An Axiomatic Theory of Fairness in Resource Allocation

Tian Lan[1] and Mung Chiang[2]
[1] Department of Electrical and Computer Engineering, George Washington University
[2] Department of Electrical Engineering, Princeton University

***Abstract*—We develop a set of five axioms for fairness measures in resource allocation: the axiom of continuity, of homogeneity, of saturation, of partition, and of starvation. We prove that there is a unique family of fairness measures satisfying the axioms, which is constructed and then shown to include Atkinson's index, α-fairness, Jain's index, entropy, and other "decomposable" global measures of fairness as special cases. We prove properties of fairness measures satisfying the axioms, including symmetry and Schur-concavity. Among the engineering implications is a generalized Jain's index that tunes the resolution of fairness measure, a decomposition of α-fair utility functions into fairness and efficiency components, and an interpretation of "larger α is more fair" and of Rawl's difference principle.**

The axiomatic theory is further extended in three directions. First, we quantify fairness of continuous-dimension inputs, where resource allocations vary over time or domain. Second, by starting with both a vector of resource allocation and a vector of user-specific weights, and modifying the axiom of partition, we derive a new family of fairness measures that are asymmetric among users. Finally, we develop a set of four axioms by removing the axiom of homogeneity to capture a fairness-efficiency tradeoff.

We present illustrative examples in congestion control, routing, power control, and spectrum management problems in communication networks, with the potential of a fairness evaluation tool explored. We also compare with other work of axiomatization in information, computer science, economics, sociology, and political philosophy.

I. Introduction

Given a vector $\mathbf{x} \in \mathbb{R}^n_+$, where x_i is the resource allocated to user i, *how fair* is it? Consider two allocations among three users: $\mathbf{x} = [1,\ 2,\ 3]$ and $\mathbf{y} = [1,\ 10,\ 100]$. Among the large variety of choices we have in quantifying fairness, we can get fairness values such as 0.33 or 0.86 for $\mathbf{x}$ and 0.01 or 0.41 for $\mathbf{y}$: $\mathbf{x}$ is viewed as 33 times more fair than $\mathbf{y}$, or just twice as fair as y. How many such "viewpoints" are there? What would disqualify a quantitative metric of fairness? Can they all be constructed from a set of simple statements taken as true for the sake of subsequent inference?

One existing approach to quantify fairness of $\mathbf{x}$ is through a fairness measure, which is a function f that maps $\mathbf{x}$ into a real number. These measures are sometimes referred to as diversity indices in statistics. Various fairness measures have been proposed throughout the years, e.g., in [1]–[6] in the field of networking and many more in other fields. These range from simple ones, e.g., the ratio between the smallest and the largest entries of $\mathbf{x}$, to more sophisticated functions, e.g., Jain's index and the entropy function. Some of these fairness measures map $\mathbf{x}$ to a normalized range between 0 and 1, where 0 denotes the minimum fairness, 1 denotes the maximum fairness (often corresponding to an $\mathbf{x}$ where all x_i are the same), and a larger value indicates more fairness. For example, min-max ratio [1] is given by the maximum ratio of any two user's resource allocation, while Jain's index [3] computes a normalized square mean. How are these fairness measures related? Is one measure "better" than any other? What other measures of fairness might be useful?

An alternative approach that has gained attention in the networking research community since [7], [8] is the optimization-theoretic approach of α-fairness and the associated utility maximization problem. Given a set of feasible allocations, a maximizer of the α-fair utility function satisfies the definition of α-fairness. Two well-known examples are as follows: a maximizer of the log utility function ($\alpha = 1$) is proportionally fair, and a maximizer of the α-fair utility function as $\alpha \to \infty$ is max-min fair. More recently, α-fair utility functions have been connected to divergence measures [9]. In [10], [11], the parameter α was viewed as a fairness measure in the sense that a fairer allocation is one that is the maximizer of an α-fair utility function with larger α – although the exact role of α in trading-off fairness and throughput can sometimes be surprising [12]. While it is often believed that $\alpha \to \infty$ is more fair than $\alpha = 1$, which is in turn more fair than $\alpha = 0$, it remains unclear what it means to say, for example, that $\alpha = 3$ is more fair than $\alpha = 2$.

Clearly, these two approaches for quantifying fairness are different. One difference is the treatment of efficiency, or magnitude, of resources. On the one hand, α-fair utility functions are continuous and strictly increasing in each entry of $\mathbf{x}$, thus its maximization results in Pareto optimal resource allocations. On the other hand, scale-invariant fairness measures (ones that map $\mathbf{x}$ to the same value as a normalized $\mathbf{x}$) are unaffected by the magnitude of $\mathbf{x}$, and $[1\ 1]$ is as fair as $[100\ 100]$. Can the two approaches be unified?

To address the above questions, we develop an axiomatic approach to fairness measures. We show that a set of five axioms, each of which simple and intuitive, thus accepted as true for the sake of subsequent inference, can lead to a useful family of fairness measures. The axioms are: the axiom of continuity, of homogeneity, of saturation, of partition, and of starvation. Starting with these five axioms, we can *generate* a family of fairness measures from generator functions g: any increasing and continuous functions that lead to a well-

This work has been supported in part by NSF CNS-0905086 and CNS-0519880.

defined "mean" function (i.e., from any Kolmogorov-Nagumo function [13], [14]). For example, using power functions with exponent β as the generator function, we derive a unique family of fairness measures f_β that include all of the following as special cases, depending on the choice of β: Jain's index, maximum or minimum ratio, entropy, and α-fair utility, and reveals new fairness measures corresponding to other ranges of β. While we start with the approach of fairness measure rather than optimization objective function, it turns out that the latter approach can also be recovered from f_β.

We also prove the uniqueness of fairness measures. First, we show that for each generator function g, its corresponding fairness measure f satisfying the five axioms is unique. In a much stronger result, generator g must belong to the following class of functions: logarithm functions $g(y) = \log y$ and power generator functions $g(y) = y^\beta$. This result is sufficient and necessary, in the sense that no other generator function is possible, and for each of the generator function within this class, we can construct a unique fairness measure f_β satisfying the five axioms.

In particular, for $\beta \leq 0$, several well-known fairness measures (e.g., Jain's index and entropy) are special cases of our construction, and we generalize Jain's index to provide a flexible tradeoff between "resolution" and "strictness" of the fairness measure. For $\beta \geq 0$, α-fair utility functions can be factorized as the product of two components: our fairness measure with $\beta = \alpha$ and a function of the total throughput that captures the scale, or efficiency, of $\mathbf{x}$. Such a factorization quantifies a tradeoff between fairness and efficiency: what is the maximum weight that can be given to fairness while still maintaining Pareto efficiency. It also facilitates an unambiguous understanding of what it means to say that a larger α is "more fair" for general $\alpha \in [0, \infty)$.

Any fairness measure satisfying the five axioms can be proven to have many properties quantifying common beliefs about fairness, including Schur-concavity [15]. Our fairness measures presents a new ordering of Lorenz curves [6], [16], different from the Gini coefficient.

We extend this axiomatic theory in three directions. First, utilizing fairness measure f that satisfies the five axioms, we derive a new class of fairness measures to quantify the degree of fairness associated with a time varying resource allocation. In Section III, we tackle the following question: Given a continuous-dimension resource allocation over time $[0, T]$, represented by $\{\mathbf{X}_t \in \mathbb{R}, \forall t \in [0, T]\}$. How do we quantify its degree of fairness using a function $\mathcal{F}$? Our result shows that fairness of continuous-dimension resource allocation can be characterized by the limit of fairness measures for n-dimensional resource allocations. Fairness $\mathcal{F}$ can be equivalently computed from resource allocation over time, cumulative distribution, or the Lorenz curve. We prove a set of interesting properties and implications, including a generalized notion of Schur-concavity.

Given a set of useful axioms, it is important to ask if other useful axiomatic systems are possible. By removing or modifying some of the axioms, what kind of fairness measures and properties would result? One limitation of the axiom of partition is that the resulting fairness measures are symmetric: labeling of users can be permuted without changing fairness value. In practice, fairness measures should incorporate asymmetric function. For example, user i has a weight w_i, which may depend on her needs or roles in the size of feasible region of allocations. In Section IV we propose an alternative set of five axioms and replace the Axiom of Partition by a new version. Starting with this new set of axioms, we can generate asymmetric fairness measures, which incorporates all special cases for symmetric fairness measure, as well as weighted p-norms and weighted geometric mean. Many properties previously proven for symmetric fairness measures are revised by the user-specific weights. In particular, equal allocation is no longer fairness-maximizing.

This family of asymmetric fairness measures lead to our discussion of a fairness evaluation tool, which can be used to perform a consistency check of fairness evaluation, model user-specific weights for different applications, and tune the tradeoff between fairness and efficiency in design objectives. In Section IV.D, we present a human-subject experiment for f_β reverse engineering and fairness evaluation's consistency check.

The development of fairness measures takes another turn towards the end of the paper in Section V. By removing the Axiom of Homogeneity, we propose another set of four axioms, which allows *efficiency* of resource allocation to be *jointly captured* in the fairness measure. We show how this alternative system connects with optimization-based resource allocation, where magnitude matters due to the feasibility constraint and an objective function that favors efficiency. The fairness measures F constructed out of the reduced set of axioms is a generalization of that out of the previous two sets of axioms. In particular, the entire α-fair utility function, rather than just its fairness component, is a special case of F now.

Starting from axioms, this paper focuses on developing a modeling "language" in quantifying fairness. Fairness measure thus constructed are illustrated in Appendix A using numerical examples, from typical networking functionalities: congestion control, routing, power control, and spectrum management, in Appendix A.

Axiomatization of key notions in bargaining, collaboration, and sharing are studied in information, science, computer science, economics, sociology, psychology, and political philosophy over the past 60 years. In Appendix B, we compare and contrast with several of them, including the well-known theories on Renyi Entropy, Lorenz Curve, Nash Bargaining Solutions, Shapley Value, Fair Division, as well as the two principles in Rawls' Theory of Justice as Fairness.

Some open questions are outlined in that appendix too. All proofs are collected in Appendix C.

II. The First Axiomatic Theory

A. A Set of Five Axioms

Let $\mathbf{x}$ be a resource allocation vector with n non-negative elements. A fairness measure is a sequence of mapping $\{f^n(\mathbf{x}), \forall n \in \mathbb{Z}_n\}$ from n-dimensional vectors $\mathbf{x}$ to real numbers, called fairness values, i.e., $\{f^n : \mathbb{R}^n_+ \to \mathbb{R}, \forall n \in \mathbb{Z}_+\}$.

Symbol	Meaning
$\mathbf{x}$	Resource allocation vector of length n
$w(\mathbf{x})$	Sum of all elements of $\mathbf{x}$
$f(\cdot), f_\beta(\cdot)$	Fairness measure (of parameter β)
β	Parameter for power function $g(y) = y^\beta$
$g(\cdot)$	Generator function
s_i	Positive weights for weighted mean
$\mathbf{1}_n$	Vector of all ones of length n
$\mathbf{x}^\uparrow$	Sorted vector with smallest element being first
$\mathbf{x} \succeq \mathbf{y}$	Vector $\mathbf{x}$ majorizes vector $\mathbf{y}$
$U_\alpha(\cdot)$	α-fair utility with parameter α
$H(\cdot)$	Shannon entropy function
$J(\cdot)$	Jain's index
$\Phi_\lambda(\cdot)$	Our multicriteria (fairness and efficiency) utility function
$\mathbf{X}_{t\in[0,T]}$	Resource allocation over time $[0,T]$
$\mathcal{F}(\cdot)$	Fairness measure for continuous-dimensional inputs
$P_{\mathbf{x}}(\cdot)$	pdf of resource allocation over $[0,T]$
$L_{\mathbf{x}}(\cdot)$	Lorenz curve of resource allocation over $[0,T]$
q_j	Positive constant weight of user j
$\mathcal{I}_i$	Set of user indices in sub-system i
$F(\cdot), F_{\beta,\lambda}(\cdot)$	Fairness-efficiency measure

TABLE I
MAIN NOTATION.

To simplify notations, we suppress n in $\{f^n\}$ and denote them simply as f. We introduce the following set of axioms about f, whose explanations are provided after the statement of each axiom.

1) *Axiom of Continuity.* Fairness measure $f(\mathbf{x})$ is continuous on $\mathbb{R}^n_+$, for all $n \in \mathbb{Z}_+$.

Axioms 1 is intuitive: A slight change in resource allocation shows up as a slight change in the fairness measure.

2) *Axiom of Homogeneity.* Fairness measure $f(\mathbf{x})$ is a homogeneous function of degree 0:

$$f(\mathbf{x}) = f(t \cdot \mathbf{x}), \quad \forall\ t > 0. \tag{1}$$

Without loss of generality, for $n = 1$, we take $|f(x_1)| = 1$ for all $x_1 > 0$, i.e., fairness is a constant for a one-user system.

Axiom 2 says that the fairness measure is independent of the unit of measurement or magnitude of the resource allocation. For an optimization formulation of resource allocation, the fairness measure $f(\mathbf{x})$ alone cannot be used as the objective function if efficiency (which depends on magnitude $\sum_i x_i$) is to be captured. In Section II-E, we will connect this fairness measure with an efficiency measure in α-fair utility function. In Section V, we will remove Axiom 2 and propose an alternative set of axioms, which make measure $f(\mathbf{x})$ dependent on both the magnitude and distribution of $\mathbf{x}$, thus capturing fairness and efficiency at the same time.

3) *Axiom of Saturation.* Equal allocation's fairness value is independent of number of users as the number of users becomes large, i.e.,

$$\lim_{n\to\infty} \frac{f(\mathbf{1}_{n+1})}{f(\mathbf{1}_n)} = 1. \tag{2}$$

This axiom is a technical condition used to help ensure *uniqueness* of the fairness measure. Note that it is *not* stating that equal allocation is most fair.

A primary motivation for quantifying fairness is to allow a comparison of fairness values. Therefore, we must ensure well-definedness of ratio of fairness measures, as the number of users in the system increases. Axiom 4 states that fairness comparison is independent of the way the resource allocation vector is reached as the system grows.

4) *Axiom of Partition.* Consider a partition of a system into two sub-systems. Let $\mathbf{x} = [\mathbf{x}^1, \mathbf{x}^2]$ and $\mathbf{y} = [\mathbf{y}^1, \mathbf{y}^2]$ be two resource allocation vectors, each partitioned and satisfying $\sum_j x^i_j = \sum_j y^i_j$ for $i = 1, 2$. There exists a *mean function* h such that their fairness ratio is the mean of the fairness ratios of the subsystems' allocations, for all partitions such that sum resources of each subsystem are the same across $\mathbf{x}$ and $\mathbf{y}$:

$$\frac{f(\mathbf{x})}{f(\mathbf{y})} = h\left(\frac{f(\mathbf{x}^1)}{f(\mathbf{y}^1)}, \frac{f(\mathbf{x}^2)}{f(\mathbf{y}^2)}\right) \tag{3}$$

where according to the axiomatic theory of mean function [13], a function h is a mean function if and only if it can be expressed as follows:

$$h = g^{-1}\left(\sum_{i=1}^{2} s_i \cdot g\left(\frac{f(\mathbf{x}^i)}{f(\mathbf{y}^i)}\right)\right), \tag{4}$$

where g is any continuous and strictly monotonic function, referred to as the Kolmogorov-Nagumo function, and s_i are positive weights such that $\sum_i s_i = 1$.

5) *Axiom of Starvation* For $n = 2$ users, we have $f(1, 0) \leq f(\frac{1}{2}, \frac{1}{2})$, i.e., starvation is no more fair than equal allocation.

Axiom 5 is the only axiom that involves a *value* statement on fairness: starvation is no more fair than equal distribution for two users. This value statement also holds for all existing fairness measures, e.g., various, spread, deviation, max-min ratio, Jain's index, α-fair utility, and entropy. We will later see in the proofs that Axiom 5 specifies an increasing direction of fairness and is used to ensure uniqueness of $f(\mathbf{x})$. We could have picked a different axiom to achieve similar effects, but the above axiom for just 2 users and involving only starvation and equal allocation is the weakest statement, thus the "strongest axiom".

With the five axioms presented, we first discuss some implications of Axiom 4. This axiom can construct fairness measure f from lower dimensional spaces. If we choose $\mathbf{y} = [w(\mathbf{x}^1),\ w(\mathbf{x}^2)]$ (where $w(\mathbf{x}^i) = \sum_j x^i_j$ is the sum of resource in sub-system i) in Axiom 4 and use the fact that $|f(w(\mathbf{x}^i))| = 1$ for scalar inputs as implied by Axiom 2, we show that Axiom 4 implies a hierarchical construction of fairness, which allows us to derive a fairness measure $f : \mathbb{R}_+^n \to \mathbb{R}$ of n users recursively (with respect to a generator function $g(y)$) from lower dimensions, $f : \mathbb{R}_+^k \to \mathbb{R}$ and $f : \mathbb{R}_+^{n-k} \to \mathbb{R}$ for integer $0 < k < n$.

Lemma 1: (Hierarchical construction of fairness.) Consider an arbitrary partition of vector $\mathbf{x}$ into two segments $\mathbf{x} = \left[\mathbf{x}^1, \mathbf{x}^2\right]$. For a fairness measure f satisfying Axiom 4, we have

$$f(\mathbf{x}) = f\left(w(\mathbf{x}^1), w(\mathbf{x}^2)\right) \cdot g^{-1}\left(\sum_{i=1}^{2} s_i \cdot g\left(\left|f(\mathbf{x}^i)\right|\right)\right). \quad (5)$$

As a special case of Lemma 1, if we denote the resource allocation at level 1 by a vector $\mathbf{z} = [w(\mathbf{x}^1), w(\mathbf{x}^2)]$ and if the resource allocation at level 2 are the same $\mathbf{x}^1 = \mathbf{x}^2 = \mathbf{y}$, it is straight forward to verify that Axioms 2 and 4 imply

$$\begin{aligned} f(\mathbf{y} \otimes \mathbf{z}) &= f(\mathbf{y} \cdot w(\mathbf{z})) \cdot g^{-1}\left(\sum_i s_i \cdot g\left(|f(\mathbf{z})|\right)\right) \\ &= f(\mathbf{y}) \cdot |f(\mathbf{z})|, \end{aligned} \quad (6)$$

where $\otimes$ is the direct product of two vectors. Intuitively, Equation (6) means that fairness of a direct product of two resource allocations equals to the product of their fairness values.

Initial:	$\sum_{i=1}^{n} x_i$	
	$\swarrow$	$\searrow$
Level 1:	$\sum_{i=1}^{k} x_i$	$\sum_{i=k+1}^{n} x_i$
	$\swarrow \quad \searrow$	$\swarrow \quad \searrow$
Level 2:	$x_1, \ldots, x_k$	$x_{k+1}, \ldots, x_n$

TABLE II
ILLUSTRATION OF THE HIERARCHICAL COMPUTATION OF FAIRNESS.

The recursive computation is illustrated by a two-level representation in Table II. Let $\mathbf{x}^1 = [x_1, \ldots, x_k]$ and $\mathbf{x}^2 = [x_{k+1}, \ldots, x_n]$. The computation is performed as follows. At level 1, since the total resource is divided into two "sub-systems", fairness across the sub-systems obtained in this level is measured by $f\left(w(\mathbf{x}^1), w(\mathbf{x}^2)\right)$. At level 2, the two sub-systems of resources are further allocated to k and $n-k$ users, achieving fairness $f(\mathbf{x}^1)$ and $f(\mathbf{x}^2)$, respectively. To compute overall fairness of the resource allocation $\mathbf{x} = [x_1, x_2, \ldots, x_n]$, we combine the fairness obtained in the two levels using (5). Axiom 4 states that the construction in (5) must be independent of the partition and generates the same fairness measure for any choice of $0 < k < n$.

The functions g giving rise to the same fairness measures f may not be unique, e.g., logarithm and power functions. The simplest case is when g is identity and $s_i = 1/n$ for all i. A natural parameterization of the weight s_i in (3) is to choose the value proportional to the sum resource of sub-systems:

$$s_i = \frac{w^\rho(\mathbf{x}^i)}{\sum_j w^\rho(\mathbf{x}^j)}, \ \forall i \quad (7)$$

where $\rho \geq 0$ is an arbitrary exponent. As shown in Section II-C, the parameter ρ can be chosen such that the hierarchical computation is independent of partition as mandated in Axiom 4.

By definition, axioms are true, as long as they are consistent. They should also be non-redundant. However, not all sets of axioms are useful: unifying known notions, discovering new measures and properties, and providing useful insights. We first demonstrate the following existence (the axioms are consistent) results.

Theorem 1: (*Existence.*) There exists a fairness measure $f(\mathbf{x})$ satisfying Axioms 1–5.

Uniqueness results contain two parts. First, we show that, from any Kolmogorov-Nagumo function $g(y)$, there is a unique $f(\mathbf{x})$ thus generated. Such an $f(\mathbf{x})$ is a well-defined fairness measure if it also satisfies Axioms 1–5. We then refer to the corresponding $g(y)$ as the *generator* of the fairness measure.

Definition 1: Function $g(y)$ is a generator if there exists a $f(\mathbf{x})$ satisfying Axioms 1–5 with respect to $g(y)$.
Second, we further show that only *logarithm* and *power* functions are possible generator functions. Therefore, we find, in closed-form, all possible fairness measures satisfying axioms 1-5. The uniqueness result is summarized as the following Theorem.

Theorem 2: (*Uniqueness.*) The family of $f(\mathbf{x})$ satisfying Axioms 1–5 is uniquely determined by logarithm and power generator functions,

$$g(y) = \log y \text{ and } g(y) = y^\beta \quad (8)$$

where $\beta \in \mathbb{R}$ is a constant.

B. Constructing Fairness Measures

Let $\text{sign}(\cdot)$ be the sign function. We derive the fairness measures in closed-form in the following theorem.

Theorem 3: (*Constructing all fairness measures satisfying Axioms 1-5.*) If $f(\mathbf{x})$ is a fairness measure satisfying Axioms 1-5 with weights in (7), $f(\mathbf{x})$ is of the form

$$f(\mathbf{x}) = \text{sign}(r) \prod_{i=1}^{n} \left(\frac{x_i}{\sum_{j=1}^n x_j}\right)^{-r\left(\frac{x_i}{\sum_{j=1}^n x_j}\right)} \quad (9)$$

$$f(\mathbf{x}) = \text{sign}(r(1-\beta r)) \cdot \left[\sum_{i=1}^{n} \left(\frac{x_i}{\sum_j x_j}\right)^{1-\beta r}\right]^{\frac{1}{\beta}} \quad (10)$$

where $r \in \mathbb{R}$ is a constant exponent. In (10), the two parameters in the mean function h have to satisfy $r\beta + \rho = 1$, where ρ is in the weights of the weighted sum in (7)and β is in the definition of the power generator function. It determines the growth rate of maximum fairness as population size n increases, as will be shown in Corollary 3:

$$f(\mathbf{1}_n) = n^r \cdot f(1). \quad (11)$$

The fairness measure in (9) can be obtained as a special case of (10) as $\beta \to 0$. Obviously, $\text{sign}(r(1-r\beta)) \to \text{sign}(r)$ as $\beta \to 0$. By removing the sign of (10) and taking a logarithm transformation, we have

$$\begin{aligned}
&\lim_{\beta\to 0} \log \left[\sum_{i=1}^{n} \left(\frac{x_i}{\sum_j x_j} \right)^{1-\beta r} \right]^{\frac{1}{\beta}} \\
&= \lim_{\beta\to 0} \frac{\log \left[\sum_{i=1}^{n} \left(\frac{x_i}{\sum_j x_j} \right)^{1-\beta r} \right]}{\beta} \\
&= \lim_{\beta\to 0} \sum_{i=1}^{n} -r \log \left(\frac{x_i}{\sum_{j=1}^{n} x_j} \right) \cdot \left(\frac{x_i}{\sum_{j=1}^{n} x_j} \right)^{1-r\beta} \\
&= \log \prod_{i=1}^{n} \left(\frac{x_i}{\sum_{j=1}^{n} x_j} \right)^{-r \frac{x_i}{\sum_{j=1}^{n} x_j}} \qquad (12)
\end{aligned}$$

where the second step uses L'Hospital's Rule. Therefore, we derive (9) by taking an exponential transformation in the last step of (12). This shows that fairness measure in (9) is a special case of (10) as $\beta \to 0$. Another way to see this is as follows: logarithm generator can be derived from the limit of the following power generators as $\beta \to 0$:

$$\log y = \lim_{\beta\to 0} \frac{y^{\beta} - 1}{\beta}. \qquad (13)$$

As a result of uniqueness in Theorem 2, their corresponding fairness measures in (9) and (10) also become equivalent as $\beta \to 0$. From here on, we consider the family of fairness measures in (10), which are parameterized by β in the power generator function.

For different values of parameter β, the fairness measures derived above are equivalent up to a constant exponent r:

$$f_{\beta,r}(\mathbf{x}) = \left[f_{\beta r,1}\right]^{r}(\mathbf{x}), \qquad (14)$$

if we denote $f_{\beta,r}$ to be the fairness measure with parameters β and r.

According to Theorem 3, r determines the growth rate of maximum fairness as population size n increases. In most of the following developments, we choose $r = 1$, so that maximum fairness per user is $f(\mathbf{1}_n)/n = n^{r-1} = 1$, independent of population size.

We now present a unified representation of the constructed fairness measures:

$$f_{\beta}(\mathbf{x}) = \text{sign}(1-\beta) \cdot \left[\sum_{i=1}^{n} \left(\frac{x_i}{\sum_j x_j} \right)^{1-\beta} \right]^{\frac{1}{\beta}}. \qquad (15)$$

In the rest of this section, we will show that this unified representation leads to many implications.

We summarize the special cases in Table III, where β sweeps from $-\infty$ to ∞ and $H(\cdot)$ denotes the entropy function. For some values of β, the corresponding mean function h has a standard name, and for some values of β, known approaches to measure fairness are recovered, while for $\beta \in (0, -1)$ and $\beta \in (-1, -\infty)$, new fairness measures are revealed. For example, when $\beta = -1$ (i.e., harmonic mean is used in Axiom 4), we recover Jain's index $J(\mathbf{x}) = f_{\beta}(\mathbf{x})/n$. For $\beta \geq 0$, we recover the component of the α-fair utility functions that is related to fairness, as we will show in Section II-E that the α-fair utility function is equal to the product of our fairness measure and a function of total resource for any $\beta = \alpha \geq 0$. The mapping from generators $g(y)$ to fairness measures $f_{\beta}(\mathbf{x})$ is illustrated in Figure 1.

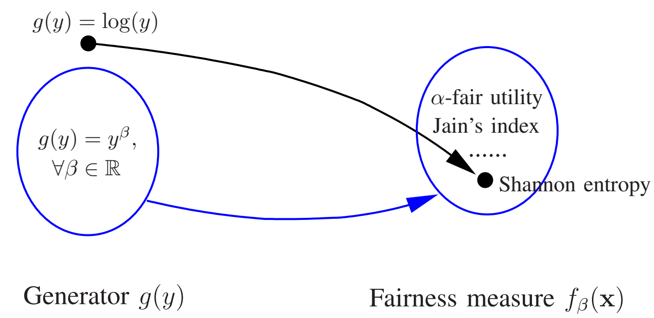

Fig. 1. The mapping from generators $g(y)$ to fairness measures $f_{\beta}(\mathbf{x})$. Among the set of all continuous and increasing functions, fairness measures satisfying axioms 1-5 can only be generated by logarithm and power functions. A logarithm generator ends up with the same form of f as a special case of those generated by a power function.

This axiomatic theory unifies many existing fairness indices and utilities. In particular, for $\beta < 1$, we recover the index-based approach in [1]–[6]. For $\beta \geq 0$, we recover the utility-based approach in [7], [8], [10]–[12], which quantify fairness for the maximizers of the α-fair utility functions. For the interesting range $\beta \in [0, 1)$, the two approaches overlap. Notice that for $\beta > 1$ the measures are in fact negative. This is a consequence of the requirement in Axiom 5 that the measure be monotonically *increasing* towards fairer solutions. As an illustration, for a fixed resource allocation vectors $\mathbf{x} = [1,\ 2,\ 3,\ 5]$ and $\mathbf{y} = [1,\ 1,\ 2.5,\ 5]$, we plot fairness $f_{\beta}(\mathbf{x})$ and $f_{\beta}(\mathbf{y})$ for different values of β in Figure 2. Different values of β clearly changes the fairness comparison ration, and may even result in different fairness orderings: $f_{\beta}(\mathbf{x}) \geq f_{\beta}(\mathbf{y})$ for $\beta \in (-\infty, 4.6]$, and $f_{\beta}(\mathbf{x}) \leq f_{\beta}(\mathbf{y})$ for $\beta \in [4.6, \infty)$.

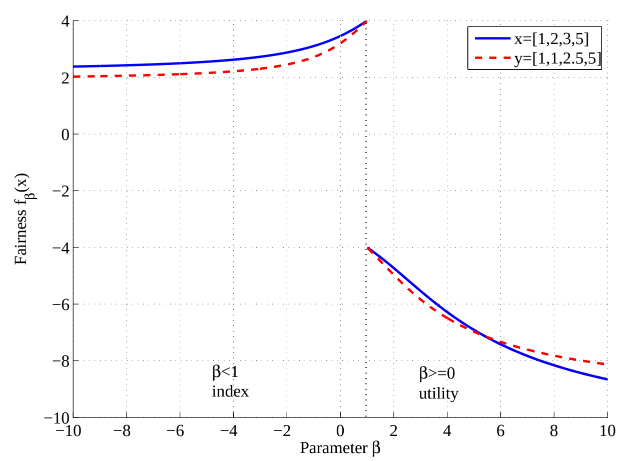

Fig. 2. For a given x, the plot of fairness $f_{\beta}(\mathbf{x})$ and $f_{\beta}(\mathbf{y})$ in (15) against β for $\mathbf{x} = [1,\ 2,\ 3,\ 5]$ and $\mathbf{y} = [1,\ 1,\ 2.5,\ 5]$: $\beta \geq 0$ recovers the utility-based approach, and $\beta < 1$ recovers the index-based approach.

Value of β	Type of Mean	Our Fairness Measure	Known Names
$\beta \to \infty$	maximum	$-\max_i \left\{ \frac{\sum_i x_i}{x_i} \right\}$	Max ratio
$\beta \in (1, \infty)$		$-\left[(1-\beta)U_{\alpha=\beta}\left(\frac{\mathbf{x}}{w(\mathbf{x})}\right)\right]^{\frac{1}{\beta}}$	α-fair utility
$\beta \in (0, 1)$		$\left[(1-\beta)U_{\alpha=\beta}\left(\frac{\mathbf{x}}{w(\mathbf{x})}\right)\right]^{\frac{1}{\beta}}$	α-fair utility
$\beta \to 0$	geometric	$e^{H\left(\frac{\mathbf{x}}{w(\mathbf{x})}\right)}$	Entropy
$\beta \in (0, -1)$		$\left[\sum_{i=1}^n \left(\frac{x_i}{w(\mathbf{x})}\right)^{1-\beta r}\right]^{\frac{1}{\beta}}$	No name
$\beta = -1$	harmonic	$\frac{(\sum_i x_i)^2}{\sum_i x_i^2} = n \cdot J(\mathbf{x})$	Jain's index
$\beta \in (-1, -\infty)$		$\left[\sum_{i=1}^n \left(\frac{x_i}{w(\mathbf{x})}\right)^{1-\beta r}\right]^{\frac{1}{\beta}}$	No name
$\beta \to -\infty$	minimum	$\min_i \left\{ \frac{\sum_i x_i}{x_i} \right\}$	Min ratio

TABLE III

PREVIOUS RESULTS OF FAIRNESS METRICS IN NETWORKING ARE RECOVERED AS SPECIAL CASES OF OUR AXIOMATIC CONSTRUCTION. FOR $\beta \in (0, -1)$ AND $\beta \in (-1, -\infty)$, NEW FAIRNESS MEASURES OF GENERALIZED JAIN'S INDEX ARE REVEALED. FOR $\beta \in [0, \infty]$ THE FAIRNESS COMPONENT OF α-FAIR UTILITY FUNCTION IS RECOVERED, SEE SECTION II-E1. IN PARTICULAR, PROPORTIONAL FAIRNESS AT $\alpha = \beta = 1$ IS OBTAINED FROM $\lim_{\beta \to 1} \frac{|f_\beta(\mathbf{x})|^\beta - n}{|1-\beta|}$.

C. Properties of Fairness Measures

In this section, we prove a set of properties for all fairness measures satisfying Axioms 1-5. We start an intuitive corollary from the five axioms that will be useful for the rest of the presentation. The first few can be readily proved directly from the axioms, without the explicit form in (15).

Corollary 1: (Symmetry.) A fairness measure satisfying Axioms 1–5 is symmetric over $\mathbf{x}$:

$$f(x_1, x_2, \ldots, x_n) = f(x_{i_1}, x_{i_2}, \ldots, x_{i_n}), \tag{16}$$

where $i_1, \ldots, i_n$ is an arbitrary permutation of indices $1, \ldots, n$.

The symmetry property shows that the fairness measure $f(\mathbf{x})$ satisfying Axioms 1–5 is independent of labeling of users. This corollary of the axioms may be viewed as objectionable by some. In Section IV, we will modify Axiom 4, so that the resulting fairness measure will depend on user-specific weights.

We now connect of our axiomatic theory to a line of work on measuring statistical dispersion by vector majorization, including the popular Gini coefficient [17]. Majorization [15] is a partial order over vectors to study whether the elements of vector $\mathbf{x}$ are less spread out than the elements of vector $\mathbf{y}$. $\mathbf{x}$ is majorized by $\mathbf{y}$, and we write $\mathbf{x} \preceq \mathbf{y}$, if $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$ (always satisfied due to Axiom 2) and

$$\sum_{i=1}^d x_i^\uparrow \le \sum_{i=1}^d y_i^\uparrow, \text{ for } d = 1, \ldots, n, \tag{17}$$

where $x_i^\uparrow$ and $y_i^\uparrow$ are the ith elements of $\mathbf{x}^\uparrow$ and $\mathbf{y}^\uparrow$, sorted in ascending order. According to this definition, among the vectors with the same sum of elements, one with the equal elements is the most majorizing vector.

Intuitively, $\mathbf{x} \preceq \mathbf{y}$ can be interpreted as $\mathbf{y}$ being a fairer allocation than $\mathbf{x}$. It is a classical result [15] that $\mathbf{x}$ is majorized by $\mathbf{y}$, if and only if, from $\mathbf{x}$ we can produce $\mathbf{y}$ by a finite sequence of Robin Hood operations.[1]

However, majorization alone cannot be used to define a fairness measure since it is a partial order and can fail to compare vectors. Still, if resource allocation $\mathbf{x}$ is majorized by $\mathbf{y}$, it is desirable to have a fairness measure f such that $f(\mathbf{x}) \le f(\mathbf{y})$. A function satisfying this property is known as Schur-concave. In statistics and economics, many measures of statistical dispersion or diversity are known to be Schur-concave, e.g., Gini Coefficient and Robin Hood Ratio [17], and we show our fairness measure also is Schur-concave. More discussions on the relationship of our axioms and other economics theories are provided in Appendix B.

Theorem 4: (*Schur-concavity.*) A fairness measure satisfying Axioms 1–5 is Schur-concave:

$$f(\mathbf{x}) \le f(\mathbf{y}), \text{ if } \mathbf{x} \preceq \mathbf{y}. \tag{18}$$

Next we present several properties of fairness measures satisfying the axioms, whose proofs rely on Schur-concavity.

Corollary 2: (Equal-allocation is most fair.) A fairness measure $f(\mathbf{x})$ satisfying Axioms 1–5 is maximized by equal-resource allocations, i.e.,

$$f(\mathbf{1}_n) = \max_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}). \tag{19}$$

Corollary 3: (Equal-allocation fairness value is independent of g.) Fairness achieved by equal-resource allocations $\mathbf{1}_n$

[1] In a Robin Hood operation, we replace two elements x_i and $x_j < x_i$ with $x_i - \epsilon$ and $x_j + \epsilon$, respectively, for some $\epsilon \in (0, x_i - x_j)$.

is independent of the choice of generator function g, i.e.,

$$f(\mathbf{1}_n) = n^r \cdot f(1), \tag{20}$$

where r is a constant exponent.

Corollary 4: (Inactive users do not change fairness.) When a fairness measure $f(\mathbf{x})$ satisfying Axioms 1–5 is generated by $\rho > 0$ in (7), adding or removing users with zero resources does not change fairness:

$$f(\mathbf{x}, \mathbf{0}_n) = f(\mathbf{x}), \quad \forall n \in \mathbb{Z}_+. \tag{21}$$

The following corollaries' proofs will directly use the unified representation in (15). First, by specifying a level of fairness, we can limit the number of starved users in a system.

Corollary 5: (Fairness value bounds the number of inactive users.) Fairness measures satisfying Axioms 1-5 count the number of inactive users in the system. When $f_\beta < 0$, $f(\mathbf{x}) \to -\infty$ if any user is assigned zero resource. When $f > 0$,

$$\text{Number of users with zero resource} \leq n - f(\mathbf{x}). \tag{22}$$

Corollary 6: (Fairness value bounds the maximum resource to a user.) Fairness measures satisfying Axioms 1-5 bound the maximum resource to a user when $f > 0$:

$$\text{Maximum resource to a user} \geq \frac{\sum_i x_i}{f(\mathbf{x})}. \tag{23}$$

Corollary 7: (Perturbation of fairness value from slight change in a user's resource.) If we increase resource allocation to user i by $\epsilon > 0$, while not changing other users' allocation, fairness measures satisfying Axioms 1-5 increase if and only if $x_i < \bar{x} = \left(\frac{\sum_j x_j}{\sum_j x_j^{1-\beta}}\right)^{\frac{1}{\beta}}$ and for all $0 < \epsilon < \bar{x} - x_i$.

This corollary implies that $\bar{x}$ serves as a threshold for identifying "poor" and "rich" users, since assigning an additional ϵ amount of resource to user i improves fairness if $x_i < \bar{x}$, while the same assignment reduces fairness if $x_i > \bar{x}$.

