Learning Competitive Equilibria in Exchange Economies with Bandit Feedback

Wenshuo Guo
University of California, Berkeley

Kirthevasan Kandasamy
University of California, Berkeley

Joseph E. Gonzalez
University of California, Berkeley

Michael I. Jordan
University of California, Berkeley

Ion Stoica
University of California, Berkeley

Abstract

The sharing of scarce resources among multiple rational agents is one of the classical problems in economics. In exchange economies, which are used to model such situations, agents begin with an initial endowment of resources and exchange them in a way that is mutually beneficial until they reach a competitive equilibrium (CE). The allocations at a CE are Pareto efficient and fair. Consequently, they are used widely in designing mechanisms for fair division. However, computing CEs requires the knowledge of agent preferences which are unknown in several applications of interest. In this work, we explore a new online learning mechanism, which, on each round, allocates resources to the agents and collects stochastic feedback on their experience in using that allocation. Its goal is to learn the agent utilities via this feedback and imitate the allocations at a CE in the long run. We quantify CE behavior via two losses and propose a randomized algorithm which achieves sublinear loss under a parametric class of utilities. Empirically, we demonstrate the effectiveness of this mechanism through numerical simulations.

1 Introduction

An exchange economy (EE) is a classical microeconomic construct used to model situations where multiple rational agents share a finite set of scarce resources. Such scenarios arise frequently for applications in operations management, urban planning, crowd sourcing, wireless networks, and sharing resources in data centers (Cohen and Cyert, 1965; Harris et al., 1982; Simonsen, 2018; Dissanayake et al., 2015; Georgiadis et al., 2006; Hussain et al., 2013). In an EE, agents share a set of resources consisting of multiple resource types. They begin with an initial endowment and then exchange these resources among themselves based on a price system. This exchange process allows two agents to trade different resource types if they find it mutually beneficial to do so. Under certain conditions, continually trading in this manner results in a *competitive equilibrium* (CE), where the allocations have desirable Pareto-efficiency and fairness properties. EEs have attracted much research attention, historically since they are tractable models to study human behavior and price determination in real-world markets, and more recently for designing multi-resource fair division mechanisms (Debreu, 1982; Crockett et al., 2008; Tiwari et al., 2009; Budish et al., 2017; Babaioff et al., 2019, 2021).

One of the most common use cases for fair division, which will be especially pertinent in this work, occurs in the context of shared computational resources. For instance, in a data center shared by an organization, we wish to allocate resources such as CPUs, memory, and GPUs to different users who wish to share this cluster in a way that is Pareto-efficient (so that the resources are put into good use) and fair (for long-term user satisfaction). Here, unlike in real world economies where agents might trade with each other until they reach an equilibrium, the equilibrium is computed using a central mechanism (e.g. a cluster manager) based on the preferences submitted by the agents to obtain an allocation with the above properties. Indeed, fair division mechanisms are a staple in many popular multi-tenant cluster management frameworks used in practice, such as Mesos (Hindman et al., 2011),

Quincy (Isard et al., 2009), Kubernetes (Burns et al., 2016), and Yarn (Vavilapalli et al., 2013). Due to this strong practical motivation, a recent line of work has studied such fair division mechanisms for resource sharing in a compute cluster (Chen et al., 2018; Ghodsi et al., 2013; Parkes et al., 2015; Ghodsi et al., 2011), with some of them based on exchange economies and their variants (Zahedi et al., 2018; Gutman and Nisan, 2012; Lai et al., 2005; Varian, 1973).

However, prior work on EEs and fair division typically assumes knowledge of the agent preferences, in the form of a utility function which maps an allocation of the m resource types to the value the agent derives from the allocation. For instance, in the above example, an application developer needs to quantify how well her application performs for each allocation of CPU/memory/GPU she receives. At best, doing so requires the laborious and often erroneous task of profiling their application (Delimitrou and Kozyrakis, 2013; Misra et al., 2021), and at worst, it can be infeasible due to practical constraints (Venkataraman et al., 2016; Rzadca et al., 2020). However, having received an allocation, application developers find it easier to report feedback about the utilities based on the performance they achieved. Moreover, in many real-world systems, this feedback scheme can often be automated (Hindman et al., 2011).

1.1 Contributions & summary of results

We study a multi-round mechanism for computing CE in an exchange economy so as to generate fair and efficient allocations when the exact utilities are *unknown* a priori. A central mechanism is used to learn the user utilities over time via feedback from the agents. At the beginning of each round, the mechanism generates allocations; at the end of the round, agents report feedback on the allocation they received. The mechanism then uses this information to better learn the preferences. In particular, we focus on applications for fair division where a centralized mechanism can compute an allocation of these resources on each round, say, by estimating the utilities and finding their equilibria.

In this pursuit, we first formalize this online learning task and construct two loss functions: the first L^{CE} directly builds on the definition of a CE, while the latter L^{PE} is motivated by the fairness and Pareto-efficiency considerations that arise in fair division. To make the learning problem tractable, we focus on a parametric class of utilities which include the constant elasticity of substitution (CES) utilities which feature prominently in the econometric literature and other application-specific utilities used in the systems literature.

We develop a randomized online mechanism which ef-

ficiently learns utilities over rounds of allocations while simultaneously striving to achieve Pareto-efficient and fair allocations. We show that this mechanism achieves $\widetilde{O}(\sqrt{T})$ loss for the two loss functions with both in-expectation and high-probability upper bounds (Theorems 4.1 and 4.2), under a general family of utility functions. To the best of our knowledge, this is the first work that studies CE without knowledge of user utilities; as such different analysis techniques are necessary. For instance, finding a CE is distinctly different from a vanilla optimization task, and common strategies in bandit optimization such as upper-confidence-bound (UCB) based algorithms do not apply (details in 4). Instead, our algorithm uses a sampling procedure to balance the exploration-exploitation trade-off. We develop new techniques both to bound the losses and to analyse the algorithm. Finally, we corroborate these theoretical insights with empirical simulations.

1.2 Related work

Our work builds on a rich line of literature at the intersection of microeconomics and machine learning. This richness is not surprising: many real world systems are economic and multi-agent in nature, where decisions taken by or for one agent are weighed against the considerations of others, especially when these agents have competing goals such as in resource allocation, matching markets, and in auction-like settings.

As in this work, several works have studied online learning formulations to handle situations where the agents' preferences are not known a priori, but can be learned from repeated interactions (Dudik et al., 2017; Kakade et al., 2010; Balcan et al., 2016; Babaioff et al., 2013; Athey and Segal, 2013; Kandasamy et al., 2020a). Our setting departs from these as we wish to learn agent preferences in an exchange economy, with a focus on designing fair division mechanisms.

Since the seminal work of Varian (1973), fair division of multiple resource types has received significant attention in the game theory, economics, and computer systems literature. One of the most common perspectives on this problem is as an exchange economy (or as a Fisher market, which is a special case of an EE). Moreover, fair allocation mechanisms have been deployed in many practical resource allocation tasks when compute resources are shared by multiple users. Due to space constraints, we defer a more detailed overview on this line of works in Appendix E.1.

Notably, in all of the above cases, an important requirement for the mechanism is that agent utilities be known ahead of time. Some work has attempted to lift this limitation by making explicit assumptions on the utility, but it is not clear that if these assumptions hold in practice (Le et al., 2020; Zahedi et al., 2018).

Recently, Kandasamy et al. (2020b) provides a general method for learning agent utilities for fair division using feedback. However, they only study a *single-resource* setting and do not do not explore multiple resource types. Crucially, in the multi-resource setting, one agent can exchange a resource of one type for a different type of resource from another user, so that both are better off after the exchange. Thus, learning in a *multi-resource* setting is significantly more challenging than the single-resource case since there is no notion of exchange, and requires new analysis techniques.

2 Background

We first present some necessary background material on exchange economies, their competitive equilibria, and fair division mechanisms.

2.1 Exchange economies

In an exchange economy, we have n agents and m divisible resource types. Each agent $i \in [n]$ has an endowment, $e_i = (e_{i1}, \ldots, e_{im})$, where e_{ij} can be viewed the amount of resource j agent i brings to the economy for trade. In the shared compute cluster example, e_i may represent agent i's contribution to this cluster. Without loss of generality we assume $\sum_{i \in [n]} e_{i1} = 1$ so that the space of resources is denoted by $[0,1]^m$.

We denote an allocation of these resources to the n agents by $x = (x_1, x_2, \ldots, x_n)$, where $x_i \in [0,1]^m$ and x_{ij} denote the amount of resource j that is allocated to agent i. The set of all feasible allocations is therefore $\mathcal{X} = \{x : \sum_{i=1}^m x_{ij} \leqslant \mathbf{1}, x_{ij} \geqslant 0, \forall i \in [n], j \in [m]\}$.

An agent's utility function is simply $u_i : [0,1]^m \to [0,1]$, where $u_i(x_i)$ represents her valuation for an allocation x_i she receives. Here u_i is non-decreasing, i.e., $u_i(x_i) \leqslant u_i(x_i')$ for all $x_i \leqslant x_i'$ element-wise (more allocations will not hurt).

In an exchange economy, agents exchange resources based on a price system. We denote a price vector by p, where $p \in \mathbb{R}_+^m$ and $\mathbf{1}^\top p = 1$ (the normalization accounts for the fact that only relative prices matter). Here p_j denotes the price for resource j. Given a price vector p, an agent i has a *budget* $p^\top e_i$, which is the *monetary* value of her endowment according to the prices in p. As this is an economy, a rational agent will then seek to maximize her utility under her budget:

$$d_i(p) = \underset{x_i \in [0,1]^m}{\arg\max} \, u_i(x_i) \quad \text{subject to } p^\top x_i \leqslant p^\top e_i. \quad (1)$$

While generally, the preferred allocations $d_i(p)$ form a set, for simplicity we will assume it is a singleton and treat d_i as a function which outputs an allocation for agent i. This is justified under very general

conditions (Mas-Colell et al., 1995; Varian and Varian, 1992). We refer to $d_i(p)$ chosen in the above manner as the agent i's demand for prices p.

Competitive equilibria – definition, existence and uniqueness: A natural way to allocate resources to agents is to set prices p for the resources, and have the agents maximize their utility under this price system. That is, we allocate $x(p) = (x_1, \ldots, x_n)$. Unfortunately, such an allocation may be infeasible, and even if it were, it may not result in an efficient allocation. However, under certain conditions, we can compute a *competitive equilibrium* (CE), where the prices have both of the desired properties:

Definition 2.1 (Competitive (Walrasian) Equilibrium). A CE is a pair of allocations and prices $(x^\star, p^\star)$ such that *(i)* the allocations are feasible and *(ii)* all agents maximize their utilities under the budget induced by prices $p^\star$. Precisely,

$$\sum_{i \in [n]} x_{i,j}^\star \leqslant \sum_{i \in [n]} e_{ij} = 1, \quad \forall j \in [m],$$
$$x_i^\star = d_i(p^\star), \quad \forall i \in [n].$$

Some definitions of a CE require that the first condition above being an exact equality (e.g., (Mas-Colell et al., 1995)). However, when the utilities are strictly increasing (which will be the case in the sequel), both definitions coincide (Varian and Varian, 1992).

Utilities. In general, CEs do always exist but may not be unique. However, one important class of utilities that guarantee this condition with much attention in the fair division literature is the constant elasticity of substitution (CES) utility. Due to its favorable properties, CES utilities are widely-studied in many fair division works, and most of the existing algorithms that generate fair and efficient allocations assume CES utilities or its sub-classes (Varian and Varian, 1992; Mas-Colell et al., 1995). CES utilities are also ubiquitous in the microeconomics literature; due to this flexibility in interpolating between perfect substitutability and complementary, they are also able to approximate several real-world utility functions. Moreover, computationally, there are efficient methods for computing a CE in the CES and related classes (Zhang, 2011; Zahedi et al., 2018). In contrast, even when CE exist, they may be hard to find under more general classes of utilities (Varian and Varian, 1992).

Example 2.2 (CES utilities). A CES utility takes the form $u_i(x) = \left(\sum_{j=1}^m \theta_{ij} x_i^\rho \right)^{1/\rho}$ where ρ is the elasticity of substitution, and $\theta_i = (\theta_{i1}, \ldots, \theta_{im})$ is an agent-specific parameter. When $\rho = 1$, this corresponds to linear utilities where goods are perfect substitutes. As $\rho \to \infty$, the utilities approach perfect complements.

2.2 Fair division

We describe exchange economies which are used in fair-division mechanisms. We first formally define the fair division problem.

In a standard mechanism for fair division when the utilities are inputs, each agent truthfully[1] submits her utility u_i to the mechanism. The mechanism then returns an allocation $x \in \mathcal{X}$ that are not only efficient but also fair, which satisfies the following two requirements: *sharing incentive* (SI) and *Pareto-efficiency* (PE). An allocation $x = (x_1, \ldots, x_m)$ satisfies SI if the utility an agent receives is at least as much as her utility when using her endowment, i.e. $u_i(x_i) \geqslant u_i(e_i)$. This simply states that she is never worse off than if she had kept her endowment to herself, so she has the incentive to participate in the fair division mechanism.

A feasible allocation x is said to be PE if the utility of one agent can be increased only by decreasing the utility of another. Rigorously, an allocation x dominates another x', if $u_i(x_i) \geqslant u_j(x'_i)$ for all $i \in [n]$ and there exists some $i \in [n]$ such that $u_i(x_i) > u_i(x'_i)$. An allocation is Pareto-efficient if it is not dominated by any other point. We denote the set of Pareto-efficient allocations by $\mathcal{PE}$. One advantage of the PE requirement, when compared to other formalisms which maximize social or egalitarian welfare, is that it does not compare the utility of one agent against that of another. The utilities are useful solely to specify an agent's preferences over different allocations.

