Bilateral and Multilateral Exchanges for Peer-Assisted Content Distribution

Christina Aperjis, Ramesh Johari, and Michael J. Freedman

Users of the BitTorrent file sharing protocol and its variants are incentivized to contribute their upload capacity in a bilateral manner: downloading is possible in return for uploading to the same user. An alternative is to use multilateral exchange to match user demand for content to available supply at other users in the system. We provide a formal comparison of peer-to-peer system designs based on bilateral exchange with those that enable multilateral exchange via a price-based market mechanism to match supply and demand.

First, we compare the two types of exchange in terms of the equilibria that arise. A multilateral equilibrium allocation is Pareto efficient, while we demonstrate that bilateral equilibrium allocations are not Pareto efficient in general. We show that Pareto efficiency represents the "gap" between bilateral and multilateral equilibria: a bilateral equilibrium allocation corresponds to a multilateral equilibrium allocation if and only if it is Pareto efficient; our proof exploits the fact that Pareto efficiency implies reversibility of an appropriately constructed Markov chain.

Second, we compare the two types of exchange through the expected percentage of users that can trade in a large system, assuming a fixed file popularity distribution. Our theoretical results as well as analysis of a BitTorrent dataset provide quantitative insight into regimes where bilateral exchange may perform quite well even though it does not always give rise to Pareto efficient equilibrium allocations.

1 Introduction

Peer-to-peer systems have been extremely successful as a disruptive technology for content distribution. Varying accounts place peer-to-peer traffic as comprising anywhere between 35% and 90% of "all" Internet traffic, with BitTorrent accounting for its large majority [6]. Early peer-to-peer systems did not provide any incentives for participation, leading to extensive *free riding*: many peers were using the resources of other peers without contributing their own. For instance, early data showed that nearly 70 percent of peers of Gnutella were sharing no files, and nearly 50 percent of all responses were returned by the top 1 percent of sharing hosts [1]. A more recent study showed that 85 percent of Gnutella peers were sharing no files [12].

The peer-to-peer community responded with mechanisms to prevent free riding by incentivizing sharing on a *bilateral barter* basis, as used by BitTorrent [8] and its variants [18, 23]. According to the BitTorrent protocol, each user splits its available upload rate among users from which it gets the highest download rates. As a result, an increase in the upload rate to one user may increase the download rate from that particular user; however, it does not increase the download rate from other users. Thus, two users are incentivized to exchange only if each has content the other wants.

The difficulties of bilateral exchange (or barter) in an economy have been long known. Several inconveniences arise in the absence of money, the most important being the improbability of coincidence between persons wanting and possessing [13]. In modern economies, the aforementioned difficulty is eliminated by the use of *money*. Money can enable multilateral exchange by serving as a medium of exchange and a common measure of value. Even though modern societies take the use of money for granted, this is not the case in peer-to-peer systems, partly because of the associated system design complexity.

Peer-to-peer systems could potentially also use *market-based multilateral exchange* to match user demand for content to available supply at other users in the system. This can be done by using virtual currency and assigning a budget to each user that decreases when downloading and increases when uploading. Monetary incentives in a virtual currency have been previously proposed to incentivize uploading in peer-to-peer systems [11, 22, 20, 3]; however such designs are usually more complex than bilateral protocols and are not widespread. Thus, there is a significant tradeoff: *bilateral exchange without money is simple; on the other hand multilateral exchange allows more users to trade*. In this paper, we provide a formal comparison of two peer-to-peer system designs: bilateral barter systems, such as BitTorrent; and a market-based exchange of content enabled by a price mechanism to match supply and demand.

We start in Section 2 with a fundamental abstraction of content exchange in systems like BitTorrent: *exchange ratios*. The exchange ratio from one user to another gives the download rate received per unit upload rate. Exchange ratios are a useful formal tool because they allow us to define and study the equilibria of bilateral exchange. In a bilateral equilibrium each user optimizes with respect to exchange ratios. In Section 3 we also define multilateral equilibria, where users optimize with respect to prices; our definition of multilateral equilibrium is the same as competitive equilibrium in economics [15].

In Section 4 we compare bilateral and multilateral peer-to-peer systems through the allocations that arise at equilibria. A multilateral equilibrium allocation is always Pareto efficient, while bilateral equilibria may be inefficient. Our main result in this section is that a bilateral equilibrium allocation is Pareto efficient if and only if it is a multilateral equilibrium allocation. This result provides formal justification of the efficiency benefits of multilateral equilibria. The proof exploits an interesting connection between equilibria and Markov chains: an important step of the proof is to show that Pareto efficiency of a bilateral equilibrium rate allocation implies reversibility of an appropriately defined Markov chain, and that this chain has an invariant distribution that corresponds to a price vector of a multilateral equilibrium.

In Section 5 we perform a quantitative comparison of bilateral and multilateral exchange. As discussed in [13], "there may be many people wanting, and many possessing those things wanted; but to allow of an act of barter, there must be a double coincidence, which will rarely happen." We quantify how rarely this double coincidence of wants occurs under different assumptions on the popularity of files in the system. We first perform an asymptotic analysis assuming that file popularity follows a power law and study two extreme regimes. We find that asymptotically all users are able to trade bilaterally when the file popularity is very concentrated. On the other hand, multilateral exchange performs significantly better than bilateral when the file popularity is not concentrated. We complement our theoretical analysis by studying file popularity from a BitTorrent dataset. Although bilateral equilibria may in general be inefficient, the gap between bilateral and multilateral exchange can be narrowed significantly if each user shares a sufficient number of files (in practice as small as ten).

Our work is related to the study of equilibria in economies where not all trades are allowed. Kakade et. al. introduce a graph-theoretic generalization of classical Arrow-Debreu economics, in which an undirected graph specifies which consumers or economies are permitted to engage in direct trade [14]. However, the inefficiencies of bilateral exchange do not arise in their model. Finally, the monetary economics literature has studied how money reduces the double coincidence problem. The implementation of a competitive equilibrium is a central theme in this literature. The superiority of monetary exchange has been studied [21], and dynamics of bilateral trading processes have been considered [17, 10]. The transactions role of money is surveyed in [16].

2 Exchange Ratios in Bilateral Protocols

Many peer-to-peer protocols enable exchange on a *bilateral* basis between users: a user i uploads to a user j if and only if user j uploads to user i in return. Of course, such an exchange is only possible if each user has something the other wants. The foremost examples of such a protocol are BitTorrent and its variants. While such protocols are traditionally studied solely through the rates that users obtain, in this section we provide an interpretation of these protocols through *exchange ratios*. As exchange ratios can be interpreted in terms of prices, these ratios allow us to compare bilateral barter-based peer-to-peer systems with multilateral price-based peer-to-peer systems.

Let r_{ij} denote the rate sent from user i to user j at a given point in time in a bilateral peer-to-peer protocol. We define the *exchange ratio* between user i and user j as the ratio $\gamma_{ij} = r_{ji}/r_{ij}$; this is the download rate received by i from j, per unit of rate uploaded to j. By definition, $\gamma_{ij} = 1/\gamma_{ji}$. Clearly, a rational user i would prefer to download from users with which he has higher exchange ratios.

The exchange ratio has a natural interpretation in terms of prices: an equivalent story emerges if we assume that users charge each other for content in a common monetary unit, but that all transactions are *settlement-free*, *i.e.*, no money ever changes hands. In this case, if user i charged user j a price p_{ij} per unit rate, the exchange of content between users i and j must satisfy:

$$p_{ij} r_{ij} = p_{ji} r_{ji}$$

We refer to p_{ij} as the *bilateral price* from i to j. Note that the preceding condition thus shows the exchange ratio is equivalent to the ratio of bilateral prices: $\gamma_{ij} = p_{ij}/p_{ji}$ (as long as the prices and rates are nonzero).

The preceding discussion highlights the fact that the rates in a bilateral peer-to-peer system can be interpreted via exchange ratios. Thus far we have assumed that *transfer rates* are given, and exchange ratios are computed from these rates. In the next section, we turn this relationship around: we explicitly consider an abstraction of bilateral peer-to-peer systems where users react to given exchange ratios, and compare the resulting outcomes to price-based multilateral exchange.

3 Bilateral and Multilateral Equilibria

In this section we define bilateral equilibrium (BE) and multilateral equilibrium (ME), i.e., the market equilibria corresponding to bilateral and multilateral exchange.

In the formal model we consider, a set of users U shares a set of files F. User i has a subset of the files $S_i \subseteq F$, and is interested in downloading files in $T_i \subseteq F - S_i$. Throughout, we use r_{ijf} to denote the rate at which user i uploads file f to user j. We then let $x_{if} = \sum_j r_{jif}$ be

Bilateral User Optimization:	**Multilateral User Optimization:**
maximize $v_i(\boldsymbol{x}_i, y_i)$	maximize $v_i(\boldsymbol{x}_i, y_i)$
subject to $x_{if} = \sum_j r_{jif}, \forall f$	subject to $x_{if} = \sum_j r_{jif} \forall f$
$y_i = \sum_{j,f} r_{ijf}$	$y_i = \sum_{j,f} r_{ijf}$
$\boldsymbol{r} \in \mathcal{X}$	$\boldsymbol{r} \in \mathcal{X}$
$\sum_f r_{jif} \leq \gamma_{ij} \sum_f r_{ijf} \forall j$	$\sum_{j,f} p_j r_{jif} \leq p_i \sum_{j,f} r_{ijf}$

Figure 1: Optimization problems for price-based exchange. The two optimization problems differ only in the last constraint (budget constraint).

the rate at which user i downloads file f. We denote the vector of download rates for user i by $\boldsymbol{x}_i = (x_{if}, f \in T_i)$. Let $y_i = \sum_{j,f} r_{ijf}$ be the total upload rate of user i. We measure the desirability of a download vector and an upload rate to user i by a *utility function* according to the following assumption.

Assumption 1 *The preference relation of a user on the set of feasible rate vectors is represented by a continuous strictly concave utility function $v_i : \Re_+^{|T_i|+1} \to \Re$, which is strictly increasing in each download rate x_{if}, $f \in T_i$; and strictly decreasing in the upload rate y_i.*

Each user is assumed to have a constraint on the available upload rate; let B_i denote this upper bound for user i. We assume that users do not face any constraint on their download rate; this is consistent with most end user connections today, where upload capacity is far exceeded by download capacity.[1]

Let

$$\mathcal{X} = \left\{ \boldsymbol{r} : \boldsymbol{r} \geq 0;\ r_{kjf} = 0 \text{ if } f \notin S_k;\ \sum_{j,f} r_{ijf} \leq B_i \ \forall i \in U \right\}$$

be the set of feasible rate vectors. In particular, this ensures that (1) all rates are nonnegative, (2) users only upload files they possess, and (3) each user does not violate his bandwidth constraint.