Corollary 8: (Box constraints of resources allocation bound fairness value.) If a resource allocation $\mathbf{x} = [x_1, x_2, \ldots, x_n]$ satisfies box-constraints, i.e., $x_{min} \leq x_i \leq x_{max}$ for all i, fairness measures satisfying Axioms 1-5 are lower bounded by a constant that only depends on β, x_{min}, x_{max}:

$$f(\mathbf{x}) \geq \text{sign}(1-\beta) \cdot \frac{\left(\mu\Gamma^{1-\beta} + 1 - \mu\right)^{\frac{1}{\beta}}}{\left(\mu\Gamma + 1 - \mu\right)^{\frac{1}{\beta}-1}}, \tag{24}$$

where $\Gamma = \frac{x_{max}}{x_{min}}$ and $\mu = \frac{\Gamma - \Gamma^{1-\beta} - \beta(\Gamma-1)}{\beta(\Gamma-1)(\Gamma^{1-\beta}-1)}$. The bound is tight when μ fraction of users have $x_i = x_{max}$ and the remaining $1-\mu$ fraction of users have $x_i = x_{min}$.

This corollary provides us with a lower bound on evaluation of the fairness measure, with which one might benchmark empirical values.

D. Implication 1: Generalizing Jain's Index

When $\beta = -1$ (i.e., harmonic mean is used in Axiom 4), f becomes a scalar multiple of the widely used Jain's index $J(\mathbf{x}) = \frac{1}{n} f(\mathbf{x})$. Upon inspection of (15) and the specific cases noted in Table III, we note that any $\beta \in (-\infty, 1)$ the range of fairness measure $f_\beta(\mathbf{x})$ lies between 1 and n. Equivalently, we can say that the fairness *per user* resides in the interval $\left[\frac{1}{n}, 1\right]$. We refer to this subclass of our family of fairness measures for $\beta \in (-\infty, 1)$ as the generalization of Jain's index.

Definition 2: $J_\beta(\mathbf{x}) = \frac{1}{n} f_\beta(\mathbf{x})$ is a generalized Jain's index parameterized by $\beta \leq 1$.

The common properties of our fairness index proven in Section II-C carry over to this generalized Jain's index. For $\beta = -1$, $J_{-1}(\mathbf{x})$ reduces to the original Jain's index. The parameter β determines the choice of generator function $g(y) = y^\beta$ in Axiom 4. We use $f_\beta(\mathbf{x})$ to denote the fairness measures in (15), parameterized by β. We first prove that, for a given resource allocation $\mathbf{x}$, fairness $f_\beta(\mathbf{x})$ is monotonic as $\beta \to 1$. Its implication is discussed next.

Theorem 5: (*Monotonicity with respect to β.*) The fairness measures in (15) is negative and decreasing for $\beta \in (1, \infty)$, and positive and increasing for $\beta \in (-\infty, 1)$:

$$\frac{\partial f_\beta(\mathbf{x})}{\partial \beta} \leq 0 \text{ for } \beta \in (1, \infty), \tag{25}$$

$$\frac{\partial f_\beta(\mathbf{x})}{\partial \beta} \geq 0 \text{ for } \beta \in (-\infty, 1). \tag{26}$$

As $\beta \to 1$, f point-wise converges to constant values:

$$\lim_{\beta \uparrow 1} f_\beta(\mathbf{x}) = n \text{ and } \lim_{\beta \downarrow 1} f_\beta(\mathbf{x}) = -n. \tag{27}$$

The monotonicity of fairness measures $f_\beta(\mathbf{x})$ on $\beta \in (-\infty, 1)$ gives an interpretation of β. Figure 3 plots fairness $f_\beta(\theta, 1-\theta)$ for resource allocation $\mathbf{x} = [\theta, 1-\theta]$ and different choices of $\beta = \{-4.0, -2.5, -1.0, 0.5\}$. The vertical bars in the figure represent the level sets of function f, for values $f_\beta(\theta_i, 1-\theta_i) = \frac{i}{10}\left(f_{max} - f_{min}\right), i = 1, 2, \ldots, 9$. For fixed resource allocations, since f increases as β approaches 1, the level sets of f are pushed toward the region with small θ (i.e., the low-fairness region), resulting in a steeper incline in the region. In the extreme case of $\beta = 1$, all level set boundaries align with the y-axis in the plot. The fairness measure f point-wise converges to step functions $f_\beta(\theta, 1-\theta) = 2$. Therefore, parameter β characterizes the shape of the fairness measures: a smaller value of $|1-\beta|$ (i.e., β closer to 1) causes the level sets to be more condensed in the low-fairness region.

Since the fairness measure must still evaluate to a number between 1 and n here, the monotonicity, and the resulting change in the resolution of fairness measure associated with varying β, also lead to differences in how unfairness is evaluated. At one extreme, as $\beta \to 1$ any solution where no user receives an allocation of zero is fairest. On the other hand, as $\beta \to -\infty$ the relationship between $f_\beta(\mathbf{x})$ and θ becomes linear, suggesting a stricter concept of fairness – for the same allocation, as $\beta \to -\infty$ more fairness is lost. Therefore, the parameter β can tune generalized Jain's index f for different tradeoffs between resolution and strictness of fairness measure.

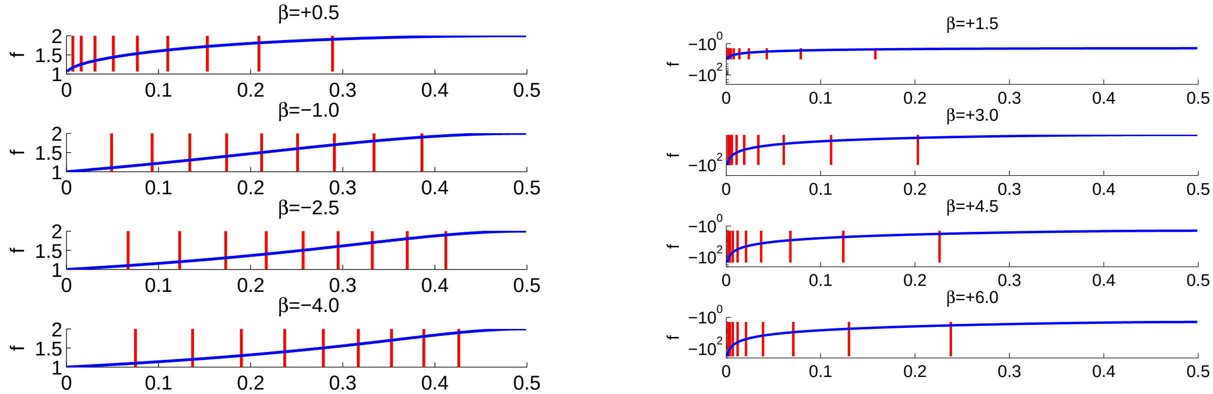

Fig. 3. Plot of the fairness measure $f_\beta(\theta, 1-\theta)$ against θ, for resource allocation $\mathbf{x} = [\theta, 1-\theta]$ and different choices of $\beta = \{-4.0, -2.5, -1.0, 0.5\}$ and $\beta = \{1.5, 3.0, 4.5, 6.0\}$, respectively. It can be observed that $f_\beta(\theta, 1-\theta)$ is monotonic as $\beta \to 1$. Further, smaller values of $|1-\beta|$ results in a steeper incline over small θ, i.e., the low-fairness region.

If the fairness measure f is used for classifying different resource allocations, a larger β is desirable, since it gives more quantization levels in low-fairness region and provides finer granularity control for unfair resource. On the other hand, if the fairness measure f is used as an objective function, a smaller β is desirable, since it has a steeper incline in the low-fairness region and give more incentive for the system to operate in the high-fairness region.

E. Implication 2: Understanding α-Fairness

Due to Axiom 2, our fairness measures are homogeneous functions of degree zero and only express desirability over the $(n-1)$-dimension subspace orthogonal to the $\mathbf{1}_n$ vector. Hence, they do not capture any notion of efficiency of an allocation. The component of resource vectors along the vector $\mathbf{1}_n$ describes another quantity used to capture the efficiency of an allocation, as a function of the sum $w(\mathbf{x})$ of resources.

We focus in this subsection on the widely applied α-fair utility function:

$$\sum_i U_\alpha(x_i), \text{ where } U_\alpha(x) = \begin{cases} \frac{x^{1-\alpha}}{1-\alpha} & \alpha \geq 0,\ \alpha \neq 1 \\ \log(x) & \alpha = 1 \end{cases}. \tag{28}$$

We first show that the α-fairness network utility function can be *factored* into two components: one corresponding to the family of fairness measures we constructed and one corresponding to efficiency. We then demonstrate that, for a fixed α, the factorization can be viewed as a single point on the optimal tradeoff curve between fairness and efficiency. Furthermore, this particular point is one where maximum emphasis is placed on fairness while maintaining Pareto optimality of the allocation. This allows us to quantitatively interpret the belief of "larger α is more fair" *across all* $\alpha \geq 0$, not just for $\alpha \in \{0, 1, \infty\}$. .

1) Factorization of α-fair Utility Function: Re-arranging the terms of the equation in Table III, we have

$$\begin{aligned} U_{\alpha=\beta}(\mathbf{x}) &= \frac{1}{1-\beta} |f_\beta(\mathbf{x})|^\beta \left(\sum_i x_i\right)^{1-\beta} \\ &= |f_\beta(\mathbf{x})|^\beta \cdot U_\beta\left(\sum_i x_i\right), \end{aligned} \tag{29}$$

where $U_\beta\left(\sum_i x_i\right)$ is the uni-variate version of the α-fair utility function with $\alpha = \beta \in [0, \infty)$. Our fairness measure $f_\beta(\mathbf{x})$ captures proportional fairness at $\alpha = \beta = 1$ and max-min fairness at $\alpha = \beta \to \infty$.

Equation (29) demonstrates that the α-fair utility functions can be factorized as the product of two components: a fairness measure, $|f_\beta(\mathbf{x})|^\beta$, and an efficiency measure, $U_\beta\left(\sum_i x_i\right)$. The fairness measure $|f_\beta(\mathbf{x})|^\beta$ only depends on the normalized distribution, $\mathbf{x}/(\sum_i x_i)$, of resources (due to Axiom 2), while the efficiency measure is a function of only the sum resource $\sum_i x_i$.

Allocation:	$\mathbf{x}$	
	$\swarrow$	$\searrow$
Factorize:	$\mathbf{x}/\sum_i x_i$	$\sum_i x_i$
	$\downarrow$	$\downarrow$
Measure:	$f_\beta\left(\mathbf{x}/\sum_i x_i\right)$	$U_\beta\left(\sum_i x_i\right)$
	$\searrow$	$\swarrow$
Combine:	$U_{\alpha=\beta}(\mathbf{x})$	

TABLE IV
ILLUSTRATION OF THE FACTORIZATION OF THE α-FAIR UTILITY FUNCTIONS INTO A FAIRNESS COMPONENT OF THE NORMALIZED RESOURCE DISTRIBUTION AND A EFFICIENCY COMPONENT OF THE SUM RESOURCE.

The factorization of α-fair utility functions is illustrated in Table IV and decouples the two components to tackle issues such as fairness-efficiency tradeoff and feasibility of $\mathbf{x}$ under a given constraint set. For example, it helps to explain the counter-intuitive throughput behavior in congestion control based on utility maximization [12]: an allocation vector that maximizes the α-fair utility with a larger α may not be less efficient, because the α-fair utility incorporates both fairness and efficiency at the same time.

The additional efficiency component in (29) can skew the optimizer (i.e., the resource allocation resulting from α-fair utility maximization) away from an equal allocation. For this to happen there must exist an allocation that is feasible (within the constraint set of realizable allocations) with a large enough gain in efficiency over all equal allocation allocations. Hence, the magnitude of this skewing depends on the fairness parameter ($\alpha = \beta$), the constraint set of $\mathbf{x}$, and the relative importance of fairness and efficiency.

Guided by the product form of (29), we consider a scalarization of the maximization of the two objectives: fairness and efficiency:

$$\Phi_\lambda(\mathbf{x}) = \lambda \ell\left(f_\beta\left(\mathbf{x}\right)\right) + \ell\left(\sum_i x_i\right), \tag{30}$$

where $\beta \in [0, \infty)$ is fixed, $\lambda \in [0, \infty)$ absorbs the exponent β in the fairness component of (29) and is a weight specifying the relative emphasis placed on the fairness, and

$$\ell(y) = \mathrm{sign}(y) \log(|y|). \tag{31}$$

The use of the log function later recovers the product in the factorization of (29) from the sum in (30).

2) What Does "Larger α is More Fair" Mean?: It is commonly believed that larger α is more fair, but it is not exactly clear what this statement means for general $\alpha \in [0, \infty)$. Guided by the factorization above and the axiomatic construction of fairness measures, we provide two interpretation of this statement that justify it from the viewpoints of Pareto optimality and geometry of the constraint set.

An allocation vector $\mathbf{x}$ is said to be Pareto dominated by $\mathbf{y}$ if $x_i \le y_i$ for all i and $x_i < y_i$ for at least some i. An allocation is called Pareto optimal if it is not Pareto dominated by any other feasible allocation. If the relative emphasis on efficiency is sufficiently high, Pareto optimality of the solution can be maintained. To preserve Pareto optimality, we require that if $\mathbf{y}$ Pareto dominates $\mathbf{x}$, then $\Phi_\lambda(\mathbf{y}) > \Phi_\lambda(\mathbf{x})$.

Theorem 6: (*Preserving Pareto optimality.*) The necessary and sufficient condition on λ, such that $\Phi_\lambda(\mathbf{y}) > \Phi_\lambda(\mathbf{x})$ if $\mathbf{y}$ Pareto dominates $\mathbf{x}$, is that λ must be no larger than a threshold, which only depends on β:

$$\lambda \le \bar{\lambda} \triangleq \left|\frac{\beta}{1-\beta}\right|. \tag{32}$$

Consider the set of maximizers of (30) for λ in the range in Theorem 6:

$$\mathbb{P} = \left\{\mathbf{x} : \mathbf{x} = \arg\max_{\mathbf{x}\in\mathbb{R}} \Phi_\lambda(\mathbf{x}),\ \forall \lambda \le \left|\frac{\beta}{1-\beta}\right|\right\}. \tag{33}$$

When weight $\lambda = 0$, the corresponding points in $\mathbb{P}$ is most efficient. When weight $\lambda = \left|\frac{\beta}{1-\beta}\right|$, it turns out that the factorization in (29) becomes the same as (30).

Summarizing the first interpretation of "larger α is more fair": α-fairness corresponds to the solution of an optimization that places the *maximum emphasis* on the fairness measure parameterized by $\beta = \alpha$ while *preserving Pareto optimality*.

Allocations in $\mathbb{P}$ corresponding to other values of λ achieve a tradeoff between fairness and efficiency, while Pareto optimality is preserved.

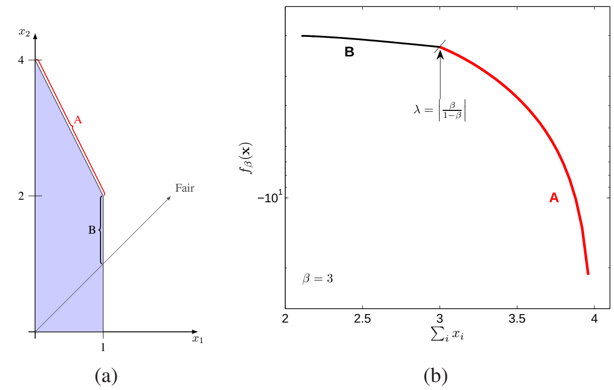

Fig. 4. (a) Feasible region (i.e., the constraint set of the utility maximization problem) where overemphasis of fairness violates Pareto dominance, and (b) its fairness-efficiency tradeoff for $\beta = 3$. Curve A corresponds to Pareto optimal solutions. Curve B is when the condition of Theorem 6 is violated, and resource allocations are more fair but no longer Pareto optimal.

Figure 4(b) illustrates an optimal fairness-efficiency tradeoff curve $\left\{\left[f_\beta(\mathbf{x}), \Sigma_i x_i\right],\ \forall \mathbf{x} = \arg\max_{\mathbf{x}\in\mathbb{R}} \Phi_\lambda(\mathbf{x}),\ \forall \lambda\right\}$ corresponding to the constraint set shown in Figure 4(a). The set of optimizers $\mathbb{P}$ in (33), which is obtained by maximizing Φ in (30), is shown by curve A in Figure 4(b). Curve B is when the condition of Theorem 6 is violated, and resource allocation is more fair but no longer Pareto optimal. Relationship with Rawl's theory of fairness will be explored in Appendix B.

3) Another Meaning of "Larger α is More Fair".: We just demonstrated the factorization (29) is an extreme point on the tradeoff curve between fairness and efficiency for fixed $\beta = \alpha$. What happens when α becomes bigger?

We denote by $\bigtriangledown_{\mathbf{x}}$ the gradient operator with respect to the vector $\mathbf{x}$. For a differentiable function, we use the standard inner product ($\langle \mathbf{x}, \mathbf{y}\rangle = \sum_i x_i y_i$) between the gradient of the function and a normalized vector to denote the directional derivative of the function.

Theorem 7: (*Monotonicity of fairness-efficiency reward ratio.*) Let allocation $\mathbf{x}$ be given. Define $\boldsymbol{\eta} = \frac{1}{n}\mathbf{1}_n - \frac{\mathbf{x}}{\sum x_i}$ as the vector pointing from the allocation to the nearest fairness maximizing solution. Then the fairness-efficiency reward ratio:

$$\frac{\left\langle \bigtriangledown_{\mathbf{x}} U_{\alpha=\beta}(\mathbf{x}), \dfrac{\boldsymbol{\eta}}{\|\boldsymbol{\eta}\|}\right\rangle}{\left\langle \bigtriangledown_{\mathbf{x}} U_{\alpha=\beta}(\mathbf{x}), \dfrac{\mathbf{1}_n}{\|\mathbf{1}_n\|}\right\rangle}, \tag{34}$$

is non-decreasing with α, i.e., higher α gives a greater relative reward for fairer solutions.

The choice of direction $\boldsymbol{\eta}$ is a result of Axiom 2 and Corollary 2, which together imply that $\boldsymbol{\eta}$ is the direction that increases fairness most and is orthogonal to increases in efficiency. Figure 5 illustrates Theorem 7 by showing the two

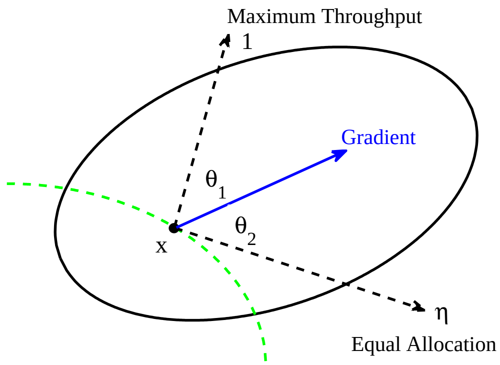

Fig. 5. This figure illustrates the two directions in Theorem 7: $\boldsymbol{\eta}$ (which points to the nearest fairness maximizing solution) and $\mathbf{1}$ (which points to the direction for maximizing total throughput).

directions: $\boldsymbol{\eta}$ (which points to the nearest fairness-maximizing solution) and $\mathbf{1}$ (which points to the direction for maximizing total throughput).

An increase in either fairness or efficiency is a "desirable" outcome. The choice of α dictates exactly *how desirable* one objective is relative to the other (for a fixed allocation). Theorem 7 states that, with a larger α, there is a larger component of the utility function gradient in the direction of fairer solutions, relative to the component in the direction of more efficiency. Economically, this could be seen as decreasing the pricing of fairer allocations relative to a fixed price on allocations of increased throughput. Notice, however, that comparison is in terms of the ratio between these two gradient components rather than the magnitude of the gradient, and both fairness and efficiency may increase simultaneously.

Theorem 7 provides the second interpretation of "larger α is more fair", again, not just for $\alpha \in \{0, 1, \infty\}$, but for any $\alpha \in [0, \infty)$. Figure 6 depicts how this ratio increases with $\alpha = \beta$ for some examples allocations.

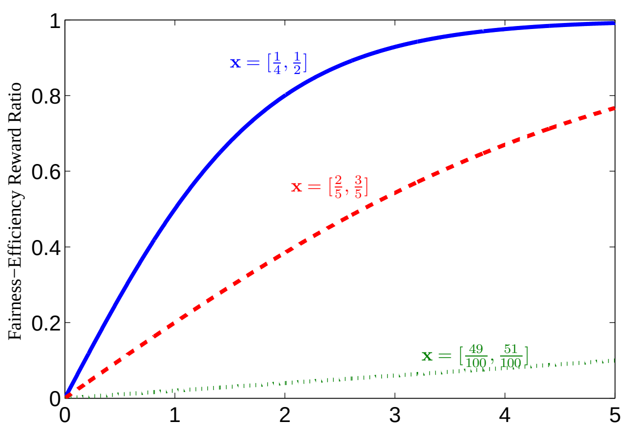

Fig. 6. Monotonic behavior of the ratio (34) as a function of α. Three fixed allocations over two users are considered, and solutions that are already more fair have a lower ratio.

This concludes the discussion of the first set of axioms of fairness. Obviously many interesting issues are missing: there is no time evolution of resource allocation, no differentiation among users, and no mentioning of efficiency in f_β. Some of the propositions, such as equal allocation maximizes fairness, may also be viewed as objectionable, thus motivating us to revisit the axioms. The next three sections carry out some of these extensions and revisions.

III. Continuous-Dimensional Allocations

A. Constructing Fairness Measures

Consider a resource allocation over time $[0, T]$, represented by $\mathbf{X}_{t\in[0,T]}$. The evolution over time may be due to a control algorithm or network dynamics. How do we quantify the fairness of $\mathbf{X}_{t\in[0,T]}$ using fairness measures?

In the previous section, we studied fairness measures $f : \mathbb{R}_+^n \to \mathbb{R}$, $\forall n \in \mathbb{Z}_+$, where resource allocation is represented by a non-negative vector $\mathbf{x}$. Utilizing these fairness measures, in this section we derive a new class of fairness measures, whose input $\mathbf{X}_{t\in[0,T]}$ is integrable over time $[0, T]$. The main idea is to divide $[0, T]$ into n subintervals, regard each subinterval as a resource element, and approximate the fairness of $\mathbf{X}_{t\in[0,T]}$ by the fairness of an n-dimensional resource vector as $n \to \infty$.

We start with fairness measure given by (15) in Section II. Due to Corollary 2 and Corollary 3, the maximum of fairness is achieved by equal resource allocations and is dependent on the number of resource elements, i.e., $f(\mathbf{1}_n) = n^r$. We normalize the fairness measure by n^r, so that $0 \leq n^{-r} \cdot f(\mathbf{x}) \leq 1$ for all $n > 0$ and non-negative $\mathbf{x}$. We have

$$\begin{aligned} & n^{-r} f(x_1, \ldots, x_n) \\ & = \text{sign}(r(1-\beta r)) \cdot n^{-r} \cdot \left[\frac{\sum_{i=1}^n x_i^{1-\beta r}}{\left(\sum_j x_j\right)^{1-\beta r}} \right]^{\frac{1}{\beta}} . \end{aligned} \quad (35)$$

We partition the interval $[0, T]$ by

$$0 = t_0 < t_1 < t_2 < \ldots < t_n = T, \quad (36)$$

where $t_i = iT/n = i\Delta$ and $\Delta = T/n$. We regard each subinterval by a resource element and approximate the resource allocation $\mathbf{X}_t$ by an n-dimensional resource vector

$$[x_1, \ldots, x_n] = [\mathbf{X}_{t_1}, \ldots, \mathbf{X}_{t_n}]. \quad (37)$$

Using (35) and (37), we have

$$\begin{aligned} & n^{-r} f(x_1, \ldots, x_n) \\ & = n^{-r} f(\mathbf{X}_{t_1}, \ldots, \mathbf{X}_{t_n}) \\ & = \text{sign}(r(1-\beta r)) \cdot n^{-r} \cdot \left[\frac{\sum_{i=1}^n \mathbf{X}_{t_i}^{1-\beta r}}{\left(\sum_j \mathbf{X}_{t_j}\right)^{1-\beta r}} \right]^{\frac{1}{\beta}} \\ & = \text{sign}(r(1-\beta r)) \cdot \frac{\Delta^r}{T^r} \cdot \left[\frac{\sum_{i=1}^n \mathbf{X}_{t_i}^{1-\beta r} \Delta^{1-\beta r}}{\left(\sum_j \mathbf{X}_{t_j} \Delta\right)^{1-\beta r}} \right]^{\frac{1}{\beta}} \\ & = \frac{\text{sign}(r(1-\beta r))}{T^r} \cdot \left[\frac{\sum_{i=1}^n \mathbf{X}_{t_i}^{1-\beta r} \Delta}{\left(\sum_j \mathbf{X}_{t_j} \Delta\right)^{1-\beta r}} \right]^{\frac{1}{\beta}} \end{aligned} \quad (38)$$

Since $\mathbf{X}_{t\in[0,T]}$ is integrable, we define a fairness measure for resource allocation over time by taking $n \to \infty$:

$$\begin{aligned}
&\mathcal{F}(\mathbf{X}_{t\in[0,T]}) \\
&\triangleq \lim_{n\to\infty} \frac{\mathrm{sign}(r(1-\beta r))}{T^r} \cdot \left[\frac{\sum_{i=1}^n \mathbf{X}_{t_i}^{1-\beta r}\Delta}{\left(\sum_j \mathbf{X}_{t_j}\Delta\right)^{1-\beta r}} \right]^{\frac{1}{\beta}} \\
&= \frac{\mathrm{sign}(r(1-\beta r))}{T^r} \cdot \left[\frac{\int_{t=0}^T \mathbf{X}_t^{1-\beta r} dt}{\left(\int_{t=0}^T \mathbf{X}_t dt\right)^{1-\beta r}} \right]^{\frac{1}{\beta}} \\
&= \mathrm{sign}(r(1-\beta r)) \cdot \left[\frac{1}{T} \int_{t=0}^T (\mathbf{X}_t/\mu)^{1-\beta r}\, dt \right]^{\frac{1}{\beta}} \quad (39)
\end{aligned}$$

where μ is the average resource over $[0,t]$ and is given by

$$\mu \triangleq \frac{1}{T}\int_{t=0}^T \mathbf{X}_t dt. \quad (40)$$

Equation (39) defines a family of fairness measures for resource allocation $\mathbf{X}_{t\in[0,T]}$, as β sweeps from $-\infty$ to ∞. It is well-defined due to integrability of $\mathbf{X}_{t\in[0,T]}$. It can be viewed as a continuous version of fairness measure (15) in Section II-B. For $\beta = -1$, fairness (39) extends Jain's index to resource allocation over time $[0,T]$, while for $\beta \geq 0$, it recovers α-fair utility functions. Similar to its counterpart (15), fairness $\mathcal{F}(\mathbf{X})$ satisfy a number of properties, which give interesting engineering implications. We first prove two equivalent representations for fairness in (39), with respect to cumulative distribution and to Lorenz curve [18] of resource allocation $\mathbf{X}_{t\in[0,T]}$:

Theorem 8: (*Equivalent representations.*) For arbitrary r, β, fairness $\mathcal{F}(\mathbf{X}$ of resource allocation $\mathbf{X}_{t\in[0,T]}$ is equivalent to the following two representations, with respect to cumulative distribution $P_{\mathbf{x}}(y)$ and to Lorenz curve $L_{\mathbf{x}}(u)$, respectively:

$$\mathcal{F} = \mathrm{sign}(r(1-\beta r)) \cdot \left[\int_{y=0}^{\infty} \left(\frac{y}{\mu}\right)^{1-\beta r} dP_{\mathbf{x}}(y) \right]^{\frac{1}{\beta}} \quad (41)$$

$$\mathcal{F} = \mathrm{sign}(r(1-\beta r)) \cdot \left[\int_{u=0}^{1} \left(\frac{P_{\mathbf{x}}^{-1}(u)}{\mu}\right)^{-\beta r} dL_{\mathbf{x}}(u) \right]^{\frac{1}{\beta}} \quad (42)$$

where cumulative distribution of $\mathbf{X}_{t\in[0,T]}$ is given by

$$P_{\mathbf{x}}(y) = \frac{1}{T} \cdot \int_{\{\mathbf{X}_t \leq y\}} dt \quad (43)$$

and Lorenz curve [18] of $\mathbf{X}_{t\in[0,T]}$ is given by

$$L_{\mathbf{x}}(u) = \frac{1}{\mu} \cdot \int_{\{P_{\mathbf{x}}(y)\leq u\}} y dP_{\mathbf{x}}(y). \quad (44)$$

This theorem shows that the fairness over time can be calculated through cumulative distribution or Lorenz curve of $\mathbf{X}_{t\in[0,T]}$. Not only does it facilitate the computation of fairness over time, it also extends majorization and Schur-concavity, which is proven in Section III-C for n-dimensional inputs, to resource allocation $\mathbf{X}_{t\in[0,T]}$. We define generalized majorization of resource allocations according to their Lorenz curves. We say that $\mathbf{X}_{t\in[0,T]}$ is majorized by $\mathbf{Y}_{t\in[0,T]}$:

$$\mathbf{X}_{t\in[0,T]} \preceq \mathbf{Y}_{t\in[0,T]}, \quad \text{if} \quad L_{\mathbf{x}}(u) \leq L_{\mathbf{y}}(u), \quad \forall u \in [0,1] \quad (45)$$

where $L_{\mathbf{x}}$ and $L_{\mathbf{y}}$ are the Lorenz curve of $\mathbf{X}_{t\in[0,T]}$, as defined by (44). According to this definition, it is easy to see that, among resource allocation over $[0,T]$, one with the equal resource over time is the most majorizing allocation.

B. Properties and Implications

We present several properties of fairness measures. Some of the proofs rely on Schur-concavity, which is proven in the following corollary.

Theorem 9: (*Schur-concavity.*) The fairness measure $\mathcal{F}$ in (39) is Schur-concave:

$$\mathcal{F}(\mathbf{X}_{t\in[0,T]}) \leq \mathcal{F}(\mathbf{Y}_{t\in[0,T]}), \text{ if } \mathbf{X}_{t\in[0,T]} \preceq \mathbf{Y}_{t\in[0,T]}. \quad (46)$$

Corollary 9: (Equal-allocation is most fair.) Fairness measure $\mathcal{F}$ is maximized by an equal allocation over time, i.e., $\{\mathbf{X}_t = c : t \in [0,T], c > 0\}$.

Corollary 10: (Equal-allocation fairness value is independent of T.) The fairness achieved by equal allocations over time is independent of the length of time T, i.e.,

$$\max_{\mathbf{X}_{t\in[0,T]}} \mathcal{F}(\mathbf{X}_{t\in[0,T]}) = \begin{cases} 1 & \text{if } 1 - r\beta \geq 0 \\ -1 & \text{o.w.} \end{cases} \quad (47)$$

Corollary 11: (Negative shift reduces fairness.) If a fixed amount $c > 0$ of the resource is subtracted for all $t \in [0,T]$, the resulting fairness measure decreases

$$\mathcal{F}(\mathbf{X}_{t\in[0,T]} - c) \leq \mathcal{F}(\mathbf{X}_{t\in[0,T]}), \quad \forall c > 0, \quad (48)$$

where $c > 0$ must be small enough such that $\mathcal{F}(\mathbf{X}_{t\in[0,T]}) - c$ is positive over $t \in [0,T]$.

Corollary 12: (Starvation time reduces fairness.) For all $r, \beta \in \mathbb{R}$, adding a time interval of length $\tau > 0$ with zero resource reduces fairness by a factor of $(1+\tau/T)^{-r}$

$$\mathcal{F}(\mathbf{X}_{t\in[0,T]}) \bigcup \mathbf{0}_{t\in[T,T+\tau]}) = (1+\tau/T)^{-r} \cdot \mathcal{F}(\mathbf{X}_{t\in[0,T]}). \quad (49)$$

Corollary 13: (Fairness value bounds the length of starvation time intervals.) When $1 - r\beta > 0$, fairness measure $\mathcal{F}$ gives an upper bound of starvation time intervals (with zero resource) in the system, i.e.,

$$\text{Length of starvation time} \leq T \cdot [|\mathcal{F}|^{-1/r} - 1]. \quad (50)$$

Corollary 14: (Fairness value bounds peak resource.) When $1 - r\beta > 0$, fairness $\mathcal{F}$ bounds the peak resource allocation over time $[0,T]$:

$$\max_{t\in[0,T]} \mathbf{X}_t \geq \frac{\mu}{2 - |\mathcal{F}|^{-1/r}} \quad (51)$$

Corollary 15: (Box constraints of resources allocation bound fairness value.) If a resource allocation $\mathbf{X}_{t\in[0,T]}$ satisfies box-constraints, i.e., $x_{min} \leq \mathbf{X}_t \leq x_{max}$ for all $t \in [0,T]$, the fairness measures $\mathcal{F}$ is lower bounded by a constant that only depends on β, rx_{min}, x_{max}:

$$\mathcal{F} \geq \text{sign}(r(1-\beta)) \left[\frac{\theta x_{max}^{1-r\beta} + (1-\theta) x_{min}^{1-r\beta}}{(\theta x_{max} + (1-\theta) x_{min})^{1-r\beta}} \right]^{\frac{1}{\beta}}, \quad (52)$$

where θ is the fraction of time that $\mathbf{X}_{t\in[0,T]}$ equals to x_{max} and is given by

$$\theta = \frac{1}{1-\rho} \cdot \left[\frac{(\Gamma-1)^\rho}{\Gamma^\rho - 1} - 1 \right] \quad (53)$$

where $\Gamma = x_{max}/x_{min}$ and $\rho = 1 - r\beta$. The bound is tight when for θ fraction of time $\mathbf{X}_{t\in[0,T]}$ equals to x_{max} and or the remaining $1-\theta$ fraction of time $\mathbf{X}_{t\in[0,T]}$ equals to x_{min}.

Corollary 16: (*Monotonicity with respect to* β.) For a fixed resource allocation $\mathbf{X}_{t\in[0,T]}$ with $T = 1$, fairness measure $\mathcal{F}$ is negative and decreasing for $\beta \in (1, \infty)$, and positive and increasing for $\beta \in (-\infty, 1)$:

$$\frac{\partial \mathcal{F}(\mathbf{X}_{t\in[0,T]})}{\partial \beta} \leq 0 \text{ for } \beta \in (1, \infty), \quad (54)$$

$$\frac{\partial \mathcal{F}(\mathbf{X}_{t\in[0,T]})}{\partial \beta} \geq 0 \text{ for } \beta \in (-\infty, 1). \quad (55)$$

The above results extend properties of f in the last section to fairness measures $\mathcal{F}$. Through Corollary 16 to Corollary 17, by specifying a level of fairness, we can limit the time of starvation. Corollary 18 implies that peak resource over time $t \in [0, T]$ is bounded by average resource and fairness - For given average resource, a smaller fairness requires an allocation with larger peak resource. Finally, Corollary 19 provides us with a lower bound on evaluation of the fairness measure, with which one may better benchmark empirical values.

C. Two-Dimensional Fairness Over Time and Users

In order to extend the model to be two-dimensional, we composite two different fairness measures among users and across time, respectively. This technique is also recently used in [44] to introduce multiscale fairness measures.

Given a resource allocation $\mathbf{X}_{t\in[0,T]} = \left[X_{1,t\in[0,T]}, \ldots, X_{n,t\in[0,T]}\right]$ over time $[0, T]$ and for n users. We can first apply $\mathcal{F}$ derived in this section to measure each user's fairness over time and then aggregate these fairness values using previous f with respect to user-specific weights $[q_1, \ldots, q_n]$. Therefore, we define the following two-dimensional fairness measures:

$$f \circ \mathcal{F}(\mathbf{X}_{t\in[0,T]}) \triangleq f\left(\left[\mathcal{F}(X_{1,t\in[0,T]}), \ldots, \mathcal{F}(X_{n,t\in[0,T]})\right], [q_1, \ldots, q_n]\right) \quad (56)$$

An alternative approach is to first use f to compute fairness for each time $t \in [0, T]$ and then apply $\mathcal{F}$ to measure its dynamic over time. This second approach gives a different definition of two-dimensional fairness measures:

$$\mathcal{F} \circ f(\mathbf{X}_{t\in[0,T]}) \triangleq \mathcal{F}\left(f\left(\left[X_{1,t\in[0,T]}, \ldots, X_{n,t\in[0,T]}\right], [q_1, \ldots, q_n]\right), \forall t\right) \quad (57)$$

A desirable property is to choose f and $\mathcal{F}$, such that the two-dimensional fairness measures in (56) and (57) become the same, i.e., $f \circ \mathcal{F}(\mathbf{X}_{t\in[0,T]}) = \mathcal{F} \circ f(\mathbf{X}_{t\in[0,T]})$, for all distributions $\mathbf{X}_{t\in[0,T]}$. Therefore, the same fairness value is guaranteed no matter which dimension of fairness is measured first.

Corollary 17: (Two-dimensional fairness measures.) For all distributions $\mathbf{X}_{t\in[0,T]}$, we have $f \circ \mathcal{F}(\mathbf{X}_{t\in[0,T]}) = \mathcal{F} \circ f(\mathbf{X}_{t\in[0,T]})$ if and only if both fairness measures f and $\mathcal{F}$ choose the same set of parameters: $1/\lambda = 1$, $r = -1$, and $\beta \in R$. The resulting two-dimensional fairness measures are given by

$$\text{sign}(1-\beta) \cdot \left[\sum_{i=1}^{n} \frac{q_i}{T} \cdot \int_{t=0}^{T} (\mathbf{X}_{i,t})^\beta \, dt \right]^{\frac{1}{\beta}}. \quad (58)$$

IV. Asymmetric Users

If we want to construct fairness measures with properties different from those proven in Section II, we have to revisit the five axioms in Section II and modify or remove some of them. For instance, one limitation of Axioms 1-5 in Section II is that the resulting fairness measures are symmetric. As a result, equal resource allocation is always most fair, irrelevant of users' requirements and contributions to the system. To be general enough in many contexts, fairness measures should incorporate asymmetric functions. Towards this end, in this section we propose a new set of five axioms, which result in a different family of asymmetric fairness measures. We start with two vectors now: $\mathbf{x}$ the resource allocation and $\mathbf{q}$ the user-specific weights, which may be arrived at from different considerations including the needs and contributions of each user.

