Privacy Auctions for Recommender Systems*

Pranav Dandekar† Nadia Fawaz‡ Stratis Ioannidis§

June 7, 2013

Abstract

We study a market for private data in which a data analyst publicly releases a statistic over a database of private information. Individuals that own the data incur a cost for their loss of privacy proportional to the differential privacy guarantee given by the analyst at the time of the release. The analyst incentivizes individuals by compensating them, giving rise to a *privacy auction*. Motivated by recommender systems, the statistic we consider is a linear predictor function with publicly known weights. The statistic can be viewed as a prediction of the unknown data of a new individual, based on the data of individuals in the database. We formalize the trade-off between privacy and accuracy in this setting, and show that a simple class of estimates achieves an order-optimal trade-off. It thus suffices to focus on auction mechanisms that output such estimates. We use this observation to design a truthful, individually rational, proportional-purchase mechanism under a fixed budget constraint. We show that our mechanism is 5-approximate in terms of accuracy compared to the optimal mechanism, and that no truthful mechanism can achieve a $2-\varepsilon$ approximation, for any $\varepsilon > 0$.

*An extended abstract of this paper appeared in the Proceedings of WINE 2012.

†Department of Management Science & Engineering, Stanford University. Email: ppd@stanford.edu. Research done in part while interning at Technicolor, Palo Alto, CA.

‡Technicolor, Palo Alto, CA. Email: nadia.fawaz@technicolor.com

§Technicolor, Palo Alto, CA. Email: stratis.ioannidis@technicolor.com

1 Introduction

Recommender systems are ubiquitous on the Internet, lying at the heart of some of the most popular Internet services, including Netflix, Yahoo, and Amazon. These systems use algorithms to predict, *e.g.*, a user's rating for a movie, her propensity to click on an advertisement or to purchase a product online. By design, such prediction algorithms rely on large training datasets, typically comprising data from thousands (often millions) of individuals. This has raised serious privacy concerns among researchers and consumer advocacy groups. Privacy researchers have shown that access to seemingly non-sensitive data (*e.g.*, movie ratings) can leak potentially sensitive information when combined with de-anonymization techniques [NS08]. Moreover, a spate of recent lawsuits [Net, Mel12, Rib12] as well as behavioral studies [JKH+09] have demonstrated the increasing reluctance of the public to allow the unfettered use of their data.

However, a widespread restriction on data used by recommender systems would be detrimental to the quality of their recommendations. One way to address this tension between the value of data and the users' need for privacy is through *incentivization*. In short, a recommender system using an individual's data ought to appropriately compensate her for the violation of her privacy, thereby incentivizing her consent to this use.

We study the issue of user incentivization through *privacy auctions*, as introduced by Ghosh and Roth [GR11]. In a privacy auction, a data analyst has access to a database $\mathbf{d} \in \mathbb{R}^n$ of private data d_i, $i = 1, \ldots, n$, each corresponding to a different individual. This data may represent information that is to be protected, such as an individual's propensity to click on an ad or purchase a product, or the number of visits to a particular website. As in Ghosh and Roth [GR11], we assume a *verified* database in which individuals cannot lie about their data. The analyst wishes to publicly release an estimate $\hat{s}(\mathbf{d})$ of a statistic $s(\mathbf{d})$ evaluated over the database. In addition, each individual incurs a privacy cost c_i upon the release of the estimate $\hat{s}(\mathbf{d})$, and must be appropriately compensated by the analyst for this loss of utility. The analyst has a budget, which limits the total compensation paid out. As such, given a budget and a statistic s, the analyst must (a) solicit the costs of individuals c_i and (b) determine the estimate $\hat{s}$ to release as well as the appropriate compensation to each individual.

Ghosh and Roth employ *differential privacy* [DMNS06] as a principled approach to quantifying the privacy cost c_i. Informally, ensuring that $\hat{s}(\mathbf{d})$ is ϵ-differentially private with respect to individual i provides a guarantee on the privacy of this individual; a small ϵ corresponds to better privacy since it guarantees that $\hat{s}(\mathbf{d})$ is essentially independent of the individual's data d_i. Privacy auctions incorporate this notion by assuming that each individual i incurs a cost $c_i = c_i(\epsilon)$, that is a function of the privacy guarantee ϵ provided by the analyst.

1.1 Our Setting

Motivated by recommender systems, we focus in this paper on a scenario where the statistic s takes the form of a *linear predictor*:

$$s(\mathbf{d}) := \langle \mathbf{w}, \mathbf{d} \rangle = \textstyle\sum_{i=1}^{n} w_i d_i, \tag{1}$$

where $\mathbf{w} \in \mathbb{R}^n$, is a publicly known vector of real (possibly negative) weights. Intuitively, the public weights w_i serve as measures of the similarity between each individual i and a new individual, outside the database. The function $s(\mathbf{d})$ can then be interpreted as a prediction of the value d for this new individual.

For the sake of concreteness, assume that each individual $i \in [n] = \{1, \ldots, n\}$ is endowed with a public vector $\mathbf{y}_i \in \mathbb{R}^m$, which includes m publicly known features about this individual. These could be, *e.g.*, demographic information such as age, gender or zip code, that the individual discloses in a public online profile. Note that, though features $\mathbf{y}_i$ are public, the data d_i is private. Let $\mathbf{Y} = [\mathbf{y}_i]_{i \in [n]} \in \mathbb{R}^{n \times m}$ be a matrix comprising public feature vectors. Consider a new individual, not belonging to the database, whose public feature profile is $\mathbf{y} \in \mathbb{R}^m$. Having access to $\mathbf{Y}$, $\mathbf{d}$, and $\mathbf{y}$, the data analyst wishes to release a prediction for the unknown value d for this new individual. In many practical cases, this prediction takes the form $s(\mathbf{d}) = \langle \mathbf{w}, \mathbf{d} \rangle$, for some $\mathbf{w} = \mathbf{w}(\mathbf{y}, \mathbf{Y})$.

> **Example.** In *k-Nearest Neighbors* (k-NN) prediction [HTF09], the predicted value is given by an average among the k nearest neighbors of the feature vector $\mathbf{y}$ of the new individual among the vectors $\mathbf{Y}$. More specifically, let $N_k(\mathbf{y}) \subset [n]$ denote the k individuals whose feature vectors $\mathbf{y}_i$ are closest to $\mathbf{y}$ under a distance metric over $\mathbb{R}^m$ (*e.g.*, the ℓ_2 norm). Then, the prediction for the new individual is given by $s(\mathbf{d}) = \frac{1}{k} \sum_{i \in \mathcal{N}_k(\mathbf{y})} d_i = \sum_i w_i d_i$, where $w_i = \frac{1}{k}$ if $i \in N_k(\mathbf{y})$ and $w_i = 0$ otherwise.

Beyond k-NN, linear predictors of the form (1) include many well-studied methods of statistical inference, such as the Nadaranya-Watson weighted average, ridge regression, and support vector machines [HTF09]. We provide a brief review of such methods in Section 5. Functions of the form (1) are thus of particular interest in the context of recommender systems [SKKR01, LSY03], as well as other applications involving predictions (*e.g.*, polling/surveys, marketing). In the sequel, we ignore the provenance of the public weights $\mathbf{w}$, keeping in mind that any of these methods apply.

1.2 Our Contributions

Our contributions are as follows:

1. **Privacy-Accuracy Trade-off.** We characterize the accuracy of the estimate $\hat{s}$ in terms of the *distortion* between the linear predictor s and $\hat{s}$ defined as $\delta(s, \hat{s}) := \max_{\mathbf{d}} \mathbb{E}\left[|s(\mathbf{d}) - \hat{s}(\mathbf{d})|^2\right]$, *i.e.*, the maximum mean square error between $s(\mathbf{d})$ and $\hat{s}(\mathbf{d})$ over all databases $\mathbf{d}$. We define a *privacy index* $\beta(\hat{s})$ that captures the amount of privacy an estimator $\hat{s}$ provides to individuals in the database. We show that any estimator $\hat{s}$ with low distortion must also have a low privacy index (Theorem 1).

2. **Laplace Estimators Suffice.** We show that a special class of *Laplace estimators* [DMNS06, Dwo06] (*i.e.*, estimators that use noise drawn from a Laplace distribution), which we call Discrete Canonical Laplace Estimator Functions (DCLEFs), exhibits an order-optimal trade-off between privacy and distortion (Theorem 2). This allows us to restrict our focus on privacy auctions that output DCLEFs as estimators of the linear predictor s.

3. **Truthful, 5-approximate Mechanism, and Lower bound.** We design a *truthful*, *individually rational*, and *budget feasible* mechanism that outputs a DCLEF as an estimator of the linear predictor (Theorem 3). Our estimator's accuracy is a 5-approximation with respect to the DCLEF output by an optimal, individually rational, budget feasible mechanism. We also prove a lower bound (Theorem 4): there is no truthful DCLEF mechanism that achieves an approximation ratio $2 - \varepsilon$, for any $\varepsilon > 0$.

In our analysis, we exploit the fact that when $\hat{s}$ is a Laplace estimator minimizing distortion under a budget resembles the knapsack problem. As a result, the problem of designing a privacy auction that outputs a DCLEF $\hat{s}$ is similar in spirit to the knapsack auction mechanism [Sin10]. However, our setting poses an additional challenge because the privacy costs exhibit *externalities*: the cost incurred by an individual is a function of which other individuals are being compensated. Despite the externalities in costs, we achieve the same approximation as the one known for the knapsack auction mechanism [Sin10].

1.3 Related Work

There is a rich literature on differential privacy beginning with the work of Dwork *et al.* [DMNS06] who introduced it as a formal definition of database privacy. Informally, an algorithm is ϵ-differentially private if changing the data of a single individual does not change the probability of any outcome by more than an $e^{\epsilon} \approx (1+\epsilon)$ multiplicative factor.