EEs in fair division: The above problem description for fair division naturally renders itself to a solution based on EEs. By treating the resource allocation environment as an exchange economy, we may compute its equilibrium to determine the allocations for each agent. Then, the SI property follows from the fact that each agent is maximizing her utility under her budget, and an agent's endowment (trivially) falls under her budget. The PE property follows from the first theorem of welfare economics (Varian and Varian, 1992; Mas-Colell et al., 1995). Several prior works have used this connection to design fair-division mechanisms for many practical applications (Varian, 1973; Crockett et al., 2008; Zahedi et al., 2018).

Computing a CE: In order to realize a CE allocation in a fair division mechanism, the mechanism needs to compute a CE given a set of utilities. One way to do this is via tatonnement (Varian and Varian, 1992). While there are general procedures, such as

tatonnement (Varian and Varian, 1992), they are not guaranteed to converge to an equilibrium even when it exists; moreover, even when they do, the rate of convergence can be slow. This has led to the development of efficient procedures for special classes of functions. One such method is proportional response dynamics (PRD) (Zhang, 2011; Zahedi et al., 2018) which converges faster under CES utilities (Zhang, 2011) and other classes of utilities (Zahedi et al., 2018) when $e_i = \alpha_i \mathbf{1}_m$ for all $i \in [n]$ (with $\sum_i \alpha_i = 1$). In fact, in our evaluations, we adopt PRD for computing a CE, which is a subroutine of the learning algorithm.

We note that in the context of fair division, the CE allocations are more pertinent than the CE prices. While the prices are used to compute fair allocations, they are not used directly in their own right.

3 Online Learning Formulation

We formalize online learning an equilibrium in an exchange economy under bandit feedback, when the exact agent utilities are unknown a priori. We consider a multi-round setting, where in each round t, the mechanism selects (x_t, p_t), where $x_t = (x_{t,1}, \ldots, x_{t,n}) \in \mathcal{X}$ are the allocations for each agent for the current round, and p_t are the prices for units of each resource.

The agents, having experienced their allocation, report stochastic feedback $\{y_{t,i}\}_{i \in [n]}$, where $y_{t,i}$ is σ sub-Gaussian and $\mathbb{E}[y_{t,i}|x_{t,i}] = u_i(x_{t,i})$. The mechanism then uses this information to compute allocations for the next round. As described in Section 1, this set up is motivated by use cases in data center resource allocations, where jobs (agents) cannot state their utility upfront, but can report feedback on their performance in an automated way.

Going forward, we slightly abuse notation when referring to the allocations. When $i \in [n]$ indexes an agent, $x_i = (x_{i1}, \ldots, x_{im}) \in [0,1]^m$ denotes the allocation to agent i. When t indexes a round, $x_t = (x_{t,1}, \ldots, x_{t,n}) \in \mathcal{X}$ will refer to an allocation to all agents, where $x_{t,i} = (x_{t,i,1}, \ldots, x_{t,i,m}) \in [0,1]^m$ denotes i's allocation in that round. The intended meaning should be clear from context.

3.1 Losses

We study two losses for this setting. The first loss is based directly on the definition of an equilibrium (Def. 2.1). For $a \in \mathbb{R}$, denote $a^+ = \max(0, a)$. We define the CE loss ℓ^{CE} of an allocation–price pair (x, p) as the sum, over all agents, of the difference between the maximum attainable utility under price p and the utility achieved by allocation x. The T-round loss L_T^{CE}

---

[1]Unlike some previous works on fair division (Parkes et al., 2015; Kandasamy et al., 2020b; Ghodsi et al., 2011), we do not study strategic considerations, where agents may attempt to manipulate outcomes in their favor by falsely submitting their utilities.

is the sum of $\ell^{\mathrm{CE}}(x_t, p_t)$ losses over T rounds. We have:

$$
\ell^{\mathrm{CE}}(x, p) \stackrel{\text{def}}{=} \sum_{i=1}^{n} \left(\max_{x_i' : p^\top x_i' \leqslant p^\top e_i} u_i(x_i') - u_i(x) \right)^+,
$$
$$
L_T^{\mathrm{CE}} \stackrel{\text{def}}{=} \sum_{t=1}^{T} \ell^{\mathrm{CE}}(x_t, p_t). \tag{2}
$$

It is straightforward to see that for a CE pair $(x^\star, p^\star)$, we have $\ell^{\mathrm{CE}}(x^\star, p^\star) = 0$. As this loss is based directly on the definition of a CE, it captures many of the properties of a CE.

Our second loss is motivated by the fair division use case. Recall from Sec. 2.2 that in fair division, while prices are useful in computing CE allocations, they have no value in their own right. Therefore, we will motivate our loss function based on the sharing incentive (SI) and Pareto-efficiency (PE) desiderata for fair division. It is composed of two parts. We define the SI loss ℓ^{SI} for an allocation x as the sum, over all agents, of how much they are worse off than their endowment utilities. We define the PE loss ℓ^{PE} for an allocation x as the minimum sum, over all agents, of how much they are worse off than some Pareto-efficient utilities. Next, we define the fair division loss ℓ^{FD} as the maximum of ℓ^{SI} and ℓ^{PE}. Finally, we define the T-round loss L_T^{FD} for the online mechanism as the sum of $\ell^{\mathrm{FD}}(x_t)$ losses over T rounds. We have:

$$
\ell^{\mathrm{SI}}(x) \stackrel{\text{def}}{=} \sum_{i=1}^{n} \left(u_i(e_i) - u_i(x_i) \right)^+,
$$
$$
\ell^{\mathrm{PE}} \stackrel{\text{def}}{=} \inf_{x' \in \mathcal{PE}} \sum_{i=1}^{n} \left(u_i(x_i') - u_i(x_i) \right)^+,
$$
$$
\ell^{\mathrm{FD}}(x) \stackrel{\text{def}}{=} \max \left(\ell^{\mathrm{PE}}(x), \ell^{\mathrm{SI}}(x) \right),
$$
$$
L_T^{\mathrm{FD}} \stackrel{\text{def}}{=} \sum_{t=1}^{T} \ell^{\mathrm{FD}}(x_t). \tag{3}
$$

Note that individually achieving either small ℓ^{SI} or ℓ^{PE} is trivial: if an agent's utility is strictly increasing, then by allocating all the resources to this agent we have zero ℓ^{PE} as such an allocation is Pareto-efficient; moreover, by simply allocating each agent their endowment we have zero ℓ^{SI}. In ℓ^{FD}, we require both to be simultaneously small which necessitates a clever allocation that accounts for agents' endowments and utilities. One intuitive interpretation of the PE loss is that it can be bounded above by the L_1 distance to the Pareto-front in utility space; i.e. denoting the set of Pareto-efficient utilities by $U_{\mathrm{PE}} = \{\{u_i(x_i)\}_{i \in [n]}; x \in \mathcal{PE}\} \subset \mathbb{R}^n$, and letting $u(x) = (u_1(x_1), \ldots, u_n(x_n)) \in \mathbb{R}^n$, we can write, $\ell^{\mathrm{PE}}(x) \leqslant \min_{u \in U_{\mathrm{PE}}} \|u - u(x)\|_1$.

The FD loss is more interpretable as it is stated in terms of the SI and PE requirements for fair division.

On the other hand, the CE loss is less intuitive. Moreover, in EEs, while prices help us determine the allocations, they do not have value on their own. Given this, the CE loss has the somewhat undesirable property that it depends on the prices p_t. That said, since the CE loss is based directly on the definition of a CE, it captures other properties of a CE that are not considered in ℓ^{FD} (see an example in Appendix E.3). It is also worth mentioning that either loss cannot be straightforwardly bounded in terms of the other.

Note that we have presented a basic version of the online learning framework as it provides a simplest platform to study the learning problem of efficient and fair allocations. For instance, one could consider richer settings where the utilities might change over time with certain contextual information. While these settings are beyond the scope of this work, we believe the analysis techniques and intuitions developed here are also insightful in analysing other variant settings.

3.2 Model and assumptions

To make the learning problem tractable, we make some additional assumptions on the problem. We consider the following parametric class of utility functions $\mathcal{P}$.

Let $\phi_j : [0, 1] \to [0, 1]$ be an increasing function which maps the allocation x_{ij} of resource j to agent i to some feature value. For brevity, we will write $\phi : [0, 1]^m \to [0, 1]^m$, such that $\phi(x_i) = (\phi_1(x_{i1}), \phi_2(x_{i2}), \ldots, \phi_m(x_{im}))$; Next, let $\mu : \mathbb{R}_+ \to [0, 1]$ be an increasing function. Finally, let $\Theta \subset \mathbb{R}_+^m$ be a set of positive parameters. Then, we consider the following class of utilities $\mathcal{P}$:

$$
\mathcal{P} = \Big\{ \{u_i\}_{i=1}^n; \ u_i(x_i) = \mu \left(\theta_i^\top \phi(x_i) \right)
$$
$$
\text{for some } \theta_i \in \Theta, \ \forall i \in [n] \Big\} \tag{4}
$$

An agent's utility then takes the form $u_i(x_i) = \mu(\theta_i^{*\top} \phi(x_i))$ where the featurization ϕ and the function μ are known, but the true parameters $\theta_i^* \in \Theta$ are unknown and need to be learned by the mechanism.

We consider the above class of functions for the following reasons. First, observe that it represents a valid class of utilities in that for all positive θ, the utilities are increasing in the allocations. Second, a CE is guaranteed to exist uniquely in this class. Third, from a practical point of view, it subsumes a majority of utilities studied in the fair division literature, such as linear utilities, the CES utilities from Example 2.2 (Crockett et al., 2008; Tiwari et al., 2009; Budish et al., 2017; Babaioff et al., 2019, 2021), and other application-specific utilities (Zahedi et al., 2018;

Venkataraman et al., 2016), Fourth, also from a practical point of view, the CE can be efficiently computed on this class (Zhang, 2011). Finally, it also allows us to leverage techniques for estimating generalized linear models in our online learning mechanism (Chen et al., 1999; Filippi et al., 2010).

We will also assume the following regularity conditions on $\mathcal{P}$ to avoid some degenerate cases in our analysis. First, μ is continuously differentiable, it is Lipschitz-continuous with constant L_μ, and $C_\mu = \inf_{\theta \in \Theta, x \in \mathcal{X}} \dot{\mu}\left(\theta^\top \phi(x)\right) > 0$. Second, $\Theta \subset [\theta_{\min}, \infty)^m$, where $\theta_{\min} > 0$. These assumptions can be relaxed (albeit with a more involved analysis), or replaced by other equivalent regularity conditions (Chen et al., 1999; Filippi et al., 2010), without affecting the main analysis ideas or take-aways in this paper. Our results also apply when μ, ϕ, and Θ can be defined separately for each agent, but we assume they are the same to simplify the exposition.

4 Algorithm and Theoretical Results

Algorithm 1 A Randomized Alg. for Learning in EEs

1: **Input:** number of initialization sub-phases $M \geqslant 1$, confidence parameters $\{\delta_t\}_{t \geqslant 1}$.
2: $t \leftarrow 0$
3: **for** $\ell = 1, \ldots, M$ **do**   // Initialization phase
4:    **for** $k = 1, \ldots, \max(m, n)$ **do**
5:      $t \leftarrow t + 1$, $x_t \leftarrow (\mathbf{0}_m, \ldots, \mathbf{0}_m)$
6:      **for** $h = 1, \ldots, \min(m, n)$ **do**
7:        **if** $m < n$ **then**
8:          $x_{t, h+k-1, j} \leftarrow 1$ for all $j \in [m]$.
9:        **else**
10:          $x_{t, i, h+k-1} \leftarrow 1$ for all $i \in [n]$.
11:      Allocate x_t and observe rewards $\{y_{t,i}\}_{i \in [n]}$.
12: **while True do**   // Round for learning phase
13:    $t \leftarrow t + 1$
14:    **for** $i = 1, 2, \ldots, n$ **do**
15:      Compute $Q_{t,i} \overset{\text{def}}{=} \sum_{s=1}^{t-1} \phi(x_{s,i}) \phi(x_{s,i})^\top$
16:      Compute
$$\bar{\theta}_{t,i} = \arg\min_{\theta \in \Theta} \left\| \sum_{s=1}^{t-1} \phi(x_{s,i})\left(\mu(\theta^\top \phi(x_{s,i})) - y_{s,i}\right) \right\|_{Q_{t,i}^{-1}}$$
17:      Sample $\theta'_{t,i} \sim \mathcal{N}(\bar{\theta}_{t,i}, \alpha_t^2 Q_{t,i}^{-1})$.   // See (5) for α_t.
18:      $\theta_{t,i} \leftarrow \arg\min_{\theta' \in \Theta} \|\theta'_{t,i} - \theta'\|$.   // Projection
19:      Choose allocations and prices $x_t, p_t = \text{CE}(\{u_{t,i}\}_{i=1}^n)$, where $u_{t,i}(\cdot) = \mu(\theta_{t,i}^\top \phi(\cdot))$
20:      Observe rewards $\{y_{t,i}\}_{i \in [n]}$.

We present a randomized online learning algorithm for learning the agents' utilities and generating fair and efficient allocations. Note that this algorithm

not only needs to learn the unknown utilities quickly, but should also simultaneoulsy find the CE allocation. This latter aspect introduces new challenges in our setting. For instance, the most popular approach for stochastic optimization under bandit feedback are based on upper-confidence-bounds (UCB). However, finding a CE cannot be straightforwardly framed as a vanilla optimization procedure and hence UCB procedures do not apply. Instead, our proposed algorithm uses a key randomized sampling step, which tradeoffs between exploration and exploration while maintaining the utilities' shape constraints in every round for computing the CE (details in proof sketch).

The algorithm, outlined in Algorithm 1, takes input parameters M and $\{\delta_t\}_{t \geqslant 1}$ whose values we will specify shortly. It begins with an initialization phase for M sub-phases (line 3), each of length $\min(n, m)$. During each sub-phase, we allocate each resource entirely to each user for at least one round. This initialization phase ensures that some matrices we define subsequently are well conditioned.