We start by considering users' behavior in bilateral schemes, given a vector of exchange ratios $(\gamma_{ij}, i, j \in U)$. User i solves the bilateral optimization problem given in Figure 1.[2] By contrast, in a *multilateral price-based exchange*, the system maintains one price per user, and users optimize with respect to these prices.[3] We denote the price of user i by p_i. Figure 1 also gives the user optimization problem in multilateral price-based exchange. Note that the first three constraints (giving download and upload rates and ensuring that the rate allocation is feasible) are identical to the bilateral user optimization. Only the last constraint is different. While the bilateral exchange implicitly requires user i to download only from those users to whom he uploads, no such constraint is imposed on multilateral exchanges: user i accrues capital for uploading, and he can spend this capital however he wishes for downloading.

For bilateral (resp., multilateral) exchange, an *equilibrium* is a combination of a rate allocation vector and an exchange ratio vector (resp., price vector) such that all users have solved their

[1] While in practice a constraint on download rate exists, we remove it for the purposes of analysis since in practice the binding constraint on user behavior is likely to be the upload rate constraint.

[2] Note that we allow users to bilaterally exchange content over multiple files, even though this is not typically supported by swarming systems like BitTorrent; in BitTorrent a single file is split into subpieces called chunks, and users exchange chunks.

[3] This is equivalent to having one price per file in our setting [2].

corresponding optimization problems. In this case, the exchange ratios (resp., prices) have exactly aligned supply and demand: for any i, j, f, the transfer rate r_{ijf} is simultaneously an optimal choice for both the uploader i and downloader j. In the next two subsections we provide formal definitions of equilibria for both models.

3.1 Bilateral Equilibrium

Definition 1 *The rate allocation $\boldsymbol{r}^* \in \mathcal{X}$ and the exchange ratios $\boldsymbol{\gamma}^* = (\gamma^*_{ij}, i, j \in U)$ with $\gamma^*_{ij} \cdot \gamma^*_{ji} = 1$ for all i, j, constitute a Bilateral Equilibrium (BE) if for each user i, $\boldsymbol{r}^*$ solves the Bilateral User Optimization problem given exchange ratios $\boldsymbol{\gamma}^*$.*

Definition 1 requires that (1) all users have optimized with respect to the exchange ratios, and (2) the market clears. Even though the market clearing condition is not explicitly stated, it is implicitly required, since the same vector $\boldsymbol{r}^*$ is an optimal solution of the bilateral optimization problems of all users.

We do not expect a BE to exist in general. For example, this is trivially the case if no pair of users has reciprocally desired files; i.e., if for every pair i, j either $S_i \cap T_j = \emptyset$ or $S_j \cap T_i = \emptyset$. To show existence we assume that every user can find every file he desires through bilateral trade. This is formalized in Assumption 2.

Assumption 2 *For every user i and every file $f \in T_i$ there exists a user j such that $f \in S_j$ and $T_j \cap S_i \neq \emptyset$.*

Proposition 1 *If Assumptions 1 and 2 hold, then a BE exists.*

3.2 Multilateral Equilibrium

We next give the definition of ME, which corresponds to the concept of competitive equilibrium in economics [15].

Definition 2 *The rate allocation $\boldsymbol{r}^*$ and the user prices $\boldsymbol{p} = (p^*_i, i \in U)$ with $p^*_i > 0$ for all $i \in U$ constitute a Multilateral Equilibrium (ME) if for each user i, $\boldsymbol{r}^*$ solves the Multilateral User Optimization problem given prices $\boldsymbol{p}$.*

Similarly to Definition 1, Definition 2 requires that (1) all users have optimized with respect to prices, and (2) the market clears. Even though the market clearing condition is not explicitly stated, it is implicitly required, since the same vector $\boldsymbol{r}$ is used in the optimization problems of all users.

Our model is closely related to exchange economies [15]. In an *exchange economy* there is a finite number of agents and a finite number of commodities. Each agent is endowed with a bundle of commodities, and has a preference relation on the set of commodity vectors. Given a price vector, each agent finds a vector of commodities to exchange that maximizes his utility. In particular, if $\boldsymbol{p}$ is the vector of prices and agent i has endowment $\boldsymbol{w}_i$, he sells it at the market and obtains wealth $\boldsymbol{p} \cdot \boldsymbol{w}_i$. Then the agent buys goods for his consumption at the same price (he may buy back some of the goods he sold).

A straightforward reformulation reveals that our model shares much in common with a standard exchange economy: it is *as if* agent i has B_i units of his own "good", priced at p_i. He can trade this for goods from other users on the open market at prices $\boldsymbol{p}$. With this interpretation, $B_i - y_i$ is the amount of his own good that he chooses to keep. However, notice that this is not a standard exchange economy, as the upload rate is not a true commodity; rather, the commodities are the rates of specific files that are uploaded. Since B_i imposes a *joint* constraint on the upload rates of these files, our model is a generalization of the standard exchange economy. In the remainder of this section, we show existence of ME for our model.

Our goal is to show that a ME exists. However, we do not expect equilibria to exist without any restrictions on the sets S_i and T_i of files being uploaded and downloaded, respectively, by user i. For example, suppose there is a file that some users want to download, but no user has available for upload. Then in general, such a file has positive demand, while supply is always be zero. Thus the excess demand for such a file is positive unless its price is sufficiently high. Setting a sufficiently high price is equivalent to considering a system without that file.

To avoid such pathological situations, we introduce a natural diversity assumption. We define the *user graph* as the directed graph $G = (V, E)$ with $V = U$, and $E = \{(i, j) : S_i \cap T_j \neq \emptyset\}$. In other words, G is a graph where nodes correspond to users. There is a directed edge from user i to user j if i has a file that j desires.

Assumption 3 *The user graph consists only of strongly connected components.*

Proposition 2 *If Assumptions 1 and 3 hold, then there exists a ME.*

4 Efficiency of Equilibria

This section rigorously analyzes the efficiency properties of bilateral and multilateral exchange. We assume users explicitly react to exchange ratios or prices, and we compare the schemes through their resulting equilibria.

A ME allocation is always Pareto efficient, *i.e.*, there is no way to increase the utility of some user without decreasing the utility of some other user; this is the content of the first fundamental theorem of welfare economics [15]. For completeness, we include the result here.

Theorem 1 *If the rate allocation $\boldsymbol{r}^*$ and the user prices $(p_i^*, i \in U)$ with $p_i^* > 0$ for all $i \in U$ constitute a ME, then the allocation $\boldsymbol{r}^*$ is Pareto efficient.*

A BE, on the other hand, may not be Pareto efficient, as the following example shows.

Example 1 *Consider a system with n users and n files, for $n > 2$. Each user i has file f_i and wants files f_{i+1} and f_{i-1}. The utility of user i is $v_i(x_{i,f_{i-1}}, x_{i,f_{i+1}}, y_i) = x_{i,f_{i-1}} + 4x_{i,f_{i+1}} + \ln(2 - y_i)$, i.e., user i wants the files of both user $i+1$ and user $i-1$, but derives a higher utility from the file of user $i+1$.*

We first consider a symmetric BE with exchange ratios $\gamma_{i,i+1}^* = 2$ and $\gamma_{i,i-1}^* = 1/2$. The equilibrium rates are $r_{i-1,i}^* = 1$ and $r_{i+1,i}^* = 1/2$, and the download rates are $x_{i,f_{i-1}}^* = 1$ and $x_{i,f_{i+1}}^* = 1/2$. The utility of each user i is $3 - \ln(2) \approx 2.3$. On the other hand, prices $p_i^* = 1$ for all i, and rates $r_{i+1,i}^* = 1.75$, $r_{i-1,i}^* = 0$ constitute a ME. The utility of each user is $7 - \ln(4) \approx 5.61$, i.e., significantly larger than the utility of a user at the BE. This demonstrates that the BE allocation is not Pareto efficient.

The previous examples show that BE may not be Pareto efficient. We next provide an example of a BE rate allocation that is Pareto efficient.

Example 2 *Consider a system with n users and n files, for $n > 2$. Each user i has file f_i and wants files f_{i+1} and f_{i-1}. The utility of user i is $v_i(x_{i,f_{i-1}}, x_{i,f_{i+1}}, y_i) = x_{i,f_{i-1}} + x_{i,f_{i+1}} + \ln(2 - y_i)$.*

We consider a symmetric BE with exchange ratios $\gamma^*_{i,i+1} = 1$ and $\gamma^*_{i,i-1} = 1$. The equilibrium rates are $r^*_{i-1,i} = 1/2$ and $r^*_{i+1,i} = 1/2$. The BE rate allocation is Pareto efficient. In particular, it corresponds to a ME: prices $p^*_i = 1$ for all i, and rates $r^*_{i+1,i} = 1/2$, $r^*_{i-1,i} = 1/2$ constitute a ME.

BE may be inefficient, while ME always have Pareto efficient allocations (Theorem 1). In Example 2 the BE rate allocation is Pareto efficient and corresponds to a ME. Our main result is that *a BE allocation is a ME allocation if and only if it is Pareto efficient.* In particular, if a BE allocation is Pareto efficient, then there exist "supporting prices", i.e., prices such that the BE rate allocation is optimal for the multilateral optimization problem of each user. Informally, Pareto efficiency represents the "gap" between BE and ME.

Proposition 3 *Assume that for every user i and any fixed $\boldsymbol{x_i}$, $v_i(\boldsymbol{x_i}, y_i) \to -\infty$ as $y_i \to B_i$. Let $(\boldsymbol{r}^*, \boldsymbol{\gamma}^*)$ be a BE. The rate allocation $\boldsymbol{r}^*$ is Pareto efficient if and only if there exists a price vector $\boldsymbol{p}$ such that $\boldsymbol{r}^*$ and $\boldsymbol{p}$ constitute a ME.*

Proposition 3 assumes that $v_i(\boldsymbol{x_i}, y_i) \to -\infty$ as $y_i \to B_i$ for every user i and every fixed $\boldsymbol{x_i}$. This assumption is needed so that upload capacity constraints do not bind at the BE. This is a reasonable assumption for a peer-to-peer setting, since we do not expect users to use all their upload capacity. We note that if the upload capacity constraint binds for some users, then there may exist Pareto efficient BE that do not correspond to ME, simply because users have already "maxed out" their available upload capacity.