Privacy of behavioral data. Differentially-private algorithms have been developed for the release of several different kinds of online user behavioral data such as click-through rates and search-query frequencies [KKMN09], as well as movie ratings [MM09]. As pointed out by McSherry and Mironov [MM09], the reason why the release of such data constitutes a privacy violation is not necessarily that, *e.g.*, individuals perceive it as embarrassing, but that it renders them susceptible to *linkage* and *de-anonymization attacks* [NS08]. Such linkages could allow, for example, an attacker to piece together an individual's address stored in one database with his credit card number or social security number stored in another database. It is therefore natural to attribute a loss of utility to the disclosure of such data.

Privacy auctions. Quantifying the cost of privacy loss allows one to study privacy in the context of an economic transaction. Ghosh and Roth initiate this study of privacy auctions in the setting where the data is binary and the statistic reported is the sum of bits, *i.e.*, $d_i \in \{0,1\}$ and $w_i = 1$ for all $i = 1, \ldots, n$ [GR11]. Unfortunately, the Ghosh-Roth auction mechanism cannot be readily generalized to asymmetric statistics such as (1), which, as discussed in Section 5, have numerous important applications including recommender systems. Our Theorems 1 and 2, which parallel the characterization of order-optimal estimators in [GR11], imply that to produce an accurate estimate of s, the estimator $\hat{s}$ *must provide different privacy guarantees to different individuals*. This is in contrast to the multi-unit procurement auction of [GR11]. In fact, as discussed the introduction, a privacy auction outputting a DCLEF $\hat{s}(\mathbf{d})$ has many similarities with a knapsack auction mechanism [Sin10], with the additional challenge of externalities introduced by the Laplacian noise (see also Section 4).

Privacy and truthfulness in mechanism design. A series of interesting results follow an orthogonal direction, namely, on the connection between privacy and truthfulness when individuals have the ability to misreport their data. McSherry and Talwar [MT07], and Nissim *et al.* [NST10] use privacy as a tool for mechanism design. Xiao [Xia11], Chen *et al.* [CCK+13] and Nissim *et al.* [NOS12] design truthful mechanisms for agents that value privacy (using differential privacy or other closely related definitions of privacy). As pointed out by Xiao [Xia11], all these papers consider an *unverified* database, *i.e.*, the mechanism designer cannot verify the data reported by individuals and therefore must incentivize them to report truthfully. Recent work on truthfully eliciting private data through a *survey* [LR12, RS12] also falls under the unverified database setting [Xia11]. In contrast, our setting, as well as that of Ghosh and Roth, and Fleischer *et al.* [FL12], is that of soliciting consent for use of data stored in a *verified* database, in which individuals cannot lie about their data. This setting is particularly relevant

to the context of online behavioral data: information on clicks, websites visited and products purchased is collected and stored in real-time and cannot be retracted after the fact.

Correlation between privacy costs and data values. An implicit assumption in privacy auctions as introduced in [GR11] is that the privacy costs c_i are *not* correlated with the data values d_i. This might not be true if, *e.g.*, the data represents the propensity of an individual to contract a disease. Ghosh and Roth [GR11] show that when the privacy costs are correlated to the data no individually rational direct revelation mechanism can simultaneously achieve non-trivial accuracy and differential privacy. As discussed in the beginning of this section, the privacy cost of the release of behavioral data is predominantly due to the risk of a linkage attack. It is reasonable in many cases to assume that this risk (and hence the cost of privacy loss) is not correlated to, *e.g.*, the user's movie ratings. Nevertheless, due to its importance in other settings such as medical data, more recent privacy auction models aim at handling such correlation [LR12, RS12, FL12]; we leave generalizing our results to such privacy auction models as future work.

2 Preliminaries

Let $[k] = \{1, \cdots, k\}$, for any integer $k > 0$, and define $\mathrm{I} := [R_{\min}, R_{\max}] \subset \mathbb{R}$ to be a bounded real interval. Consider a database containing the information of $n > 0$ individuals. In particular, the database comprises a vector $\mathbf{d}$, whose entries $d_i \in \mathrm{I}$, $i \in [n]$, represent the private information of individual i. Each entry d_i is *a priori* known to the database administrator, and therefore individuals do not have the ability to lie about their private data. A data analyst with access to the database would like to publicly release an estimate of the statistic $s(\mathbf{d})$ of the form (1), *i.e.* $s(\mathbf{d}) = \sum_{i\in[n]} w_i d_i$, for some publicly known weight vector $\mathbf{w} = (w_1, \ldots, w_n) \in \mathbb{R}^n$. For any subset $H \subseteq [n]$, we define $w(H) := \sum_{i\in H} |w_i|$, and denote by $W := w([n]) = \sum_{i=1}^{n} |w_i|$ the ℓ_1 norm of vector $\mathbf{w}$. We denote the length of interval I by $\Delta := R_{\max} - R_{\min}$, and its midpoint by $\bar{R} := (R_{\min} + R_{\max})/2$. Without loss of generality, we assume that $w_i \neq 0$ for all $i \in [n]$; if not, since entries for which $w_i = 0$ do not contribute to the linear predictor, it suffices to consider the entries of $\mathbf{d}$ for which $w_i \neq 0$.

2.1 Differential Privacy and Distortion

Similar to [GR11], we use the following generalized definition of differential privacy:

Definition 1. (Differential Privacy). *A (randomized) function $f : \mathrm{I}^n \to \mathbb{R}^m$ is $(\epsilon_1, \ldots, \epsilon_n)$-differentially private if for each individual $i \in [n]$ and for any pair of data vectors $\mathbf{d}, \mathbf{d}^{(i)} \in \mathrm{I}^n$ differing in only their i-th entry, ϵ_i is the smallest value such that $\mathbb{P}[f(\mathbf{d}) \in S] \leq e^{\epsilon_i}\mathbb{P}[f(\mathbf{d}^{(i)}) \in S]$ for all $S \subset \mathbb{R}^m$.*

This definition differs slightly from the usual definition of ϵ-differential privacy [DMNS06, Dwo06], as the latter is stated in terms of the *worst case* privacy across all individuals. More specifically, according to the notation in [DMNS06, Dwo06], an $(\epsilon_1, \ldots, \epsilon_n)$-differentially private function is ϵ-differentially private, where $\epsilon = \max_i \epsilon_i$.

Given a deterministic function f, a well-known method to provide ϵ-differential privacy is to add random noise drawn from a Laplace distribution to this function [DMNS06, Dwo06]. This readily extends to $(\epsilon_1, \ldots, \epsilon_n)$-differential privacy.

Lemma 1 ([DMNS06, Dwo06]). *Consider a deterministic function $f : \mathrm{I}^n \to \mathbb{R}$. Define $\hat{f}(\mathbf{d}) := f(\mathbf{d}) + Lap(\sigma)$, where $Lap(\sigma)$ is a random variable sampled from the Laplace distribution with parameter σ. Then, $\hat{f}$ is $(\epsilon_1, \ldots, \epsilon_n)$-differentially private, where $\epsilon_i = S_i(f)/\sigma$, and $S_i(f) := \max_{\mathbf{d}, \mathbf{d}^{(i)} \in \mathrm{I}^n} |f(\mathbf{d}) - f(\mathbf{d}^{(i)})|$, is the sensitivity of f to the i-th entry d_i, $i \in [n]$.*

Intuitively, the higher the variance σ of the Laplace noise added to f, the smaller ϵ_i, and hence, the better the privacy guarantee of $\hat{f}$. Moreover, for a fixed σ, entries i with higher sensitivity $S_i(f)$ receive a worse privacy guarantee (higher ϵ_i).

There is a natural tradeoff between the amount of noise added and the accuracy of the perturbed function $\hat{f}$. To capture this, we introduce the notion of *distortion* between two (possibly randomized) functions:

Definition 2. (Distortion). *Given two functions $f : \mathrm{I}^n \to \mathbb{R}$ and $\hat{f} : \mathrm{I}^n \to \mathbb{R}$, the* distortion, *$\delta(f, \hat{f})$, between f and $\hat{f}$ is given by*

$$\delta(f, \hat{f}) := \max_{\mathbf{d} \in \mathrm{I}^n} \mathbb{E}\left[|f(\mathbf{d}) - \hat{f}(\mathbf{d})|^2 \right].$$

In our setup, the data analyst wishes to disclose an *estimator function* $\hat{s} : \mathrm{I}^n \to \mathbb{R}$ of the linear predictor s. Intuitively, a good estimator $\hat{s}$ should have a small distortion $\delta(s, \hat{s})$, while also providing good differential privacy guarantees.

2.2 Privacy Auction Mechanisms

Each individual $i \in [n]$ has an associated cost function $c_i : \mathbb{R}_+ \to \mathbb{R}_+$, which determines the cost $c_i(\epsilon_i)$ incurred by i when an $(\epsilon_1, \ldots, \epsilon_n)$-differentially private estimate $\hat{s}$ is released by the analyst. As in [GR11], we consider linear cost functions, *i.e.*, $c_i(\epsilon) = v_i \epsilon$, for all $i \in [n]$. We refer to v_i as the *unit-cost* of individual i. The unit-costs v_i are not *a priori* known to the data analyst. Without loss of generality, we assume throughout the paper that $v_1 \leq \ldots \leq v_n$.

Given a weight vector $\mathbf{w} = (w_1, \ldots, w_n) \in \mathbb{R}^n$, let M_s be a mechanism compensating individuals in $[n]$ for their loss of privacy from the release of an estimate $\hat{s}$ of the linear predictor $s(\mathbf{d})$. Formally, M_s takes as input a vector of reported unit-costs $\mathbf{v} = (v_1, \ldots, v_n) \in \mathbb{R}_+^n$ and a budget B, and outputs

1. a payment $p_i \in \mathbb{R}_+$ for every $i \in [n]$, and
2. an estimator function $\hat{s} : I^n \to \mathbb{R}_+$.

Assume that the estimator $\hat{s}$ satisfies $(\epsilon_1, \ldots, \epsilon_n)$-differential privacy. A mechanism is *budget feasible* if $\sum_{i \in [n]} p_i \leq B$, *i.e.*, the payments made by the mechanism are within the budget B. Moreover, a mechanism is *individually rational* if for all $i \in [n]$, $p_i \geq c_i(\epsilon_i) = v_i \epsilon_i$, *i.e.*, payments made by the mechanism exceed the cost incurred by individuals. Finally, a mechanism is *truthful* if for all $i \in [n]$, $p_i(v_i, v_{-i}) - v_i \epsilon_i(v_i, v_{-i}) \geq p_i(v_i', v_{-i}) - v_i \epsilon_i(v_i', v_{-i})$, *i.e.*, no individual can improve her utility by misreporting her private unit-cost.