After the initialization phase, the algorithm operates on each of the remaining rounds as follows. For each user, it first computes quantities $Q_{t,i} \in \mathbb{R}^{m \times m}$ and $\bar{\theta}_{t,i} \in \mathbb{R}^m$ as defined in lines 15, and 16. As we explain shortly, $\bar{\theta}_{t,i}$ can be viewed as an estimate of θ_i^* based on the data from the first $t-1$ rounds. The algorithm then samples $\theta'_{t,i} \in \mathbb{R}^m$ from a normal distribution with mean $\bar{\theta}_{t,i}$ and co-variance $\alpha_t^2 Q_{t,i}$, where, α_t is defined as:

$$\alpha_t^2 = 4 \frac{\kappa^2 \sigma^2}{C_\mu^2} m \log(t) \log\left(\frac{m}{\delta_t}\right),$$
$$\kappa = 3 + 2\log\left(1 + 2\|\phi(\mathbf{1})\|_2^2\right). \tag{5}$$

The sampling distribution, which is centered at our estimate $\bar{\theta}_{t,i}$, is designed to balance the exploration-exploitation trade-off on this problem. Next, it projects the sampled $\theta'_{i,t}$ onto Θ to obtain $\theta_{t,i}$.

In line (19), the algorithm obtains an allocation and price pair x_t, p_t by computing the CE on the $\theta_{t,i}$ values obtained above, i.e. by pretending that $u_{t,i}(\cdot) = \mu(\theta_{t,i}^\top \phi(\cdot))$ is the utility for user i.

It is important to note that the computation of the CE happens as a *subroutine* of the mechanism, and users will simply receive the allocations x_t. The mechanism collects the rewards $\{y_{t,i}\}_{i \in [n]}$ from each user and then repeats the same for the remaining rounds. As we discussed in Sec. 2.2, there are different ways to compute a CE efficiently in our setting, including tatonnement or the proportional response dynamics (PRD) algorithm (Zhang, 2011) which we implemented. Given that our algorithm focus on learning the efficient and fair allocations, we do not focus on the computation

complexity of CE in this work. Empirically, we find PRD converges quickly in the simulations.

Computation of $\bar{\theta}_{it}$: It is worth explaining steps 15–16 used to obtain the estimate $\bar{\theta}_{it}$ for user i's parameter θ_i^*. Recall that for each agent i, the mechanism receives stochastic rewards $y_{t,i}$ where $y_{t,i}$ is a σ sub-Gaussian random variable with $\mathbb{E}[y_{t,i}] = u_i(x_{t,i})$ in round t. Therefore, given the allocation-reward pairs $\{(x_{s,i}, y_{s,i})\}_{s=1}^{t-1}$, the maximum quasi-likelihood estimator $\widehat{\theta}_{t,i}^{\mathrm{MLE}}$ for θ_i is defined as the maximizer of the quasi-likelihood $\mathcal{L}(\theta) = \sum_{s=1}^{t-1} \log p_\theta(y_{s,i}|x_{s,i})$, where $p_\theta(y_i|x_i)$ is as defined below. Here, $\mu(\nu) = \frac{\partial b(\nu)}{\partial \nu}$ and $c(\cdot)$ is a normalising term. We have:

$$p_\theta(y_i|x_i) = \exp\left(y_i \theta^\top \phi(x_{si}) - b(\theta^\top \phi(x_{si})) + c(y_i)\right). \quad (6)$$

Upon differentiating, we have that $\widehat{\theta}_{t,i}^{\mathrm{MLE}}$ is the unique solution of the estimating equation:

$$\sum_{s=1}^{t-1} \phi(x_{si})\left(\mu\left(\left(\widehat{\theta}_{t,i}^{\mathrm{MLE}}\right)^\top \phi(x_{s,i})\right) - y_{si}\right) = 0.$$

In other words, $\theta_{t,i}^{\mathrm{MLE}}$ would be the maximum likelihood estimate for θ_i^* if the rewards $y_{t,i}$ followed an exponential family likelihood as shown in (6). Our assumptions are more general; we only assume the rewards are sub-Gaussian centred at $\mu(\theta_i^{*\top}\phi(x_{t,i}))$. However, this estimate is known to be consistent under very general conditions, including when the rewards are sub-Gaussian (Chen et al., 1999; Filippi et al., 2010). Since $\widehat{\theta}_{t,i}^{\mathrm{MLE}}$ might be outside of the set of feasible parameters Θ, this motivates us to perform the projection in the $Q_{t,i}^{-1}$ norm to obtain $\bar{\theta}_{t,i}$ as defined in line 16. Here, $Q_{t,i}$, defined in line (15), is the design matrix obtained from the data in the first $t-1$ steps.

On the algorithm design: It is worth comparing the design of our algorithm against prior work in the bandit literature under similar parametric assumptions (Dani et al., 2008; Filippi et al., 2010; Li et al., 2017; Rusmevichientong and Tsitsiklis, 2010). For instance, in a CE, each agent is maximizing their utility under a budget constraint. Therefore, a seemingly natural idea is to adopt a UCB based procedure, which is the most common approach for stochastic optimization under bandit feedback (Auer, 2002). However, adopting a UCB-style method for our problem proved to be unfruitful. Consider using a UCB of the form $\mu(\bar{\theta}_{it}^\top \phi(\cdot)) + U_{it}(\cdot)$, where U_{it} quantifies the uncertainty in the current estimate. Unfortunately, a CE is not guaranteed to exist for utilities of the above form, which means that finding a suitable allocation can be difficult. An alternative idea is to consider UCBs of the form $\mu(\widehat{\theta}_{t,i}^\top \phi(\cdot))$ where $\widehat{\theta}_{t,i}$ is an upper

confidence bound on θ_i^* (recall that both θ_i^* and ϕ are non-negative). While CEs are guaranteed to exist for such UCBs, $\widehat{\theta}_{t,i}$ is not guaranteed to uniformly converge to θ_i^*, resulting in linear loss.

Instead, our algorithm takes inspiration from classical Thompson sampling (TS) procedure for multi-armed bandits in the Bayesian paradigm (Thompson, 1933). The sampling step in line 17 is akin to sampling from the posterior beliefs in TS. It should be emphasized that the sampling distributions on each round cannot be interpreted as the posterior of some prior belief on θ_i^*. In fact, they were designed so as to put most of their mass inside a frequentist confidence set for θ_i^*.

4.1 Upper bounds on the loss

The following two theorems are the main results bounding the loss terms $L^{\mathrm{FD}}, L^{\mathrm{CE}}$ for Algorithm 1. In the first theorem, we are given a target failure probability of at most δ. By choosing δ_t appropriately, we obtain an infinite horizon algorithm for which both loss terms are $\widetilde{O}(\sqrt{T})$ with probability at least $1 - \delta$. In the second theorem, with a given time horizon T, we obtain an algorithm whose expected losses are $\widetilde{O}(\sqrt{T})$.

Theorem 4.1. *Assume the conditions in Section 3.2. Let $\delta > 0$ be given. Choose $\delta_t = \frac{2\delta}{n\pi^2 t^2}$. Then, the following upper bounds on $L^{\mathrm{FD}}, L^{\mathrm{CE}}$ hold for Algorithm 1 with probability at least $1 - \delta$.*

$$L^{\mathrm{FD}}(T), L^{\mathrm{CE}}(T) \in O\left(n\left(m + \frac{m^2}{\sqrt{M}}\right)\sqrt{T}\left(\log(nT/\delta) + \log(T)\right)\right).$$

Theorem 4.2. *Assume the conditions in Section 3.2. Let $T > M\max(m^2, n)$ be given. Choose $\delta_t = \frac{1}{T}$. Then, the follow upper bounds on $L^{\mathrm{FD}}, L^{\mathrm{CE}}$ hold for Algorithm 1.*

$$\mathbb{E}[L^{\mathrm{FD}}(T)], \mathbb{E}[L^{\mathrm{CE}}(T)] \in O\left(n\left(m + \frac{m^2}{\sqrt{M}}\right)\sqrt{T}\left(\log(T)\right)\right),$$

Above, probabilities and expectations are with respect to both the randomness in the observations and the sampling procedure. Both theorems show that we can learn with respect to both losses at $\sqrt{T}$ rate. Note that the rates depend on the number of initialization subphases M. By choosing $M = m^2$, we get a $\widetilde{O}(nm\sqrt{T})$ bound. However, this also requires a large initialization phase, which may not be feasible in practice. We can instead choose M to be small, but this leads to correspondingly worse asymptotic bounds.

Proof sketch. Our proof uses some prior martingale concentration results from the bandit literature (Rusmevichientong and Tsitsiklis, 2010; Filippi et al., 2010), and additionally, we use some high level
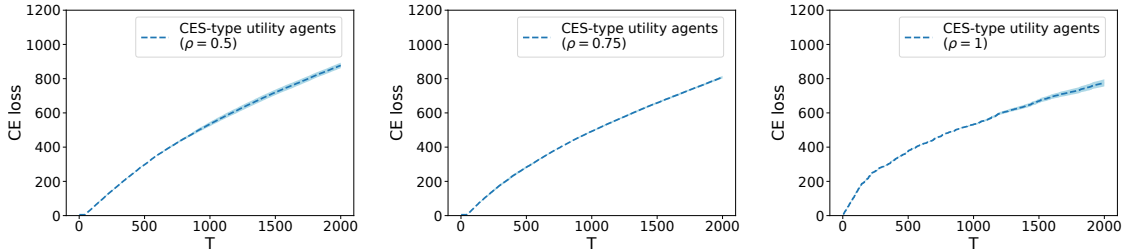
Figure 1: The CE loss L_T^{CE} vs the number of rounds T, evaluated with $m = 3$ resource types and $n = 5$ agents with CES utilities. We present results for $\rho = 0.5$, $\rho = 0.75$, and $\rho = 1$ respectively (see Example 2.2). All figures show results which are averaged over 10 runs, and the shaded region shows the standard error at each time T.

intuitions from prior frequentist analyses of Thompson sampling (Kaufmann et al., 2012; Agrawal and Goyal, 2012; Mutnỳ and Krause, 2019). At the same time, we also require novel techniques, both to bound the loss terms, and analyse the algorithm. Our proof for bounding L_T^{CE} first defines high probability events $A_{t,i}, B_{t,i}$ for each agent i and round t. $A_{t,i}$ captures the event that the estimated $\bar{\theta}_{t,i}$ is close to θ_i^* in $Q_{t,i}$ norm. We upper bound $\mathbb{P}(A_{t,i}^c)$ using the properties of the maximum quasi-likelihood estimator on GLMs (Chen et al., 1999; Filippi et al., 2010) and a martingale argument. $B_{t,i}$ captures the event that the sampled $\theta_{t,i}$ is close to $\bar{\theta}_{t,i}$ in $Q_{t,i}$ norm. Given these events, we then bound the instantaneous losses $\ell^{\mathrm{CE}}(x_t, p_t)$ by a super martingale with bounded differences. The final bound is obtained by an application of the Azuma inequality.

Another key ingredient in this proof is to show that the sampling step also explores sufficiently–the B_{it} event only captures exploitation; since the sampling distribution is a multi-variate Gaussian, this can be conveniently argued using an upper bound on the standard normal tail probability. While bounding L^{FD} uses several results and techniques as above, it cannot be directly related to L^{CE}, and requires a separate analysis.

5   Experiments

We evaluated Algorithm 1 with simulations. To the best of our knowledge, this is the first online algorithm studying fair and efficient allocations with unknown utilities with multiple heterogeneous resource types, and there are no existing natural baselines. There is also no straightforward adaptation of the method described in Kandasamy et al. (2020b) for single resource types since they do not consider the exchange of resources. We evaluated based on two types of utilities.

1. CES utilities: Described in Example 2.2.

2. Amdahl's utilities: The Amdahl's utility function, described in Zahedi et al. (2018), is used to model the performance of jobs distributed across heteroge-

neous machines in a data center. This utility is motivated by Amdahl's Law (Amdahl, 2013), which models a job's speed up in terms of the fraction of work that can be parallelized. Let $0 < f_{ij} < 1$ denote the parallel fraction of user i's job on machine type j. Then, an agent's Amdahl utility is: $u_i(x) = \sum_{j=1}^{m} \theta_{ij} \phi_{ij}(x_{ij})$, where $\phi_{ij}(x_{ij}) = x_{ij}/f_i + (1-f_i)x_{ij}$. $\phi_{ij}(x_{ij})$ is the relative speedup produced by allocation x_{ij}. Both CES and Amdahl utilities belong to our class $\mathcal{P}$ given in (4).

We focus our evaluation on the CE loss; computing the FD loss is computationally expensive as it requires taking an infimum over the Pareto-front (more details in Appendix E). Our first set of experiments consider an environment with $m = 3$ resource types and $n = 5$ agents, all of whom have CES utilities. We conduct three experiments with different values for the elasticity of substitution ρ. Our second set of experiments consider an environment with $m = 2$ resource types and $n = 8$ agents, all of whom have Amdahl's utilities, where the results are similar and thus included in Appendix D. We conduct three experiments with different values for the parallel fraction f_{ij}. All experiments are run for $T = 2000$ rounds, where we set $\delta = \frac{1}{T}$. The results are given in Figure 1. They show that the CE loss grows sublinearly with T which indicates that the algorithm is able to learn utilities and compute a CE.

To compute the CE at line 19 of Algorithm 1, we use the proportional response dynamics procedure from Zhang (2011) with 20 iterations. To compute L^{CE}, we need to maximize each agent's utility subject to a budget. Full experimental details and additional results are included in Appendix D.

6   Conclusion

We introduced and studied the problem of online learning a competitive equilibrium in an exchange economy, without a priori knowledge of agents' utilities. We quantify the learning performance via two losses, the first motivated from the definition of an equilibrium, and the second by fairness and Pareto-efficiency considerations in fair division. We develop a random-

ized algorithm which achieves $\tilde{O}(nm\sqrt{T})$ loss after T rounds under both losses, and corroborate these theoretical results with simulations. While our work takes the first step towards sequentially learning a market equilibrium in exchange economies, an interesting avenue for future work would be to study learning approaches in broader classes of agent utilities and market dynamics.