We provide an overview of the proof of Proposition 3, which demonstrates an interesting connection between equilibria and Markov chains; the details of the proof are provided in the appendix. From a BE rate allocation $\boldsymbol{r}^*$ we construct a transition rate matrix $\boldsymbol{Q}$ such that $Q_{ij} = \sum_f r^*_{ijf}$ if $i \neq j$; and $Q_{ii} = -\sum_{j,f} r^*_{ijf}$. We first observe that $\boldsymbol{\pi Q} = 0$ implies that the multilateral budget constraint is satisfied with price vector $\boldsymbol{\pi}$; therefore for any invariant distribution $\boldsymbol{\pi}$, $\boldsymbol{r}^*$ is feasible for the multilateral optimization problem of every user when prices are equal to $\boldsymbol{\pi}$. We then show that there exists an invariant distribution of $\boldsymbol{Q}$ (say $\boldsymbol{p}$) such that $\boldsymbol{r}^*$ is an optimal solution of the multilateral optimization problem of each user when the prices are equal to $\boldsymbol{p}$. We conclude that $\boldsymbol{r}^*$ and $\boldsymbol{p}$ constitute a ME.

A key step of the proof is to show that Pareto efficiency of $\boldsymbol{r}^*$ implies *reversibility* of $\boldsymbol{Q}$. Let $\boldsymbol{\pi}$ be an invariant distribution of $\boldsymbol{Q}$. $\boldsymbol{Q}$ is reversible if and only if $\gamma^*_{ij} = \pi_i/\pi_j$ for all pairs of users i and j that trade at the BE. This means that if $\boldsymbol{Q}$ is reversible, then $\boldsymbol{r}^*$ solves the multilateral optimization problem of each user given prices $\boldsymbol{\pi}$ *if the user is restricted to trade with peers it trades at the BE.* The matrix corresponding to the BE allocation of Example 1 is not reversible, which implies that the BE allocation is not Pareto efficient. On the other hand, the matrix corresponding to the BE allocation of Example 2 is reversible, and the BE allocation is Pareto efficient and corresponds to a ME allocation.

5 Quantitative Comparison

Bilateral exchange may be particularly restrictive because a pair of users can exchange only if each has a file that the other wants. On the other hand, allowing multilateral exchange significantly increases the number of possible exchanges, and potentially increases the number of users that can trade, but is also associated with increased complexity. In this section we compare bilateral and multilateral exchange through the corresponding percentages of users that can trade. Though distinct from Pareto efficiency, this metric provides quantitative insight into the comparison of the two types of exchange. We characterize regimes where bilateral exchange performs very well with respect to this metric, and for which — as a result — it may not be worth the effort to use multilateral exchange.

We first perform an asymptotic analysis assuming that file popularity follows a power law. We find that asymptotically all users are able to trade bilaterally when the file popularity is very concentrated. We complement our theoretical analysis by studying file popularity from a BitTorrent dataset. We find that a very large percentage of users is able to trade bilaterally if each user is sharing a sufficiently large number of files; e.g., over 96% of users can trade if each user has 10 files (for the number of users in the dataset).

We start by formally defining the quantities we compare. For a given peer-to-peer system, we define the *system profile* to consist of the specification of which files each user possesses and desires, i.e., $\mathcal{P} = \{T_i, S_i, i \in U\}$. For simplicity, we consider settings where each user is interested in downloading one file, i.e., $|T_i| = 1$ for all $i \in U$. This assumption significantly simplifies the analysis, since we do not need to consider how a user's utility function depends on different files. We thus abstract from specific utility functions and focus on how much bilateral exchange restricts trade.

We say that user i *can trade bilaterally under* $\mathcal{P}$ if there exists some user j such that $S_i \cap T_j \neq \emptyset$ and $S_j \cap T_i \neq \emptyset$, i.e., if i and j have reciprocally desired files. Given a system profile $\mathcal{P}$, let $\rho_{BE}(\mathcal{P})$ be the percentage of users that cannot trade bilaterally. We note that $\rho_{BE}(\mathcal{P})$ is equal to the percentage of users that need to be removed from the system so that a BE exists for $\mathcal{P}$ (if Assumption 1 holds). The condition $\rho_{BE}(\mathcal{P}) = 0$ is equivalent to Assumption 3 when each user possesses one file, as we assume in this section.

Similarly, we say that user i *can trade multilaterally under* $\mathcal{P}$ if there exist users $k_1, k_2, ..., k_n$ such that $S_{k_j} \cap T_{k_{j+1}} \neq \emptyset$ for $j = 1, ..., n$; $S_i \cap T_{k_1} \neq \emptyset$ and $S_{k_n} \cap T_i \neq \emptyset$. In words, user i is able to trade multilaterally if and only if there exists a cycle of users starting (and ending) at i such that each user possesses a file that is desired by the next user in the cycle. Clearly, if user i can trade bilaterally under $\mathcal{P}$, then he can also trade multilaterally under $\mathcal{P}$. Let $\rho_{ME}(\mathcal{P})$ be the percentage of users that cannot trade multilaterally. We note that $\rho_{ME}(\mathcal{P})$ is equal to the percentage of users that need to be removed from the system so that a ME exists for $\mathcal{P}$ (if Assumption 1 holds). The condition $\rho_{ME}(\mathcal{P}) = 0$ is weaker than Assumption 3; however, it is sufficient for ME existence when each user possesses one file, as we assume in this section.

We assume that the system profile $\mathcal{P}$ is chosen according to some distribution that depends on the popularity of different files. We are interested in comparing the expected values of $\rho_{BE}(\mathcal{P})$ and $\rho_{ME}(\mathcal{P})$. In Section 5.1 we consider a large system and perform an asymptotic analysis. In Section 5.2 we use a BitTorrent dataset to derive file popularity distributions, and then compare the expected values of $\rho_{BE}(\mathcal{P})$ and $\rho_{ME}(\mathcal{P})$ through simulations.

5.1 Asymptotic Analysis

In this section we theoretically study bilateral and multilateral exchange in large systems. We compare the two types of exchange through the expected percentages of users that cannot trade. We focus on large systems, and consider the asymptotic regime as the number of files and users in the system becomes large.

We assume the files that users possess and desire are drawn from a Zipf file popularity distribution independently and identically for each user. Our motivation to study this distribution comes from the fact that Zipf's law has been observed in many settings, and has been suggested as a good model for file popularity (e.g., [7]). Zipf's law states that the popularity of the r-th largest occurrence is inversely proportional to its rank. We adjust this definition to our setting.

Definition 3 *File popularity has a* Zipf distribution *with parameter s if the r-th most popular file has probability proportional to r^{-s}.*

For example, if S_i and T_i are singletons, a user desires the i-th most popular file and possesses the j-th most popular file with probability $(i \cdot j)^{-s}/(\sum_{k' \neq k}(k' \cdot k)^{-s})$.

Note that $s = 0$ corresponds to the uniform distribution. On the other hand, as s increases the distribution becomes more concentrated.

We are interested in the expected percentage of users that cannot trade. This is a function of the number of users N, the number of files K and the Zipf exponent s. Let $\overline{\rho}_{BE}(K, N, s)$ and $\overline{\rho}_{ME}(K, N, s)$ be the expected percentages of users that cannot trade bilaterally and multilaterally. In particular, $\overline{\rho}_{BE}(K, N, s)$ (resp., $\overline{\rho}_{ME}(K, N, s)$) is the expected value of $\rho_{BE}(\mathcal{P})$ (resp., $\rho_{ME}(\mathcal{P})$) over system profiles.

We consider a sequence of peer-to-peer systems indexed by N. The Nth system has N users and $K(N)$ files, where $K(N)$ is a nondecreasing function of N. The function $K(N)$ represents how the number of files scales with the number of users. We study an asymptotic regime where $N \to \infty$.

Since the number of users that cannot trade bilaterally is always greater than or equal to the number of users that cannot trade multilaterally, we have $\overline{\rho}_{BE}(K, N, s) - \overline{\rho}_{ME}(K, N, s) \geq 0$. The following propositions imply that in a large system $\overline{\rho}_{BE}(K, N, s) - \overline{\rho}_{ME}(K, N, s)$ may be significant when $s = 0$, but is always negligible when $s > 1$.

Proposition 4 *Assume $s = 0$, i.e., files are chosen uniformly. Moreover, $|S_i| = \sigma$ and $|T_i| = 1$ for all $i \in U$.*

1. *If $K(N) \geq \sigma\sqrt{N}$ for large N, then there exists $\overline{N}$ such that $\overline{\rho}_{BE}(K(N), N, 0) \geq 1/e$ for $N \geq \overline{N}$.*

2. *If there exists $\varepsilon > 0$ such that $K(N) \leq \sigma N^{1/2-\varepsilon}$ for large N, then $\overline{\rho}_{BE}(K(N), N, 0) \to 0$ as $N \to \infty$.*

3. *If there exists $\varepsilon > 0$ such that $K(N) \leq \sigma N^{1-\varepsilon}$ for large N, then $\overline{\rho}_{ME}(K(N), N, 0) \to 0$ as $N \to \infty$.*

The case $K(N) \in [\sigma\sqrt{N}, \sigma N^{1-\varepsilon}]$ is of particular interest. According to Proposition 4 in this case $\overline{\rho}_{BE}(K(N), N, 0) \geq 1/e \approx 0.37$ for large N, while $\overline{\rho}_{ME}(K(N), N, 0) \to 0$ as $N \to \infty$. In words, when the system is large more than one third of the users cannot trade bilaterally, while almost all users can trade multilaterally. We conclude that if $\sigma N^{1-\varepsilon} \geq K(N) \geq \sigma\sqrt{N}$ and files are chosen uniformly, then multilateral exchange performs significantly better than bilateral exchange in terms of the number of users that can trade.