Ghosh and Roth motivate the linear dependence of costs c_i on the privacy guarantees ϵ_i by linking it directly to expected utility loss due to data disclosure [GR11]. Nevertheless, modeling the privacy cost of an individual as a linear function of her privacy guarantee is not the only possible choice, and has been criticized Chen *et al.* [CCK+13], and Nissim *et al.* [NOS12]. For example, when data across multiple individuals are correlated, the cost function can depend on the privacy guarantees provided to other individuals. Moreover, ϵ measures the worst-case

effect on privacy, while the typical effect on an agent might be lower than the above choice. On these grounds, the above works argue that it is more appropriate to treat such linear functions as upper bounds on the costs incurred by privacy loss, which however further complicates the design of incentive compatible mechanisms.

2.3 Outline of Our Approach

We denote by $\delta_{M_s} := \delta(s, \hat{s})$ the distortion between s and the function output by the mechanism M_s. Ideally, a mechanism should output an estimator that has small distortion. However, the smaller the distortion, the higher the privacy violation and, hence, the more money the mechanism needs to spend. As such, the objective of this paper is to design a mechanism with minimal distortion, subject to the constraints of truthfulness, individual rationality, and budget feasibility.

To address this question, in Section 3, we first establish a privacy-distortion tradeoff for differentially-private estimators of the linear predictor. We then introduce a family of estimators, Discrete Canonical Laplace Estimator Functions (DCLEFs), and show that they achieve a near-optimal privacy-distortion tradeoff. This result allows us to limit our attention to DCLEF privacy auction mechanisms, *i.e.*, mechanisms that output a DCLEF $\hat{s}$. In Section 4, we present a mechanism that is truthful, individually rational, and budget feasible, while also being near-optimal in terms of distortion.

The above approach mirrors the approach followed by Ghosh and Roth [GR11] in the case where the statistic s is the sum of bits. Ghosh and Roth also establish a privacy vs. accuracy tradeoff among differentially private estimators for such statistics. Furthermore, they show that a particular class of estimators achieve a close-to-optimal privacy vs. accuracy tradeoff, and subsequently focus on the design of mechanisms outputing estimators from this class. Nevertheless, the tradeoff and class of estimators developed by Ghosh and Roth cannot be readily extended to privacy auctions for general linear predictors. The asymmetry of such functions requires the introduction of the more general class of DCLEFs. Finally, in contrast to the multi-unit procurement auction of [GR11], DCLEF privacy auction mechanisms are similar to the knapsack mechanism of [Sin10], which makes distortion minimization a more challenging combinatorial task.

3 Privacy-Distortion Tradeoff and Laplace Estimators

Recall that a good estimator should exhibit low distortion and simultaneously give good privacy guarantees. In this section, we establish the privacy-distortion tradeoff for differentially-private estimators of the linear predictor. Moreover, we introduce a family of estimators that exhibits a near-optimal tradeoff between privacy and distortion. This will motivate our focus on privacy auction mechanisms that output estimators from this class in Section 4.

3.1 Privacy-Distortion Tradeoff

There exists a natural tension between privacy and distortion, as highlighted by the following two examples.

> **Example 1.** Consider the estimator $\hat{s} := \bar{R} \sum_{i=1}^{n} w_i$, where recall that $\bar{R} = (R_{\min} + R_{\max})/2$. This estimator guarantees perfect privacy (*i.e.*, $\epsilon_i = 0$), for all individuals. However, $\delta(s, \hat{s}) = (W\Delta)^2/4$.

Example 2. Consider the estimator function $\hat{s} := \sum_{i=1}^{n} w_i d_i$. In this case, $\delta(s, \hat{s}) = 0$. However, $\epsilon_i = \infty$ for all $i \in [n]$.

In order to formalize this tension between privacy and distortion, we define the *privacy index* of an estimator as follows.

Definition 3. *Let $\hat{s} : \mathrm{I}^n \to \mathbb{R}$ be any $(\epsilon_1, \ldots, \epsilon_n)$-differentially private estimator function for the linear predictor. We define the* privacy index*, $\beta(\hat{s})$, of $\hat{s}$ as*

$$\beta(\hat{s}) := \max \left\{ w(H) : H \subseteq [n] \text{ and } \sum_{i \in H} \epsilon_i < 1/2 \right\}. \tag{2}$$

The index $\beta(\hat{s})$ captures the weight of the individuals that have been guaranteed good privacy by $\hat{s}$. Next we characterize the impossibility of having an estimator with a low distortion but a high privacy index. Note that for Example 1, $\beta(\hat{s}) = W$, *i.e.*, the largest value possible, while for Example 2, $\beta(\hat{s}) = 0$. We stress that the selection of $1/2$ as an upper bound in (2) is arbitrary; Theorems 1 and 2 still hold if another value is used, though the constants involved will differ.

Our first main result, which is proved in Appendix A, establishes a trade-off between the privacy index and the distortion of an estimator.

Theorem 1 (Trade-off between Privacy-index and Distortion)**.** *Let $0 < \alpha < 1$. Let $\hat{s} : \mathrm{I}^n \to \mathbb{R}$ be an arbitrary estimator function for the linear predictor. If $\delta(s, \hat{s}) \leq (\alpha W \Delta)^2/48$ then $\beta(\hat{s}) \leq 2\alpha W$.*

In other words, if an estimator has low distortion, the weight of individuals with a good privacy guarantee (*i.e.*, a small ϵ_i) can be at most an α fraction of $2W$.

3.2 Laplace Estimator Functions

Consider the following family of estimators for the linear predictor $\hat{s} : \mathrm{I}^n \to \mathbb{R}$:

$$\hat{s}(\mathbf{d}; \mathbf{a}, \mathbf{x}, \sigma) := \sum_{i=1}^{n} w_i d_i x_i + \sum_{i=1}^{n} w_i a_i (1 - x_i) + \mathrm{Lap}(\sigma) \tag{3}$$

where $x_i \in [0, 1]$, and each $a_i \in \mathbb{R}$ is a constant independent of the data vector $\mathbf{d}$. This function family is parameterized by $\mathbf{x}, \mathbf{a}$ and σ, and is a generalization of the Laplace estimators considered by Ghosh and Roth [GR11]. The estimator $\hat{s}$ results from distorting s in two ways: (a) a randomized distortion by the addition of the Laplace noise, and (b) a deterministic distortion through a linear interpolation between each entry d_i and some constant a_i. Intuitively, the interpolation parameter x_i determines the extent to which the estimate $\hat{s}$ depends on entry d_i. Using Lemma 1 and the definition of distortion, it is easy to characterize the privacy and distortion properties of such estimators.

Lemma 2. *Given w_i, $i \in [n]$, let $s(\mathbf{d})$ be the linear predictor given by (1), and $\hat{s}$ an estimator of s given by (3). Then,*

1. *$\hat{s}$ is $(\epsilon_1, \ldots, \epsilon_n)$-differentially private, where $\epsilon_i = \frac{\Delta |w_i| \, x_i}{\sigma}$, $i \in [n]$.*
2. *The distortion satisfies $\delta(s, \hat{s}) \geq \left(\frac{\Delta}{2} \sum_{i=1}^{n} |w_i|(1 - x_i) \right)^2 + 2\sigma^2$, with equality attained when $a_i = \bar{R}$, for all $i \in [n]$.*

The proof of this lemma can be found in Appendix B. Recall that $\hat{s}$ interpolates between the data d_i and a fixed value a_i. The first statement of Lemma 2 implies that the constants a_i do not affect the differential privacy properties of $\hat{s}$. The second statement implies that, among all estimators with given $\mathbf{x}$, the distortion $\delta(s, \hat{s})$ is minimized when $a_i = \bar{R} = \frac{1}{2}(R_{\min} + R_{\max})$ for all $i \in [n]$. Together, these statements imply that it is always preferable to set all a_i values to $\bar{R}$: this selection has the minimum distortion among all estimators of the form (3) under the same privacy guarantees. This motivates us to define the family of Laplace estimator functions as follows.

Definition 4. *Given w_i, $i \in [n]$, the* Laplace estimator function family *(LEF) for the linear predictor s is the set of functions $\hat{s} : \mathrm{I}^n \to \mathbb{R}$, parameterized by $\mathbf{x}$ and σ, such that*

$$\hat{s}(\mathbf{d}; \mathbf{x}, \sigma) = \sum_{i=1}^{n} w_i d_i x_i + \bar{R} \sum_{i=1}^{n} w_i (1 - x_i) + Lap(\sigma) \tag{4}$$

We call a LEF *discrete* if $x_i \in \{0, 1\}$. Furthermore, we call a LEF *canonical* if the Laplace noise added to the estimator has a parameter of the form

$$\sigma = \sigma(\mathbf{x}) := \Delta \sum_{i=1}^{n} |w_i|(1 - x_i) \tag{5}$$

Recall that x_i controls the dependence of $\hat{s}$ on the entry d_i; thus, intuitively, the standard deviation of the noise added in a canonical Laplace estimator is proportional to the "residual weight" of data entries. Note that, by Lemma 2, the distortion of a canonical Laplace estimator $\hat{s}$ has the following simple form:

$$\delta(s, \hat{s}) = \frac{9}{4}\Delta^2 \big(\sum_{i=1}^{n} |w_i|(1 - x_i) \big)^2 = \frac{9}{4}\Delta^2 \big(W - \sum_{i=1}^{n} |w_i| x_i \big)^2. \tag{6}$$

Our next result establishes that there exists a discrete canonical Laplace estimator function (DCLEF) with a small distortion and a high privacy index.