Acknowledgments

References

Milton Abramowitz, Irene A Stegun, and Robert H Romer. Handbook of mathematical functions with formulas, graphs, and mathematical tables, 1988.

Shipra Agrawal and Navin Goyal. Analysis of thompson sampling for the multi-armed bandit problem. In *Conference on Learning Theory*, pages 39–1, 2012.

Gene M Amdahl. Computer architecture and amdahl's law. *Computer*, 46(12):38–46, 2013.

Susan Athey and Ilya Segal. An efficient dynamic mechanism. *Econometrica*, 81(6):2463–2485, 2013.

Peter Auer. Using confidence bounds for exploitation-exploration trade-offs. *Journal of Machine Learning Research*, 3(Nov):397–422, 2002.

Moshe Babaioff, Robert Kleinberg, and Aleksandrs Slivkins. Multi-parameter mechanisms with implicit payment computation. In *Proceedings of the Fourteenth ACM Conference on Electronic Commerce*, pages 35–52, 2013.

Moshe Babaioff, Noam Nisan, and Inbal Talgam-Cohen. Fair allocation through competitive equilibrium from generic incomes. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, pages 180–180, 2019.

Moshe Babaioff, Noam Nisan, and Inbal Talgam-Cohen. Competitive equilibrium with indivisible goods and generic budgets. *Mathematics of Operations Research*, 46(1):382–403, 2021.

Maria-Florina F Balcan, Tuomas Sandholm, and Ellen Vitercik. Sample complexity of automated mecha-

nism design. In *Conference on Neural Information Processing Systems*, pages 2083–2091, 2016.

Eric Budish, Gérard P Cachon, Judd B Kessler, and Abraham Othman. Course match: A large-scale implementation of approximate competitive equilibrium from equal incomes for combinatorial allocation. *Operations Research*, 65(2):314–336, 2017.

Brendan Burns, Brian Grant, David Oppenheimer, Eric Brewer, and John Wilkes. Borg, omega, and kubernetes. *ACM Queue*, 14:70–93, 2016. URL http://queue.acm.org/detail.cfm?id=2898444.

Kani Chen, Inchi Hu, Zhiliang Ying, et al. Strong consistency of maximum quasi-likelihood estimators in generalized linear models with fixed and adaptive designs. *Annals of Statistics*, 27(4):1155–1163, 1999.

Li Chen, Shuhao Liu, Baochun Li, and Bo Li. Scheduling jobs across geo-distributed datacenters with max-min fairness. *IEEE Transactions on Network Science and Engineering*, 6(3):488–500, 2018.

Kalman J Cohen and Richard Michael Cyert. Theory of the firm; resource allocation in a market economy. Technical report, Prentice-Hall, 1965.

Sean Crockett, Stephen Spear, and Shyam Sunder. Learning competitive equilibrium. *Journal of Mathematical Economics*, 44(7-8):651–671, 2008.

Varsha Dani, Thomas P Hayes, and Sham M Kakade. Stochastic linear optimization under bandit feedback. 2008.

Gerard Debreu. Existence of competitive equilibrium. *Handbook of Mathematical Economics*, 2:697–743, 1982.

Christina Delimitrou and Christos Kozyrakis. Paragon: Qos-aware scheduling for heterogeneous datacenters. *ACM SIGPLAN Notices*, 48(4):77–88, 2013.

Indika Dissanayake, Jie Zhang, and Bin Gu. Task division for team success in crowdsourcing contests: Resource allocation and alignment effects. *Journal of Management Information Systems*, 32(2):8–39, 2015.

Danny Dolev, Dror G Feitelson, Joseph Y Halpern, Raz Kupferman, and Nathan Linial. No justified complaints: On fair sharing of multiple resources. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, pages 68–75, 2012.

Miroslav Dudik, Nika Haghtalab, Haipeng Luo, Robert E Schapire, Vasilis Syrgkanis, and Jennifer Wortman Vaughan. Oracle-efficient online learning and auction design. In *IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, pages 528–539. IEEE, 2017.

Sarah Filippi, Olivier Cappe, Aurélien Garivier, and Csaba Szepesvári. Parametric bandits: The generalized linear case. In *Conference on Neural Information Processing Systems*, volume 23, pages 586–594, 2010.

Leonidas Georgiadis, Michael J Neely, and Leandros Tassiulas. *Resource Allocation and Cross-layer Control in Wireless Networks*. Now Publishers Inc, 2006.

Ali Ghodsi, Matei Zaharia, Benjamin Hindman, Andy Konwinski, Scott Shenker, and Ion Stoica. Dominant resource fairness: Fair allocation of multiple resource types. In *NSDI*, volume 11, pages 24–24, 2011.

Ali Ghodsi, Matei Zaharia, Scott Shenker, and Ion Stoica. Choosy: Max-min fair sharing for datacenter jobs with constraints. In *Proceedings of the 8th ACM European Conference on Computer Systems*, pages 365–378, 2013.

Avital Gutman and Noam Nisan. Fair allocation without trade. 2012. URL https://www.cs.huji.ac.il/~noam/notrade.pdf.

Milton Harris, Charles H Kriebel, and Artur Raviv. Asymmetric information, incentives and intrafirm resource allocation. *Management Science*, 28(6): 604–620, 1982.

Benjamin Hindman, Andy Konwinski, Matei Zaharia, Ali Ghodsi, Anthony D Joseph, Randy H Katz, Scott Shenker, and Ion Stoica. Mesos: A platform for fine-grained resource sharing in the data center. In *NSDI*, volume 11, pages 22–22, 2011.

Hameed Hussain, Saif Ur Rehman Malik, Abdul Hameed, Samee Ullah Khan, Gage Bickler, Nasro Min-Allah, Muhammad Bilal Qureshi, Limin Zhang, Wang Yongji, Nasir Ghani, et al. A survey on resource allocation in high performance distributed computing systems. *Parallel Computing*, 39(11): 709–736, 2013.

Michael Isard, Vijayan Prabhakaran, Jon Currey, Udi Wieder, Kunal Talwar, and Andrew Goldberg. Quincy: fair scheduling for distributed computing clusters. In *Proceedings of the ACM SIGOPS 22nd Symposium on Operating Systems Principles*, pages 261–276, 2009.

Sham M Kakade, Ilan Lobel, and Hamid Nazerzadeh. An optimal dynamic mechanism for multi-armed bandit processes. *arXiv preprint arXiv:1001.4598*, 2010.

Kirthevasan Kandasamy, Joseph E Gonzalez, Michael I Jordan, and Ion Stoica. Mechanism design with bandit feedback. *arXiv preprint arXiv:2004.08924*, 2020a.

Kirthevasan Kandasamy, Gur-Eyal Sela, Joseph E Gonzalez, Michael I Jordan, and Ion Stoica. Online learning demands in max-min fairness. *arXiv preprint arXiv:2012.08648*, 2020b.

Emilie Kaufmann, Nathaniel Korda, and Rémi Munos. Thompson sampling: An asymptotically optimal finite-time analysis. In *International Conference on Algorithmic Learning Theory*, pages 199–213. Springer, 2012.

Kevin Lai, Lars Rasmusson, Eytan Adar, Li Zhang, and Bernardo A Huberman. Tycoon: An implementation of a distributed, market-based resource allocation system. *Multiagent and Grid Systems*, 1 (3):169–182, 2005.

Tan N Le, Xiao Sun, Mosharaf Chowdhury, and Zhenhua Liu. Allox: compute allocation in hybrid clusters. In *Proceedings of the Fifteenth European Conference on Computer Systems*, pages 1–16, 2020.

Jin Li and Jingyi Xue. Egalitarian division under leontief preferences. *Economic Theory*, 54(3):597–622, 2013.

Lihong Li, Yu Lu, and Dengyong Zhou. Provably optimal algorithms for generalized linear contextual bandits. In *International Conference on Machine Learning*, pages 2071–2080. PMLR, 2017.

Andreu Mas-Colell, Michael Dennis Whinston, Jerry R Green, et al. *Microeconomic Theory*, volume 1. Oxford university press New York, 1995.

Ujval Misra, Richard Liaw, Lisa Dunlap, Romil Bhardwaj, Kirthevasan Kandasamy, Joseph E Gonzalez, Ion Stoica, and Alexey Tumanov. Rubberband: cloud-based hyperparameter tuning. In *Proceedings of the Sixteenth European Conference on Computer Systems*, pages 327–342, 2021.

Mojmír Mutný and Andreas Krause. Efficient high dimensional bayesian optimization with additivity and quadrature fourier features. *Conference on Neural Information Processing Systems 31*, pages 9005–9016, 2019.

David C Parkes, Ariel D Procaccia, and Nisarg Shah. Beyond dominant resource fairness: Extensions, limitations, and indivisibilities. *ACM Transactions on Economics and Computation*, 3(1):1–22, 2015.

Paat Rusmevichientong and John N Tsitsiklis. Linearly parameterized bandits. *Mathematics of Operations Research*, 35(2):395–411, 2010.

Krzysztof Rzadca, Pawel Findeisen, Jacek Swiderski, Przemyslaw Zych, Przemyslaw Broniek, Jarek Kusmierek, Pawel Nowak, Beata Strack, Piotr Witusowski, Steven Hand, et al. Autopilot: workload autoscaling at google. In *Proceedings of the Fifteenth European Conference on Computer Systems*, pages 1–16, 2020.

William Simonsen. *Citizen participation in resource allocation*. Routledge, 2018.

William R Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3/4):285–294, 1933.

Mayank Tiwari, Theodore Groves, and Pamela C Cosman. Competitive equilibrium bitrate allocation for multiple video streams. *IEEE Transactions on Image Processing*, 19(4):1009–1021, 2009.

Hal R Varian. Equity, envy, and efficiency. 1973.

Hal R Varian and Hal R Varian. *Microeconomic Analysis*, volume 3. Norton New York, 1992.

Vinod Kumar Vavilapalli, Arun C Murthy, Chris Douglas, Sharad Agarwal, Mahadev Konar, Robert Evans, Thomas Graves, Jason Lowe, Hitesh Shah, Siddharth Seth, et al. Apache hadoop yarn: Yet another resource negotiator. In *Proceedings of the 4th Annual Symposium on Cloud Computing*, page 5. ACM, 2013.

Shivaram Venkataraman, Zongheng Yang, Michael Franklin, Benjamin Recht, and Ion Stoica. Ernest: Efficient performance prediction for large-scale advanced analytics. In *13th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 16)*, pages 363–378, 2016.

Martin J Wainwright. *High-dimensional Statistics: a Non-asymptotic Viewpoint*, volume 48. Cambridge University Press, 2019.

Rich Wolski, James S Plank, John Brevik, and Todd Bryan. Analyzing market-based resource allocation strategies for the computational grid. *The International Journal of High Performance Computing Applications*, 15(3):258–281, 2001.

Seyed Majid Zahedi, Qiuyun Llull, and Benjamin C Lee. Amdahl's law in the datacenter era: a market for fair processor allocation. In *2018 IEEE International Symposium on High Performance Computer Architecture (HPCA)*, pages 1–14. IEEE, 2018.

Li Zhang. Proportional response dynamics in the fisher market. *Theoretical Computer Science*, 412 (24):2691–2698, 2011.

A   Technical Lemmas

We first provide some useful technical lemmas.

Lemma A.1. *Suppose that Z is a χ_m^2 random variable, i.e. $Z = \sum_{k=1}^m Z_k^2$, where for all k, Z_k are i.i.d. random variables drawn from $\mathcal{N}(0,1)$. Then,*

$$P(Z > m + \alpha) = \begin{cases} e^{-\frac{\alpha}{8}} & \alpha > m \\ e^{-\frac{\alpha^2}{8m}} & \alpha \leqslant m. \end{cases}$$

Proof. Suppose that X is sub-exponential random variable with parameters (ν, b) and expectation μ. Applying well known tail bounds for sub-exponential random variables (e.g. (Wainwright, 2019)) yields:

$$P(X > \mu + \alpha) = \begin{cases} e^{-\frac{\alpha^2}{2\nu^2}} & 0 \leqslant \alpha \leqslant \frac{\nu^2}{b}, \text{ and} \\ e^{-\frac{\alpha}{2b}} & \alpha > \frac{\nu^2}{b}. \end{cases}$$

The lemma follows from the fact that a χ_m^2 random variable is sub-exponential with parameters $(\nu, b) = (2, 4)$. ∎

Lemma A.2. *(Lower bound for normal distributions) Let Z be a random variable $Z \sim \mathcal{N}(0,1)$, then $P(Z > t) \geqslant \frac{1}{t+\sqrt{t^2+4}} \sqrt{\frac{2}{\pi}} e^{-\frac{t^2}{2}}$.*

Proof. First, from Abramowitz et al. (1988) (7.1.13) we have,

$$e^{x^2} \int_x^\infty e^{-t^2} dt \geqslant \frac{1}{x + \sqrt{x^2 + 2}}.$$

Set $t = \sqrt{2}x$, then the above equation yields:

$$P(Z > t) = \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-\frac{x^2}{2}} dx \geqslant \frac{1}{t + \sqrt{t^2 + 4}} \sqrt{\frac{2}{\pi}} e^{-\frac{t^2}{2}},$$

which completes the proof. ∎

Lemma A.3. *(Azuma-Hoeffding inequality(Wainwright, 2019)) Let $(Z_s)_{s \geqslant 0}$ be a super martingale w.r.t. a filtration $(\mathcal{F}_t)_{t \geqslant 0}$. Let $(B_t)_{t \geqslant 0}$ be predictable processes w.r.t. $(\mathcal{F}_t)_{t \geqslant 0}$, such that $|Z_s - Z_{s-1}| \leqslant B_s$ for all $s \geqslant 1$ almost surely. Then for any $\delta > 0$,*

$$P\left(Z_T - Z_0 \leqslant \sqrt{2 \log\left(\frac{1}{\delta}\right) \sum_{t=1}^T B_t^2} \right) \geqslant 1 - \delta.$$

Lemma A.4. $\forall x \in [0, c], c > 0$, *we have* $x \leqslant \frac{c}{\log(1+c)} \log(1 + x)$.

