, t_N) and (s_N, t_1) in Fig. 6 always have *non-zero* cost and impose *non-zero* prices. In particular, the side link (s_1, t_N) has price $p_1(v_1) = a_1 v_1$ for $a_1 > 0$ and the side link (s_N, t_1) has price $p_N(v_N) = a_N v_N$ for $a_N > 0$.

Clearly, by sending remedy packets over side link (s_1, t_N), user 1 is *helping* user N to decode the encoded packets it may receive. However, due to non-zero cost at the side links, user 1 will

be *charged* for sending these remedy packets. A similar statement is true for user N when it sends remedy packets on the side link (s_N, t_1). Therefore, users 1 and N may decide to *reduce* the rate at which they send the remedy packets on the side links, compared to the rate at which they send their own data packets to node i. In other words, they may decide to have *partial* or *no* participation in network coding. Users 1 and N can inform node i about their decision using a simple *packet marking* scheme, e.g., by using a *flag* in the packet header. Let y_1 and z_1 denote the rate at which source s_1 sends data to node i *marked* for routing and network coding, respectively. Similarly, let y_N and z_N denote the data rate at which source s_N sends data to node i *marked* for routing and network coding, respectively. Node i may jointly encode only those packets which are marked for network coding. If the packet is marked for routing, then node i simply forwards the packet without modifying its content. Furthermore, let v_1 and v_N denote the data rates at which sources s_1 and s_N send remedy packets on side links (s_1, t_N) and (s_N, t_1), respectively. The routing users $2, \ldots, N-1$ just send routing packets at rates $y_2, \ldots, y_{N-1}$, respectively.

Given the above data rates, intermediate node i encodes packets at rate $\min(z_1, z_N)$ and *forwards* the rest of packets at rate $\sum_{n=1}^{N} y_n + |z_1 - z_N|$. As a result, the total rate on link (i, j) becomes $\sum_{n=1}^{N} y_n + \max(z_1, z_N)$. Let $\boldsymbol{y} = (y_1, \ldots, y_N)$, $\boldsymbol{z} = (z_1, z_N)$, and $\boldsymbol{v} = (v_1, v_N)$. For the butterfly network in Fig. 6, the network aggregate surplus maximization problem becomes

Problem 3 (Surplus Maximization with Network Coding and Non-Zero-Cost Side Links):

$$\underset{\boldsymbol{y},\boldsymbol{z},\boldsymbol{v}}{\text{maximize}} \quad \sum_{n=2}^{N-1} U_n(y_n) + U_1(y_1 + \min(z_1, v_N)) + U_N(y_N + \min(z_N, v_1))$$

$$- C\left(\sum_{n=1}^{N} y_n + \max(z_1, z_N)\right) - C_1(v_1) - C_N(v_N)$$

$$\text{subject to} \quad y_n \geq 0, \quad n = 1, \ldots, N, \qquad z_1, z_N, v_1, v_N \geq 0.$$

We can see that the objective function in Problem 3 is more complex than the one in Problem 2 and includes the cost of side links (s_1, t_N) and (s_N, t_1).

Following a discriminatory pricing model as in Section III-A, we can define a resource allocation game for the network setting in Fig. 6, when users are price anticipators:

Game 3: (Resource Allocation Game with Inter-session Network Coding and Routing Flows and Non-zero Costs for Side Links)

- *Players*: Users in set $\mathcal{N}$.

- *Strategies*: Transmission rates $\boldsymbol{y}$, $\boldsymbol{z}$, and $\boldsymbol{v}$.

- *Payoffs*: $W_n(\cdot)$ for each user $n \in \mathcal{N}$, where

$$W_1(y_1, z_1, v_1; \boldsymbol{y}_{-1}, z_N, v_N) = U_1\left(y_1 + \min(z_1, v_N)\right) - v_1 p_1(v_1)$$
$$- \left(y_1 + z_1 - (1-\beta)\min(z_1, z_N)\right) p\left(\sum_{r=1}^{N} y_r + \max(z_1, z_N)\right),$$
(27)

$$W_N(y_N, z_N, v_N; \boldsymbol{y}_{-N}, z_1, v_1) = U_N\left(y_N + \min(z_N, v_1)\right) - v_N p_N(v_N)$$
$$- \left(y_N + z_N - (1-\beta)\min(z_1, z_N)\right) p\left(\sum_{r=1}^{N} y_r + \max(z_1, z_N)\right),$$
(28)

$$W_n(y_n; \boldsymbol{y}_{-n}) = U_n(y_n) - y_n p\left(\sum_{r=1}^{N} y_r + \max(z_1, z_N)\right), \qquad n \in \mathcal{N}\backslash\{1, N\}.$$
(29)

Here, for each user $n \in \mathcal{N}$, we have $\boldsymbol{y}_{-n} = (y_1, \ldots, y_{n-1}, y_{n+1}, \ldots, y_N)$. Next, we study the efficiency and the worst-case efficiency (i.e., the PoA) at Nash equilibria of Game 3.

B. Users' Best Responses

For network coding user 1, the best response is in form of $\left(y_1^B(\boldsymbol{y}_{-1}, z_N, v_N),\ z_1^B(\boldsymbol{y}_{-1}, z_N, v_N),\right.$ $\left. v_1^B(\boldsymbol{y}_{-1}, z_N, v_N)\right)$ which is obtained as the solution of the following optimization problem

$$\left(y_1^B(\boldsymbol{y}_{-1}, z_N, v_N), z_1^B(\boldsymbol{y}_{-1}, z_N, v_N), v_1^B(\boldsymbol{y}_{-1}, z_N, v_N)\right) = \underset{y_1 \geq 0,\, z_1 \geq 0,\, v_1 \geq 0}{\arg\max}\ W_1(y_1, z_1, v_1; \boldsymbol{y}_{-1}, z_N, v_N).$$

The best response for network coding user N, denoted by $\left(y_N^B(\boldsymbol{y}_{-N}, z_1, v_1),\ z_N^B(\boldsymbol{y}_{-N}, z_1, v_1),\right.$ $\left. v_N^B(\boldsymbol{y}_{-N}, z_1, v_1)\right)$ is obtained similarly. We can show the following for users 1 and N.

Proposition 3: Users 1 and N *always* send zero remedy packets at the best responses of Game 3. That is, we always have $v_1^B(\boldsymbol{y}_{-1}, z_N, v_N) = 0$ and $v_N^B(\boldsymbol{y}_{-N}, z_1, v_1) = 0$.

Proposition 3 can be proved by noticing that the payoff $W_1(y_1, z_1, v_1; \boldsymbol{y}_{-1}, z_N, v_N)$ is *decreasing* in v_1 and $W_N(y_N, z_N, v_N; \boldsymbol{y}_{-N}, z_1, v_1)$ is *decreasing* in v_N. Clearly, if the network coding users do *not* receive the remedy data from the side links, they *cannot* decode any encoded packet they may receive through the shared link (i, j). In fact, we can further show the following.

Proposition 4: Users 1 and N *always* send zero network coding packets to intermediate node i as the best responses of Game 3. That is, $z_1^B(\boldsymbol{y}_{-1}, z_N, v_N) = 0$ and $z_N^B(\boldsymbol{y}_{-N}, z_1, v_1) = 0$.

Notice that if $v_N = 0$, then $\min(z_1, v_N) = 0$ and the payoff function for user 1 reduces to $U_1(y_1) - v_1 p_1(v_1) - \left(y_1 + z_1 - (1-\beta)\min(z_1, z_N)\right) p(\sum_{r=1}^{N} y_r + \max(z_1, z_N))$. In that case, the payoff function is *decreasing* in z_1. A similar statement is true for network coding user N.

C. Nash Equilibrium and Price-of-Anarchy

Given the results on users' best responses in Propositions 3 and 4, we can conclude that at any Nash equilibrium of Game 3, denoted by $(\boldsymbol{y}^*, \boldsymbol{z}^*, \boldsymbol{v}^*)$, we should indeed have

$$z_1^* = z_N^* = v_1^* = v_N^* = 0. \tag{30}$$

In other words, at a Nash equilibrium of Game 3, *none* of the users perform network coding. In that case, the Nash equilibria of Game 3 would be closely related to the Nash equilibria of Game 1. In fact, for any choice of system parameters, if $\boldsymbol{x}^*$ is a Nash equilibrium of Game 1, then $\boldsymbol{y}^* = \boldsymbol{x}^*$, $\boldsymbol{z}^* = \boldsymbol{0}$, and $\boldsymbol{v}^* = \boldsymbol{0}$ would be a Nash equilibrium of Game 3 for the same choice of system parameters. From this, together with the results in Theorem 1(a), we can conclude that Game 3 always has a *unique* Nash equilibrium. This leads to the following theorem.

Theorem 9: The *worst-case* efficiency of Game 3 occurs when the utility functions are *linear*.

The proof of Theorem 9 is similar to that of [8, Lemma 4]. From Theorem 9, to obtain the PoA for Game 3 for *arbitrary* choices of utility functions (as long as the utility functions satisfy Assumption 1), it is *enough* to only analyze the case where all utility functions are *linear*. Furthermore, we notice that if the side links have a very large cost compared to the cost of the bottleneck link, the optimal performance is achieved if network coding is *not* applied. In that case, the efficiency can be obtained by using Theorem 1. Notice that in this case, the *optimal* network aggregate surplus for Problem 3 is the same as the *optimal* network aggregate surplus for Problem 1. In addition, the network aggregate surplus is the same at the Nash equilibrium of Game 3 and Game 1. However, for general choices of $a_1 > 0$ and $a_N > 0$, obtaining the PoA requires further investigation of the optimal solution of Problem 3.

Theorem 10: Let $\boldsymbol{y}^S = (y_1^S, \ldots, y_N^S)$, $\boldsymbol{z}^S = (z_1^S, z_N^S)$, and $\boldsymbol{v}^S = (v_1^S, v_N^S)$ denote the optimal solution for Problem 3. Assume that all utility functions are linear with slope γ_n for each user $n \in \mathcal{N}$. Define $\gamma_{\max} = \max_{n \in \mathcal{N}} \gamma_n$ and $\mathcal{M} = \{n : \gamma_n = \gamma_{\max}\}$ with size $M = |\mathcal{M}|$.
(a) If $\gamma_1 + \gamma_N \geq \left(1 + \frac{a_1 + a_N}{a}\right) \gamma_{\max}$, then

$$z_1^S = z_N^S = v_1^S = v_N^S = \frac{\gamma_1 + \gamma_N}{a + a_1 + a_N}, \qquad y_n^S = 0, \quad \forall n \in \mathcal{N}. \tag{31}$$

(b) If $\gamma_{\max} \leq \gamma_1 + \gamma_N \leq \left(1 + \frac{a_1 + a_N}{a}\right) \gamma_{\max}$, then

$$z_1^S = z_N^S = v_1^S = v_N^S = \frac{\gamma_1 + \gamma_N - \gamma_{\max}}{a_1 + a_N}, \qquad y_n^S = \begin{cases} \frac{(a + a_1 + a_N)\gamma_{\max} - a(\gamma_1 + \gamma_N)}{aM(a_1 + a_N)}, & \text{if } n \in \mathcal{M}, \\ 0, & \text{if } n \in \mathcal{N} \backslash \mathcal{M}. \end{cases} \tag{32}$$

(c) If $\gamma_{\max} \geq \gamma_1 + \gamma_N$, then

$$z_1^S = z_N^S = v_1^S = v_N^S = 0, \qquad y_n^S = \begin{cases} \frac{\gamma_{\max}}{aM}, & \text{if } n \in \mathcal{M}, \\ 0, & \text{if } n \in \mathcal{N} \backslash \mathcal{M}. \end{cases} \tag{33}$$

Recall that the *linear* pricing parameters a, a_1, and a_N are defined in Assumptions 2 and 4, respectively. The proof of Theorem 10 is given in Appendix I. Combining (30)-(33) with the results in [8, Section II], we can show the following results on the efficiency loss in Game 3.

Theorem 11: Consider a network coding system as in Fig. 6 and assume that $N \geq 2$.
(a) We have

$$\text{PoA (Game 3, Problem 3)} = \frac{1}{5}. \tag{34}$$

(b) The worst-case efficiency occurs when $N \to \infty$.