Theorem 2 (DCLEFs suffice)**.** *Let $0 < \alpha < 1$. Let*

$$\hat{s}^* := \operatorname*{argmax}_{\hat{s}: \delta(s,\hat{s}) \le (\alpha W \Delta)^2/48} \beta(\hat{s})$$

be an estimator with the highest privacy index among all $\hat{s}$ for which $\delta(s, \hat{s}) \le (\alpha W \Delta)^2/48$. There exists a DCLEF $\hat{s}^\circ : \mathrm{I}^n \to \mathbb{R}$ such that $\delta(s, \hat{s}^\circ) \le (9/4)(\alpha W \Delta)^2$, and $\beta(\hat{s}^\circ) \ge \frac{1}{2}\beta(\hat{s}^*)$.

In other words, there exists a DCLEF that is within a constant factor, in terms of both its distortion and its privacy index, from an optimal estimator $\hat{s}^*$. Theorem 2 is proved in Appendix C and has the following immediate corollary:

Corollary 1. *Consider an arbitrary estimator $\hat{s}$ with distortion $\delta(s, \hat{s}) < (W\Delta)^2/48$. Then, there exists a DCLEF $\hat{s}^\circ$ such that $\delta(s, \hat{s}^\circ) \le 108\delta(s, \hat{s})$ and $\beta(\hat{s}^\circ) \ge \frac{1}{2}\beta(\hat{s})$.*

Proof. Apply Theorem (2) with $\alpha = \sqrt{48\delta(s, \hat{s})}/(W\Delta)$. In particular, for this α and $\hat{s}$ as in the theorem statement, we have that $\hat{s}^* := \operatorname{argmax}_{\hat{s}': \delta(s,\hat{s}') \le \delta(s,\hat{s})} \beta(\hat{s}')$, hence $\beta(\hat{s}^*) \ge \beta(\hat{s})$. Therefore, there exists a DCLEF $\hat{s}^\circ$ such that $\delta(s, \hat{s}^\circ) \le (9/4)(\alpha W \Delta)^2 \le 108\delta(s, \hat{s})$, and $\beta(\hat{s}^\circ) \ge \frac{1}{2}\beta(\hat{s}^*) \ge \frac{1}{2}\beta(\hat{s})$. □

Theorems 1 and 2 imply that, when searching for estimators with low distortion and high privacy index, it suffices (up to constant factors) to focus on DCLEFs. Similar results were derived in [GR11] for estimators of unweighted sums of bits.

4 Privacy Auction Mechanism

Motivated by Theorems 1 and 2, we design a truthful, individually rational, budget-feasible DCLEF mechanism (*i.e.*, a mechanism that outputs a DCLEF) and show that it is 5-approximate in terms of accuracy compared with the optimal, individually rational, budget-feasible DCLEF mechanism. Note that a DCLEF is fully determined by the vector $\mathbf{x} \in \{0,1\}^n$. Therefore, we will simply refer to the output of the DCLEF mechanisms described below as $(\mathbf{x}, \mathbf{p})$, as the latter characterize the released estimator and the compensations to individuals.

4.1 An Optimal DCLEF Mechanism

Consider the problem of designing a DCLEF mechanism M that is individually rational and budget feasible (but not necessarily truthful), and minimizes δ_M. Given a DCLEF $\hat{s}$, define $H(\hat{s}) := \{i : x_i = 1\}$ to be the set of individuals that receive non-zero differential privacy guarantees. Eq. (6) implies that $\delta(s, \hat{s}) = \frac{9}{4}\Delta^2(W - w(H(\hat{s})))^2$. Thus, minimizing $\delta(s, \hat{s})$ is equivalent to maximizing $w(H(\hat{s}))$. Let $(\mathbf{x}_{opt}, \mathbf{p}_{opt})$ be an optimal solution to the following problem:

$$\begin{aligned} \text{maximize} \quad & S(\mathbf{x}; \mathbf{w}) = \sum_{i=1}^{n} |w_i| x_i \\ \text{subject to:} \quad & p_i \geq v_i \epsilon_i(\mathbf{x}), \quad \forall i \in [n], \quad \text{(individual rationality)} \\ & \sum_{i=1}^{n} p_i \leq B \qquad \text{(budget feasibility)} \\ & x_i \in \{0,1\}, \ \forall i \in [n] \quad \text{(discrete estimator function)} \end{aligned} \tag{7}$$

where, by Lemma 2 and (5),

$$\epsilon_i(\mathbf{x}) = \frac{\Delta |w_i| x_i}{\sigma(\mathbf{x})} = \frac{|w_i| x_i}{\sum_i |w_i|(1 - x_i)} \text{ (canonical property).} \tag{8}$$

A mechanism M_{opt} that outputs $(\mathbf{x}_{opt}, \mathbf{p}_{opt})$ will be an optimal, individually rational, budget feasible (but not necessarily truthful) DCLEF mechanism. Let $OPT := S(\mathbf{x}_{opt}; \mathbf{w})$ be the optimal objective value of (7). We use OPT as the benchmark to which we compare the (truthful) mechanism we design below. Without loss of generality, we make the following assumption:

Assumption 1. *For all* $i \in [n]$, $|w_i| v_i / (W - |w_i|) \leq B$.

Observe that if an individual i violates this assumption, then $c_i(\epsilon_i(\mathbf{x})) > B$ for any $\mathbf{x}$ output by a DCLEF mechanism that sets $x_i = 1$. In other words, no DCLEF mechanism (including M_{opt}) can compensate this individual within the analyst's budget and, hence, will set $x_i = 0$. Therefore, it suffices to focus on the subset of individuals for whom the assumption holds.

4.2 A Truthful DCLEF Mechanism

To highlight the challenge behind designing a truthful DCLEF mechanism, observe that if the privacy guarantees were given by $\epsilon_i(\mathbf{x}) = x_i$ rather than (8), the optimization problem (7) would be identical to the budget-constrained mechanism design problem for knapsack studied by Singer [Sin10]. In the reverse-auction setting of [Sin10], an auctioneer purchases items valued

Algorithm 1 FairInnerProduct$(\mathbf{v}, \mathbf{w}, B)$

Let k be the largest integer such that $\frac{B}{w([k])} \geq \frac{v_k}{W - w([k])}$.
Let $i^* := \text{argmax}_{i \in [n]} |w_i|$.
Let $\hat{p}$ be as defined in (9).
if $|w_{i^*}| > \sum_{i \in [k] \setminus \{i^*\}} |w_i|$ **then**
  Set $O = \{i^*\}$.
  Set $p_{i^*} = \hat{p}$ and $p_i = 0$ for all $i \neq i^*$.
else
  Set $O = [k]$.
  Pay each $i \in O,\ p_i = |w_i| \min\{\frac{B}{w([k])}, \frac{v_{k+1}}{W - w([k])}\}$, and for $i \notin O,\ p_i = 0$.
end if
Set $x_i = 1$ if $i \in O$ and $x_i = 0$ otherwise.

at fixed costs v_i by the individuals that sell them. Each item i is worth $|w_i|$ to the auctioneer, while the auctioneer's budget is B. The goal of the auctioneer is to maximize the total worth of the purchased set of items, *i.e.*, $S(\mathbf{x}; \mathbf{w})$. Singer presents a truthful mechanism that is 6-approximate with respect to OPT. However, in our setting, the privacy guarantees $\epsilon_i(\mathbf{x})$ given by (8) introduce *externalities* into the auction. In contrast to [Sin10], the ϵ_i's couple the cost incurred by an individual i to the weight of other individuals that are compensated by the auction, making the mechanism design problem harder. This difficulty is overcome by our mechanism, which we call FairInnerProduct, described in Algorithm 1.

The mechanism takes as input the budget B, the weight vector $\mathbf{w}$, and the vector of unit-costs $\mathbf{v}$, and outputs a set $O \subset [n]$, that receive $x_i = 1$ in the DCLEF, as well as a set of payments for each individual in O. Our construction uses a greedy approach similar to the Knapsack mechanism in [Sin10]. In particular, it identifies users that are the "cheapest" to purchase. To ensure truthfulness, it compensates them within budget based on the unit-cost of the last individual that was not included in the set of compensated users. As in greedy solutions to knapsack, this construction does not necessarily yield a constant approximation w.r.t. OPT; for that, the mechanism needs to sometimes compensate only the user with the highest absolute weight $|w_i|$. In such cases, the payment of the user of the highest weight is selected so that she has no incentive to lie about here true unit cost.

Recall that $v_1 \leq \ldots \leq v_n$. The mechanism defines $i^* := \text{argmax}_{i \in [n]} |w_i|$ as the individual with the largest $|w_i|$, and k as the largest integer such that $\frac{B}{w([k])} \geq \frac{v_k}{W - w([k])}$. Subsequently, the mechanism either sets $x_i = 1$ for the first k individuals, or, if $|w_{i^*}| > \sum_{i \in [k] \setminus \{i^*\}} |w_i|$, sets $x_{i^*} = 1$. In the former case, individuals $i \in [k]$ are compensated *in proportion to their absolute weights* $|w_i|$. If, on the other hand, only $x_{i^*} = 1$, the individual i^* receives a payment $\hat{p}$ defined as follows: Let

$$S_{-i^*} := \Big\{ t \in [n] \setminus \{i^*\} : \frac{B}{\sum_{i \in [t] \setminus \{i^*\}} |w_i|} \geq \frac{v_t}{W - \sum_{i \in [t] \setminus \{i^*\}} |w_i|} \text{ and} \sum_{i \in [t] \setminus \{i^*\}} |w_i| \geq |w_{i^*}| \Big\}.$$

If $S_{-i^*} \neq \emptyset$, then let $r := \min\{i : i \in S_{-i^*}\}$. Define

$$\hat{p} := \begin{cases} B, & \text{if } S_{-i^*} = \emptyset \\ \frac{|w_{i^*}| v_r}{W - |w_{i^*}|}, & \text{otherwise} \end{cases} \tag{9}$$

The next theorem states that FairInnerProduct has the properties we desire.

Theorem 3. *FairInnerProduct is truthful, individually rational and budget feasible. It is 5-approximate with respect to OPT. Further, it is 2-approximate when all weights are equal.*

The theorem is proved in Appendix D. We note that the truthfulness of the knapsack mechanism in [Sin10] is established via Myerson's characterization of truthful single-parameter auctions (*i.e.*, by showing that the allocation is monotone and the payments are threshold). In contrast, because of the coupling of costs induced by the Laplace noise in DCLEFs, we are unable to use Myerson's characterization and, instead, give a direct argument about truthfulness.