Proof. The result follows immediately from the fact that the function $f(x) \overset{\text{def}}{=} \frac{x}{\log(1+x)}$ is non-decreasing on $(0, \infty)$. ∎

Lemma A.5. *(Lemma 1, Filippi et al. (2010)) Let $(\mathcal{F}_k, k \geqslant 0)$ be a filtration, $(m_k; k \geqslant 0)$ be an $\mathbb{R}^d$-valued stochastic process adapted to $(\mathcal{F}_k)$. Assume that η_k is conditionally sub-Gaussian in the sense that there exists some $R > 0$ such that for any $\gamma \geqslant 0$, $k \geqslant 1$, $\mathbb{E}[\exp(\gamma \eta_k)|\mathcal{F}_{k-1}] \leqslant \exp\left(\frac{\gamma^2 R^2}{2}\right)$ almost surely. Then, consider the martingale $\xi_t = \sum_{k=1}^t m_{k-1} \eta_k$ and the process $M_t = \sum_{k=1}^t m_{k-1} m_{k-1}^\top$. Assume that with probability one, the smallest eigenvalue of M_d is lower bounded by some positive constant λ_0, and that $\|m_k\|_2 \leqslant c_m$ almost surely for any $k \geqslant 0$. Then, the following holds true: for any $0 < \delta < \min(1, d/e)$ and $t > \max(d, 2)$, with probability at least $1 - \delta$,*

$$\|\xi_t\|_{M_t^{-1}} \leqslant \kappa R \sqrt{2d \log(t) \log(d/\delta)},$$

where $\kappa = \sqrt{3 + 2\log(1 + 2\frac{c_m^2}{\lambda_0})}$.

B Bounding L^{CE}

First, consider any round t. We will let $\mathcal{F}_t \overset{\text{def}}{=} \sigma\left(\{(x_{is}, y_{is})_{i=1,s=1}^{n,t-1}\}\right)$ denote the σ-algebra generated by the observations in the first $t-1$ rounds. Clearly, $\{\mathcal{F}_t\}_{t\geqslant 0}$ is a filtration. We will denote $\mathbb{E}_t[\cdot|\mathcal{F}_t] = \mathbb{E}_t[\cdot]$ to be the expectation when conditioning on the past observations up to round $t-1$. Similarly, define $P_t(\cdot) \overset{\text{def}}{=} P(\cdot|\mathcal{F}_t) = \mathbb{E}[\mathbb{I}(\cdot)|\mathcal{F}_t]$.

Recall that $\{\delta_t\}_{t\geqslant 0}$ are inputs to the algorithm. Similarly, let $\{\delta_{2t}\}_{t\geqslant 0}$ be a sequence. We will specify values for both sequences later in this proof. Given these, further define the following quantities on round t:

$$\beta_{1t} \overset{\text{def}}{=} \frac{2}{C_\mu}\kappa\sigma\sqrt{2m\log(t)}\sqrt{\log\left(\frac{m}{\delta_t}\right)}$$

$$\beta_{2t} \overset{\text{def}}{=} \sqrt{\alpha_t(m+\gamma_{2t})} \qquad \text{where,} \qquad \gamma_{2t} \overset{\text{def}}{=} \max\left(8\log\left(\frac{1}{\delta_{2t}}\right), \sqrt{8m\log\left(\frac{1}{\delta_{2t}}\right)}\right)$$

$$\beta_{3t} \overset{\text{def}}{=} L_\mu(\beta_{1t} + \beta_{2t}).$$

Here, recall that L_μ is the Lipschitz constant of $\mu(\cdot)$, C_μ is such that $C_\mu \overset{\text{def}}{=} \inf_{\theta\in\Theta, x\in\mathcal{X}} \dot{\mu}\left(\theta^\top\phi(x)\right)$, and α_t is a sequence that is defined and used in Algorithm 1.

Next, we consider the following two events:

$$A_{it} \overset{\text{def}}{=} \{\|\theta_i^* - \bar{\theta}_{it}\|_{Q_{it}} \leqslant \beta_{1t}\},$$
$$B_{it} \overset{\text{def}}{=} \{\|\bar{\theta}_{it} - \theta_{it}\|_{Q_{it}} \leqslant \beta_{2t}\}.$$

where $Q_{it} \overset{\text{def}}{=} \sum_{s=1}^{t-1} \phi(x_{is})\phi(x_{is})^\top$ is a design matrix that corresponding to the first $t-1$ steps.

Lastly, define

$$\rho_{it}(x) \overset{\text{def}}{=} \|\phi(x)\|_{Q_{it}^{-1}} = \sqrt{\phi^\top(x)Q_{it}^{-1}\phi(x)},$$

and

$$S_{it} \overset{\text{def}}{=} \{x \in \mathcal{X} : u_i(x_{it}^*) - u_i(x) \geqslant \beta_{3t}\rho_{it}(x)\},$$

where $x_{it}^* = \arg\max_{y\in\mathcal{X}, p_t^\top y \leqslant p_t^\top e_i} u_i(y)$. Here, we used $\mathcal{X}$ to denote the set of feasible allocations for one agent: $\{x \in \mathbb{R}^m : 0 \leqslant x \leqslant \mathbf{1}\}$.

Intuitively, $\tilde{x}_{it}$ is the best true optimal affordable allocation for agent i in round t under the price function p_t. Since the set $\{y \in \mathcal{X}, p_t^\top y \leqslant p_t^\top e_i\}$ is a compact set, the maximum is well defined.

Now we begin our analysis with the following lemmas.

Lemma B.1. *For any round $t > t_0$, $P_t(A_{it}) \geqslant 1 - \delta_{1t}$.*

Proof. Define function $g_{it}(\theta) = \sum_{s=1}^{t-1} \mu\left(\theta^\top\phi(x_{is})\right)\phi(x_{is})$. Then by the fundamental theorem of calculus, we have

$$g_{it}(\theta_{it}) - g_{it}(\bar{\theta}_{it}) = G_{it}(\theta_i^* - \bar{\theta}_{it}),$$

where $G_{it} = \int_0^1 \nabla g_{it}\left(s\theta_i^* + (1-s)\bar{\theta}_{it}\right)ds$, and

$$\nabla g_{it}(\theta) = \sum_{s=1}^{t-1} \phi(x_{is})\phi(x_{is})^\top\mu'(\theta^\top\phi(x_{is})).$$

By the definition of C_μ and Q_{it}, we have that $G_{it} \succeq C_\mu Q_{it} \succeq M \cdot I$, where the last inequality follows due to the initialisation scheme. Therefore, G_{it} is invertible and moreover,

$$G_{it}^{-1} \preceq \frac{1}{C_\mu}Q_{it}^{-1}. \tag{7}$$

We can write,

$$\theta_i^* - \bar{\theta}_{it} = G_{it}^{-1}\left(g_{it}(\theta_{it}) - g_{it}(\bar{\theta}_{it})\right). \tag{8}$$

Therefore, we have,

$$
\begin{aligned}
&\left(\theta_i^* - \bar{\theta}_{it}\right)^\top Q_{it}\left(\theta_i^* - \bar{\theta}_{it}\right) \\
&= \left(g_{it}(\theta_{it}) - g_{it}(\bar{\theta}_{it})\right)^\top G_{it}^{-1} Q_{it} G_{it}^{-1}\left(g_{it}(\theta_{it}) - g_{it}(\bar{\theta}_{it})\right) \\
&\leqslant \frac{1}{C_\mu^2}\left(g_{it}(\theta_{it}) - g_{it}(\bar{\theta}_{it})\right) Q_{it}^{-1}\left(g_{it}(\theta_{it}) - g_{it}(\bar{\theta}_{it})\right) \\
&= \frac{1}{C_\mu^2}\left\|\left(g_{it}(\theta_{it}) - g_{it}(\bar{\theta}_{it})\right)\right\|_{Q_{it}^{-1}},
\end{aligned}
$$

where the first equality follows from Eq (7), and the inequality follows from Eq (8).

Therefore,

$$
\begin{aligned}
\|\theta_i^* - \bar{\theta}_{it}\|_{Q_{it}} &\leqslant \frac{1}{C_\mu}\left\|\left(g_{it}(\theta_{it}) - g_{it}(\bar{\theta}_{it})\right)\right\|_{Q_{it}^{-1}} \\
&\leqslant \frac{2}{C_\mu}\left\|\left(g_{it}(\theta_{it}) - g_{it}(\bar{\theta}_{it}^{MLE})\right)\right\|_{Q_{it}^{-1}} \\
&= \frac{2}{C_\mu}\left\|\sum_{s=1}^{t-1}\phi(x_{is})(Y_{is} - \mu\left(\phi(x_{is})^\top\theta_i^*\right))\right\|_{Q_{it}^{-1}}.
\end{aligned}
$$

where the second inequality is from the triangle inequality, and the last equality is from the definition of θ_{it}^{MLE} and g_{it}.

Let A_{it} denote the event that

$$\left\|\sum_{s=1}^{t-1}\phi(x_{is})(Y_{is} - \mu\left(\phi(x_{is})^\top\theta_i^*\right))\right\|_{Q_{it}^{-1}} \leqslant \kappa\sigma\sqrt{2m\log(t)}\sqrt{\log\left(\frac{d}{\delta_t}\right)},$$

then we have A_{it} holds with probability at least δ_t by Lemma A.5.

∎

Lemma B.2. *For any round $t > t_0$, $P_t(B_{it}) \geqslant 1 - \delta_{2t}$.*

Proof. First, recall that $B_{it} = \{\|\bar{\theta}_{it} - \theta_{it}\|_{Q_{it}} \leqslant \beta_{2t}\}$. We can now write,

$$
\begin{aligned}
P(B_{it}^c) &= P(\|\theta_{it} - \bar{\theta}_{it}\|_{Q_{it}} > \beta_{2t}) \\
&\leqslant P(\|\theta_{it}' - \bar{\theta}_{it}\|_{Q_{it}} > \beta_{2t}) \\
&= P(\|\theta_{it}' - \bar{\theta}_{it}\|_{\alpha_t^{-1}Q_{it}} > \alpha_t^{-\frac{1}{2}}\beta_{2t}) \\
&= P(\sqrt{Z} > \sqrt{M_t\gamma_{2t}}),
\end{aligned}
$$

where $Z = \left(\theta_{it} - \bar{\theta}_{it}\right)^\top \alpha_t^{-2} Q_{it}\left(\theta_{it} - \bar{\theta}_{it}\right)$. The first step simply uses the fact that since $\bar{\theta}_{it}$ is already inside Θ (see line 17 in Algorithm 1), projecting θ_{it}' to be inside Θ after sampling only brings it *even closer* to $\bar{\theta}_{it}$.

Note that Z is a χ_m^2 random variable. This follows from the fact that

$$\theta_{it} \sim \mathcal{N}(\bar{\theta}_{it}, \alpha_t^2 Q_{it}^{-1}),$$

therefore we have

$$\alpha_t^{-1}Q_{it}^{1/2}\left(\theta_{it} - \bar{\theta}_{it}\right) \sim \mathcal{N}(0, I_m).$$

Denote $y = \alpha_t^{-1} Q_{it}^{1/2} \left(\theta_{it} - \bar{\theta}_{it} \right)$, then $Z = y^\top y$ is a χ_m^2 random variable.

Therefore, by Lemma A.1, and the definition that $\gamma_{2t} = \max \left(8 \log(\frac{1}{\delta_{2t}}), \sqrt{\log(\frac{1}{\delta_{2t}})} \right)$, we have

$$P(B_{it}^c) = P(Z > n + \gamma_{2t}) \leqslant \delta_{2t},$$

which completes the proof. ∎

Lemma B.3. *Let x be arbitrary such that $x \in \mathcal{X}$. Then,*

$$P_t(u_{it}(x) > u_i(x)|A_{it}) \geqslant q_0,$$

with $q_0 = \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{2}+\sqrt{6}} \frac{1}{e} \approx 0.075$.

Proof. First, notice that

$$
\begin{aligned}
u_{it}(x) > u_i(x) &\iff \mu\left(\theta_{it}^\top \phi(x) \right) > \mu\left((\theta_i^*)^\top \phi(x) \right) \\
&\iff \theta_{it}^\top \phi(x) > (\theta_i^*)^\top \phi(x) \\
&\iff \frac{\left(\theta_{it} - \bar{\theta}_{it} \right)^\top \phi(x)}{\alpha_t \rho_{it}(x)} > \frac{\left(\theta_i^* - \bar{\theta}_{it} \right)^\top \phi(x)}{\alpha_t \rho_{it}(x)}.
\end{aligned}
$$

Since $\theta_{it} \sim \mathcal{N}(\bar{\theta}_{it}, \alpha_t^2 Q_{it}^{-1})$, we have

$$\left(\theta_{it} - \bar{\theta}_{it} \right)^\top \phi(x) \sim \mathcal{N}(0, \alpha_t^2 \phi(x)^\top Q_{it}^{-1} \phi(x)), \implies \left(\theta_{it} - \bar{\theta}_{it} \right)^\top \phi(x) \sim \mathcal{N}(0, \alpha_t^2 \rho_{it}^2(x)).$$

$$\implies \frac{\left(\theta_{it} - \bar{\theta}_{it} \right)^\top \phi(x)}{\alpha_t \rho_{it}(x)} \sim \mathcal{N}(0, 1).$$

From the above we have that

$$P_t\left(u_{it}(x) > u_i(x)|A_{it} \right) = P_t\left(Z > \frac{\left(\theta_{it} - \bar{\theta}_{it} \right)^\top \phi(x)}{\alpha_t \rho_{it}(x)} \middle| A_{it} \right),$$

where $Z \sim \mathcal{N}(0, 1)$ is sampled independently of the observations, since the randomness in Algorithm 1 can be assumed to be independent of the randomness in the observations. Therefore, under the event A_{it},

$$
\begin{aligned}
\left| \frac{\left(\theta_{it} - \bar{\theta}_{it} \right)^\top \phi(x)}{\alpha_t \rho_{it}(x)} \right| &= \left| \frac{\left(\theta_{it} - \bar{\theta}_{it} \right)^\top Q_{it}^{\frac{1}{2}} Q_{it}^{-\frac{1}{2}} \phi(x)}{\alpha_t \rho_{it}(x)} \right| \\
&\leqslant \frac{\|\theta_{it} - \bar{\theta}_{it}\|_{Q_{it}} \|\phi(x)\|_{Q_{it}^{-1}}}{\alpha_t \rho_{it}(x)} \\
&\leqslant \frac{\beta_{it}}{\alpha_t} = \frac{\sqrt{8}}{\alpha_0}.
\end{aligned}
$$

Here, the first inequality follows from the definition of the matrix norm and the definition of A_{it}, and the second inequality follows from the definition of β_{1t}.