A. A Second Set of Five Axioms

Given two n-dimensional vectors, $\mathbf{x}$ and $\mathbf{q}$, let fairness measure $\{f^n(\mathbf{x}, \mathbf{q}), \forall n \in \mathbb{Z}_+\}$ be a sequence of mapping from $(\mathbf{x}, \mathbf{q})$ to real numbers, i.e., $\{f^n : \mathbb{R}_+^n \times \mathbb{R}_+^n \to \mathbb{R}, \forall n \in \mathbb{Z}_+\}$. Again we suppress n for notational simplicity. In the following, we propose a new set of Axioms by modifying Axiom 2 (i.e., the Axiom of Partition) to introduce user-specific weights q_j for each user j in the axiom, which provides a value statement on relative importance of users in quantifying fairness of the system. Axiom 5 is restated for two users with equal weights, since it may not hold in general when fairness depends on user-specific weights.

The Axioms of Continuity, Homogeneity, and Saturation in Section II remains unchanged.

1′) *Axiom of Continuity.* Fairness measure $f(\mathbf{x}, \mathbf{q})$ is continuous on $\mathbb{R}_+^n$ for all $n \in \mathbb{Z}_+$.

2′) *Axiom of Homogeneity.* Fairness measure $f(\mathbf{x}, \mathbf{q})$ is a homogeneous function of degree 0:

$$f(\mathbf{x}, \mathbf{q}) = f(t \cdot \mathbf{x}, \mathbf{q}), \quad \forall\ t > 0. \quad (59)$$

Without loss of generality, for a single user, we take $|f(x_1, q_1)| = 1$ for all $q_1, x_1 > 0$, i.e., fairness is a constant for $n = 1$.

3′) *Axiom of Saturation.* Fairness value of equal allocation and equal weights is independent of number of users as the number of users becomes large, i.e.,

$$\lim_{n\to\infty} \frac{f(\mathbf{1}_{n+1}, \mathbf{1}_{n+1})}{f(\mathbf{1}_n, \mathbf{1}_n)} = 1. \qquad (60)$$

4′) *Axiom of Partition.* Consider an arbitrary partition of a system into two sub-systems. Let $\mathbf{x} = \left[\mathbf{x}^1, \mathbf{x}^2\right]$ and $\mathbf{y} = \left[\mathbf{y}^1, \mathbf{y}^2\right]$ be two resource allocation vectors, each partitioned and satisfying $\sum_j x_j^i = \sum_j y_j^i$, for $i = 1, 2$. Suppose $\mathbf{q} = [\mathbf{q}^1, \mathbf{q}^2]$ and $\mathbf{p} = [\mathbf{p}^1, \mathbf{p}^2]$ are corresponding weights for $\mathbf{x}$ and $\mathbf{y}$ and also satisfy $\sum_j q_j^i = \sum_j p_j^i$, for $i = 1, 2$. There exists a mean function [13] generated by g, such that their fairness ratio is the mean of the fairness ratios of the subsystems' allocations, for all possible partitions:

$$\frac{f(\mathbf{x}, \mathbf{q})}{f(\mathbf{y}, \mathbf{p})} = g^{-1}\left(\sum_{i=1}^{2} s_i \cdot g\left(\frac{f(\mathbf{x}^i, \mathbf{q}^i)}{f(\mathbf{y}^i, \mathbf{p}^i)}\right)\right) \qquad (61)$$

where g is any continuous and strictly monotonic function. s_i are positive weights given by

$$s_i = \frac{1}{c} \sum_{j\in\mathcal{I}_i} q_j \cdot w^{\rho}(\mathbf{x}^i) \text{ for } i = 1, 2. \qquad (62)$$

where $1/c$ is a normalization parameter, such that $\sum_i s_i = 1$, and $\mathcal{I}_i$ is the set of user indices in subsystem i.

The weights s_i for $i = 1, 2$ in (62) are constants for a given partition of sub-systems. Axiom 4′ is similar to Axiom 4 in Section II, except that a positive constant weight q_j is assigned to each user j, and that weights s_i in the generalized mean value in (61) are now proportional to the product of sum resource and total user-specific weights.

Similar to Section II, we regard g as a generator function for fairness if the resulting f satisfies all the axioms. . We show that Axiom 4′ also implies a hierarchical construction of fairness,

$$\begin{aligned} f(\mathbf{x}, q) = \quad & f\left([w(\mathbf{x}^1), w(\mathbf{x}^2)], [w(\mathbf{1}^1), w(\mathbf{q}^2)]\right) \\ & \cdot g^{-1}\left(\sum_{i=1}^{2} s_i \cdot g\left(\left|f(\mathbf{x}^i, \mathbf{q}^i)\right|\right)\right), \end{aligned} \qquad (63)$$

which allows us to derive a fairness measure $f : \mathbb{R}_+^n \to \mathbb{R}$ of n users recursively (with respect to a generator function $g(y)$) from lower dimensions, $f : \mathbb{R}_+^k \to \mathbb{R}$ and $f : \mathbb{R}_+^{n-k} \to \mathbb{R}$ for integer $0 < k < n$.

5′) *Axiom of Starvation* For $n = 2$ users with $q_1 = q_2 = q$, we have $f([1, 0], [q, q]) \le f(\left[\frac{1}{2}, \frac{1}{2}\right], [q, q])$, i.e., starvation is no more fair than equal allocation under equal weights.

Since fairness measures for $n = 2$ users with $q_1 = q_2$ becomes symmetric, Axiom 5′ is the weakest statement, thus the strongest axiom, that is consistent with Axiom 5 in Section II. We will later see that it specifies an increasing direction of fairness and is used to ensure uniqueness of $f(\mathbf{x}, \mathbf{q})$.

We demonstrate the following existence and uniqueness results. Similar to Section II, we prove that, from logarithm or power generator function $g(y)$, we can generate a unique $f(\mathbf{x}, \mathbf{q})$, satisfying Axioms 1′–5′. No other generator function $g(y)$ can satisfy axioms 1′-5′ with any $f(\mathbf{x}, \mathbf{q})$.

Theorem 10: (*Existence.*) There exists a fairness measure $f(\mathbf{x}, \mathbf{q})$ satisfying Axioms 1′–5′.

Theorem 11: (*Uniqueness.*) The family of $f(\mathbf{x}, \mathbf{q})$ satisfying Axioms 1′–5′ is uniquely determined by logarithm and power generator functions,

$$g(y) = \log y \text{ and } g(y) = y^{\beta} \qquad (64)$$

where $\beta \in \mathbb{R}$ is an arbitrary constant.

B. Constructing Asymmetric Fairness Measures

Without loss of generality, we assume that user-specific weights are normalized $\sum_j q_j = 1$. We derive the fairness measures in closed-form in the following theorem.

Theorem 12: (*Constructing all fairness measures satisfying Axioms* 1′-5′.) If $f(\mathbf{x}, \mathbf{q})$ is a fairness measure satisfying Axioms 1′-5′ with weight in (62), $f(\mathbf{x}, \mathbf{q})$ is of the form

$$f(\mathbf{x}, \mathbf{q}) = \text{sign}(-r) \prod_{i=1}^{n} \left(\frac{x_i}{\sum_{j=1}^{n} x_j}\right)^{-rq_i} \qquad (65)$$

$$f(\mathbf{x}, \mathbf{q}) = \text{sign}(-r(1 + r\beta)) \left[\sum_{i=1}^{n} q_i \left(\frac{x_i}{\sum_{j=1}^{n} x_j}\right)^{-r\beta}\right]^{\frac{1}{\beta}} \qquad (66)$$

where $r \in \mathbb{R}$ is a constant exponent. In (66), the two parameters in the mean function h have to satisfy $r\beta + \rho = 0$, where ρ is in the weights of the weighted sum and β is in the definition of the power generator function. It determines the fairness of equal resource allocation, which only depends on the number of users:

$$f(\mathbf{1}_n, \mathbf{1}_n) = n^r \cdot f(1, 1). \qquad (67)$$

The fairness measure in (65) can be obtained as a special case of (66) as $\beta \to 0$. Obviously, $\text{sign}(-r(1 + r\beta)) \to \text{sign}(-r)$ as $\beta \to 0$. By removing the sign of (66) and taking a logarithm transformation, we have

$$\begin{aligned} & \lim_{\beta\to 0} \log \left[\sum_{i=1}^{n} q_i \left(\frac{x_i}{\sum_{j=1}^{n} x_j}\right)^{-r\beta}\right]^{\frac{1}{\beta}} \\ & = \lim_{\beta\to 0} \frac{\log\left[\sum_{i=1}^{n} q_i \left(\frac{x_i}{\sum_{j=1}^{n} x_j}\right)^{-r\beta}\right]}{\beta} \\ & = \lim_{\beta\to 0} \sum_{i=1}^{n} -rq_i \log\left(\frac{x_i}{\sum_{j=1}^{n} x_j}\right) \cdot \left(\frac{x_i}{\sum_{j=1}^{n} x_j}\right)^{-r\beta} \\ & = \log \prod_{i=1}^{n} \left(\frac{x_i}{\sum_{j=1}^{n} x_j}\right)^{-rq_i} \end{aligned} \qquad (68)$$

where the second step uses L'Hospital's Rule. Therefore, we derive (65) by taking an exponential transformation in the last step of (68). This shows that fairness measure in (65) is a special case of (66) as $\beta \to 0$.

The mapping from generators $g(y)$ to fairness measures $f_\beta(\mathbf{x})$ is illustrated in Figure 7. Among the set of all continuous and increasing functions, fairness measures satisfying axioms $1'$-$5'$ can only be generated by logarithm and power functions. A logarithm generator ends up with the same form of f of a power function as the exponent $\beta \to 0$.

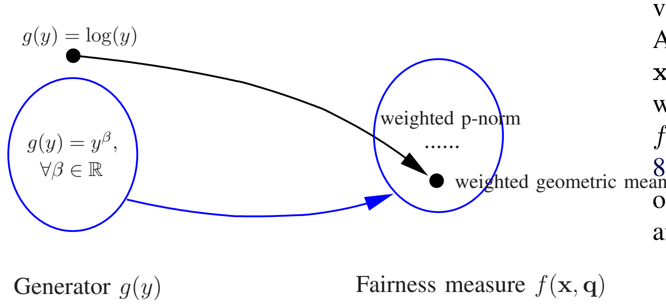

Fig. 7. The mapping from generators $g(y)$ to fairness measures $f(\mathbf{x}, \mathbf{q})$. Among the set of all continuous and increasing functions, fairness measures satisfying axioms $1'$-$5'$ can only be generated by logarithm and power functions. A logarithm generator ends up with the same form of f as a special case of those generated by a power function.

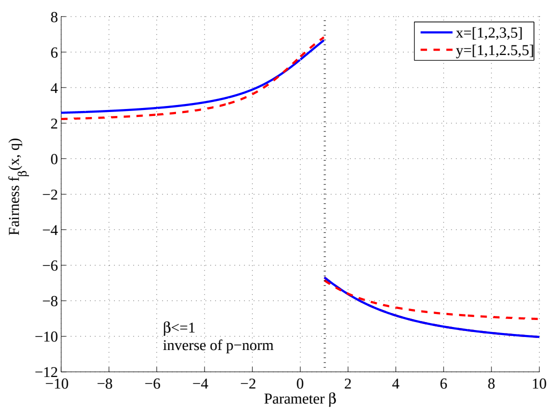

Fig. 8. For a given $\mathbf{x}$, the plot of fairness $f_\beta(\mathbf{x}, \mathbf{q})$ in (69) for different values of β: $\beta \le -1$ recovers inverse of p-norms for $p = -\beta$. Different values of β may result in same fairness ordering: $f_\beta(\mathbf{x}, \mathbf{q}) \ge f_\beta(\mathbf{y}, \mathbf{q})$ for $\beta \in (-\infty, -0.8] \cup [1, 1.8]$, and $f_\beta(\mathbf{x}, \mathbf{q}) \le f_\beta(\mathbf{y}, \mathbf{q})$ for $\beta \in [-0.8, 1] \cup [1.8, \infty)$.

We denote $f_{\beta,r}$ to be the fairness measure with parameters β and r. According to Theorem 12, r determines the growth rate of maximum fairness as population size n increases. Similar to Section II, we choose $r = 1$ and derive a unified representation of the constructed fairness measures:

$$f_\beta(\mathbf{x}, \mathbf{q}) = \mathrm{sign}(-1-\beta) \cdot \left[\sum_{i=1}^{n} q_i \left(\frac{x_i}{\sum_{j=1}^{n} x_j} \right)^{-\beta} \right]^{\frac{1}{\beta}}. \quad (69)$$

From here on, in this section we consider the family of fairness measures in (69) parameterized by β. Notice that (69) includes (multiplicative) inverse of weighted p-norm as a special case for $\beta \le -1$. In particular, for $\beta = -1$, we recover inverse of taxicab norm; for $\beta = -2$, we recover inverse of Euclidean norm; and for $\beta = -\infty$, we recover inverse of infinity norm or maximum norm.

Axioms $1'$-$5'$ differ from Axioms 1-5 in Section II mainly in that generalized mean value in Axioms $1'$-$5'$ depends on sum resource, as well as total user-specific weights, while in Axioms 1-5 it depends only on sum resource of partition. Table V summarizes the two choices of weights and the resulting fairness measures. The axiomatic theory now provides a unification of both symmetric and asymmetric fairness. As an illustration, for a fixed resource allocation vectors $\mathbf{x} = [1,\ 2,\ 3,\ 5]$ and $\mathbf{y} = [1,\ 1,\ 2.5,\ 5]$, and user-specific weights $\mathbf{q} = [0.4,\ 0.2,\ 0.2,\ 0.2]$, we plot fairness $f(\mathbf{x}, \mathbf{q})$ and $f(\mathbf{y}, \mathbf{q})$ in the unified representation (69) against β in Figure 8. Again, different values of β may result in different fairness orderings: $f_\beta(\mathbf{x}, \mathbf{q}) \ge f_\beta(\mathbf{y}, \mathbf{q})$ for $\beta \in (-\infty, -0.8] \cup [1, 1.8]$, and $f_\beta(\mathbf{x}, \mathbf{q}) \le f_\beta(\mathbf{y}, \mathbf{q})$ for $\beta \in [-0.8, 1] \cup [1.8, \infty)$.

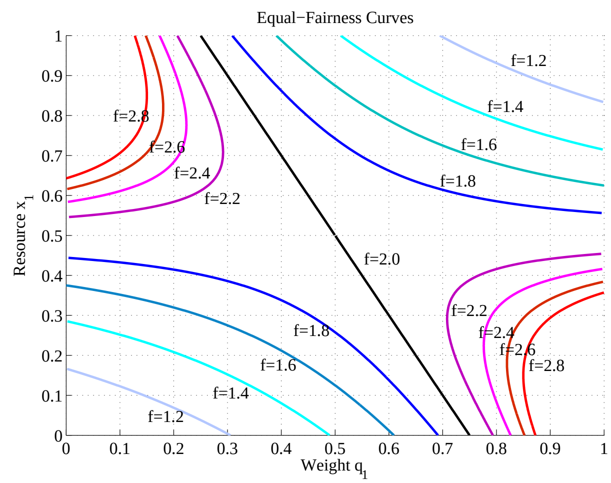

Fig. 9. Plot of equal fairness curves for $n = 2$ users. Different combinations of resource allocation $\mathbf{x} = [x_1, 1-x_1]$ and weight $\mathbf{q} = [q_1, 1-q_1]$ achieve the same fairness value for $f = 1.2$, 1.4, 1.6, 1.8, 2.0, 2.2, 2.4, 2.6, 2.8.

Different combinations of resource allocation $\mathbf{x}$ and weight $\mathbf{q}$ can achieve the same fairness value $f_\beta(\mathbf{x}, \mathbf{q})$. To illustrate this, Figure 9 plots equal fairness curves for $n = 2$ users with $\mathbf{x} = [x_1, 1-x_1]$ and $\mathbf{q} = [q_1, 1-q_1]$, i.e.,

$$\{(x_1, q_1) : f_\beta([x_1, 1-x_1], [q_1, 1-q_1]) = f\}, \quad (70)$$

where $\beta = -2$ and $f = 1.2,\ 1.4,\ \ldots,\ 2.8$. Unlike symmetric fairness measures in Section II, equal fairness curves for asymmetric fairness measures offer an extra degree of design freedom: we can choose different user-specific weights from various considerations including the needs and contributions of each user.

C. Properties and Implications

We prove the following properties for fairness measures derived in this section. Some properties are similar to their counterparts in Section II-C.

Corollary 18: (Weighted Symmetry.) Let π be a permutation of n elements. For a fairness measure satisfying Axioms $1'$–$5'$, we have

$$f(\mathbf{x}, \mathbf{q}) = f(\pi(\mathbf{x}), \pi(\mathbf{q})),$$

Axioms	Weight s_i	Generators	Fairness Measures
1–5	$w^{\rho}(\mathbf{x}^i)$	logarithm	$\text{sign}(r)\prod_{i=1}^n \left(\frac{x_i}{\sum_{j=1}^n x_j}\right)^{-r\left(\frac{x_i}{\sum_{j=1}^n x_j}\right)}$
		power	$\text{sign}(r(1-\beta r)) \cdot \left[\sum_{i=1}^n \left(\frac{x_i}{\sum_j x_j}\right)^{1-\beta r}\right]^{\frac{1}{\beta}}$
$1'$–$5'$	$(\sum_{j\in\mathcal{I}_i} q_j)\cdot w^{\rho}(\mathbf{x}^i)$	logarithm	$\text{sign}(-r)\prod_{i=1}^n \left(\frac{x_i}{\sum_{i=j}^n x_j}\right)^{-rq_i}$
		power	$\text{sign}(-r(1+r\beta)) \cdot \left[\sum_{i=1}^n q_i \left(\frac{x_i}{\sum_{j=1}^n x_j}\right)^{-r\beta}\right]^{\frac{1}{\beta}}$

TABLE V
COMPARING FAIRNESS MEASURES SATISFYING AXIOMS 1-5 AND AXIOMS $1'$-$5'$, WITH DIFFERENT CHOICES OF WEIGHTS IN GENERALIZED MEAN VALUE. NOTE THAT WEIGHTS IN THIS TABLE REMAINS TO BE NORMALIZED, SO THAT $\sum_i s_i = 1$.

for all permutation π.

Corollary 19: (Robin Hood Operations.) For small enough ϵ, replacing two elements x_i and x_j with $x_i-\epsilon$ and $x_j+\epsilon$ improves fairness if and only if $x_i/q_i^{\frac{1}{r\beta+1}} < x_j/q_j^{\frac{1}{r\beta+1}}$ for fairness measures satisfying Axioms $1'$–$5'$.

Corollary 20: (Most Fair Allocation.) Fairness measure satisfying Axioms $1'$–$5'$ is optimized by a resource allocation vector $[q_1^{\frac{1}{r\beta+1}}, q_2^{\frac{1}{r\beta+1}}, \ldots, q_n^{\frac{1}{r\beta+1}}]$.

Corollary 21: (Perturbation of fairness value from slight change in a user's resource.) For fairness measures satisfying Axioms $1'$–$5'$, if we increase resource allocation to user i by a small amount ϵ, while not changing other users' allocation, the fairness measure increases if and only if $x_i < q_i^{\frac{1}{r\beta+1}}\cdot\bar{x} = \left(q_i\sum_j x_j/\sum_j q_j x_j^{-r\beta}\right)^{\frac{1}{r\beta+1}}$ and $0<\epsilon<q_i^{\frac{1}{r\beta+1}}\cdot\bar{x}-x_i$.

Corollary 22: (*Monotonicity with respect to* β.) The fairness measures in (69) is negative and decreasing for $\beta \in (-1,\infty)$, and positive and increasing for $\beta\in(-\infty,-1)$:

$$\frac{\partial f_\beta(\mathbf{x},\mathbf{q})}{\partial\beta} \le 0 \text{ for } \beta\in(-1,\infty), \tag{71}$$

$$\frac{\partial f_\beta(\mathbf{x},\mathbf{q})}{\partial\beta} \ge 0 \text{ for } \beta\in(-\infty,-1). \tag{72}$$

Most properties have similar forms to their counterparts in Section II-C, while their implications become user-dependent. The properties of fairness measure (65) can be derived as a special case from that of fairness measure (66) as $\beta\to 0$. The weighted-symmetric property illustrates that user differentiation in Axioms $1'$-$5'$ is introduced in terms of user-specific weights. Users with equal weights are treated equally for fairness. Corollary 22 is an extension of the Robin Hood Operation and Schur-concavity property proven in Section II.

D. Towards A Fairness Evaluation Tool

Our axiomatic study might lead to a fairness evaluation tool, which turns the axiomatic theory into an easy-to-understand user-interface. We outline some possible functionalities of such a tool here.

The tool can (a) perform a consistency check of fairness based on discovering a person's underlying fairness function; (b) models user-specific weights $\mathbf{q}$ for different applications and settings; and (c) tune the tradeoff, via λ, between fairness and efficiency in system design objectives.

In this section, we present a human-subject experiment as a step towards using fairness theory as a consistency checking tool. The tool can be accessed from http://www.princeton.edu/∼chiangm/fairnesstool.html. In a typical variation of the experiment, each subject is told that there are two workers, A and B. A works twice as fast as B, and their monthly incomes are represented as a 1×2 vector. Through a sequence of simple Y/N questions, we will discover the underlying β and $\mathbf{q}$ used by the subject. We run the following three steps in our experiment.

1) Ask the subject a sequence of binary questions: is vector $\mathbf{x}$ (a particular income vector for workers A and B) more fair than $\mathbf{y}$ (another income vector)? For a given answer to the first question, plot the region of possible (q_1,β) region compatible with that answer, i.e., if $\mathbf{x}$ is more fair than $\mathbf{y}$, we choose

$$\{(q_1,\beta): f_\beta(\mathbf{x},\mathbf{q}) > f_\beta(\mathbf{x},\mathbf{q}), \mathbf{q}=[q_1,1-q_1]\} \tag{73}$$

Then adaptively pick another vector that can help shrink that region further. Iterate till we have a very thin slice of feasible region in (q_1,β) plane compatible with all the answers.
2) Ask the person what range of weights would she give to user 1 (and she may say, eg, $[0.6, 0.7]$). Then we know the range of β that is compatible with the answers so far.
3) Pick extreme values of beta in the above range, and tell the person that, according to her implicit fairness measure, vector $\mathbf{x}$ must be more fair than vector $\mathbf{y}$. Ask her if this consequence sounds alright to her? Based on her answer, further shrink the (q_1,β) range. This can be viewed as a mathematical expression of the principle of "reflective equilibrium" in political philosophy's study of distributive fairness [33], [55]. Finally, we can nail

down to a narrow range of (q_1, β) as her underlying fairness function.

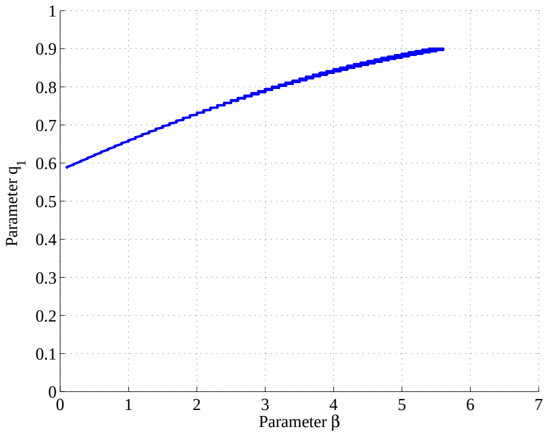

Fig. 10. Plot of all possible fairness parameter β and user-specific weight $\mathbf{q} = [q_1, 1 - q_1]$, such that the resulting fairness measures in (69)) satisfy constraints (74)-(78). Steps 2 and 3 could be repeated to further shrink the (q_1, β) range.

Consider the the family of fairness measures $f_\beta(\mathbf{x}, \mathbf{q})$ in (69). We take 8 different samplings of feasible allocation vectors, which are assigned the following fairness orders by a human subject:

$$[40,\ 60] \text{ is less fair than } [60,\ 40], \quad (74)$$

$$[50,\ 50] \text{ is less fair than } [58,\ 42], \quad (75)$$

$$[67,\ 33] \text{ is less fair than } [50,\ 50], \quad (76)$$

$$q_1 \in [0.5,\ 0.9] \quad (77)$$

$$[54,\ 46] \text{ is less fair than } [62,\ 38]. \quad (78)$$

We use an exhaustive search to find out which fairness measures (i.e., which fairness parameter β and user-specific weight $\mathbf{q} = [q_1, 1 - q_1]$ in (69)) satisfy fairness constraints (74)-(78). The resulting set of $\{(\beta, q_1)\}$ is plotted in Figure 10. It shows that the human subject implicitly assumes a fairness parameter $\beta \in [0.1, 5.5]$ and a weight vector $\mathbf{q} = [q_1, 1 - q_1]$ satisfying $q_1 \in [0.59, 0.9]$ in the test. No other fairness measures in (69) is consistent with the answers. Steps 2 and 3 could be repeated to further shrink the (q_1, β) range. In many experiments, we observe that the set of compatible fairness measures can be efficiently narrowed down using information provided by only 5 questions.

Through the above experiment, we do not need to ask explicitly for a subject's underlying β, which would have been a practically difficult question to answer. Instead, through a sequence of binary questions comparing two given allocation vectors, we can reverse-engineer the fairness function implicitly used by the subject. For example, a small scale experiment was recently run with the staff and students in Chiang's research group, and showed that the β value of these people concentrate between 0 and 1, which turns out to be the "busiest" part of the β spectrum.

Even when fairness measures are put on an axiomatic ground, different people will still prefer different fairness measures. What the unique, axiomatic construction, together with the discovery of β, can do is to enforce consistency of using the same β by the same person as she evaluates different resource allocations.

V. Fairness-Efficiency Unification

By removing some of the axioms — for example, Axiom (2 and $2'$) of Homogeneity that decouples efficiency from fairness — what kind of new fairness measures would result? We now develop an alternative set of fewer axioms that lead to the construction of fairness measures and do not automatically decouple fairness from efficiency and feasibility of resource allocation.

A. A Third Set of Four Axioms

We propose a set of alternative axioms, which includes Axioms 1–5 and Axioms $1'$–$5'$ as special cases[2]. Let $F : \mathbb{R}^n_+ \to \mathbb{R}$ for all $n \in \mathbb{Z}_+$ be a sequence of fairness measures satisfying four axioms as follows.

$1''$) *Axiom of Continuity.* Fairness measure $F(\mathbf{x})$ is continuous on $\mathbb{R}^n_+$ for all integer $n \in \mathbb{Z}_+$.

$2''$) *Axiom of Saturation.* Fairness measure $F(\mathbf{x})$ of equal resource allocations eventually becomes independent of the number of users:

$$\lim_{n\to\infty} \frac{F(\mathbf{1}_{n+1})}{F(\mathbf{1}_n)} = 1. \quad (79)$$

$3''$) Consider an arbitrary partition of a system into two sub-systems. Let $\mathbf{x} = \left[\mathbf{x}^1, \mathbf{x}^2\right]$ and $\mathbf{y} = \left[\mathbf{y}^1, \mathbf{y}^2\right]$ be two resource allocation vectors, each partitioned. There exists a mean function [13] generated by g, such that their fairness ratio is the mean of the fairness ratios of the subsystems' allocations, for all possible partitions and positive $t > 0$:

$$\frac{F(t \cdot \mathbf{x})}{F(t \cdot \mathbf{y})} = g^{-1}\left(\sum_{i=1}^{2} s_i \cdot g\left(\frac{F(\mathbf{x}^i)}{F(\mathbf{y}^i)}\right)\right) \quad (80)$$

where g is any continuous and strictly monotonic function. Weights s_i are chosen according to (7) and (62).

$4''$) *Axiom of Starvation* For $n = 2$ users, we have $F(1, 0) \leq F(\frac{1}{2}, \frac{1}{2})$, i.e., starvation is no more fair than equal allocation.

Axioms $1''$, $2''$, and $4''$ remain the same Axioms 1, 3, 5 and Axioms $1'$, $3'$, and $5'$ in previous sections. Since the Axiom 2 and ($2'$) of Homogeneity is removed, fairness measure $F(\mathbf{x})$ may depend on the absolute magnitude of resource vector $\mathbf{x}$. Axiom $3''$ requires (80) to be satisfied by all partitions and positive scaling factor t. It defines a hierarchical structure of resource allocation in a stronger sense than Axiom 4 in Section II and Axiom $4'$ in Section IV, which are special cases of Axiom $3''$ for $t = 1$.

[2]Weights $\mathbf{q}$ Axioms $1'$–$5'$ are suppressed for a unified representation of fairness F.

B. Deriving a Fairness-Efficiency Unification

Using Axiom $3''$, we can prove that $F(\mathbf{x})$ is a homogeneous function of real degree. Furthermore, if the order of homogeneity is zero, Axioms $1''$-$4''$ is equivalent to the two axiomatic systems, Axioms 1-5 in Section II and Axioms $1'$-$5'$ in Section IV for different choice of s_i in (7) and (62), respectively. This means that the new axiomatic system is more general than previous ones.

Theorem 13: (*Existence and Uniqueness.*) If fairness $F(\mathbf{x})$ satisfies Axioms $1''$-$4''$ with weights in (7) or (62), it is of the form

$$F(\mathbf{x}) = f(\mathbf{x}) \cdot \left(\sum_i x_i \right)^{\frac{1}{\lambda}}, \tag{81}$$

where $\frac{1}{\lambda} \in \mathbb{R}$ is the degree of homogeneity, and $f(\mathbf{x})$ is a symmetric fairness measure satisfying Axioms 1-5 in Section II or an symmetric fairness measure satisfying Axioms $1'$-$5'$ in Section IV.

It is easy to verify that some properties, like that of symmetry in Section II-C also hold for fairness measure $F(\mathbf{x})$ when weight s_i is chosen by (7), yet some properties of fairness measures satisfying Axioms 1–5 are changed in the generalization. In particular, the new fairness measure may not be Schur-concave, and equal allocation may not be fairness-maximizing. For instance, it can be readily verified that $F(1, 1) < F(0.2, 10)$ for $\beta = 2$ and $\frac{1}{\lambda} = 10$.

From Theorem 13 and uniqueness results in Sections 2 and 4, power generators $g(y) = y^\beta$ and logarithm generators $g(y) = \log y$ (viewed as a special case of power generators as $\beta \to 0$) are unique. We can thus construct a family of fairness measures $F_{\beta,\lambda}(\mathbf{x})$, which is parameterized by both λ and β,

$$F_{\beta,\lambda}(\mathbf{x}) = f_\beta(\mathbf{x}) \cdot \left(\sum_i x_i \right)^{\frac{1}{\lambda}}. \tag{82}$$

where $f_\beta(\mathbf{x})$ is the unified representation (15) in Section II and (69) in Section IV.

This recovers our results in previous sections: Generalized Jain's index is a special case of $F_{\beta,\lambda}(\mathbf{x})$ for s_i in (7), $1/\lambda = 0$, and $\beta < 1$; fairness measure $f_\beta(\mathbf{x})$ is a subclass of $F_{\beta,\lambda}(\mathbf{x})$ for s_i in (7) and $\lambda = 0$; inverse of p-norm is another subclass of $F_{\beta,\lambda}(\mathbf{x})$ for s_i in (62) and $\beta \leq -1$; and α-utility is obtained for s_i in (7), $1/\lambda = \beta/(1-\beta)$, and $\beta > 0$. The degree of homogeneity $1/\lambda$ determines how $F_{\beta,\lambda}(\mathbf{x})$ scales as throughput increases. The decomposition of fairness and efficiency in Section II-E is now an immediate consequence from Axioms $1''$-$4''$.

There is a useful connection with the characterization of α-fair utility function in Section III. The absolute value $|\lambda|$ is equivalent to the parameter used for defining the utility function (30) in Section II-E. From Theorem 6, we conclude that fairness measure $F_{\beta,\lambda}(\mathbf{x})$ is Pareto optimal if and only if

$$\frac{1}{|\lambda|} \geq \left| \frac{1-\beta}{\beta} \right|. \tag{83}$$

For a given β, there is a minimum degree of homogeneity such that Pareto optimality can be achieved. When inequality (83) is not satisfied, $F_{\beta,\lambda}(\mathbf{x})$ loses Pareto optimality and produces less throughput-efficient solutions if it is used as an objective function in utility maximization.

The degree of homogeneity of a fairness measure satisfying Axioms $1''$-$4''$ parameterizes a tradeoff between fairness and efficiency. Moreover, when power functions are used as generating functions, the degree of homogeneity is equivalent to $\frac{1}{\lambda}$ in (30). Therefore, the intuition behind our result on a maximum $|\lambda|$ (minimum degree of homogeneity) to ensure Pareto optimality can be extended to the general optimization-theoretic approach to fairness, i.e., for a fairness measure F generated from any g, there is a minimum degree of homogeneity $\frac{1}{\lambda}$ to produce a Pareto optimal solution. While our first set of axioms generalizes Jain's index and reveals new fairness measures, the third set of axioms offers a rich family of objective functions for resource allocation based on optimization formulations.

This family of fairness measures unifies a wide range of existing fairness indices and utilities in diverse fields briefly discussed in Appendix B. The unification is illustrated in Figure 11, including all known fairness measures that are global (i.e., mapping a given allocation vector to a single scalar and decomposable (i.e., subsystems' fairness values can be somehow collectively mapped into the overall system's fairness value).

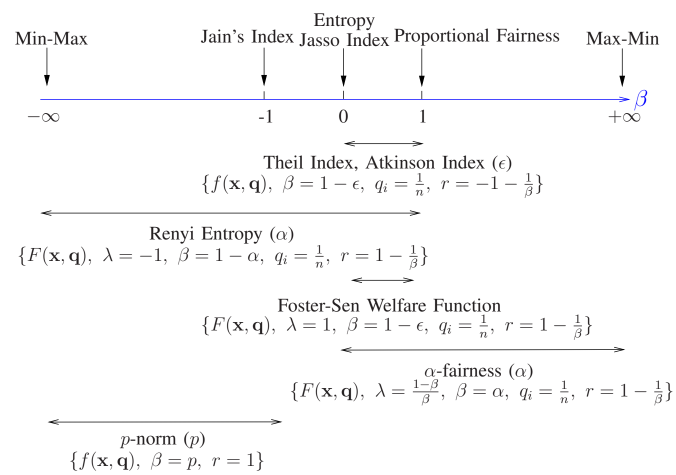

Fig. 11. F unifies different fairers measures from diverse disciplines.

VI. Conclusion

For a given resource allocation vector, different people will have different evaluations of how fair it is. These evaluations may correspond to image of the same fairness measure F but with different parameters β and λ. Realizing which parameters are implicitly used helps to ensure consistency where the same $F_{\beta,\lambda}$ is used to evaluate fairness of other allocation vectors.

An axiomatic quantification to the fundamental concepts of fairness also illuminates many issues in resource allocation. Extending a basic system of five axioms in Section II, we have shown that one way to re-examine axioms of fairness is to refute their more controversial corollaries. For example, the first set of axioms imply that fairness is always maximized by equal resource allocation and has no user-specific value or the notion of efficiency, a simplification that clearly is inadequate in many scenarios. In Section IV and Section V, we provide two alternative sets of axioms, which incorporates asymmetric fairness measures (by introducing user-specific weights) and unifies the notion of efficiency (i.e., making more resource available by scaling $\mathbf{x}$ in all coordinates could lead to larger fairness value), respectively. Another approach is to expand the data structure.

We started with an extremely simple one: given a vector, look for a scalar-valued function. In Section III, we extended fairness measures to continuous dimensional inputs, where resource allocations become non-negative functions over time. In Section V, we start with both a vector of resources and a vector of user weights, which may be arrived at from various considerations. Time-varying and user-specific values of resource can also be modeled via Sections IV and V.

There are in general two main approaches to fairness evaluation: *binary* (is it fair according to a particular criterion, e.g., proportional fair, no-envy?) or *continuous* (quantifying how fair is it, and how fair is it relative to another alloation?). In the latter case, there are three sub-approaches:

- (A) *A system-wide, global measure*: (A1) $f(\mathbf{x})$ where f is our fairness function, or (A2) $f(U_1, U_2, \ldots, U_n)$ where U_i is a utility function for each user that may depend on the entire $\mathbf{x}$.
- (B) *Individual, global measures*: the set of $\{f_i(\mathbf{x})\} = \{f_{\beta_i}(\mathbf{x})\}$ (user i cares about the entire allocation).
- (C) *Individual, local measures*: the set of $\{\tilde{f}_i(x_i)\}$ (user i only cares about her resource via some function $\tilde{f}$).

Approaches B and C may also involve some notion of expectation, based on need or contribution, and user-specific utility evaluation of resource.

We have so far been focusing on approach (A1) , whereas in fairness studies in computer science, economics, and sociology, discussed in Appendix B, approaches B and C have also been studied, in which one can think of two ways to aggregate the individual evaluations:

- *Scalarization*: $W(f_1, f_2, \ldots, f_n)$ for some increasing function W, with a special case being $\sum_i f_i$, and another special case being (A2), if U_i is f_i for all i.
- *Binary summary*: e.g., envy-freeness, which reduces to a *binary* evaluation criterion that is true if no user wants to change her resource with any other user.

In what ways will system-wide fairness evaluation and an aggregate of individual fairness evaluations be compatible? What will be $(W, \mathbf{q})$ choice's impact on the difference between $f_\beta(\mathbf{x}, \mathbf{q})$ and $W(f_1, f_2, \ldots, f_n)$? Addressing this question will also be a step towards bringing individual strategic action together with an allocator's decision over time.

Another distinction can be drawn between two types of fairness measures:

- *Global fairness measures* that assume "decomposability": the fairness value of a vector can be somehow reconstructed from the fairness values of the subvectors. This is the case in our (f, F), as well as in other cases in Appendix B such as Renyi entropy and Atkin's index.
- *Local fairness measures* that take an axiom of "local-closedness": an evaluation about x_1 and x_2 should be independent of x_3. These fairness measures include Gini index, Nash Bargaining Solution, and Shapley Value as typical examples.

