Private Pareto Optimal Exchange*

Sampath Kannan[1], Jamie Morgenstern[2], Ryan Rogers[1], and Aaron Roth[1]

[1] *Computer Science Department, The University of Pennsylvania*
[2] *Computer Science Department, Carnegie Mellon University*

July 9, 2014

Abstract

We consider the problem of implementing an individually rational, asymptotically Pareto optimal allocation in a barter-exchange economy where agents are endowed with goods and have preferences over the goods of others, but may not use money as a medium of exchange. Because one of the most important instantiations of such economies is kidney exchange – where the "input" to the problem consists of sensitive patient medical records – we ask to what extent such exchanges can be carried out while providing formal privacy guarantees to the participants. We show that individually rational allocations cannot achieve any non-trivial approximation to Pareto optimality if carried out under the constraint of differential privacy – or even the relaxation of *joint* differential privacy, under which it is known that asymptotically optimal allocations can be computed in two-sided markets, where there is a distinction between buyers and sellers and we are concerned only with privacy of the buyers [Hsu et al., 2014]. We therefore consider a further relaxation that we call *marginal* differential privacy – which promises, informally, that the privacy of every agent i is protected from every other agent $j \neq i$ so long as j does not collude or share allocation information with other agents. We show that, under marginal differential privacy, it is possible to compute an individually rational and asymptotically Pareto optimal allocation in such exchange economies.

*Kannan was partially supported by NSF grant NRI-1317788. email: kannan@cis.upenn.edu. Morgenstern was partially supported by NSF grants CCF-1116892 and CCF-1101215, as well as a Simons Award for Graduate Students in Theoretical Computer Science. Contact information: J. Morgenstern, Computer Science Department, Carnegie Mellon University, jamiemmt@cs.cmu.edu. Roth was partially supported by an NSF CAREER award, NSF Grants CCF-1101389 and CNS-1065060, and a Google Focused Research Award. Email: aaroth@cis.upenn.edu.

1 Introduction

Consider the following exchange problem: for each $i \in [n]$, agent i arrives at a market endowed with a good of type g_i from among a finite collection of types of goods $\mathcal{G}$, as well as a *preference over types of goods*, represented by a total ordering $\succ_i$ over $\mathcal{G}$. In settings where money can be used as the medium of transaction (and for which people have cardinal preferences over goods), this would represent an exchange economy for which we could compute market clearing prices, resulting in a Pareto optimal allocation. However, in some settings – most notably markets for kidney exchanges[1] Roth et al. [2005] – the use of money is not permitted, and agents are limited to participating in an exchange – essentially a permutation of the goods amongst the players, with no additional payments. In such settings, we may still require that:

1. The exchange be *individually rational* – i.e. every agent weakly prefers the good that they receive to the good that they were endowed with, and
2. the exchange be (approximately) *Pareto optimal* – that there should not be any other permutation of the goods that some agent (or, in the approximate case, many agents) strictly prefers, unless there exists another agent who strictly prefers the original allocation.

Because one of the key applications of efficient barter-exchange involves computation on extremely sensitive medical data, this paper investigates to what extent it can be accomplished while guaranteeing a formal notion of *privacy* to the agents participating in the market, without compromising the above two desiderata of individual rationality and (approximate) Pareto optimality. We wish to give algorithms that protect the privacy of agents' initial endowment as well as their preferences over goods from the other agents in the market, using the tools of differential privacy.

In this respect, our paper continues a recent line of work exploring the power of algorithms satisfying various privacy definitions (relaxations of *differential privacy*) in different kinds of exchange problems. To understand this question, note that the *input* to an algorithm clearing an exchange economy is partitioned amongst the n agents (each agent reports their initial endowment g_i and their preference ordering $\succ_i$), as is the output (the mechanism reports to each agent the type of good they will receive, g'_i.) This allows us to parameterize privacy guarantees in terms of adversaries who can see differing sets of outputs of the mechanism. The standard notion of *differential privacy*, requires that we guarantee privacy even against an adversary who can see all n outputs of the mechanism – i.e. the type of good that is allocated to each of the n agents. It is intuitively clear that nothing non-trivial can be done while guaranteeing differential privacy in allocation problems – "good" allocations must give individuals what they want, and this is exactly what we must keep private. An adversary who is able to see the allocation given to player i by any mechanism which guarantees individual rationality would immediately learn the relative ranking of the good that agent i was allocated compared to the good that he was endowed with. However, this does not rule out the possibility of protecting the privacy of agent i against an adversary who can only see the allocation of *some* agents– notably, not the allocation given to agent i.

1.1 Our Results

The question of privately computing allocations where agents' preferences are sensitive data was first studied by Hsu et al. [2014], who showed that in two sided allocation problems (with distinguished *buyers* and *sellers*) with monetary transfers, no non-trivial allocation can be computed under the constraint of differential privacy, when we must protect the privacy of the buyers. However, Hsu et al. [2014] showed that near-optimal results can be achieved under *joint differential privacy*, which informally requires that for every player i simultaneously, the joint distribution on allocations given to players $j \neq i$ be differentially private in the data of agent i. This corresponds to privacy against an adversary who can observe the allocation of *all other*

[1] Kidney exchange forms one of the most notable "barter" markets, but it is not the only example. A number of startups such as "TradeYa" and "BarterQuest" act as market makers for barter exchange of consumer goods.

players $j \neq i$, but who cannot observe agent i's own allocation when trying to violate agent i's privacy.

The allocation problem we study in this paper is distinct from the two sided problem studied in Hsu et al. [2014] in that there are no distinguished buyers and sellers – in our barter exchange problem, every agent both provides a good and receives one, and so we must protect the privacy of every agent in the market. We insist on algorithms that always guarantee *individually rational* allocations, and ask how well they can approximate *Pareto optimality*. (Informally, an allocation π is α-approximately Pareto optimal if for every other allocation π' that is strictly preferred by an α-fraction of agents, there must be some agent who strictly prefers π to π'.) We start by showing that even under the relaxed notion of *joint* differential privacy, no individually rational mechanism can achieve any nontrivial approximation to Pareto optimality (and, since joint differential privacy is a relaxation of differential privacy, neither can any differentially private mechanism).

Theorem (Informal) *No ϵ-jointly differentially private algorithm for the exchange problem that guarantees individually rational allocations with high probability can guarantee that the resulting allocation will be α-approximately Pareto optimal for:*

$$\alpha \leq 1 - \frac{e^\epsilon}{e^\epsilon + 1}$$

Given this impossibility result, we consider a further relaxation of differential privacy, which we call *marginal differential privacy*. In contrast to joint differential privacy, marginal differential privacy requires that, simultaneously for every pair of agents $i \neq j$, the *marginal* distribution on player j's allocation be differentially private in player i's data. This corresponds to privacy from an adversary who has the ability only to look at a single other agent's allocation (equivalently – privacy from the other agents, assuming they do not collude). Our main result is a marginally-differentially private algorithm that simultaneously guarantees individually rational and approximately Pareto optimal allocations, showing a separation between marginal and joint differential privacy for the exchange problem:

Theorem (Informal) *There exists an ϵ-marginally differentially private algorithm that solves the exchange problem with n agents and $k = |\mathcal{G}|$ types of goods by producing an allocation which is individually rational and, with high probability, α-approximately Pareto optimal for*

$$\alpha = O\left(\frac{\text{poly}(k)}{\epsilon n}\right)$$

Note that the approximation to Pareto optimality depends polynomially on the number of *types* of goods in the market, but decreases linearly in the number of participants in the market n. Hence, fixing k, and letting the market size n grow, this mechanism is asymptotically Pareto optimal.

It is natural to ask whether this bound can be improved so that the dependence on k is only $\alpha = O\left(\frac{\text{polylog}(k)}{\epsilon n}\right)$, which is the dependence on the number of distinct types of goods achieved in the approximation to optimality in Hsu et al. [2014] (again, under *joint* differential privacy, for a two-sided market). We show that this is not the case.

Theorem (Informal) *For every ϵ-marginally differentially private algorithm that on every instance of the exchange problem with n agents and $k = |\mathcal{G}|$ types of goods, produces an individually rational allocation that with high probability is α-approximately Pareto optimal, we have:*

$$\alpha = \Omega\left(\frac{k}{n}\left(1 - \frac{e^\epsilon}{e^\epsilon + 1}\right)\right)$$

Our work continues the study of which kinds of problems can be solved under parameterized relaxations of differential privacy. Taken together with the results of Hsu et al. [2014], our results

suggest an intriguing direction for future work – characterizing the power of relaxations which lie between marginal and joint differential privacy, in which privacy is protected from subsets of c colluding agents $j \neq i$, for $1 < c < n - 1$.

1.2 Related Work

Differential privacy, introduced by Dwork et al. [2006] has become a standard "privacy solution concept" over the last decade, and has spawned a vast literature too large to summarize. We here discuss only the most closely related work.