The proof of Theorem 11 is given in Appendix J. Here, PoA (Game 3, Problem 3) denotes the *lowest* (i.e., *worst-case*) ratio of the network surplus at a Nash equilibrium of Game 3 to the network surplus at the optimal solution of Problem 3. Comparing Theorem 11 and Theorem 8, we can see that a non-zero cost at the side links can further reduce the PoA in a network resource allocation game as the users do *not* perform network coding in this case. If the side link price parameters a_1 and a_N are significantly greater than the bottleneck link price parameter a, then network coding is not an optimal solution and the efficiency loss follows from the results in Theorem 1. This is shown in Fig. 7. For the results in this figure, the network topology is assumed to be as in Fig. 6, where $N \to \infty$, $\gamma_1 = \gamma_N = 1$, $a = 1$, and $\gamma_n = \frac{4}{5}$ for all $n \in \mathcal{N} \backslash \{1, N\}$. The side link price parameters $a_1 = a_N$ vary from 0 (*non-inclusive*) to 10. If $a_1 > 0$ and $a_N > 0$ tend to zero, the efficiency becomes as low as $\frac{1}{5} = 0.2$ as Theorem 11 suggests. As $a_1 = a_N$ increases and tends to infinity, Problem 3 becomes equivalent to Problem 1 (in terms of the optimal network aggregate surplus) and Game 3 becomes equivalent to Game 1 (in terms of network aggregate surplus at Nash equilibrium) which leads to an efficiency higher than $\frac{2}{3} \approx 0.67$ as Theorem 1 suggests (for the choice of parameters in Fig. 7, the efficiency approaches $\frac{4}{5} = 0.8$). Numerical results on the efficiency of the Nash equilibrium of Game 3 for 200 random scenarios with different choices of system parameters in the *two-user* case are also shown in Fig. 8. We can see that the simulations confirm Theorem 11.

V. CONCLUSION

This paper represents a first step towards understanding the impact of strategic network coding users on the network resource allocation efficiency. To gain insights, we focus on the case of the well-known butterfly network topology where there is a single bottleneck link in the network shared by several users. Two of the users have the capability of performing inter-session network coding, and the rest are routing only users. Even with this simple setup, the results are dramatically different from the traditional routing-only case. In particular, there can be many (even infinite) Nash equilibria in the resulting resource allocation game. This is in sharp contrast to a similar game setting with traditional packet forwarding where the Nash equilibrium is always unique. Furthermore, we showed that the efficiency loss can be significantly more severe than for the case without network coding. The precise values of the PoA and the efficiency loss depend on the pricing scheme used by the links. Compared with the traditional single pricing approach, a novel discriminatory pricing, which charges encoded and forwarded packets differently can improve efficiency. However, regardless of the discriminatory pricing scheme being used, the PoA is still worse than for the case when network coding is not applied. This implies that, although inter-session network coding can improve performance compared to routing, it is significantly *more sensitive* to users' strategic behaviors. For example, in a butterfly network when the side links have *zero* cost, the efficiency at certain Nash equilibria can be as low as 25%. If the side links have *non-zero* cost, then the efficiency at some Nash equilibria further reduces to only 20%. These results generalize the well-known result of guaranteed 67% worst-case efficiency, shown by Johari and Tsitsiklis in [8] for traditional packet forwarding networks. This motivates our ongoing work of mechanism design to encourage the strategic users to perform network coding, e.g., by using a combination of reward and punishment in a dynamic game setting.

APPENDIX

A. Proof of Theorem 2

Let $\tilde{\boldsymbol{x}} = (\tilde{x}_1, \ldots, \tilde{x}_N)$ denote *any* arbitrary feasible solution for Problem 2 such that $\tilde{x}_1 \neq \tilde{x}_N$. Without loss of generality, we assume that $\tilde{x}_1 > \tilde{x}_N$. We then define $\hat{\boldsymbol{x}} = (\hat{x}_1, \ldots, \hat{x}_N)$ as *another* feasible solution such that $\hat{x}_n = \tilde{x}_n$ for all $n \in \mathcal{N} \backslash \{1, N\}$ and $\hat{x}_1 = \hat{x}_N = \max(\tilde{x}_1, \tilde{x}_N) = \tilde{x}_1$.

Since for all users, the utility functions are strictly increasing in transmission rates, we have

$$\sum_{n=1}^{N} U_n(\hat{x}_n) = \sum_{n=1}^{N-1} U_n(\tilde{x}_n) + U_N(\hat{x}_N) > \sum_{n=1}^{N-1} U_n(\tilde{x}_n) + U_N(\tilde{x}_N) = \sum_{n=1}^{N} U_n(\tilde{x}_n). \tag{35}$$

On the other hand, since $\max(\hat{x}_1, \hat{x}_N) = \max(\tilde{x}_1, \tilde{x}_N)$,

$$C\left(\sum_{n=2}^{N-1} \hat{x}_n + \max(\hat{x}_1, \hat{x}_N)\right) = C\left(\sum_{n=2}^{N-1} \tilde{x}_n + \max(\tilde{x}_1, \tilde{x}_N)\right). \tag{36}$$

From (35) and (36), the feasible solution $\hat{\boldsymbol{x}}$ results in strictly higher network aggregate surplus compared to $\tilde{\boldsymbol{x}}$. Thus, vector $\tilde{\boldsymbol{x}}$ *cannot* be an optimal solution for Problem 2. ∎

B. Proof of Theorem 3

We prove the existence of Nash equilibrium for Game 2 by using Rosen's existence theorem for N-person games [29, Theorem 1]. In this regard, we need to show that for each user $n \in \mathcal{N}$, the payoff function Q_n is *continuous* and *concave* in data rate x_n. This is *not* trivial, particularly for network coding users due to the complexity of the payoff functions.

For user 1, given $\bar{\boldsymbol{x}}_{-1} = (\bar{x}_2, \ldots, \bar{x}_N)$, we can rewrite the payoff function Q_1 as

$$Q_1(x_1; \bar{\boldsymbol{x}}_{-1}) = \begin{cases} G_1(x_1; \bar{\boldsymbol{x}}_{-1}), & \text{if } x_1 \leq \bar{x}_N, \\ H_1(x_1; \bar{\boldsymbol{x}}_{-1}), & \text{if } x_1 > \bar{x}_N, \end{cases} \tag{37}$$

where

$$G_1(x_1; \bar{\boldsymbol{x}}_{-1}) = U_1(x_1) - \beta x_1 \, p\left(\sum_{n=2}^{N-1} \bar{x}_n + \bar{x}_N\right),$$

and

$$H_1(x_1; \bar{\boldsymbol{x}}_{-1}) = U_1(x_1) - (x_1 - (1-\beta)\bar{x}_N) \, p\left(\sum_{n=2}^{N-1} \bar{x}_n + x_1\right).$$

Clearly, $G_1(\bar{x}_N; \bar{\boldsymbol{x}}_{-1}) = H_1(\bar{x}_N; \bar{\boldsymbol{x}}_{-1})$. That is, $G_1(x_1; \bar{\boldsymbol{x}}_{-1})$ is continuous at $x_1 = \bar{x}_N$. Thus, since $G_1(\bar{x}_N, \bar{\boldsymbol{x}}_{-1})$ and $H_1(\bar{x}_N, \bar{\boldsymbol{x}}_{-1})$ are also continuous functions, function $Q_1(x_1, \bar{\boldsymbol{x}}_{-1})$ is continuous in x_1. However, $Q_1(x_1; \bar{\boldsymbol{x}}_{-1})$ may *not* always be differentiable. Next, we show that $Q_1(x_1; \bar{\boldsymbol{x}}_{-1})$ is *concave*. We need to show that for each $\tilde{x}_1 \geq 0$ and $\hat{x}_1 \geq 0$ and for any $0 \leq \theta \leq 1$, we have

$$Q_1(\theta \hat{x}_1 + (1-\theta)\tilde{x}_1; \bar{\boldsymbol{x}}_{-1}) \geq \theta Q_1(\hat{x}_1; \bar{\boldsymbol{x}}_{-1}) + (1-\theta) Q_1(\tilde{x}_1; \bar{\boldsymbol{x}}_{-1}). \tag{38}$$

Without loss of generality, we assume that $\tilde{x}_1 \leq \hat{x}_1$. We notice that if $\tilde{x}_1 \leq \hat{x}_1 \leq \bar{x}_N$ or $\bar{x}_N \leq \tilde{x}_1 \leq \hat{x}_1$, then (38) is directly obtained from the fact that G_1 and H_1 are concave. Thus, we only consider the case where $\tilde{x}_1 \leq \bar{x}_N \leq \hat{x}_1$. We consider two different scenarios for choice of θ:

$$0 \leq \theta \leq \frac{\bar{x}_N - \tilde{x}_1}{\hat{x}_1 - \tilde{x}_1}, \tag{39}$$

and

$$\frac{\bar{x}_N - \tilde{x}_1}{\hat{x}_1 - \tilde{x}_1} \leq \theta \leq 1. \tag{40}$$

If (39) holds, then $\theta\hat{x}_1 + (1-\theta)\tilde{x}_1 \leq \bar{x}_N$. In addition, since $\hat{x}_1 \geq \bar{x}_N$, we have

$$(1-\beta)\bar{x}_N \leq (1-\beta)\hat{x}_1 \qquad \Rightarrow \qquad \beta\hat{x}_1 \leq \hat{x}_1 - (1-\beta)\bar{x}_N, \tag{41}$$

and

$$p\left(\sum_{n=2}^{N-1} \bar{x}_n + \hat{x}_1\right) \geq p\left(\sum_{n=2}^{N-1} \bar{x}_n + \bar{x}_N\right). \tag{42}$$

From (41) and (42), we have

$$G_1(\hat{x}_1) \geq H_1(\hat{x}_1). \tag{43}$$

Therefore,

$$\begin{aligned}
Q_1(\theta\hat{x}_1 + (1-\theta)\tilde{x}_1; \bar{\boldsymbol{x}}_{-1}) &= G_1(\theta\hat{x}_1 + (1-\theta)\tilde{x}_1; \bar{\boldsymbol{x}}_{-1}) \\
&\geq \theta G_1(\hat{x}_1; \bar{\boldsymbol{x}}_{-1}) + (1-\theta)G_1(\tilde{x}_1; \bar{\boldsymbol{x}}_{-1}), \\
&= \theta G_1(\hat{x}_1) + (1-\theta)Q_1(\tilde{x}_1; \bar{\boldsymbol{x}}_{-1}), \\
&\geq \theta H_1(\hat{x}_1; \bar{\boldsymbol{x}}_{-1}) + (1-\theta)Q_1(\tilde{x}_1; \bar{\boldsymbol{x}}_{-1}), \\
&= \theta Q_1(\hat{x}_1; \bar{\boldsymbol{x}}_{-1}) + (1-\theta)Q_1(\tilde{x}_1; \bar{\boldsymbol{x}}_{-1}),
\end{aligned} \tag{44}$$

where the second line results from concavity of G_1, the third line results from the fact that $Q_1(\tilde{x}_1; \bar{\boldsymbol{x}}_{-1}) = G_1(\tilde{x}_1; \bar{\boldsymbol{x}}_{-1})$, the fourth line results from (43), and the fifth line is due to $Q_1(\hat{x}_1; \bar{\boldsymbol{x}}_{-1}) = H_1(\hat{x}_1; \bar{\boldsymbol{x}}_{-1})$. Next assume that (40) holds. In that case, $\theta\hat{x}_1 + (1-\theta)\tilde{x}_1 \geq \bar{x}_N$. In addition, since $\tilde{x}_1 \leq \bar{x}_N$, we have

$$(1-\beta)\bar{x}_N \geq (1-\beta)\tilde{x}_1 \qquad \Rightarrow \qquad \beta\tilde{x}_1 \geq \tilde{x}_1 - (1-\beta)\bar{x}_N, \tag{45}$$

and

$$p\left(\sum_{n=2}^{N-1} \bar{x}_n + \tilde{x}_1\right) \leq p\left(\sum_{n=2}^{N-1} \bar{x}_n + \bar{x}_N\right). \tag{46}$$