A summary of the unification of these global, decomposable fairness measures into $F_{\beta,\lambda}(\mathbf{x}, \mathbf{q})$ is shown in Figure 11.

This paper presents *an* axiomatic theory, and is far from the end of axiomatic *theories* of fairness. For instance, we have assumed that there is only one type of resource, this resource x is infinitely divisible, and fairness measure is not a function of the feasible region of allocations. The following two axioms have remained unchanged throughout the paper: the Axiom of Continuity and the Axiom of Saturation, thus becoming interesting targets to investigate. Furthermore, we have ignored the process of resource allocation, such as the issues of who makes the allocation and whether it is made with autonomity of users. Time and time again, centralized systems with unchecked power claims to achieve fairness-maximizing allocation and actually produces extreme unfairness. We have also ignored the long-term evolution of the system where users, intelligent and sometimes irrational, anticipate and react to the allocation at any given time, which further influences the long term efficiency and fairness of the system. Sharpened connections with other axiomatic theories in information, economics, and politics are thus called for.

VII. Acknowledgement

We are grateful to David Kao and Ashu Sabharwal for collaborating on an earlier version of the axioms and fairness-efficiency unification [40], to Rob Calderbank, Sanjeev Kulkarni, Xiaojun Lin, Linda Ness, Jennifer Rexford, and Kevin Tang for helpful comments, and to Paschalis Tsiaflakis for data on the spectrum management examples.

Appendix

A. Numerical Illustrations in Networking

This paper was originally motivated by problems of resource allocation in communication networks. We use classical problems in congestion control, routing, power control, and spectrum management to numerically illustrate some of the key points in the paper, including the use of generalization of α-fairness, fairness-efficiency tradeoff, fairness during transient behavior of iterative algorithms, and the impact of user-specific weights.

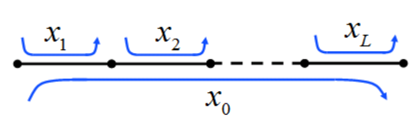

Fig. 12. A linear network with L links and $n = L + 1$ flows. All links have the same capacity of 1 unit.

1) Congestion Control: TCP congestion control in the Internet has been studied via a utility maximization model since [7], [20], with an extensive research literature including papers focusing on α-fairness. Overviews can be found in e.g., [21], [22]. The typical optimization problem modeling congestion control is as follows:

$$\max_{\mathbf{x}} \quad \sum_{i=1}^{n} U_i(x_i) \tag{84}$$
$$\text{s.t.} \quad \mathbf{x} \in \mathcal{R}$$

where U is a utility function and $\mathcal{R}$ is a set of all feasible resource allocation vectors. We consider the classical example of a linear network with L links, indexed by $l = 1, \ldots, L$, and $n = L + 1$ flows, indexed by $i = 0, 1, \ldots, L$, shown in Figure 12. Flow $i = 0$ goes through all the links and sources $i \geq 1$ go through links $l = i$. All links have the same capacity of 1 unit. We denote x_i to be the rate of flow i. We will illustrate two points: how a given $\mathbf{x}$ can be evaluated by different fairness measures, and how F acting as the objective function broadens the trade-off between efficiency and fairness.

The first example compares different fairness measures for a given resource allocation $\mathbf{x}$ in the linear network with $L = 5$:

$$\mathbf{x} = \left[\frac{1}{3}, \frac{2}{3}, \frac{2}{3}, \frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right]. \tag{85}$$

We illustrate the unified representation $F_{\beta,\lambda}(\mathbf{x})$ in (82), which includes symmetric and asymmetric fairness measures developed in Sections II and IV as subclasses. Figure 13 plots fairness $F_{\beta,\lambda}(\mathbf{x})$ against β for $1/\lambda = 0, 1, 2$. Larger λ puts more emphasis on fairness relative to efficiency. For asymmetric fairness, user specific weights are given by $q_0 = 0.5$ and $q_i = 0.1$ for $i \geq 1$.

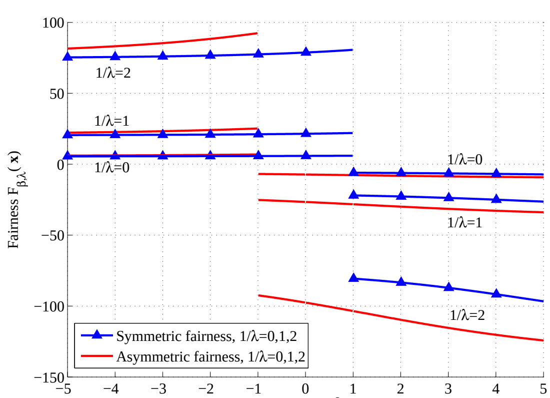

Fig. 13. Fairness $F_{\beta,\lambda}(\mathbf{x})$ derived in (82) in Section V, for $\beta \in [-4, 4]$ and $1/\lambda = 0, 1, 2$. Both Symmetric and asymmetric fairness measures are monotonically increasing on $\beta < -1$ and decreasing on $\beta > 1$.

We observe that symmetric fairness measures change sign at $\beta = 1$, while asymmetric fairness measures change sign at $\beta = -1$. Both symmetric and asymmetric fairness measures are monotonically increasing on $\beta < -1$ and decreasing on $\beta > 1$. For $-1 < \beta < 1$, they have different signs and monotonicities. As in Theorem 13 in Section V, fairness $F_{\beta,\lambda}(\mathbf{x})$ can be factorized as the product of a fairness component (i.e., $f_\beta(\mathbf{x})$ or $f_\beta(\mathbf{x}, \mathbf{q})$) and an efficiency component $(\sum_i x_i)^{\frac{1}{\lambda}}$. For fixed resource allocation $\mathbf{x}$, parameter β determines the value of the fairness component, while $1/\lambda$ decides how $F_{\beta,\lambda}(\mathbf{x})$ scales as efficiency $\sum_i x_i$ increases.

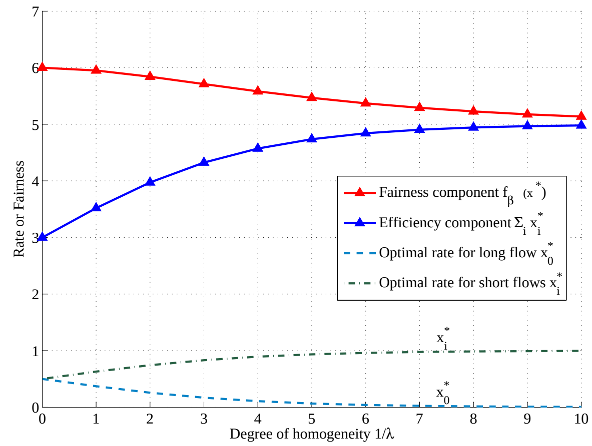

Fig. 14. Optimal rate allocations, their fairness components, and efficiency components, for congestion control problem (86) with different degrees of homogeneity $1/\lambda$ of the fairness measure. As $1/\lambda$ grows, efficiency component $\sum_i x_i^*$ increases and skews the optimal rate allocation $\mathbf{x}^*$ away from an equal allocation, while fairness component $f_\beta(\mathbf{x}^*)$ is monotonically decreasing.

Next, we consider a congestion control problem in linear network, which is formulated as a maximization of symmetric fairness $F_{\beta,\lambda}(\mathbf{x})$, a generalization of α-fair utility function, under link capacity constraints:

$$\max_{\mathbf{x}} \quad F_{\beta,\lambda}(\mathbf{x}) = f_\beta(\mathbf{x}) \cdot \left(\sum_i x_i\right)^{\frac{1}{\lambda}} \tag{86}$$
$$\text{s.t.} \quad x_0 + x_i \leq 1, \text{ for } i = 1, \ldots, L.$$

For $\beta \geq 0$ and $1/\lambda = \beta/(1-\beta)$, the optimal rate allocation maximizing (86) achieves α-fairness for $\alpha = \beta$. For $1/\lambda \geq \beta/(1-\beta)$, problem (86) is concave after a logarithm change of objective function, i.e., $\log F_{\beta,\lambda}(\mathbf{x})$. We fix $\beta = 1/2$ and solve (86) for different values of $1/\lambda$. Figure 14 plots optimal rate allocations $\mathbf{x}^*$ (in terms of x_0^* and x_i^* for $i \geq 1$), their fairness components $f_\beta(\mathbf{x}^*)$, and their efficiency components $\sum_i x_i^*$, all against $1/\lambda$. As $1/\lambda$ grows, efficiency component $\sum_i x_i^*$ increases and skews the optimal rate allocation away from the equal allocation: the long flow x_0 in the linear network gets penalized, while short flows x_i are favored. At the same time, fairness component of the objective function decreases. These observations illustrate the results in Section II-E and in Section V: degree of homogeneity $1/\lambda$ serves as a tradeoff factor between fairness and efficiency in congestion control. Different values of $1/\lambda$ result in various source algorithms in dual-based TCP congestion protocols [21], [22].

2) Multi-Commodity Flow Routing: We consider one classical type of routing problems: maximum multi-commodity flow on an undirected graph $G = (V, E)$, with given capacity $c(u, v)$ on each edge $(u, v) \in E$. A multi-commodity flow instance on G is a set of ordered pairs of vertices (s_i, d_i) for $i = 1, 2, \ldots, n$. Each pair (s_i, d_i) represents a commodity

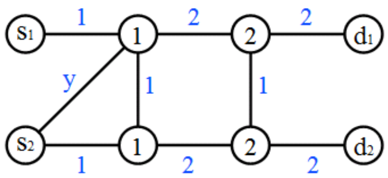

Fig. 15. An undirected graph $G = (V, E)$ with $n = 2$ commodities, (s_1, d_1) and (s_2, d_2). The number next to each edge represents its capacity. Due to the existence of edge $(s_2, 1)$, source s_2 can contribute more to maximum multi-commodity throughput than source s_1.

with source s_i and destination d_i. The objective of a maximum multi-commodity flow problem is to maximize the total throughput of flows traveling from the sources to the corresponding destinations, subject to edge capacity constraints. Let x_i be the throughput of commodity (s_i, d_i) and $\phi_i(u, v)$ denote its flow along edge $(u, v) \in E$. Maximum multi-commodity flow problem is the following one extensively studied since the 1950s:

$$\begin{aligned} \max_{\mathbf{x},\phi} \quad & \sum_{i=1}^{n} x_i \\ \text{s.t.} \quad & \sum_{i=1}^{n} \phi_i(u, v) \le c(u, v), \ \forall (u, v) \in E \\ & \phi_i(u, v) = -\phi_i(v, u), \ \forall (u, v) \in E \\ & \sum_{u \in V} \phi_i(u, v) = 0, \ \text{if } v \neq s_i, d_i \\ & \sum_{u \in V}^{n} \phi_i(s_i, v) = \sum_{u \in V}^{n} \phi_i(u, d_i) = x_i. \ \forall i \end{aligned} \tag{87}$$

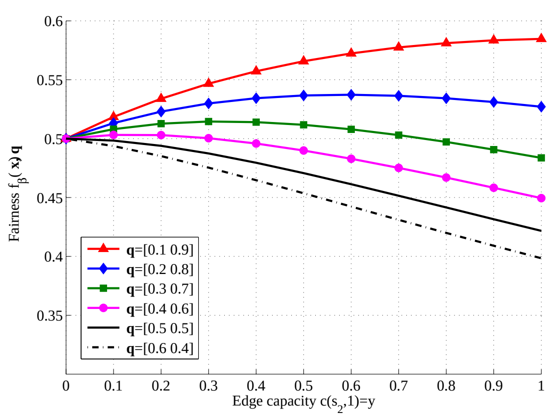

Fig. 16. Fairness of the maximum multi-commodity flow solutions against edge capacity $c(s_2, 1) = y \in [0, 1]$. Asymmetric fairness measures $f_\beta(\mathbf{x}, \mathbf{q})$ in (65) are employed, with parameters $r = -1$, $\beta = -2$, and commodity specific weights $q_1 = 1 - q_2$ and $q_2 = 0.9, 0.8, 0.7, 0.6, 0.5, 0.4$. Increasing y improves fairness only if the weight q_2 assigned to commodity 2 is large enough. For each choice of weights, the point that achieves maximum fairness (over choices of y) are marked by triangles.

For the graph with $n = 2$ commodities shown in Figure 15, we vary capacity y from 0 to 1 on edge $(s_2, 1)$, to create an asymmetry between the two commodities. We solve the maximum multi-commodity flow problem (87) and evaluate fairness of the resulting solutions $\mathbf{x} = [x_1, x_2]$, using asymmetric fairness measures $f_\beta(\mathbf{x}, \mathbf{q})$ derived in (65) in Section IV with $r = -1$, $\beta = -2$, and commodity specific weights $q_1 = 1 - q_2$ and $q_2 = 0.9, 0.8, 0.7, 0.6, 0.5, 0.4$. The results are plotted in Figure 16.

Due to the existence of edge $(s_2, 1)$, source s_2 contributes more to maximum multi-commodity throughput than source s_1. Figure 16 shows that this contribution improves fairness only if the weight q_2 assigned to commodity 2 is large enough. When $q_1 \ge q_2$, increasing capacity y always reduces fairness, an objectionable outcome. For $y \in [0, 1]$, we derive the solution to the maximum commodity flow problem (87) as $\mathbf{x} = [1, \ 1 + y]$. Using Corollary 24 in Section IV, in order to have a monotonic relationship between fairness $f_\beta(\mathbf{x}, \mathbf{q})$ and capacity $y \in [0, t]$, we need to choose weights $[q_1, q_2]$ satisfying

$$\frac{q_2}{q_1} \ge (1 + t)^{1 + r\beta}. \tag{88}$$

This closed-form bound quantifies the argument that sources contributing more to maximum multi-commodity throughput should be given higher weights, if we want their contribution to have a positive impact on fairness. Larger the range of capacity y for which this needs to hold, larger the ratio q_2/q_1 needs to be.

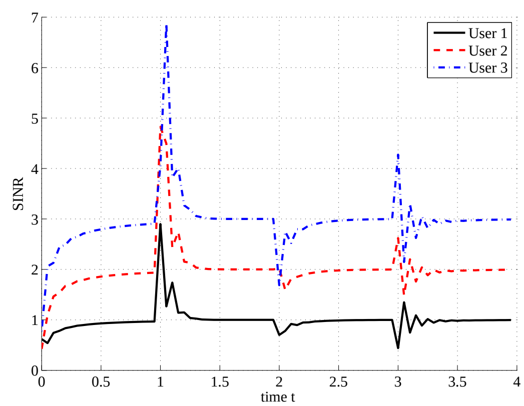

Fig. 17. Evolution of users' SINR over time t, using the distributed power control algorithm [19]. Transmit power is updated according to (90) every 0.05 second. Channel matrix G is updated every 1 second.

3) Power Control: Transmit power control is a key design challenge in wireless cellular networks since the 1990s, with a vast literature on optimization problem formulations and algorithms, e.g., as surveyed in [23]. We use different fairness measures to evaluate a wireless power control system. Consider a wireless cellular network with n users, each served by a different logical link from the mobile device to the base station. Let p_j be the transmit power of user j and G_{ij} denote the channel gain from the transmitter of logical channel j to the receiver of logical channel i. Therefore, interference caused from user j to user i has power $G_{ij} \cdot p_j$. Signal-To-Interference-

Noise-Ratio (SINR) for user i's link is defined as:

$$\gamma_i = \frac{G_{ii} \cdot p_i}{\sum_{j \neq i} G_{ij} \cdot p_j + \sigma^2}, \text{ for } i = 1, \ldots, n, \quad (89)$$

where σ^2 denotes noise power on this link. Much of the study on cellular network power control started in 1992 with a series of results that minimize total transmit power $\sum_i p_i$ for a fixed, feasible SINR target γ^*.

We consider the widely adopted distributed power control algorithm [19], which iteratively updates user's transmit power to achieve SINR target γ^*. In this iterative algorithm with each iteration indexed by integer k, each user only needs to monitor its individual received SINR and then update its transmit power by

$$p_i[k] = \frac{\gamma_i^*}{\gamma[k-1]} p_i[k-1], \text{ for } i = 1, \ldots, n, \quad (90)$$

independently and possibly asynchronously. Intuitively, each user i increases its power when its SINR is below target γ_i^* and decreases it otherwise. We run this algorithm, for $n = 3$ and a channel matrix G draw from a i.i.d. Gaussian distribution with 0 mean and variance 1. The channels are changed every second due to variations of the wireless propagation environment. We choose SINR target $\gamma^* = [1, 2, 3]$ and update transmit power according to (90) every 0.05 second. The resulting evolution of users' SINR over time t is illustrated in Figure 17.

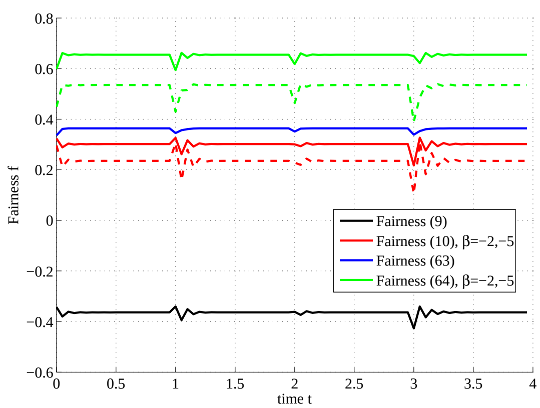

Fig. 18. Evaluation of a wireless power control system, which runs a iterative power update algorithm for a fixed and feasible SINR assignment [19], using different fairness measures. Fairness for system snapshots are plotted against time t. As $|\beta|$ increases, fairness value becomes more sensitive to resource changes during system transit states.

First, we take a snapshot of the network at each time t and evaluate fairness of the instantaneous SINR assignment, i.e., $f(\gamma)$ and $f(\gamma, \mathbf{q})$. We choose $r = -1$ and $\beta = -2, -5$ for fairness measures (10) and (66), generated by power functions. For asymmetric fairness measures (65) and (66), we choose weight $\mathbf{q} = [1/6, 1/3, 1/2]$ and $\mathbf{q} = [0.03, 0.22, 0.75]$ in (65) and (66) respectively, so that fairness $f(\gamma, \mathbf{q})$ are maximized at target SINR using Corollary 23.

As shown in Figure 18, at time $t = 1, 2, 3$, fairness of the system fluctuates due to the fluctuation of channel matrix G. Fairness measures (10) and (66) with $\beta = -5$ are more sensitive during transit states than that of $\beta = -2$, while fairness measures (9) and (65) with $\beta = 0$ are least sensitive. This illustrates the analysis in Section II-D that smaller parameter β suggests a stricter notion of fairness: fairness value becomes more sensitive to resource change. If we choose user-specific weights in asymmetric fairness measures so that fairness is maximized at target SINR, the distributed power control algorithm can be viewed as maximizing (asymmetric) fairness of the network.

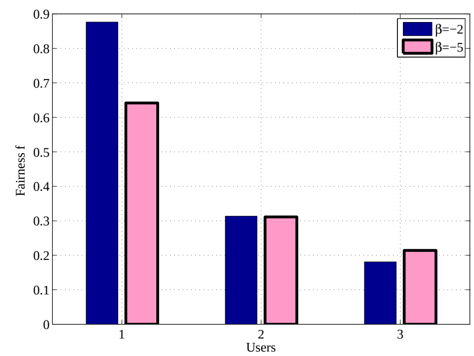

Fig. 19. Use $\mathcal{F}(\gamma_{i,t\in[0,T]})$ in (39) to quantify fairness of users' SINR distributions over time. We choose $r = 1$ and compute fairness $\mathcal{F}$ for users $i = 1, 2, 3$ for $\beta = -2$ and $\beta = -5$. User 1's SINR distribution over time is most fair.

Another approach to analyze fairness values aggregated over time is to start with the time series of users' SINR and quantify fairness of SINR distributions over time. Let $\gamma_{i,t\in[0,T]}$ be user i's SINR distribution over time $[0, T]$. We use fairness measure (39) in Section III to evaluate $\gamma_{i,t\in[0,T]}$:

$$\mathcal{F}(\gamma_{i,t\in[0,T]}) = \text{sign}(r(1-\beta r)) \cdot \left[\frac{1}{T} \int_{t=0}^{T} (\gamma_{i,t}/\mu)^{1-\beta r} \, dt \right]^{\frac{1}{\beta}}.$$

We choose $r = 1$ and compute fairness $\mathcal{F}$ for users $i = 1, 2, 3$ for two different choices of parameters: $\beta = -2$ and $\beta = -5$. Fairness measure $\mathcal{F}$ for this range is generalized Jain's index with continuous-dimensional resource distributions. Fairness value for different users' SINR distributions are shown in Figure 19. In this case, what matters is how stable (small magnitude of fluctuation) is γ during the transient stage after channel change. User 1's SINR is most stable over the transient stage in Figure 17. Therefore, it chieves the maximum fairness of $\mathcal{F} = 0.89$ for $\beta = -2$ among the three users in the network. User 3's SINR is most fluctuating over time. It is most unfair with $\mathcal{F} = 0.18$. Although $\beta = -2$ and $\beta = -5$ gives different fairness values, the fairness ordering of users remains unchanged.

4) Rate-Power Tradeoff in Spectrum Management: When power control takes place over multiple frequency bands, it is often called spectrum management. It arises in both Orthogonal Frequency-Division Multiplexing (OFDM) wireless networks and Digital-Subscriber-Line (DSL) broadband access networks. In the past decade, it has become one of the most effective solutions in engineering next generation DSL [24].

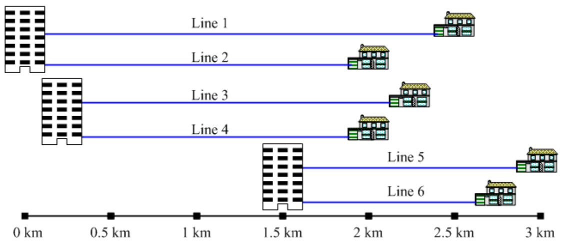

Fig. 20. A network consisting of $n = 6$ DSL users (i.e., lines). Near-far crosstalk occurs when a user (e.g., line 6) enjoying a good channel close to the receiver interferes with the received signal of a user (e.g., line 4) further away having a worse channel. Crosstalk may be mitigated by power control over the frequency bands, i.e., spectrum management.

We consider a network consisting of n interfering DSL users (i.e., lines), depicted in Figure 20, with a standard synchronous discrete multi-tone (DMT) modulation of K tones (i.e., frequency bands). Let p_i^k be the transmit power of user i on tone k and $p_i = \sum_k p_i^k$ be his total transmit power. We denote $\mathbf{G}^k$ to be an $n \times n$ channel gain matrix, whose (i, j)'th element is the channel gain from transmitter modem i to receiver modem j on tome k. The diagonal elements of $\mathbf{G}^k$ are the direct channels and the off-diagonal elements are the crosstalk channels. The data rate of user i on tone k is defined as:

$$b_i^k = \psi_i^k(\mathbf{G}^k, \mathbf{p}^k) \triangleq \log\left(1 + \frac{1}{\Gamma}\frac{G_{ii}^k \cdot p_i^k}{\sum_{j\neq i} G_{ij}^k \cdot p_j^k + \sigma_i^k}\right), \quad (91)$$

where $\psi_i^k(\cdot)$ is a data rate function for user i on tone k, σ_i^k denotes the corresponding noise power, and Γ is the SNR-gap to capacity, which is a function of bit error rate, coding gain, and noise margin. Let η be the DMT symbol rate. Total achievable data rate of user i is given by

$$R_i = \eta \cdot \sum_{k=1}^{K} b_i^k = \eta \cdot \sum_{k=1}^{K} \psi_i^k(\mathbf{G}^k, \mathbf{p}^k) \quad (92)$$

The goal of DSL spectrum management is to allocate each user's transmit powers over K tones to optimize certain design objectives and satisfy certain constraints. One of the recent spectrum management challenges is to strike a proper tradeoff between power-consumption and total rate of the network. In order to reduce power consumed, rates achieved by the users need to be decreased, and decreased by different amounts for different users. Since rates are interference-limited as in (91) and location of users put them into different interference environment (i.e., different G_{ij}^k), an interesting question is to quantify fairness of rate-reduction. In [25], the following optimization problem is formulated and solved to enforce a certain notion of fairness:

$$\begin{aligned} \max_{\{p_i^k, R_i\}} \quad & \sum_{i=1}^{n} R_i \\ \text{s.t.} \quad & p_i = \sum_{k=1}^{K} p_i^k \le p_{i,\max}, \quad \text{for } i = 1, \dots, n \\ & R_i = \eta \cdot \sum_{k=1}^{K} \psi_i^k(\mathbf{G}^k, \mathbf{p}^k), \quad \text{for } i = 1, \dots, n \\ & \frac{R_i}{R_i^o} / \frac{p_i}{p_{i,\max}} = \theta, \quad \text{for } i = 1, \dots, n \end{aligned} \quad (93)$$

where $[R_1^o, R_2^o, \dots, R_n^o]$ is a given rate allocation vector somewhere on the boundary of feasible rate region, $p_{i,\max}$ is the given transmit power budget of user i, and $\theta > 0$ is a positive constant.

The notion of fairness, enforced by the third constraint in optimization problem (93), is based on the following ad hoc argument: the ratio of each user's percentage data rate decrement to her percentage transmit power decrement should be the same across the users.

We use the algorithm and data in [25] to illustrate further uses of the axiomatically constructed fairness measures in the paper. The topology is as shown in Figure 20. The channel model in [25] has the following parameters: Twisted pair lines have a diameter of 0.5 mm. Transmit power budget is $p_{i,\max}$ =20.4 dBm for each user. SNR gap Γ is 12.9 dB, corresponding to a coding gain of 3 dB, a noise margin of 6 dB, and a target symbol error probability of 10^{-7}. Tone spacing is 4.3125 kHz. DMT symbol rate is η =4 kHz.

There are three points we will illustrate. First, the "proportional" style fairness in (93) can be quantified using fairness measure f. Second, using f we can quantify the tradeoff between reducing power and the fairness of doing so. Third, asymmetric fairness measures are usefuls in weighing different users based on her location.

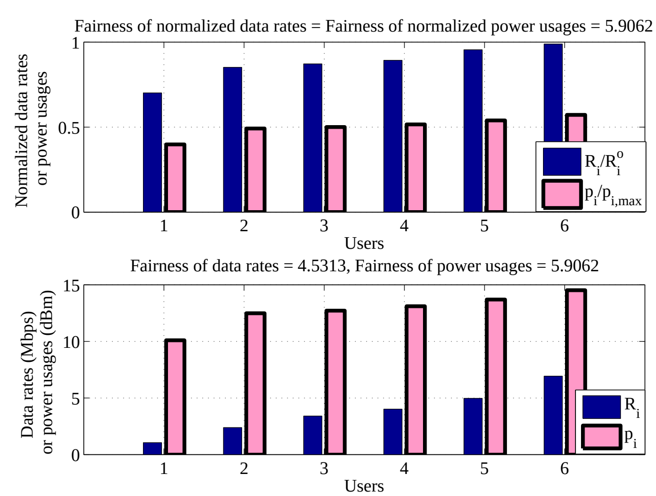

Fig. 21. Plot optimal (normalized and unnormalized) data rates and power usages, solving (93). Normalized data rates and power usages achieve a large fairness value $f = 5.9062$, while fairness of data rates is smaller, due to an near-far effect in the network.

First, we solve problem (93) with $\rho = 1.75$ and plot the distributions of the optimal data rate and transmit power across different users in Figure 21. The ratio of normalized data rates to normalized power usage is proportional for all users, due to the third constraint in (93). The optimal solution to (93) achieves the same fairness for the rate reduction vector and the power reduction vector , i.e.,

$$f\left(\frac{R_1}{R_1^o},\ldots,\frac{R_1}{R_1^o}\right) = f\left(\frac{p_1}{p_{1,\max}},\ldots,\frac{p_n}{p_{n,\max}}\right) = 5.9062 \quad (94)$$

where we choose f to be the symmetric fairness measure (10) in Section II with $\beta = -2$. This fairness value is only 0.1 (i.e., 2%) away from maximum fairness $f = 6$. It is (almost) $\beta = -2$ fair. We also observe that fairness of data rate distribution is much smaller:

$$f(R_1,\ldots,R_n) = 4.5313. \quad (95)$$

This is due to the near far near-far effect in the DSL network.

Next, we solve (93) with different constant $\rho > 0$ and let the resulting normalized total power $\sum_i p_i / \sum_i p_{i,\max}$ in the optimal solution vary between 0.2 and 1. Figure 22 plots the corresponding fairness of (normalized and unnormalized) data rate distributions in each optimal solution, using the same symmetric fairness measure with $\beta = -2$. It illustrates a tradeoff between fairness and the amount of power reduction: fairness improves as normalized total power $\sum_i p_i / \sum_i p_{i,\max}$ increases from 0.2 (i.e., an 80% power reduction) to 1 (i.e., no power reduction). Again, data rates are less fair than normalized data rates due to the near-far effect.

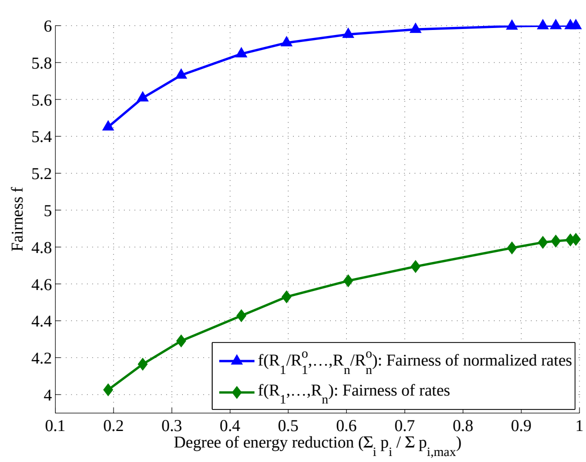

Fig. 22. Tradeoff between reducing power and fairness in doing so, i.e., fairness of (normalized and unnormalized) data rate distributions vs. normalized total power $\sum_i p_i / \sum_i p_{i,\max}$. Fairness improves as normalized total power increases from 0.2 (80% power reduction) to 1 (no power reduction).

Finally, consider R_i^*, user i's optimal data rate in problem (93). It is larger for the more advantageously-located users. If we consider the family of asymmetric fairness measures (66) in Section IV, according to Corollary 23, $\{R_i^*\}$ also maximizes the following asymmetric fairness measure $f([R_1,\ldots,R_n],[q_1,\ldots,q_n])$, where a higher weight is assigned to a user with larger optimal data rate, with the following choice of user specific weights:

$$q_i = (R_i^*)^{r\beta+1}, \quad \text{for } i = 1,\ldots,n. \quad (96)$$

B. Related Axiomatic Theories and Fairness Studies

Forcing metrics of important notions to be the inevitable consequence of simple statements is a practice commonly found in the mathematical way of thinking. Here we compare and contrast with several axiomatic theories in information, economics, and political philosophy, some of which are quantitative. We show that fairness measures in this paper is a generalization of Renyi entropy, and a special case of Lorenz curve ordering. Axioms here are contrasted with those in Nash bargaining solution and in Shapley value. We also survey several studies of fairness in other disciplines.

1) Renyi entropy: Renyi entropy is a family of functionals quantifying the uncertainty or randomness of generalized probability distributions, developed in 1960 [26]. Renyi entropy is derived from a set of five axioms as follows:

1) Symmetry.
2) Continuity.
3) Normalization.
4) Additivity.
5) Mean-value property.

Comparing Renyi's axioms to ours, we notice that the Axiom of Continuity and the Axiom of Normalization are equivalent to our Axiom of Continuity and Axiom of Homogeneity, respectively. The Axiom of Symmetry becomes Corollary of Symmetry in Section II, due to our Axiom of Partition. Next, the Axiom of Additivity and Axiom of Mean-value are replaced by our Axiom of Partition. More precisely, the Axiom of Additivity can be directly derived from our Axiom of Partition as in (6) for direct product. The Axiom of Mean-value, which states that the entropy of the union of two incomplete distributions is the weighted mean value of the entropies of the two distributions, plays a role similar to our recursive construction of f in the Axiom of Partition. The key difference is that in our Axiom of Partition and its implications, fairness of a vector is constructed by the product of a "global" fairness across partitions, and a weighted mean of "local" fairness within each partition. The Axiom of Saturation and the Axiom of Starvation are unique to our system.

For $\beta \le 1$, it is straightforward to verify that our fairness measure satisfying the Axioms $1''$-$5''$ in Section V includes Renyi entropy as a special case. If we choose $\beta = 1-\alpha$ and $\lambda = -1$ in our fairness measure $F_{\beta,\lambda}$, then $F_{\beta,\lambda}$ equals to the exponential of Renyi entropy with parameter α, i.e.,

$$F_{\beta=1-\alpha,\lambda=-1}(\mathbf{x}) = e^{h_\alpha(\mathbf{x})}. \quad (97)$$

Our proof for the existence and uniqueness of fairness measures in Theorems 1-3 in Appendix C extends those of Renyi entropy [26] and Shannon entropy [27]: A key step in the proofs is to show that any function satisfying the (fairness or entropy) axioms is an additive number-theoretical function [26]. For fairness we show that $\log f(\mathbf{1}_n)$ is an additive number-theoretical function over integer $n \geq 1$, i.e.,

$$\log \frac{f(\mathbf{1}_{mn})}{f(1)} = \log \frac{f(\mathbf{1}_n)}{f(1)} + \log \frac{f(\mathbf{1}_m)}{f(1)}, \quad (98)$$

$$\lim_{n\to\infty} \log \frac{f(\mathbf{1}_{n+1})}{f(1)} - \frac{\log f(\mathbf{1}_n)}{f(1)} = 0. \quad (99)$$

where $\mathbf{1}_n$ is an all-one resource allocation vector of length n. Conditions (98) and (99) follow from the Axiom of Partition and the Axiom of Saturation, respectively. Using the result in [26] we have $\log f(\mathbf{1}_n) = r \log n$, which, together with the Axiom of Homogeneity, leads to an explicit expression of fairness measure $f(\mathbf{x})$ for the case of two users and resource vectors of rational numbers. The result is then generalized to resource vectors of real numbers using the Axiom of Continuity, and to arbitrary number of users based on Axiom of Partition. The Axiom of Saturation and the Axiom of Starvation are then used to ensure uniqueness of fairness measures.

Renyi entropy, a generalization of Shannon entropy, has been further extended by others since the 1960s, e.g., smooth Renyi entropy for information sources [28] and quantum Renyi entropy [29]. This work is a different generalization of Renyi entropy: first along asymmetric users with user-specific values and then along parameter plane of general λ and $\beta > 1$. Renyi entropy and α-fairness, which are now special cases of our axiomatic system (for $\beta \leq 1$ and $\beta \geq 0$ respectively), are largely "complementary" along β axis, but also with a range of overlap for $\beta \leq 1$.

2) Lorenz Curve: Schur-concavity of fairness measures proven in Theorem 3 for discrete-dimensional inputs and in Theorem 12 for continuous-dimensional inputs is a critical property for justifying fairness measures by establishing an ordering on the set of Lorenz curves [18]. Let $P_{\mathbf{x}}(y)$ be the cumulative distribution of a resource allocation $\mathbf{x}$. Its Lorenz curve L_x, defined by

$$L_{\mathbf{x}}(u) = \frac{1}{\mu} \cdot \int_{\{P_{\mathbf{x}}(y)\leq u\}} y dP_{\mathbf{x}}(y), \quad (100)$$

is a graphical representation of the distribution of $\mathbf{x}$, and has used to characterize the social welfare distributions and relative income differences in economics.

In 2001, an axiomatic characterization of Lorenz curve orderings was proposed based on a set of four axioms [18]:

1) Order. (The ordering is transitive and complete.)
2) Dominance. (The ordering is Shur-concavity.)
3) Continuity.
4) Independence.

It is shown that a Lorenz curve orderings $L_{\mathbf{x}} \succeq L_{\mathbf{y}}$ satisfies the four axioms above if and only if there exists a continuous and non-increasing real function $p(u)$ defined on the unit interval, such that

$$L_{\mathbf{x}} \succeq L_{\mathbf{y}} \;\Leftrightarrow\; \int_0^1 p(u) dL_{\mathbf{x}}(u) \geq \int_0^1 p(u) dL_{\mathbf{y}}(u) \quad (101)$$

Our fairness measures have an equivalent representation in Section III.1, i.e.,

$$\mathcal{F} = \text{sign}(r(1-\beta r)) \cdot \left[\int_{u=0}^1 \left(\frac{P_{\mathbf{x}}^{-1}(u)}{\mu} \right)^{-\beta r} dL_{\mathbf{x}}(u) \right]^{\frac{1}{\beta}} \quad (102)$$

Comparing (102) to (101), it is easy to see that our fairness measure defines a Lorenz curve ordering for all $\beta \geq 0$ and $r \geq 0$, by setting

$$p(u) = \left(\frac{P_{\mathbf{x}}^{-1}(u)}{\mu} \right)^{-\beta r}, \quad (103)$$

which is continuous and non-increasing over $u \in [0, 1]$. Therefore, the Lorenz curve ordering defined by our fairness measure satisfies the four axioms in [18], including Shur-concavity property proven in Section II and Section III.

3) Cooperative Economic Theories: A number of theories have been developed to study the collective decisions of groups. Many of these theories have also been uniquely associated with sets of axioms [30]. While applications of these theories range from political theory to voting methods, here we focus on two well-known axiomatic constructions: the the Nash bargaining solution in 1950 [31] and the Shapley value in 1953 [32].

The Nash bargaining solution was developed from a set of four axioms:

1) Invariance to affine transformation.
2) Pareto optimality.
3) Independence of irrelevant alternatives (IIA).
4) Symmetry.

Symmetry is shown as a corollary in our theory, and Pareto optimality not imposed as an axiom given our focus on fairness, but rather used as a criterion in balancing fairness and efficiency in Section V. Nash's axiom of IIA contributes most to his uniqueness result, and it is also often considered as a value statement. It has been shown by many others, that replacing the "local-closedness" axiom of IIA with other value statements may result in solution classes different from the bargaining solution. Given a feasible region of individual utilities, the Nash bargaining solution is also equivalent to a maximization of the proportional fairness utility function.