Although the majority of the differential privacy literature has considered numeric valued and continuous optimization problems, a small early line of work including Nissim et al. [2007] and Gupta et al. [2010] study combinatorial optimization problems. The dearth of work in this area in large part stems from the fact that many optimization problems cannot be nontrivially solved under the constraint of differential privacy, which requires that the entire output be insensitive to any input. This problem was first observed by Gupta et al. [2010] in the context of set cover and vertex cover, who also noted that if the notion of a solution is slightly relaxed to include private "instructions" which can be given to the agents, allowing them (together with their own private data) to reconstruct a solution, then more is possible. Similar ideas are also present in McSherry and Mironov [2009], in the context of recommendation systems.

Joint differential privacy, which can be viewed as a generalization of the "instructions" based solution of Gupta et al. [2010], was formalized by Kearns et al. [2014], who showed that, although correlated equilibria in large games could not be computed to any nontrivial accuracy under differential privacy, they can be computed quite accurately under joint differential privacy. A similar result was shown by Rogers and Roth [2014] for *Nash* equilibria in congestion games. Hsu et al. [2014] subsequently studied a two-sided allocation problem, in which buyers with private valuation functions over bundles of goods must be allocated goods to maximize social welfare (note that here buyers are allocated goods, but do not provide them, unlike the problem we study in this work). They also show that, although the allocation problem cannot be solved to nontrivial accuracy under differential privacy, it can be solved accurately (when buyers preferences satisfy the *gross substitutes* condition) under joint differential privacy.

In the present paper, we continue the study of private allocation problems, and consider the exchange problem in which n agents both supply and receive the goods to be allocated. This problem was first studied by Shapley and Scarf [1974], who also proposed the Top Trading Cycles algorithm for solving it (attributing this algorithm to David Gale). We show that this problem is strictly harder from a privacy perspective than the two sided allocation problem: it cannot be solved non-trivially even under joint differential privacy, but can be solved under a weaker notion which we introduce, that of marginal differential privacy. Our solution involves a privacy-preserving modification of the top trading cycles algorithm. To the best of our knowledge, we are the first to give *marginal* differential privacy a name and to demonstrate a separation from joint differential privacy. The solution concept has, however appeared in other works – for example in He and Mu [2014], in the context of privacy preserving and incentive compatible recommendation systems.

2 Model

We study the setting of trading within an exchange market. There are k kinds of goods in the market. We denote this set of types of goods as $\mathcal{G}$. The set of agents will be denoted as N where $|N| = n$. Each $i \in N$ has one copy of some type of good $g_i \in \mathcal{G}$, and some linear preference over all goods in $\mathcal{G}$, $\succ_i$. Let $N_j = \{i : g_i = j\}$ and $n_j = |N_j|$ denote the number of people who bring good $j \in \mathcal{G}$ to the market. Since each agent brings exactly one good to the market, $\sum_{j \in \mathcal{G}} n_j = n$. An instance of an exchange market is given as $\mathbf{x} = (x_i)_{i \in N}$ where each $x_i = (g_i, \succ_i)$. Our goal will be to find beneficial trades amongst the agents in the market.

Definition 1 (Allocation)**.** *An allocation is a mapping $\pi : N \to \mathcal{G}$ where we have $|\{i \in N : \pi(i) = j\}| = n_j$ for each $j \in \mathcal{G}$.*

In particular, one feasible allocation is the one which matches each individual with the good they started with $\pi(i) = g_i$ (where no agents trade goods). However, we prefer allocations that

result in agents having goods that they prefer to their initial endowments. We only consider algorithms which produce *individually rational* allocations:

Definition 2 (IR)**.** *An allocation π is Individually Rational (IR) if $\pi(i) \succeq_i g_i \ \forall \ i \in N$.*

The previously mentioned identity allocation (with no trade) is trivially IR. Subject to the IR constraint, we wish to find a Pareto optimal allocation. An allocation is Pareto optimal if it cannot be changed to improve some agent i's utility without harming the utility of some other agent j. Under privacy constraints, it will be impossible to obtain exact Pareto optimality, and so we instead ask for an approximate version.

Definition 3 (α-Approximate Pareto Optimality)**.** *An allocation π is α-approximately Pareto optimal (or just α-Pareto optimal) if for any other allocation π', if there exists a set $S \subset N$ with $|S| > \alpha n$ such that $\pi'(i) \succ_i \pi(i), \forall i \in S$, then there must be some $j \in N \backslash S$ such that $\pi(j) \succ_j \pi'(j)$. In other words, an allocation is α-approximately Pareto optimal if strictly improving the allocation for more than an α-fraction of agents necessarily requires strictly harming the allocation of at least 1 agent.*

We say that an algorithm guarantees α-Pareto optimality if, on every exchange problem instance, it outputs an α-Pareto optimal allocation. If $\alpha = \alpha(n)$ is a function of the number of agents n, we say that an algorithm is asymptotically Pareto optimal if it guarantees $\alpha(n)$-Pareto optimality, and $\alpha(n) = o(1)$.

We wish to compute such allocations while guaranteeing a formal notion of privacy to the participating agents. The notions of privacy we consider will all be relaxations of *differential privacy*, which has become a standard privacy "solution concept". We borrow standard notation from game theory: given a vector $\mathbf{t} \in T^n$, we write $\mathbf{t}_{-i} \in T^{n-1}$ to denote the vector $\mathbf{t}$ with the i'th coordinate removed, and given $t'_i \in T$ we write $(t'_i, \mathbf{t}_{-i}) \in T^n$ to denote the vector $\mathbf{t}$ with its i'th coordinate replaced by t'_i.

Definition 4 (Differential Privacy, [Dwork et al., 2006])**.** *A mechanism $M : T^n \to R$ satisfies (ϵ, δ)- differential privacy if for every $i \in [n]$, for any two types $t_i, t'_i \in T$, any tuple of types $\mathbf{t}_{-i} \in T^{n-1}$, and any $B \subseteq R$, we have*

$$\mathbb{P}\left(M(t_i, \mathbf{t}_{-i}) \in B\right) \leq e^{\epsilon}\mathbb{P}\left(M(t'_i, \mathbf{t}_{-i}) \in B\right) + \delta$$

Here R denotes an arbitrary range of the mechanism. Note that the definition of differential privacy assumes that the *input* to the mechanism $\mathbf{t}$ is explicitly partitioned amongst n agents. In the problem we consider, the input $\mathbf{x}$ is partitioned and the range of the mechanism is also naturally partitioned between n agents (the output of the mechanism can be viewed as n messages, one to each agent i, telling them the type of good they are receiving, $\pi(i)$). In such cases, we can consider relaxations of differential privacy informally parameterized by the maximum size of collusion that we are concerned about. Joint differential privacy, defined by Kearns et al. [2014], asks that the mechanism simultaneously protect the privacy of every agent i from arbitrary collusions of up to $n-1$ agents $j \neq i$ (who can share their own allocations, but cannot see the allocation of agent i):

Definition 5 (Joint Differential Privacy, [Kearns et al., 2014])**.** *A mechanism $M : T^n \to O^n$ satisfies (ϵ, δ)-joint differential privacy if for any player $i \in [n]$, any two types $t_i, t'_i \in T$, any tuple of types $\mathbf{t}_{-i} \in T^{n-1}$, and any $B_{-i} \subseteq O^{n-1}$, we have*

$$\mathbb{P}\left(M(t_i, \mathbf{t}_{-i})_{-i} \in B_{-i}\right) \leq e^{\epsilon}\mathbb{P}\left(M(t'_i, \mathbf{t}_{-i})_{-i} \in B_{-i}\right) + \delta$$

As we will show, it is not possible to find individually rational and asymptotically Pareto optimal allocations under joint differential privacy, and so we introduce a further relaxation which we call marginal differential privacy. Informally, marginal differential privacy requires that the mechanism simultaneously protect the privacy of every agent i from every other agent $j \neq i$, assuming that they do not collude (i.e. it requires the differential privacy condition only on the *marginal* distribution of allocations to other agents, not on the joint distribution).

Definition 6 (Marginal Differential Privacy)**.** *A mechanism $M : T^n \to O^n$ satisfies (ϵ, δ)-marginal differential privacy if $\forall i \neq j$, $\forall \mathbf{t}_{-i} \in T^{n-1}$, $\forall t_i, t_i' \in T$, and $\forall S \subset O$, we have*

$$\mathbb{P}(M(t_i, \mathbf{t}_{-i})_j \in S) \leq e^{\epsilon}\mathbb{P}(M(t_i', \mathbf{t}_{-i})_j \in S) + \delta$$

3 Lower Bounds

In this section, we show lower bounds on how well the barter-exchange problem can be solved subject to privacy constraints. We first show that under the constraint of joint-differential privacy, there does not exist any individually rational and asymptotically Pareto optimal mechanism. This motivates our relaxation of studying exchange problems subject to marginal differential privacy. We then show that under marginal differential privacy, any mechanism producing individually rational and α-approximately Pareto optimal allocations must have $\alpha = \Omega(k/n)$, where k is the number of distinct types of goods. (i.e. a linear dependence on k is necessary). We complement these impossibility results in Section 4 where we show that under marginal differential privacy, it is indeed possible to achieve both individual rationality and asymptotic Pareto optimality simultaneously, if $k = o(n^{1/6})$ (the number of goods grows slower than $n^{1/6}$, where n is the number of agents).