Proposition 5 *If $s > 1$, then $\overline{\rho}_{BE}(K(N), N, s) \to 0$ as $N \to \infty$ for any nondecreasing $K(N)$.*

Since $\overline{\rho}_{ME}(K(N), N, s) \leq \overline{\rho}_{BE}(K(N), N, s)$, we conclude that if files are chosen according to a Zipf distribution with $s > 1$ then both $\overline{\rho}_{BE}(K(N), N, s) \to 0$ and $\overline{\rho}_{ME}(K(N), N, s) \to 0$ as $N \to \infty$. We note that this result does not depend on the number of files that users possess. When $s > 1$, bilateral exchange performs very well asymptotically even if each user only possesses one file.

This is an interesting result: even though bilateral exchange significantly restricts trade compared to multilateral exchange, in expectation almost all users can trade under both types of exchange when the system is large and file popularity follows a Zipf distribution with exponent $s > 1$. The intuition behind this result is that when s is large, the popularity distribution is more concentrated, i.e., the most popular files are chosen with relatively high probability. As a result, for any user i both T_i and S_i probably consist of one of the most popular files, and it is more likely that there exists a user j, such that i and j have reciprocally desired files.

5.2 Data Analysis

In this section we quantitatively compare bilateral and multilateral exchange using a BitTorrent dataset collected by Piatek et. al. from the University of Washington [19]. We find that a significant percentage of users cannot trade bilaterally when each user is sharing one file; however, the percentage becomes negligible as peers share more files.

The dataset consists of 1,364,734 downloads, 679,523 users and 7,323 files. We use the number of downloads of each file in the dataset to estimate the probability that a given file is selected. We then use the estimated probabilities to generate system profiles and compute the percentages of users that cannot trade bilaterally and multilaterally. We assume that there are 7,323 files with the given distribution, and vary the number of users in the system.

We first assume that each user possesses and desires exactly one file, i.e., $|T_i| = |S_i| = 1$ for every $i \in U$. Figure 2 shows the percentages of users that cannot trade bilaterally and multilaterally from simulations for various numbers of users in the system.[4] We observe that the percentage of users that cannot trade bilaterally is significantly larger than the percentage of users that cannot trade multilaterally. Moreover, the latter is close to 0, indicating that most users are able to trade multilaterally. Finally, as the number of users increases, the percentages of users that can trade increase for both bilateral and multilateral exchange. For instance, when there are 200,000 users in the system, about 93% cannot trade bilaterally while only 0.7% cannot trade multilaterally. When the number of users increases to 1,000,000, about 75% cannot trade bilaterally and 0.06% cannot trade multilaterally.

We next assume each user desires one file and possesses multiple files. As the number of files that each user has increases, the number of possible trades increases, and as a result the percentage of users that can trade bilaterally increases. In Figure 3 we show the percentages of users that cannot trade bilaterally when each user desires one file ($|T_i| = 1$) and possesses multiple files (from simulations).

For the first figure in Figure 3 we assume that all users possess the same number of files, i.e., $|S_i| = |S_j|$ for all $i, j \in U$. The figure shows the percentages of users that cannot trade multilaterally for various values of $|S_i|$. The case $|S_i| = 1$ has already been considered in Figure 2. In Figure 3 we

[4]The algorithm we use to compute ρ_{BE} is exact: for every user i we check whether there is some user j such that i and j have reciprocally desired files. Computing the exact value of ρ_{ME} for a large system seems computationally intractable. Therefore, we use an approximation algorithm to compute ρ_{ME}: we recursively remove users that possess files not desired by others or desire files not possessed by others. Simulations for small numbers of users suggest that this algorithm provides a very good approximation for ρ_{ME}.

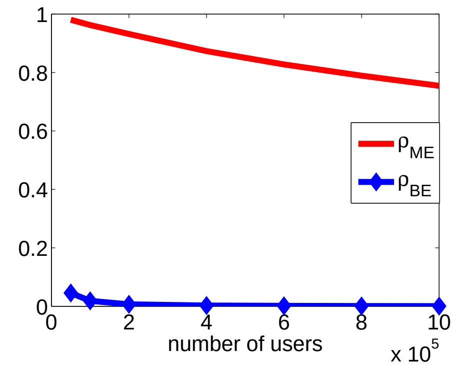

Figure 2: Percentages of users (from simulations) that cannot trade bilaterally and multilaterally when users desire and possess one file, i.e., $|T_i| = |S_i| = 1$ for all i. The horizontal axis shows the number of users in the system.

also consider $|S_i| \in \{2, 5, 10, 20\}$. We observe a significant decrease in the percentage of users that cannot trade when $|S_i|$ increases from 1 to 20. We can illustrate this by considering the minimum required number of users in the system so that at most 10% are not able to trade: at least 1,000,000 users are required when each user has 5 files; at least 200,000 users are needed in the system when each possesses ten files; at least 50,000 users are needed when each user possesses 20 files. We observe that there is a significant decrease in the required number of users. Moreover, when each user possesses 20 files, the percentage of users that cannot trade bilaterally is very small when there are more than 200,000 users in the system: 2.3% cannot trade with 200,000 users; only 0.28% cannot trade with 1,000,000 users.

Our simulations up to now have assumed that all users in the system possess the same number of files, i.e., $|S_i| = |S_j|$ for all i, j. We next assume that the number of files that users possess vary across different users. We are interested in whether the percentage of users that can trade bilaterally increase as the variance of the distribution of $|S_i|$ increases (assuming that the mean remains the same). At first it may seem plausible that users with very large $|S_i|$ would be able to accommodate a lot of trades and as a result ρ_{BE} should increase as the $|S_i|$'s become more dispersed. However, this is not the case as we discuss next.

We infer the distribution of $|S_i|$ from the dataset.[5] The mean value of $|S_i|$ in the dataset is 2.0084. Therefore, we are interested in whether ρ_{BE} increases compared to the case that $|S_i| = 2$ for all i. The second figure in Figure 3 shows the percentage of users that cannot trade bilaterally when the number of files that users possess ($|S_i|$) follows the distribution from the dataset. For comparison, we also show the percentages that cannot trade when $|S_i|$ is equal to 1 or 2 for all users. We observe that when $|S_i|$ is drawn from the distribution, the percentages are between the case $|S_i| = 1$ and $|S_i| = 2$ — even though the expected value of $|S_i|$ is slightly greater than two. In particular, about 81% of users cannot trade bilaterally when $|S_i|$ is drawn from the distribution, while 77% cannot trade when $|S_i| = 2$ for all users. The percentage of users with large $|S_i|$ that cannot trade bilaterally significantly decreases: 26% of users with $|S_i| > 10$ cannot trade bilaterally. However, only 2% of the users belong in this category; 67 % of all users have $|S_i| = 1$ and 88% of them cannot trade bilaterally.

[5] We assume that the number of files that a user possesses is equal to the number of files he downloads.

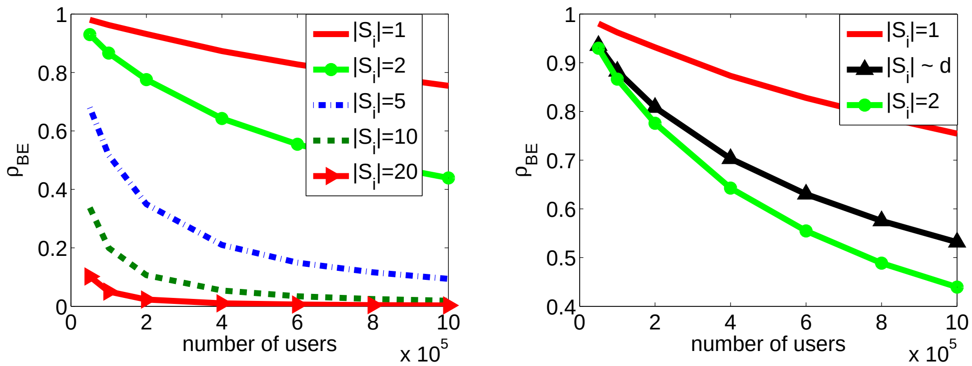

Figure 3: Percentages of users (from simulations) that cannot trade bilaterally when each user desires one file ($|T_i| = 1$) and possesses multiple files. The legend shows $|S_i|$ for each line. The horizontal axis shows the number of users in the system. In the left figure all users possess the same number of files ($|S_i| \in \{1, 2, 5, 10, 20\}$). In the right figure the number of files that a user possesses ($|S_i|$) is drawn from the dataset distribution (denoted by "d"). We also include the cases $|S_i| = 1$ and $|S_i| = 2$ for all users in the right figure.

To further examine the conjecture that a more dispersed distribution of $|S_i|$ further restricts bilateral trade, we run simulations assuming that $|S_i| = 1$ with probability 0.5 and $|S_i| = 3$ with probability 0.5, so that the mean value of $|S_i|$ is equal to 2. This distribution is less dispersed than the distribution we obtain from the data, but more dispersed than when $|S_i| = 2$ for all users. Indeed, the percentage of users than cannot trade bilaterally when there are 200,000 users is 78%, which is greater than the percentage of users that cannot trade when $|S_2| = 1$ for all users (77%), but smaller than the percentage of users than cannot trade under the data distribution (80%).

6 Conclusion

This paper provides a formal comparison of two peer-to-peer system designs: bilateral barter systems, such as BitTorrent; and a market-based exchange of content enabled by a price mechanism to match supply and demand. Our results demonstrate that even though bilateral equilibria are not in general Pareto efficient, bilateral exchange may perform very well in terms of the expected percentage of users that can trade for certain file probability distributions. Moreover, our data analysis shows a significant increase in the percentage of users that can trade bilaterally when each user shares multiple files. This suggests that bilateral incentives in BitTorrent would be much stronger if the protocol considered exchanges across different files.

Acknowledgments This work was partially supported by the National Science Foundation under grant 0428868. We are grateful to Michael Piatek et. al. for providing the BitTorrent dataset [19].

References

[1] E. Adar and B. Huberman. Free riding on Gnutella. *First Monday*, 5(10), 2000.

[2] C. Aperjis, M. J. Freedman, and R. Johari. Peer-assisted content distribution with prices. In *ACM CoNEXT*, 2008.

[3] C. Aperjis and R. Johari. A peer-to-peer system as an exchange economy. In *GameNets*, 2006.

[4] C. Aperjis, R. Johari, and M. J. Freedman. Bilateral and multilateral exchanges for peer-assisted content distribution. Technical report, Stanford University, 2009.