Another well-known concept in economic study of groups is the Shapley value [32]. In a coalition game, individual decide whether or not to form coalitions in order to increase the maximum utility of the group, while ensuring that their share of the group utility is maximized. The Shapley value concerns an operator that maps the structure of the game, to a set of allocations of utility in the overall group. Given a coalition game, four axioms uniquely define the Shapley value:

1) Pareto Optimality.
2) Symmetry.
3) Dummy.
4) Additivity.

Again, Pareto optimality is included as an axiom. Although the structure of a coalition game is different than a simple division of resources, some parallels are apparent. For instance, the Dummy axiom refers to a scenario where an individual does not increase any groups utility by joining it. In this case, the Shapley value assigns a utility of zero to that individual. This is not imposed in our theory as an axiom, but is similar to Corollary 4, where an inactive user has no impact on fairness value. Shapley's axiom of additivity provides a method of building up a single coalition game with potentially many individuals, from smaller games, which may have only two players. This is similar to our Axiom if Partition, wherein the fairness measure is recursively constructed from the fairness attained by subsets of the overall allocation.

In both Nash bargaining solution and Shapley value, efficiency is an axiom in defining the solution. In contrast, the first family of fairness measures, f in Section II, is confined to homogeneous functions of degree zero. One might suspect that it is possible to extend our axiomatic structure to the optimization theoretic fairness approach by relaxing the Axiom of Homogeneity to homogeneous functions of arbitrary degree. This is indeed the case as developed in Section V.

The distinction between fairness measures with local-closedness axiom and those with decomposability axiom has been discussed in the Conclusion. It remains to be seen if the "fair" operating point as defined by NBS/Shapley can be incorporated into a user reaction model together with our fairness measure.

4) User Reaction Models and Fair Taxation: In [83], [84], a dynamic system model was developed to study the long-term impact and equilibrium behaviors of users planning economic activities in anticipation of tax policies, defined as the median voter action. Fairness is an important part of the model, and is defined as the effort-based (rather than corruption-based component in the sum of total income) and measured by variance. Again, there is some ideal point of operation that is fair based on the source of income.

A more general notion of fairness may be inserted in the control dynamic system. What would be the equilibrium properties (existence, number, income distribution, efficiency) if our fairness measure is used, with $\mathbf{q}$ modeling the user's relative position in effort-based vs. corruption-based income? This is also related to optimal taxation from a fairness point of view.

The standard theory of optimal taxation maximizes a social welfare function subject to a range of constraints [85]. The objective function is almost always the sum utilities, and often linear utilities. We briefly discuss some implications of our fairness theory on taxation. There are many more questions to be addressed beyond the initial study below.

We may quantify the impact of tax on fairness by analyzing fairness of pre-tax, post-tax resource allocations, and taxes themselves, using fairness measures derived from Axioms 1-5. Let $\mathbf{x} = [x_1, \ldots, x_n]$ be a pre-tax resource allocation vector and $c : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be a tax policy. The corresponding post-tax resource allocation $[x_1 - c(x_1), \ldots, x_n - c(x_n)]$ will be denoted by $\mathbf{x} - c(\mathbf{x})$ with a slight abuse of notation. We consider a tax policy c satisfying the following two assumptions:

(a) Zero tax are collected from inactive users, i.e., $c(0) = 0$.
(b) Users' resource ranking remains unchanged, i.e., $x_i - c(x_i) > x_j - c(x_j)$ for all $x_i > x_j$.

It is easy to see that these two assumptions guarantee positivity of post-tax allocation, i.e., $x_i - c(x_i) > 0$ for all $x_i \in \mathbb{R}_+$.

For a given fairness measure f satisfying Axioms 1-5, we say that a tax policy is *fair* if $f(\mathbf{x} - c(\mathbf{x})) \geq f(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n_+$ and *unfair* if $f(\mathbf{x} - c(\mathbf{x})) \leq f(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n_+$. We derive the following corollaries for characterizing fairness of pre-tax and post-tax resource allocations that hold for all f_β

Corollary 23: (Fair tax policy.) A tax policy c is fair if and only if $c(z)/z$ is non-decreasing for $z \in \mathbb{R}_+$.

Corollary 24: (Unfair tax policy.) A tax policy c is unfair if and only if $c(z)/z$ is non-increasing for $z \in \mathbb{R}_+$.

Corollary 23 shows that when fairness of a tax policy c is evaluated using f_{beta}, it coincides with the intuition that fair tax must be progressive. It is straightforward to verify that any convex and non-decreasing function c satisfies Corollary 23 and is therefore fair. On the other hand, Corollary 24 shows that any tax policy c less progressive than a linear taxation is unfair. A special case is collecting a constant tax from all users. More precisely, if a constant amount $t > 0$ of the resource is subtracted from each user (i.e., $x_i - t$ for all i), the resulting fairness decreases, i.e., $f(\mathbf{x} - t \cdot \mathbf{1}_n) \leq f(\mathbf{x})$, where $t > 0$ is small enough such that $x_i - t$ is non-negative.

Corollary 25: (Fairness of c.) A tax policy c is fair only if $f(c(\mathbf{x})) \leq f(\mathbf{x})$. Similarly, c is unfair only if $f(c(\mathbf{x})) \geq f(\mathbf{x})$.

It is necessary to have the taxation vector as less fair than the pre-tax resource vector. But this condition is only necessary. $f(c(\mathbf{x})) \leq f(\mathbf{x})$ does not always imply that tax policy c is fair. For instance, we have $f(c(\mathbf{x})) < f(\mathbf{x})$ and $f(\mathbf{x} - c(\mathbf{x})) \leq f(\mathbf{x})$, when $\mathbf{x} = [6, 3, 2]$, $c(\mathbf{x}) = 3, 1, 1$, and $\beta = 2.5$.

Corollary 26: (Ordering of tax policies.) Let c_1 and c_2 be two tax policies. We say that c_1 is more fair than c_2 if $f(\mathbf{x} - c_1(\mathbf{x})) \geq f(\mathbf{x} - c_2(\mathbf{x}))$ for all $\mathbf{x} \in \mathbb{R}^n_+$. This holds if $c_1(\mathbf{x})/c_2(\mathbf{x})$ is non-decreasing, plus any of the following three sufficient conditions:

(a) c_1 is fair and c_2 is unfair.
(b) Both c_1 and c_2 are fair and $c_1(z) \geq c_2(z)$ for all $z \in \mathbb{R}_+$.
(c) Both c_1 and c_2 are unfair and $c_1(z) \leq c_2(z)$ for all $z \in \mathbb{R}_+$.

This corollary establishes a complete ordering of tax policies with respect to fairness of post-tax allocations. Given that tax policy c_1 is more progressive than tax policy c_2, three sufficient conditions for c_1 to be more fair than c_2 are derived. When two tax policies are both fair or both unfair, their impact on fairness is affected by their magnitudes.

5) Ultimatum Game, Cake-Cutting, and Fair Division: In Ultimatum Game [81], player A divides a resource into two parts, one part for herself and the other for player B. Player B can then choose to accept the division, or reject it, in which case neither player receives any resource. Without a prior knowledge about player B's reaction, player A may divide anywhere between $[0.5,\ 0.5]$ and $[1,\ 0]$. Running this game as a social experiment in different cultures have lead to debates about the exact implications of the results on people's perception on fairness: how fair does it take for player B's perception, and player A's guess of that, to accept player A's division? This has also been contrasted with the perception of fairness in the related Dictator Game [75], [76], where player B has no option but to accept the division by player A.

A classic generalization of Ultimatum Game that has received increasing attention in the past decade is the cake-cutting problem. As reviewed in books [74], [80], the cake is a measure space, and each player uses a countably-additive, non-atomic measure to evaluate different parts of the cake.

Among the work studying the cake-cutting problem, e.g., [79], the primary focus has been on two criteria: efficiency (Pareto optimality) and fairness (envy-freeness). Achievability results for 4 users or more are still challenging.

As reviewed in [82], cake-cutting can be generalized to "fair division" problem, a typical example being the joint problem of room-assignment and rent-division. Each user contributes to the resource pool, and as explained in [77], there are many surprises in the solution structures if envy-freeness is again taken as the definition of fairness.

Fairness in cake-cutting and fair-division is traditionally defined as envy-freeness. It is a binary summary based on each individual's local evaluation: the allocation of cake is fair if no user wants to trade her piece with another piece. In [78], this restrictive viewpoint on fairness is expanded to include proportional allocation of the left-over piece after each user gets $1/n$th of the cake (in her own evaluation). It is shown that Pareto optimality and proportional sense of fairness may not be compatible for 3 players or more.

Equipped with the axiomatically constructed fairness measures f_β, we see that some questions are in order. Is there a 2-person, 1-cut procedure (with the help of a referee and possibly not strategy proof) that is envy-free, Pareto optimal, and fairness-value-y achieving ? For 3-person, is there a procedure that is Pareto optimal and fairness-value-y achieving? For what β in $f_\beta(x)$ would envy-free be compatible with fairness? More generally, if envy-freeness is replaced by fairness in the sense of achieving a target value of $f_\beta(U_1(x_1), \ldots, U_n(x_n))$, where U_i is user i's own preference/utility function, will existence results and constructive procedures in efficient and fair cutting of cake be more readily achieved?

6) Rawls' Theory of Justice and Distributive Fairness: In 20th century political philosophy, the work of Rawls on fairness has been arguably the most influential and provocative since its original publication in 1971 [33]. Starting with the "original position" behind the "veil of ignorance", the arguments posed by Rawls are based on two fundamental principles (i.e., axioms stated in English), as stated in his 2001 restatement [55]:

1) "Each person is to have an equal right to the most extensive scheme of equal basic liberties compatible with a similar scheme of liberties for others."
2) "Social and economic inequalities should be arranged so that they are both (a) to the greatest benefit of the least advantaged persons, and (b) attached to offices and positions open to all under conditions of equality of opportunity."

The first principle governs the distribution of liberties and has priority over the second principle. One may interpret it as a principle of distributive fairness in allocating limited resources among users. It can now be captured as a theorem (rather than an axiom) that says any fairness measures satisfying Axioms 1-5 will satisfy the following property: adding an equal amount of resource to each user will only increase fairness value, i.e.,

$$f_\beta\left(\mathbf{x}+c\cdot\mathbf{1}_n,\mathbf{q}\right) \geq f_\beta\left(\mathbf{x},\mathbf{q}\right), \text{ for } q_i=\frac{1}{n}, \ \forall c\geq 0, \ \forall\beta,$$

where equal weights $q_i = \frac{1}{n}$ can be viewed as a quantification of "equal right" in the first principle of Rawls' theory.

First part of Rawls' second principle concerns with the distribution of opportunity while the second part is the celebrated "difference principle": an approach different from strict egalitarianism (since it is on the absolute value of the least advantaged user rather than the relative value) and utilitarianism (when narrowly interpreted where the utility function does not capture fairness). Now it is axiomatically constructed as a special case of a continuum of generalized notions trading-off fairness with efficiency. Our results can be viewed as an axiomatic quantification of Rawls' theory, especially as a mathematical extension and representation of the equal rights principle and the difference principle. This is best illustrated by annotating Rawls' own graph in Figure 23. The point representing the difference principle is now seen as the consequence of concatenating two steps of pushing to the *extremum* on both β and λ: $\beta\to\infty$, and λ as large as possible while retaining Pareto-efficiency, i.e., $\lambda\to\bar{\lambda}$.

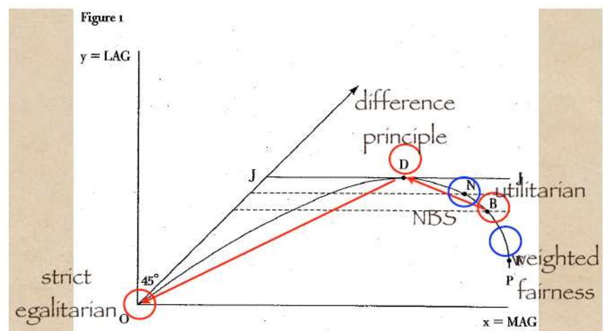

Fig. 23. Annotated version of the only quantitative graph in Rawls [55] on p.62. The colored circles and arrows are added by the PI. Since the efficiency frontier between the More Advantaged Group (MAG) user and the Less Advantaged Group (LAG) user does not pass through 45 degree line, strict egalitarian point is the origin. Difference principle generates a point D on the efficient frontier, whereas utility maximization generates point B. Nash Bargaining Solution point N may lie between D and B. α-fair utility function, being the point on the efficient frontier closest to the 45 degree line, will produce D as well. Fairness function $F_{\beta,\lambda}(\mathbf{x},\mathbf{q})$ can generate any point on the efficient frontier when λ is restricted to be less than the threshold for Pareto efficiency to retain.

Among the other main branches of fairness in philosophy, the utilitarian or welfare-based principle [56], [58] advocates utility maximization, but the form of utility function can be broad, indeed, as broad as $F_{\beta,\lambda}$ that captures both fairness and efficiency. Still, specification and implementation can be challenging, there as well as in our fairness theory: e.g., user-specific utility modeling, inter-user utility "conversion rate", computational challenge of solving constrained optimization, and potential non-uniqueness of optimal solution.

Desert-based principles [60], [62] provide helpful angles for us to model the user-weights $\mathbf{q}$, which can be based on contribution, effort, or compensation. Libertarian principles offer other viewpoints on fairness. One of these is the 3-part entitlement theory by Nozick [63], complemented by the "Lockean proviso" [64], where an exclusive acquisition of a resource is just if there is "enough and as good left in common for others" afterwards. The theory becomes less clear in case of scarce resources where the constraint set does not allow such a scenario [65]

7) Normative Economics and Welfare Theory: Rawls' theory also has intricate interactions with normative economics, where many results are analytic in nature [68]. In addition to stochastic dominance, Arrow's Impossibility Theorem, and the cooperation game theories of Nash and of Shapley, there are several major branches [57], [69], [71]. Another set is the ethical axioms of transfers. For example, the Pigou-Dalton principle states that inequality decreases via Robin Hood operation that does not reverse relative ranking. The principle of proportional transfers states that what the donor gives and the beneficiary receives should be proportional to their initial positions.

Bergson-Samuelson social welfare function [66], [72] $W(U_1(\mathbf{x}),\ldots,U_n(\mathbf{x}))$ aims at enabling complete and consistent social welfare judgement on top of individual preference-based utility functions U_i. Its existence raises the question of comparing $\sum_i w_i x_i$ with $f_\beta(x)\cdot(\sum_i x_i)^\lambda$ in our theory.

Kolm's theory of fair allocation [70] uses the criterion of equity as "no-envy", and it is well- known that competitive equilibrium with equal budget is the only Pareto-efficient and envy-free allocation if preferences are sufficiently diverse and form a continuum [59]. It is unclear if it we can replace envy-freeness with a system-wide fairness criterion like $f_\beta(x)$.

8) Connections with Sociology and Psychology: Inequality Indices. Quantifying inequality/injustice/unfairness using individual, local measures, has been pursued in sociology. For example, Jasso in 1980 [46] advocated justice evaluation index as log of the ratio between actual allocation and "just" allocation. Allocation can be done either in quantity or in quality (in which case ranking quantifies the quality allocation). Many properties were derived in theory and experimented with in data about income distribution in different countries [48]. In particular, probability distribution of the index is induced by the probability distribution of the allocation. This index is derived based on two principles and three laws in the paper, including "equal allocation maximizes justice" and "aggregate justice is arithmetic mean of individual ones". These are special cases of the corollaries or axioms in our development, and the index can indeed be viewed as a special case of f_β: a logarithm function with a constant offset.

In [47], two other injustice indexes, JI1 and JI2, were developed based on the above. One interesting feature is that JI1 differentiates between under-reward and over-reward as two types of injustice. Another useful feature is the decomposition of the total amount of perceived injustice into injustice due to scarcity and injustice due to inequality. This is similar to the two components on efficiency and fairness in $F_{\beta,\lambda}(\mathbf{x})$. In [49], inequality between two individuals is decomposed into inequality within the groups they belong to and inequality within each group. This decomposition is generalized in our recursive construction of fairness measures implied by Axiom of Partition.

They are further unified with Atkinson's measure of inequality [57]: 1 minus the ratio of geometric mean and arithmetic mean. At the heart of these indices is the approach of taking combinations of arithmetic and geometric means of an allocation to quantify the its spread. In this view, it is not surprising that Atkinson's index is also a special case of our fairness measures, as shown in Figure 11. The list of axioms behind Atkinson's index, however, is both longer and more restrictive.

One may be able to view actual allocation as the optimizer of $F_{\beta,\lambda}(\mathbf{x})$ and the just allocation as say NBS. Then the justice evaluation by each individual thus calculated may be substituted into dynamic system describing the allocation and reaction process.

Procedural Fairness. Fairness may be as much about the outcome as about the procedure. Procedural fairness has been extensively studied in legal system and has generated much interest in psychology and organization studies since the 1960s. It is sometimes coupled with user-weight assignment in distributive fairness, e.g., in [50] equity means proportionality between resource allocation and individual contributions.

Substantial amount of work in procedural fairness initiated with Thibaut and Walker's book [54] in 1975, with specific focus on dispute resolution via a third party, as well as psychological and cultural underpinnings of preferences for various abjudication procedures. Methodological issues in laboratory studies since then have been surveyed in Lind and Tyler's book [51].

In two publications [52], [53] on general theory of procedural fairness, Leventhal advocated six criteria: consistency, bias suppression, accuracy of information, correctability, representativeness, and ethicality. Some of these elements are clearly context and culture-dependent. Several interesting questions are in order. Is it possible to combine distributive and procedural justice to reach general assessment of fairness? When will procedural justice concerns override distributive justice assessment? Can procedural fairness measures be axiomatically constructed out of the key criteria and maximized using practical protocols?

C. Collection of Proofs

1) Proof of Theorems 1, 2, and 3: We prove that fairness measures satisfying Axioms 1-5 can only be generated by logarithm and power generator functions. Then, for each choice of logarithm and power generator function, a fairness measure exist and is uniquely given by either (9) or (10). A fairness measure is a sequence of mapping $\{f^n:\mathbb{R}^n_+\to\mathbb{R},\forall n\ge 1\}$. To simplify notations, we suppress n in fairness measures and denote them by f. It is easy to verify that if $f>0$ satisfies Axioms 1-4, so does $-f<0$. We first assume that $f>0$ during this proof and then determine sign to fairness measures from Axioms 5.

Our proof starts with a lemma, which states that fairness achieved by equal-resource allocations $\mathbf{1}_n$ is independent of the choice of g. This Lemma will be used repeatedly in the proof.

Lemma 2: From Axioms 1–5, fairness achieved by equal-resource allocations $\mathbf{1}_n$ is given by

$$f(\mathbf{1}_n)=n^r\cdot f(1),\quad \forall\ n\ge 1, \tag{104}$$

where $r\in\mathbb{R}$ is a constant exponent.

Proof:

We first prove that $f(\mathbf{1}_{mn}) \cdot f(1) = f(\mathbf{1}_n) \cdot f(\mathbf{1}_m)$ for all $m, n \in \mathbb{Z}_+$. For $m = 2$ and arbitrary $n \in \mathbb{Z}_+$, we have

$$\begin{aligned} \tfrac{f(\mathbf{1}_{2n})}{f(\mathbf{1}_2)} &= \frac{f(\mathbf{1}_{2n})}{f(n \cdot \mathbf{1}_2)} \\ &= g^{-1}\left(\sum_{i=1}^{2} s_i \cdot g\left(\frac{f(\mathbf{1}_n)}{f(n)}\right)\right) \\ &= g^{-1}\left(g\left(\frac{f(\mathbf{1}_n)}{f(1)}\right)\right) \\ &= \frac{f(\mathbf{1}_n)}{f(1)}, \end{aligned} \tag{105}$$

where the first step uses homogeneity in Axiom 2, the second step uses Axiom 4 with $\mathbf{x} = [\mathbf{1}_n, \mathbf{1}_n]$ and $\mathbf{y} = [n, n]$, and the third step uses $\sum_i s_i = 1$.

We use induction and assume that for $m \le k$ and $n \in \mathbb{Z}_+$, we have

$$f(\mathbf{1}_{mn}) \cdot f(1) = f(\mathbf{1}_n) \cdot f(\mathbf{1}_m) \tag{106}$$

For $m = k + 1$ and $n \in \mathbb{Z}_+$, we derive that

$$\begin{aligned} \tfrac{f(\mathbf{1}_{(k+1)\cdot n})}{f(\mathbf{1}_{k+1})} &= \frac{f(\mathbf{1}_{(k+1)\cdot n})}{f(n \cdot \mathbf{1}_{k+1})} \\ &= g^{-1}\left(s_1 \cdot g\left(\frac{f(\mathbf{1}_{kn})}{f(n \cdot \mathbf{1}_k)}\right) + s_2 \cdot g\left(\frac{f(\mathbf{1}_n)}{f(n)}\right)\right) \\ &= g^{-1}\left((s_1 + s_2) \cdot g\left(\frac{f(\mathbf{1}_n)}{f(1)}\right)\right) \\ &= \frac{f(\mathbf{1}_n)}{f(1)}, \end{aligned} \tag{107}$$

where the second step uses Axiom 4 with $\mathbf{x} = [\mathbf{1}_{kn}, \mathbf{1}_n]$ and $\mathbf{y} = [n \cdot \mathbf{1}_k, n]$, and the third step uses (106) with $m = k$. This means that $f(\mathbf{1}_{mn}) \cdot f(1) = f(\mathbf{1}_n) \cdot f(\mathbf{1}_m)$ holds for $m \le k+1$ and $n \in \mathbb{Z}_+$. Therefore, it holds for all $m, n \in \mathbb{Z}_+$.

From (106), we can show that $\log f(\mathbf{1}_{mn})/f(1)$ is an additive number-theoretical function [34], i.e.,

$$\log \frac{f(\mathbf{1}_{mn})}{f(1)} = \log \frac{f(\mathbf{1}_n)}{f(1)} + \log \frac{f(\mathbf{1}_m)}{f(1)} \tag{108}$$

Further, from Axiom 3, we derive

$$\lim_{n \to \infty} \left[\log \frac{f(\mathbf{1}_{n+1})}{f(1)} - \log \frac{f(\mathbf{1}_n)}{f(1)}\right] = 0 \tag{109}$$

Using the result in [34], (108) and (109) implies that $\log f(\mathbf{1}_n)/f(1)$ must be a logarithmic function. We have

$$\log \frac{f(\mathbf{1}_n)}{f(1)} = r \log n, \tag{110}$$

where r is a real constant. This is (104) after taking an exponential on both sides. ■

(Uniqueness of logarithm and power generator functions.) We first show that logarithm and power generator functions are necessary. In other words, no other fairness can possibly satisfy Axioms 1-5 with weight s_i in (15).

We consider arbitrary positive $X_1 + X_2 = 1$, which can be written as the ρ-th power of rational numbers, i.e., $X_1 = (a_1/b_1)^\rho$ and $X_2 = (a_2/b_2)^\rho$, for some positive integers a_1, b_1, a_2, b_2. Let u, v be two positive integers and $K = (u/v)^{1/\rho}$ be a rational number to be used as a short-hand notation later. We choose the sum resource of the two sub-systems to be $\mathcal{W}_1 = a_1 b_2 v$ and $\mathcal{W}_2 = a_2 b_1 v$, so that

$$X_i = \frac{\mathcal{W}_i^\rho}{\mathcal{W}_1^\rho + \mathcal{W}_2^\rho}, \ i = 1, 2. \tag{111}$$

Consider the following two resource allocation vectors: For resource allocation vector $\mathbf{x}$, sub-system 1 has $t = b_1 b_2 u$ users, each with resource $\mathcal{W}_1/t$; sub-system 2 has $t = b_1 b_2 u$ users, each with resource $\mathcal{W}_2/t$. In other words, we have

$$\mathbf{x}^1 = [\underbrace{\frac{\mathcal{W}_1}{t}, \ldots, \frac{\mathcal{W}_1}{t}}_{t \text{ times}}] \text{ and } \mathbf{x}^2 = [\underbrace{\frac{\mathcal{W}_2}{t}, \ldots, \frac{\mathcal{W}_2}{t}}_{t \text{ times}}] \tag{112}$$

For resource allocation vector $\mathbf{y}$, sub-system 1 has $a_1 b_2 v$ users, each with resource 1; sub-system 2 has $a_2 b_1 v$ users, each with resource 1. This means

$$\mathbf{y}^1 = [\underbrace{1, \ldots, 1}_{a_1 b_2 v \text{ times}}] \text{ and } \mathbf{y}^2 = [\underbrace{1, \ldots, 1}_{a_2 b_1 v \text{ times}}] \tag{113}$$

The sum resource of the two sub-systems $\mathcal{W}_i = v t a_i / b_i$ is in-variant. Let $\xi_i = \mathcal{W}_i^\rho / (\mathcal{W}_1^\rho + \mathcal{W}_2^\rho)$, for $i = 1, 2$, be an auxiliary variable. It follows from Axiom 4 and (104) in Lemma 2 that

$$\begin{aligned} &\frac{f(\mathbf{y}^1, \mathbf{y}^2)}{f(\mathbf{x}^1, \mathbf{x}^2)} \\ &= g^{-1}\left(\sum_{i=1}^{2} s_i \cdot g\left(\frac{f(\mathbf{y}^i)}{f(\mathbf{x}^i)}\right)\right) \\ &= g^{-1}\left(\xi_1 \cdot g\left(\frac{a_1^r b_2^r v^r}{t^r}\right) + \xi_2 \cdot g\left(\frac{a_2^r b_1^r v^r}{t^r}\right)\right) \\ &= g^{-1}\left(X_1 \cdot g\left(\frac{a_1^r v^r}{b_1^r u^r}\right) + X_2 \cdot g\left(\frac{a_2^r v^r}{b_2^r u^r}\right)\right) \\ &= g^{-1}\left(X_1 \cdot g\left((X_1/K)^{\frac{r}{\rho}}\right) + X_2 \cdot g\left((X_2/K)^{\frac{r}{\rho}}\right)\right) \end{aligned} \tag{114}$$

where the second step uses Axiom 2 and (104) in Lemma 2, i.e.,

$$f(\mathbf{x}^i) = f\left(\frac{\mathcal{W}_i}{t} \cdot \mathbf{1}_t\right) = f(\mathbf{1}_t) = t^r \tag{115}$$

and

$$f(\mathbf{y}^1) = f\left(\frac{\mathcal{W}_1}{t} \cdot \mathbf{1}_{a_1 b_2 v}\right) = (a_1 b_2 v)^r \tag{116}$$

$$f(\mathbf{y}^2) = f\left(\frac{\mathcal{W}_2}{t} \cdot \mathbf{1}_{a_2 b_1 v}\right) = (a_2 b_1 v)^r \tag{117}$$

Next, we apply (5) in Lemma 1 to $f(\mathbf{x}^1, \mathbf{x}^2)$ and obtain

$$\begin{aligned} &f(\mathbf{x}^1, \mathbf{x}^2) \\ &= f\left(\frac{\mathcal{W}_1 + \mathcal{W}_1}{t} \cdot \mathbf{1}_t\right) \cdot g^{-1}\left(g\left(f\left(\frac{\mathcal{W}_1}{t}, \frac{\mathcal{W}_2}{t}\right)\right)\right) \\ &= f(\mathbf{1}_t) \cdot f(\mathcal{W}_1, \mathcal{W}_2) \\ &= t^r \cdot f(\mathcal{W}_1, \mathcal{W}_2) \\ &= t^r \cdot f(X_1^{\frac{1}{\rho}}, X_2^{\frac{1}{\rho}}) \end{aligned} \tag{118}$$

where the entire system is partitioned into t sub-systems[3], each with a resource allocation vector $[\mathcal{W}_1/t, \mathcal{W}_1/t]$. The last step uses (111). Again, using Axiom 2 and (5) in Lemma 1, we derive

$$f(\mathbf{y}^1, \mathbf{y}^2) = f\left(\mathbf{1}_{a_1 b_2 v + a_2 b_1 v}\right) = (a_1 b_2 v + a_2 b_1 v)^r \quad (119)$$

Combining (114), (118), and (119), we derive

$$\begin{aligned} & \frac{(X_1^{\frac{1}{\rho}} + X_2^{\frac{1}{\rho}})^r}{f(X_1^{\frac{1}{\rho}}, X_2^{\frac{1}{\rho}})} K^{-\frac{r}{\rho}} \\ & = \frac{(a_1 b_2 v + a_2 b_1 v)^r}{f(X_1^{\frac{1}{\rho}}, X_2^{\frac{1}{\rho}}) \cdot t^r} \\ & = g^{-1}\left(\sum_{i=1}^{2} X_i \cdot g\left(\left(\frac{X_i}{K}\right)^{\frac{r}{\rho}}\right)\right) \end{aligned} \quad (120)$$

For different $K > 0$, (120) further implies that

$$\begin{aligned} & g^{-1}\left(\sum_{i=1}^{2} X_i \cdot g\left(\left(\frac{X_i}{K}\right)^{\frac{r}{\rho}}\right)\right) \\ & = K^{\frac{r}{\rho}} \cdot g^{-1}\left(\sum_{i=1}^{2} X_i \cdot g\left(\left(\frac{X_i}{K}\right)^{\frac{r}{\rho}}\right)\right) \end{aligned} \quad (121)$$

for rational $X_1, X_2, K > 0$ and arbitrary ρ, r. Due to Axiom 1, it is easy to show that (121) also holds for real $X_1, X_2, K > 0$. Without loss of generality, we assume that $g(1) = 0$, because we may replace $g(y)$ by $g(y) - g(1)$ in (121), and the equation still holds.

It then follows from Theorem 4.3 in [35] that

$$g(X \cdot K) = p(K)g(X) + q(K) \quad (122)$$

where $p(K), q(K)$ are functions of K for real $X, K > 0$. Using $g(1) = 0$ and choosing $X = 1$, we derive

$$g(K) = q(K) \quad (123)$$

Substituting it into (122), we find that

$$g(X \cdot K) = p(K)g(X) + g(K) \quad (124)$$

for all positive $X, K > 0$. Similarly, we have

$$g(X \cdot K) = p(X)g(K) + g(X). \quad (125)$$

Combining (124) and (125) gives

$$\frac{p(X) - 1}{g(X)} = \frac{p(K) - 1}{g(K)} \quad (126)$$

for all $X, K > 0$[4]. Each of these functions must reduces to a constant c, so that $p(X) - 1 = c \cdot g(X)$. It then follows from (124) that

$$g(X \cdot K) = cg(X) \cdot g(K) + g(X) + g(K). \quad (127)$$

[3]Although Axiom 4 is defined for partitioning into 2 subsystems, it can be extended to partitioning into t identical sub-systems, using the same induction in proof of Lemma 2.

[4]It holds only for $X \neq 1$ and $K \neq 1$. However, since (127) is true if $X = 1$ or $K = 1$, the exception is irrelevant for proving (126).

To determine possible generator functions g, we need to consider two cases. (a) If $c = 0$, (127) reduces to $g(X \cdot K) = g(X) + g(K)$, whose solution is

$$g(y) = z \cdot \log(y) \quad (128)$$

where $z \neq 0$ is an arbitrary constant. (b) If $c \neq 0$, we put $c \cdot g(y) + 1 = h(y)$ and (127) reduces to $h(X \cdot K) = h(X) \cdot h(K)$, whose general solution is $h(y) = y^\beta$. Hence, we have

$$g(y) = \frac{y^\beta - 1}{c} \quad (129)$$

where $\beta \in \mathbb{R}$ is a constant exponent.

(Existence of fairness measures for logarithm generator.) We derive fairness measure for the logarithm generator function in (128) and show that it satisfies Axioms 1-5. By choosing $K = u/v = 1$ and a change of variable $x_i = X_i^{\frac{1}{\rho}} = a_i/b_i$ in (120), we plug in the logarithm generator (128) and find

$$\begin{aligned} f(x_1, x_2) & = \frac{(x_1 + x_2)^r}{g^{-1}\left(\sum_{i=1}^{2} \frac{x_i^\rho}{x_1^\rho + x_2^\rho} \cdot g(x_i^r)\right)} \\ & = \prod_{i=1}^{2} \left(\frac{x_i}{x_1 + x_2}\right)^{-r \frac{x_i^\rho}{x_1^\rho + x_2^\rho}} \end{aligned} \quad (131)$$

To generate fairness measure for $n > 2$ users, we use induction and assume that for $n = k$ users, we have

$$f(x_1, \ldots, x_k) = \prod_{i=1}^{k} \left(\frac{x_i}{\sum_{j=1}^{k} x_j}\right)^{-r \frac{x_i^\rho}{\sum_j x_j^\rho}} \quad (132)$$

which holds for $n = k = 2$. Let $u_k = \sum_{i=1}^{k} x_i$. For $n = k+1$, we apply (5) in Lemma 1 and derive (130), where the second step uses Axiom 2 with $f(x_{k+1}) = f(1) = 1$. Comparing the last step of (130) to (132) with $n = k+1$, induction of fairness measure from $n = k$ to $n = k + 1$ holds if and only if

$$u_k^\rho = \left(\sum_{i=1}^{k} x_i\right)^\rho = \sum_{i=1}^{k} x_i^\rho. \quad (133)$$

Therefore, we must have $\rho = 1$. Fairness measure generated by logarithm generator in (128) can only have the following form

$$f(x_1, \ldots, x_k) = \prod_{i=1}^{k} \left(\frac{x_i}{\sum_{j=1}^{k} x_j}\right)^{-r \frac{x_i}{\sum_j x_j}} \quad (134)$$

which is (9) in Theorem 3.

The proof so far assumes positive fairness measure $f > 0$. To determine the sign of f, we add $\mathrm{sign}(r)$ in (134), apply Axiom 5, and find that for $r > 0$

$$f(0, 1) = 1 < 2^r = f(\frac{1}{2}, \frac{1}{2}), \quad (135)$$

and for $r < 0$:

$$f(0, 1) = -1 < -2^r = f(\frac{1}{2}, \frac{1}{2}). \quad (136)$$

Therefore, Axiom 5 is satisfied. We derive the fairness measure in (9).

$$f(x_1,\ldots,x_k,x_{k+1})$$
$$= f\left(\sum_{i=1}^{k} x_i, x_{k+1}\right)\cdot [f(x_1,\ldots,x_k)]^{\frac{(\sum_{i=1}^k x_i)^\rho}{(\sum_{i=1}^k x_i)^\rho + x_{k+1}^\rho}} \cdot [f(x_{k+1})]^{\frac{x_{k+1}^\rho}{(\sum_{i=1}^k x_i)^\rho + x_{k+1}^\rho}}$$
$$= f\left(\sum_{i=1}^{k} x_i, x_{k+1}\right)\cdot [f(x_1,\ldots,x_k)]^{\frac{u_k^\rho}{u_k^\rho + x_{k+1}^\rho}} \cdot 1$$
$$= \left(\frac{\sum_{i=1}^k x_i}{\sum_{i=1}^{k+1} x_i}\right)^{\frac{-r u_k^\rho}{u_k^\rho + x_{k+1}^\rho}} \cdot \left(\frac{x_{k+1}}{\sum_{i=1}^{k+1} x_i}\right)^{\frac{-r x_{k+1}^\rho}{u_k^\rho + x_{k+1}^\rho}} \cdot \prod_{i=1}^{k}\left(\frac{x_i}{\sum_{j=1}^k x_j}\right)^{-r\frac{x_i^\rho}{\sum_j x_j^\rho}\cdot\frac{u_k^\rho}{u_k^\rho + x_{k+1}^\rho}} \quad (130)$$

It remains to prove that fairness measure (134) satisfies Axioms 1-4 with arbitrary $r \in \mathbb{R}$. It is easy to see that Axioms 1–3 are satisfied. To verify Axiom 4, we consider two arbitrary vectors $\mathbf{x} = [\mathbf{x}^1, \mathbf{x}^2]$ and $\mathbf{y} = [\mathbf{y}^1, \mathbf{y}^2]$ with $w(\mathbf{x}^i) = w(\mathbf{y}^i)$. Since generator g is logarithm function, we show that

$$\frac{f(\mathbf{x}^1,\mathbf{x}^2)}{f(\mathbf{y}^1,\mathbf{y}^2)} = \frac{g^{-1}\left(\sum_{i=1}^2 s_i\cdot g\left(f(\mathbf{x}^i)\right)\right)}{g^{-1}\left(\sum_{i=1}^2 s_i\cdot g\left(f(\mathbf{y}^i)\right)\right)}$$
$$= \exp\left\{\sum_{i=1}^2 s_i\cdot\left[\log\left(f(\mathbf{x}^i)\right) - \log\left(f(\mathbf{y}^i)\right)\right]\right\}$$
$$= \exp\left\{\sum_{i=1}^2 s_i\cdot\left[\log\left(\frac{f(\mathbf{x}^i)}{f(\mathbf{y}^i)}\right)\right]\right\}$$
$$= g^{-1}\left(\sum_{i=1}^2 s_i\cdot g\left(\frac{f(\mathbf{x}^i)}{f(\mathbf{y}^i)}\right)\right) \quad (137)$$

which establishes Axiom 4 for any partition. This completes the proof for logarithm generators.

(Existence of fairness measures for power generator.) We derive fairness measure for power generator function in (129) and show that it satisfies Axioms 1-5. By choosing $K = u/v = 1$ and a change of variable $x_i = X_i^{\frac{1}{\rho}} = a_i/b_i$ in (120), we plug in the power generator (129) and find

$$f(x_1,x_2) = \frac{(x_1+x_2)^r}{g^{-1}\left(\sum_{i=1}^2 \frac{x_i^\rho}{x_1^\rho + x_2^\rho}\cdot g(x_i^r)\right)}$$
$$= \frac{(x_1+x_2)^r\,(x_1^\rho + x_2^\rho)^{\frac{1}{\beta}}}{\left(x_1^{\rho+\beta r} + x_2^{\rho+\beta r}\right)^{\frac{1}{\beta}}}. \quad (138)$$

To derive the fairness measure for three users, we consider two different partitions of the resource allocation vector $[x_1,x_2,x_3]$ as $[x_1,x_2],[x_3]$ and $[x_1],[x_2,x_3]$. Using (5) in Lemma 1, we obtain two equivalent form of the fairness measure for $n = 3$ users in (139) and (140).