We prove a 5-approximation by using the optimal solution of the fractional relaxation of (7). This technique can also be used to show that the knapsack mechanism in [Sin10] is 5-approximate instead of 6-approximate. FairInnerProduct generalizes the Ghosh-Roth mechanism; in the special case when all weights are equal FairInnerProduct reduces to the Ghosh-Roth mechanism, which, by Theorem 3, is 2-approximate with respect to OPT. In fact, our next theorem, proved in Appendix E, states that the approximation ratio of a truthful mechanism is lower-bounded by 2.

Theorem 4 (Impossibility of Approximation)**.** *For all $\varepsilon > 0$, there is no truthful, individually rational, budget feasible DCLEF mechanism that is also $2-\varepsilon$-approximate with respect to OPT.*

Our benchmark OPT is stricter than that used in [GR11]. In particular, Ghosh and Roth show that their mechanism is optimal among all truthful, individually rational, budget-feasible, and *envy-free* mechanisms. In fact, the example we use to show hardness of approximation is a uniform weight example, implying that the lower-bound also holds for uniform weight case. Indeed, the mechanism in [GR11] is 2-approximate with respect to OPT, although it is optimal among individually rational, budget feasible mechanisms that are also truthful *and* envy free.

5 Discussion on Linear Predictors

As discussed in the introduction, a statistic $s(\mathbf{d})$ of the form (1) can be viewed as a *linear predictor* and is thus of particular interest in the context of recommender systems. We elaborate on this interpretation in this section.

Recall that each individual $i \in [n] = \{1, \ldots, n\}$ is endowed with a public vector $\mathbf{y}_i \in \mathbb{R}^m$, which includes m publicly known features about this individual, and let $\mathbf{Y} = [\mathbf{y}_i]_{i\in[n]} \in \mathbb{R}^{n\times m}$ be a matrix comprising the public feature vectors. Consider again a new individual, not belonging to the database, whose public feature profile is $\mathbf{y} \in \mathbb{R}^m$. Beyond, k-NN, there are several predictors that take the form (1), for weights $\mathbf{w} = \mathbf{w}(\mathbf{y}, \mathbf{Y})$:

- *Nadaranya-Watson Weighted Average.* In contrast to k-NN, the Nadaranya-Watson weighted average leverages all data in the database, weighing more highly data closer to $\mathbf{y}$. The general form of the prediction is $s(\mathbf{d}) = \sum_{i=1}^n K(\mathbf{y}, \mathbf{y}_i) d_i / \sum_{i'=1}^n K(\mathbf{y}, \mathbf{y}_{i'})$ where the *kernel* $K : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}_+$ is a function decreasing in the distance between its argument (*e.g.*, $K(\mathbf{y}, \mathbf{y}') = e^{-\|\mathbf{y}-\mathbf{y}'\|_2^2}$).

- *Ridge Regression.* In ridge regression, the analyst first fits a linear model to the data, *i.e.*, solves the optimization problem

$$\min_{\mathbf{b}\in\mathbb{R}^m} \sum_{i=1}^n \left(d_i - \langle \mathbf{y}_i, \mathbf{b}\rangle\right)^2 + \lambda\|\mathbf{b}\|_2^2, \tag{10}$$

where $\lambda \geq 0$ is a regularization parameter, enforcing that the vector $\mathbf{b}$ takes small values. The prediction is then given by the inner product $\langle \mathbf{y}, \mathbf{b} \rangle$. The solution to (10) is given by $\mathbf{b} = (\mathbf{Y}^T\mathbf{Y} + \lambda \mathbf{I})^{-1}\mathbf{Y}^T\mathbf{d}$; as such, the predicted value for a new user with feature vector $\mathbf{y}$ is given by $s(\mathbf{d}) = \langle \mathbf{y}, \mathbf{b} \rangle = \mathbf{y}^T(\mathbf{Y}^T\mathbf{Y} + \lambda \mathbf{I})^{-1}\mathbf{Y}^T\mathbf{d}$.

- *Support Vector Machines.* A more general regression model assumes that the private values d_i can be expressed in terms of the public vectors $\mathbf{y}_i$ as a linear combination of a set of basis functions $h_\ell : \mathbb{R}^m \to \mathbb{R}$, $\ell = 1, \ldots, L$, *i.e.*, the analyst first solves the optimization problem

$$\min_{\mathbf{b} \in \mathbb{R}^L} \sum_{i=1}^n \left(d_i - \sum_{\ell=1}^L b_\ell h_\ell(\mathbf{y}_i) \right)^2 + \lambda \|\mathbf{b}\|_2^2 \tag{11}$$

  For $\mathbf{y}, \mathbf{y}' \in \mathbb{R}^m$, denote by $K(\mathbf{y}, \mathbf{y}') = \sum_{\ell=1}^L h_\ell(\mathbf{y}) h_\ell(\mathbf{y})$ the kernel of the space spanned by the basis functions. Let $\mathbf{K}(\mathbf{Y}) = [K(\mathbf{y}_i, \mathbf{y}_j)]_{i,j \in [n]} \in \mathbb{R}^{n \times n}$ be the $n \times n$ matrix comprising the kernel values evaluated at each pair of feature vectors in the database, and $\mathbf{k}(\mathbf{y}, \mathbf{Y}) = [K(\mathbf{y}, \mathbf{y}_i)]_{i \in [n]} \in \mathbb{R}^n$ the kernel values w.r.t. the new user. The solution to (11) yields a predicted value for the new individual of the form: $s(\mathbf{d}) = (\mathbf{k}(\mathbf{y}, \mathbf{Y}))^T(\mathbf{K}(Y) + \lambda \mathbf{I})^{-1}\mathbf{d}$.

In all four examples, including k-NN, the prediction $s(\mathbf{d})$ is indeed of the form (1). Note that the weights are non-negative in our first two examples, but may assume negative values in the latter two.

6 Conclusion and Future Work

We considered the setting of an auction, where a data analyst wishes to buy, from a set of n individuals, the right to use their private data $d_i \in \mathbb{R}$, $i \in [n]$, in order to *cheaply* obtain an *accurate* estimate of a statistic. Motivated by recommender systems and, more generally, prediction problems, the statistic we consider is a linear predictor with publicly known weights. The statistic can be viewed as a prediction of the unknown data of a new individual based on the database entries. We formalized the trade-off between privacy and accuracy in this setting; we showed that obtaining an accurate estimate necessitates giving poor differential privacy guarantees to individuals whose cumulative weight is large. We showed that DCLEF estimators achieve an order-optimal trade-off between privacy and accuracy, and, consequently, it suffices to focus on DCLEF mechanisms. We use this observation to design a truthful, individually rational, budget feasible mechanism under the constraint that the analyst has a fixed budget. Our mechanism can be viewed as a proportional-purchase mechanism, *i.e.*, the privacy ϵ_i guaranteed by the mechanism to individual i is proportional to her weight $|w_i|$. We show that our mechanism is 5-approximate in terms of accuracy compared to an optimal (possibly non-truthful) mechanism, and that no truthful mechanism can achieve a $2 - \varepsilon$ approximation, for any $\varepsilon > 0$.

Our formalization of a tradeoff between privacy and accuracy builds upon the work of Ghosh and Roth, and led to the introduction of the distortion of an estimator, as well as its privacy index. As a worst-case expected variance, our definition of distortion naturally generalizes to estimators of non-linear statistics. The privacy index is on the other hand the sum of weights among individuals receiving a sufficiently good privacy guarantee. Intuitively, such weights capture the effect that each individual's data has in the evaluation of the linear statistic. A natural question to ask if whether this definition of a privacy index, as well as the corresponding tradeoff between privacy and distortion, can be extended to a more general class of statistics. For

example, it would be interesting to see if such tradeoffs can be established, *e.g.*, for a privacy index in which weights are replaced by worst or average case marginal contributions to the statistic.

Finally, it is natural to ask if our present results apply under weaker assumptions on how costs are associated with differential privacy guarantees. In light of recent results [CCK+13, NOS12], a possible starting point for such an investigation is considering costs that are bounded by–rather than equal to–functions that depend linearly on the privacy guarantees.

References

[CCK+13] Yiling Chen, Stephen Chong, Ian A. Kash, Tal Moran, and Salil P. Vadhan. Truthful mechanisms for agents that value privacy. In *Proceedings of the 14th ACM Conference on Electronic Commerce*, EC '13, 2013.

[DMNS06] Cynthia Dwork, Frank Mcsherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *Proc. Theory of Cryptography Conference*, 2006.

[Dwo06] Cynthia Dwork. Differential privacy. In *Proc. ICALP*, pages 1–12, 2006.

[FL12] Lisa Fleischer and Yu-Han Lyu. Approximately optimal auctions for selling privacy when costs are correlated with data. *CoRR*, abs/1204.4031, 2012.

[GR11] Arpita Ghosh and Aaron Roth. Selling privacy at auction. In *Proc. ACM EC*, pages 199–208, 2011.

[HTF09] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. Springer, 2nd edition, 2009.

[JKH+09] Joseph Joseph, Jennifer King, Chris Jay Hoofnagle, Amy Bleakley, and Michael Hennessy. Americans reject tailored advertising and three activities that enable it, 2009. http://ssrn.com/abstract=1478214.

[KKMN09] Aleksandra Korolova, Krishnaram Kenthapadi, Nina Mishra, and Alexandros Ntoulas. Releasing search queries and clicks privately. In *WWW*, 2009.

[LR12] Katrina Ligett and Aaron Roth. Take it or Leave it: Running a Survey when Privacy Comes at a Cost. In *Proceedings of the 8th Workshop on Internet and Network Economics*, WINE '12, page To appear, 2012.

[LSY03] G. Linden, B. Smith, and J. York. Amazon.com recommendations: item-to-item collaborative filtering. *Internet Computing, IEEE*, 7(1):76 – 80, jan/feb 2003.

[Mel12] John P. Mello. Facebook hit with lawsuit alleging privacy wrongs. *PCWorld*, May 18 2012.