Our impossibility result for joint differential privacy is based on a reduction to the following well-known claim.

Claim 1. *There is no (ϵ, δ)- differentially private mechanism $M : \{0,1\} \to \{0,1\}$ such that $\mathbb{P}(M(b) = b) > \frac{e^{\epsilon}+\delta}{e^{\epsilon}+1}$ for both $b = 0, 1$.*

Proof. Let M be (ϵ, δ)-differentially private such that $\mathbb{P}(M(0) = 0) > \frac{e^{\epsilon}+\delta}{e^{\epsilon}+1}$ and $\mathbb{P}(M(1) = 1) > \frac{e^{\epsilon}+\delta}{e^{\epsilon}+1}$. By the definition of ϵ-differential privacy, we have

$$\begin{aligned}\frac{e^{\epsilon}+\delta}{e^{\epsilon}+1} &< \mathbb{P}(M(0) = 0) \leq e^{\epsilon}\mathbb{P}(M(1) = 0) + \delta \\ &= e^{\epsilon}(1 - \mathbb{P}(M(1) = 1)) + \delta < e^{\epsilon}\left(1 - \frac{e^{\epsilon}+\delta}{e^{\epsilon}+1}\right) + \delta,\end{aligned}$$

giving a contradiction. □

In the setting of exchange markets, we will let $\mathcal{X} = \mathcal{G} \times \mathcal{T}$ where $\mathcal{G}$ is the set of types of goods and $\mathcal{T}$ is the set of linear orderings over $\mathcal{G}$ for a single player. Now, we show that it will not be possible to guarantee privacy, IR, and $\alpha = o(1)$-Pareto optimality (considering ϵ as a constant).

Theorem 1. *If we have an (ϵ, δ)-joint differentially private mechanism $M_J : \mathcal{X}^n \to \mathcal{G}^n$ that guarantees an α-Pareto optimal allocation with probability at least $1 - \beta$ and always gives an IR allocation then*

$$\alpha \geq 1 - \frac{e^{\epsilon}+\delta}{(1-\beta)(e^{\epsilon}+1)}$$

Proof. The proof proceeds by reduction to Claim 1. Suppose we had such an (ϵ, δ)-jointly differentially private mechanism M_J, with $\alpha < 1 - \frac{e^{\epsilon}+\delta}{(1-\beta)(e^{\epsilon}+1)}$. We show that we could use it to construct an ϵ-differentially private mechanism $M : \{0,1\} \to \{0,1\}$ that contradicts Claim 1. We design an exchange market parameterized by the input bit b received by mechanism M (see Figure 1). The market has two types of goods, g_0 and g_1 and $2n$ agents partitioned into two sets, N_0 and N_1 of size n each. The agents $j \in N_1$ are endowed with good g_1 and strictly prefer good g_0 (i.e. all such agents j have preference $g_0 \succ_j g_1$). The agents in N_0 are endowed with good g_0. We assume $n-1$ of them strictly prefer good g_1 (i.e. all such agents j have preference $g_1 \succ_j g_0$). A distinguished agent, $i \in N_0$, selected among the n agents in N_0 uniformly at random, has preference determined by bit b: $g_b \succ_i g_{1-b}$. (i.e. the i'th agent wishes to trade if $b = 1$, but prefers keeping her own good if $b = 0$.) We denote the vector of linear preferences that depends

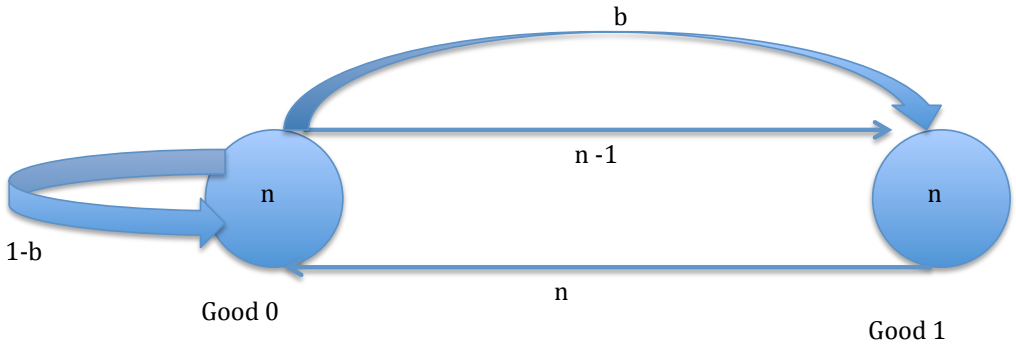

Figure 1: Depicting the exchange market considered in the proof of Claim 2.

on whether i's bit b as $\succ(b)$. We will refer to this exchange market as $\mathbf{x}(b) = (\mathbf{g}, \succ(b)) \in \mathcal{X}^{2n}$ where $\mathbf{g} \in \{0,1\}^{2n}$. We remark that when $b = 1$, the agents in N_0 are identical.

Let $M_J : \mathcal{X}^{2n} \to \mathcal{G}^{2n}$ be the (ϵ, δ)-joint differentially private mechanism given in the statement of the theorem. Note first that by the definition of joint differential privacy, the mechanism $M' : \{0,1\} \to \mathcal{G}^{n-1}$ defined as $M'(b) = M_J(\mathbf{x}(b))_{-i}$ (which takes as input b and outputs the allocation of all $n-1$ players $j \neq i$) is (ϵ, δ)-differentially private.

Note that in this construction, when $b = 0$, the IR constraint requires that $M_J(\mathbf{x}(0))_i = 0$ with probability 1. Note also that from the output of $M'(b)$, we can determine whether agent i engaged in trade: if not, we must have:

$$|\{j \in N_0 \backslash i : M'(b)_j = g_1\}| = |j \in N_1 : M'(b)_j = g_0|$$

but if so, we must have:

$$|\{j \in N_0 \backslash i : M'(b)_j = g_1\}| = |j \in N_1 : M'(b)_j = g_0| - 1$$

Define $f : \mathcal{G}^{2n-1} \to \{0,1\}$ to be the indicator function of the event $|\{j \in N_0 \backslash i : M'(b)_j = g_1\}| \neq |j \in N_1 : M'(b)_j = g_0|$. Define $M : \{0,1\} \to \{0,1\}$ to be $M(b) = f(M'(b))$. Note that M is (ϵ, δ)-differentially private by the post-processing guarantee of differential privacy, and is an indicator variable determining whether agent i has traded in our exchange economy.

If $M(b) = 1$, therefore, it must be that $b = 1$ by the individual rationality guarantee of M_J. (i.e. $\mathbb{P}(M(0) = 1) = 0 \implies \mathbb{P}(M(0) = 0) = 1$). Hence, by Claim 1 we must have $\mathbb{P}(M(1) = 1) \leq \frac{e^\epsilon + \delta}{e^\epsilon + 1}$. We know by hypothesis that M_J finds an α-approximate Pareto optimal allocation with probability $1 - \beta$. If $b = 1$ then every agent wishes to trade in $\mathbf{x}(b)$ and hence with probability $1 - \beta$, M_J must produce an allocation in which at least $2n(1-\alpha)$ people trade. Since all players in N_1 are identical, and i was selected uniformly at random, it must therefore be that agent i engages in trade with probability at least $(1-\beta)(1-\alpha)$. Thus, we have

$$\frac{e^\epsilon + \delta}{e^\epsilon + 1} \geq \mathbb{P}(M(1) = 1) \geq (1-\beta)(1-\alpha)$$

which gives us the conclusion of the theorem.

□

We also show that, even using just marginal differential privacy, the number of goods must be much smaller than the number of players to get asymptotically exact Pareto optimality, as we show in the theorem below.

Theorem 2. *If we have an (ϵ, δ)-marginal differentially private mechanism $M_S : \mathcal{X}^n \to \mathcal{G}^n$ that guarantees an α-approximate Pareto optimal allocation with probability at least $1 - \beta$ and*

always gives an IR allocation then

$$\alpha \geq \frac{k(1-\beta)}{n}\left(1 - \frac{e^\epsilon + \delta}{1+e^\epsilon}\right)$$

Proof. Suppose we have an (ϵ, δ)-marginally differentially private mechanism M_S which is IR and α-Pareto optimal. We will use it to construct some (ϵ, δ)-differentially private mechanism M, which will give a lower bound on α by Claim 1.