[5] L. M. Ausubel and R. J. Deneckere. A generalized theorem of the maximum. Technical Report 899, Northwestern University, Center for Mathematical Studies in Economics and Management Science, 1990.

[6] E. Bangeman. P2P responsible for as much as 90 percent of all 'Net traffic. *ArsTechnica*, Sep 3 2007.

[7] M. Cha, H. Kwak, P. Rodriguez, Y.-Y. Ahn, and S. Moon. I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system. In *IMC '07: Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, pages 1–14, New York, NY, USA, 2007. ACM.

[8] B. Cohen. Incentives build robustness in BitTorrent. In *Workshop on Economics of Peer-to-Peer Systems*, 2003.

[9] C. Cooper and A. Frieze. The size of the largest strongly connected component of a random digraph with a given degree sequence. *Combinatorics, Probability and Computing*, 2004.

[10] A. M. Feldman. Bilateral trading processes, pairwise optimality, and pareto optimality. *The Review of Economic Studies*, 40(4):463–473, 1973.

[11] M. Gupta, P. Judge, and M. Ammar. A reputation system for peer-to-peer networks. In *NOSSDAV '03: Proceedings of the 13th international workshop on Network and operating systems support for digital audio and video*, pages 144–152, New York, NY, USA, 2003. ACM Press.

[12] D. Hughes, G. Coulson, and J. Walkerdine. Free riding on Gnutella revisited: The bell tolls? *IEEE Distributed Systems Online*, 6(6), 2005.

[13] W. S. Jevons. *Money and the Mechanism of Exchange.* Kegan Paul, Trench, 1885.

[14] S. M. Kakade, M. Kearns, and L. E. Ortiz. Graphical economics. In *Learning Theory*, volume 3120/2004 of *Lecture Notes in Computer Science*, pages 17–32. Springer Berlin / Heidelberg, 2004.

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

[16] J. M. Ostroy and R. M. Starr. The transactions role of money. In B. M. Friedman and F. H. Hahn, editors, *Handbook of Monetary Economics.*

[17] J. M. Ostroy and R. M. Starr. Money and the decentralization of exchange. *Econometrica*, 42(6):1093–1113, 1974.

[18] M. Piatek, T. Isdal, T. Anderson, A. Krishnamurthy, and A. Venkataramani. Do incentives build robustness in BitTorrent? In *NSDI*, 2007.

[19] M. Piatek, T. Isdal, A. Krishnamurthy, and T. Anderson. One hop reputations for peer-to-peer file sharing workloads. In *NSDI*, 2008.

[20] M. Sirivianos, J. H. Park, X. Yang, and S. Jarecki. Dandelion: Cooperative content distribution with robust incentives. In *USENIX Technical*, 2007.

[21] R. M. Starr. The structure of exchange in barter and monetary economies. *The Quarterly Journal of Economics*, 86(2):290–302, 1972.

[22] V. Vishnumurthy, S. Chandrakumar, and E. G. Sirer. Karma: A secure economic framework for peer-to-peer resource sharing. In *Workshop on Economics of Peer-to-Peer Systems*, 2003.

[23] F. Wu and L. Zhang. Proportional response dynamics leads to market equilibrium. In *STOC*, 2007.

7 Proofs

Proof of Proposition 1: We first define the concept of restricted BE and show that such an equilibrium always exists. We then use the exchange ratios of the restricted BE to construct a BE according to Definition 1.

The rate allocation $\boldsymbol{r}^*$ and the exchange ratios $\boldsymbol{\gamma}^*$ constitute a *restricted BE* if

1. $\gamma_{ij}^* = 0$ if $S_i \cap T_j = \emptyset$ or $S_j \cap T_i = \emptyset$;
   $\gamma_{ij}^* \cdot \gamma_{ij}^* = 1$ otherwise.

2. For each user i, $\boldsymbol{r}^*$ solves the Bilateral User Optimization problem given exchange ratios $\boldsymbol{\gamma}^*$.

At a restricted BE all exchange ratios between peers that cannot trade bilaterally are set to zero. We show that a restricted BE exists under Assumption 1.

Let $E = \{(i,j) : T_i \cap S_j \neq \emptyset, T_j \cap S_i \neq \emptyset\}$ be the set of tuples of users with reciprocally desired files. According to Definition 1 $\gamma_{ij} > 0$ if and only if $(i,j) \in E$ at a BE. We consider a price p_{ij} for every tuple $(i,j) \in E$. This is the price that user j pays to download from user i (see Section 2). For this proof, let $\boldsymbol{p} = (p_{ij}, (i,j) \in E)$. The exchange ratio between i and j is $\gamma_{ij} = p_{ij}/p_{ji}$. In particular, without loss of generality we assume that the budget constraint in the bilateral user optimization of user i is replaced by

$$p_{ji} \sum_f r_{jif} = p_{ij} \sum_f r_{ijf}.$$

We ignore pairs of users that are not in E (since by definition such users cannot trade bilaterally) and show that it is possible to have some $\boldsymbol{p} \gg 0$ such that the market clears.

For the purposes of this proof, let $\boldsymbol{r}^i(\boldsymbol{p})$ be the optimal solution for the bilateral optimization problem of user i when the exchange ratios are equal to $\gamma_{ij} = p_{ij}/p_{ji}$. If $\boldsymbol{r}$ and $\boldsymbol{p}$ constitute a BE, then $\boldsymbol{r} \in \boldsymbol{r}^i(\boldsymbol{p})$ for all $i \in U$. We note that each r^i_{ijf} is in general a correspondence. We define excess demand for each $(i,j) \in E$ as

$$z_{ij}(\boldsymbol{p}) = \sum_f r^j_{ijf}(\boldsymbol{p}) - \sum_f r^i_{ijf}(\boldsymbol{p}).$$

We first show that the excess demand $\boldsymbol{z}$ has the following properties:

(i) For every $\boldsymbol{p}$ and $\boldsymbol{z} \in \boldsymbol{z}(\boldsymbol{p})$, $\boldsymbol{p} \cdot \boldsymbol{z}(\boldsymbol{p}) = 0$.

(ii) $\boldsymbol{z}(\cdot)$ is convex-valued

(iii) $\boldsymbol{z}(\cdot)$ is homogeneous of degree 0

(iv) $\boldsymbol{z}(\cdot)$ is upper-hemicontinuous

(v) There is $s > R$ such that $z_{ij} > -s$ for any $\boldsymbol{z} \in \boldsymbol{z}(\boldsymbol{p})$ and $\boldsymbol{p}$.

(vi) If $\boldsymbol{p}^n \to \boldsymbol{p} \neq 0$, $\boldsymbol{z}^n \in \boldsymbol{z}(\boldsymbol{p}^n)$ and $p_{ij} = 0$, $p_{ji} > 0$ for some $(i,j) \in E$, then

$$\max\{z^n_{ij} : (i,j) \in E\} \to \infty.$$

By Assumption 1, the budget constraint of each user binds. The budget constraint of user i is

$$p_{ji} \sum_f r^i_{jif}(\boldsymbol{p}) = p_{ij} \sum_f r^i_{ijf}(\boldsymbol{p}).$$

By summing over all users, we obtain Property (i).

Fix a price vector $\boldsymbol{p} \gg 0$. By Assumption 1 $v(\cdot)$ is strictly concave; therefore $r^i_{ijf}(\boldsymbol{p})$ and $r^j_{ijf}(\boldsymbol{p})$ are convex-valued. Thus the aggregate excess demand $\boldsymbol{z}(\cdot)$ is a convex valued correspondence (Property (ii)).

Consider a price vector $\boldsymbol{p} \gg 0$, and fix a constant $t > 0$. It is clear that the feasible region of the bilateral user optimization problem remains unchanged if we replace the price vector $\boldsymbol{p}$ by $t\boldsymbol{p}$. Thus the aggregate excess demand is homogeneous of degree zero (Property (iii)).

By Assumption 1 $v(\cdot)$ is a continuous function. From the Theorem of the Maximum [5] it follows that $r^i_{ijf}(\boldsymbol{p})$ and $r^i_{jif}(\boldsymbol{p})$ are upper hemicontinuous correspondences. The aggregate excess demand for $(i,j) \in E$ is a linear combination of the rates $r^j_{ijf}(\boldsymbol{p})$ and $r^i_{ijf}(\boldsymbol{p})$, and therefore is also upper hemicontinuous (Property (iv)).

The upload rate of any user i is upper bounded by his upload rate constraint B_i, so the total supply is upper bounded and the excess demand is bounded from below (Property (v)).

Suppose that $\boldsymbol{p}^n \to \boldsymbol{p} \neq 0$, and $p_{ij} = 0$, $p_{ji} > 0$ for some $(i,j) \in E$. Let $f \in T_i \cap S_{i+1}$. As $\boldsymbol{p}^m \to \boldsymbol{p}$ and the amount of f that user i can afford goes to infinity. On the other hand, the total possible supply is bounded above by the upload rate constraint of user j. Thus $\max\{z^m_{ij} : (i,j) \in E\} \to \infty$, establishing Property (vi).

Using properties (i)-(vi) we show that there exists a BE. Let

$$\Delta = \{\boldsymbol{p} \in R^{|E|}_+ : p_{ij} + p_{ji} = 1, (i,j) \in E\}$$

$$\Delta^n = \{\boldsymbol{p} \in \Delta : p_{ij} \geq 1/n, (i,j) \in E\}$$

We observe that Δ^n is compact. Then (from property (iv)) for each n, there exists $\boldsymbol{r}^n > 0$ such that $\boldsymbol{z}(\boldsymbol{p}) \subset [-\boldsymbol{r}^n, \boldsymbol{r}^n]^{|E|}$. For each n, define $\boldsymbol{f}^n : \Delta^n \times [-\boldsymbol{r}^n, \boldsymbol{r}^n]^{|E|} \to \Delta^n \times [-\boldsymbol{r}^n, \boldsymbol{r}^n]^{|E|}$ by

$$\boldsymbol{f}^n(\boldsymbol{p}, \boldsymbol{z}) = \{\boldsymbol{q} \in \Delta^n : \boldsymbol{z} \cdot \boldsymbol{q} \geq \boldsymbol{z} \cdot \boldsymbol{q}', \forall \boldsymbol{q}' \in \Delta^n\} \times \boldsymbol{z}(\boldsymbol{p}).$$

For each n, the correspondence $\boldsymbol{f}^n$ is convex-valued and upper-hemicontinuous. We can now apply Kakutani's theorem to conclude that for each n, $\boldsymbol{f}^n(\cdot)$ has a fixed point, which we denote by $(\boldsymbol{p}^n, \boldsymbol{z}^n)$.