[MM09] Frank McSherry and Ilya Mironov. Differentially private recommender systems: building privacy into the net. In *Proc. ACM KDD*, pages 627–636, 2009.

[MT90] S. Martello and P. Toth. *Knapsack problems: algorithms and computer implementations*. Wiley-Interscience series in discrete mathematics and optimization. 1990.

[MT07] Frank McSherry and Kunal Talwar. Mechanism design via differential privacy. In *Proc. FOCS*, 2007.

[Net] Netflix Privacy Litigation. www.videoprivacyclass.com.

[NOS12] Kobbi Nissim, Claudio Orlandi, and Rann Smorodinsky. Privacy-Aware Mechanism Design. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, EC '12, 2012.

[NS08] Arvind Narayanan and Vitaly Shmatikov. Robust De-anonymization of Large Sparse Datasets. In *IEEE Symposium on Security and Privacy*, pages 111–125, 2008.

[NST10] Kobbi Nissim, Rann Smorodinsky, and Moshe Tennenholtz. Approximately optimal mechanism design via differential privacy. 2010.

[Rib12] John Ribeiro. Google faces class-action lawsuits over new privacy policy. *PCWorld*, Mar 22 2012.

[RS12] Aaron Roth and Grant Schoenebeck. Conducting truthful surveys, cheaply. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, EC '12, pages 826–843, New York, NY, USA, 2012. ACM.

[Sin10] Yaron Singer. Budget feasible mechanisms. In *Proc. FOCS*, 2010.

[SKKR01] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on World Wide Web*, WWW '01, pages 285–295, New York, NY, USA, 2001. ACM.

[Xia11] David Xiao. Is privacy compatible with truthfulness? Technical Report 2011/005, Cryptology ePrint Archive, 2011.

A Proof of Theorem 1 (Trade-off between Privacy-index and Distortion)

By Definition 3, the privacy index $\beta(\hat{s})$ for an estimator $\hat{s}$ is the optimal objective value of the following optimization problem: maximize $\sum_{i=1}^{n} |w_i| x_i$ where $\sum_{i=1}^{n} \epsilon_i x_i < \frac{1}{2}$ and for all $i \in [n]$, $x_i \in \{0,1\}$.

Interpreting $|w_i|$ as the value, and ϵ_i as the size of object i, the above problem can be viewed as a 0/1 knapsack problem where the size of the knapsack is 1/2. Assume for this proof, without loss of generality, that $\frac{\epsilon_1}{|w_1|} \leq \ldots \leq \frac{\epsilon_n}{|w_n|}$. We define some notation that is needed in the proof. Let $h(\hat{s}) := \max\left\{j \in [n] : \frac{\epsilon_j}{|w_j|} < \frac{1}{2w([j])}\right\}$ if $\frac{\epsilon_1}{|w_1|} < \frac{1}{2|w_1|}$ and $h(\hat{s}) := 0$ otherwise. Observe that $0 \leq h(\hat{s}) \leq n$. Next, define

$$\hat{i} := \operatorname*{argmax}_{i \in [n]: \epsilon_i < 1/2} |w_i|, \quad \text{and} \quad H(\hat{s}) := \begin{cases} [h(\hat{s})], & \text{if } w([h(\hat{s})]) \geq |w_{\hat{i}}|, \\ \{\hat{i}\}, & \text{otherwise.} \end{cases} \tag{12}$$

The following then holds.

Lemma 3. $2w(H(\hat{s})) \geq \beta(\hat{s})$.

Proof. $H(\hat{s})$ is a 2-approximate greedy solution to the 0/1 knapsack problem given by [MT90, Section 2.4]. We include a proof for completeness. Observe that $\sum_{i=1}^{h(\hat{s})} \epsilon_i < \frac{1}{2}$, and $\sum_{i=1}^{h(\hat{s})+1} \epsilon_i \geq \frac{1}{2}$. Also, since the "objects" are arranged in increasing order of size per value, it follows that $w([h(\hat{s})]) + w_{h(\hat{s})+1} \geq OPT$. Therefore,

$$H(\hat{s}) = \max\{w([h(\hat{s})]), w_{\hat{i}}\} \geq \frac{w([h(\hat{s})]) + w_{\hat{i}}}{2} \geq \frac{w([h(\hat{s})]) + w_{h(\hat{s})+1}}{2} \geq \frac{OPT}{2} \qquad \square$$

Now we are ready to prove that if the distortion $\delta(s, \hat{s})$ is small, then $w(H(\hat{s}))$ is also small, which, together with Lemma 3, proves the theorem. In our proof, we make use of the notion of k-accuracy defined in [GR11, Definition 2.6]. For $\hat{s} : \mathrm{I}^n \to \mathbb{R}$, let

$$k_{\hat{s}} := \min\left\{k \in \mathbb{R}_+ : \forall \mathbf{d} \in \mathrm{I}^n, \mathbb{P}[|s(\mathbf{d}) - \hat{s}(\mathbf{d})| \geq k] \leq \frac{1}{3}\right\} \tag{13}$$

Lemma 4. *Let $0 < \alpha < 1$. If $w(H(\hat{s})) > \alpha W$ then $k_{\hat{s}} > \alpha W \Delta/4$.*

Proof. Assume for the sake of contradiction that $w(H(\hat{s})) > \alpha W$ and $k_{\hat{s}} \leq \alpha W\Delta/4$. For a data vector $\mathbf{d}$, let $z = s(\mathbf{d}) = \sum_i w_i d_i$ and $\hat{z} = \hat{s}(\mathbf{d})$. Also, let $S := \{y \in \mathbb{R} : |y - z| < k_{\hat{s}}\}$. Then, by (13), $\mathbb{P}[\hat{z} \in S] \geq 2/3$.

The set $H(\hat{s})$ can be partitioned as follows: $H(\hat{s}) = H^+(\hat{s}) \cup H^-(\hat{s})$, with $H^+(\hat{s}) \cap H^-(\hat{s}) = \{\emptyset\}$, where the disjoint subsets $H^+(\hat{s})$ and $H^-(\hat{s})$ are defined by

$$\begin{aligned} H^+(\hat{s}) &= \{i \in [n] : d_i \leq \bar{R} \text{ and } w_i \leq 0\} \cup \{i \in [n] : d_i > \bar{R} \text{ and } w_i > 0\}, \\ H^-(\hat{s}) &= \{i \in [n] : d_i \leq \bar{R} \text{ and } w_i > 0\} \cup \{i \in [n] : d_i > \bar{R} \text{ and } w_i \leq 0\}. \end{aligned} \tag{14}$$

Then $w(H(\hat{s})) = w(H^+(\hat{s})) + w(H^-(\hat{s}))$. Thus, one of the subsets $H^+(\hat{s})$ and $H^-(\hat{s})$ must have a total weight greater or equal to $w(H(\hat{s}))/2$. Without loss of generality, assume that $w(H^+(\hat{s})) \geq w(H(\hat{s}))/2$.

Consider another data vector $\mathbf{d}'$ where $d'_i = d_i$ if $i \in [n] \setminus H^+(\hat{s})$, while if $i \in H^+(\hat{s})$,

$$d'_i = \begin{cases} d_i + \frac{\Delta}{2}, & \text{if } d_i \leq \bar{R} \text{ and } w_i \leq 0 \\ d_i - \frac{\Delta}{2}, & \text{if } d_i > \bar{R} \text{ and } w_i > 0 \end{cases} \tag{15}$$

Let $z' := s(\mathbf{d}') = \sum_{i=1}^n w_i d'_i$ and let $\hat{z}' = \hat{s}(\mathbf{d}')$. Also, let $S' := \{y \in \mathbb{R} : |y - z'| < k_{\hat{s}}\}$. From eq. (15), we have

$$|z - z'| = \Big|\sum_{i \in H^+(\hat{s})} w_i(d_i - d'_i)\Big| = \Big|\sum_{i \in H^+(\hat{s})} |w_i|\Delta/2\Big| = \frac{\Delta}{2} w(H^+(\hat{s})) \geq \frac{\Delta}{4} w(H(\hat{s})) > \alpha \frac{\Delta}{4} W. \tag{16}$$

Since $k_{\hat{s}} \leq \alpha W\Delta/4$, eq. (16) implies that S and S' are disjoint.

Since $\hat{s}$ is $(\epsilon_1, \ldots, \epsilon_n)$-differentially private, and $\mathbf{d}$ and $\mathbf{d}'$ differ in exactly the entries in $H^+(\hat{s})$, $\mathbb{P}[\hat{z}' \in S] \geq \exp\left(-\sum_{i \in H^+(\hat{s})} \epsilon_i\right) \mathbb{P}[\hat{z} \in S] \geq \exp\left(-\sum_{i \in H^+(\hat{s})} \epsilon_i\right) \frac{2}{3}$. Note that $\sum_{i \in [h(\hat{s})]} \epsilon_i < \sum_{i \in [h(\hat{s})]} \frac{|w_i|}{2w([h(\hat{s})])} = \frac{1}{2}$, and also $\epsilon_{\hat{i}} < 1/2$. Therefore, $\sum_{i \in H(\hat{s})} \epsilon_i < 1/2$. Since $H^+(\hat{s}) \subset H(\hat{s})$, we have $\sum_{i \in H^+(\hat{s})} \epsilon_i \leq \sum_{i \in H(\hat{s})} \epsilon_i < 1/2$.