Therefore, by Lemma A.2, we have

$$
\begin{aligned}
&P_t(u_{it}(x) > u_i(x)|A_{it}) \\
&= P_{Z \sim \mathcal{N}(0,1)}(Z > \frac{\sqrt{8}}{\alpha_0}) \\
&\geqslant \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{8/\alpha_0^2} + \sqrt{4 + \sqrt{8/\alpha_0^2}}} e^{-\frac{4}{\alpha_0^2}}.
\end{aligned}
$$

Setting $\alpha_0^2 = 4$, we have

$$P_t\left(u_{it}(x) > u_i(x)|A_{it} \right) \geqslant \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{2} + \sqrt{6}} \frac{1}{e} \approx 0.075,$$

which completes the proof. ∎

Lemma B.4. *Let $\theta_1, \theta_2 \in \Theta \in \mathbb{R}^m$. Let $Q \succeq 0, Q \in \mathbb{R}^{m \times m}$ be a positive semi-definite matrix, and $\rho_Q(x) = \sqrt{\phi(x)^\top Q^{-1} \phi(x)}$. Then,*

$$\left| \mu(\theta_1^\top \phi(x)) - \mu(\theta_2^\top \phi(x)) \right| \leqslant L_\mu \|\theta_1 - \theta_2\|_Q \cdot \rho_Q(x).$$

Proof. This follows from the Lipschitz properties of μ and the following simple calculations:

$$
\begin{aligned}
\left| \mu(\theta_1^\top \phi(x)) - \mu(\theta_2^\top \phi(x)) \right| &\leqslant L_\mu |(\theta_1 - \theta_2)^\top \phi(x)| \\
&= L_\mu |(\theta_1 - \theta_2)^\top Q^{\frac{1}{2}} Q^{-\frac{1}{2}} \phi(x)| \\
&\leqslant L_\mu \|\theta_1 - \theta_2\|_Q \|\phi(x)\|_{Q^{-1}} \\
&= L_\mu \|\theta_1 - \theta_2\|_Q \cdot \rho_Q(x).
\end{aligned}
$$

∎

Lemma B.5. *For any round $t > t_0$, $P_t(x_{it} \notin \mathcal{S}_{it}) \geqslant q_0(1 - \delta_t) - \delta_{2t}$.*

Proof. First, when event B_{it} holds, by lemma B.4, we have that for all x,

$$|u_{it}(x) - \bar{u}_{it}(x)| \leqslant L_\mu \beta_{2t} \rho_{it}(x).$$

Note that by definition, $u_{it}(x) = \mu\left((\theta_{it})^\top \phi(x)\right)$, and $\bar{u}_{it}(x) = \mu\left((\bar{\theta}_{it})^\top \phi(x)\right)$. Therefore,

$$\bar{u}_{it}(x) - u_{it}(x) > -L_\mu \beta_{2t} \rho_{it}(x). \tag{9}$$

On the other side, under event A_{it}, by lemma B.4, we have that for all x,

$$|u_i(x) - \bar{u}_{it}(x)| \leqslant L_\mu \beta_{1t} \rho_{it}(x). \tag{10}$$

Moreover, recall that by definition for any $x \in \mathcal{S}_{it}$,

$$u_i(x_{it}^*) - u_i(x) \geqslant \beta_{3t} \rho_{it}(x). \tag{11}$$

Therefore, consider any $x \in \mathcal{S}_{it}$, and under the condition that $A_{it} \cap B_{it} \cap \{u_{it}(x_{it}^*) > u_i(x_{it}^*)\}$, we have

$$
\begin{aligned}
u_{it}(x_{it}^*) - u_{it}(x) &> u_i(x_{it}^*) - u_{it}(x) \\
&= (u_i(x_{it}^*) - u_i(x)) + (u_i(x) - \bar{u}_{it}(x)) + (\bar{u}_{it}(x) - u_{it}(x)) \\
&> 0,
\end{aligned}
\tag{12}
$$

where the last inequality follows from combining equations Eq (9), Eq (10), Eq (11) and the definition of β_{3t}. Hence, Eq (12) implies that, under the same condition, $x_{it} \notin \mathcal{S}_{it}$ since by construction, x_{it} maximizes u_{it} under the budget, thus

$$u_{it}(x_{it}) \geqslant u_{it}(x_{it}^*),$$

This further implies that,

$$
\begin{aligned}
P_t(x_{it} \notin \mathcal{S}_{it}) &\geqslant P_t\left(u_{it}(x_{it}^*) > u_{it}(x), \forall x \in \mathcal{S}_{it}\right) \\
&\geqslant P_t\left(u_{it}(x_{it}^*) > u_{it}(x), \forall x \in \mathcal{S}_{it} | A_{it} \cap B_{it}\{u_i(x_{it}^*) > u_i(x)\}\right) \\
&\qquad \times P(A_{it} \cap B_{it} \cap \{u_i(x_{it}^*) > u_i(x)\}) \\
&= P(A_{it} \cap B_{it} \cap \{u_i(x_{it}^*) > u_i(x)\}) \\
&\geqslant P(A_{it} \cap \{u_i(x_{it}^*) > u_i(x)\}) - P(B_{it}^c) \\
&= P(\{u_i(x_{it}^*) > u_i(x)\} | A_{it}) P(A_{it}) - P(B_{it}^c) \\
&\geqslant q_0(1 - \delta_t) + \delta_{2t}.
\end{aligned}
$$

Here, the second and third inequality both from the law of total probability and rearranging terms, and the last inequality follows from Lemma B.1, Lemma B.2 and Lemma B.3, which completes the proof. ∎

Lemma B.6. *For $t \geqslant \max(t_0, t_0')$,*

$$\mathbb{E}_t[\ell_{it}] \leqslant \frac{5}{q_0}\beta_{3t}\mathbb{E}_t[\rho_{it}(x_{it})] + \delta_t + \delta_{2t}.$$

Here, $\ell_{it} = (u_i(x_{it}^*) - u_i(x_{it}))^+$, and t_0' is chosen such that, $\forall t > t_0'$, $\delta_t < \frac{1}{4}$, $\delta_{2t} < \frac{q_0}{4}$.

Proof. First, define

$$x_{it}' = \underset{x:p_t^\top \leqslant p_t^\top e_i, x \notin \mathcal{S}_{it}}{\arg\min} \rho_{it}(x).$$

This implies that,

$$\mathbb{E}_t[\rho_{it}(x_{it})] \geqslant \mathbb{E}_t[\rho_{it}(x_{it})|x_{it} \notin \mathcal{S}_{it}]P(x \notin \mathcal{S}_{it}) \geqslant \rho_{it}(x_{it}')P(x \notin \mathcal{S}_{it}).$$

Therefore, by Lemma B.5, we have

$$\rho_{it}(x_{it}') \leqslant \frac{\mathbb{E}_t[\rho_{it}(x_{it})]}{q_0(1-\delta_t) - \delta_{2t}}.$$

Select t_0' such that, $\forall t \geqslant t_0'$, $\delta_t = \frac{1}{4}$, and $\delta_{2t} \leqslant \frac{q_0}{4}$, then we have:

$$\rho_{it}(x_{it}') \leqslant \frac{2}{q_0}\mathbb{E}_t[\rho_{it}(x_{it})].$$

Also, under $A_{it} \cap B_{it}$,

$$\|\theta_i^* - \theta_{it}\|_{Q_{it}} \leqslant \|\theta_i^* - \bar{\theta}_{it}\|_{Q_{it}} + \|\bar{\theta}_{it} - \theta_{it}\|_{Q_{it}} \leqslant \beta_{1t} + \beta_{2t},$$

where the first inequality follows from triangle inequality, and the second one follows from the definitions of A_{it} and B_{it}. Hence,

$$|\mu_{it}(x) - \mu_i(x)| \leqslant L_\mu(\beta_{1t} + \beta_{2t})\rho_{it}(x) = \beta_{3t}\rho_{it}(x).$$

Therefore, we have

$$\begin{aligned}
\ell_{it} &= u_i(x_{it}^*) - u_i(x_{it}) \\
&= u_i(x_{it}^*) - u_i(x_{it}') + u_i(x_{it}') - u_i(x_{it}) \\
&\leqslant 2\beta_{3t}\rho_{it}(x_{it}') + \beta_{3t}\rho_{it}(x_{it}) \\
&\leqslant \frac{4}{q_0}\beta_{3t}\mathbb{E}[\rho_{it}(x_{it})] + \beta_{3t}\rho_{it}(x_{it}).
\end{aligned}$$

which further yields

$$\begin{aligned}
\mathbb{E}_t[\ell_{it}] &\leqslant \mathbb{E}_t[\ell_{it}|A_{it} \cap B_{it}] + \mathbb{E}_t[\ell_{it}|A_{it}^c \cup B_{it}^c]P(A_{it}^c \cup B_{it}^c) \\
&\leqslant \frac{4}{q_0}\beta_{3t}\mathbb{E}[\rho_{it}(x_{it})] + \beta_{3t}\mathbb{E}_t[\rho_{it}(x_{it})] + \delta_t + \delta_{2t} \\
&\leqslant \frac{5}{q_0}\beta_{3t}\mathbb{E}_t[\rho_{it}(x_{it})] + \delta_t + \delta_{2t},
\end{aligned}$$

(13)

which completes the proof. ∎

Lemma B.7. *Let $\delta' > 0$. Define $L_{iT} = \sum_{t=1}^T \ell_{it}$. Then, with probability at least $1 - \delta'$,*

$$L_{iT} \leqslant \sum_{t=1}^T (\delta_t + \delta_{2t}) + \tilde{O}\left(\frac{m^2}{\sqrt{M}}\sqrt{T}\left(\log(T) + \log\left(\frac{1}{\delta'}\right)\right)\right)$$

Proof. First, define for $s > 1$,

$$u_{is} = \ell_{is} - \frac{5\beta_{3s}}{q_0}\rho_{is}(x_{is}) - (\delta_s + \delta_{2s}),$$

and $v_{it} = \sum_{s=1}^t u_{is}$, with $v_{i0} = 0$ and $u_{i0} = 0$. We show that $\{v_{it}\}, t \geqslant 0$ is a super-martingale with respect to the filtration $(\mathcal{F}_t)_{t \geqslant 0}$.

First,

$$\mathbb{E}_t[u_{it}] = \mathbb{E}_t[\ell_{it}] - \frac{5\beta_{3t}}{q_0}\mathbb{E}_t[\rho_{it}(x_{it})] - (\delta_t + \delta_{2t}) \leqslant 0.$$

Moreover,

$$|v_{it} - v_{i,t-1}| \leqslant |\ell_{it}| + \frac{5\beta_{3t}}{q_0}|\rho_{it}(x_{it})| + (\delta_t + \delta_{2t})$$

$$\leqslant 1 + \frac{5\beta_{3t}}{q_0}\frac{\|\phi(1)\|_2}{\sqrt{M}} + 1$$

$$\leqslant \frac{7\beta_{3t}}{q_0}\frac{\|\phi(1)\|_2}{\sqrt{M}} \triangleq D_t.$$

Therefore, by Lemma A.3, with probability at least $1 - \delta'$, we have that

$$v_{iT} - v_{i,0} \leqslant \sqrt{s\log\left(\frac{1}{\delta'}\sum_{t-1}^{T}D_t^2\right)} \leqslant \frac{7\beta_{3t}}{q_0}\|\phi(1)\|_2\sqrt{\frac{2\log(\frac{1}{\delta'})}{M}T}.$$

Therefore, we have

$$L_{iT} = \sum_{s=1}^{T}\ell_{is} \leqslant \sum_{t=1}^{T}(\delta_t + \delta_{2t}) + \frac{5\beta_{3t}}{q_0}\sum_{t=1}^{T}\rho_{it}(x_t) + \frac{7\beta_{3t}}{q_0}\|\phi(1)\|_2\sqrt{\frac{2\log(\frac{1}{\delta'})}{M}T}. \tag{14}$$

Now it remains to bound $\sum_{t=1}^{T}\rho_{it}(x_t)$. Since $Q_{it} \succeq MI$, by the definition of $\rho_{it}(x_{it})$, we have

$$\rho_{it}x_{it} \leqslant \frac{\|\phi(1)\|_2}{\sqrt{M}}.$$

Hence, by Lemma A.4 and rearranging terms, we have

$$\sum_{t=t_0}^{T}\rho_{it}^2(x_{it}) \leqslant \frac{\|\phi(1)\|_2^2}{M}\frac{1}{\log\left(1 + \frac{\|\phi(1)\|_2}{\sqrt{M}}\right)}\sum_{t=t_0}^{T}\log\left(1 + \phi^\top(x_{it})Q_{it}^{-1}\phi(x_{it})\right). \tag{15}$$