The sequence $\boldsymbol{p}^n$ in Δ has a subsequence that converges, because Δ is compact. By (v) and the fact that $\boldsymbol{z}^n$ is bounded, the limit must be in the interior of Δ. Therefore, by taking a subsequence if necessary, we can assume that $\boldsymbol{p}^n \to \boldsymbol{p}^*$ and $\boldsymbol{z}^n \to \boldsymbol{z}^*$, where $\boldsymbol{p}^*$ is in the interior of Δ. The limit $\boldsymbol{p}^*$ is a BE price vector.

We have shown that a restricted BE exists under Assumption 1. We know show how to construct a BE (according to Definition 1) when Assumption 2 holds. Suppose $\tilde{\boldsymbol{r}}$ and $\tilde{\boldsymbol{\gamma}}$ constitute a restricted BE. Let $\boldsymbol{r}^* = \tilde{\boldsymbol{r}}$. For pairs of users i, j such that $S_i \cap T_j \neq \emptyset$ and $S_j \cap T_i \neq \emptyset$, set $\gamma^*_{ij} = \tilde{\gamma}_{ij}$. Having set exchange ratios for all pairs of users that can trade bilaterally, we now consider users that cannot trade bilaterally. If $S_i \cap T_j = \emptyset$ and $S_j \cap T_i = \emptyset$, set $\gamma^*_{ij} = \gamma^*_{ji} = 1$. If $S_i \cap T_j \neq \emptyset$ and $S_j \cap T_i = \emptyset$, set

$$\gamma^*_{ij} = \varepsilon + \max_{k: S_k \cap T_i \neq \emptyset, S_i \cap T_k \neq \emptyset} \{\gamma^*_{kj}\},$$

and $\gamma^*_{ji} = 1/\gamma_{ij}$. If Assumption 2 holds, then $\boldsymbol{r}^*$ solves the bilateral optimization problem of every user with respect to exchange ratios $\boldsymbol{\gamma}^*$. In particular, user i can find every file in T_i through bilateral trade at the same exchange ratios as in the restricted BE. Exchange ratios with users that i cannot trade with bilaterally are set so that they do not affect i's optimization problem.

■

Proof of Proposition 2: If Assumption 3 holds, then either the user graph is strongly connected, or the system can be decomposed to subsystems for which the user graphs are strongly connected. Therefore, without loss of generality, in this proof we assume that the user graph is strongly connected.

For the purposes of this proof, let $\boldsymbol{r}^i(\boldsymbol{p})$ be the optimal solution of the multilateral optimization problem of user i when the price vector is $\boldsymbol{p} = (p_i, i \in U)$. If $\boldsymbol{r}$ and $\boldsymbol{p}$ constitute a ME, then $\boldsymbol{r} \in \boldsymbol{r}^i(\boldsymbol{p})$ for all $i \in U$.

We define excess demand for the upload rate of each user $i \in U$ as

$$z_i(\boldsymbol{p}) = \sum_{f,j} r^j_{ijf}(\boldsymbol{p}) - \sum_{f,j} r^i_{ijf}(\boldsymbol{p}).$$

We show that the aggregate excess demand correspondence $\boldsymbol{z}(\cdot)$ defined on $(0, \infty)^{|U|}$ satisfies the following properties:

1. For every $\boldsymbol{p} \gg 0$ and $\boldsymbol{z} \in \boldsymbol{z}(\boldsymbol{p})$, $\boldsymbol{p} \cdot \boldsymbol{z} = 0$.

2. $\boldsymbol{z}(\cdot)$ is convex-valued.

3. $\boldsymbol{z}(\cdot)$ is homogeneous of degree 0.

4. $\boldsymbol{z}(\cdot)$ is upper hemicontinuous.

5. There is an $s > 0$ such that $z_j > -s$ for any $\boldsymbol{z} \in \boldsymbol{z}(\boldsymbol{p})$, for every file $j \in F$ and every price vector $\boldsymbol{p} \gg 0$.

6. If $\boldsymbol{p}^m \to \boldsymbol{p} \neq \boldsymbol{0}$, $\boldsymbol{z}^m \in \boldsymbol{z}(\boldsymbol{p}^m)$ and $p_j = 0$ for some j, then $\max\{z_j^m : j \in F\} \to \infty$.

Then the existence of a ME follows from standard results in microeconomics; see, e.g., [15], Exercise 17.C.2.

By Assumption 1, the budget constraint of each user binds. The budget constraint of user i is

$$\sum_{j,f} p_j r^i_{jif}(\boldsymbol{p}) = p_i \sum_{j,f} r^i_{ijf}(\boldsymbol{p}).$$

By summing over all users, we obtain Property 1.

Fix a price vector $\boldsymbol{p} \gg 0$. By Assumption 1 $v(\cdot)$ is strictly concave; therefore $r^i_{ijf}(\boldsymbol{p})$ and $r^j_{ijf}(\boldsymbol{p})$ are convex-valued. Thus the aggregate excess demand $\boldsymbol{z}(\cdot)$ is a convex valued correspondence (Property 2).

Consider a price vector $\boldsymbol{p} \gg 0$, and fix a constant $t > 0$. It is clear that the feasible region of the multilateral user optimization problem remains unchanged if we replace the price vector $\boldsymbol{p}$ by $t\boldsymbol{p}$. Thus the aggregate excess demand is also homogeneous of degree zero (Property 3).

We now show that the aggregate excess demand correspondence is upper hemicontinuous. By Assumption 1 $v(\cdot)$ is a continuous function. From the Theorem of the Maximum [5] it follows that $r^i_{ijf}(\boldsymbol{p})$ and $r^i_{jif}(\boldsymbol{p})$ are upper hemicontinuous correspondences. The aggregate excess demand for user i is a linear combination of the rates $r^j_{ijf}(\boldsymbol{p})$ and $r^i_{ijf}(\boldsymbol{p})$, and therefore is also upper hemicontinuous (Property 4).

The upload rate of any user i is upper bounded by his upload rate constraint B_i, so the total supply is upper bounded and the excess demand is bounded from below (Property 5).

If $\boldsymbol{p}^m \to \boldsymbol{p} \neq 0$ and $p_j = 0$, then $p_k > 0$ for some k. Because of Assumption 3, there is a sequence of users $1, 2, ..., n \in U$ such that $T_i \cap S_{i+1} \neq \emptyset$. Thus, there is a user i such that p_i approaches a strictly positive limit and p_{i+1} approaches zero. Let $f \in T_i \cap S_{i+1}$. The budget of user i approaches a strictly positive limit as $\boldsymbol{p}^m \to \boldsymbol{p}$ and the amount of f he can afford goes to infinity. On the other hand, the total possible supply is bounded above by the upload rate constraints of user $i+1$. Thus $\max\{z_j^m : j \in F\} \to \infty$, establishing Property 6. ■

Proof of Theorem 1: Suppose that $\boldsymbol{r} \in \mathcal{X}$ is a Pareto improvement. Then some user i strictly prefers $\boldsymbol{r}$ to $\boldsymbol{r}^*$. Since $\boldsymbol{r}$ is not an optimal solution for user i under $\boldsymbol{p}$, it must be that

$$\sum_{j,f} p_j r_{jif} > p_i \sum_{j,f} r_{ijf}.$$

All users $k \neq i$ are at least as well off under $\boldsymbol{r}$ as under $\boldsymbol{r}^*$. This implies that

$$\sum_{j,f} p_j r_{jkf} \geq p_k \sum_{j,f} r_{kjf},$$

because the utilities are increasing in the total rates of files that users are interested in. In particular, consider a user k who gets exactly the same utility under $\boldsymbol{r}$ and $\boldsymbol{r}^*$: if $\sum_{j,f} p_k r_{jkf} < p_k \sum_{j,f} r_{kjf}$, then there is a rate allocation that satisfies k's budget constraint and k strictly prefers to $\boldsymbol{r}$, which would implies that $\boldsymbol{r}^*$ is not optimal.

Summing over all users,

$$\sum_k \sum_{j,f} p_j r_{jkf} > \sum_k p_k \sum_{j,f} r_{kjf},$$

which is a contradiction. We conclude that a ME allocation is Pareto efficient. ■

Proof of Proposition 3: Define $r^*_{ij} \equiv \sum_f r^*_{ijf}$ (the total rate that user i sends to user j). We define the matrix $\boldsymbol{Q}$ such that $Q_{ij} = r^*_{ij}$ if $i \neq j$; and $Q_{ii} = -\sum_j r^*_{ij}$. By construction, $\boldsymbol{Q}$ is a transition rate matrix with no transient subclasses, since $r^*_{ij} > 0$ implies that $r^*_{ji} > 0$ (by the definition of BE). In what follows we consider the communicating classes of $\boldsymbol{Q}$: if $r^*_{ij} > 0$, then users i and j are in the same communicating class. For the purposes of this proof, let $\mathcal{N}_i(\boldsymbol{r}^*)$ be the set of peers with which i trades under $\boldsymbol{r}^*$, i.e., $\mathcal{N}_i(\boldsymbol{r}^*) = \{j \in U : r^*_{ji} > 0\}$.

As noted above, we first observe that $\boldsymbol{\pi Q} = \boldsymbol{0}$ implies that the multilateral budget constraint is satisfied; therefore for any invariant distribution $\boldsymbol{\pi}$, $\boldsymbol{r}^*$ is feasible for the multilateral optimization problem of every user when prices are equal to $\boldsymbol{\pi}$. We show that for some invariant distribution of $\boldsymbol{Q}$ (say $\boldsymbol{p}$), $\boldsymbol{r}^*$ and $\boldsymbol{p}$ constitute a ME. In particular, we show that for each user i, $\boldsymbol{r}^*$ solves the multilateral optimization problem under $\boldsymbol{p}$.