This implies $\mathbb{P}[\hat{z}' \in S] \geq \exp\left(-\sum_{i \in H^+(\hat{s})} \epsilon_i\right) \frac{2}{3} > \exp\left(-\frac{1}{2}\right) \frac{2}{3} = \frac{2}{3\sqrt{e}} > \frac{1}{3}$. Given that S and S' are disjoint, $\mathbb{P}[\hat{z}' \in S] > 1/3$ implies that $\mathbb{P}[\hat{z}' \notin S'] > 1/3$, which contradicts the assumption that $k_{\hat{s}} \leq \alpha W\Delta/4$. $\square$

Next we relate $k_{\hat{s}}$-accuracy to the distortion $\delta(s,\hat{s})$:

Lemma 5. *For $s(\mathbf{d})$ as defined in (1) and a function $\hat{s}:\mathrm{I}^n\to\mathbb{R}$, $k_{\hat{s}}\le\sqrt{3\delta(s,\hat{s})}$.*

Proof. Observe that for all $k\ge\sqrt{3\delta(s,\hat{s})}$, $\mathbb{P}[|s(\mathbf{d})-\hat{s}(\mathbf{d})|\ge k]\le\mathbb{P}[|s(\mathbf{d})-\hat{s}(\mathbf{d})|\ge\sqrt{3\delta(s,\hat{s})}]\le\frac{\mathbb{E}[|s(\mathbf{d})-\hat{s}(\mathbf{d})|^2]}{3\delta(s,\hat{s})}\le\frac{1}{3}$ where the second step follows from Markov's inequality. This implies $k_{\hat{s}}\le\sqrt{3\delta(s,\hat{s})}$. □

Corollary 2. *If $w(H(\hat{s}))>\alpha W$ then $\delta(s,\hat{s})>(\alpha W\Delta)^2/48$.*

Proof. The corollary follows from Lemma 4 and Lemma 5. □

Thus from Corollary 2, we have that if $\delta(s,\hat{s})\le(\alpha W\Delta)^2/48$, then $w(H(\hat{s}))\le\alpha W$. Since $w(H(\hat{s}))\ge\frac{1}{2}\beta(\hat{s})$ (from Lemma 3), it implies if $\delta(s,\hat{s})\le(\alpha W\Delta)^2/48$, then $\frac{1}{2}\beta(\hat{s})\le\alpha W$. This concludes the proof of Theorem 1. □

B Proof of Lemma 2

For the first part of this lemma, observe that the sensitivity of $\sum_i w_i[x_id_i+(1-x_i)a_i]$ w.r.t. i is $S_i(\hat{s})=\Delta|w_i|x_i$. The differential privacy guarantee therefore follows from Lemma 1.

To obtain the lower bound on the distortion, observe that substituting the expressions for s and $\hat{s}$ in the expression for $\delta(s,\hat{s})$, we get

$$\begin{aligned}
\delta(s,\hat{s}) &= \max_{\mathbf{d}\in\mathrm{I}^n}\mathbb{E}[|s(\mathbf{d})-\hat{s}(\mathbf{d};\mathbf{a},\mathbf{x},\sigma)|^2]\\
&= \max_{\mathbf{d}\in\mathrm{I}^n}\mathbb{E}\big[\big(\sum_{i=1}^n w_id_i(1-x_i)-\sum_{i=1}^n w_ia_i(1-x_i)-z\big)^2\big]\ \text{(where } z\sim\mathrm{Lap}(\sigma)\text{)}\\
&= \max_{\mathbf{d}\in\mathrm{I}^n}\big(\sum_{i=1}^n w_i(1-x_i)(d_i-a_i)\big)^2+2\sigma^2\ \text{(since } \mathbb{E}[z]=0;\mathbb{E}[z^2]=2\sigma^2\text{)}\\
&= 2\sigma^2+\max_{\mathbf{d}\in\mathrm{I}^n}\big(\sum_{i=1}^n\gamma_i(d_i-a_i)\big)^2\ \text{(where } \gamma_i:=w_i(1-x_i)\text{)}\\
&= 2\sigma^2+\big(\max_{\mathbf{d}\in\mathrm{I}^n}\big|\sum_{i=1}^n\gamma_i(d_i-a_i)\big|\big)^2
\end{aligned}$$

Observe that $\max_{\mathbf{d}\in\mathrm{I}^n}|f(\mathbf{d})|=\max\big\{\big|\max_{\mathbf{d}\in\mathrm{I}^n}f(\mathbf{d})\big|,\big|\min_{\mathbf{d}\in\mathrm{I}^n}f(\mathbf{d})\big|\big\}$ for any continuous function $f:\mathrm{I}^n\to\mathbb{R}$. Therefore,

$$\begin{aligned}
\delta(s,\hat{s}) &= 2\sigma^2+\big(\max\big\{\big|\max_{\mathbf{d}\in\mathrm{I}^n}\sum_{i=1}^n\gamma_i(d_i-a_i)\big|,\big|\min_{\mathbf{d}\in\mathrm{I}^n}\sum_{i=1}^n\gamma_i(d_i-a_i)\big|\big\}\big)^2\\
&= 2\sigma^2+\big(\max\big\{\big|\gamma^{(+)}R_{\max}+\gamma^{(-)}R_{\min}-\sum_{i=1}^n\gamma_ia_i\big|,\big|\gamma^{(+)}R_{\min}+\gamma^{(-)}R_{\max}-\sum_{i=1}^n\gamma_ia_i\big|\big\}\big)^2,
\end{aligned}$$

where $\gamma^{(+)}:=\sum_{i:\gamma_i\ge0}^n\gamma_i$, and $\gamma^{(-)}:=\sum_{i:\gamma_i<0}^n\gamma_i$. Observe that, for any $a,b,c,\in\mathbb{R}$, it is true that $\max(|a-c|,|b-c|)\ge\frac{|a-b|}{2}$ with equality attained at $c=\frac{a+b}{2}$. Applying this for $a=\gamma^{(+)}R_{\max}+\gamma^{(-)}R_{\min}$, $b=\gamma^{(+)}R_{\min}+\gamma^{(-)}R_{\max}$ and $c=\sum_{i=1}^n\gamma_ia_i$ we get $\min_{\mathbf{a}\in\mathbb{R}^n}\delta(s,\hat{s})\ge2\sigma^2+\frac{(\gamma^+-\gamma^-)(R_{\max}-R_{\min})}{2}=2\sigma^2+\big(\frac{\Delta}{2}\sum_{i=1}^n|w_i|(1-x_i)\big)^2$, with equality attained when $\sum_i\gamma_ia_i=(\gamma^++\gamma^-)(R_{\max}+R_{\min})/2=\sum_i\gamma_i\bar{R}$, which holds for $a_i=\bar{R}$. □

C Proof of Theorem 2 (DCLEFs Suffice)

Consider the function

$$\hat{s}^\circ(\mathbf{d}) := \sum_{i \notin H^\circ} w_i d_i + \bar{R} \sum_{i \in H^\circ} w_i + Lap(w(H^\circ)),$$

where

$$H^\circ := \underset{H: H \subseteq [n] \text{ and } w(H) \leq \alpha W}{\operatorname{argmax}} w(H),$$

is a subset of $[n]$ with maximal weight among all sets of weight at most αW. We can write $\hat{s}^\circ$ as

$$\hat{s}^\circ(\mathbf{d}; \mathbf{x}) := \sum_{i=1}^{n} w_i d_i x_i + \bar{R} \sum_{i=1}^{n} w_i (1 - x_i) + \mathrm{Lap}(w(H^\circ)),$$

where $x_i = 0$ for all $i \in H^\circ$ and $x_i = 1$ otherwise. Hence, $\hat{s}^\circ$ is a DCLEF and

$$\delta(s, \hat{s}^\circ) \overset{\text{Lem. 2}}{=} \frac{9}{4}\Delta^2 \left(\sum_{i=1}^{n} |w_i|(1 - x_i) \right)^2 = \frac{9}{4}\Delta^2 \left(w(H^\circ)\right)^2 \leq \frac{9}{4}(\alpha W \Delta)^2.$$

Let $\hat{s}^*$ be the estimator with the highest privacy index defined at the statement of the theorem. Since, by the definition of $\hat{s}^*$, $\delta(s, \hat{s}^*) \leq (\alpha W \Delta)^2/48$, it follows from Lemma 5 that $k_{\hat{s}^*} \leq \alpha W \Delta/4$. Let $H(\hat{s}^*)$ be as defined in (12). Then, $w(H(\hat{s}^*)) \leq \alpha W$; otherwise, Lemma 4 would imply that $k_{\hat{s}^*} > \alpha W \Delta/4$, a contradiction. Moreover, $w(H^\circ) \geq w(H(\hat{s}^*)) \geq \frac{1}{2}\beta(\hat{s}^*)$: the first inequality follows by the definition of H° and the fact that $w(H(\hat{s}^*)) \leq \alpha W$, and the second from Lemma 3. On the other hand, $\beta(\hat{s}^\circ) \geq w(H^\circ)$, as $\hat{s}^\circ$ is 0-differentially private for every $i \in H^\circ$. It thus follows that $\beta(\hat{s}^\circ) \geq \frac{1}{2}\beta(\hat{s}^*)$. □

D Proof of Theorem 3

D.1 Truthfulness, Individual Rationality, and Budget Feasibility

In this section, we prove that FairInnerProduct is truthful, individually rational, and budget feasible. We first define

$$S_1 := \left\{ t \in [n] \setminus \{i^*\} : \frac{B}{\sum_{i \in [t] \setminus \{i^*\}} |w_i|} \geq \frac{v_t}{W - \sum_{i \in [t] \setminus \{i^*\}} |w_i|} \right\}$$

and

$$S_2 := \left\{ t \in [n] \setminus \{i^*\} : \sum_{i \in [t] \setminus \{i^*\}} |w_i| \geq |w_{i^*}| \right\}.$$

Observe that $S_{-i^*} = S_1 \cap S_2$.