Suppose there are k types of goods in the market, $1, \ldots, k$. Let g_i represent the type of good with which player i is initially endowed. Table 3 shows the favorite and second favorite goods of all n players[2]. We will refer to this exchange market as $\mathbf{x}(b) = (\mathbf{g}, \succ(b)) \in \mathcal{X}^n$ where $\mathbf{g} \in \{1, \cdots, k\}^n$. A single player k has preferences which are determined by bit b: if $b = 0$, her favorite good is $g_k = k$ her own, and if $b = 1$, her favorite good is good 1. In the case that $b = 1$, players $1, \ldots, k$ form a cycle with their favorite preferences: player $i = 1, \cdots, k-1$ gets good $i+1$ and player k gets good 1, which would give each player her favorite good. If, on the other hand, $b = 0$, the uniquely IR trade is the $\pi(i) = g_i$, or no trade, since that instance contains no cycles, other thank self loops.

Player	Endowment	Favorite Good	Second Favorite Good
1	1	2	1
2	2	3	2
...	...	...	...
$k-1$	$k-1$	k	$k-1$
k (if $b=0$)	k	k	Does not matter
k (if $b=1$)	k	1	k
$k+1 \ldots n$	k	k	Does not matter

Figure 2: The endowments and preferences for Theorem 2

Consider the mechanism $M_i(b) = M_S(\mathbf{x}(b))_i$ for $i \neq k$ (which is (ϵ, δ)-differentially private, since M_S is marginally differentially private). Let $f_1 : \mathcal{G} \to \{0, 1\}$ be the indicator function for the event that 1 receives good 2; e.g. $f_1(1) = 0$ and $f_1(2) = 1$. Then, define $M'(b) = f_1(M_1(b))$, which is also (ϵ, δ)-differentially private, due to f being a post processing function. By individual rationality, $M'(0) = 0$ with probability 1. Thus, by Claim 1,

$$\mathbb{P}[M'(1) = 1] \leq \frac{e^\epsilon + \delta}{e^\epsilon + 1}$$

Thus,

$$\mathbb{P}[M'(1) = 0] \geq 1 - \frac{e^\epsilon + \delta}{e^\epsilon + 1} \tag{1}$$

When $M'(1) = 0$, $M_i(1) = g_i$ (each player was allocated her initial endowment). Consider the allocation $\pi(i) = i + 1 (\mod k)$ for $i \in [k]$ and $\pi(i) = g_i = k$ for $i = k+1, \cdots, n$. Players $1, \ldots, k$ prefer π to $M_S(\mathbf{x}(1))$ when $M'(1) = 0$, and all other players are indifferent. Since M_S is α-Pareto optimal, but there exists some π which k agents prefer and no agent likes less,

$$\alpha \geq (1-\beta)\frac{k}{n}\mathbb{P}[M'(1) = 0] \geq (1-\beta)\frac{k}{n}\left(1 - \frac{e^\epsilon + \delta}{e^\epsilon + 1}\right)$$

by Equation 1, which completes the proof. □

[2]Note that, since we are only considering IR mechanisms, preferences need only be specified to the level where i ranks g_i, and each player i in our example has good g_i as her first or second choice.

4 Private Top Trading Cycles

The algorithm PTTC (given in Algorithm 2 in Appendix **??**) takes an exchange market $\mathbf{x} = (g_i, \succ_i)_{i=1}^n$ as input and allocates each person a good. PTTC begins by considering the complete directed graph $G = (V, A)$ with each node $u \in V$ corresponding to a type of good in $\mathcal{G}$. For each arc $e = (u, v) \in A$ we define the set $P_{(u,v)}$ to be the set of people who have good u and consider v their favorite remaining good (and thus want to trade along arc $e = (u, v)$): $P_{(u,v)} = \{i \in N : g_i = u \quad \& \quad v \succeq_i g \quad \forall g \in V\}$.

We then define the arc weight w_e for arc $e \in A$ as $w_e = |P_e|$. Let N_u be the set of people that are endowed with good u, so $N = \cup_{u \in V} N_u$. The node weights n_u for $u \in V$ are then $n_u = |N_u| = \sum_{e:e=(u,v)} w_e$. PTTC works as follows (the formal algorithm is in Appendix **??**):

1. (Arc Weight Noise) Add Laplacian noise Z_e with parameter $1/\epsilon'$ to each arc weight where $\epsilon' = \frac{\epsilon}{4k^3}$. We let E denote a high probability upper bound on all realizations of the noise Z_e, which is $E = \frac{\log(2k^3/\beta)}{\epsilon'} = \frac{4k^3 \log(2k^3/\beta)}{\epsilon}$, see Claim 2.
   If our noise is ever bigger than E then we output that our algorithm failed and make no trades. We want to ensure that our noisy estimates undercount the number of people on each edge, while still having a lower bound on the difference between the actual edge weight and our noisy estimate. To do this we subtract $2E$ from our estimate after we add noise, e.g. $\hat{w}_e = w_e + Z_e - 2E$.
2. (Clear Cycle) If there is a cycle in the graph where all the arcs along it have weight $\lfloor \hat{w}_e \rfloor > 0$ then call it C. Let $\hat{W} = \min_{e \in C}\{\lfloor \hat{w}_e \rfloor\}$ be defined to be the weight of C. Otherwise there is no cycle.
3. (Trade) If there is a cycle C with weight $\hat{W}$, then for each $e \in C$ select $\hat{W}$ people uniformly at random from P_e to trade along e and denote this set of people S_e (S for satisfied). We then update, for every $e = (u, v) \in C$, that
$$P_e \leftarrow P_e \backslash S_e \quad w_e \leftarrow w_e - \hat{W} \quad \hat{w}_e \leftarrow \hat{w}_e - \hat{W}$$
   e.g., remove the players served from the arc and counters), and for each $i \in S_e$, we set $\pi(i) = v$ where $g_i = u$ and $e = (u, v)$ (they get the good they requested). Return to Step 2.
4. (No Cycle - Deletion) If there is no cycle with positive weight, then there must be a node v such that $\hat{n}_v = \sum_{e:e=(v,w)} \hat{w}_e < k$. Note that this means that there will be fewer than D copies n_v of good type v (see Lemma 3) if we set $D = (3E + 1)k$.
   We then update $V \leftarrow V \backslash \{v\}$. The arcs that were ingoing $IN_v = \{e \in A : e = (u, v) \text{ some } u \in V\}$ and outgoing from v, $OUT_v = \{e \in A : e = (v, u) \text{ some } u \in V\}$ are also deleted: $A \leftarrow A \backslash \{OUT_v \cup IN_v\}$.
   We then move all agents $i \in P_e$ for each $e \in IN_v$ to the arc that corresponds to trade between g_i and their next most preferred good among the remaining vertices. For every $e = (u, w) \in A$ we define these additional agents as
$$B_{(u,w)} = \{i \in P_{(u,v)} : w \succeq_i g \quad \forall g \in V\}$$
   (where $v \notin V$) and update the quantities $w_e \leftarrow w_e + |B_e|$ and $P_e \leftarrow P_e \cup B_e$.
   We then assign those that were remaining at the deleted node v their original good type: $\pi(i) = g_i = v$ for all i such that $i \in P_e$ where $e \in OUT_v$. Return to Step 1.

We state the formal version of our main theorem below.

Theorem 3. *PTTC is (ϵ, δ)-marginally differentially private, IR with certainty, and with probability $1 - \beta$ outputs an allocation which is α-approximately Pareto Optimal for*
$$\alpha = O\left(\frac{k^6}{\epsilon n} \log(k/\beta)\right).$$

We make several claims about our algorithm PTTC, from which Theorem 3 follows.

Lemma 1. *The allocation that PTTC produces is IR.*

Proof. Note that even if the algorithm returns that it failed (when there is a noise term $|Z_e^t| > E$) then everyone gets their original good type back, which is IR trivially. We will then assume that the algorithm runs to completion but $\exists i$, such that $g_i \succ_i \pi(i) = v$ for $v \neq g_i$. This means that at a round t, at the τ'th cycle C_t^τ, there is an arc $e = (g_i, v) \in C_t^\tau$ where $v \neq g_i$. For i to be part of the satisfied set $S_e^{t,\tau} \subseteq P_e^{t,\tau}$ at this round, we then know that for $e = (g_i, v)$ then $v \succ_i g_i$ by construction. This gives our contradiction. □

Claim 2. *(Low Error) With probability $1 - \beta$, all noise terms we add in* `PTTC` *satisfies the following*

$$\max_{e \in A, t \in [k]} |Z_e^t| \leq E = \frac{8k^3 \log(2k^3/\beta)}{\epsilon}.$$

Hence, with probability $1 - \beta$,

$$E \leq w_e - \hat{w}_e \leq 3E. \tag{2}$$

Proof. Recall that for a random variable $Z \sim \text{Lap}(b)$ we have

$$\mathbb{P}(|Z| \geq tb) \leq e^{-t}$$

We then have, for $t = \log(k^3/\beta)$ and $b = \frac{1}{\epsilon'}$

$$\mathbb{P}\left(|Z| \geq \frac{\log(k^3/\beta)}{\epsilon'}\right) \leq \frac{\beta}{k^3}.$$