From (45) and (46), we have

$$G_1(\tilde{x}_1; \bar{\boldsymbol{x}}_{-1}) \leq H_1(\tilde{x}_1; \bar{\boldsymbol{x}}_{-1}), \tag{47}$$

Therefore, we can show that

$$Q_1(\theta\hat{x}_1 + (1-\theta)\tilde{x}_1; \bar{\boldsymbol{x}}_{-1}) = H_1(\theta\hat{x}_1 + (1-\theta)\tilde{x}_1; \bar{\boldsymbol{x}}_{-1})$$

$$\geq \theta H_1(\hat{x}_1; \bar{\boldsymbol{x}}_{-1}) + (1-\theta)H_1(\tilde{x}_1; \bar{\boldsymbol{x}}_{-1}),$$

$$= \theta Q_1(\hat{x}_1; \bar{\boldsymbol{x}}_{-1}) + (1-\theta)H_1(\tilde{x}_1; \bar{\boldsymbol{x}}_{-1}), \tag{48}$$

$$\geq \theta Q_1(\hat{x}_1; \bar{\boldsymbol{x}}_{-1}) + (1-\theta)G_1(\tilde{x}_1; \bar{\boldsymbol{x}}_{-1}),$$

$$= \theta Q_1(\hat{x}_1; \bar{\boldsymbol{x}}_{-1}) + (1-\theta)Q_1(\tilde{x}_1; \bar{\boldsymbol{x}}_{-1}).$$

From (44) and (48), payoff $Q_1(x_1, \bar{\boldsymbol{x}}_{-1})$ always satisfies (38). Thus, $Q_1(x_1, \bar{\boldsymbol{x}}_{-1})$ is concave. Similarly, we can show that $Q_N(x_N, \bar{\boldsymbol{x}}_{-N})$ is continuous and concave. The same statement is evidently true for all $n \in \mathcal{N}\backslash\{1, N\}$. Thus, Game 2 is a concave *N-person* game and the existence of Nash equilibria follows from the Rosen's existence theorem [29, Theorem 1]. ∎

C. Proof of Proposition 1

For each user $n \in \mathcal{N}\backslash\{1, N\}$, the KKT optimality conditions for problem (10) become

$$U'_n(x_n) - a(\textstyle\sum_{r=2,r\neq n}^{N-1} x_r + x_n + \max(x_1, x_N)) - ax_n = -\lambda_n, \tag{49}$$

and $\lambda_n x_n = 0$. Here, $\lambda_n \geq 0$ denotes the *Lagrange multiplier* corresponding to inequality constraint $x_n \geq 0$. If $x_n > 0$, then $\lambda_n = 0$ and (49) reduces to (11). Otherwise, (11) cannot hold for any positive x_n and we have $x_n = 0$. Recall that the best response is bounded below by zero. ∎

D. Proof of Proposition 2

For user 1, we can write the KKT optimality conditions for optimization problem (12) as

$$U'_1(x_1) - a(\textstyle\sum_{n=2}^{N-1} x_n + x_1) - a(x_1 - (1-\beta)x_N) = -\lambda_1, \tag{50}$$

and $\lambda_1 (x_1 - x_N) = 0$. Here, $\lambda_1 \geq 0$ denotes the *Lagrange multiplier* corresponding to constraint $x_1 \geq x_N$. If $x_1 > x_N$, then $\lambda_1 = 0$ and (50) reduces to (14). Otherwise, x_1 takes its lower bound value, i.e., $x_1 = x_N$. It is also easy to verify that (15) is simply the KKT condition for convex optimization problem (13), as long as the utility function $U_1(x_1)$ is *non-linear* (i.e., concave, but *not* linear). Therefore, it should hold for user 1's best response data rate. If the utility function $U_1(x_1)$ is *linear*, then the objective function in optimization problem (13) becomes

$$x_1 \left(U'_1(x_1) - \beta\, a \left(\textstyle\sum_{n=2}^{N-1} x_n + x_N\right)\right), \tag{51}$$

where $U'_1(x_1) - \beta a(\sum_{n=2}^{N-1} x_n + x_N)$ is a *constant*. In this case, the best response depends on the *sign* of the multiplier term $U'_1(x_1) - \beta a(\sum_{n=2}^{N-1} x_n + x_N)$ as explained in Proposition 2. ∎

E. Proof of Theorem 4

At Nash equilibrium of Game 2, we have $x_n^B(\boldsymbol{x}_{-n}^*) = x_n^*$ for all users $n \in \mathcal{N}$. From this, together with (11), (14), and (15) and also due to $\beta \leq 1$, we can show that

$$\sum_{n=1}^{N} U_n'(x_n^*)x_n^* - C\left(\sum_{n=2}^{N-1} x_n^* + \max(x_1^*, x_N^*)\right) \geq 0. \tag{52}$$

That is, the network aggregate surplus is always *non-negative* at any Nash equilibrium of Game 2. On the other hand, using concavity of the utility functions, for each user $n \in \mathcal{N}$, we have

$$U_n(x_n^S) \leq U_n(x_n^*) + U_n'(x_n^*)\left(x_n^S - x_n^*\right). \tag{53}$$

Again, by concavity, $U_n(0) \leq U_n(x_n^*) + U_n'(x_n^*)(0 - x_n^*)$. Since according to Assumption 1, $U_n(0) \geq 0$, we have $U_n'(x_n^*)x_n^* \leq U_n(x_n^*)$. Applying this to all users $n \in \mathcal{N}$, we have

$$\sum_{n=1}^{N} \left(U_n(x_n^*) - U_n'(x_n^*)x_n^*\right) \geq 0. \tag{54}$$

For notational simplicity, we define

$$C^* = C\left(\sum_{n=2}^{N-1} x_n^* + \max(x_1^*, x_N^*)\right) \quad \text{and} \quad C^S = C\left(\sum_{n=2}^{N-1} x_n^S + \max(x_1^S, x_N^S)\right). \tag{55}$$

We notice that

$$\begin{aligned}
\sum_{n=1}^{N} U_n'(x_n^*)x_n^S - C^S &= \sum_{n=2}^{N-1} U_n'(x_n^*)x_n^S + (U_1'(x_1^*) + U_N'(x_N^*))\, x_1^S - C^S \\
&\leq \sigma\left(\sum_{n=2}^{N-1} x_n^S + \max(x_1^S, x_N^S)\right) - C^S \\
&\leq \max_{\tilde{q} \geq 0}\left[\sigma\tilde{q} - C(\tilde{q})\right],
\end{aligned} \tag{56}$$

where $\tilde{q}$ is an *auxiliary* variable, the equality results from Theorem 2, the first inequality is due to the definition of σ, and the second inequality results from (55). Clearly, $\max_{\tilde{q} \geq 0}\left[\sigma\tilde{q} - C(\tilde{q})\right]$ in (56) provides an *upper bound* for the *optimal* network aggregate surplus when the utility functions are *linear* as in (16). First assume that $\sigma = U_1'(x_1^*) + U_N'(x_N^*)$. In this case, if we select $x_n^S = 0$ for all $n \in \mathcal{N} \setminus \{1, N\}$ and select $x_1^S = x_N^S = \tilde{q}^S$, where $\tilde{q}^S = \arg\max_{\tilde{q} \geq 0}\left[\sigma\tilde{q} - C(\tilde{q})\right]$, then the network aggregate surplus reaches its upper bound $\max_{\tilde{q} \geq 0}\left[\sigma\tilde{q} - C(\tilde{q})\right]$. Next, assume that $\sigma = U_m'(x_m^*)$ for some $m \in \mathcal{N} \setminus \{1, N\}$. In that case, if we select $x_n^S = 0$ for all $n \in \mathcal{N} \setminus \{m\}$ and select $x_m^S = \tilde{q}^S$, then the network aggregate surplus reaches its upper bound $\max_{\tilde{q} \geq 0}\left[\sigma\tilde{q} - C(\tilde{q})\right]$.

Therefore, we can conclude that $\max_{\tilde{q} \geq 0} [\sigma\tilde{q} - C(\tilde{q})]$ is indeed equal to the optimal objective value of Problem 2 for the special case of having *linear* utility functions. Finally, from (53)-(56),

$$
\begin{aligned}
1 &\geq \frac{\sum_{n=1}^{N} U_n(x_n^*) - C^*}{\sum_{n=1}^{N} U_n(x_n^S) - C^S} \\
&= \frac{\sum_{n=1}^{N} \left(U_n(x_n^*) - U_n'(x_n^*)x_n^*\right) + \sum_{n=1}^{N} U_n'(x_n^*)x_n^* - C^*}{\sum_{n=1}^{N} U_n(x_n^S) - C^S} \\
&\geq \frac{\sum_{n=1}^{N} \left(U_n(x_n^*) - U_n'(x_n^*)x_n^*\right) + \sum_{n=1}^{N} U_n'(x_n^*)x_n^* - C^*}{\sum_{n=1}^{N} \left(U_n(x_n^*) - U_n'(x_n^*)x_n^*\right) + \sum_{n=1}^{N} U_n'(x_n^*)x_n^S - C^S} \\
&\geq \frac{\sum_{n=1}^{N} \left(U_n(x_n^*) - U_n'(x_n^*)x_n^*\right) + \sum_{n=1}^{N} U_n'(x_n^*)x_n^* - C^*}{\sum_{n=1}^{N} \left(U_n(x_n^*) - U_n'(x_n^*)x_n^*\right) + \max_{\tilde{q}\geq 0}[\sigma\tilde{q} - C(\tilde{q})]} \\
&\geq \frac{\sum_{n=1}^{N} U_n'(x_n^*)x_n^* - C^*}{\max_{\tilde{q}\geq 0}[\sigma\tilde{q} - C(\tilde{q})]} \geq 0,
\end{aligned}
\tag{57}
$$

where the third line results from (53), the fourth line results from (56), and the last line results from (52), (54), and (56) and also the fact that $\max_{\tilde{q}\geq 0}[\sigma\tilde{q} - C(\tilde{q})]$ is always non-negative. ∎

F. Proof of Theorem 6

We notice that since the utility functions are linear, we have: $U_n'(\cdot) = \gamma_n$ for all $n \in \mathcal{N}$. Next, we prove that due to $\gamma_1 \geq \gamma_N$, we should always have $x_1^* \geq x_N^*$. To show thus, assume that $x_1^* < x_N^*$. Since $U_1'(x_1) = \gamma_1$, from Proposition 2, the inequality $x_1^* < x_N^*$ implies that

$$
\gamma_1 \leq \beta a \left(q^* + x_N^*\right).
\tag{58}
$$

Furthermore, since $U_N'(x_N) = \gamma_N$, from (14) and after reordering the terms, we have[3]

$$
\gamma_N = aq^* + 2ax_N^* - a(1 - \beta)x_1^*.
\tag{59}
$$

From (58), (59), and due to $\gamma_N \leq \gamma_1$, it is required that

$$
\begin{aligned}
aq^* + 2ax_N^* - a(1 - \beta)x_1^* &\leq \beta a(q^* + x_N^*) \\
\Rightarrow \quad q^*(1 - \beta) + x_N^*(2 - \beta) - (1 - \beta)x_1^* &\leq 0.
\end{aligned}
\tag{60}
$$

If $\beta = 1$, then the inequality in (60) reduces to $x_N^* \leq 0$ which *contradicts* the assumption that $x_N^* > x_1^* \geq 0$. On the other hand, if $0 < \beta < 1$, then we can further show that

$$
q^*(1 - \beta) + x_N^*(2 - \beta) - (1 - \beta)x_1^* \geq q^*(1 - \beta) + (x_N^* - x_1^*)(1 - \beta) > 0,
\tag{61}
$$

---

[3]Notice that (14) in its current format is formulated for user 1; however, it can be reformulated similarly for user N.

where the last (strict) inequality results from the assumption that $x_N^* > x_1^*$. It is clear that (61) *contradicts* (60). Thus, for any $0 < \beta \leq 1$, the data rate x_N^* *cannot* be greater than x_1^* and we *always* have $x_N^* \leq x_1^*$. Next, we prove each part of the theorem separately.