Also notice that,

$$\sum_{t=t_0}^{T}\log\left(1 + \phi^\top(x_{it})Q_{it}^{-1}\phi(x_{it})\right) = \log\Pi_{t=t_0}^{T}\left(1 + \|\phi(x_{it})\|_{Q_{it}^{-1}}^2\right) = \log\frac{\det(Q_{iT})}{\det(Q_{i,t_0})}.$$

Note that the trace of $Q_{i,t+1}$ is upper-bounded by $t \cdot \|\phi(1)\|_2$, then given that the trace of the positive definite matrix Q_{iT} is equal to the sum of its eigenvalues, we have that $\det(Q_{iT}) \leqslant \left(t\|\phi(1)\|_2^2\right)^m$. Moreover, $\det(Q_{i,t_0}) \geqslant (M)^m$, therefore,

$$\sum_{t=t_0}^{T}\log\left(1 + \phi^\top(x_{it})Q_{it}^{-1}\phi(x_{it})\right) \leqslant m\log\left(\frac{\|\phi(1)\|_2^2 T}{M}\right).$$

Combining with Eq (15), and applying Cauchy-Schwartz inequality, we have

$$\sum_{t=t_0}^{T}\rho_{it}(x_{it}) \leqslant \sqrt{T\sum_{t=t_0}^{T}\rho_{it}^2(x_{it})} \leqslant \sqrt{T}\sqrt{\frac{\|\phi(1)\|_2^2}{M}\frac{m}{\log\left(1 + \frac{\|\phi(1)\|_2}{\sqrt{M}}\right)}\log\left(\frac{\|\phi(1)\|_2^2 T}{M}\right)}. \tag{16}$$

Putting together Eq (14) and Eq (16), with the fact that $\|\phi(\mathbf{1})\|_2 = O(\sqrt{m})$, e.g. $\|\phi(\mathbf{1})\|_2 \leqslant c_\phi \sqrt{m}$, we have that with probability at least $1 - \delta'$,

$$
\begin{aligned}
L_{iT} &\leqslant \sum_{t=1}^{T} (\delta_t + \delta_{2t}) + \frac{\beta_{3t}}{q_0} \frac{\sqrt{T}}{\sqrt{M}} \left(5c_\phi m \sqrt{\frac{1}{\log\left(1 + \frac{\|\phi(\mathbf{1})\|_2}{m}\right)} \log\left(\frac{\|\phi(\mathbf{1})\|_2^2 T}{M}\right)} + 7c_\phi \sqrt{m} \sqrt{2\log(\frac{1}{\delta'})} \right) \\
&= \sum_{t=1}^{T} (\delta_t + \delta_{2t}) + \tilde{O}\left(\frac{m^2}{\sqrt{M}} \sqrt{T} \left(\log(T) + \log\left(\frac{1}{\delta'}\right) \right) \right),
\end{aligned}
$$

where the last step comes from the fact that $\beta_{3t} = \tilde{O}(m)$. This completes the proof. ∎

B.1 Proof of Theorem 4.1 for L^{CE}

Corollary B.8. *With probability at least $1 - \delta'$,*

$$
L_T^{\mathrm{CE}} \leqslant n \sum_{t=1}^{T} (\delta_{1t} + \delta_{2t}) + \tilde{O}\left(n \frac{m^2}{\sqrt{M}} \sqrt{T} \left(\log(T) + \log(\frac{1}{\delta'}) \right) \right)
$$

Proof. This a direct result from Lemma B.7 and the definition of L_T^{CE}: With probability at least $1 - \delta'$,

$$
\begin{aligned}
L_T^{\mathrm{CE}} &= \sum_{i=1}^{n} \sum_{t=1}^{T} \left(\max_{y:p(y) \leqslant p(e_i)} u_i(y) - u_i(x_{it}) \right)^{+} \\
&\leqslant n \sum_{t=1}^{T} (\delta_t + \delta_{2t}) + \tilde{O}\left(n \frac{m^2}{\sqrt{M}} \sqrt{T} \left(\log(T) + \log\left(\frac{1}{\delta'}\right) \right) \right).
\end{aligned}
$$

∎

Proof. Choose $\delta_t = \delta_{2t} = \frac{2\delta}{n\pi^2 t^2}$. Then,

$$
\sum_{t=1}^{T} (\delta_t + \delta_{2t}) \leqslant \frac{2}{3} \delta.
$$

Also choose $\delta' = \frac{\delta}{3}$, then by Corollary B.8, with probability at least $1 - \delta$,

$$
L_T^{\mathrm{CE}} = O\left(n \frac{m^2}{\sqrt{M}} \sqrt{T} \left(\log\left(\frac{\delta}{3}\right) + \log(T) \right) \right).
$$

∎

B.2 Proof of Theorem 4.2 for L^{CE}

Proof. Choose $\delta_{1t} = \delta_{2t} = \delta' = \frac{1}{T}$. Denote the event where $L_T^{\mathrm{CE}} = O\left(n \frac{m^2}{\sqrt{M}} \sqrt{T} \left(\log\left(\frac{\delta}{3}\right) + \log(T) \right) \right)$ holds as $\mathcal{E}$. Then, by Lemma B.7,

$$
\mathbb{E}[L_T^{\mathrm{CE}}] = \mathbb{E}[L_T^{\mathrm{CE}}|\mathcal{E}] + \mathbb{E}[L_T^{\mathrm{CE}}|\mathcal{E}^c] P(\mathcal{E}^c) \leqslant 2 + O\left(n \frac{m^2}{\sqrt{M}} \sqrt{T} \left(\log\left(\frac{\delta}{3}\right) + \log(T) \right) \right).
$$

where $M \geqslant m$. This completes the proof. ∎

C   Bounding L^{FD}

Recall that the definition of ℓ^{FD} is directly based on the requirements of Pareto efficiency and fair share: $\ell^{\mathrm{FD}}(x) \stackrel{\mathrm{def}}{=} \max\left(\ell^{\mathrm{PE}}(x), \ell^{\mathrm{SI}}(x)\right)$, where $\ell^{\mathrm{PE}}(x) = \min_{x' \in \mathcal{PE}} \sum_{i=1}^{n} \left(u_i(x_i') - u_i(x_i)\right)^+$; and $\ell^{\mathrm{SI}}(x) = \sum_{i=1}^{n} \left(u_i(e_i) - u_i(x_i)\right)^+$.

To bound L^{FD}, we first provide a useful lemma which shows that ℓ^{SI} is a weaker notion than ℓ^{CE}.

Lemma C.1. *For any allocation x and price p, $\ell^{\mathrm{SI}}(x) \leqslant \ell^{\mathrm{CE}}(x, p)$.*

Proof. This simply uses the fact that an agent's endowment is always affordable under any price vector p. Therefore,

$$
\begin{aligned}
\ell^{\mathrm{SI}}(x) &= \sum_{i=1}^{n} \left(u_i(e_i) - u_i(x_i)\right)^+ \\
&\leqslant \sum_{i=1}^{n} \left(\max_{y:p^\top y \leqslant p^\top e_i} u_i(y) - u_i(x)\right)^+, \forall p \\
&= \ell^{\mathrm{CE}}(x, p),
\end{aligned}
$$

∎

Having lemma C.1 at hand, the key remaining task is to bound ℓ^{PE}. We will show that this can be achieved by an analogous analysis as in Section B.1, but with some key differences.

First, we define $\tilde{\mathcal{S}}_{it}$ (in comparison to $\mathcal{S}_{it}$ used in Section B.1):

$$
\tilde{\mathcal{S}}_{it} \stackrel{\mathrm{def}}{=} \left\{x \in \mathcal{X} : u_i(x_i^*) - u_i(x) \geqslant \beta_{3t} \rho_{it}(x)\right\},
$$

where $x^* \in \mathbb{R}^{n \times m}$ is the unique equilibrium allocation. Note that $\tilde{\mathcal{S}}_{it}$ shares a similar spirit as $\mathcal{S}_{it}$, which is used in Section B.1, but with a different referencing point x^*.

We show a key lemma which provides a lower bound on $P(x \notin \tilde{\mathcal{S}}_{it})$.

Lemma C.2. *For any round $t > t_0$, $P_t(x_{it} \notin \tilde{\mathcal{S}}_{it}) \geqslant q_0(1 - \delta_t) - \delta_{2t}$.*

Proof. First, when event B_{it} holds, by lemma B.4, we have that for all x,

$$
|u_{it}(x) - \bar{u}_{it}(x)| \leqslant L_\mu \beta_{2t} \rho_{it}(x).
$$

Note that by definition, $u_{it}(x) = \mu\left((\theta_{it})^\top \phi(x)\right)$, and $\bar{u}_{it}(x) = \mu\left((\bar{\theta}_{it})^\top \phi(x)\right)$. Therefore,

$$
\bar{u}_{it}(x) - u_{it}(x) > -L_\mu \beta_{2t} \rho_{it}(x). \tag{17}
$$

On the other side, under event A_{it}, by lemma B.4, we have that for all x,

$$
|u_i(x) - \bar{u}_{it}(x)| \leqslant L_\mu \beta_{1t} \rho_{it}(x). \tag{18}
$$

Moreover, recall that by definition for any $x \in \tilde{\mathcal{S}}_{it}$,

$$
u_i(x_i^*) - u_i(x) \geqslant \beta_{3t} \rho_{it}(x). \tag{19}
$$

Therefore, consider any $x \in \tilde{\mathcal{S}}_{it}$, and under the condition that $A_{it} \cap B_{it} \cap \{u_{it}(x_i^*) > u_i(x_i^*)\}$, we have

$$
\begin{aligned}
u_{it}(x_i^*) - u_{it}(x) &> u_i(x_i^*) - u_{it}(x) \\
&= (u_i(x_i^*) - u_i(x)) + (u_i(x) - \bar{u}_{it}(x)) + (\bar{u}_{it}(x) - u_{it}(x)) \\
&> 0,
\end{aligned} \tag{20}
$$

where the last inequality follows from combining equations Eq (17), Eq (18), Eq (19) and the definition of β_{3t}. Moreover, recall that x_{it} maximizes u_{it} under the budget, thus

$$u_{it}(x_{it}) \geqslant u_{it}(x_i^*),$$

Therefore, Eq (20) implies that, $x_{it} \notin \tilde{\mathcal{S}}_{it}$. This further implies,

$$
\begin{aligned}
P_t(x_{it} \notin \tilde{\mathcal{S}}_{it}) &\geqslant P_t\left(u_{it}(x_i^*) > u_{it}(x), \forall x \in \tilde{\mathcal{S}}_{it}\right) \\
&\geqslant P_t\left(u_{it}(x_i^*) > u_{it}(x), \forall x \in \tilde{\mathcal{S}}_{it} | A_{it} \cap B_{it}\{u_i(x_{it}^*) > u_i(x)\}\right) \\
&\quad \cdot P(A_{it} \cap B_{it} \cap \{u_i(x_{it}^*) > u_i(x)\}) \\
&= P(A_{it} \cap B_{it} \cap \{u_i(x_i^*) > u_i(x)\}) \\
&\geqslant P(A_{it} \cap \{u_i(x_i^*) > u_i(x)\}) - P(B_{it}^c) \\
&= P(\{u_i(x_i^*) > u_i(x)\} | A_{it}) P(A_{it}) - P(B_{it}^c) \\
&\geqslant q_0(1 - \delta_t) + \delta_{2t}.
\end{aligned}
$$

Here, the second and third inequality both from the law of total probability and rearranging terms, and the last inequality follows from Lemma B.1, Lemma B.2 and Lemma B.3, which completes the proof. ∎

Lemma C.3. *At any round $t > t_0$, define $x_{it}'' \overset{\text{def}}{=} \arg\min_{x \notin \tilde{\mathcal{S}}_{it}, p_t^\top y \leqslant p_t^\top e_i} \rho_{it}(x_{it})$, then we have*

$$\ell^{\mathrm{PE}}(x_t) \leqslant \sum_{i \in [n]} b_{it}^{\mathrm{PE}},$$

with probability at least $1 - \delta_t - \delta_{2t}$, and $b_{it}^{\mathrm{PE}} = 2\beta_{3t}\rho_{it}(x_{it}'') + \beta_{3t}\rho_{it}(x_{it})$.

Proof. Begin with the definition of ℓ^{PE}, we have

$$
\begin{aligned}
\ell^{\mathrm{PE}}(x_t) &= \min_{x \in \mathcal{PE}} \sum_{i \in [n]} (u_i(x_i) - u_i(x_{it})) \\
&\leqslant \sum_{i \in [n]} (u_i(x_i^*) - u_i(x_{it})) \\
&= \sum_{i \in [n]} (u_i(x_i^*) - u_i(x_{it}'')) + \sum_{i \in [n]} (u_i(x_i'') - u_i(x_{it})) \\
&\leqslant \beta_{3t} \sum_{i \in [n]} \rho_{it}(x_{it}'') + \sum_{i \in [n]} (u_i(x_i'') - u_i(x_{it})),
\end{aligned}
$$

where the last inequality follows from this definition. Moreover, we have

$$u_i(x_{it}'') \leqslant u_{it}(x_{it}'') + \beta_{3t}\rho_{it}(x_{it}''),$$

and

$$u_i(x_{it}) \geqslant u_{it}(x_{it}) - \beta_{3t}\rho_{it}(x_{it}).$$

under the event $A_{it} \cap B_{it}$, by Eq (17) and Eq (18). Putting these together yields

$$\ell^{\mathrm{PE}}(x_t) \leqslant 2\beta_{3t} \sum_{i \in [n]} \rho_{it}(x_{it}'') + \beta_{3t} \sum_{i \in [n]} \rho_{it}(x_{it}) \overset{\text{def}}{=} \sum_{i \in [n]} b_{it}^{\mathrm{PE}},$$

which completes the proof. ∎

Now we show that the above result leads to the lemma below, which shows a analogous guarantee as we obtained in lemma B.6.