Now our algorithm will sample a new Laplace random variable Z_e^t at most k^3 times - for each of the $O(k^2)$ edges, one Laplace random variable is sampled in each round and there are at most k rounds. Hence, we can obtain the following bound,

$$\mathbb{P}\left(|Z_e^t| \leq \frac{\log(k^3/\beta)}{\epsilon'} \quad \forall e \in A, \forall t \in [k]\right) \geq \left(1 - \frac{\beta}{k^3}\right)^{k^3} \geq 1 - \beta.$$

Thus, if we set E as in 2,

$$E = \frac{\log(k^3/\beta)}{\epsilon'} = \frac{4k^3 \log(k^3/\beta)}{\epsilon}$$

with probability $1 - \beta$, we have a bound on the largest the noise can be. Hence we can then lower bound and upper bound the difference between the error in the noisy arc weights and the actual arc weights to get the relation in (2):

$$E \leq w_e - \hat{w}_e = 2E - Z_e \leq 3E.$$

□

Lemma 2. *Assuming the condition in (2), for a fixed round t, there can be at most k^2 cycles selected, i.e. $\tau \leq k^2$ in the implementation of* `PTTC`*.*

Proof. At round t, `PTTC` finds the τ'th cycle with noisy weight ≥ 1. Then there is some arc e along the cycle that gets depleted, i.e. $\hat{w}_e(t, \tau) \leftarrow \hat{w}_e(t, \tau - 1) - \lfloor \hat{W}_t^\tau \rfloor < 1$ and hence there will never be another cycle in the same round t that uses this arc e. Thus, each cycle that is cleared in a round t can be charged to a unique arc in the graph. There are at most k^2 arcs in any round. Hence, we can clear at most k^2 cycles before there are no cycles left at round t. □

Lemma 3. *Assuming the condition in (2), if* `PTTC` *cannot find a cycle with noisy weight ≥ 1, then a node v has $n_v < D$ for $D = (3E + 1)k$.*

Proof. If there is no cycle of noisy weight at least 1, then there must be some node v such that every outgoing arc has $\hat{w}_e < 1$. From our Low Error Claim, we then know that the exact arc weight $w_e < 3E + 1$ from (2) for every outgoing arc of v. Hence we can count how may people are on the node $n_v = \sum_{e:e=(v,z)} w_e < k(3E + 1) = D$. □

We first give a lemma that will be useful in proving marginal differential privacy, given an intermediate step that is differentially private.

Lemma 4. *Let $M' : T^n \to O$ be ϵ_1-differentially private. Let $M''_j : \times T^{n-1} \times T \times O \to R$ for $j = 1, \cdots, n$ be ϵ_2- differentially private in its first argument. Then the mechanism $M : T^n \to R^n$ is $\epsilon_1 + \epsilon_2$-marginally differentially private where*

$$M(\mathbf{t}) = (M''_j(\mathbf{t}_{-j}, t_j, M'(\mathbf{t})))_{j=1}^n$$

Proof. To prove marginal differential privacy we need to prove that $M(\mathbf{t}_{-j}, t_j)_j$ is differentially private in its first argument, for every j. Fix any index $i \neq j$, $\mathbf{t}_{-i} \in T^{n-1}$, $t'_i \neq t_i$, and $S \subset R$, then we can use the composition properties of differentially private mechanisms to get the following

$$\begin{aligned}
&\mathbb{P}\left(M''_j[\mathbf{t}_{-j}, t_j, M'(\mathbf{t})] \in S\right) = \int_O \mathbb{P}(M''_j((\mathbf{t}_{-(i,j)}, t_i), t_j, o) \in S|o)\mathbb{P}(M'(\mathbf{t}_{-i}, t_i) = o)do \\
&\leq \int_O \mathbb{P}(M''_j((\mathbf{t}_{-(i,j)}, t_i), t_j, o) \in S|o)e^{\epsilon_1}\mathbb{P}(M'(\mathbf{t}_{-i}, t'_i) = o)do \\
&\leq \int_O e^{\epsilon_2}\mathbb{P}(M''_j((\mathbf{t}_{-(i,j)}, t'_i), t_j, o) \in S|o)e^{\epsilon_1}\mathbb{P}(M'(\mathbf{t}_{-i}, t'_i) = o)do \\
&= e^{\epsilon_1+\epsilon_2}\mathbb{P}(M''_j[(\mathbf{t}_{-(i,j)}, t'_i), t_j, M'(\mathbf{t}_{-i}, t'_i)] \in S)
\end{aligned}$$

□

Theorem 4. *PTTC is (ϵ, β)-marginally differentially private when $k \geq 2$.*

Before we prove this theorem, we first state the definition of sensitivity of a function, which will help when we define the Laplace Mechanism [Dwork et al., 2006] which is a subroutine of PTTC.

Definition 7 (Sensitivity [Dwork et al., 2006])**.** *A function $\phi : T^n \to \mathbb{R}^m$ has sensitivity $\delta(\phi)$ defined as*

$$\delta(\phi) = \max_{\substack{i \in [n] \\ \mathbf{t}_{-i} \in T^{n-1} \\ t_i \neq t'_i}} ||\phi(t_i, \mathbf{t}_{-i}) - \phi(t'_i, \mathbf{t}_{-i})||_1.$$

The Laplace Mechanism, given in Algorithm 1, answers a numeric query $\phi : T^n \to \mathbb{R}^m$ by adding noise to each component of ϕ's output.

Algorithm 1 Laplace Mechanism

Input: : Type profile $\mathbf{t}$.
Output: : An approximate value for ϕ
  procedure $M_L(\epsilon, g)(\mathbf{t})$
    $\hat{\phi}(\mathbf{t}) = \phi(\mathbf{t}) + (Z_1, \cdots, Z_m) \qquad Z_i \overset{i.i.d.}{\sim} \text{Lap}(\delta(\phi)/\epsilon)$
    return $\hat{\phi}(\mathbf{t})$.
  end procedure

We now state, without proof, that the Laplace Mechanism M_L is ϵ-differentially private.

Theorem 5. *$M_L(\epsilon, \phi)$ is ϵ-differentially private for any $\phi : T^n \to \mathbb{R}^n$ with bounded sensitivity [Dwork et al., 2006].*

Our algorithm PTTC uses the Laplace Mechanism as a subroutine to modify the arc and node weights. We are now ready to prove Theorem 3.

Proof for Theorem 3. Fix any player i and pair of types $x_i = (g_i, \succ_i)$ and $x_i' = (g_i', \succ_i')$.

We will condition first on the event that the condition in Claim 2 holds. We then want to bound the following ratio for $i \neq j$, $\mathbf{x}_{-i} \in \mathcal{X}^{n-1}$, for any $g \in \mathcal{G}$ for $i \neq j$ and $x_i \neq x_i'$:

$$\frac{\mathbb{P}\left(\texttt{PTTC}(x_i, \mathbf{x}_{-i})_j = g\right)}{\mathbb{P}\left(\texttt{PTTC}(x_i', \mathbf{x}_{-i})_j = g\right)}. \tag{3}$$

In order to match the setting of Lemma 4, we first construct a differentially private mechanism M'. At each round t the algorithm publishes, say on a billboard, the noisy weights for the τ'th cycle $\hat{w}_e(t,\tau)$ for each arc. We define $\phi_A : \mathcal{X}^n \to \mathbb{R}^{k^5}$ as

$$\phi_A(\mathbf{x}) = (w_e(t,\tau))_{e\in A, t\in[k], \tau\in[k^2]}.$$

Note that the sensitivity of ϕ_A is $2k^3$ because a person can change the weight by 1 on at most 2 arcs at every round $t = 1, \cdots, k$ and cycle $\tau = 1, \cdots, k^2$. Thus running $M_L\left(\phi_A, \frac{\epsilon}{2}\right)(\mathbf{x})$ gives a $\frac{\epsilon}{2}$-differentially private output. In PTTC, we are adding Laplacian noise to each edge weight with parameter $\frac{\delta(\phi_A)}{(\epsilon/2)} = \frac{4k^3}{\epsilon} = 1/\epsilon'$ from the definition of $\epsilon' = \frac{\epsilon}{4k^3}$. We then define $M' : \mathcal{X}^n \to O$ where $O \subset \mathbb{R}^{k^5}$ and

$$M'(\mathbf{x}) = M_L(\phi_A, \epsilon/2)(\mathbf{x})$$

which is $\epsilon_1 = \frac{\epsilon}{2}$ -differentially private.