Part (a): We first prove that $x_1^* = x_N^*$ by contradiction. If $x_1^* \neq x_N^*$, then since $x_1^* \geq x_N^*$, we should have $x_1^* > x_N^*$. In that case, from Proposition 2,

$$x_1^* = \frac{\gamma_1 - aq^* + a(1 - \beta)x_N^*}{2a} > x_N^* \quad \Rightarrow \quad \gamma_1 > (1 + \beta)ax_N^* + aq^*, \tag{62}$$

and

$$\gamma_N \leq \beta ax_1^* + \beta aq^*. \tag{63}$$

From (62) and (63),

$$\gamma_N \leq \beta aq^* + \frac{\beta}{2}(\gamma_1 - aq^* + a(1 - \beta)x_N^*)$$

$$< \beta aq^* + \frac{\beta}{2}\left(\gamma_1 - aq^* + (1 - \beta)\frac{\gamma_1 - aq^*}{1 + \beta}\right) = \frac{\beta}{1 + \beta}(\gamma_1 + \beta aq^*). \tag{64}$$

This implies that

$$\left(1 + \frac{1}{\beta}\right)\gamma_N - \beta aq^* < \gamma_1. \tag{65}$$

However, this contradicts the assumption in this scenario that $\gamma_1 \leq (1 + 1/\beta)\gamma_N - \beta aq^*$. Thus, we indeed have $x_1^* = x_N^*$. From this, together with Proposition 2, we have

$$\gamma_1 \leq \beta ax_1^* + aq^* + ax_1^* = aq^* + (1 + \beta)ax_1^* \quad \Rightarrow \quad \frac{\gamma_1 - aq^*}{a(1 + \beta)} \leq x_1^*, \tag{66}$$

and

$$\gamma_N \geq \beta aq^* + \beta ax_1^* \quad \Rightarrow \quad x_1^* \leq \frac{\gamma_N - \beta aq^*}{\beta a}. \tag{67}$$

Notice that the transmission rates are *non-negative*. Thus, the best responses are as in (20).

Part (b): We first notice that the condition in this scenario holds if and only if

$$\left(1 + \frac{1}{\beta}\right)\gamma_N - \beta aq^* \leq \frac{2}{\beta}\gamma_N - aq^* \quad \Rightarrow \quad \gamma_N \geq \beta aq^*. \tag{68}$$

On the other hand, since $\gamma_1 \leq \frac{2}{\beta}\gamma_N - aq^*$, we have $\frac{2}{\beta}\gamma_N - \gamma_1 - aq^* \geq 0$ and x_N^* in (21) is *non-negative*. Since $x_1^* \geq x_N^*$, this also implies non-negativity of x_1^*. Next, we consider two cases:

Case I) Assume that $x_N^* > 0$. Following similar steps as in the proof of Part (a), we can show that in this case, $x_1^* > x_N^*$. From this, together with Proposition 2,

$$x_1^* = \frac{\gamma_1 - aq^* + a(1 - \beta)x_N^*}{2a}, \tag{69}$$

and

$$\gamma_N = \beta a q^* + \beta a x_1^* \qquad \Rightarrow \qquad x_1^* = \frac{\gamma_N - \beta a q^*}{\beta a} = \frac{\gamma_N}{\beta a} - q^*. \tag{70}$$

Replacing (70) in (69), we have

$$\frac{\gamma_N}{\beta a} - q^* = \frac{\gamma_1 - a q^* + a(1-\beta)x_N^*}{2a} \qquad \Rightarrow \qquad x_N^* = \frac{\frac{2}{\beta}\gamma_N - \gamma_1 - a q^*}{a(1-\beta)}. \tag{71}$$

Case II) Assume that $x_N^* = 0$. In that case, from Proposition 2, we have

$$x_1^* = \frac{\gamma_1 - a q^*}{2a}, \tag{72}$$

and

$$\gamma_N \leq \beta a q^* + \beta a x_1^*. \tag{73}$$

Replacing (72) in (73), we have

$$\gamma_N \leq \frac{\beta}{2} + \frac{\beta a q^*}{2} \qquad \Rightarrow \qquad \gamma_1 \geq \frac{2}{\beta}\gamma_N - a q^*. \tag{74}$$

From (74) and knowing that $\gamma_1 \leq \frac{2}{\beta}\gamma_N - a q^*$, it is required that

$$\gamma_1 = \frac{2}{\beta}\gamma_N - a q^* \qquad \overset{\text{By (72)}}{\Rightarrow} \qquad x_1^* = \frac{\gamma_N}{\beta a} - q^*. \tag{75}$$

Thus, in both Case I and Case II, the best responses are as in (21).

Part (c): We consider two cases:

Case I) Assume that $x_1^* > x_N$. In that case, (62) and (63) hold. First, assume that $\gamma_N = \beta a q^* + \beta a x_1^*$. Thus, $x_1^* = \frac{\gamma_N}{\beta a} - q^*$. Replacing this in (62), the data rate for user N is obtained as

$$x_N^* = \frac{\frac{2}{\beta}\gamma_N - \gamma_1 - a q^*}{a(1-\beta)} \leq 0, \tag{76}$$

where the inequality results from the fact that $\gamma_1 \geq \frac{2}{\beta}\gamma_N - a q^*$. Clearly, since the data rates are non-negative, the above implies that $x_N^* = 0$. Replacing this in (62), we have

$$x_1^* = \frac{\gamma_1 - a q^*}{2a}. \tag{77}$$

Next, assume that (63) holds as a *strict* inequality. That is,

$$\gamma_N < \beta a q^* + \beta a x_1^* \qquad \Rightarrow \qquad x_N^* = 0. \tag{78}$$

Replacing this in (62), the data rate model in (22) is obtained.

Case II) Assume that $x_1^* = x_N^*$. In that case, from Proposition 2, we have

$$\gamma_1 \leq (1+\beta)a x_1^* + a q^*, \tag{79}$$

and

$$\gamma_N \geq \beta a q^* + \beta a x_1^* \quad \Rightarrow \quad \frac{2}{\beta}\gamma_N - aq^* \geq aq^* + 2ax_1^*. \tag{80}$$

From (79) and (80) and since by assumption $\gamma_1 \geq \frac{2}{\beta}\gamma_N - aq^*$, it is required that

$$(1+\beta)ax_1^* + aq^* \geq aq^* + 2ax_1^* \quad \Rightarrow \quad a(1-\beta)x_1^* \leq 0. \tag{81}$$

Thus, either $x_1^* = x_N^* = 0$ and $\gamma_1 = aq^*$ or $\beta = 1$ and $\gamma_1 = 2ax_1^* + aq^*$. In the latter case,

$$x_1^* = \frac{\gamma_1 - aq^*}{2a}. \tag{82}$$

In addition, from Proposition 2, we have

$$x_1^* = \frac{\gamma_1 - aq^* + a(1-\beta)x_N^*}{2a}. \tag{83}$$

From (82) and (83), $x_1^* = x_N^* = 0$ and $\gamma_1 = aq^*$. Clearly, these results satisfy (22).

Part (d): For each node $n \in \mathcal{N} \backslash \{1, N\}$, at each Nash equilibrium $\boldsymbol{x}^* \in \mathcal{X}^*$ of Game 2, it is required that $x_n^B(\boldsymbol{x}_{-n}^*) = x_n^*$. Thus, in case of having linear utility functions, the derivative of the objective function in problem (11) with respect to variable x_n is obtained as $\gamma_n - a(q^* + x_1^*) - x_n a$. If $\gamma_n \leq a(q^* + x_1^*)$, the derivative is always *non-positive* and the objective function is *decreasing* in data rate x_n. In that case, $x_n^* = 0$. Otherwise, i.e., if $\gamma_n \geq a(q^* + x_1^*)$, then since the objective function is *convex*, we have $x_n^* = \frac{\gamma_n}{a} - q^* - x_1^*$. Together, these two cases result in (23). ∎

G. Proof of Theorem 7

We first obtain the optimal solution of Problem 2 when $N = 2$. Since Problem 2 is convex, we can use the KKT optimality conditions and confirm that the *optimal* data rates are

$$x_1^S = x_2^S = \frac{\gamma_1 + \gamma_2}{a}. \tag{84}$$

Thus, at optimality, the network aggregate surplus becomes

$$\gamma_1 x_1^S + \gamma_2 x_2^S - \frac{a}{2}\left(\max\{x_1^S, x_2^S\}\right)^2 = \frac{(\gamma_1+\gamma_2)^2}{a} - \frac{a}{2}\left(\frac{\gamma_1+\gamma_2}{a}\right)^2 = \frac{(\gamma_1+\gamma_2)^2}{2a}. \tag{85}$$

Next, we examine the efficiency for all the scenarios in Theorem 6(a), (b), (c), where $q^* = 0$ as there is *no* routing user in the network. First, we assume that

$$\gamma_2 \leq \gamma_1 \leq \left(1 + \frac{1}{\beta}\right)\gamma_2. \tag{86}$$

From Theorem 6(a), the Nash equilibria are as in (20). Since there are *multiple* Nash equilibria,

the *worst-case* efficiency for Game 2 is obtained by solving the following optimization problem

$$\underset{x_1^*}{\text{minimize}} \quad \frac{\left(\gamma_1 + \gamma_N\right) x_1^* - \frac{a}{2} x_1^{*2}}{\frac{(\gamma_1+\gamma_2)^2}{2a}} \tag{87}$$

$$\text{subject to} \quad \frac{\gamma_1}{(1+\beta)a} \le x_1^* \le \frac{\gamma_2}{\beta a}.$$

Problem (87) is a concave *minimization* (not maximization) problem. Thus, the optimality occurs at one of the *boundary* points. That is, either at $x_1^* = \frac{\gamma_2}{\beta a}$ or at $x_1^* = \frac{\gamma_1}{(1+\beta)a}$. The derivative of the objective function in Problem (87) with respect to x_1^* can be written as

$$\frac{2a}{(\gamma_1 + \gamma_2)^2} \left((\gamma_1 + \gamma_2) - ax_1^*\right). \tag{88}$$

The sign of the derivative in (88) depends on the choice of parameter β. We assume that $\frac{1}{2} \le \beta \le 1$. This includes the two cases of $\beta = \frac{1}{2}$ and $\beta = 1$. From (86), we have

$$\gamma_1 \beta \ge \gamma_2(1 - \beta) \quad \Rightarrow \quad \gamma_1 + \gamma_2 \ge \frac{\gamma_2}{\beta} \ge ax_1^*, \tag{89}$$

where the last inequality is due to constraint $x_1^* \le \frac{\gamma_2}{\beta a}$ in (87). From (89), the derivative of the objective function in Problem (87) is *positive*. Thus, the objective function is *increasing* in x_1^* and the minimum occurs at *lower-bound* $x_1^* = \frac{\gamma_1}{a(1+\beta)}$. In this case, the efficiency ratio becomes

$$\frac{2\gamma_1}{(1+\beta)\left(\gamma_1+\gamma_2\right)^2} \left(\gamma_1+\gamma_2-\frac{\gamma_1}{2(1+\beta)}\right) = \frac{2}{1+\beta}\left(\frac{1}{1+\gamma_2/\gamma_1}\right) - \frac{1}{(1+\beta)^2}\left(\frac{1}{1+\gamma_2/\gamma_1}\right)^2.$$

For the ease of exposition, we define $\Gamma = \frac{1}{1+\gamma_2/\gamma_1}$. Since $\gamma_1 \ge \gamma_2$, we have

$$\frac{\gamma_2}{\gamma_1} \le 1 \;\Rightarrow\; 1 + \frac{\gamma_2}{\gamma_1} \le 2 \;\Rightarrow\; \frac{1}{1+\gamma_2/\gamma_1} \ge \frac{1}{2} \;\Rightarrow \Gamma \ge \frac{1}{2}. \tag{90}$$

On the other hand, from the lower and upper bounds in problem (87), we also have

$$1 + \frac{\gamma_2}{\gamma_1} \ge \frac{1}{1+1/\beta} + 1 \;\Rightarrow\; \Gamma \le \frac{1+\beta}{1+2\beta}. \tag{91}$$