This is done in three steps. First, we show that if $\boldsymbol{r}^*$ is Pareto efficient, then $\boldsymbol{Q}$ corresponds to a reversible Markov chain. This implies that if $\boldsymbol{\pi}$ is a strictly positive invariant distribution of $\boldsymbol{Q}$, then $\gamma^*_{ij} = \pi_i/\pi_j$ whenever $r^*_{ij} > 0$, and as a result $\boldsymbol{r}^*$ solves the multilateral optimization problem of user i given prices $\boldsymbol{\pi}$ if user i is restricted only with users in $\mathcal{N}_i(\boldsymbol{r}^*)$ (Step 1). We then show that if user i is restricted to trade with users in the same communicating class under prices $\boldsymbol{\pi}$, then $\boldsymbol{r}^*$ is an optimal solution of the multilateral optimization problem (Step 2). Step 2 completes the proof if $\boldsymbol{Q}$ consists of one communicating class. Finally, we show that if there are multiple communicating classes, there exists an invariant distribution $\boldsymbol{p}$ (where the invariant distribution of each communicating class is scaled appropriately) such that $\boldsymbol{r}^*$ is an optimal solution of the multilateral optimization problem of each user (Step 3).

We show each of these steps by demonstrating that otherwise there exists a rate vector $\boldsymbol{r}$ that Pareto improves $\boldsymbol{r}^*$. Suppose $\boldsymbol{r}^*$ solves the bilateral optimization problem of user i under $\boldsymbol{\gamma}^*$. Let $(x^*_{if}, f \in T_i)$ and y_i be the corresponding download and upload rates for user i. Consider a rate allocation $\boldsymbol{r}$ where $(x_{if}, f \in T_i)$ and y_i are the corresponding download and upload rates for user i. Assuming that $x_{if} - x^*_{if}$ and $y_i - y^*_i$ are sufficiently small, we can use Taylor's approximation to conclude that user i is strictly better off under $\boldsymbol{r}$ if

$$(x_{if} - x^*_{if})\frac{\partial v_i(\boldsymbol{x}^*_i, y^*_i)}{\partial x_{if}} > (y_i - y^*_i)\frac{\partial v_i(\boldsymbol{x}^*_i, y^*_i)}{\partial y_i}. \tag{1}$$

Suppose that $r^*_{jif} > 0$ for some j, which implies that $x^*_{if} > 0$. The optimality conditions for the bilateral optimization problem of user i give that

$$\frac{\partial v_i(x^*_{ig}, g \in T_i)}{\partial x_{if}} = \frac{1}{\gamma^*_{ij}}\frac{\partial v_i(\boldsymbol{x}^*_i, y^*_i)}{\partial y_i}.$$

Combining this with (1), we see that user strictly prefers $\boldsymbol{r}$ to $\boldsymbol{r}^*$ if

$$\frac{x_{if} - x^*_{if}}{y_i - y^*_i} > \gamma^*_{ij}. \tag{2}$$

Step 1: Let $\boldsymbol{\pi}$ be a strictly positive invariant distribution of $\boldsymbol{Q}$, i.e., $\boldsymbol{\pi} \gg 0$ and $\boldsymbol{\pi} \cdot \boldsymbol{Q} = \boldsymbol{0}$. Then, for every user i, $\sum_k (\pi_k/\pi_i) r^*_{ki} = \sum_k r^*_{ik}$. On the other hand, the budget constraint of the bilateral optimization problem of user i implies that $\gamma^*_{ki} r^*_{ki} = r^*_{ik}$. Summing over k and substituting, we conclude that

$$\sum_k \gamma^*_{ki} r^*_{ki} = \sum_k \frac{p_k}{p_i} r^*_{ki}. \tag{3}$$

If $\boldsymbol{Q}$ is reversible, then the detailed balance equations hold for every $i, j \in U$, i.e., $\pi_i r^*_{ij} = \pi_j r^*_{ji}$. We note that the detailed balance equations trivially hold if $r^*_{ij} = 0$, because then also $r^*_{ji} = 0$. We show that Pareto efficiency of $\boldsymbol{r}^*$ implies reversibility of $\boldsymbol{Q}$.

Assume that $\boldsymbol{Q}$ is not reversible. Then $\pi_i r^*_{ij} < \pi_j r^*_{ji}$ for some i, j with $r^*_{ij} > 0$. Moreover, since $\boldsymbol{\gamma}^*$ and $\boldsymbol{r}^*$ constitute a BE, we have $\gamma^*_{ji} = r^*_{ij}/r^*_{ji}$ whenever $r^*_{ji} > 0$. Thus, if $\boldsymbol{Q}$ is not reversible, then $\pi_j/\pi_i > \gamma^*_{ji}$ for some i, j with $r^*_{ij} > 0$. Without loss of generality, we relabel i to be $j+1$. Then, by (3), there exists some user k such that $\pi_{j+1}/\pi_k > \gamma^*_{j+1,k}$ and $r^*_{j+1,k} > 0$. We relabel k to be user $j+2$, and then $\pi_{j+1}/\pi_{j+2} > \gamma^*_{j+1,j+2}$. Applying this reasoning inductively, we can find a sequence of users $1, 2, ..., K, K+1$ such that $1 \equiv K+1$ and $\pi_k/\pi_{k+1} > \gamma^*_{k,k+1}$ for all k.

We show how the utility of each user in $S = \{1, 2, ..., K\}$ can increase while the rate allocation to users outside S remains the same. In particular, we increase $r^*_{k,k-1}$ and y^*_k by a_k for all $k \in S$, as illustrated in the first part of Figure 4 (for $K = 3$). We note that users' upload capacity constraints do not bind at the BE, a consequence of the assumption that $v_i(\boldsymbol{x}_i, y_i) \to \infty$ as $y_i \to B_i$. Therefore, it is feasible to slightly increase the upload rates of all users. Applying (2), user k is better off if

$$\frac{a_{k+1}}{a_k} > \gamma^*_{k,k+1}.$$

Since $\pi_k/\pi_{k+1} > \gamma^*_{k,k+1}$, it follows $\prod_k \gamma^*_{k,k+1} < 1$. Then, it is possible to make all users in the set better off by choosing δ and ε small enough, and setting $a_1 = \delta$; $a_{k+1} = \gamma_{k,k+1} \cdot a_k + \varepsilon$, for all $k \in S$.

We conclude that if $\boldsymbol{r}^*$ is the rate allocation of a BE and is Pareto efficient, then $\boldsymbol{Q}$ is reversible, and $\gamma^*_{ij} = \pi_i/\pi_j$ whenever $r^*_{ij} > 0$. This means that $\boldsymbol{r}^*$ solves the multilateral optimization problem of user i given prices $\boldsymbol{\pi}$ if he is restricted to trade with peers in $\mathcal{N}_i(\boldsymbol{r}^*)$. The remainder of the proof shows that for some invariant distribution $\boldsymbol{p}$, $\boldsymbol{r}^*$ is optimal for the multilateral optimization problem of every user under $\boldsymbol{p}$.

Step 2: Let $\boldsymbol{\pi}$ be a strictly positive invariant distribution of $\boldsymbol{Q}$, and consider the multilateral user optimization problems when prices are given by $\boldsymbol{\pi}$. We already showed in Step 1 that $\boldsymbol{r}^*$ is feasible. Suppose that $\boldsymbol{r}^*$ is not optimal for the multilateral optimization problem of some user i. Then by Step 1 there must exist a user j such that $r^*_{j,i} = 0$ with which i wants to exchange under $\boldsymbol{\pi}$.

In this step we consider the case that i and j are in the same communicating class. Then, there exists a sequence of users between i and j such that each two consecutive users trade at the BE. Without loss of generality we relabel user i by K, user j by 1, and the users in the sequence by $2, 3, ..., K-1$. Then, $r^*_{j,j-1} > 0$ for $j = 2, 3, ..., K$. We show that there is a Pareto improvement, where the utilities of all users in the set $S = \{1, 2, ..., K\}$ strictly increase, while utilities of users outside S remain the same.

Let a_j be the amount by which we increase rate $r_{j,j-1}$. We assume that all users in the set increase this rate by increasing their upload rates. In particular, user j increases his upload rate by a_j, and gets a_{j+1} more from user $j+1$. This is illustrated for $K = 3$ in the second part of Figure 4. Applying (2), user $j \neq K$ is better off if $a_{j+1}/a_j > \gamma^*_{j,j+1} \equiv \pi_j/\pi_{j+1}$ (the last part follows from the reversibility of $\boldsymbol{Q}$). To conclude this step, we show that user K is better off is $a_1/a_K > \pi_K/\pi_1$. Then, as in Step 1, it is possible to find a_i for $i \in S$, such that all users in S are better off.

Now consider user K. Let $f \in T_K$ be a file that user K wants to get from user 1 under prices $\boldsymbol{\pi}$. There are two cases to consider, depending on whether user K downloads file f at the BE.

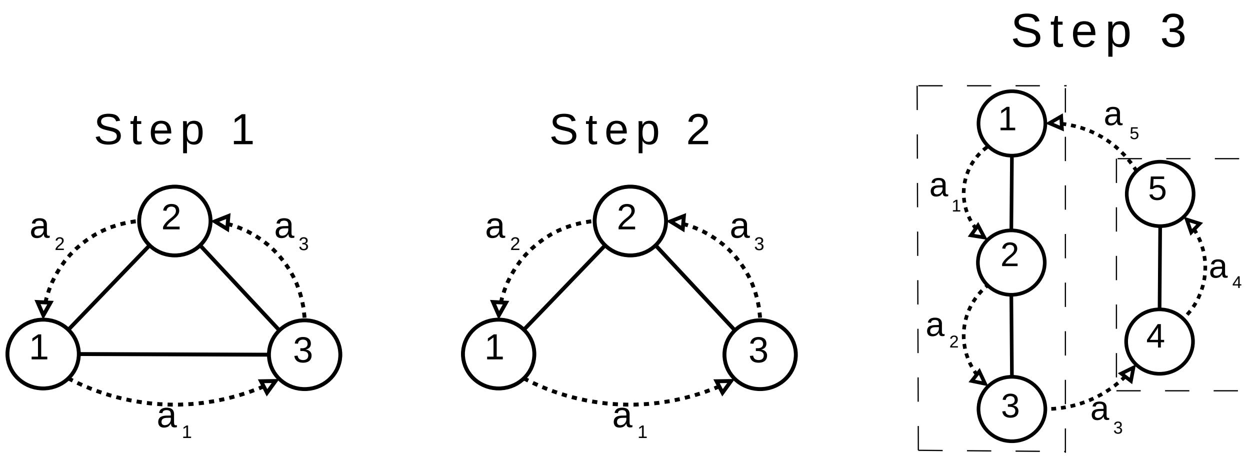

Figure 4: Pareto improvements when the BE allocation does not correspond to a ME allocation for Steps 1, 2, and 3 of the proof of Proposition 3 respectively. A pair of users that trade at the BE is connected with a solid line. Dotted arrows show the rates that increase for the Pareto improvement: user i increases his upload rate and the rate he sends to user $i-1$ by a_i. In the third figure (Step 3) there are two communicating classes — each class is included in a dashed box.