Our algorithm PTTC then uses the output of M' coupled with the original data to assign a type of good to every person. In order to fit the notation of Lemma 4 to the context of our algorithm PTTC, we let $M_j'' : \mathcal{X}^{n-1} \times \mathcal{X} \times O \to \mathcal{G}$ be the allocation to the j'th person. At each round t given the τ'th cycle to clear and the approximate edge weights $\hat{w}_e(t,\tau)$, PTTC selects individuals to trade uniformly at random from those that wanted to trade, i.e. from $P_e^{t,\tau}$ for all e in the cycle C_t^τ. Note that the number of trades we make from one good u to another v is $\hat{W}_t^\tau$ from the $w_e(t,\tau)$ people that currently want to trade along that edge $e = (u, v)$.

We now bound the ratio between the probabilities that j is selected at the τ^{th} cycle given the edge weights $\hat{w}_e(t,\tau)$ when $i \neq j$ changes her type. We will write $w_e'(t,\tau)$ to denote the actual number of people on edge e at round t, cycle τ when i has type x_i', instead of x_i. Let $M_j^{t,\tau} : \mathcal{X}^{n-1} \times \mathcal{X} \times O \to \{0\} \cup \mathcal{G}$, where

$$M_j^{t,\tau}(\mathbf{x}_{-i}, x_i, o) = \begin{cases} g & \text{If } j \in S_e^{t,\tau} \text{ for some } e = (g_j, g) \in C_t^\tau \\ 0 & \text{else} \end{cases}$$

This gives us

$$\begin{aligned} \frac{\mathbb{P}\left(M_j^{t,\tau}(\mathbf{x}_{-i}, x_i, o) = g\right)}{\mathbb{P}\left(M_j^{t,\tau}(\mathbf{x}_{-i}, x_i', o) = g\right)} &= \frac{\frac{\hat{W}_t^\tau}{w_e(t,\tau)}}{\frac{\hat{W}_t^\tau}{w_e'(t,\tau)}} \leq \frac{\frac{\hat{W}_t^\tau}{w_e(t,\tau)}}{\frac{\hat{W}_t^\tau}{w_e(t,\tau)+1}} \\ &= \frac{w_e(t,\tau)+1}{w_e(t,\tau)} \leq 1 + \frac{1}{E} \leq e^{1/E} \end{aligned} \tag{4}$$

The first inequality follows because i can at most move to edge e after she changed preferences. The last inequality comes from the fact that the cycle always has an edge with $\hat{w}_e = w_e + Z_e - 2E \geq 1 \implies w_e \geq E + 1$. We now consider the case when j is not selected at round t, cycle τ when his edge is on the cycle.

$$\begin{aligned} \frac{\mathbb{P}\left(M_j^{t,\tau}(\mathbf{x}, o) = 0\right)}{\mathbb{P}\left(M_j^{t,\tau}((x_i', \mathbf{x}_{-i}), o) = 0\right)} &= \frac{1 - \frac{\hat{W}_t^\tau}{w_e(t,\tau)}}{1 - \frac{\hat{W}_t^\tau}{w_e'(t,\tau)}} \leq \frac{1 - \frac{\hat{W}_t^\tau}{w_e(t,\tau)}}{1 - \frac{\hat{W}_t^\tau}{w_e(t,\tau)-1}} \\ &= \frac{(w_e(t,\tau) - \hat{W}_t^\tau)(w_e(t,\tau)-1)}{w_e(t,\tau) \cdot (w_e(t,\tau) - \hat{W}_t^\tau - 1)} \leq \frac{w_e(t,\tau) - \hat{W}_t^\tau}{w_e(t,\tau) - \hat{W}_t^\tau - 1}. \end{aligned} \tag{5}$$

The first inequality holds because person i could at most move off of e if she changes preferences. Recall that we have $w_e(t,\tau) - \hat{w}_e(t,\tau) \geq E$ from (2) $\implies w_e(t,\tau) - \hat{W}_t^\tau \geq E$. Hence, we can further bound our ratio in (5) by

$$\frac{\mathbb{P}\left(M_j^{t,\tau}(\mathbf{x}, o) = 0\right)}{\mathbb{P}\left(M_j^{t,\tau}((x_i', \mathbf{x}_{-i}), o) = 0\right)} \leq \frac{E}{E-1} \leq 1 + \frac{2}{E} \leq e^{2/E}. \tag{6}$$

We then define $M_j : \mathcal{X}^{n-1} \times \mathcal{X} \times O \to \{\mathcal{G} \cup \{0\}\}^{k^3}$ as

$$M_j(\mathbf{x}_{-i}, x_i, o) = \left((M_j^{t,\tau}(\mathbf{x}_{-i}, x_i, o))_{\tau \in [k^2]}\right)_{t \in [k]} \qquad i \neq j.$$

Because the output of M_j is nonzero in at most one entry of its output, we can apply $f : \{\mathcal{G} \cup \{0\}\}^{k^3} \to \mathcal{G}$ to the output of M_j to get the good that j is allocated, where

$$f(M_j(\mathbf{x}_{-i}, x_i, o)) = \begin{cases} g & \text{if there is a nonzero entry of } M_j(\mathbf{x}_{-i}, x_i, o) \text{ with value } g \\ g_i & \text{if all entries of } M_j(\mathbf{x}_{-i}, x_i, o) \text{ are zero} \end{cases}$$

We now define the mechanism $M_j'' : \mathcal{X}^{n-1} \times \mathcal{X} \times O \to \mathcal{G}$ as

$$M_j''(\mathbf{x}_{-i}, x_i, o) = f(M_j(\mathbf{x}_{-i}, x_i, o)) \qquad i \neq j$$

Next, we show that M_j'' is differentially private in its first argument. We need to only consider the randomization in $M_j^{t,\tau}$ for every t and τ. If j is not on an arc e in cycle C_t^τ then $i \neq j$ changing type is not going to effect the probability of j being chosen to trade, which is zero. Further, we know that once a player is allocated a good, she is eliminated from further consideration. Hence, we must consider the case that j is on an arc contained in C_t^τ at each round t and each cycle τ but never gets chosen and hence j gets her original good g_j. We use (6) to examine this outcome, which gives us,

$$\frac{\mathbb{P}\left(M_j^{t,\tau}(\mathbf{x}_{-i}, x_i, o) = 0, \quad \forall t, \tau\right)}{\mathbb{P}\left(M_j^{t,\tau}\mathbf{x}_{-i}, x_i', o) = 0, \quad \forall t, \tau\right)} \leq \exp\left(\frac{2k^3}{E}\right)$$

Thus, M_j'' is ϵ_2-differentially private in its first argument for $\epsilon_2 = \frac{2k^3}{E}$.

We now invoke Lemma 4 to say that $\texttt{PTTC}(\mathbf{x}) = \left(M_j''(\mathbf{x}, M'(\mathbf{x}))\right)_{j=1}^n$ is $\epsilon_1 + \epsilon_2$-marginally differentially private where

$$\begin{aligned} \epsilon_1 + \epsilon_2 &= \epsilon/2 + 2k^3/E = \epsilon/2 + \frac{2k^3\epsilon'}{\log(2k/\beta)} \\ &\leq \epsilon/2 + 2k^3\epsilon' = \epsilon \end{aligned}$$

where we are using $E = \frac{\log(2k^3/\beta)}{\epsilon'}$ and $\epsilon' = \frac{\epsilon}{4k^3}$.

The above analysis holds when the algorithm does not fail prior to allocating a good to every person, so we get ϵ-marginal differential privacy with probability $1 - \beta$. However, we cannot bound the ratio of the distributions in this case, hence we get the additive β term to conclude that PTTC is (ϵ, β)-marginally differentially private. □

It remains to show that the resulting allocation is asymptotically Pareto optimal.

Theorem 6. *PTTC outputs an allocation that is α-approximately Pareto Optimal for $\alpha = \tilde{O}\left(\frac{k^6}{\epsilon n}\right)$, with probability $1 - \beta$. Hence, PTTC is asymptotically Pareto optimal.*

Proof. We refer the reader to the formal description of our algorithm (Algorithm 2) in Appendix **??**. Recall that N is the set of agents in the market. We relabel the goods in the order in which they are deleted: good type $g = 1$ to be the first good eliminated in PTTC, good $g = 2$

the second, and so on, where the good types $g \in \mathcal{G} = [k]$. We will compare the allocation π from PTTC with any other allocation π' that Pareto dominates π, i.e. $\pi'(i) \succeq_i \pi(i)$ and $\exists j$ such that $\pi'(j) \succ \pi(j)$. We will count the number of agents who could possibly prefer π'

First, we count the total number of goods which the algorithm can "ignore" by giving them to their owners (who have Pareto improving trades available according to π). This happens only at the end of a round, when the shifted noisy supply of at least one good falls below 0. Thus, by Claim 2 and Lemma 3, there are at most $D = k(3E + 1)$ copies of goods for which this is the case, with probability at least $1 - \beta$. We condition on this bounded error for the remainder of the proof.

Recall that $S_e^{t,\tau}$ is the set of people that traded that were along edge e at round t when the τ'th cycle C_t^τ was selected in PTTC. We define the set $S(t)$ to be all the people that traded at round t: $S(t) = \cup_{\tau \in [k^2]} \cup_{e \in C_t^\tau} S_e^{t,\tau}$.