Thus, the worst-case efficiency is obtained by solving the following problem over variable Γ:

$$\underset{\Gamma}{\text{minimize}} \quad \frac{2}{1+\beta}\Gamma - \frac{1}{(1+\beta)^2}\Gamma^2 \tag{92}$$

$$\text{subject to} \quad \frac{1}{2} \le \Gamma \le \frac{1+\beta}{1+2\beta}.$$

The derivative of the objective function in problem (92) can be obtained as

$$\frac{2}{1+\beta} - \frac{2}{(1+\beta)^2}\Gamma = \frac{2}{1+\beta}\left(1 - \frac{\Gamma}{1+\beta}\right). \tag{93}$$

On the other hand, we can show that

$$\Gamma \le \frac{1+\beta}{1+2\beta} \quad \Rightarrow \quad \frac{\Gamma}{1+\beta} \le \frac{1}{1+2\beta} \quad \Rightarrow \quad 1 - \frac{\Gamma}{1+\beta} \ge 1 - \frac{1}{1+2\beta} = \frac{2\beta}{1+2\beta} \ge 0. \quad (94)$$

Thus, the derivative in (93) is always *positive* and minimum efficiency occurs at the *lower bound* $\Gamma = \frac{1}{2}$. Therefore, the worst-case efficiency when (86) holds is obtained as

$$\frac{2}{1+\beta} \times \frac{1}{2} - \frac{1}{(1+\beta)^2} \times \frac{1}{4} = \frac{1}{1+\beta} - \frac{1}{4(1+\beta)^2}. \quad (95)$$

If $\beta = 1$, then (95) becomes

$$1/2 - 1/16 = 7/16 \approx 0.438. \quad (96)$$

On the other hand, if $\beta = \frac{1}{2}$, then (95) becomes

$$6/9 - 1/9 = 5/9 \approx 0.556. \quad (97)$$

Next, we assume that $\beta < 1$ and

$$(1 + 1/\beta)\, \gamma_2 \le \gamma_1 \le \frac{2}{\beta}\gamma_2. \quad (98)$$

From Theorem 6, the data rates of users 1 and 2 at Nash equilibrium are as in (21) where $q^* = 0$. Assuming that γ_2 is *fixed*, the worst-case efficiency is obtained by solving the following problem

$$\underset{\gamma_1}{\text{minimize}} \quad \frac{\gamma_2}{(\gamma_1+\gamma_2)^2} \left(2\left(\frac{1}{\beta} - \frac{1}{1-\beta} \right) \gamma_1 + \frac{1}{\beta}\left(\frac{4}{1-\beta} - \frac{1}{\beta} \right) \gamma_2 \right)$$

$$\text{subject to} \quad (1 + 1/\beta)\, \gamma_2 \le \gamma_1 \le \frac{2}{\beta}\gamma_2. \quad (99)$$

The derivative of the objective function in (99) can be obtained as

$$-\frac{1}{(\gamma_1+\gamma_2)^3} \left(\Phi\,(\gamma_1 - \gamma_2) + 2\Psi \right), \quad (100)$$

where

$$\Phi = 2\gamma_2 \left(\frac{1}{\beta} - \frac{1}{1-\beta} \right) \qquad \text{and} \qquad \Psi = \frac{1}{\beta}\left(\frac{4}{1-\beta} - \frac{1}{\beta} \right) \gamma_2^2. \quad (101)$$

If $\beta = \frac{1}{2}$, then $\Phi = 0$ and $\Psi = 12\gamma_2^2$. Thus, the derivative in (100) is *negative* and the minimum in (99) occurs at *upper bound* $\gamma_1 = \frac{2}{\beta}\gamma_2 = 4\gamma_2$. Replacing this in the objective function in (99), the worst-case efficiency when $\beta = \frac{1}{2}$ and (98) holds is obtained as

$$\frac{12\gamma_2^2}{(4\gamma_2 + \gamma_2)^2} = \frac{12\gamma_2^2}{(5\gamma_2)^2} = \frac{12}{25} = 0.48. \quad (102)$$

Finally, we assume that

$$\frac{2}{\beta}\gamma_2 \leq \gamma_1. \tag{103}$$

From Theorem 6(c), the Nash equilibrium is as in (22) and the worst-case efficiency is obtained by solving the following optimization problem

$$\underset{\gamma_1,\gamma_2}{\text{minimize}} \quad \frac{\gamma_1\frac{\gamma_1}{2a} - \frac{a}{2}\left(\frac{\gamma_1}{2a}\right)^2}{\frac{(\gamma_1+\gamma_2)^2}{2a}} \tag{104}$$

$$\text{subject to} \quad 0 \leq \frac{2}{\beta}\gamma_2 \leq \gamma_1.$$

The objective function in problem (104) is *decreasing* in γ_2. Thus, the minimum occurs at *upper-bound* $\gamma_2 = \frac{\beta}{2}\gamma_1$. The worst-case efficiency when (103) holds is obtained as

$$\frac{\frac{1}{2a} - \frac{a}{2}\left(\frac{1}{2a}\right)^2}{\frac{\left(1+\frac{\beta}{2}\right)^2}{2a}} = \frac{3}{4}\left(\frac{1}{1+\beta/2}\right)^2 = \frac{3}{(2+\beta)^2}. \tag{105}$$

If $\beta = 1$, then (105) becomes

$$\frac{3}{(2+1)^2} = \frac{1}{3} \approx 0.33. \tag{106}$$

On the other hand, if $\beta = \frac{1}{2}$, then (105) becomes

$$\frac{3}{\left(2+\frac{1}{2}\right)^2} = \frac{12}{25} = 0.48. \tag{107}$$

Considering all the possible choices of system parameters, if pricing parameter $\beta = 1$, then

$$\text{PoA (Game 2, Problem 2)} \underset{\text{and (106)}}{\overset{\text{By (96)}}{=}} \min\left\{\frac{7}{16}, \frac{1}{3}\right\} = \frac{1}{3}. \tag{108}$$

On the other hand, if $\beta = \frac{1}{2}$, then

$$\text{PoA (Game 2, Problem 2)} \underset{\text{and (107)}}{\overset{\text{By (97), (102),}}{=}} \min\left\{\frac{5}{9}, \frac{12}{25}, \frac{12}{25}\right\} = \frac{12}{25}. \tag{109}$$

This concludes the proof. ∎

H. Proof of Theorem 8

Using the KKT optimality conditions, the optimal network aggregate surplus, i.e., the optimal solution of Problem 2, for linear utilities is obtained as $\sigma^2/(2a)$. Next, we study two cases:

Case I) We assume that $\gamma_1 + \gamma_N = \sigma$. Similar to the proof of Theorem 7, here we obtain the PoA by examining all the scenarios in Theorem 6(a), (b), (c). First, assume that

$$\gamma_N \leq \gamma_1 \leq (1 + 1/\beta)\gamma_N - \beta a q^*, \tag{110}$$

and

$$\gamma_1 \geq aq^*. \tag{111}$$

In that case, Nash equilibrium is obtained as in (20). We notice that

$$0 \overset{\text{By (111)}}{\leq} \frac{\gamma_1 - aq^*}{a(1+\beta)} \overset{\text{By (110)}}{\leq} \frac{\gamma_N - \beta aq^*}{\beta a}. \tag{112}$$

To obtain the worst-case efficiency for this scenario, we need to solve the following problem

$$\underset{\boldsymbol{x}^*, \boldsymbol{\gamma}, a, N, q^*}{\text{minimize}} \quad \frac{\sigma x_1^* + \sum_{n=2}^{N-1} \gamma_n x_n^* - \frac{a}{2}(q^* + x_1^*)^2}{\sigma^2/(2a)} \tag{113}$$

$$\text{subject to} \quad \gamma_n = a\left(q^* + x_n^* + x_1^*\right), \quad \text{if } x_n^* > 0, \quad n = 2, \ldots, N-1, \tag{114}$$

$$\gamma_n \leq a(q^* + x_1^*), \quad \text{if } x_n^* = 0, \quad n = 2, \ldots, N-1, \tag{115}$$

$$\sum_{n=2}^{N-1} x_n^* = q^*, \tag{116}$$

$$\gamma_1 + \gamma_N = \sigma, \tag{117}$$

$$\gamma_1 \geq aq^*, \tag{118}$$

$$0 < \gamma_n \leq \sigma, \quad n = 2, \ldots, N-1, \tag{119}$$

$$\gamma_N \leq \gamma_1 \leq (1 + 1/\beta)\gamma_N - \beta aq^*, \tag{120}$$

$$\frac{\gamma_1 - aq^*}{a(1+\beta)} \leq x_1^* = x_N^* \leq \frac{\gamma_N - \beta aq^*}{\beta a}, \tag{121}$$

$$x_n^* \geq 0, \quad n = 1, \ldots, N. \tag{122}$$

Here, $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_N)$ denotes the vector of utility parameters for all users. We assume, without loss of generality, that $\gamma_n = a(q^* + x_n^* + x_1^*)$ for all users $n \in \mathcal{N}\backslash\{1, N\}$. In fact, if for a feasible efficiency, we have $x_n^* > 0$ for some $n \in \mathcal{N}\backslash\{1, N\}$, then this assumption simply implies constraint (114). On the other hand, if for a feasible efficiency, we have $x_n^* = 0$ for some $n \in \mathcal{N}\backslash\{1, N\}$, then assuming $\gamma_n = a(q^* + x_n^* + x_1^*)$ does *not* have any impact on the objective function as the term $\gamma_n x_n^* = 0$, regardless of the value of γ_n. Therefore, we can restrict our attention only to those feasible solutions for which we have $\gamma_n = a(q^* + x_n^* + x_1^*)$ for all $n = 2, \ldots, N-1$. Having done so, to solve problem (113)-(122), we first assume that all variables, except for $x_1 = x_N$, are *fixed*. The derivative of the objective function in (113) with respect to variable x_1 is obtained as

$$\frac{2a}{\sigma^2}\left(\sigma - a(q^* + x_1^*)\right). \tag{123}$$

For any $\frac{1}{2} \leq \beta \leq 1$ and due to the fact that $\gamma_1 \geq \gamma_N$, we have $\sigma = \gamma_1 + \gamma_N \geq \frac{\gamma_N}{\beta}$. By adding $-a(q^* + x_1^*)$ to both sides of this inequality, we have

$$\sigma - a(q^* + x_1^*) \geq \frac{\gamma_N}{\beta} - a(q^* + x_1^*) \overset{\text{By (110)}}{\geq} 0. \tag{124}$$

Therefore, the derivative in (123) is always *non-negative* and the worst-case efficiency occurs at *lower bound* $x_1^* = x_N^* = \frac{\gamma_1 - aq^*}{a(1+\beta)}$. From this, together with (114), for each $n = 2, \ldots, N-1$,

$$\gamma_n = ax_n^* + a(q^* + x_1^*) = ax_n^* + a\left(\frac{\gamma_1}{a(1+\beta)} - \frac{q^*}{1+\beta} + q^*\right) = ax_n^* + a\left(\frac{\gamma_1}{a(1+\beta)} + q^*\left(\frac{\beta}{1+\beta}\right)\right). \tag{125}$$

Thus, the objective function in (113) becomes

$$\frac{2a}{\sigma^2}\left[\sigma\left(\frac{\gamma_1 - aq^*}{a(1+\beta)}\right) + \sum_{n=2}^{N-1} a\left(\frac{\gamma_1}{a(1+\beta)} + q^*\left(\frac{\beta}{1+\beta}\right) + x_n^*\right)x_n^* - \frac{a}{2}\left(\frac{\gamma_1}{a(1+\beta)} + q^*\left(\frac{\beta}{1+\beta}\right)\right)^2\right]. \tag{126}$$

On the other hand, after reordering the terms, constraint (119) becomes

$$x_n^* \leq \frac{\sigma}{a} - \frac{\gamma_1}{a(1+\beta)} - q^*\left(\frac{\beta}{1+\beta}\right), \qquad n = 2, \ldots, N-1. \tag{127}$$