As in Axiom 4, the fairness measure is independent of partition. Hence, (139) and (140) should be equivalent for all $x_1,x_2,x_3 \geq 0$. Comparing the terms in (139) and (140), we must have $r = 0$ or $\rho + \beta r = 1$. Since $r = 0$ is a special case $\rho+\beta r = 1$ by choosing $\rho = 1$ and $\beta \neq 0$, we only need to consider a unified representation for $\rho + \beta r = 1$. We conclude that the fairness measure for $n = 3$ must have the following form

$$f(x_1,x_2,x_3) = \frac{\left(\sum_{i=1}^3 x_i^{1-\beta r}\right)^{\frac{1}{\beta}}}{\left(\sum_{i=1}^3 x_i\right)^{\frac{1}{\beta}-r}}, \quad (141)$$

where $r = \frac{1-\rho}{\beta}$ is a proper exponent. Let $u_k = \sum_{i=1}^k x_i$ be the sum of the first k elements in vector $[x_1,\ldots,x_k,x_{k+1}]$. Then, using (5) in Lemma 1 inductively, we obtain

$$f(x_1,\ldots,x_k,x_{k+1})$$
$$= f(\sum_{i=1}^k x_i, x_{k+1})\cdot g^{-1}\left(s_1 g\left(f(x_1,\ldots,x_k)\right) + s_2 g(1)\right)$$
$$= \frac{\left(u_k^{1-\beta r} + x_{k+1}^{1-\beta r}\right)^{\frac{1}{\beta}}}{(u_k + x_{k+1})^{\frac{1}{\beta}-r}}\cdot\left[\frac{u_k^\rho\frac{\sum_{i=1}^k x_i^{1-\beta r}}{u_k^{1-\beta r}} + x_{k+1}^\rho}{u_k^\rho + x_{k+1}^\rho}\right]^{\frac{1}{\beta}}$$
$$= \frac{\left(u_k^{1-\beta r} + x_{k+1}^{1-\beta r}\right)^{\frac{1}{\beta}}}{(u_k + x_{k+1})^{\frac{1}{\beta}-r}}\cdot\left[\frac{\sum_{i=1}^k x_i^{1-\beta r} + x_{k+1}^{1-\beta r}}{u_k^{1-\beta r} + x_{k+1}^{1-\beta r}}\right]^{\frac{1}{\beta}}$$
$$= \frac{\left(\sum_{i=1}^{k+1} x_i^{1-\beta r}\right)^{\frac{1}{\beta}}}{\left(\sum_{i=1}^{k+1} x_i\right)^{\frac{1}{\beta}-r}}, \quad (142)$$

where induction holds from $n = k$ to $n = k+1$. The last step of (142) is almost (10) in Theorem 3, except for a term $\mathrm{sign}(r(1-r\beta))$. The term is added in order to satisfy Axiom 5. To show this, we consider four possible cases: for $1-r\beta > 0$ and $r > 0$

$$f(0,1) = 1 < 2^r = f(\frac{1}{2},\frac{1}{2}), \quad (143)$$

and for $1-r\beta > 0$ and $r < 0$

$$f(0,1) = -1 < -2^r = f(\frac{1}{2},\frac{1}{2}). \quad (144)$$

When $1-r\beta < 0$, we have $r > 0 \Leftrightarrow \beta > 0$ and $r < 0 \Leftrightarrow \beta < 0$. We find that for $1-r\beta < 0$ and $r > 0$

$$f(0,1) = -\infty < -2^r = f(\frac{1}{2},\frac{1}{2}), \quad (145)$$

and for $1-r\beta > 0$ and $r < 0$

$$f(0,1) = 0 < 2^r = f(\frac{1}{2},\frac{1}{2}). \quad (146)$$

$$
\begin{aligned}
f(x_1,x_2,x_3) &= f(x_1+x_2,x_3)\cdot g^{-1}\left(s_1 g\left(f(x_1,x_2)\right)+s_2 g\left(1\right)\right)\\
&= \frac{(x_1+x_2+x_3)^r\left((x_1+x_2)^\rho+x_3^\rho\right)^{\frac{1}{\beta}}}{\left((x_1+x_2)^{\rho+\beta r}+x_3^{\rho+\beta r}\right)^{\frac{1}{\beta}}}\cdot\left[\frac{\frac{(x_1+x_2)^{\rho+\beta r}(x_1^\rho+x_2^\rho)}{x_1^{\rho+\beta r}+x_2^{\rho+\beta r}}+x_3^\rho}{(x_1+x_2)^\rho+x_3^\rho}\right]^{\frac{1}{\beta}}\\
&= \frac{(x_1+x_2+x_3)^r}{\left((x_1+x_2)^{\rho+\beta r}+x_3^{\rho+\beta r}\right)^{\frac{1}{\beta}}}\cdot\left[\frac{(x_1+x_2)^{\rho+\beta r}(x_1^\rho+x_2^\rho)}{x_1^{\rho+\beta r}+x_2^{\rho+\beta r}}+x_3^\rho\right]^{\frac{1}{\beta}} \qquad (139)
\end{aligned}
$$

$$
\begin{aligned}
f(x_1,x_2,x_3) &= f(x_1,x_2+x_3)\cdot g^{-1}\left(s_1 g\left(f(1)\right)+s_2 g\left(x_2,x_3\right)\right)\\
&= \frac{(x_1+x_2+x_3)^r\left(x_1^\rho+(x_2+x_3)^\rho\right)^{\frac{1}{\beta}}}{\left(x_1^{\rho+\beta r}+(x_2+x_3)^{\rho+\beta r}\right)^{\frac{1}{\beta}}}\cdot\left[\frac{\frac{(x_2+x_3)^{\rho+\beta r}(x_2^\rho+x_3^\rho)}{x_2^{\rho+\beta r}+x_3^{\rho+\beta r}}+x_1^\rho}{(x_2+x_3)^\rho+x_1^\rho}\right]^{\frac{1}{\beta}}\\
&= \frac{(x_1+x_2+x_3)^r}{\left(x_1^{\rho+\beta r}+(x_2+x_3)^{\rho+\beta r}\right)^{\frac{1}{\beta}}}\cdot\left[\frac{(x_2+x_3)^{\rho+\beta r}(x_2^\rho+x_3^\rho)}{x_2^{\rho+\beta r}+x_3^{\rho+\beta r}}+x_1^\rho\right]^{\frac{1}{\beta}} \qquad (140)
\end{aligned}
$$

It is straightforward to check that fairness measure in (142) satisfies Axioms 1-4. This completes the proof for power generators.

2) Proof of Corollary 1: We prove symmetry from Axioms 1-5 directly, without resorting to the fairness expression in (9) and (10). For $n = 2$ users, symmetry follows directly from (120), which implies

$$f(x_1,x_2) = f(x_2,x_1),\ \forall x_1,x_2 \ge 0. \qquad (147)$$

To start the induction, assume symmetry holds for n users. Let $\mathbf{x} = [x_1,\dots,x_n,x_{n+1}]$ be a resource allocation vector and $i_1,\dots,i_n,i_{n+1}$ be an arbitrary permutation of the indices $1,\dots,n,n+1$. When $i_{n+1}\neq 1$, applying (5) in Lemma 1, we derive (148).

Next, we can use (148) to show that

$$f(x_{i_1},\dots,x_{i_n},x_{i_{n+1}}) = f(x_1,\dots,x_n,x_{n+1}). \qquad (149)$$

When $i_{n+1}=1$, using the same technique, we have

$$
\begin{aligned}
f(x_{i_1},\dots,x_{i_n},x_{i_{n+1}}) &= f(x_{i_1},\dots,x_{i_{n+1}},x_{i_n})\\
&= f(x_1,\dots,x_n,x_{n+1}). \qquad (150)
\end{aligned}
$$

Then symmetry also holds for $n+1$ users.

3) Proof of Theorem 4: Because vector $\mathbf{x}$ is majorized by vector $\mathbf{y}$, if and only if, from $\mathbf{x}$ we can produce $\mathbf{y}$ by a finite sequence of Robin Hood operations [15], where we replace two elements x_i and $x_j < x_i$ with $x_i-\epsilon$ and $x_j+\epsilon$, respectively, for some $\epsilon \in (0, x_i - x_j)$, it is necessary and sufficient to show that such an Robin Hood operation always improves a fairness measure defined by Axioms 1–5.

Toward this end, we consider partitioning a resource allocation vector $\mathbf{x}$ of n users into two segments: $\mathbf{x}^1 = [x_i, x_j]$ and $\mathbf{x}^2 = [x_1,\dots,x_{i-1},x_{i+1},\dots,x_{j-1},x_{j+1},\dots,x_n]$. Let $\mathbf{y} = [\mathbf{y}^1,\mathbf{x}^2]$ where $\mathbf{y}^1 = [x_i-\epsilon, x_j+\epsilon]$ be the vector obtained from $\mathbf{x}^1$ by the Robin Hood operation. Using Axiom 4, we have

$$
\begin{aligned}
\frac{f(\mathbf{x})}{f(\mathbf{y})} &= \frac{f(\mathbf{x}^1,\mathbf{x}^2)}{f(\mathbf{y}^1,\mathbf{x}^2)}\\
&= g^{-1}\left(s_1\cdot g\left(\frac{f(\mathbf{x}^1)}{f(\mathbf{y}^1)}\right)+s_2\cdot g\left(\frac{f(\mathbf{x}^2)}{f(\mathbf{x}^2)}\right)\right)\\
&= g^{-1}\left(s_1\cdot g\left(\frac{f(x_i,x_j)}{f(x_i-\epsilon,x_j+\epsilon)}\right)+s_2\cdot 1\right)\\
&\le 1,
\end{aligned}
$$

where the last step follows form the monotonicity of g and the monotonicity of fairness measure with two-users in (120), which gives

$$f(x_i,x_j) \le f(x_i-\epsilon,x_j+\epsilon). \qquad (151)$$

Therefore, if $\mathbf{x}$ is majorized by $\mathbf{y}$, then we have $f(\mathbf{x}) \le f(\mathbf{y})$. The fairness measure is Schur-concave.

4) Proof of Corollary 2: The proof for Corollary 2 is immediate from Theorem 4, because among the vectors with the same sum of elements, one with the equal elements is the most majorizing vector, i.e., $\frac{1}{n}\cdot\mathbf{1}_n$ majorized all $\mathbf{x}$.

5) Proof of Corollary 3: This corollary is proven by Lemma 2, in the proof of Theorems 1, 2, and 3.

6) Proof of Corollary 4: Let $\mathbf{x}$ be an arbitrary resource allocation vector, $t>0$ be a positive number, and $z=\sum_i x_i$ be an axillary variable. From Axiom 1, we have

$$
\begin{aligned}
&f(\mathbf{x},\mathbf{0}_n)\\
&= \lim_{t\to\infty} f\left(\mathbf{x},\frac{1}{t}\mathbf{1}_n\right)\\
&= \lim_{t\to\infty} f\left(z,\frac{n}{t}\right)\cdot g^{-1}\left(\frac{z^\rho g\left(f(\mathbf{x})\right)}{z^\rho+\left(\frac{n}{t}\right)^\rho}+\frac{\left(\frac{n}{t}\right)^\rho g\left(f(\mathbf{1}_n)\right)}{z^\rho+\left(\frac{n}{t}\right)^\rho}\right)\\
&= \lim_{t\to\infty} f\left(\sum_i x_i,\frac{n}{t}\right)\cdot g^{-1}\left(g\left(f(\mathbf{x}_n)\right)\right)\\
&= f(\mathbf{x}),
\end{aligned}
$$

$$
\begin{aligned}
&f(x_{i_1},\ldots,x_{i_n},x_{i_{n+1}})\\
&= f\left(\sum_{j=1}^{n} x_{i_j}, x_{i_{n+1}}\right)\cdot g^{-1}\left(s_1\cdot g\left(f(x_{i_1},\ldots,x_{i_n})\right)+s_2 g\left(f(x_{i_{n+1}})\right)\right)\\
&= f\left(\sum_{j=1}^{n} x_{i_j}, x_{i_{n+1}}\right)\cdot g^{-1}\left(s_1\cdot g\left(f(x_1,\ldots,x_{i_{n+1}-1},x_{i_{n+1}+1},\ldots,x_{n+1})\right)+s_2 g\left(f(x_{i_{n+1}})\right)\right)\\
&= f(x_1,\ldots,x_{i_{n+1}-1},x_{i_{n+1}+1},\ldots,x_{n+1},x_{i_{n+1}})\\
&= f\left(\sum_{j=1}^{i_{n+1}-1} x_j, \sum_{j=i_{n+1}}^{n+1} x_j\right)\cdot g^{-1}\left(s_1\cdot g\left(f(x_1,\ldots,x_{i_{n+1}-1})\right)+s_2 g\left(f(x_{i_{n+1}+1},\ldots,x_{i_{n+1}})\right)\right)\\
&= f\left(\sum_{j=1}^{i_{n+1}-1} x_j, \sum_{j=i_{n+1}}^{n+1} x_j\right)\cdot g^{-1}\left(s_1\cdot g\left(f(x_1,\ldots,x_{i_{n+1}-1})\right)+s_2 g\left(f(x_{i_{n+1}},,\ldots,x_{n+1})\right)\right)\\
&= f(x_1,\ldots,x_{i_{n+1}-1},x_{i_{n+1}},\ldots,x_{n+1})
\end{aligned} \tag{148}
$$

where the third step follows from (5) in Lemma 1.

7) Proof of Corollaries 5 and 6: When $f<0$ is negative, it is easy to show that $f(\mathbf{x})\to-\infty$ if $x_i\to 0$. When $f>0$, suppose that k users are inactive. From (104) and Corollaries 1 and 3, we have

$$f(\mathbf{x})\le f(\mathbf{1}_{n-k})=n-k. \tag{152}$$

which gives $k\le n-f(\mathbf{x})$. Further, since the number of active users $n-k$ is upper bounded by $f(\mathbf{x})$, the maximum resource is lower bounded by $\sum_i x_i/f(\mathbf{x})$.

8) Proof of Corollary 7: Let $k(\mathbf{x})=\sum_{i=1}^{n}\left(\frac{x_i}{\sum_j x_j}\right)^{1-\beta}$ be an auxiliary function, such that

$$f(\mathbf{x})=\text{sign}(1-\beta)\cdot k^{\frac{1}{\beta}}(\mathbf{x}). \tag{153}$$

It is easy to check that fairness measures $f(\mathbf{x})$ in (9) and (10) are differentiable. We have

$$\frac{\partial f(\mathbf{x})}{\partial x_i}=\frac{1}{\beta}k^{\frac{1}{\beta}-1}(\mathbf{x})\cdot\frac{|1-\beta|}{(\sum_j x_j)^{1-\beta}}\left[x_i^{-\beta}-\frac{\sum_j x_j^{1-\beta}}{\sum_j x_j}\right]$$

Because $k(\mathbf{x})>0$ is positive, $\frac{\partial f(\mathbf{x})}{\partial x_i}$ has a single root at

$$x_i=\bar{x}=\left(\frac{\sum_j x_j}{\sum_j x_j^{1-\beta}}\right)^{\frac{1}{\beta}}. \tag{154}$$

It is straightforward to show that for any $\beta\ne 1$, we have

$$\frac{\partial f(\mathbf{x})}{\partial x_i}>0,\ \text{if}\ x_i>\bar{x}\quad\text{and}\quad\frac{\partial f(\mathbf{x})}{\partial x_i}<0,\ \text{if}\ x_i<\bar{x}$$

Therefore, when x_j $\forall j\ne i$ are fixed, $f(\mathbf{x})$ is maximized by $x_i=\bar{x}$.

9) Proof of Corollary 8: To derive an lower bound on $f(\mathbf{x})$ under the box constraints $x_{min}\le x_i\le x_{max}$ $\forall i$, we first argue that $f(\mathbf{x})$ is minimized only if users are assigned resource x_{min} or x_{max}. Using the box constraints and Corollary 8, we have

$$
\begin{aligned}
\bar{x} &= \left(\frac{\sum_j x_j}{\sum_j x_j^{1-\beta}}\right)^{\frac{1}{\beta}}\\
&= \left(\sum_i \frac{x_i}{\sum_j x_j}\cdot x_i^{-\beta}\right)^{-\frac{1}{\beta}}\\
&\ge \left(\sum_i \frac{x_i}{\sum_j x_j}\cdot x_{min}^{-\beta}\right)^{-\frac{1}{\beta}}\\
&= x_{min}.
\end{aligned} \tag{155}
$$

Similarly, we can show

$$\bar{x}\le x_{max}. \tag{156}$$

According to Corollary 8, $f(\mathbf{x})$ is increasing on $x_i\in[x_{min},\bar{x}]$ and decreasing on $x_i\in[\bar{x},x_{max}]$. Hence, $f(\mathbf{x})$ is minimized only if all x_i take the boundary values in the box constraints, i.e.,

$$x_i=x_{min}\ \text{or}\ x_i=x_{max}. \tag{157}$$

Let $\Gamma=\frac{x_{max}}{x_{min}}$ and μ be fraction of users who receive x_{max}. By relaxing the constraint $\mu\in\left\{\frac{i}{n},\forall i\right\}$ to $\mu\in[0,1]$, we derive an lower bound on $f(\mathbf{x})$ as follows

$$
\begin{aligned}
&\min_{x_i\in[x_{min},x_{max}],\forall i} f(\mathbf{x})\\
&= \min_{\mu\in\left\{\frac{i}{n},\forall i\right\}} \text{sign}(1-\beta)\cdot n\left[\frac{\mu\Gamma^{1-\beta}+(1-\mu)}{(\mu\Gamma+1-\mu)^{1-\beta}}\right]^{\frac{1}{\beta}}\\
&\ge \min_{\mu\in[0,1]} \text{sign}(1-\beta)\cdot n\left[\frac{\mu\Gamma^{1-\beta}+(1-\mu)}{(\mu\Gamma+1-\mu)^{1-\beta}}\right]^{\frac{1}{\beta}}
\end{aligned} \tag{158}
$$

To find the minimizer in the last optimization problem above, we first recognize that at the two boundary points $\mu=0$ and $\mu=1$ (i.e., all users receive the same amount of resource), $f(\mathbf{x})=n$ achieves its maximum value. Therefore, the minimum value is achieved by some $\mu\in(0,1)$. If μ^*

is the minimizer of (158), it is necessary that the first order derivative of the right hand side of (158) is zero, i.e.,

$$\frac{\partial\left[\frac{\mu\Gamma^{1-\beta}+(1-\mu)}{(\mu\Gamma+1-\mu)^{1-\beta}}\right]}{\partial\mu}=0. \tag{159}$$

Solving the above equation, we obtain

$$\begin{aligned}&(\Gamma-1)(1-\beta)\left[\left(\Gamma^{1-\beta}-1\right)\mu+1\right]\\&\quad=\left(\Gamma^{1-\beta}-1\right)\left[(\Gamma-1)\mu+1\right].\end{aligned}$$

Because this equation is a linear in μ, its root μ^* is the unique minimizer of (158):

$$\mu^*=\frac{\Gamma-\Gamma^{1-\beta}-\beta(\Gamma-1)}{\beta(\Gamma-1)(\Gamma^{1-\beta}-1)}. \tag{160}$$

The lower bound in Corollary 8 follows by plugging μ^* into (158).

10) Proof of Corollary 9: We first prove the following lemma, which will be used repeatedly in proofs of Corollaries 9,10,11,12.

Lemma 3: If $\mathbf{u},\mathbf{v}\in\mathbb{R}^n_+$ are two vectors satisfying $\frac{u_i^\uparrow}{v_i^\uparrow}\le\frac{u_j^\uparrow}{v_j^\uparrow}$ for all $i\ge j$, then $\mathbf{u}$ is majorized by $\mathbf{v}$ and $f(\mathbf{u})\le f(\mathbf{v})$. $u_i^\uparrow$ and $v_i^\uparrow$ are the ith smallest elements of $\mathbf{u}$ and $\mathbf{v}$, respectively.

Proof: Due to Schur-concavity of fairness measure f in Theorem 3, it is sufficient to prove that $\mathbf{u}$ is majorized by $\mathbf{v}$ under the given conditions. For any integer $1\le d\le n$, we have

$$\begin{aligned}\frac{\sum_{i=1}^d u_i^\uparrow}{\sum_{i=1}^n u_i^\uparrow}&=\frac{\sum_{i=1}^d u_i^\uparrow}{\sum_{i=1}^d u_i^\uparrow+\sum_{i=d+1}^n u_i^\uparrow}\\&=\frac{\sum_{i=1}^d u_i^\uparrow}{\sum_{i=1}^d u_i^\uparrow+\sum_{i=d+1}^n v_i^\uparrow\cdot\frac{u_i^\uparrow}{v_i^\uparrow}}\\&\le\frac{\sum_{i=1}^d u_i^\uparrow}{\sum_{i=1}^d u_i^\uparrow+\frac{\sum_{j=1}^d u_j^\uparrow}{\sum_{j=1}^d v_j^\uparrow}\cdot\sum_{i=d+1}^n v_i^\uparrow}\\&=\frac{\sum_{i=1}^d v_i^\uparrow}{\sum_{i=1}^d v_i^\uparrow+\sum_{i=d+1}^n v_i^\uparrow}\\&=\frac{\sum_{i=1}^d v_i^\uparrow}{\sum_{i=1}^n v_i^\uparrow},\end{aligned} \tag{161}$$

where the third step uses $\frac{\sum_{j=1}^d u_j^\uparrow}{\sum_{j=1}^d v_j^\uparrow}\le\frac{u_i^\uparrow}{v_i^\uparrow}$ for all $i\ge d$, since $u_i^\uparrow/v_i^\uparrow$ is non-decreasing over its index i. (161) implies that $\mathbf{u}$ is majorized by $\mathbf{v}$ according to subsection II-C. ■

To prove Corollary 9, sufficiency is immediate from Lemma 3, by setting $\mathbf{u}=\mathbf{x}$ and $\mathbf{v}=\mathbf{x}-c(\mathbf{x})$, which are positive due to the assumptions of tax policies in subsection B4. To verify the condition in Lemma 3, we have

$$\frac{u_i^\uparrow}{v_i^\uparrow}=\frac{x_i^\uparrow}{x_i^\uparrow-c(x_i^\uparrow)}=\left[1-\frac{c(x_i^\uparrow)}{x_i^\uparrow}\right]^{-1}, \tag{162}$$

which is non-decreasing over i due to the monotonicity of $c(z)/z$ over $z\in\mathbb{R}$. The first step in (162) uses the assumption that a tax policy does not change users' resource ranking. Therefore, $f(\mathbf{x})\le f(\mathbf{x}-c(\mathbf{x}))$ according to Lemma 1.

To prove necessity, we consider $\mathbf{x}=[x_1,x_2]$ with $x_1\le x_2$. When c is a fair tax policy, it implies that $f(x_1-c(x_1),x_2-c(x_2))\ge f(x_1,x_2)$ and $x_1-c(x_1)\le x_2-c(x_2)$ due to the assumption that a tax policy does not change users' resource ranking. From Schur-concavity of fairness measure f in Theorem 3, we have

$$\frac{x_1}{x_2}\le\frac{x_1-c(x_1)}{x_2-c(x_2)}. \tag{163}$$

A simple manipulation of (163) shows that $c(x_1)/x_1\le c(x_2)\le x_2$ for all $x_1\le x_2$. This completes the proof of Corollary 9.

11) Proof of Corollary 10: We use Lemma 3 with $\mathbf{u}=\mathbf{x}-c(\mathbf{x})$ and $\mathbf{v}=\mathbf{x}$ to prove sufficiency. We have

$$\frac{u_i^\uparrow}{v_i^\uparrow}=\frac{x_i^\uparrow}{x_i^\uparrow-c(x_i^\uparrow)}=\left[1-\frac{c(x_i^\uparrow)}{x_i^\uparrow}\right]^{-1}, \tag{164}$$

which is non-decreasing over i due to the monotonicity of $c(z)/z$ over $z\in\mathbb{R}$. Therefore, from Lemma 3, we conclude that $f(\mathbf{x})\le f(\mathbf{x}-c(\mathbf{x}))$.

For necessity, we again consider $\mathbf{x}=[x_1,x_2]$ with $x_1\le x_2$. When c is a fair tax policy and $f(x_1,x_2)\ge f(x_1-c(x_1),x_2-c(x_2))$. From Schur-concavity of fairness measure f in Theorem 3, we have

$$\frac{x_1-c(x_1)}{x_2-c(x_2)}\le\frac{x_1}{x_2}. \tag{165}$$

A simple manipulation of (165) shows that $c(x_1)/x_1\ge c(x_2)\le x_2$ for all $x_1\le x_2$. This completes the proof of Corollary 10.

12) Proof of Corollary 11: Corollary 11 is a direct consequence of Lemma 3.

13) Proof of Corollary 12: We prove the three sufficient conditions one-by-one. When c_1 is fair and c_2 is unfair, we have $f(\mathbf{x}-c_1(\mathbf{x}))\ge f(\mathbf{x})\ge f(\mathbf{x}-c_2(\mathbf{x}))$. When c_1 and c_2 are both fair, choosing $\mathbf{u}=\mathbf{x}-c_2(\mathbf{x})$ and $\mathbf{u}=\mathbf{x}-c_1(\mathbf{x})$, we have

$$\begin{aligned}\frac{u_i^\uparrow}{v_i^\uparrow}&=\frac{x_i^\uparrow-c_2(x_i^\uparrow)}{x_i^\uparrow-c_1(x_i^\uparrow)}\\&=1+\frac{c_1(x_i^\uparrow)/x_i^\uparrow-c_2(x_i^\uparrow)/x_i^\uparrow}{1-c_1(x_i^\uparrow)/x_i^\uparrow}\\&=\left[1-\frac{c_2(x_i^\uparrow)}{c_1(x_i^\uparrow)}\right]\cdot\left[x_i^\uparrow/c_1(x_i^\uparrow)-1\right]^{-1}.\end{aligned} \tag{166}$$

It is easy to see that $1-\frac{c_2(x_i^\uparrow)}{c_1(x_i^\uparrow)}$ in (166) is positive and non-decreasing over index i, since $c_1(z)\ge c_2(z)$ and $c_1(z)/c_2(z)$ is non-decreasing over $z\in\mathbb{R}$. We also have $\left[x_i^\uparrow/c_1(x_i^\uparrow)-1\right]^{-1}$ in (166) is also positive and non-decreasing over index i, since tax policy c_1 is fair. Therefore, the quantity in (166) is non-decreasing over index i. From Lemma 3, we conclude that $f(\mathbf{x}-c_1(\mathbf{x}))\ge f(\mathbf{x}-c_2(\mathbf{x}))$. When c_1 and c_2 are both unfair, by setting $\mathbf{u}=\mathbf{x}-c_1(\mathbf{x})$ and $\mathbf{u}=\mathbf{x}-c_2(\mathbf{x})$, the proof is almost the same and will not be repeated here.

14) Proof of Theorem 5: We first prove the monotonicity of $f_\beta(\mathbf{x})$ for $\beta \in (-\infty, 0)$. Consider two different values $0 > \beta_1 \geq \beta_2$. We define the a function $\phi(y) = y^{\frac{\beta_2}{\beta_1}}$ for $y \in \mathbb{R}_+$. Since $\beta_2/\beta_1 \geq 1$, the function $\phi(y)$ is convex in y. Therefore, we have

$$
\begin{aligned}
f_{\beta_2}(\mathbf{x}) &= \left[\sum_{i=1}^{n}\left(\frac{x_i}{\sum_j x_j}\right)^{1-\beta_2}\right]^{\frac{1}{\beta_2}} \\
&= \left[\sum_{i=1}^{n}\frac{x_i}{\sum_j x_j}\cdot\phi\left(\left(\frac{x_i}{\sum_j x_j}\right)^{-\beta_1}\right)\right]^{\frac{1}{\beta_2}} \\
&\leq \left[\phi\left(\sum_{i=1^n}\frac{x_i}{\sum_j x_j}\left(\frac{x_i}{\sum_j x_j}\right)^{-\beta_1}\right)\right]^{\frac{1}{\beta_2}} \\
&= \left[\phi\left(\sum_{i=1}^{n}\left(\frac{x_i}{\sum_j x_j}\right)^{1-\beta_1}\right)\right]^{\frac{1}{\beta_2}} \\
&= \left[\sum_{i=1}^{n}\left(\frac{x_i}{\sum_j x_j}\right)^{1-\beta_1}\right]^{\frac{1}{\beta_1}} \\
&= f_{\beta_1}(\mathbf{x}),
\end{aligned} \tag{167}
$$

where the third step follows from Jensen's inequality and $\beta_2 < 0$. This shows that $f_\beta(\mathbf{x})$ is increasing on $(-\infty, 0)$.

The rest of proof is similar. For $\beta \in (0, 1)$, we consider $1 > \beta_1 \geq \beta_2 > 0$. Function $\phi(y) = y^{\frac{\beta_2}{\beta_1}}$ is concave. We have

$$
\begin{aligned}
f_{\beta_2}(\mathbf{x}) &= \left[\sum_{i=1}^{n}\left(\frac{x_i}{\sum_j x_j}\right)^{1-\beta_2}\right]^{\frac{1}{\beta_2}} \\
&= \left[\sum_{i=1}^{n}\frac{x_i}{\sum_j x_j}\cdot\phi\left(\left(\frac{x_i}{\sum_j x_j}\right)^{-\beta_1}\right)\right]^{\frac{1}{\beta_2}} \\
&\leq \left[\phi\left(\sum_{i=1^n}\frac{x_i}{\sum_j x_j}\left(\frac{x_i}{\sum_j x_j}\right)^{-\beta_1}\right)\right]^{\frac{1}{\beta_2}} \\
&= f_{\beta_1}(\mathbf{x}).
\end{aligned} \tag{168}
$$

where the third step follows from Jensen's inequality and $\beta_2 > 0$. Therefore, $f_\beta(\mathbf{x})$ is increasing on $(0, 1)$.

For $\beta \in (1, \infty)$, we consider $\beta_1 \geq \beta_2 > 1$. Function $\phi(y) = y^{\frac{\beta_2}{\beta_1}}$ is concave. We have

$$
\begin{aligned}
f_{\beta_2}(\mathbf{x}) &- \left[\sum_{i=1}^{n}\left(\frac{x_i}{\sum_j x_j}\right)^{1-\beta_2}\right]^{\frac{1}{\beta_2}} \\
&= -\left[\sum_{i=1}^{n}\frac{x_i}{\sum_j x_j}\cdot\phi\left(\left(\frac{x_i}{\sum_j x_j}\right)^{-\beta_1}\right)\right]^{\frac{1}{\beta_2}} \\
&\geq -\left[\phi\left(\sum_{i=1^n}\frac{x_i}{\sum_j x_j}\left(\frac{x_i}{\sum_j x_j}\right)^{-\beta_1}\right)\right]^{\frac{1}{\beta_2}} \\
&= f_{\beta_1}(\mathbf{x}).
\end{aligned} \tag{169}
$$

where the third step follows from Jensen's inequality and $\beta_2 > 0$. Therefore, $f_\beta(\mathbf{x})$ is decreasing on $(1, \infty)$. This completes the proof of Theorem 5.

15) Proof of Theorem 6: We first assume $\beta > 1$ (which implies $f_\beta(\cdot) < 0$) and show that the condition $\lambda \leq \left|\frac{\beta}{1-\beta}\right|$ is necessary and sufficient for preserving Pareto optimality. The case where $\beta < 1$ can be shown using an analogous proof.

To show that the condition $\lambda \leq \left|\frac{\beta}{1-\beta}\right|$ is sufficient, we consider an allocation $\mathbf{x}$ and a vector $\boldsymbol{\gamma}$ such that $\gamma_i \geq 0$ for all i and $\sum_i \gamma_i = \sum_i x_i$. Clearly, $\mathbf{x}' = \mathbf{x} + \delta\boldsymbol{\gamma}$ Pareto dominates $\mathbf{x}$ for $\delta > 0$. We now consider the difference between the function (30) evaluated for these two allocations. First assume $\beta > 1$, which implies $f_\beta(\cdot) < 0$, and

$$
\begin{aligned}
&\Phi_\lambda(\mathbf{x}') - \Phi_\lambda(\mathbf{x}) \\
&\quad = \lambda\left(\ell\left(f_\beta\left(\mathbf{x}'\right)\right) - \ell\left(f_\beta\left(\mathbf{x}\right)\right)\right) + \ell\left(\sum_i x_i'\right) - \ell\left(\sum_i x_i\right) \\
&\quad = -\lambda\left(\log\left|f_\beta\left(\mathbf{x}'\right)\right| - \log\left|f_\beta\left(\mathbf{x}\right)\right|\right) \\
&\qquad\quad + \log\left((1+\delta)\sum_i x_i\right) - \log\left(\sum_i x_i\right) \\
&\quad = -\lambda\left(\log\left|f_\beta\left(\mathbf{x}'\right)\right| - \log\left|f_\beta\left(\mathbf{x}\right)\right|\right) + \log\left(1+\delta\right).
\end{aligned} \tag{170}
$$

If $\mathbf{x}'$ is also more fair than $\mathbf{x}$, then showing

$$
-\lambda\left(\log\left|f_\beta\left(\mathbf{x}'\right)\right| - \log\left|f_\beta\left(\mathbf{x}'\right)\right|\right) > 0 \tag{171}
$$

is trivial, and the difference between the objective evaluated at the two allocations is strictly positive. Therefore, we consider the case where $\mathbf{x}'$ is less fair.

Continuing from (170) and applying the definition in (15) yields

$$
\begin{aligned}
&\Phi_\lambda(\mathbf{x}') - \Phi_\lambda(\mathbf{x}) \\
&\quad = -\lambda\log\left(\left[\sum_{i=1}^{n}\left(\frac{x_i'}{\sum_j x_j'}\right)^{1-\beta}\right]^{\frac{1}{\beta}}\right) + \log\left(1+\delta\right) \\
&\qquad\quad + \lambda\log\left(\left[\sum_{i=1}^{n}\left(\frac{x_i}{\sum_j x_j}\right)^{1-\beta}\right]^{\frac{1}{\beta}}\right) \\
&\quad = -\frac{\lambda}{\beta}\log\left(\frac{\sum_{i=1}^{n}\left(x_i'\right)^{1-\beta}}{\sum_{i=1}^{n}\left(x_i\right)^{1-\beta}}\right) + \left(1 - \lambda\frac{\beta-1}{\beta}\right)\log\left(1+\delta\right).
\end{aligned} \tag{172}
$$

Because $x_i' \geq x_i$ for all i, we know that for $\beta > 1$, $(x_i')^{1-\beta} \leq (x_i)^{1-\beta}$, which implies

$$
-\frac{\lambda}{\beta}\log\left(\frac{\sum_{i=1}^{n}\left(x_i'\right)^{1-\beta}}{\sum_{i=1}^{n}\left(x_i\right)^{1-\beta}}\right) > 0. \tag{173}
$$

Consequently, for the entire difference to be positive, it is sufficient that

$$
1 - \lambda\frac{\beta-1}{\beta} \geq 0, \tag{174}
$$

or, equivalently,

$$
\lambda \leq \frac{\beta}{\beta-1}. \tag{175}
$$

Next, we prove that the condition $\lambda \leq \left|\frac{\beta}{1-\beta}\right|$ is necessary. Suppose $\beta > 1$ as before and $\lambda > \left|\frac{\beta}{1-\beta}\right|$ for constructing a contradiction. We show that there exists two vectors $\mathbf{x}$ and $\mathbf{x}'$, such that $\Phi_\lambda(\mathbf{x}') - \Phi_\lambda(\mathbf{x}) < 0$, while $\mathbf{x}'$ Pareto dominates $\mathbf{x}$.

Consider a resource allocation vector $\mathbf{x}$ of length $n+1$, such that $x_i = 1$ for $i = 1, \ldots, n$ and $x_{n+1} = n$. Clearly, $\mathbf{x}$ is Pareto dominated by another vector $\mathbf{x}'$, where $x'_i = x_i$ for $i = 1, \ldots, n$ and $x'_{n+1} = x_i + \delta\left(\sum_i x_i\right)$, for some positive $\delta > 0$. Let $\xi = \log(1+\delta) > 0$ be an axillary variable. From the last step of (172), we have

$$
\begin{aligned}
&\Phi_\lambda(\mathbf{x}') - \Phi_\lambda(\mathbf{x}) \\
&\quad= -\frac{\lambda}{\beta}\log\left(\frac{\sum_{i=1}^{n+1}(x'_i)^{1-\beta}}{\sum_{i=1}^{n+1}(x_i)^{1-\beta}}\right) + \left(1 - \lambda\frac{\beta-1}{\beta}\right)\xi \\
&\quad= -\frac{\lambda}{\beta}\log\left(\frac{n + (n+2n\delta)^{1-\beta}}{n + n^{1-\beta}}\right) + \left(1 - \lambda\frac{\beta-1}{\beta}\right)\xi \\
&\quad\leq -\frac{\lambda}{\beta}\log\left(\frac{n}{n + n^{1-\beta}}\right) + \left(1 - \lambda\frac{\beta-1}{\beta}\right)\xi \\
&\quad= -\frac{\lambda}{\beta}\log\left(1 + n^{-\beta}\right) + \left(1 - \lambda\frac{\beta-1}{\beta}\right)\xi
\end{aligned}
$$

It is straight forward to verify that $\Phi_\lambda(\mathbf{x}') - \Phi_\lambda(\mathbf{x}) < 0$, if we set

$$\delta = \frac{1}{2}\left[\left(1 + n^{-\beta}\right)^{\frac{\lambda}{\lambda(\beta-1)-\beta}}\right] > 0. \tag{176}$$

As a result, the condition (32) of the theorem is sufficient and necessary for ensuring Pareto optimality of the solution.