Some agents are not cleared at any round of PTTC, and these agents receive the good they were endowed with. We refer to those people that were never selected as $S(n) = N \backslash \{\cup_{t=1}^k S(t)\}$. We then have a partition of $N = \cup_{t=1}^k S(t) \cup S(n)$. It suffices to bound the following quantity: $\Delta = |\{i \in N : \pi'(i) \succ_i \pi(i)\}|$.

We now denote the quantity $\Delta_r = |\{i \in S(r) : \pi'(i) \succ_i \pi(i)\}|$ for $r = 1, \cdots, k, n$: note that $\sum_{r \in \{1,\ldots,k,n\}} \Delta_r = \Delta$. Further, we would like to refer to the number of goods of type g allocated to agents in $S(r)$ in π' but not in π, which we define as $\Delta_r(g)$. Note that for $g \geq r$ we know that $\Delta_r(g) = 0$ because agents cleared by PTTC receive their favorite good among the one's remaining at the round in which they are cleared, and all goods $g \geq r$ are available at round r:

$$\Delta_r(g) = |\{i \in S(r) : \pi'(i) = g \succ_i \pi(i)\}| \text{ for } r = 1, \cdots, k, \quad g < r.$$

Thus we can write: $\Delta_r = \sum_{g=1}^{r-1} \Delta_r(g)$

Note that at the first round $r = 1$ that $\Delta_1 = 0$ because everyone that was selected gets their favorite type of good. Let us also define $\Delta_r(g, h)$ to be the number of goods of type g allocated to agents in $S(r)$ in π' who received good h in π:

$$\Delta_r(g, h) = |\{i \in S(r) : \pi'(i) = g \succ_i h = \pi(i)\}| \text{ for } r = 1, \cdots, k, \quad h \geq r \quad g < r.$$

$$\Delta_r(g) = \sum_{h=r}^{k} \Delta_r(g, h)$$

We denote the number of initial goods of type g as $n(g)$ for $g \in [k]$: i.e. $n(g) = |i : g_i = g|$. We define $n_t(g)$ as the number of goods of type g that are not allocated to members of $\cup_{r=1}^t S(r)$ in π'. Our notation uses the round as a subscript and the good as an argument in parentheses.

$$n_t(g) = n(g) - \sum_{r=1}^{t} |\{i \in S(r) : \pi'(i) = g\}|.$$

We will now bound the quantities $n_t(g)$. Note that

$$n_1(1) = n(1) - |\{i \in S(1) : \pi'(i) = 1| \leq D$$

because each person in $S(1)$ got their favorite good in PTTC, and since π' Pareto dominates π, π' must have made the same allocation as π for $S(1)$, except for the at most D agents PTTC never selected that had a good of type 1. The agents that get selected later can Pareto improve because they could have selected good type 1, but PTTC has deleted that good for future rounds. These "lost" copies of good 1 can potentially be used in allocation π' to improve the outcome of agents selected at future rounds by PTTC. We will account for these improvements by keeping track of the "extra" copies of good 1 $n_t(1)$ remaining, where

$$n_t(1) = n_{t-1}(1) - \Delta_t(1) \qquad \text{for } t = 2, \cdots, k.$$

We then continue in this fashion with good type 2 in order to bound $n_2(2)$. We know that π' allocates $\Delta_2(1)$ goods of type 1 to agents in $S(2)$. Because π' Pareto dominates π, it must be that π' makes the same allocations as π among agents in $S(2)$, except for the agents that π'

matches to good type 1 that got good type 2 in π (These are the only agents who are possibly not getting their favorite good among those "remaining" in π').

$$n_2(2) \leq D + \Delta_2(1,2)$$
$$n_r(2) = n_{r-1}(2) - \Delta_2(1) \qquad r > 2.$$

We now consider $n_3(3)$, the number of goods of type 3 that π' has not allocated to members in $S(1), S(2)$, and $S(3)$. This is the same as the number of goods of type 3 that π will ever give to selected people (at most D) in addition to the goods that π gave good type 3 to people in $S(2)$ and $S(3)$ that π' gave a different good to, i.e. the $\Delta_2(1,3)$ people that got good 3 at round 2 in π, but π' gave them good 1, along with the $\Delta_3(1,3)$ and $\Delta_3(2,3)$ people that got good 3 at round 3 in π, but π' gave them good 1 and 2 respectively. This gives us

$$n_3(3) \leq D + \Delta_2(1,3) + \Delta_3(1,3) + \Delta_3(2,3)$$
$$n_r(2) = n_{r-1}(3) - \Delta_3(1) \qquad r > 3$$

We then have the relation for $r \geq 3$:

$$n_r(r) \leq D + \sum_{\ell=2}^{r} \sum_{g=1}^{\ell-1} \Delta_\ell(g,r) \tag{7}$$
$$n_t(r) = n_{t-1}(r) - \Delta_t(r) \qquad t > r. \tag{8}$$

Because the number of goods remaining at each round must be nonnegative, we must have

$$\Delta_t(r) \leq n_{t-1}(r). \tag{9}$$

We know that $\Delta_1 = 0$ because of the people selected at round one, they all got their favorite good. We also have: $\Delta_2 = \Delta_2(1) \leq n_1(1) \leq D$. For round $3 \leq t \leq k$, we use (7), (8), and (9) to get:

$$\begin{aligned}
\Delta_t &= \sum_{g=1}^{t-1} \Delta_t(g) \underbrace{\leq}_{(9)} \sum_{g=1}^{t-1} n_{t-1}(g) \underbrace{=}_{(8)} \sum_{r=1}^{t-1} n_r(r) - \sum_{g=1}^{t-2} \sum_{r=g+1}^{t-1} \Delta_r(g) \\
&\underbrace{\leq}_{(7)} (t-1)D + \sum_{r=2}^{t-1} \sum_{\ell=2}^{r} \sum_{g=1}^{\ell-1} \Delta_\ell(g,r) - \sum_{g=1}^{t-2} \sum_{r=g+1}^{t-1} \Delta_r(g) \\
&= (t-1)D + \sum_{\ell=2}^{t-1} \sum_{g=1}^{\ell-1} \underbrace{\sum_{r=\ell}^{t-1} \Delta_\ell(g,r)}_{\leq \Delta_\ell(g)} - \sum_{r=2}^{t-1} \sum_{g=1}^{r-1} \Delta_r(g) \\
&\leq (t-1)D
\end{aligned}$$

We next bound Δ_n. All the people that were never selected could get a better good in π', which is not more than the total goods that π' never to allocated to any of the selected people. Thus we have

$$\begin{aligned}
\Delta_n &\leq \sum_{g=1}^{k} n_k(g) = \sum_{r=1}^{k} n_r(r) - \sum_{g=1}^{k-1} \sum_{r=g+1}^{k} \Delta_r(g) \\
&\leq kD + \sum_{\ell=2}^{k} \sum_{g=1}^{\ell-1} \underbrace{\sum_{r=\ell}^{k} \Delta_\ell(g,r)}_{= \Delta_\ell(g)} - \sum_{r=2}^{k} \sum_{g=1}^{r-1} \Delta_r(g) \\
&= kD
\end{aligned}$$

We then sum over all the Δ_t's to get Δ.

$$\Delta_1 + \Delta_2 + \Delta_3 + \cdots + \Delta_k + \Delta_n \leq \sum_{j=1}^{k} jD = \mathcal{O}\left(k^2 D\right) = \mathcal{O}\left(\frac{k^6 \log(k/\beta)}{\epsilon}\right)$$

□

Lemma 5. *The bound of $\mathcal{O}(k^2 D)$ for α in Theorem 6 is tight.*

Proof. We will assume that at each round of PTTC, it leaves exactly its upper bound D copies of each good unallocated. As in the previous proof, we assume that the label of the type of good is the same as the ordering in which they are depleted. We use the same notation as in the proof above. We will assume that

$$\Delta_t = \Delta_t(t-1) = \Delta_t(t-1, t) \qquad t = 2, \cdots, k. \tag{10}$$

Further, we will assume that $\forall i \in S(t)$ such that $\pi'(i) \succ_i \pi(i)$ we have $\pi'(i) = t-1$ and $\pi(i) = t = g_i$. Note that this fits the assumption in (10).

What our assumptions are saying is that of the people selected at each round t that can do better in π', they only prefer the good that was depleted at the previous round $t-1$. In addition, these people that can improve were all selected at round t to get their own good type in π that they were endowed with t, i.e. the good depleted at that current round.