The right hand side in (127) is *non-negative*. In fact, since $\sigma \geq \gamma_1$ and due to (118), we have

$$(\sigma - \gamma_1) + \beta(\sigma - aq^*) \geq 0 \quad \Rightarrow \quad \sigma(1+\beta) \geq \gamma_1 + a\beta q^* \quad \Rightarrow \quad \frac{\sigma}{a} \geq \frac{\gamma_1}{a(1+\beta)} + q^*\left(\frac{\beta}{1+\beta}\right). \tag{128}$$

Replacing (125), (126), and (127) in problem (113)-(122), it reduces to the following problem

$$\underset{\gamma, a, N, q^*}{\text{minimize}} \quad \frac{2a}{\sigma^2}\left[\frac{\sigma(\gamma_1 - aq^*)}{a(1+\beta)} + \sum_{n=2}^{N-1} a\left(\frac{\gamma_1}{a(1+\beta)} + \frac{q^*\beta}{1+\beta} + x_n^*\right)x_n^* - \frac{a}{2}\left(\frac{\gamma_1}{a(1+\beta)} + \frac{q^*\beta}{1+\beta}\right)^2\right] \tag{129}$$

$$\text{subject to} \quad \sum_{n=2}^{N-1} x_n^* = q^*, \tag{130}$$

$$\gamma_1 + \gamma_N = \sigma, \tag{131}$$

$$\gamma_1 \geq aq^*, \tag{132}$$

$$0 < \gamma_n \leq \sigma, \quad n = 2, \ldots, N-1, \tag{133}$$

$$\gamma_N \leq \gamma_1 \leq \left(1 + \frac{1}{\beta}\right)\gamma_N - \beta aq^* \tag{134}$$

$$x_n^* \leq \frac{\sigma}{a} - \frac{\gamma_1}{a(1+\beta)} - q^*\left(\frac{\beta}{1+\beta}\right), \quad n = 2, \ldots, N-1, \tag{135}$$

$$x_n^* \geq 0, \quad n = 2, \ldots, N-1. \tag{136}$$

Problem (129)-(136) is *symmetric* in $x_2^*, \ldots, x_N^*$. Therefore, the worst-case efficiency occurs when at Nash equilibrium we have

$$x_2^* = x_3^* = \ldots = x_{N-2}^* = x_{N-1}^* = \frac{q^*}{N-2}. \tag{137}$$

From (135) and (127), it is required that

$$\frac{q^*}{N-2} \le \frac{\sigma}{a} - \frac{\gamma_1}{a(1+\beta)} - q^*\left(\frac{\beta}{1+\beta}\right). \tag{138}$$

By replacing (137) in (129), the objective function in (129) reduces to

$$\frac{2a}{\sigma^2}\left[\sigma\left(\frac{\gamma_1-aq^*}{a(1+\beta)}\right) + \sum_{n=2}^{N-1} a\left(\frac{\gamma_1}{a(1+\beta)} + \frac{q^*\beta}{1+\beta} + \frac{q^*}{N-2}\right)\frac{q^*}{N-2} - \frac{a}{2}\left(\frac{\gamma_1}{a(1+\beta)} + \frac{q^*\beta}{1+\beta}\right)^2\right]. \tag{139}$$

The objective function in (139) is *increasing* in N and constraint (138) becomes *less restrictive* as N *increases*. Thus, the worst-case efficiency occurs as $N \to \infty$. We also notice that

$$\lim_{N\to\infty} \sum_{n=2}^{N-1} a\left(\frac{\gamma_1}{a(1+\beta)} + \frac{q^*\beta}{1+\beta} + \frac{q^*}{N-2}\right)\frac{q^*}{N-2} = aq^*\left(\frac{\gamma_1}{a(1+\beta)} + \frac{q^*\beta}{1+\beta}\right). \tag{140}$$

On the other hand, as $N \to \infty$, constraint (138) becomes $0 \le \frac{\sigma}{a} - \frac{\gamma_1}{a(1+\beta)} - \frac{q^*\beta}{1+\beta}$. However, we already know that this constraint always holds as shown in (126). Thus, we can simply *eliminate* this constraint. By replacing (140) in (139), the objective function becomes

$$\frac{2a}{\sigma^2}\left[\sigma\left(\frac{\gamma_1-aq^*}{a(1+\beta)}\right) + aq^*\left(\frac{\gamma_1}{a(1+\beta)} + \frac{q^*\beta}{1+\beta}\right) - \frac{a}{2}\left(\frac{\gamma_1}{a(1+\beta)} + \frac{q^*\beta}{1+\beta}\right)^2\right]. \tag{141}$$

For notational simplicity, we define $\bar{q}^* = aq^*$. After reordering the terms, (141) becomes

$$\frac{2}{\sigma^2(1+\beta)}\left(\gamma_1\sigma - \frac{\gamma_1^2}{2(1+\beta)} + \bar{q}^{*2}\beta\left(1 - \frac{\beta}{2(1+\beta)}\right) + \bar{q}^*\left(-\sigma + \gamma_1\left(\frac{1}{1+\beta}\right)\right)\right). \tag{142}$$

By looking at the remaining constraints, we can see that constraint (118) can be written as

$$\gamma_1 \ge \bar{q}^* \ge 0. \tag{143}$$

Given the definition of $\bar{q}^*$ and by combining constraints (117) and (120), we also have

$$\frac{\sigma}{2} \le \gamma_1 \le \frac{1+\beta}{1+2\beta}\sigma - \frac{\beta^2\bar{q}^*}{1+2\beta}. \tag{144}$$

Next, we take the derivative of the objective function in (142) with respect to γ_1 which yeilds

$$\frac{2}{\sigma^2(1+\beta)}\left(\sigma - \frac{\gamma_1}{1+\beta} + \frac{\bar{q}^*}{1+\beta}\right) \ge \frac{2}{\sigma^2(1+\beta)}\left(\sigma - \gamma_1 + \frac{\bar{q}^*}{1+\beta}\right) \ge 0, \tag{145}$$

where the last inequality is due to $\sigma \ge \gamma_1$ and $\bar{q}^* \ge 0$. Since the derivative is *non-negative*, the

worst-case efficiency occurs at the *lower bound* of γ_1. Comparing constraints (143) and (144),

$$\frac{\sigma}{2} \le \bar{q}^* \quad \Rightarrow \quad \gamma_1 = \bar{q}^*, \tag{146}$$

and

$$\frac{\sigma}{2} \ge \bar{q}^* \quad \Rightarrow \quad \gamma_1 = \frac{\sigma}{2}. \tag{147}$$

If (146) holds, problem (129)-(136) becomes

$$\underset{\bar{q}^*}{\text{minimize}} \quad \frac{\bar{q}^{*2}}{\sigma^2} \tag{148}$$

$$\text{subject to} \quad \frac{\sigma}{2} \le \bar{q}^* \le \frac{\sigma}{1+\beta}. \tag{149}$$

On the other hand, if (147) holds, problem (129)-(136) becomes

$$\underset{\bar{q}^*}{\text{minimize}} \quad \frac{2}{\sigma^2(1+\beta)} \left(\frac{3}{8(1+\beta)}\sigma^2 + \bar{q}^{*2}\frac{\beta(2+\beta)}{2(1+\beta)} - \sigma\frac{1+2\beta}{2(1+\beta)} \right) \tag{150}$$

$$\text{subject to} \quad 0 \le \bar{q}^* \le \frac{\sigma}{2}. \tag{151}$$

The objective function in (148) is *increasing* in $\bar{q}^*$, while the objective function in (150) is *decreasing* in $\bar{q}^*$. Thus, for both optimization problems (148)-(149) and (150)-(151), the worst-case efficiency occurs at $\bar{q}^* = \frac{\sigma}{2}$. Exploiting this in (148) and (150), the worst-case efficiency when $\gamma_1 + \gamma_N = \sigma$ and inequalities (110) and (111) hold is obtained as

$$\frac{1}{\sigma^2}\left(\frac{\sigma}{2}\right)^2 = \frac{1}{4}. \tag{152}$$

Interestingly, the worst-case efficiency in this scenario does *not* depend on the choice of pricing parameter β. Next, assume that condition (110) holds and we have

$$\gamma_1 \le aq^*. \tag{153}$$

From Theorem 6(a), the Nash equilibria are obtained as $0 \le x_1^* = x_N^* \le \frac{\gamma_N - \beta aq^*}{\beta a}$. We can show that, in this scenario, the worst-case efficiency occurs if $N \to \infty$ and we have $x_1^* = x_N^* = 0$ and $aq^* = \frac{1+2\beta\sigma}{2\beta^2+4\beta+3}$. Thus, the worst-case efficiency when (110) and (153) hold is obtained as

$$\frac{2}{2\beta^2 + 4\beta + 3}. \tag{154}$$

Details are omitted for brevity. Notice that if $\beta = \frac{1}{2}$, then (154) becomes $\frac{4}{11} \approx 0.36$.

Next, we assume that

$$\left(1 + \frac{1}{\beta}\right)\gamma_N - \beta aq^* \le \gamma_1 \le \frac{2}{\beta}\gamma_N - aq^*. \tag{155}$$

From Theorem 6(b), at Nash equilibrium we have $x_1^* = \frac{\gamma_N}{\beta a} - q$ and $x_N^* = \frac{\frac{2}{\beta}\gamma_N - \gamma_1}{a(1-\beta)} - \frac{q^*}{1-\beta}$. We can show that the worst-case efficiency when (155) holds is again as in (154). Finally, we assume

$$\gamma_1 \geq \frac{2}{\beta}\gamma_N - aq^*. \tag{156}$$

In that case, from Theorem 6, if $\gamma \geq aq^*$, then $x_1^* = \frac{\gamma_1}{2a} - \frac{q^*}{2}$ and $x_N^* = 0$. On the other hand, if $\gamma \leq aq^*$, then $x_1^* = x_N^* = 0$. In either case, the worst-case efficiency is still obtained as in (154).

Case II) We assume that $\gamma_1 + \gamma_N < \sigma$. Without loss of generality, we also assume that $\gamma_2 = \sigma$. In that case, we can show that the worst-case efficiency in this scenario becomes

$$\frac{2}{3} \approx 0.67. \tag{157}$$

In fact, the results in this case are similar to the results in Theorem 1 (see [8, Theorem 3]).

Combining the results from Case I and Case II, if pricing parameter $\beta = 1$, then

$$\text{PoA (Game 2, Problem 2)} \overset{\substack{\text{By (152), (154),}\\\text{and (157)}}}{=} \min\left\{\frac{1}{4}, \frac{2}{9}, \frac{2}{3}\right\} = \frac{2}{9}. \tag{158}$$

On the other hand, if $\beta = \frac{1}{2}$, then

$$\text{PoA (Game 2, Problem 2)} \overset{\substack{\text{By (152), (154),}\\\text{and (157)}}}{=} \min\left\{\frac{1}{4}, \frac{4}{11}, \frac{2}{3}\right\} = \frac{1}{4}. \tag{159}$$

This concludes the proof. ∎

I. Proof of Theorem 10

We first notice that at optimality, we always have $z_1^S = z_N^S = v_1^S = v_N^S$. This can be easily proved by contradiction. Therefore, we can rewrite the objective function of Problem 3 as

$$\sum_{n=1}^{N} \gamma_n y_n + z_1(\gamma_1 + \gamma_N) - \frac{a_1 + a_N}{2}z_1^2 - \frac{a}{2}\left(\sum_{n=1}^{N} y_n + z_1\right)^2. \tag{160}$$

The above objective function is *concave* in $y_1, \ldots, y_N$. Using KKT optimality conditions, we can show that for each $n \in \mathcal{N}\backslash\mathcal{M}$ we have $x_n^S = 0$ and for each $n \in \mathcal{M}$, we have

$$y_n^S = \begin{cases} \frac{1}{M}\left(\frac{\gamma_{\max}}{a} - z_1^S\right), & \text{if } \frac{\gamma_{\max}}{a} \geq z_1^S, \\ 0, & \text{otherwise.} \end{cases} \tag{161}$$

Part (a): Examining both cases $\frac{\gamma_{\max}}{a} \geq z_1$ and $\frac{\gamma_{\max}}{a} \leq z_1$, we can show that optimality occurs if $\frac{\gamma_{\max}}{a} \leq z_1$. In that case, $y_n^S = 0$ for all $n \in \mathcal{N}$ and the objective function in (160) becomes

$$z_1(\gamma_1 + \gamma_N) - \frac{a + a_1 + a_N}{2}z_1^2, \tag{162}$$

which is *concave* in z_1. Thus, at optimality, we have $z_1^S = \frac{\gamma_1 + \gamma_N}{a + a_1 + a_N}$.