16) Proof of Theorem 7: From the definition of α-fair utility, we compute the numerator and denominator:

$$
\begin{aligned}
\left\langle \nabla_{\mathbf{x}} U_{\alpha=\beta}(\mathbf{x}), \frac{\boldsymbol{\eta}}{\|\boldsymbol{\eta}\|}\right\rangle &= \sum_i x_i^{-\beta}\frac{\frac{1}{N} - \frac{x_i}{\sum_i x_i}}{\sqrt{\frac{\|\mathbf{x}\|^2}{(\sum_i x_i)^2} - \frac{1}{N}}} \\
&= \frac{1}{\sqrt{\frac{\|\mathbf{x}\|^2 N}{(\sum_i x_i)^2} - 1}} \cdot \frac{1}{\sqrt{N}} \\
&\quad \cdot \sum_i x_i^{-\beta}\left(1 - \frac{x_i}{\sum_j x_j}N\right),
\end{aligned} \tag{177}
$$

and

$$\left\langle \nabla_{\mathbf{x}} U_{\alpha=\beta}(\mathbf{x}), \frac{\mathbf{1}}{\|\mathbf{1}\|}\right\rangle = \sum_i x_i^{-\beta}\frac{1}{\sqrt{N}}. \tag{178}$$

Notice that both values are positive. The ratio between these then is

$$\frac{\left\langle \nabla_{\mathbf{x}} U_{\alpha=\beta}(\mathbf{x}), \frac{\boldsymbol{\eta}}{\|\boldsymbol{\eta}\|}\right\rangle}{\left\langle \nabla_{\mathbf{x}} U_{\alpha=\beta}(\mathbf{x}), \frac{\mathbf{1}}{\|\mathbf{1}\|}\right\rangle} = \frac{1}{\sqrt{\frac{\|\mathbf{x}\|^2 N}{(\sum_i x_i)^2} - 1}}\left(1 - \frac{\sum_i \frac{x_i}{\sum_j x_j} x_i^{-\beta}}{\sum_i \frac{1}{N} x_i^{-\beta}}\right). \tag{179}$$

It is easily shown that the factor in front of the bracket is strictly positive. The only component that varies with β is the ratio between two weighted averages of the same vector with different weights:

$$\frac{\sum_i \frac{x_i}{\sum_j x_j} x_i^{-\beta}}{\sum_i \frac{1}{N} x_i^{-\beta}}. \tag{180}$$

The sum in the numerator places more weight ($\frac{x_i}{\sum_j x_j} > \frac{1}{N}$) on elements that decrease more rapidly (or increase more slowly for the case $x_i < 1$) with β, implies that the overall numerator decreases more rapidly (or increases more slowly) than the denominator as β increases. Therefore, (180) is monotonically non-increasing, and Theorem 7 is true.

17) Proof of Theorem 8: To prove the equivalent fairness representation in (41), we start with the integral in (39):

$$
\begin{aligned}
&\frac{1}{T}\int_{t=0}^{T} (\mathbf{X}_t/\mu)^{1-\beta r}\, dt \\
&\quad= \sum_{i=1}^{n} \frac{1}{T}\int_{\{y_i \leq \mathbf{X}_t < y_{i+1}\}} (\mathbf{X}_t/\mu)^{1-\beta r}\, dt \\
&\quad= \sum_{i=1}^{n} (\bar{y}_i/\mu)^{1-\beta r}\frac{1}{T}\int_{\{y_i \leq \mathbf{X}_t < y_{i+1}\}} dt \\
&\quad= \sum_{i=1}^{n} (\bar{y}_i/\mu)^{1-\beta r}\left[P_{\mathbf{x}}(y_{i+1}) - P_{\mathbf{x}}(y_i)\right]
\end{aligned} \tag{181}
$$

where $0 = y_1 < y_2 < \ldots < y_{n+1} < \infty$ is a partition of the domain of $P_{\mathbf{x}}$, and $y_i \leq \bar{y}_i \leq y_{i+1}$ is a proper value in step 2. The last step uses the definition of $P_{\mathbf{x}}$ in (43). Let $n \to \infty$. Due to continuity, we derive

$$\frac{1}{T}\int_{t=0}^{T} (\mathbf{X}_t/\mu)^{1-\beta r}\, dt = \int_{y=0}^{\infty} (y/\mu)^{1-\beta r}\, dP_{\mathbf{x}}(y) \tag{182}$$

which proves (41).

Similarly, to prove the equivalent fairness expression in (42), we start with (182):

$$
\begin{aligned}
&\int_{y=0}^{\infty} (y/\mu)^{1-\beta r}\, dP_{\mathbf{x}}(y) \\
&\quad= \sum_{i=1}^{n} \int_{\{u_i \leq P_{\mathbf{x}}(y) < u_{i+1}\}} (y/\mu)^{1-\beta r}\, dP_{\mathbf{x}}(y) \\
&\quad= \sum_{i=1}^{n} \left(P_{\mathbf{x}}^{-1}(\bar{u})/\mu\right)^{-\beta r}\int_{\{y_i \leq \mathbf{X}_t < y_{i+1}\}} \frac{y}{\mu} dP_{\mathbf{x}}(y) \\
&\quad= \sum_{i=1}^{n} \left(P_{\mathbf{x}}^{-1}(\bar{u})/\mu\right)^{-\beta r}\left[L_{\mathbf{x}}(u_{i+1}) - L_{\mathbf{x}}(u_i)\right]
\end{aligned} \tag{183}
$$

where where $0 = y_1 < y_2 < \ldots < y_{n+1} = 1$ is a partition of the domain of $L_{\mathbf{x}}$, and $u_i \leq \bar{u}_i \leq u_{i+1}$ is a proper value in step 2. $P_{\mathbf{x}}^{-1}$ exists, since cumulative distribution $P_{\mathbf{x}}$ is monotone. The last step uses the definition of $L_{\mathbf{x}}$ in (44). Let $n \to \infty$. Due to continuity, we derive

$$\int_{y=0}^{\infty}\left(\frac{y}{\mu}\right)^{1-\beta r} dP_{\mathbf{x}}(y) = \int_{u=0}^{1}\left(\frac{P_{\mathbf{x}}^{-1}(u)}{\mu}\right)^{-\beta r} dL_{\mathbf{x}}(u) \tag{184}$$

which, combined with (182), proves (42).

18) Proof of Theorem 9: From (42), we find that $\mathcal{F}(\mathbf{X}_{t\in[0,T]}) \leq \mathcal{F}(\mathbf{Y}_{t\in[0,T]})$ holds if

$$\int_{u=0}^{1}\left(\frac{P_{\mathbf{x}}^{-1}(u)}{\mu}\right)^{-\beta r} dL_{\mathbf{x}}(u) \leq \int_{v=0}^{1}\left(\frac{P_{\mathbf{y}}^{-1}(v)}{\mu}\right)^{-\beta r} dL_{\mathbf{y}}(v)$$

Since above inequality is an integral of a non-negative function over Lorenz curve $dL_{\mathbf{x}}$ and $dL_{\mathbf{x}}$, the axiomatic theory of Lorenz curve ordering in [18] implies that this inequality defines a *complete order* over all Lorenz curves. From the Theorem 9 in [18], we have

$$\mathcal{F}(\mathbf{X}_{t\in[0,T]}) \leq \mathcal{F}(\mathbf{Y}_{t\in[0,T]}), \text{ if } L_{\mathbf{x}}(u) \leq L_{\mathbf{y}}(u) \ \forall u \in [0,1].$$

This equation, together with (45), completes the proof.

19) Proof of Corollaries 13 and 14: From the construction of $\mathcal{F}$ in Section III,

$$\mathcal{F}(\mathbf{X}_{t\in[0,T]}) = \lim_{n\to\infty} n^{-r} f(\mathbf{X}_{t_1}, \ldots, \mathbf{X}_{t_n}), \tag{185}$$

where $0 = t_0 < t_1 < t_2 < \ldots < t_n = T$ is a partition of $[0,T]$. Due to Corollary 2 and Corollary 3 in Section II, f is maximized by equal resource allocation and has maximum value $|f| = n^r$. We have

$$n^{-r} f(\mathbf{X}_{t_1}, \ldots, \mathbf{X}_{t_n}) \leq 1, \tag{186}$$

for all dimension n. Taking $n \to \infty$ in (186), we find that $\mathcal{F}(\mathbf{X}_{t\in[0,T]}) \leq 1$. When $\mathcal{F}$ is positive, it is easy to verify using (39) that optimal value $\mathcal{F}(\mathbf{X}_{t\in[0,T]}) = 1$ can be achieved by equal resource allocation over time, i.e., $\{\mathbf{X}_t = c : t \in [0,T], c > 0\}$. The proof for negative $\mathcal{F}$ is similar.

20) Proof of Corollary 15: We define $\mathbf{Y}_t = \mathbf{X}_t - c$ for $c > 0$ and all $t \in [0,T]$ as the shifted resource allocation. From the definition of cumulative distribution in (43), we find

$$P_{\mathbf{y}}(y) = \frac{1}{T} \cdot \int_{\{\mathbf{Y}_t \leq y\}} dt = \frac{1}{T} \cdot \int_{\{\mathbf{X}_t \leq y+c\}} dt = P_{\mathbf{x}}(y+c) \tag{187}$$

Using the definition of Lorenz curve in (44), we derive

$$\begin{aligned} L_{\mathbf{y}}(u) &= \frac{1}{\mu - c} \cdot \int_{\{P_{\mathbf{y}}(y) \leq u\}} y dP_{\mathbf{y}}(y) \qquad (188)\\ &= \frac{1}{\mu - c} \cdot \int_{\{P_{\mathbf{x}}(y+c) \leq u\}} y dP_{\mathbf{x}}(y+c) \\ &= \frac{1}{\mu - c} \cdot \int_{\{P_{\mathbf{x}}(z) \leq u, z > c\}} (z-c) dP_{\mathbf{x}}(z) \\ &= \frac{1}{\mu} \cdot \int_{\{P_{\mathbf{x}}(z) \leq u, z > c\}} \frac{\mu(z-c)}{\mu - c} dP_{\mathbf{x}}(z) \\ &= \frac{1}{\mu} \cdot \int_{\{P_{\mathbf{x}}(z) \leq u\}} \frac{\mu(z-c)}{\mu - c} dP_{\mathbf{x}}(z) \\ &= \frac{1}{\mu} \cdot \int_{\{P_{\mathbf{x}}(z) \leq u\}} \left(z + \frac{c}{\mu - c}(z - \mu) \right) dP_{\mathbf{x}}(z) \end{aligned}$$

The second step uses (187). The third step changes variable from y to $z - c$. The fifth step follows from the fact that $P_{\mathbf{x}}(z) = 0$ for $z \leq c$, since $c \leq \mathbf{X}_t$ for all $t \in [0,T]$ (otherwise we have $\mathbf{Y}_t < 0$, which is not possible). Since $z = P_{\mathbf{x}}^{-1}(u)$ is non-decreasing over u, we have

$$\mu = \int_{\{P_{\mathbf{x}}(z) \leq 1\}} z dP_{\mathbf{x}}(z) \geq \frac{1}{u} \int_{\{P_{\mathbf{x}}(z) \leq u\}} z dP_{\mathbf{x}}(z) \tag{189}$$

which follows from the fact that average of a non-decreasing function $P_{\mathbf{x}}^{-1}$ over $[0,u]$ is non-decreasing. Using this inequality, for the last term in (189), we have

$$\int_{\{P_{\mathbf{x}}(z) \leq u\}} (z-\mu) dP_{\mathbf{x}}(z) = \int_{\{P_{\mathbf{x}}(z) \leq u\}} z dP_{\mathbf{x}}(z) - \mu u \leq 0.$$

Combine above equation and (189). We find

$$L_{\mathbf{y}}(u) \leq \int_{\{P_{\mathbf{x}}(z) \leq u\}} z dP_{\mathbf{x}}(z) = L_{\mathbf{x}}(u) \ \forall u. \tag{190}$$

From Schur-concavity property in Theorem 13, we conclude that $\mathcal{F}(\mathbf{X}_{t\in[0,T]} - c) \leq \mathcal{F}(\mathbf{X}_{t\in[0,T]})$.

21) Proof of Corollary 16: Let ω be the average of $\mathbf{X}_{t\in[0,T]} \bigcup \mathbf{0}_{t\in[T,T+\tau]}$. It is easy to see that

$$\omega = \frac{1}{T+\tau} \int_{t=0}^{T} \mathbf{X}_t dt = \frac{T}{T+\tau} \cdot \mu. \tag{191}$$

Let $s = \text{sign}(r(1-r\beta))$ be the sign of fairness measure $\mathcal{F}$. When $1 - \beta r > 0$, from the definition of fairness $\mathcal{F}$ in (39), we derive

$$\begin{aligned} &\mathcal{F}(\mathbf{X}_{t\in[0,T]} \bigcup \mathbf{0}_{t\in[T,T+\tau]}) \\ &= s \left[\frac{1}{T+\tau} \int_{t=0}^{T} (\mathbf{X}_t/\omega)^{1-\beta r} dt + 0 \right]^{\frac{1}{\beta}} \\ &= s \left[\frac{T^{r\beta}}{(T+\tau)^{r\beta}} \cdot \frac{1}{T} \int_{t=0}^{T} (\mathbf{X}_t/\mu)^{1-\beta r} dt + 0 \right]^{\frac{1}{\beta}} \\ &= (1+\tau/T)^{-r} \cdot s \left[\frac{1}{T} \int_{t=0}^{T} (\mathbf{X}_t/\mu)^{1-\beta r} dt + 0 \right]^{\frac{1}{\beta}} \\ &= (1+\tau/T)^{-r} \cdot \mathcal{F}(\mathbf{X}_{t\in[0,T]}). \end{aligned} \tag{192}$$

Adding a time interval of length $\tau > 0$ with zero resource reduces fairness by a factor of $(1+\tau/T)^{-r}$.

22) Proof of Corollary 17: Suppose that $\mathbf{X}_t > 0$ for all $t \in [0,T]$. We consider a resource allocation

$$\mathbf{Y}_{t\in[0,T+\tau]} \triangleq \mathbf{X}_{t\in[0,T]} \bigcup \mathbf{0}_{t\in[T,T+\tau]}, \tag{193}$$

whose length of inactive time intervals is exactly τ. First, we consider $r > 0$, so that fairness $\mathcal{F}$ is positive. From Corollary 16 and 17, we have

$$\mathcal{F}(\mathbf{Y}_{t\in[0,T+\tau]}) = (1+\tau/T)^{-r} \cdot \mathcal{F}(\mathbf{X}_{t\in[0,T]}) \leq (1+\tau/T)^{-r}$$

Rearranging the terms, we conclude that

$$\tau \leq T \cdot [(\mathcal{F}(\mathbf{Y}_{t\in[0,T+\tau]}))^{-1/r} - 1], \tag{194}$$

When $r < 0$ and fairness $\mathcal{F}$ is negative, we have

$$\begin{aligned} \mathcal{F}(\mathbf{Y}_{t\in[0,T+\tau]}) &= (1+\tau/T)^{-r} \cdot \mathcal{F}(\mathbf{X}_{t\in[0,T]}) \\ &\leq -(1+\tau/T)^{-r} \end{aligned} \tag{195}$$

Rearranging the terms and considering that $-r > 0$, we conclude that

$$\tau \leq T \cdot [(-\mathcal{F}(\mathbf{Y}_{t\in[0,T+\tau]}))^{-1/r} - 1], \tag{196}$$

These inequalities give an upper bound of inactive time intervals (with zero resource) for $1 - r\beta > 0$ and all $r \in \mathbb{R}$.

23) Proof of Corollary 18: Suppose resource allocation $\mathbf{X}_{t\in[0,T]}$ has fairness $\mathcal{F}$. Since μ is the average resource, we find

$$\begin{aligned} \mu &= \frac{1}{T} \int_{\{\mathbf{X}_t > 0\}} \mathbf{X}_t dt \\ &\leq (\max_{t\in[0,T]} \mathbf{X}_t) \cdot \frac{1}{T} \int_{\{\mathbf{X}_t > 0\}} dt \\ &\leq (\max_{t\in[0,T]} \mathbf{X}_t) \cdot (2 - |\mathcal{F}|^{-1/r}) \end{aligned} \tag{197}$$

where the last step uses Corollary 17 - the length of inactive time intervals is upper bounded by $[|\mathcal{F}|^{-1/r} - 1]$. If $2-|\mathcal{F}|^{-1/r} > 0$, we conclude that the peak resource allocation over time $[0, T]$ is lower bounded by

$$\max_{t\in[0,T]} \mathbf{X}_t \geq \frac{\mu}{2 - |\mathcal{F}|^{-1/r}} \tag{198}$$

24) Proof of Corollary 19: We prove a lower bound on fairness $\mathcal{F}$ for box constraints. Define

$$h(\mathbf{X}_{t\in[0,T]}) = \int_{t=0}^{T} (\mathbf{X}_t/\mu)^{1-\beta r}\, dt \tag{199}$$

as an auxiliary function, such that

$$\mathcal{F}(\mathbf{X}_{t\in[0,T]}) = \text{sign}(r(1-\beta r)) \cdot \left[\frac{1}{T} h(\mathbf{X}_{t\in[0,T]})\right]^{\frac{1}{\beta}}. \tag{200}$$

In the following proof, we will only consider $r(1 - r\beta) > 0$ and $\beta > 0$, while the proofs for other cases are similar and omitted for simplicity. When $r(1 - r\beta) > 0$ and $\beta > 0$, to minimize fairness $\mathcal{F}$ over resource allocation $\mathbf{X}_{t\in[0,T]}$, we formulate the following minimization problem of $h(\mathbf{X}_{t\in[0,T]})$ for average resource constraint μ and the box constraint $x_{min} \leq \mathbf{X}_t \leq x_{max}$:

$$\begin{aligned} \min_{\mathbf{X}_{t\in[0,T]}} \quad & h(\mathbf{X}_{t\in[0,T]}) \\ s.t. \quad & x_{min} \leq \mathbf{X}_t \leq x_{max},\ \forall t \\ & \mu = (1/T) \cdot \int_{t=0}^{T} \mathbf{X}_t dt \end{aligned} \tag{201}$$

We introduce a Lagrangian multiplier λ for the average resource constraint, and two Lagrangian multiplier functions, ϕ_t and ψ_t, for the two boundaries of the box constraint. We formulate an Lagrangian for the minimization problem:

$$\begin{aligned} L \ &= \int_{t=0}^{T} (\mathbf{X}_t/\mu)^{1-\beta r} + \lambda \left(\mu - \frac{1}{T}\int_{0}^{T} \mathbf{X}_t dt\right) \\ &+ \int_{0}^{T} (\mathbf{X}_t - x_{max})\phi_t dt + \int_{0}^{T} (x_{min} - \mathbf{X}_t)\psi_t dt. \end{aligned}$$

Since KKT conditions are necessary for optimality, we derive the first order condition, i.e., the derivative of L with respect to $\mathbf{X}_t$ is zero:

$$\frac{\partial L}{\partial \mathbf{X}_t} = \frac{\mathbf{X}_t^{-r\beta}}{\mu^{1-r\beta}}(1 - r\beta) - \frac{\lambda}{T} + \phi_t - \psi_t = 0, \tag{202}$$

and complementary slackness conditions:

$$(\mathbf{X}_t - x_{max})\phi_t = 0,\ \forall t, \tag{203}$$
$$(x_{min} - \mathbf{X}_t)\psi_t = 0,\ \forall t. \tag{204}$$

Let $\mathbf{X}_t^*$ be an optimizer. We first prove that $\mathbf{X}_t^*$ is either x_{min} or x_{max} for all $t \in [0, T]$. Suppose $x_{min} < \mathbf{X}_t^* < x_{max}$ holds for some $t \in \mathcal{T}$, from (203) and (204), we conclude that $\phi_t = 0$ and $\psi_t = 0$. Substituting this result into (202), we find that when $t \in \mathcal{T}$, $\mathbf{X}_t^*$ is a constant, which is denoted by x_0 and satisfies

$$\frac{x_0^{-r\beta}}{\mu^{1-r\beta}}(1 - r\beta) - \frac{\lambda}{T} = 0. \tag{205}$$

Because $\mathbf{X}_t^{-r\beta}$ is strictly increasing over $\mathbf{X}_t$, we find that when $\mathbf{X}_t^* = x_{max}$

$$\begin{aligned} \psi_t - \phi_t \ &= \frac{x_{max}^{-r\beta}}{\mu^{1-r\beta}}(1 - r\beta) - \frac{\lambda}{T} \\ &> \frac{x_0^{-r\beta}}{\mu^{1-r\beta}}(1 - r\beta) - \frac{\lambda}{T} = 0 \end{aligned} \tag{206}$$

where the first step uses the first order condition for $\mathbf{X}_t^* = x_{max}$. Since $\phi_t, \psi_t \geq 0$ are non-negative, inequality (206) implies that $\psi_t > \phi_t \geq 0$. Inequality $\psi_t > 0$ and $\mathbf{X}_t^* = x_{max} > x_{min}$ contradict with complementary slackness condition (204). Therefore, we conclude that $x_{min} < \mathbf{X}_t^* < x_{max}$ is impossible. $\mathbf{X}_t^*$ must take boundary values of the box constraint, i.e., it is either x_{min} or x_{max} for all $t \in [0, T]$.

Define $\Gamma = x_{max}/x_{min}$ and θ to be the fraction of time the $\mathbf{X}_t^* = x_{max}$. We rewrite the auxiliary function h by

$$\begin{aligned} h(\mathbf{X}_t^*) \ &= \int_{t=0}^{T} (\mathbf{X}_t^*/\mu)^{1-\beta r}\, dt \\ &= \frac{\theta T x_{max}^{1-r\beta} + (1-\theta) T x_{min}^{1-r\beta}}{(\theta x_{max} + (1-\theta)x_{min})^{1-r\beta}} \end{aligned} \tag{207}$$

We recognize that at the two boundary points $\theta = 0$ and $\theta = 1$ (i.e. all users receive the same amount of resource), $\mathcal{F} = 1$ achieves its maximum value. Therefore, the minimum fairness is achieved by some $\theta \in (0, 1)$. We set $\partial h/\partial\theta = 0$,

$$\begin{aligned} \tfrac{\partial h}{\partial \theta} h(\mathbf{X}_t^*) \ &= \frac{\partial}{\partial \theta}\left[\frac{\theta T x_{max}^{1-r\beta} + (1-\theta) T x_{min}^{1-r\beta}}{(\theta x_{max} + (1-\theta)x_{min})^{1-r\beta}}\right] \\ &= 0. \end{aligned} \tag{208}$$

Solving the last equality in (208), we obtain

$$\begin{aligned} &(x_{max}^{1-r\beta} - x_{min}^{1-r\beta}) z^{1-r\beta} \\ &\quad = (\theta x_{max}^{1-r\beta} + (1-\theta)x_{min}^{1-r\beta})(1 - r\beta) z^{r\beta}(x_{max} - x_{min}). \end{aligned}$$

where $z = \theta x_{max} + (1-\theta)x_{min}$. Simplify the equation above, we find a unique optimizer θ as follows:

$$\theta = \frac{1}{1-\rho} \cdot \left[\frac{(\Gamma - 1)^{\rho}}{\Gamma^{\rho} - 1} - 1\right] \tag{209}$$

where $\Gamma = x_{max}/x_{min}$ and $\rho = 1 - r\beta$. This complete the proof.

25) Proof of Corollary 20: We first prove the monotonicity of $\mathcal{F}_\beta(\mathbf{X}_{t\in[0,T]})$ for $\beta \in (-\infty, 0)$. Consider two different values $0 > \beta_1 \geq \beta_2$. We define the a function $\phi(y) = y^{\frac{\beta_2}{\beta_1}}$ for $y \in \mathbb{R}_+$. Since $\beta_2/\beta_1 \geq 1$, the function $\phi(y)$ is convex in y. Since $T = 1$ and $\mathcal{F}_\beta(\mathbf{X}_{t\in[0,T]})$ is positive for $\beta \in (-\infty, 0)$, we have

$$\begin{aligned}\mathcal{F}_{\beta_2}(\mathbf{X}_{t\in[0,T]}) &= \left[\int_{t=0}^{T} (\mathbf{X}_t/\mu)^{1-\beta_2}\, dt\right]^{\frac{1}{\beta_2}} \\ &= \left[\int_{t=0}^{T} (\mathbf{X}_t/\mu)\cdot \phi\left((\mathbf{X}_t/\mu)^{-\beta_1}\right) dt\right]^{\frac{1}{\beta_2}} \\ &\le \left[\phi\left(\int_{t=0}^{T} (\mathbf{X}_t/\mu)\cdot (\mathbf{X}_t/\mu)^{-\beta_1}\, dt\right)\right]^{\frac{1}{\beta_2}} \\ &= \left[\phi\left(\int_{t=0}^{T} (\mathbf{X}_t/\mu)^{1-\beta_1}\, dt\right)\right]^{\frac{1}{\beta_2}} \\ &= \left[\int_{t=0}^{T} (\mathbf{X}_t/\mu)^{1-\beta_1}\, dt\right]^{\frac{1}{\beta_1}} \\ &= \mathcal{F}_{\beta_1}(\mathbf{X}_{t\in[0,T]}),\end{aligned} \tag{210}$$

where the third step follows from Jensen's inequality and $\beta_2 < 0$. This proves that $\mathcal{F}_\beta(\mathbf{X}_{t\in[0,T]})$ is increasing on $(-\infty, 0)$.

The rest of proof is similar. For $\beta \in (0,1)$, we consider $1 > \beta_1 \ge \beta_2 > 0$, such that function $\phi(y) = y^{\frac{\beta_2}{\beta_1}}$ is concave. For $\beta \in (1,\infty)$, we consider $\beta_1 \ge \beta_2 > 1$, such that function $\phi(y) = y^{\frac{\beta_2}{\beta_1}}$ is concave. Apply the chains on inequality (210) to $\beta \in (0,1)$ and $\beta \in (1,\infty)$ respectively, we can prove the monotonicity of $\mathcal{F}_\beta(\mathbf{X}_{t\in[0,T]})$ for $\beta \in \mathbb{R}$.

26) Proof of Theorems 10, 11, and 12: This proof is an extension of the proof in C.1. We prove that fairness measures satisfying Axioms $1'$-$5'$ can only be generated by logarithm and power generator functions. Then, for each choice of logarithm and power generator function, a fairness measure exist and is uniquely given by either (65) or (66). It is easy to verify that if $f > 0$ satisfies Axioms $1'$-$5'$, so does $-f$. We first assume that $f > 0$ during this proof and then determine sign to fairness measures from Axioms $5'$. Since the proof of Lemma 2 does not depend on the definition of s_i, it holds for all $\mathbf{q} \in \mathbb{Z}_+^n$. Therefore, Lemma 2 becomes

$$f(\mathbf{1}_n, \mathbf{q}) = n^r \cdot f(1,1), \quad \forall \mathbf{q} \in \mathbb{Z}_+,\ n \ge 1, \tag{211}$$

(Uniqueness of logarithm and power generator functions.) We first show that logarithm and power generator functions are necessary. In other words, no other fairness can possibly satisfy Axioms $1'$-$4'$ with weight s_i in (62). Consider arbitrary positive $X_1 + X_2 = 1$, which can be written as the ρ-th power of rational numbers, i.e., $X_1 = (a_1/b_1)^\rho$ and $X_2 = (a_2/b_2)^\rho$, for some positive integers a_1, b_1, a_2, b_2. Similarly, we choose arbitrary positive $Q_1 + Q_2 = 1$, which can be written as the ρ-th power of rational numbers, i.e., $Q_1 = (c_1/d_1)^\rho$ and $X_2 = (c_2/d_2)^\rho$, for some positive integers c_1, d_1, c_2, d_2. Let u, v be two positive integers and $K = (u/v)^{1/\rho}$ be rational number. We choose the sum resource of the two sub-systems by $\mathcal{W}_1 = a_1 b_2 c_2 d_1 v$ and $\mathcal{W}_2 = a_2 b_1 c_1 d_2 v$, so that

$$\begin{aligned}\frac{Q_1\mathcal{W}_1^\rho}{Q_1\mathcal{W}_1^\rho + Q_2\mathcal{W}_2^\rho} &= \frac{(c_1 d_2 a_1 b_2 c_2 d_1 v)^\rho}{(c_1 d_2 a_1 b_2 c_2 d_1 v)^\rho + (c_2 d_1 a_2 b_1 c_1 d_2 v)^\rho} \\ &= \frac{(a_1/b_1)^\rho}{(a_1/b_1)^\rho + (a_2/b_2)^\rho} \\ &= \frac{X_1}{X_1 + X_2} \\ &= X_1\end{aligned} \tag{212}$$

where the last step uses $X_1 + X_2 = 1$. Similarly, we have

$$\frac{Q_2\mathcal{W}_2^\rho}{Q_1\mathcal{W}_1^\rho + Q_2\mathcal{W}_2^\rho} = X_2 \tag{213}$$

Consider a partition and two resource allocation vectors as follows: for $\mathbf{x}$, sub-system 1 has $t = b_1 b_2 c_1 c_2 u$ users, each with weight Q_1/t and resource $\mathcal{W}_1/t$; sub-system 2 also has $t = b_1 b_2 c_1 c_2 u$ users, each with weight Q_2/t and resource $\mathcal{W}_2/t$. In other words, we have

$$\mathbf{x}^1 = [\underbrace{\frac{\mathcal{W}_1}{t}, \ldots, \frac{\mathcal{W}_1}{t}}_{t\ \text{times}}] \text{ and } \mathbf{x}^2 = [\underbrace{\frac{\mathcal{W}_2}{t}, \ldots, \frac{\mathcal{W}_2}{t}}_{t\ \text{times}}], \tag{214}$$

$$\mathbf{q}^1 = [\underbrace{\frac{Q_1}{t}, \ldots, \frac{Q_1}{t}}_{t\ \text{times}}] \text{ and } \mathbf{q}^2 = [\underbrace{\frac{Q_2}{t}, \ldots, \frac{Q_2}{t}}_{t\ \text{times}}]. \tag{215}$$

For $\mathbf{y}$, sub-system 1 has $u^1 = a_1 b_2 c_2 d_1 v$ users, each with weight Q_1/u^1 and resource 1; sub-system 2 also has $u^2 = a_2 b_1 c_1 d_2 v$ users, each with weight Q_2/u^2 and resource 1. This means

$$\mathbf{y}^1 = \underbrace{[1, \ldots, 1]}_{u^1\ \text{times}} \text{ and } \mathbf{y}^2 = \underbrace{[1, \ldots, 1]}_{u^2\ \text{times}}, \tag{216}$$

$$\mathbf{p}^1 = [\underbrace{\frac{Q_1}{u^1}, \ldots, \frac{Q_1}{u^1}}_{u^1\ \text{times}}] \text{ and } \mathbf{p}^2 = [\underbrace{\frac{Q_2}{u^2}, \ldots, \frac{Q_2}{u^2}}_{u^2\ \text{times}}]. \tag{217}$$

Since $w(\mathbf{x}^i) = w(\mathbf{y}^i)$ and $w(\mathbf{q}^i) = w(\mathbf{p}^i)$, (218) holds due to Axiom $4'$ and (211). The second step of (218) uses Axiom 2 and (211), i.e.,

$$\begin{aligned}f(\mathbf{x}^i, \mathbf{q}^i) &= f\left(\frac{\mathcal{W}_i}{t}\cdot \mathbf{1}_t, \frac{\mathcal{Q}_i}{t}\cdot \mathbf{1}_t\right) \\ &= f\left(\mathbf{1}_t, \frac{\mathcal{Q}_i}{t}\cdot \mathbf{1}_t\right) \\ &= t^r\end{aligned} \tag{219}$$

and similarly,

$$f(\mathbf{y}^1) = (a_1 b_2 c_2 d_1 v)^r, \tag{220}$$

$$f(\mathbf{y}^2) = (a_2 b_1 c_1 d_2 v)^r. \tag{221}$$

Next, we apply (63) to $f(\mathbf{x}, \mathbf{q})$ and obtain

$$\begin{aligned}&f(\mathbf{x}, \mathbf{q}) \\ &= f\left(\frac{\mathcal{W}_1 + \mathcal{W}_2}{t}\cdot \mathbf{1}_t, \frac{Q_1 + Q_2}{t}\cdot \mathbf{1}_t\right) \\ &\quad \cdot g^{-1}\left(g\left(f\left(\left[\frac{\mathcal{W}_1}{t}, \frac{\mathcal{W}_2}{t}\right], \left[\frac{Q_1}{t}, \frac{Q_2}{t}\right]\right)\right)\right) \\ &= t^r \cdot f([\mathcal{W}_1, \mathcal{W}_2], [Q_1/t, Q_2/t]) \\ &= t^r \cdot f([(X_1/Q_1)^{\frac{1}{\rho}}, (X_2/Q_2)^{\frac{1}{\rho}}], [Q_1/t, Q_2/t])\end{aligned} \tag{222}$$

$$
\begin{aligned}
\frac{f(\mathbf{y},\mathbf{p})}{f(\mathbf{x},\mathbf{q})} &= g^{-1}\left(\sum_{i=1}^{2} s_i \cdot g\left(\frac{f(\mathbf{y}^i,\mathbf{p}^i)}{f(\mathbf{x}^i,\mathbf{q}^i)}\right)\right) \\
&= g^{-1}\left(\frac{Q_1\mathcal{W}_1^\rho}{Q_1\mathcal{W}_1^\rho+Q_2\mathcal{W}_2^\rho}\cdot g\left(\frac{a_1^r b_2^r c_2^r d_1^2 v^r}{t^r}\right)+\frac{Q_2\mathcal{W}_2^\rho}{Q_1\mathcal{W}_1^\rho+Q_2\mathcal{W}_2^\rho}\cdot g\left(\frac{a_2^r b_1^r c_1^r d_2^r v^r}{t^r}\right)\right) \\
&= g^{-1}\left(X_1\cdot g\left(\frac{a_1^r d_1^r v^r}{b_1^r c_1^r u^r}\right)+X_2\cdot g\left(\frac{a_2^r d_2^r v^r}{b_2^r c_2^r u^r}\right)\right) \\
&= g^{-1}\left(X_1\cdot g\left(\left(\frac{X_1}{Q_1K}\right)^{\frac{r}{\rho}}\right)+X_2\cdot g\left(\left(\frac{X_2}{Q_2K}\right)^{\frac{r}{\rho}}\right)\right) \qquad (218)
\end{aligned}
$$

where the entire system is partitioned into t sub-systems[5], each with a resource allocation vector $[\mathcal{W}_1/t, \mathcal{W}_1/t]$. The last step uses (212) and (213). Again, using Axiom $2'$ and (211), we derive

$$
\begin{aligned}
f([\mathbf{y}^1,\mathbf{y}^2],\mathbf{p}) &= f\left(\mathbf{1}_{a_1b_2c_2d_1v+a_2b_1c_1d_2v}\right) \\
&= (a_1b_2c_2d_1v+a_2b_1c_1d_2v)^r \qquad (223)
\end{aligned}
$$

Combining (218), (222), and (223), we derive

$$
\begin{aligned}
&\frac{\left((X_1/Q_1)^{\frac{1}{\rho}}+(X_2/Q_2)^{\frac{1}{\rho}}\right)^r}{f([(X_1/Q_1)^{\frac{1}{\rho}},(X_2/Q_2)^{\frac{1}{\rho}}],[Q_1/t,Q_2/t])}K^{-\frac{r}{\rho}} \\
&= \frac{(a_1b_2c_2d_1v+a_2b_1c_1d_2v)^r}{f([(X_1/Q_1)^{\frac{1}{\rho}},(X_2/Q_2)^{\frac{1}{\rho}}],[Q_1/t,Q_2/t])\cdot t^r} \\
&= g^{-1}\left(\sum_{i=1}^{2} X_i\cdot g\left(\left(\frac{X_i}{Q_iK}\right)^{\frac{r}{\rho}}\right)\right) \qquad (224)
\end{aligned}
$$

For different $K>0$, (224) further implies that

$$
\begin{aligned}
&g^{-1}\left(\sum_{i=1}^{2} X_i\cdot g\left(\left(\frac{X_i}{Q_iK}\right)^{\frac{r}{\rho}}\right)\right) \\
&= K^{\frac{r}{\rho}}\cdot g^{-1}\left(\sum_{i=1}^{2} X_i\cdot g\left(\left(\frac{X_i}{Q_iK}\right)^{\frac{r}{\rho}}\right)\right) \qquad (225)
\end{aligned}
$$

for rational $X_1, X_2, Q_1, Q_2, K > 0$ and arbitrary ρ, r. Due to Axiom $1'$, it is easy to show that (225) also holds for real $X_1, X_2, Q_1, Q_2, K > 0$. It follows from [35] and the argument in the proof C.1 that there exist only two possible generator functions:

$$
g(y) = z\cdot\log(y) \qquad (226)
$$

where $z \neq 0$ is an arbitrary constant, and

$$
g(y) = \frac{y^\beta - 1}{c} \qquad (227)
$$

where $\beta \in \mathbb{R}$ is a constant exponent.

(Existence of fairness measures for logarithm generator.) We derive fairness measure for the logarithm generator function in (226) and show that it satisfies Axioms $1'$-$4'$. By choosing $K = u/v = 1$ and a change of variable $x_i = X_i^{\frac{1}{\rho}} = a_i/b_i$ in (224), we plug in the logarithm generator (226) and find

$$
\begin{aligned}
f([x_1,x_2],[q_1,q_2]) &= \frac{(x_1+x_2)^r}{g^{-1}\left(\sum_{i=1}^2 \frac{q_i x_i^\rho}{q_1x_1^\rho+q_2x_2^\rho}\cdot g(x_i^r)\right)} \\
&= \prod_{i=1}^{2}\left(\frac{x_i}{x_1+x_2}\right)^{-r\frac{q_i x_i^\rho}{q_1x_1^\rho+q_2x_2^\rho}} \qquad (228)
\end{aligned}
$$

To generate fairness measure for $n > 2$ users, we use induction assume that for $n = k$ users, we have

$$
f(\mathbf{x},\mathbf{q}) = \prod_{i=1}^{k}\left(\frac{x_i}{\sum_{j=1}^k x_j}\right)^{-r\frac{q_i x_i^\rho}{\sum_j q_j x_j^\rho}} \qquad (230)
$$

which holds for $n = k = 2$. Let $u_k = \sum_{i=1}^k x_i$ and $Q_k = \sum_{i=1}^k q_i$ For $n = k+1$, we apply (63) and show that (229) holds. The first step of (229) uses Axiom $2'$ with $f(x_{k+1}, q_{k+1}) = f(1,1) = 1$. Comparing the last step of (229) to (230) with $n = k+1$, induction of fairness measure from $n = k$ to $n = k+1$ holds if and only if

$$
Q_k u_k^\rho = (\sum_{i=1}^k q_i)(\sum_{i=1}^k x_i)^\rho = \sum_{i=1}^k q_i x_i^\rho. \qquad (231)
$$

Therefore, we must have $\rho = 0$. Fairness measure generated by logarithm generator in (226) can only have the following form

$$
f(\mathbf{x},\mathbf{q}) = \prod_{i=1}^{k}\left(\frac{x_i}{\sum_{j=1}^k x_j}\right)^{-rq_i} \qquad (232)
$$

which is (65) in Theorem 10.