Lemma C.4. *For* $t \geqslant \max(t_0, t'_0)$,

$$\mathbb{E}_t[\sum_{i \in [n]} \ell_{it}^{\mathrm{FD}}] \leqslant \frac{5}{q_0} \beta_{3t} \mathbb{E}_t[\sum_{i \in [n]} \rho_{it}(x_{it})] + n(\delta_t + \delta_{2t}).$$

Here, t'_0 *is chosen such that,* $\forall t > t'_0$, $\delta_t < \frac{1}{4}$, $\delta_{2t} < \frac{q_0}{4}$.

Proof. First, note that, by the definition of x''_{it},

$$\mathbb{E}_t[\rho_t(x_{it})] \geqslant \mathbb{E}_t[\rho_t(x_{it}) | x_{it} \notin \mathcal{S}_{it}] P(x \notin \mathcal{S}_{it}) \geqslant \rho_{it}(x''_{it}) P(x \notin \mathcal{S}_{it}).$$

Moreover, combining the above with Lemma C.2, we have

$$\rho_{it}(x''_{it}) \leqslant \frac{\mathbb{E}_t[\rho_{it}(x_{it})]}{q_0(1 - \delta_t) - \delta_{2t}}.$$

Select t'_0 such that, $\forall t \geqslant t'_0$, $\delta_t = \frac{1}{4}$, and $\delta_{2t} \leqslant \frac{q_0}{4}$, then we have:

$$\rho_{it}(x''_{it}) \leqslant \frac{2}{q_0} \mathbb{E}_t[\rho_{it}(x_{it})].$$

Also, under $A_{it} \cap B_{it}$,

$$\|\theta_i^* - \theta_{it}\|_{Q_{it}} \leqslant \|\theta_i^* - \bar{\theta}_{it}\|_{Q_{it}} + \|\bar{\theta}_{it} - \theta_{it}\|_{Q_{it}} \leqslant \beta_{1t} + \beta_{2t},$$

where the first inequality follows from triangle inequality, and the second one follows from the definitions of A_{it} and B_{it}. Hence,

$$|\mu_{it}(x) - \mu_i(x)| \leqslant L_\mu(\beta_{1t} + \beta_{2t})\rho_{it}(x) = \beta_{3t}\rho_{it}(x).$$

Therefore, we have

$$\sum_{i \in [n]} b_{it}^{\mathrm{PE}} \leqslant 2\beta_{3t} \sum_{i \in [n]} \rho_{it}(x''_{it}) + \beta_{3t} \sum_{i \in [n]} \rho_{it}(x_{it})$$

$$\leqslant \frac{4}{q_0} \beta_{3t} \mathbb{E}[\sum_{i \in [n]} \rho_{it}(x_{it})] + \beta_{3t} \sum_{i \in [n]} \rho_{it}(x_{it}).$$

Moreover, by Lemma C.1 and eq (13) which holds under the same condition of $A_{it} \cap B_{it}$, we have

$$\sum_{i \in [n]} \ell_{it}^{\mathrm{FD}} \leqslant \sum_{i \in [n]} \max\{\ell_{it}^{\mathrm{PE}}, \ell_{it}^{\mathrm{SI}}\}$$

$$\leqslant 2\beta_{3t} \sum_{i \in [n]} \rho_{it}(x''_{it}) + \beta_{3t} \sum_{i \in [n]} \rho_{it}(x_{it})$$

$$\leqslant \frac{4}{q_0} \beta_{3t} \mathbb{E}[\sum_{i \in [n]} \rho_{it}(x_{it})] + \beta_{3t} \sum_{i \in [n]} \rho_{it}(x_{it}).$$

Therefore, we have

$$\mathbb{E}_t[\sum_{i \in [n]} \ell_{it}^{\mathrm{FD}}] \leqslant \mathbb{E}_t[\sum_{i \in [n]} b_{it}^{\mathrm{PE}}] \leqslant \frac{4}{q_0} \beta_{3t} \mathbb{E}[\sum_{i \in [n]} \rho_{it}(x_{it})] + \beta_{3t} \mathbb{E}_t[\sum_{i \in [n]} \rho_{it}(x_{it})] + n(\delta_t + \delta_{2t})$$

$$\leqslant \frac{5}{q_0} \beta_{3t} \mathbb{E}_t[\sum_{i \in [n]} \rho_{it}(x_{it})] + n(\delta_t + \delta_{2t}),$$

which completes the proof. ∎

With the above lemmas at hand, we are now ready to provide a proof of Theorem 4.1 for L_T^{FD}.

C.1 Proof of Theorem 4.1 for L^{FD}

Proof. Lemma C.4 shows a analog guarantee as we obtained in lemma B.6 for the L^{FD} loss function. Therefore, following the same steps in lemma B.7, we have that with probability at least $1 - \delta'$ where δ' will be specified momentarily,

$$
\begin{aligned}
L_T^{\text{FD}} &= \sum_{t=t_0}^{T} \sum_{i=1}^{n} \ell_{it}^{\text{FD}} \\
&\leqslant n \sum_{t=1}^{T} (\delta_t + \delta_{2t}) + \frac{n\beta_{3t}}{q_0} \frac{\sqrt{T}}{\sqrt{M}} \left(5c_\phi m \sqrt{\frac{1}{\log\left(1 + \frac{\|\phi(\mathbf{1})\|_2}{\sqrt{M}}\right)} \log\left(\frac{\|\phi(\mathbf{1})\|_2^2 T}{M}\right)} \right. \\
&\qquad \left. + 7c_\phi^2 \sqrt{m} \sqrt{2\log(\frac{1}{\delta'})} \right) \\
&= n \sum_{t=1}^{T} (\delta_t + \delta_{2t}) + \tilde{O}\left(n \frac{m^2}{\sqrt{M}} \sqrt{T} \left(\log(T) + \log\left(\frac{1}{\delta'}\right) \right) \right),
\end{aligned} \tag{21}
$$

Choose $\delta_t = \delta_{2t} = \frac{2\delta}{n\pi^2 t^2}$. Then,

$$
\sum_{t=1}^{T} (\delta_t + \delta_{2t}) \leqslant \frac{2}{3} \delta.
$$

Also choose $\delta' = \frac{\delta}{3}$, then by Eq (21), with probability at least $1 - \delta$,

$$
L_T^{\text{FD}} = O\left(\frac{nm^2}{\sqrt{M}} \sqrt{T} \left(\log\left(\frac{\delta}{3}\right) + \log(T) \right) \right).
$$

∎

C.2 Proof of Theorem 4.2 for L^{FD}

Proof. The theorem results follow from eq (21). Choose $\delta_{1t} = \delta_{2t} = \delta' = \frac{1}{T}$ and denote the event where $L_T^{\text{FD}} = O\left(\frac{nm^2}{\sqrt{M}} \sqrt{T} \left(\log\left(\frac{\delta}{3}\right) + \log(T) \right) \right)$ holds as $\mathcal{E}$. Then,

$$
\mathbb{E}[L_T^{\text{FD}}] = \mathbb{E}[L_T^{\text{PE}}|\mathcal{E}] + \mathbb{E}[L_T^{\text{FD}}|\mathcal{E}^c]P(\mathcal{E}^c) \leqslant O\left(\frac{nm^2}{\sqrt{M}} \sqrt{T} \left(\log\left(\frac{\delta}{3}\right) + \log(T) \right) \right),
$$

where $M \geqslant m$. This completes the proof.

∎

D Additional Experimental Details and Results

In this section, we present the simulation results with the Amdahl utilities, as described in Section 5, and additional implementation details.

Figure 2 presents the simulation results for agents with the Amdahl's utilities.

Empirically, we compute L^{CE} by maximizing each agent's utility subject to the budget constraint. We approximate this by randomly sampling feasible allocations y from a simplex, accept those that cost no more price than the agent's endowment, lastly take the maximum. We sampled up to 50 accepted samples in each round. All experiments are run on a AWS EC2 p3.2xlarge machine.
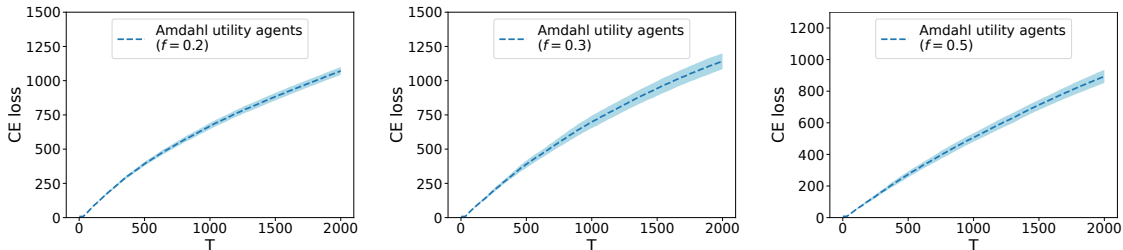
Figure 2: The CE loss L_T^{CE} vs the number of rounds T. evaluated with $m = 2$ resource types and $n = 8$ agents and Amdahl's utilities. The three figures correspond to $f_i = 0.2$, $f_i = 0.3$, and $f_i = 0.5$ (as in Section 5). All figures show results which are averaged over 10 runs, and the shaded region shows the standard error at each time T.

E   Further Discussions

E.1   Further Background on Fair Division and Exchange Economies

Since the seminal work of Varian (1973), fair division of multiple resource types has received significant attention in the game theory, economics, and computer systems literature. We provide more background on the related works in the fair division literature and their applications in this section.

Among the theoretical works in fair division, one of the most common perspectives on this problem is as an exchange economy (or as a Fisher market, which is a special case of an EE) (Varian, 1973; Mas-Colell et al., 1995; Gutman and Nisan, 2012; Crockett et al., 2008; Tiwari et al., 2009; Budish et al., 2017; Babaioff et al., 2019, 2021).

Fair allocation mechanisms have been deployed in many practical resource allocation tasks when compute resources are shared by multiple users (Zahedi et al., 2018). There have also been applications of other market-based resource allocation schemes for data centers and power grids (Lai et al., 2005; Wolski et al., 2001; Hindman et al., 2011; Burns et al., 2016; Vavilapalli et al., 2013). One line of work in this setting studied fair division when the resources in question are perfect complements; some examples include dominant resource fairness and its variants (Ghodsi et al., 2011; Parkes et al., 2015; Gutman and Nisan, 2012; Ghodsi et al., 2013; Li and Xue, 2013; Dolev et al., 2012). Although the assumption of perfect complement resource types leads to computationally simple mechanisms, in many practical applications, there is ample substitutability between resources, and hence the above mechanisms can be inappropriate. For example, in compute clusters, CPUs and GPUs are often interchangeable for many jobs, albeit with different performance characteristics. Indeed, in this work, we in particular focused on the applications of EE and fair division mechanisms for computing resource allocation.

E.2   Computation of the FD loss

We note that one main challenge for computing the FD loss is that we need to approximate the Pareto Front and then take a minimum over it. To approximate it, even in the simplest two agent two resource set up, this requires grid search on a 4D space which can be computationally prohibitive. In our experiments, the dimensions are 15 (3 resources by 5 agents) and 16 (2 resources by 8 agents), for which grid search is not feasible. Given that efficient computations over the Pareto frontier have remain a technical challenge, we focused the evaluations on the CE loss in this work.

E.3   On the loss functions

We provide an example to demonstrate that the FD loss (3), while more interpretable than the CE loss (2), may not capture all properties of an equilibrium.

For this consider the following example with $n = 3$ agents and $m = 2$ resources where the endowments of agent 1, agent 2, and agent 3 are $e_1 = (0.45, 0.05)$, $e_2 = (0.45, 0.05)$, and $e_3 = (0.1, 0.9)$ respectively. Their utilities are:

$$u_1(x_1) = 0.1x_{11} + x_{12}, \qquad u_2(x_2) = 0.2x_{21} + x_{22}, \qquad u_3(x_3) = x_{31} + 0.1x_{32}.$$

The utilities of the three users if they were to simply use their endowment is, $u_1(e_1) = 0.1 \times 0.45 + 0.05 = 0.095$, $u_2(e_2) = 0.14$, and $u_3(e_3) = 0.19$. We find that while agents 1 and 2 benefit more from the second resource, they have more of the first resource in their endowments and vice versa for agent 3. By exchanging resources, we can obtain a more efficient allocation.

The unique equilibrium prices for the two goods are $p^\star = (1/2, 1/2)$ and the allocations are $x_1^\star = (0, 0.5)$ for agent 1, $x_2^\star = (0, 0.5)$ for agent 2, and $x_3^\star = (1.0, 0.0)$ for agent 3. The utilities of the agents under the equilibrium allocations are $u_1(x_1) = 0.5$, $u_2(x_2) = 0.5$, and $u_3(x_3) = 1.0$. Here, we find that by the definition of CE, $\ell^{\mathrm{PE}}(x^\star, p^\star) = 0$. It can also be verified that $\ell^{\mathrm{FD}}(x^\star, p^\star) = 0$.

In contrast, consider the following allocation for the 3 users: $x_1 = (0.35, 0.49)$ for agent 1, $x_2 = (0.35, 0.49)$ for agent 2, and $x_3 = (0.3, 0.02)$ for agent 3. Here, the utilities are $u_1(x_1) = 0.1 \times 0.35 + 0.49 = 0.525$, $u_2(x_2) = 0.56$, and $u_3(x_3) = 0.3002$. This allocation is both PE (as the utility of one user can only be increased by taking resources from someone else), and SI (as all three users are better off than having their endowments). Therefore, $\ell^{\mathrm{FD}}((x_1, x_2, x_3)) = 0$. However, user 3 might complain that their contribution of resource 2 (which was useful for users 1 and 2) has not been properly accounted for in the allocation. Specifically, there do not exist a set of prices p for which $\ell^{\mathrm{PE}}(x, p) = 0$. This example illustrates the role of prices in this economy: it allows us to value the resources relative to each other based on the demand.