Part (b): We can show that in this case, optimality occurs when $\frac{\gamma_{\max}}{a} \geq z_1$. Replacing (161) in (160), the objective function becomes

$$z_1 \left(\gamma_1 + \gamma_N - \gamma_{\max} \right) - \frac{a_1 + a_N}{2} z_1^2 + \frac{\gamma_{\max}^2}{2a}, \tag{163}$$

which is *concave* in z_1. Thus,

$$z_1^S = \frac{\gamma_1 + \gamma_N - \gamma_{\max}}{a_1 + a_N}. \tag{164}$$

By replacing (164) in (161), the data rates in (32) are resulted.

Part (c): Again, we can show that in this case, optimality occurs when $\frac{\gamma_{\max}}{a} \geq z_1$. The objective function is the same as that in (163). However, since $\gamma_1 + \gamma_N \leq \gamma_{\max}$, at optimality we have $z_1^S = 0$. Replacing this in (161), the data rates in (33) are obtained. ∎

J. Proof of Theorem 11

From (30), and given the payoff functions in Game 3, for each user $n \in \mathcal{N}$, we have

$$y_n^* = \begin{cases} \left(\gamma_n - a \sum_{r=1, r \neq n}^N y_r^* \right) / (2a), & \text{if } \gamma_n > a \sum_{r=1, r \neq n}^N y_r^*, \\ 0, & \text{if } \gamma_n \leq a \sum_{r=1, r \neq n}^N y_r^*. \end{cases} \tag{165}$$

First assume that

$$\gamma_1 + \gamma_N \geq \left(1 + \frac{a_1 + a_N}{a} \right) \gamma_{\max}. \tag{166}$$

In that case, from (31), the maximum network aggregate surplus becomes $(\gamma_1 + \gamma_N)^2 / (2(a + a_1 + a_N))$. The worst-case efficiency is obtained by solving the following optimization problem

$$\underset{\boldsymbol{y}^*, \boldsymbol{\gamma}, a, a_1, a_N, N, q^*}{\text{minimize}} \quad \frac{\sum_{n=1}^N \gamma_n y_n^* - \frac{a}{2} q^{*2}}{\frac{(\gamma_1 + \gamma_N)^2}{2(a + a_1 + a_N)}} \tag{167}$$

$$\text{subject to} \quad \gamma_n = aq^* + ay_n^*, \quad \text{if } y_n > 0, \quad n = 1, \ldots, N, \tag{168}$$

$$\gamma_n \leq aq^*, \quad \text{if } y_n = 0, \quad n = 1, \ldots, N, \tag{169}$$

$$\sum_{n=1}^N y_n^* = q^* \geq 0, \tag{170}$$

$$0 \leq \gamma_n \leq \gamma_{\max}, \quad n = 1, \ldots, N, \tag{171}$$

$$\gamma_1 + \gamma_N \geq \left(1 + \frac{a_1 + a_N}{a} \right) \gamma_{\max}, \tag{172}$$

$$y_n^* \geq 0, \quad n = 1, \ldots, N. \tag{173}$$

$$a, a_1, a_N > 0. \tag{174}$$

Since the objective function in (167) is *increasing* in $a_1 > 0$ and $a_N > 0$, the minimum occurs if $a_1 \to 0$ and $a_N \to 0$. We can also assume, without loss of generality, that $\gamma_n = aq^* + ay_n^*$ for each $n \in \mathcal{N}$. By replacing (168) in (167), (171), and (172), problem (167)-(174) becomes

$$\underset{\boldsymbol{y}^*, a, N, q^*}{\text{minimize}} \quad \frac{2a}{(2aq^* + ay_1^* + ay_N^*)^2} \left(a(q^* + y_1^*)y_1^* + a(q^* + y_N^*)y_N^* + \sum_{n=2}^{N-1} a(q^* + y_n^*)y_n^* - \frac{a}{2}q^{*2} \right) \quad (175)$$

$$\text{subject to} \quad \sum_{n=2}^{N-1} y_n^* = q^* - y_1^* - y_N^* \geq 0, \quad (176)$$

$$0 \leq y_n^* \leq \frac{\gamma_{\max}}{a} - q^*, \quad n = 1, \ldots, N, \quad (177)$$

$$aq^* \leq \gamma_{\max}, \quad (178)$$

$$q^* \geq y_1^* + y_N^*, \quad (179)$$

$$2q^* + y_1^* + y_N^* \geq \frac{\gamma_{\max}}{a}, \quad (180)$$

$$a > 0. \quad (181)$$

Problem (175)-(181) is *symmetric* in $y_2^*, \ldots, y_{N-1}^*$. Thus, the worst-case efficiency occurs if

$$y_2^* = y_3^* = \ldots = y_{N-2}^* = y_{N-1}^* = \frac{q^* - y_1^* - y_N^*}{N - 2}. \quad (182)$$

Furthermore, problem (175)-(181) is *decreasing* in N. Thus, the worst-case efficiency occurs when $N \to \infty$. From this, together with (182), optimization problem (175)-(181) reduces to

$$\underset{y_1^*, y_N^*, a, q^*}{\text{minimize}} \quad \frac{2(y_1^{*2} + y_N^{*2} + \frac{q^{*2}}{2})}{(2q^* + y_1^* + y_N^*)^2} \quad (183)$$

$$\text{subject to} \quad y_1^* + y_N^* \leq q^* \leq \frac{\gamma_{\max}}{a}, \quad (184)$$

$$2q^* + y_1^* + y_N^* \geq \frac{\gamma_{\max}}{a} \quad (185)$$

Problem (183)-(185) is *symmetric* in y_1 and y_N. In fact, we can show that the worst-case efficiency occurs at $y_1^* = y_N^* = \frac{q^*}{4}$. Replacing this in the objective function in (183), the worst-case efficiency when (166) holds is obtained as

$$2 \left(\frac{2 \times \frac{1}{16} + \frac{1}{2}}{(\frac{5}{2})^2} \right) = \frac{1}{5}. \quad (186)$$

We can also show that if $\gamma_{\max} \leq \gamma_1 + \gamma_N \leq \left(1 + \frac{a_1 + a_N}{a}\right) \gamma_{\max}$ or $\gamma_{\max} \geq \gamma_1 + \gamma_N$, then the worst-case efficiency is *equal* or *higher* (i.e., better) than $\frac{1}{5}$. In particular, if $\gamma_{\max} \geq \gamma_1 + \gamma_N$, then the worst-case efficiency is $\frac{2}{3}$ which resembles the results in [8]. Details are omitted here

for brevity. In summary, we have

$$\text{PoA}\,(\text{Game 3}, \text{Problem 3}) \quad \overset{\text{By (186)}}{=} \quad \min\left\{\frac{1}{5}, \frac{2}{3}\right\} = \frac{1}{5}. \tag{187}$$

This concludes the proof. ∎

REFERENCES

[1] R. Ahlswede, N. Cai, S. Li, and R. Yeung, "Network information flow," *IEEE Trans. on Information Theory*, vol. 46, pp. 1204–1216, Apr. 2000.

[2] L. Chen, T. Ho, S. Low, M. Chiang, and J. Doyle, "Optimization based rate control for multicast with network coding," in *Proc. of IEEE INFOCOM*, Anchorage, AK, May 2007.

[3] D. Traskov, N. Ratnakar, D. S. Lun, R. Koetter, and M. Medard, "Network coding for multiple unicasts: An approach based on linear optimization," in *Proc. of IEEE ISIT*, Seattle, WA, July 2006.

[4] A. Eryilmaz and D. S. Lun, "Control for inter-session network coding," in *Proc. of NetCod*, San Diego, CA, Jan. 2007.

[5] C. C. Wang and N. B. Shroff, "Beyond the butterfly: Graph-theoretic characterization of the feasibility of network coding with two simple unicast sessions," in *Proc. of IEEE ISIT*, Nice, France, June 2007.

[6] F. P. Kelly, A. Maulloo, and D. Tan, "Rate control for communication networks: Shadow prices, proportional fairness and stability," *Journal of Operations Research Society*, vol. 49, pp. 237–252, Mar. 1998.

[7] R. Johari and J. N. Tsitsiklis, "Efficiency loss in a network resource allocation game," *Mathematics of Operations Research*, vol. 29, pp. 407–435, Aug. 2004.

[8] ——, "A scalable network resource allocation mechanism with bounded efficiency loss," *IEEE J. on Selected Areas in Commun.*, vol. 24, no. 5, pp. 992–999, May 2006.

[9] Y. Chen and J. Zhang, "Design of price mechanisms for network resource allocation via price of anarchy," Technical Report, IEOR Department, University of California at Berkeley, 2008.

[10] T. Harks and K. Miller, "Efficiency and stability of Nash equilibria in resource allocation games," Technical Report, Department of Mathematics, Technische Universitt Berlin, 2008.

[11] S. Sanghavi and B. Hajek, "Optimal allocation of a divisible good to strategic buyers," in *Proc. of IEEE Conference on Decision and Control (CDC)*, Atlantis, Bahamas, Dec. 2004.

[12] D. Acemoglu and A. Ozdaglar, "Competition in parallel-serial networks," *IEEE J. on Sel. Areas in Commun.*, vol. 25, pp. 1180–1192, Aug. 2007.

[13] T. Roughgarden and E. Tardos, "How bad is selfish routing," *Journal of the ACM*, vol. 49, pp. 236–259, Feb. 2002.

[14] R. Johari, S. Mannor, and J. N. Tsitsiklis, "Efficiency loss in a network resource allocation game: The case of elastic supply," *IEEE Trans. on Automatic Control*, vol. 50, no. 11, pp. 1712–1724, Nov. 2005.

[15] S. Yang and B. Hajek, "VCG-Kelly mechanisms for allocation of divisible goods: Adapting VCG mechanisms to one-dimensional signals," in *Proc. of CISS*, Princeton, NJ, Mar. 2006.

[16] S. Bhadra, S. Shakkottai, and P. Gupta, "Min-cost selfish multicast with network coding," *IEEE Trans. on Information Theory*, vol. 52, pp. 5077–5087, Nov. 2006.

[17] Z. Li, "Cross-monotonic multicast," in *Proc. of IEEE INFOCOM*, Phoenix, AZ, Apr. 2008.

[18] X. Liang, "Matrix games in the multicast networks: Maximum information flows with network switching," *IEEE/ACM Trans. on Networking*, vol. 52, pp. 2433–2466, June 2006.

[19] Z. Li, "Min-cost multicast of selfish information flows," in *Proc. of IEEE INFOCOM*, Anchorage, AK, May 2007.

[20] X. Zhang and B. Li, "Dice: A game theoretic framework for wireless multipath network coding," in *Proc. of ACM MobiHoc*, Hong Kong, China, May 2008.

[21] J. Price and T. Javidi, "Network coding games with unicast flows," *IEEE Journal on Selected Areas in Commun.*, vol. 26, pp. 1302–1316, Sept. 2008.

[22] D. Samet and Y. Tauman, "The determination of marginal cost prices under a set of axioms," *Econometrica*, vol. 59, pp. 895–909, 1982.

[23] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. on Networking*, vol. 8, pp. 556–567, Oct. 2000.

[24] N. Shetty, G. Schwartz, and J. Walrand, "Network neutrality: Avoiding the extremes," in *Proc. of 46th Annual Allerton Conference on Communications, Control, and Computing*, Urbana-Champaign, IL, Sept. 2008.

[25] A. Mas-Colell, M. D. Whinston, and J. R. Green, *Microeconomic Theory*.   Oxford University Press, 1995.