The proof so far assumes positive fairness measure $f > 0$. To determine the sign of f, we add $\mathrm{sign}(-r)$ in (232), apply Axiom 5, and find that for $r > 0$

$$
f([0,1],[1,1]) = 0 < 2^r = f\left(\left[\frac{1}{2},\frac{1}{2}\right],[1,1]\right), \qquad (233)
$$

and for $r < 0$:

$$
f([0,1],[1,1]) = -\infty < -2^r = f\left(\left[\frac{1}{2},\frac{1}{2}\right],[1,1]\right). \qquad (234)
$$

Therefore, Axiom $5'$ is satisfied. We derive the fairness measure in (65).

It remains to prove that fairness measure (232) satisfies Axioms $1'$-$4'$ with arbitrary $r \in \mathbb{R}$. It is easy to see that Axioms $1'$-$3'$ are satisfied. To verify Axiom $4'$, we use the

[5]Although Axiom $4'$ is defined for partitioning into 2 subsystems, it can be extended to partitioning into t identical sub-systems, using the same induction in proof of Lemma 2.

$$
\begin{aligned}
&f([x_1,\ldots,x_k,x_{k+1}],[q_1,\ldots,q_k,q_{k+1}])\\
&= f\left(\left[\sum_{i=1}^{k} x_i, x_{k+1}\right], [\sum_{i=1}^{k} q_i, q_{k+1}]\right)\cdot [f([x_1,\ldots,x_k],[q_1,\ldots,q_k])]^{\frac{Q_k u_k^\rho}{Q_k u_k^\rho + q_{k+1}x_{k+1}^\rho}}\cdot 1\\
&= \left(\frac{\sum_{i=1}^{k} x_i}{\sum_{i=1}^{k+1} x_i}\right)^{\frac{-rQ_k u_k^\rho}{Q_k u_k^\rho + q_{k+1}x_{k+1}^\rho}}\cdot\left(\frac{x_{k+1}}{\sum_{i=1}^{k+1} x_i}\right)^{\frac{-rq_{k+1}x_{k+1}^\rho}{Q_k u_k^\rho + q_{k+1}x_{k+1}^\rho}}\cdot\prod_{i=1}^{k}\left(\frac{x_i}{\sum_{j=1}^{k} x_j}\right)^{-r\frac{q_i x_i^\rho}{\sum_j q_j x_j^\rho}\cdot\frac{Q_k u_k^\rho}{Q_k u_k^\rho + q_{k+1}x_{k+1}^\rho}}
\end{aligned}
\tag{229}
$$

same argument (137) in the proof of Theorems 1, 2, and 3 for logarithm generator function. Since (137) is independent of the choice of s_i, we have

$$
\frac{f(\mathbf{x},\mathbf{q})}{f(\mathbf{y},\mathbf{p})} = g^{-1}\left(\sum_{i=1}^{2} s_i\cdot g\left(\frac{f(\mathbf{x}^i,\mathbf{q}_i)}{f(\mathbf{y}^i,\mathbf{p}_i)}\right)\right) \tag{235}
$$

which establishes Axiom $4'$ for any partition. This completes the proof for logarithm generators.

(Existence of fairness measures for power generator.) We derive fairness measures for power generator function in (227) and show that it satisfies Axioms $1'$-$5'$. By choosing $K = u/v = 1$ and a change of variable $x_i = X_i^{\frac{1}{\rho}} = a_i/b_i$ in (224), we plug in the power generator (227) and find

$$
\begin{aligned}
f([x_1,x_2],[q_1,q_2]) &= \frac{(x_1+x_2)^r}{g^{-1}\left(\sum_{i=1}^{2}\frac{q_i x_i^\rho}{q_1x_1^\rho+q_2x_2^\rho}\cdot g(x_i^r)\right)}\\
&= \frac{(x_1+x_2)^r(q_1x_1^\rho+q_2x_2^\rho)^{\frac{1}{\beta}}}{(q_1x_1^{\rho+r\beta}+q_2x_2^{\rho+r\beta})^{\frac{1}{\beta}}}
\end{aligned}
\tag{236}
$$

To derive the fairness measure for three users, we consider two different partitions of the resource allocation vector $[x_1,x_2,x_3]$ as $[x_1,x_2],[x_3]$ and $[x_1],[x_2,x_3]$. Let $q_{1,2}=q_1+q_2$, $q_{2,3}=q_2+q_3$, and $u=x_1+x_2+x_3$ be auxiliary variables. Using the construction in (63), we obtain two equivalent form of the fairness measure for $n=3$ users in (237) and (238).

As in Axiom $4'$, the fairness measure is independent to partition. Similar to the proof steps in C.1, we will consider two different partitions of three uses to derive a sufficient and necessary condition on parameters ρ, r, and β. We find that (237) and (238) should be equivalent for all $x_1,x_2,x_3\ge 0$. Comparing the terms in (237) and (238), we find that (237) and (238) are equivalent if and only if

$$
\rho = -r\cdot\beta. \tag{239}
$$

Similar to the proof in C.1, (239) is a sufficient and necessary for fairness measures to be independent of partition for $n=3$. Use (239) to remove ρ in (237) and (238) and simply the terms. We derive the expression of fairness measure for three users as follows

$$
f([x_1,x_2,x_3],[q_1,q_2,q_3]) = \left[\sum_{i=1}^{3} q_i\left(\frac{x_i}{\sum_j x_j}\right)^{-r\beta}\right]^{\frac{1}{\beta}} \tag{240}
$$

where $\beta\neq 0$ and r can be arbitrary.

To generate fairness measure for $n>3$ users, we assume that (231) holds for $n=k$ users. We define axillary variables $u_k=\sum_{i=1}^{k} x_i$, $Q_k=\sum_{i=1}^{k} q_i$, and $\xi_k=\sum_{i=1}^{k} q_i x_i^{-r\beta}$. Let $\mathbf{x}=[x_1,\ldots,x_k]$ and $\mathbf{q}=[q_1,\ldots,q_k]$. For $n=k+1$, we apply (3) in Lemma 1 and use (240) to derive

$$
\begin{aligned}
&f([x_1,\ldots,x_k,x_{k+1}],[q_1,\ldots,q_k,q_{k+1}])\\
&= f([u_k,x_{k+1}],[Q_k,q_{k+1}])\cdot\left(\frac{Q_k u_k^\rho f(\mathbf{x},\mathbf{q})^\beta + q_{k+1}x_{k+1}^\rho}{Q_k u_k^\rho + q_{k+1}x_{k+1}^\rho}\right)^{\frac{1}{\beta}}\\
&= \frac{(Q_k u_k^\rho + q_{k+1}x_{k+1}^\rho)^{\frac{1}{\beta}}}{u_{k+1}^{-r}}\cdot\left(\frac{u_k^{\rho-r\beta}\xi_k + q_{k+1}x_{k+1}^\rho}{Q_k u_k^\rho + q_{k+1}x_{k+1}^\rho}\right)^{\frac{1}{\beta}}\\
&= \frac{(\sum_{i=1}^{k+1} q_i x_i^{-r\beta})^{\frac{1}{\beta}}}{u_{k+1}^{-r}}\\
&= \left[\sum_{i=1}^{k+1} q_i\left(\frac{x_i}{\sum_j x_j}\right)^{-r\beta}\right]^{\frac{1}{\beta}}
\end{aligned}
\tag{241}
$$

where the third step uses $\rho - r\beta = 0$ and the last step uses $u_k=\sum_{i=1}^{k} x_i$. Therefore, (241) holds for $n=k+1$ due to induction.

Notice that (241) is almost (66), except for a term $\text{sign}(-r(1+r\beta))$. The term is added in order to satisfy Axiom $5'$. To show this, we first consider the case of $r\beta>0$, which implies $(1+r\beta)>0$. For $r>0$ and $\beta>0$, we find

$$
f([0,1],[1,1]) = -\infty < -2^r = f(\left[\frac{1}{2},\frac{1}{2}\right],[1,1]), \tag{242}
$$

and for $r<0$ and $\beta<0$,

$$
f([0,1],[1,1]) = 0 < 2^r = f(\left[\frac{1}{2},\frac{1}{2}\right],[1,1]). \tag{243}
$$

Next, when $r\beta<0$, we have

$$
\begin{aligned}
f([0,1],[1,1]) &= -\text{sign}(1+r\beta)\cdot q_2^{\frac{1}{\beta}}\\
&= -\text{sign}(1+r\beta)\cdot 2^{-\frac{1}{\beta}},
\end{aligned}
\tag{244}
$$

where the last step uses the statement of equal weights $q_1=q_2=1/2$ in Axiom $5'$. We also obtain

$$
f(\left[\frac{1}{2},\frac{1}{2}\right],[1,1]) = -\text{sign}(1+r\beta)\cdot 2^r \tag{245}
$$

Comparing (244) and (245) for $r>-1/\beta$ and $r<-1/\beta$, we conclude that $f(\left[\frac{1}{2},\frac{1}{2}\right],[1,1]) \ge f([0,1],[1,1])$ holds for $r\beta<0$. Therefore, we derive fairness measure (66). It is straightforward to check that fairness measure in (142) satisfies Axioms 1-4. This completes the proof for power generators.

$$f([x_1,x_2,x_3],[q_1,q_2,q_3]) \tag{237}$$
$$= f([x_1+x_2,x_3],[q_1+q_2,q_3])\cdot g^{-1}\left(\frac{(q_{1,2})(x_1+x_2)^{\rho}\cdot g\left(f([x_1,x_2],[q_1,q_2])\right)+q_3x_3^{\rho}\cdot g\left(1\right)}{(q_{1,2})(x_1+x_2)^{\rho}+q_3x_3^{\rho}}\right)$$
$$= \frac{u^r((q_{1,2})(x_1+x_2)^{\rho}+q_3x_3^{\rho})^{\frac{1}{\beta}}}{((q_{1,2})(x_1+x_2)^{\rho+r\beta}+q_3x_3^{\rho+r\beta})^{\frac{1}{\beta}}}\cdot\left(\frac{(q_{1,2})(x_1+x_2)^{\rho}\cdot\frac{(x_1+x_2)^{r\beta}(q_1x_1^{\rho}+q_2x_2^{\rho})}{(q_1x_1^{\rho+r\beta}+q_2x_2^{\rho+r\beta})}+q_3x_3^{\rho}\cdot}{(q_{1,2})(x_1+x_2)^{\rho}+q_3x_3^{\rho}}\right)$$

$$f([x_1,x_2,x_3],[q_1,q_2,q_3]) \tag{238}$$
$$= f([x_1+x_2,x_3],[q_1+q_2,q_3])\cdot g^{-1}\left(\frac{(q_{2,3})(x_2+x_3)^{\rho}\cdot g\left(f([x_2,x_3],[q_2,q_3])\right)+q_1x_1^{\rho}\cdot g\left(1\right)}{(q_{2,3})(x_2+x_3)^{\rho}+q_1x_1^{\rho}}\right)$$
$$= \frac{u^r((q_{2,3})(x_2+x_3)^{\rho}+q_1x_1^{\rho})^{\frac{1}{\beta}}}{((q_{2,3})(x_2+x_3)^{\rho+r\beta}+q_1x_1^{\rho+r\beta})^{\frac{1}{\beta}}}\cdot\left(\frac{(q_{2,3})(x_2+x_3)^{\rho}\cdot\frac{(x_2+x_3)^{r\beta}(q_2x_2^{\rho}+q_3x_3^{\rho})}{(q_2x_2^{\rho+r\beta}+q_3x_3^{\rho+r\beta})}+q_1x_1^{\rho}\cdot}{(q_{2,3})(x_2+x_3)^{\rho}+q_1x_1^{\rho}}\right)$$

27) Proof of Corollary 21: For $n=2$ users, the weighted symmetry is immediate from (222), which implies

$$f([x_1,x_2],[q_1,q_2]) = f([x_2,x_1],[q_2,q_1]). \tag{246}$$

We use induction to prove weighted symmetry for $n>2$ users. The remaining proof is same as that of Corollary 1.

28) Proof of Corollaries 22, 23, and 24: We prove Corollaries 22, 23, and 24 for (66) when $f>0$ is positive. The proof for $f<0$ is the same. Let $\kappa = \sum_{i=1}^{n} q_i\left(\frac{x_i}{\sum_j x_j}\right)^{-r\beta}$ be an auxiliary function. We have

$$\frac{\partial f(\mathbf{x},\mathbf{q})}{\partial x_i} = \frac{-r\beta\cdot\kappa^{\frac{1}{\beta}-1}}{(\sum_j x_j)^{-r}}\left[q_i x_i^{-r\beta-1} - \frac{\sum_j q_j x_j^{-r\beta}}{\sum_j x_j}\right] \tag{247}$$

Because $\kappa>0$ is positive, $\frac{\partial f(\mathbf{x},\mathbf{q})}{\partial x_i}$ has a single root at

$$x_i = q_i^{\frac{1}{r\beta+1}}\cdot\bar{x} = q_i^{\frac{1}{r\beta+1}}\cdot\left(\frac{\sum_j x_j}{\sum_j q_j x_j^{-r\beta}}\right)^{\frac{1}{\beta}}. \tag{248}$$

If a small amount of resource Δx is moved from user i to user j, we have

$$\Delta f = \frac{-r\beta\kappa^{\frac{1}{\beta}-1}}{(\sum_m x_m)^{-r}}\left[q_i x_i^{-r\beta-1} - q_j x_j^{-r\beta-1}\right]\cdot\Delta x + o(\Delta x)$$

which is positive if and only if $x_i/q_i^{\frac{1}{r\beta+1}} < x_j/q_j^{\frac{1}{r\beta+1}}$. This proves Corollary 22. From (247) and (248), it is straightforward to show that

$$\frac{\partial f(\mathbf{x},\mathbf{q})}{\partial x_i} > 0, \text{ if } x_i > q_i^{\frac{1}{r\beta+1}}\cdot\bar{x}, \tag{249}$$
$$\frac{\partial f(\mathbf{x},\mathbf{q})}{\partial x_i} < 0, \text{ if } x_i < q_i^{\frac{1}{r\beta+1}}\cdot\bar{x}, \tag{250}$$

which proves Corollary 24. Further, fairness measure is maximized by resource allocation $[q_1^{\frac{1}{r\beta}}, q_2^{\frac{1}{r\beta+1}},\ldots,q_n^{\frac{1}{r\beta+1}}]$. Corollary 23 is proven.

29) Proof of Corollary 25: For $\beta\in(-\infty,-1)$, $f_\beta(\mathbf{x})$ is equivalent to inverse of p-norm with $p=-\beta$. Monotonicity of $f_\beta(\mathbf{x},\mathbf{q})$ is straightforward from monotonicity of p-norm. For $\beta\in(-1,0)$, we consider two axillary functions $h_1(y)=y^{-\beta}$ and $h_2(y)=y^{1/\beta}$, which are both monotonically increasing over β for $y\in(0,1)$. Therefore, $f_\beta(\mathbf{x},\mathbf{q}) = -h_2(\sum_i q_i\cdot h_1(x_i))$ is monotonically decreasing.

To prove monotonicity of $f_\beta(\mathbf{x},\mathbf{q})$ for $\beta\in(0,\infty)$. Consider two different values $0<\beta_1\le\beta_2$. We define the a function $\phi(y)=y^{\frac{\beta_2}{\beta_1}}$ for $y\in\mathbb{R}_+$. Since $\beta_2/\beta_1\ge 1$, the function $\phi(y)$ is convex in y. Therefore, we have

$$\begin{aligned} f_{\beta_2}(\mathbf{x}) &= -\left[\sum_{i=1}^{n}\left(q_i\cdot\frac{x_i}{\sum_j x_j}\right)^{-\beta_2}\right]^{\frac{1}{\beta_2}} \\ &= -\left[\sum_{i=1}^{n} q_i\cdot\phi\left(\left(\frac{x_i}{\sum_j x_j}\right)^{-\beta_1}\right)\right]^{\frac{1}{\beta_2}} \\ &= \left[\phi\left(\sum_{i=1}^{n} q_i\cdot\left(\frac{x_i}{\sum_j x_j}\right)^{-\beta_1}\right)\right]^{\frac{1}{\beta_2}} \\ &= \left[\sum_{i=1}^{n}\left(\frac{x_i}{\sum_j x_j}\right)^{-\beta_1}\right]^{\frac{1}{\beta_1}} \\ &= f_{\beta_1}(\mathbf{x},\mathbf{q}), \end{aligned} \tag{251}$$

where the third step follows from Jensen's inequality and $\beta_2>0$. This shows that $f_\beta(\mathbf{x},\mathbf{q})$ is monotonically decreasing on $\beta\in(0,\infty)$. This completes the proof.

30) Proof of Theorem 13: We need to prove that (81) is sufficient and necessary. We first show that if $F(\mathbf{x})$ satisfies Axioms $1''$-$4''$, its normalization, denoted by

$$f(\mathbf{x}) = F(\mathbf{x})\cdot\left(\sum_i x_i\right)^{-\frac{1}{\lambda}}, \tag{252}$$

must satisfy Axioms 1-5 for s_i in (7) and satisfying Axioms $1'$-$5'$ for s_i in (62). Then, we verify that fairness $F(\mathbf{x})$ given by (252) indeed satisfies Axioms $1''$-$4''$.

($\Leftarrow$) The continuity of $f(\mathbf{x})$ follows directly from that of $F(\mathbf{x})$ in Axioms $1''$. Let $z>0$ be a positive real number and $\mathbf{v}$ be a vector of arbitrary length. To prove homogeneity, we make use of Axiom $3''$ for $\mathbf{x}=[z\mathbf{v},z\mathbf{v}]$, $\mathbf{y}=[1,1]$, and $t=1$,

$$\begin{aligned}\frac{F(z\mathbf{v},z\mathbf{v})}{f(1,1)} &= g^{-1}\left(s_1\cdot g\left(\frac{F(z\mathbf{v})}{F(1)}\right)+s_2\cdot g\left(\frac{F(z\mathbf{v})}{F(1)}\right)\right) \\ &= \frac{F(z\mathbf{v})}{F(1)}\end{aligned} \tag{253}$$

where the second step uses the fact that $\sum_i s_i = 1$. Next, we apply Axiom $3''$ for $\mathbf{x}=[\mathbf{v},\mathbf{v}]$, $\mathbf{y}=[1,1]$, and $\mathbf{z}=[z,z]$,

$$\begin{aligned}\frac{F(z\mathbf{v},z\mathbf{v})}{F(t,t)} &= g^{-1}\left(s_1\cdot g\left(\frac{F(\mathbf{v})}{F(1)}\right)+s_2\cdot g\left(\frac{F(\mathbf{v})}{F(1)}\right)\right) \\ &= \frac{F(\mathbf{v})}{F(1)}\end{aligned} \tag{254}$$

and for $\mathbf{x}=[t,t]$, $\mathbf{y}=[1,1]$, and $t=1$,

$$\begin{aligned}\frac{F(t,t)}{f(1,1)} &= g^{-1}\left(s_1\cdot g\left(\frac{F(t)}{F(1)}\right)+s_2\cdot g\left(\frac{F(t)}{F(1)}\right)\right) \\ &= \frac{F(t)}{F(1)}\end{aligned} \tag{255}$$

Comparing the above three equations, (253), (254), and (255), we have

$$F(z\mathbf{v}) = F(z)\cdot F(\mathbf{v}). \tag{256}$$

When $\mathbf{v}$ is a scalar, using the result in [34], (256) implies that $\log F(z) = \frac{1}{\lambda}\log(z)$ must be a logarithmic function with an exponent $\frac{1}{\lambda}$. We have

$$F(z\mathbf{v}) = z^{\frac{1}{\lambda}}\cdot F(\mathbf{v}), \tag{257}$$

which is a homogenous function of order $\frac{1}{\lambda}$. Therefore, normalization $f(\mathbf{x})$ in (252) is a homogenous function of degree zero and satisfies Axiom 2 and Axiom $2'$.

Using the homogeneity property (257) and Axioms $2''$, we obtain

$$\begin{aligned}\lim_{n\to\infty}\frac{f(\mathbf{1}_{n+1})}{f(\mathbf{1}_n)} &= \lim_{n\to\infty}\frac{F(\mathbf{1}_{n+1})}{F(\mathbf{1}_n)}\cdot\left(1+\frac{1}{n}\right)^{-\frac{1}{\lambda}} \\ &= \lim_{n\to\infty}\frac{F(\mathbf{1}_{n+1})}{F(\mathbf{1}_n)} \\ &= 1.\end{aligned} \tag{258}$$

This proves Axiom 3 and Axiom $3'$. It is easy to see that Axiom 4 and Axiom $4'$) are special cases of Axiom $3''$ for $t=1$. From the homogeneity property (257) and Axiom $3''$, we find that

$$f(\tfrac{1}{2},\tfrac{1}{2}) = F(\frac{1}{2},\frac{1}{2})\cdot 1^{-\frac{1}{\lambda}} \ge F(1,0) = f(1,0), \tag{259}$$

which proves Axiom 5 and Axiom $5'$, i.e., starvation is no more fair than equal allocation.

($\Leftarrow$) Suppose that $F(x)$ is homogenous function of order $\frac{1}{\lambda}$ and is given by

$$F(\mathbf{x}) = f(\mathbf{x})\cdot\left(\sum_i x_i\right)^{\frac{1}{\lambda}}. \tag{260}$$

where $f(\mathbf{x})$ satisfies Axioms 1-5 for s_i in (7) and Axioms $1'$-$5'$ for s_i in (62). We need to prove that such $F(x)$ satisfies Axioms $1''$-$4''$.

Axiom $1''$ is a direct result from continuity of $f(\mathbf{x})$ in Axiom 1 and Axiom $1\prime$. Axiom $2''$ can be obtained from Axiom 3 and Axiom $3'$ due to

$$\begin{aligned}\lim_{n\to\infty}\frac{F(\mathbf{1}_{n+1})}{F(\mathbf{1}_n)} &= \lim_{n\to\infty}\frac{f(\mathbf{1}_{n+1})}{f(\mathbf{1}_n)}\cdot\left(1+\frac{1}{n}\right)^{\frac{1}{\lambda}} \\ &= \lim_{n\to\infty}\frac{f(\mathbf{1}_{n+1})}{f(\mathbf{1}_n)} \\ &= 1.\end{aligned} \tag{261}$$

To prove Axiom $3''$, from Axiom 4 and Axiom $4'$, we have

$$\frac{f(\mathbf{x})}{f(\mathbf{y})} = g^{-1}\left(\sum_{i=1}^{2} s_i\cdot g\left(\frac{f(\mathbf{x}^i)}{f(\mathbf{y}^i)}\right)\right) \tag{262}$$

where $\mathbf{x}=[\mathbf{x}^1,\mathbf{x}^2]$ and $\mathbf{y}=[\mathbf{y}^1,\mathbf{y}^2]$ are two resource allocation vectors satisfying $w(\mathbf{x}^i)=w(\mathbf{y}^i)$. We conclude that

$$\frac{f(\mathbf{x}^i)}{f(\mathbf{y}^i)} = \frac{w^{\frac{1}{\lambda}}(\mathbf{x}^i)\cdot f(\mathbf{y}^i)}{w^{\frac{1}{\lambda}}(\mathbf{x}^i)\cdot f(\mathbf{y}^i)} = \frac{F(\mathbf{x}^i)}{F(\mathbf{y}^i)} \tag{263}$$

where the last step uses (260). Similarly, we derive

$$\frac{f(\mathbf{x})}{f(\mathbf{y})} = \frac{t^{\frac{1}{\lambda}}\cdot w^{\frac{1}{\lambda}}(\mathbf{x})\cdot f(\mathbf{y}^i)}{t^{\frac{1}{\lambda}}\cdot w^{\frac{1}{\lambda}}(\mathbf{x}^i)\cdot f(\mathbf{y})} = \frac{F(t\mathbf{x}^i)}{F(t\mathbf{y}^i)}. \tag{264}$$

Plugging (263) and (264) into (262), we prove (80) and Axiom $3''$. Finally, Axiom $4''$ follows from (260), Axioms 5, and Axioms $5'$:

$$F(\frac{1}{2},\frac{1}{2}) = f(\frac{1}{2},\frac{1}{2})\cdot 1^{\frac{1}{\lambda}} \ge f(1,0) = F(1,0). \tag{265}$$

Therefore, $F(\mathbf{x})$ given by (260) satisfies Axioms $1''$-$4''$. This completes the proof.

References

[1] M.A. Marson and M. Gerla, "Fairness in Local Computing Networks," *in Proceeding of IEEE ICC*, 1982.

[2] J.W. Wong and S.S. Lam, "A Study Of Fairness In Packet Switching Networks," *in Proceedings of IEEE ICC*, 1982.

[3] R. Jain, D. Chiu, and W. Hawe, "A Quantitative Measure of Fairness And Discrimination for Resource Allocation in Shared Computer System," *DEC Technical Report 301*, 1984.

[4] C. E. Koksal, H. I. Kassab, and H. Balakrishnan, " An Analysis of Short-term Fairness in Wireless Media Access Protocols," *in Proceedings of ACM SIGMETRICS*, 2000.

[5] M. Bredel and M. Fidler, "Understanding Fairness and Its Impact on Quality of Service in IEEE 802.11," *in Proceedings of IEEE INFOCOM*, 2009.

[6] M. Dianati, X. Shen, and S. Naik, "A New Fairness Index for Radio Resource Allocation in Wireless Networks," *in Proceedings of WCNC*, 2005.

[7] F. P. Kelly, A. Maulloo and D. Tan, "Rate Control in Communication Networks: Shadow Prices, Proportional Fairness and Stability," *Journal of the Operational Research Society*, vol. 49, pp. 237-252, 1998.

[8] J. Mo and J.Walrand, "Fair End-to-end Window-based Congestion Control," *IEEE/ACM Transactions Networking*, vol. 8, no. 5, pp. 556-567, Oct. 2000.

[9] M. Uchida and J. Kurose, " An Information-Theoretic Characterization of Weighted α-Proportional Fairness," *in Proceedings of IEEE INFOCOM*, 2009.

[10] T. Bonald and L. Massoulie, "Impact of Fairness on Internet Performance," *in Proceedings of ACM Sigmetrics*, 2001.

[11] L. Massoulie and J. Roberts, "Bandwidth Sharing: Objectives and Algorithms," *IEEE/ACM Transactions Networking*, vol. 10, no. 3, pp. 320-328, Jun. 2002.

[12] A. Tang, J. Wang and S. Low, "Counter-intuitive Behaviors in Networks under End-to-end Control," *IEEE /ACM Transactions on Networking*, vol. 14, no. 2, pp. 355-368, April 2006.

[13] A. Kolmogoroff, "Sur la notion de la moyenne", *Atti della R.Academia nazionale dei Lincei*, vol. 12, 1930.

[14] J. Aczél, "On Mean Values," *Bullitin of American Mathematics Socioty*, vol. 54, pp. 392-400, 1948.

[15] A. W. Marshall and I. Olkin, "Inequalites: Theory of Majorization and Its Applications," *Academic Press*, 1979.

[16] R. Bhargava, A. Goelt, and A. Meyerson, "Using Approximate Majorization to Characterize Protocol Fairness," *in Proceedings of ACM Sigmetrics*, 2001.

[17] C. W. Gini, "Variability and Mutability, Contribution to The Study of Statistical Distribution and Relaitons", *Studi Economico-Giuricici della R*, Universita de Cagliari, 1912.

[18] R. Aaberge, "Axiomatic Characterization of the Gini Coefficient and Lorenz Curve Orderings", *Journal of Economic Theory* vol.101, pp.115-132, 2001.

[19] G. J. Foschini and Z. Miljanic, "A Simple Distributed Autonomous Power Control Algorithm and its Convergence", *IEEE Transactions on Vehicular Technology*, no. 4, pp. 641-646, April 1993.

[20] R. Mazumdar, L. G. Mason, and C. Douligeris, "Fairness in Network Optimal Control: Optimality of Product Flow Forms." *IEEE Transactions on Communications*, vol. 39, no. 5, May 1991.

[21] R. Srikant, "The Mathematics of Internet Congestion Control." *Birkhauser*, 2004

[22] S.H. Low, "A Duality Model of TCP and Queue Management Algorithms", *IEEE/ACM Transactions on Networking*, vol. no. 4, pp. 525-536, Aug 2003.

[23] M. Chiang, P. Hande, T. Lan, and C.W. Tan. "Power Control in Cellular Networks", *Foundations and Trends in Networking*, Now Publishers Inc, 2008.

[24] T. Starr, M. Sorbora, J.M. Cioffi, and P.J. Silverman, *DSL Advances*, Prentice Hall, 2003.

[25] P. Tsiaflakis., Y. Yi, M. Chiang, and M. Moonen, "Dynamic Spectrum Management for Fast and Fair Green DSL", *submitted to IEEE Trasanctions on Signal Processing*, 2010.

[26] A. Renyi, "On Measures of Information and Entropy", *in Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability*, 1960.

[27] D. K. Fadeev, "Zum Begriff der Entrlopie ciner endlichen Wahrseheililhlikeitsschemas," *Arbeiten zur Informationstheorie I, Berlini, Deutscher Verlag der Wissensehaf ten*, pp. 85-90, 1957.

[28] R. Renner, "Security of Quantum Key Distribution", *PhD Dissertation at ETH*, no. 16242, September 2005.

[29] R. Renner and S. Wolf, "Smooth Renyi Entropy and Applications", *in Proceedings of IEEE ISIT*, 2004.

[30] H. Moulin, "Axioms of Cooperative Decision Making", *Cambridge University Press*, 1991.

[31] J. F. Nash, "The Bargaining Problem", *Econometrica*, vol. 18, pp. 155-162, 1950.

[32] L. S. Shapley, "A Value for n-person Games" (in "Contributions to the Theory of Games", Volumn II, by H.W. Kuhn and A.W. Tucker, editors), *Annals of Mathematical Studies* vol. 28, pp. 307-317, 1953.

[33] J. Rawls, "A Theory of Justice", *Cambridge: Harvard University Press*, 1971.

[34] P. Erdos, "On The Distribution Function of Additive Functions", *Annals of Mathematics*, Vol. 47 pp. 1-20, 1946.

[35] G. H. Hardy, J.E. Littlewood, G. Polya, "Inequalities", *Cambridge Univ Press*, 1973.

[36] J. C. A. van der Lubbet, "An Axiomatic Theory of Heterogeneity and Homogeneity", *Metrika*, vol. 33, pp. 223-245, 1986.

[37] M. Osborne and A. Rubinstein, "A Course in Game Theory", *The MIT Press*, 1994.

[38] S. Boyd and L. Vandenberghe, "Convex optimization", *Cambridge University Press*, 2005.

[39] G. Wang and M. Jiang, "Axiomatic characterization of nonlinear homomorphic means", *Journal of Mathematical Analysis and Applications*, vol. 303, pp. 350-363, 2005.

[40] T. Lan, D. Kao, M. Chiang, and A. Sabharwal "An Axiomatic Approach to Fairness in Network Resource Allocation", *in Proceedings of IEEE INFOCOM*, 2010.

[41] H. Theil, "Economics and Information Theory". *Chicago: Rand McNally and Company*, 1967.

[42] A. B. Atkinson, "On the measurement of economic inequality". *Journal of Economic Theory*, Vol. 2, no. 3, pp. 244-263, 1970.

[43] J. E. Foster and A. Sen, "On Economic Inequality", *expanded edition with annexe*, ISBN 0-19-828193-5, 1996.

[44] E. Altman, K. Avrachenkov, and S. Ramanath, "Multiscale Fairness and its Application to Dynamic Resource Allocation in Wireless Networks", *INRIA Technical Report*, 2010.

[45] N. G. Makiw, M. Weinzierl, and D. Yagan, "Optimal taxation in theory and practice", *Journal of Economic Perspectives*, vol. 23, pp. 147-174, 2009.

[46] G. Jasso, "A new theory of distributive justice", *American Sociological Review*, vol. 45, pp. 3-32, 1980.

[47] G. Jasso, "How much injustice is there in the world? Two new justice indexes", *American Sociological Review*, vol. 64, pp. 133-158, 1999.

[48] G. Jasso, "Theoretical unification in justice and beyond", *Social Justice Research*, vol. 20, pp. 336-371, 2007.

[49] G. Jasso, "Two types of inequality: Inequality between persons and inequality between sub-groups", *Sociological Methods and Research*, vol. 37, pp. 31-74, 2008.

[50] J. S. Adams, "Toward an understanding of inequity", *Journal of Abnormal and Social Psychology*, vol. 67, pp. 422-436, 1963.

[51] E. A. Lind and T. R. Tyler, "The Social Psychology of Procedural Justice", Plenum Press, 1988.

[52] G. S. Leventhal, "What should be done with equity theory? New approaches to the study of fairenss in social relationships", *Social Exchanges: Advances in Theory and Research*, Plenum Press 1980.

[53] G. S. Leventhal, J. Karuza, and W. R. Fry, "Beyond fairness: A theory of allocation preference", *Justice and Social Interaction*, Springer-Verlag, 1980.

[54] J. Thibaut and L. Walker, "Procedural Justice: A Psychological Analysis", Erlbaum, 1975.

[55] J. Rawls, "Justice as Fairness: A Restatement", *Harvard University Press*, 2001.

[56] N. Rescher, "Distributive Justice: A Constructive Critique of the Utilitarian Theory of Distribution", *Bobbs-Merrill*, 1966.

[57] A. B. Atkinson, "On the measurement of inequality", *Journal of Economic Theory*, vol. 2, pp. 244-263, 1970.

[58] A. Sen and W. Bernard, "Utilitarianism and Beyond", *Cambridge University Press*, 1982.

[59] H. Varian, "Two problems in the theory of fairness", *Journal of Public Economics*, vol. 5, pp. 249-260, 1976.

[60] J. S. Mill, "Principles of Political Economy", 1948.

[61] N. Daniels, "Equality of what: Welfare, resources, or capabilities?", *Philosophy and Phenomenological Research*, vol. 50, pp. 273-306, 1990.

[62] J. Dick, "How to justify a distribution of earning?", *Philosophy and Public Affairs*, vol. 4, pp. 248-272, 1975.

[63] R. Nozick, "Anarchy, State and Utopia", *Basic Books*, 1974.

[64] J. H. Bogart, "Lockean provisos and state of nature theories", *Ethics*, vol. 5, pp. 824-836, 1985.

[65] J. Lamont and C. Favor, "Distributive justice", *Stanford Encyclopedia of Philosophy*, 2007.

[66] R. Dworkin, "What is equality? Part 1: Equality of resoruces", *Philosophy and Public Affairs*, vol. 10, pp. 185-246, 1981.

[67] R. Dworkin, "What is equality? Part 2: Equality of welfare", *Philosophy and Public Affairs*, vol. 10, pp. 283-345, 1981.

[68] M. Fleurbaey, "Economics and economic justice", *Stanford Encyclopedia of Philosophy*, 2008.

[69] S. C. Kolm, "The optimal production of social justice", *Public Economics*, Macmillan, 1969.

[70] S. C. Kolm, "Justice and Equity", MIT Press 1972.

[71] A. K. Sen, "On Economic Inequality", *Clareondon Press*, 1973.

[72] A. Bergson, "A reformulation of certain aspects of welfare economics", *Quarterly Journal of Economics*, vol. 52, pp. 310-334, 1938.

[73] P. A. Samuelson, "Foundations of Economic Analysis", *Harvard University Press*, 1947.

[74] J. Barbanel, "The Geometry of Effcient Fair Division", *Cambridge University Press*, 2005.

[75] A. Ben-Ner, L. Putterman, F. Kong, and D. Magan, "Reciprocity in a two-part dictator game", *Journal of Economic Behavior and Organization*, vol. 53, pp. 333-352, 2004.

[76] G. E. Bolton, E. Katok, and R. Zwick, "Dictator game giving: Rules of fairness versus acts of kindness", *International Journal of Game Theory*, vol. 27, pp. 269-299, 1998.

[77] S. J. Brams, P. H. Edelman, and P. C. Fishburn, "Paradoxes of fair division", *Journal of Philosophy*, vol. 98, pp. 300-314, 2001.

[78] S. J. Brams, M. A. Jones, and C. Klamler, "Better ways to cut a cake", *Notices of the AMS*, vol. 53, pp. 1314-1321, 2006.

[79] S. J. Brams and A. D. Taylor, "An envy-free cake division protocol", *The American Mathematical Monthly*, vol. 102, pp. 9-18, 1995.

[80] S. J. Brams and A. D. Taylor, "Fair Division: From Cake-Cutting to Dispute Resolution", Cambridge University Press, 1996.
[81] M. A. Nowak, K. M. Page, and K. Sigmund, "Fairness versus reason in the ultimatum game", *Science*, vol. 289, pp. 1773-1775, 2000.
[82] F. E. Su, "Sperner's lemma in fair division", *The American Mathematical Monthly*, vol. 106, pp. 930-942, 1999.
[83] A. Alesina and G. M. Angeletos, "Corruption, inequality, and fairness", *Journal of Monetary Economics*, vol. 52, 2005.
[84] A. Alesina and G. M. Angeletos, "Fairenss and redistribution", *American Economic Review*, vol. 95, 2005.
[85] N. G. Makiw, M. Weinzierl, and D. Yagan, "Optimal taxation in theory and practice", *Journal of Economic Perspectives*, vol. 23, pp. 147-174, 2009.