- $r^*_{jKf} > 0$ for some j. Then, by (2), we conclude that user K is better off if $a_1/a_K > \gamma_{Kj}$. Moreover, since K prefers to get f from 1 under $\boldsymbol{\pi}$ it must be that $\pi_K/\pi_1 > \gamma_{Kj}$. Thus, it suffices that $a_1/a_K > \pi_K/\pi_1$.

- $\sum_j r^*_{jKf} = 0$, i.e., K does not download file f at rate allocation $\boldsymbol{r}^*$. Under $\boldsymbol{\pi}$, user K is strictly better off downloading a positive amount of f from user 1. This implies that

$$\frac{\partial v_i(x^*_{ig}, g \in T_i)}{\partial x_{if}} > \frac{\pi_1}{\pi_K} \frac{\partial v_i(\boldsymbol{x}^*_i, y^*_i)}{\partial y_i}.$$

  Combining this with (1) we conclude that user K is better off if $a_1/a_K > \pi_K/\pi_1$.

In either case, user K is better off if $a_1/a_K > \pi_K/\pi_1$. This shows that at any optimal solution of the multilateral optimization problem for user i under $\boldsymbol{\pi}$, $r_{jif} = 0$ if $r^*_{jif} = 0$ and i, j are in the same communicating class.

Step 3: We now extend the result of Step 2 across communicating classes. Let $\boldsymbol{\pi}_c$ be the invariant distribution for communicating class c. We show that there exist coefficients ρ_c such that $\boldsymbol{r}^*$ is optimal for the multilateral optimization problem of each user under $\boldsymbol{p} \equiv \sum_c \rho_c \boldsymbol{\pi}_c$.

We start by deriving the conditions that the coefficients ρ_c need to satisfy. Consider two communicating classes c and c'. If $(\cup_{i\in c} T_i) \cap (\cup_{j \in c'} S_j) \neq \emptyset$, then some users from class c are interested in files that are possessed by users in class c'. To ensure that $\boldsymbol{r}^*$ is optimal for the multilateral optimization problems of these users, the ratio $\rho_{c'}/\rho_c$ should be sufficiently large. We denote this lower bound by $\xi_{c',c}$.

Suppose that there do not exist coefficients ρ_c such that $\boldsymbol{r}^*$ is an optimal solution of the multilateral optimization problem of each peer. Then, there exists a directed cycle of classes such that (1) $(\cup_{i\in c} T_i) \cap (\cup_{j \in c'} S_j) \neq \emptyset$ for each two consecutive classes in the cycle, and (2) the product of $\xi_{c',c}$ along the cycle is strictly greater than 1. This implies the existence of a vector $\boldsymbol{\rho}$ such that

$\rho_{c'}/\rho_c < \xi_{c',c}$ for every pair of consecutive classes along the directed cycle. In particular, when prices are $\boldsymbol{p} \equiv \sum_c \rho_c \boldsymbol{\pi}_c$, for each pair of consecutive classes along the cycle c and c', there is a user n_c in class c that wants to trade with user $m_{c'}$ from class c'. We construct a set S that includes users n_c, m_c as well as the users between them, i.e., users $i_{c1}, ..., i_{cl}$ such that $n_c \equiv i_{c1}$, $m_c \equiv i_{cl}$ and $r^*_{i_{cj}, i_{c,j+1}} > 0$. We relabel users in S by $\{1, 2, ..., K\}$ such that if i and $i+1$ are in different communicating classes (say c and c') then $i = n_c$ and $i + 1 = m_{c'}$, i.e., user i wants to trade with user $i + 1$.

We demonstrate a Pareto improvement where user $i \in S$ increases his upload rate and the rate he sends to user $i - 1$ by a_i. In Figure 4 we illustrate an example with two communicating classes. We demonstrate that it is possible to reallocate rates in a way that strictly increases the utilities of all users in S and does not change the utilities of users outside S. From (2) we see that a user $j \neq n_c$ can be made better off if $a_{j+1}/a_j > p_j/p_{j+1}$. A user $j \equiv n_c$ for some i can be made better off if $a_{j+1}/a_j > p_j/p_{j+1}$ (this can be shown by applying the same argument we used for user K in Step 2). As in Steps 1 and 2, since the product of all left hand sides is equal to 1 while the product of all right hand sides is strictly less than 1, it is possible to find a vector $\boldsymbol{a}$ that satisfies all these inequalities. ■

Proof of Proposition 4: If $s = 0$, files are chosen uniformly. If there are K files in the system, the probability that a given user has files $\{f_1, ..., f_\sigma\}$ and wants file g is equal to $1/(\binom{K}{\sigma}(K - \sigma))$ for any set of distinct files $\{f_1, ..., f_\sigma, g\}$. A given user i can trade bilaterally with user j with probability

$$\frac{\sigma \binom{K-2}{\sigma-1}}{\binom{K}{\sigma}(K - \sigma)} = \frac{\sigma^2}{K(K-1)}$$

Thus a user cannot trade bilaterally with probability

$$\overline{\rho}_{BE}(K, N, 0) = \left(1 - \frac{\sigma^2}{K(K-1)}\right)^{N-1}.$$

We observe that $\overline{\rho}_{BE}(K, N, 0)$ is increasing in K. If $K(N) = \sigma\sqrt{N}$, then $\overline{\rho}_{BE}(K(N), N, 0) \to 1/e$ as $N \to \infty$. Thus, if $K(N) \geq \sigma\sqrt{N}$, then $\overline{\rho}_{BE}(K(N), N, 0) \geq 1/e$ for large N. On the other hand, if there exists $\varepsilon > 0$ such that $K(N) \leq \sigma N^{1/2-\varepsilon}$ for large N, then $\overline{\rho}_{BE}(K(N), N, 0) \to 0$ as $N \to \infty$.

We now consider multilateral exchange. We do not have a closed form formula to compute $\overline{\rho}_{ME}(K, N, 0)$. Instead, we reduce this to a random graph problem, and we use the results from [9], which studies the size of a strongly connected component of a random digraph. We consider the user graph that was defined in Section 3. Recall that this is the directed graph $G = (V, E)$ with $V = U$, and $E = \{(i, j) : S_i \cap T_j \neq \emptyset\}$. If this graph is strongly connected, then all users participate in the multilateral exchange. When users choose file uniformly, there is a directed edge from user i to user j with probability σ/K; this is the probability that user i has the file user j wants. On the other hand, there are N nodes in this graph. Applying the results of [9], the size of the strongly connected component is $\approx K$ if and only if $\equiv (1 - \sigma/K)^N \to 0$. This is the case if $N^{1-\varepsilon} \geq \sigma K$ for some small $\varepsilon > 0$. ■

Proof of Proposition 5:

The expected percentage of users that cannot trade bilaterally when there are K files and N

users is

$$\overline{\rho}_{BE}(K,N,s) = \sum_{i\neq j: i,j\in\{1,\ldots,K(N)\}} \frac{(ij)^{-s}}{\sum_{i\neq j: i,j\in\{1,\ldots,K(N)\}}(ij)^{-s}} \left(1 - \frac{(ij)^{-s}}{\sum_{i\neq j: i,j\in\{1,\ldots,K(N)\}}(ij)^{-s}}\right)^{N-1}.$$

Let $A_N \equiv \overline{\rho}_{BE}(K(N),N,s)$. We are interested in the limit of A_N as $N \to \infty$.

We observe that

$$1 - \frac{(ij)^{-s}}{\sum_{i\neq j: i,j\in\{1,\ldots,K(N)\}}(ij)^{-s}} < 1,$$

and thus each term in the sum approaches 0 as $N \to \infty$.

We first assume that $K(N) \not\to \infty$ as $N \to \infty$. Then A_N is the sum of a finite number of terms, each of which $\to 0$ as $N \to \infty$. Thus, $A_N \to 0$ as $N \to \infty$.

Now assume that $K(N) \to \infty$ as $N \to \infty$ and let

$$\sigma(s) \equiv \sum_{i\neq j: i,j\in\{1,2,\ldots\}} (ij)^{-s}.$$

Since $s > 1$, $\sigma(s)$ is finite.

$$A_N \le \frac{1}{\sum_{i\neq j: i,j\in\{1,\ldots,K(N)\}}(ij)^{-s}} \sum_{i\neq j: i,j\in\{1,\ldots,K(N)\}} (ij)^{-s}\left(1 - \frac{(ij)^{-s}}{\sigma(s)}\right)^{N-1}$$

Since

$$\frac{1}{\sum_{i\neq j: i,j\in\{1,\ldots,K(N)\}}(ij)^{-s}} \to \frac{1}{\sigma(s)} < \infty,$$

it suffices to show that

$$B_N \equiv \sum_{i\neq j: i,j\in\{1,\ldots,K(N)\}} (ij)^{-s}\left(1 - \frac{(ij)^{-s}}{\sigma(s)}\right)^{N-1} \to 0 \text{ as } N \to \infty.$$

In particular, it suffices to show that for every $\varepsilon > 0$ there exists $\overline{N}$ such that $B_N \le \varepsilon$ for all $N < \overline{N}$. We observe that for any N_1,

$$B_N < \sum_{i\neq j: i\cdot j\le N_1} \left(1 - \frac{(ij)^{-s}}{\sigma(s)}\right)^{N-1} + \sum_{i\neq j: i\cdot j > N_1} (ij)^{-s}.$$

For $\varepsilon > 0$, we choose $N_1(\varepsilon)$ such that

$$\sum_{i\neq j: i\cdot j > N_1} (ij)^{-s} < \varepsilon/2.$$

We choose $\overline{N}$ such that

$$|\{(i,j) : i\cdot j \le N_1\}| \cdot \left(1 - \frac{N_1^{-s}}{\sigma(s)}\right)^{\overline{N}-1} < \varepsilon/2.$$

Then $B_N < \varepsilon$ for all $N \ge \overline{N}$. ∎