[26] A. Khreishah, C. C. Wang, and N. B. Shroff, "Optimization based rate control for communication networks with inter-session network coding," in *Proc. of IEEE INFOCOM*, Phoenix, AZ, Apr. 2008.

[27] S. E. Rouayheb, A. Sprintson, and C. Georghiades, "On the relation between the index coding and the network coding problem," in *Proc. of IEEE ISIT*, Toronto, Canada, July 2008.

[28] S. Boyd and L. Vandenberghe, *Convex Optimization*.   Cambridge University Press, 2004.

[29] J. B. Rosen, "Existence and uniqueness of equilibrium points for concave *n*-person games," *Econometrica*, vol. 33, pp. 347–351, 1965.

[30] D. Fudenberg and J. Tirole, *Game Theory*.   The MIT Press, 1991.

TABLE I

Networking Setting	Routing Only	Network Coding and Routing with *Zero* Side Link Costs	Network Coding and Routing with *Non-zero* Side Link Costs
Optimization	Problem 1	Problem 2	Problem 3
Game	Game 1	Game 2	Game 3
Number of Nash Equilibria	One (Unique)	Can be *infinite*	One (Unique)
Price-of Anarchy† (PoA)	$\frac{2}{3}$	$\frac{1}{4}$	$\frac{1}{5}$
Theorem	Theorem 1	Theorem 8	Theorem 11
Reference	[8]	This Paper	

$\dagger$ Here, the PoA for the network coding scenario with *zero* side link costs is calculated based on the assumption that we use *price discrimination* with parameter $\beta = 0.5$. If single pricing is used, i.e., if $\beta = 1$, then the PoA can be less (i.e., worse) than $\frac{1}{4}$, as discussed in Theorem 7.

TABLE II

LIST OF KEY NOTATIONS

PoA (Game Ω, Problem Ω)	Price-of-anarchy for Game Ω with respect to Problem Ω, where $\Omega = 1, 2, 3$.
$\mathcal{N}$	Set of all users in the network.
N	Number of all users in the network.
s_n, t_n	Transmitter and receiver nodes of user $n \in \mathcal{N}$, respectively.
(i, j)	Shared bottleneck link between intermediate nodes i and j.
x_n	Data rate of user $n \in \mathcal{N}$ in the networks in Figs. 1 and 2.
$\boldsymbol{x}_{-n}$	Vector of data rates of all users other than user n in the networks in Figs. 1 and 2.
$\boldsymbol{x}$	Vector of data rates of all users in the networks in Figs. 1 and 2.
$U_n(\cdot)$	Utility function of user $n \in \mathcal{N}$.
γ_n	Slope of linear utility function of user $n \in \mathcal{N}$.
$C(\cdot)$	Cost function of shared bottleneck link (i, j).
$p(\cdot)$	Price function of shared bottleneck link (i, j).
a	Price parameter, $p(q) = aq$.
μ	Price value for routed packets.
δ	Price value for network coded packets.
β	Price discrimination parameter.
P_n	Payoff function of user $n \in \mathcal{N}$ in Game 1.
Q_n	Payoff function of user $n \in \mathcal{N}$ in Game 2.
$\boldsymbol{x}^S$	Optimal solution for Problems 1 and 2.
$\boldsymbol{x}^*$	Nash equilibrium for Games 1 and 2.
$x_n^B(\cdot)$	Best response data rate for user $n \in \mathcal{N}$ in Games 1 and 2.
X_1, X_N	Packets/symbols sent from source nodes s_1 and s_N, respectively.
$X_1 \oplus X_N$	Packet/symbol obtained by joint encoding of packets/symbols X_1 and X_N.
$C_1(\cdot), C_N(\cdot)$	Cost functions of side links (s_1, t_N) and (s_N, t_1) in the network in Fig. 6.
$p_1(\cdot), p_N(\cdot)$	Price functions of side links (s_1, t_N) and (s_N, t_1) in the network in Fig. 6.
a_1, a_N	Price parameters, $p_1(q) = a_1 q$ and $p_N(q) = a_N q$.
y_n	Data rate for routed packets of user $n \in \mathcal{N}$ in the network in Fig. 6.
z_1, z_N	Data rate for encoded packets of users 1 and N on link (i, j) in the network in Fig. 6.
v_1, v_N	Data rate for remedy packets of users 1 and N on the side links in the network in Fig. 6.
$\boldsymbol{y}_{-n}$	Vector of data rates for routed packets of all users other than user n in the network in Fig. 6.
W_n	Payoff function of user $n \in \mathcal{N}$ in Game 3.
$\boldsymbol{y}^S, \boldsymbol{z}^S, \boldsymbol{v}^S$	Optimal solution for Problem 3.
$\boldsymbol{y}^*, \boldsymbol{z}^*, \boldsymbol{v}^*$	Nash equilibrium for Game 3.
$y_n^B(\cdot)$	Best response data rate for routed packets of user $n \in \mathcal{N}$ in Game 3.
$z_1^B(\cdot), z_N^B(\cdot)$	Best response data rate for encoded packets of users 1 and N in Game 3.
$v_1^B(\cdot), v_N^B(\cdot)$	Best response data rate for remedy packets of users 1 and N in Game 3.

Fig. 1. A single wireline bottleneck link shared by N *routing* flows [8]. Transmission of packets over bottleneck link at each unit of data rate incurs at cost of $C(\sum_{n=1}^{N} x_n)$ and imposes price $p(\sum_{n=1}^{N} x_n)$. Elastic data rates $x_1, \ldots, x_N$ are selected by users $1, \ldots, N$, respectively.

Fig. 2. A butterfly network (cf. [5], [26]) with a single bottleneck link shared by N flows from N users and two side links. Transmission of packets over bottleneck link at each unit of data rate incurs at cost of $C\left(\sum_{n=2}^{N-1} x_n + \max(x_1, x_N)\right)$ and imposes a price $p\left(\sum_{n=2}^{N-1} x_n + \max(x_1, x_N)\right)$. Since the source node s_1 of user 1 is located closer to the destination node t_N of user N (and vice versa), users 1 and N can jointly perform *inter-session network coding* and reduce the traffic load on the shared bottleneck link (i, j) (and reduce cost C). The side links (s_1, t_N) and (s_N, t_1) are assumed to be free of charge in this setting. Here, x_1 and x_N denote the data rates of source nodes s_1 and s_N, respectively. On the other hand, X_1 and X_N denote the actual packets/symbols sent from nodes s_1 and s_N, respectively. The notation $X_1 \oplus X_N$ indicates a network coded packet/symbol obtained by jointly encoding packets X_1 and X_N.

(a) β = 1

(b) β = 0.5

Fig. 3.   Nash equilibria for the resource allocation game in Fig. 2, i.e., Game 2, when there are $N = 2$ (network coding) users: (a) Without price discrimination, i.e., $\beta = 1$, (b) With price discrimination when $\beta = \frac{1}{2}$. The rest of the parameters are as follows: $a = 1$, $\gamma_1 \geq \gamma_2$, and $\gamma_2 = 1$. If γ_1 and γ_2 are close (e.g., $\gamma_2 \leq \gamma_1 \leq 2\gamma_2$ for $\beta = 1$ and $\gamma_2 \leq \gamma_1 \leq 3\gamma_2$ for $\beta = \frac{1}{2}$), there are multiple Nash equilibria as indicated by the *shaded* area, following the model in (20) with $q^* = 0$. For example, if $\beta = \frac{1}{2}$ and $\gamma_1 = \gamma_2 = 1$, then *any* choice of data rates $x_1^* = x_2^*$ between $\frac{2}{3}$ and 2 is a Nash equilibrium. Therefore, in this case, there is an *infinite* number of Nash equilibria. Recall from Theorem 1(a) that in Game 1, i.e., the resource allocation with routing-only users, the Nash equilibrium is unique. This is one of the key differences between Game 1 and Game 2.

Fig. 4. Efficiency at a Nash equilibrium of 200 randomly generated resource allocation game scenarios when the network topology is as in Fig. 2 and the number of users $N = 2$. Notice that for each scenario, the efficiency is obtained as the ratio of the network aggregate surplus at Nash equilibrium of Game 2 and the network aggregate surplus at optimal solution of Problem 2. Here, we set either $\beta = 1$ or $\beta = \frac{1}{2}$, where β is the price discriminating parameter. For each scenario, the pricing parameter $a \in (0, 10)$ is selected randomly. The utility functions U_1 and U_2 are chosen to be α-fair (cf. [23]) with randomly selected choices of utility parameter $\alpha \in (0, 1)$. Note that the PoA is defined as the *worst-case* efficiency. We can see that if $\beta = 1$, the PoA is equal to $\frac{1}{3} \approx 0.33$. On the other hand, if $\beta = \frac{1}{2}$, then the PoA is equal to $\frac{12}{25} = 0.48$. These results confirm Theorem 7.

Fig. 5. Nash equilibria for the resource allocation game in Fig. 2 when $N = 3$, $a = 1$, $\beta = \frac{1}{2}$, $\gamma_1 \geq \gamma_3$, $\gamma_3 = 1$, and $\gamma_2 = 3$. Users 1 and 3 jointly perform inter-session network coding and user 2 is a routing user: (a) Transmission rates for users 1 and 3, (b) Transmission rates for user 2. If $\gamma_1 \leq 2$, then $\gamma_2 \geq \gamma_1 + \gamma_3$. In that case, at optimality of Problem 2, link (i, j) should only carry the packets from routing user 2. If $\gamma_1 > 2$, then $\gamma_2 < \gamma_1 + \gamma_3$. In that case, at optimality, link (i, j) should only carry the packets from network coding users 1 and 3.

Fig. 6. A single link shared by N flows. Users 1 and N perform inter-session network coding. The side links (s_1, t_N) and (s_N, t_1) have non-zero cost. Thus, users 1 and N need to pay for their transmissions over links (s_1, t_N) and (s_N, t_1), respectively. Here, y_1 and z_1 denote the data rate at which source s_1 sends data to intermediate node i *marked* for routing and network coding, respectively. Similarly, y_N and z_N denote the data rate at which source s_N sends data to node i *marked* for routing and network coding, respectively. Node i jointly encodes only those packets which are marked for network coding. On the other hand, v_1 and v_N denote the data rates at which sources s_1 and s_N send remedy packets over side links (s_1, t_N) and (s_N, t_1), respectively. The routing users $2, \ldots, N-1$ send packets to node i all marked for routing with rates $y_2, \ldots, y_{N-1}$, respectively. Notation $X_1 \oplus X_N$ indicates a network coded packet/symbol obtained by jointly encoding packets X_1 and X_N.

Fig. 7.   Efficiency at Nash equilibrium of Game 3 when the network topology is as in Fig. 6 and we have: $N \to \infty$, $\gamma_1 = \gamma_N = 1$, $\gamma_n = \frac{4}{5}$ for all $n \in \mathcal{N} \backslash \{1, N\}$, and $a = 1$. Side link price parameters $a_1 = a_N$ vary from 0 (*non-inclusive*) to 10. Notice that for each choice of parameters $a_1 = a_N$, the efficiency is obtained as the ratio of the network surplus at Nash equilibrium of Game 3 and the network surplus at optimal solution of Problem 3. We can see that the worst-case efficiency in this case is equal to $\frac{1}{5}$, as expected from Theorem 11. As the prices on the side links increase, network coding becomes less beneficial and the efficiency increases to a value (in this case $\frac{4}{5} = 0.8$) which is higher than $\frac{2}{3}$, as expected from Theorem 1.

Fig. 8. Efficiency at Nash equilibrium of 200 randomly generated resource allocation game scenarios when the network topology is as in Fig. 6 and where $N = 2$. For each scenario, the pricing parameters $a \in (0, 10)$, $a_1 \in (0, 5)$, and $a_2 \in (0, 5)$ are selected randomly. The utility functions U_1 and U_2 are chosen to be α-fair (cf. [23]) with randomly selected choices of utility parameter $\alpha \in (0, 1)$. Note that the PoA is defined as the *worst-case* efficiency. We can see that the efficiency of Game 3 is always lower bounded by $\frac{1}{5} = 0.2$, i.e., the PoA is $\frac{1}{5}$, as expected from Theorem 11.