All the people that can improve from those that were selected at round t must then use up all $n_{t-1}(t-1)$ goods of type $t-1$ that π' did not allocate to any of the people selected in rounds up to $t-1$. This gives us that

$$\Delta_t = n_{t-1}(t-1) \qquad t = 2, \cdots, k.$$

From (7) we then have

$$n_t(t) = D + n_{t-1}(t-1) = tD \qquad t = 1, \cdots, k.$$

We now need to make sure that we can perform trades where everyone in $S(t)$ can get the good that was depleted at round $t-1$. To do this, we simply have all the people that were never selected $S(n)$ prefer good k (the last one depleted) the most. This will create a cycle with everyone that could improve from π and thus the trades can be made by following the cycle. Figure 3 gives a visual depiction of this example. □

Note that the D people that had good type k that were never selected actually do prefer their type of good in this construction. Thus $\Delta_n = (k-1)D$. We then have

$$\Delta_1 + \Delta_2 + \cdots + \Delta_k + \Delta_n = \sum_{t=2}^{k} (t-1)D + \Delta_n = \frac{k(k+1)}{2} D - D$$

5 Conclusion/Open Problems

In this paper we have continued the study of the accuracy to which *allocation problems* can be solved under parameterized relaxations of *differential privacy*. Generically, these kinds of problems cannot be solved under the standard constraint of differential privacy. Unlike two sided allocation problems which can be solved under *joint*-differential privacy, we show that pareto optimal exchanges cannot be solved even under this relaxation, but can be solved asymptotically exactly under marginal differential privacy whenever the number of types of goods $k = o(n^{1/6})$. (We note that in many applications, such as kidney exchange, k will be constant).

Our understanding of this family of relaxations of differential privacy remains incomplete: what problems can be solved under the constraints that an adversary trying to learn about i's data can view the portion of the output shared with c agents other than player i? Joint differential privacy corresponds to $c = n-1$ and marginal differential privacy corresponds to $c = 1$. For $1 < c < n-1$, the problem remains largely unexplored.

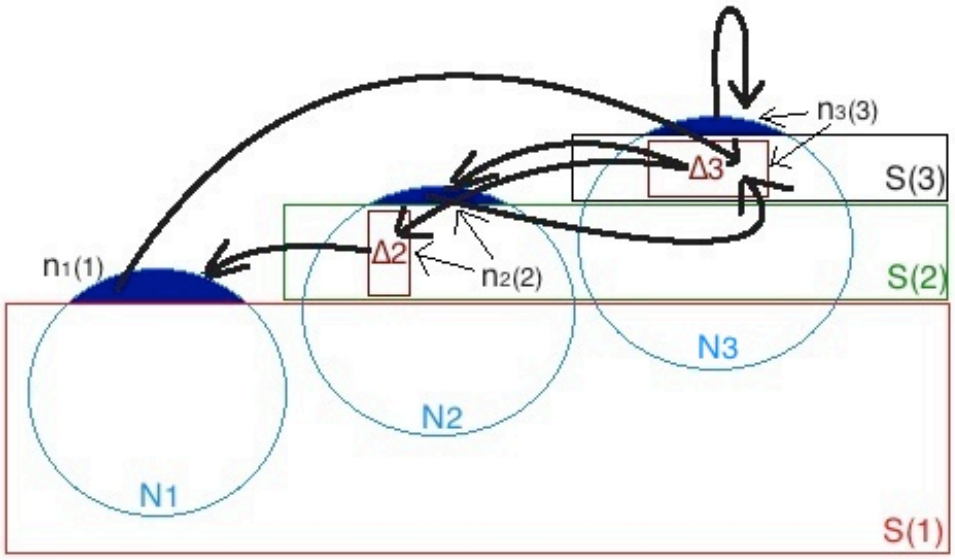

Figure 3: Depicting the argument in the proof of Lemma 5 with 3 types of goods. After the $S(1)$ people are selected at round 1, there are D goods of type 1 left unallocated (shown in blue). The trades that can improve upon the allocation in π are shown as a bold directed arc. There are Δ_2 people at round two that can trade with the D goods of type 1, but were selected to have good 2. This leaves $n_2(2) = D + n_1(1)$. Similarly $n_3(3) = D + n_2(2)$. We then have the D remaining people in N_1, N_2, and N_3 trade with good type 3.

References

Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *TCC '06*, pages 265–284, 2006.

Anupam Gupta, Katrina Ligett, Frank McSherry, Aaron Roth, and Kunal Talwar. Differentially private combinatorial optimization. In *Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1106–1125. Society for Industrial and Applied Mathematics, 2010.

Kevin He and Xiaosheng Mu. Differentially private and incentive compatible recommendation system for the adoption of network goods. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 949–966. ACM, 2014.

Justin Hsu, Zhiyi Huang, Aaron Roth, Tim Roughgarden, and Zhiwei Steven Wu. Private matchings and allocations. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, STOC '14, pages 21–30, New York, NY, USA, 2014. ACM. ISBN 978-1-4503-2710-7. doi: 10.1145/2591796.2591826. URL `http://doi.acm.org/10.1145/2591796.2591826`.

Michael Kearns, Mallesh Pai, Aaron Roth, and Jonathan Ullman. Mechanism design in large games: Incentives and privacy. In *Proceedings of the 5th ACM SIGact Innovations in Theoretical Computer Science (ITCS)*, 2014.

Frank McSherry and Ilya Mironov. Differentially private recommender systems: building privacy into the net. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 627–636. ACM, 2009.

Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith. Smooth sensitivity and sampling in private data analysis. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 75–84. ACM, 2007.

Ryan M Rogers and Aaron Roth. Asymptotically truthful equilibrium selection in large congestion games. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 771–782. ACM, 2014.

Alvin E Roth, Tayfun Sönmez, and M Utku Ünver. Pairwise kidney exchange. *Journal of Economic Theory*, 125(2):151–188, 2005.

Lloyd Shapley and Herbert Scarf. On cores and indivisibility. *Journal of Mathematical Economics*, 1(1):23–37, March 1974. URL `http://ideas.repec.org/a/eee/mateco/v1y1974i1p23-37.html`.

A Appendix - Formal Algorithm PTTC

Algorithm 2 Private Top Trading Cycles

Input: : Exchange Market $\mathbf{x} = (\mathbf{g}, \succ)$.

procedure PTTC$(\mathbf{g}, \succ)$

Set $\epsilon' = \frac{\epsilon}{4k^3}$ and $E = \frac{4k^3 \log(2k/\beta)}{\epsilon}$.

Initialize $P_e^{1,0} \leftarrow P_e$, $t \leftarrow 1$, and $\tau \leftarrow 0$.

1. (Arc Weight Noise) For each $e \in A$ we set

$$\hat{w}_e(t,\tau) = w_e(t,\tau) + Z_e^t - 2E \qquad \text{where } Z_e^t \sim Lap(1/\epsilon').$$

if $|Z_e^t| > E$ **return** FAIL and then $\pi(i) = g_i \qquad \forall i \in N$.

2. (Clear Cycle) **while** there is a cycle with positive weight **do** $\tau \leftarrow \tau + 1$

Denote the cycle by C_t^τ and let $\hat{W}_t^\tau \leftarrow \min_{e \in C_t^\tau}\{\lfloor \hat{w}_e(t,\tau) \rfloor\}$.

3. (Trade) **for** $e = (u,w) \in C_t^\tau$ **do**

Set $S_e^{t,\tau}$ as the $\hat{W}_t^\tau$ people chosen uniformly at random from $P_e^{t,\tau-1}$. Update:

$$P_e^{t,\tau} \leftarrow P_e^{t,\tau-1} \backslash S_e^{t,\tau}$$

$$w_e(t,\tau) \leftarrow w_e(t,\tau-1) - \hat{W}_t^\tau \quad \& \quad \hat{w}_e(t,\tau) \leftarrow \hat{w}_e(t,\tau-1) - \hat{W}_t^\tau$$

$$\pi(i) = v \quad \forall i \in S_e^{t,\tau} \text{ where } g_i = u.$$

4. (No Cycle - Deletion) If there is no cycle with positive rounded down noisy weight, then there exists a node v s.t. $\hat{n}_v(t,\tau) = \sum_{e:e=(v,w)} \hat{w}_e(t,\tau) < k$ (Lemma 3). Let

$$IN_v = \{e \in A : e = (u,v) \text{ some } u \in V^t\} \quad \& \quad OUT_v = \{e \in A : e = (v,u) \text{ some } u \in V^t\}$$

We then update:

$$V^{t+1} \leftarrow V^t \backslash \{v\} \qquad \& \qquad A \leftarrow A \backslash \{OUT_v \cup IN_v\}.$$

for all $e = (u,w) \in A$ **do**

Define $B_e^t = \{i \in P_{(u,v)}^{t,\tau} : w \succeq_i g \quad \forall g \in V^{t+1}\}$ and update:

$$w_e(t+1,0) \leftarrow w_e(t,\tau) + |B_e^t| \quad \& \quad P_e^{t+1,0} \leftarrow P_e^{t,\tau} \cup B_e^t$$

and assign goods

$$\pi(i) = g_i \qquad \forall i \in P_e^{t,\tau} \text{ where } e \in OUT_v.$$

Set $t \leftarrow t+1$ and $\tau \leftarrow 0$. Return to Step 1.

Output: π

end procedure