Proposition 1. *FairInnerProduct is budget feasible.*

Proof. When $O = \{i^*\}$ and $\hat{p} = B$, the mechanism is trivially budget feasible. If $\hat{p} = \frac{|w_{i^*}| v_r}{W - |w_{i^*}|}$ then observe that since $r \in S_{-i^*}$, this implies $r \in S_1$ and $r \in S_2$. Therefore, $\hat{p} = \frac{|w_{i^*}| v_r}{W - |w_{i^*}|} \leq \frac{|w_{i^*}| v_r}{W - \sum_{i \in [r] \setminus \{i^*\}} |w_i|} \leq \frac{|w_{i^*}| B}{\sum_{i \in [r] \setminus \{i^*\}} |w_i|} \leq B$ where the second inequality holds because $r \in S_1$ and the last inequality because $r \in S_2$. When $O = [k]$, the sum of the payments made by the mechanism is given by $\sum_{i \leq k} p_i \leq \sum_{i \leq k} |w_i| \frac{B}{w([k])} = \frac{B}{w([k])} \sum_{i \leq k} |w_i| = B$. □

Proposition 2. *If $i^* > k+1$ and $|w_{i^*}| > \sum_{i\in[k]\setminus\{i^*\}} |w_i|$, then $S_{-i^*} = \emptyset$.*

Proof. Observe that if $i^* > k+1$ and $|w_{i^*}| > \sum_{i\in[k]\setminus\{i^*\}} |w_i|$, then $S_1 = [k]$ and $S_2 \cap [k] = \emptyset$. □

Proposition 3. *If $|w_{i^*}| > \sum_{i\in[k]\setminus\{i^*\}} |w_i|$ and $S_{-i^*} \neq \emptyset$, then $r > i^*$.*

Proof. From Proposition 2, $S_{-i^*} \neq \emptyset$ implies either $i^* \leq k+1$ or $|w_{i^*}| \leq \sum_{i\in[k]\setminus\{i^*\}} w_i$. Since the latter is false, it must be that $i^* \leq k+1$. In that case, $S_2 \cap [k] = \emptyset$. Therefore $r > k$. If $i^* = k+1$, then for all $j \in S_2,\ j \geq k+2$. Therefore $r \geq k+2$. □

Proposition 4. *FairInnerProduct is individually rational.*

Proof. We divide the proof into two cases:
Case I: $O = [k]$. We know that $B/w([k]) \geq v_k/(W - w([k]))$ (by construction) and $v_{k+1} \geq v_k$ (by definition). Therefore, for all $i \leq k$, $p_i \geq \frac{|w_i|v_k}{W-w([k])} \geq \frac{|w_i|v_i}{W-w([k])} = c_i(\epsilon_i)$.
Case II: $O = \{i^*\}$. If $p_{i^*} = B$, then the mechanism is individually rational by Assumption 1. If $p_{i^*} = \frac{|w_{i^*}|v_r}{W-|w_{i^*}|}$, then, by Proposition 3, $v_r \geq v_{i^*}$ and therefore the mechanism is individually rational. □

Proposition 5. *FairInnerProduct is dominant-strategy truthful.*

Proof. Fix any $\mathbf{v}$ and assume that user i reports a value $z \neq v_i$, while the remaining values $\mathbf{v}_{-i}$ remain the same. Let $\mathbf{u}$ be the resulting vector of values, *i.e.*, $u_i = z$ and $u_j = v_j$, for $j \neq i$. The vector $\mathbf{u}$ induces a new ordering of the users in terms of their reported values u_i, $i \in [n]$; let $\pi : [n] \to [n]$ be the permutation indicating the position of users under the new ordering. That is, π is 1-1 and onto such that if $u_j < u_{j'}$ then $\pi(j) < \pi(j')$, for all $j, j' \in [n]$. For given $j \in [n]$, we denote the set of users preceding j under this ordering by $P_j = \{j' : \pi(j') \leq \pi(j)\}$. Note that al $j' \in P_j$ satisfy $u'_j \leq u_j$. Observe that if $z > v_i$ then

$$w(P_j) = \begin{cases} w([j]), & \text{for all } j < i \\ w([j]) - |w_i|, & \text{for all } j > i \text{ s.t. } \pi(j) < \pi(i) \\ w([i]) + w(\{\ell : \ell > i \wedge \pi(\ell) < \pi(i)\}), & \text{for } j = i \\ w([i]), & \text{for all} j > i \text{ s.t. } \pi(j) > \pi(i) \end{cases} \tag{17}$$

while if $z < v_i$ then

$$w(P_j) = \begin{cases} w([j]), & \text{for all } j < i \text{ s.t. } \pi(j) < \pi(i) \\ w([i]) - w(\{\ell : \ell < i \wedge \pi(\ell) > \pi(i)\}), & \text{for } j = i \\ w([j]) + |w_i|, & \text{for all } j < i \text{ s.t. } \pi(j) > \pi(i) \\ w([i]), & \text{for all } j > i \end{cases} \tag{18}$$

Let $M_\pi = \left\{j \in [n] : \frac{B}{w(P_j)} \geq \frac{u_j}{W-w(P_j)}\right\}$ where $W = w([n])$. Then, by (17), if $z > v_i$ then $w(P_j) \leq w([j])$ for $j \neq i$ while $w(P_i) \geq w([i])$. As a result, if $z > v_i$, then

$$\text{for } j \neq i, \text{if } j \in [k], \text{ then } j \in M_\pi \tag{19a}$$
$$\text{if } i \notin [k], \text{ then } i \notin M_\pi \tag{19b}$$

Similarly, from (18), if $z < v_i$, then

$$\text{for } j \neq i, \text{if } j \notin [k], \text{ then } j \notin M_\pi \tag{20a}$$

$$\text{if } i \in [k], \text{ then } i \in M_\pi \tag{20b}$$

Observe that, given the value vector $\mathbf{u}$, the mechanism will output $O_\pi = \{i^*\}$, if $|w_i^*| > w(M_\pi \setminus \{i^*\})$, and $O_\pi = M_\pi$ otherwise. If $O_\pi = M_\pi$, users $j \in M_\pi$ are compensated by $p_j = w_j \min\left\{\frac{B}{w(M_\pi)}, \frac{\min_{\ell:\ell\notin M_\pi} u_\ell}{W - w(M_\pi)}\right\}$. If $O_\pi = \{i^*\}$, the latter is compensated by $\hat{p}$ given by (9). We consider the following cases:

Case I: $O_\pi = M_\pi$. If $i \notin M_\pi$, then $p_i = \epsilon_i = 0$, so since FairInnerProduct is individually rational, i has no incentive to report z. Suppose thus that $i \in M_\pi$. We consider the following subcases:

Case I(a): $i \notin [k]$. Then $v_i \geq v_{k+1}$. Since $i \in M_\pi$ but $i \notin [k]$, (19) implies that $z < v_i$. By (20) $k+1 \notin M_\pi$. Thus $p_i \leq |w_i| v_{k+1}/w(M_\pi) \leq |w_i| v_i / w(M_\pi)$.

Case I(b): $i \in [k]$. We will first show that $M_\pi \setminus [k] = \emptyset$. Suppose, for the sake of contradiction, that $M_\pi \setminus [k] \neq \emptyset$. Then $M_\pi \setminus [k]$ must contain an element different than i; this, along with (20) implies that $z > v_i$. If $\pi(i) < \pi(k+1)$, then by (17) $w(P_j) = w([j])$ and $j \notin M_\pi$ for all $j \geq k+1$, which contradicts that $M_\pi \setminus [k]$ is non-empty. Hence, $\pi(i) > \pi(k+1)$; this however implies that $w(P_i) \geq w([k+1])$, by (17), and that $z \geq v_{k+1}$. Thus $\frac{B}{w(P_i)} \leq \frac{B}{w([k+1])} < \frac{v_{k+1}}{W - w([k+1])} \leq \frac{z}{W - w(P_i)}$, so $i \notin M_\pi$, a contradiction. Hence $M_\pi \setminus [k] = \emptyset$.

Next we will show that the original output $O = [k]$. Suppose, for the sake of contradiction, that $O = \{i^*\}$. Then $|w_{i^*}| > w([k] \setminus \{i^*\})$ while $|w_{i^*}| \leq w(M_\pi \setminus \{i^*\})$. Thus, $M_\pi \setminus [k] \neq \emptyset$, a contradiction. Thus, $O = [k]$.

If $O_\pi = M_\pi = [k]$, then since $O = [k]$, user i receives the same payoff, so it has no incentive to report z. Suppose that $M_\pi \neq [k]$. Since $M_\pi \setminus [k] = \emptyset$, it must be that $[k] \setminus M_\pi \neq \emptyset$. By (19), this implies $z < v_i$. If $i < k$, (18) implies that $k \in M_\pi$ and so do all j s.t. $\pi(j) < \pi(k)$. Thus, $[k] = M_\pi$, a contradiction. If $i = k$ and $z < v_i$, then it is possible that $j \notin M_\pi$ for some $j < k$. Thus, $p_i \leq \frac{|w_i| v_k}{w(M_\pi)} = \frac{|w_i| v_i}{w(M_\pi)}$ and so i has no incentive to report z.

Case II. $O_\pi = \{i^*\}$. If $i \neq i^*$, then i's payoff is obviously zero, so it has no incentive to report z. Suppose thus that $i = i^*$. We consider the following two subcases.

Case II(a). $O = \{i^*\}$. Observe that S_{-i^*} and $\hat{p}$ do not depend on v_{i^*}. Thus, since $O = \{i^*\}$, i receives the same payment $\hat{p}$, so it has no incentive to misreport its value.

Case II(b) $O = [k]$. Then $|w_{i^*}| \leq w([k] \setminus \{i^*\})$ while $|w_{i^*}| > w(M_\pi \setminus \{i^*\})$. Thus, $[k] \setminus M_\pi$ must contain an element different than i^*. From (19), this implies that $z < v_i$. If $i < k$, (18) implies that $k \in M_\pi$ and so do all j s.t. $\pi(j) < \pi(k)$. Thus, $[k] = M_\pi$, a contradiction.

Assume thus that $i \geq k$. Then $v_i \geq v_k$. Let $j^* = k$ if $i > k$ and $j^* = k - 1$ if $i = k$. Observe that $j^* \in S_{-i^*}$: indeed, it is in S_1 since $i \in [k]$, by the definition of k, and it is in S_2 because $|w_{i^*}| \leq w([k] \setminus \{i^*\})$. Hence $\hat{p} \leq \frac{|w_i| v_{j^*}}{W - |w_i|} \leq \frac{w_i v_k}{W - |w_i|} \leq \frac{|w_i| v_i}{W - |w_i|}$ so i's payoff is at most zero, so it has no incentive to misreport its value. □

D.2 Approximation Ratio

In this section we prove that FairInnerProduct is 5-approximate with respect to OPT.

D.2.1 Optimal Continuous Canonical Laplace Mechanism

We first characterize an individually rational, budget feasible, continuous canonical Laplace mechanism that has optimal distortion. Consider the fractional relaxation of (7).

$$\text{maximize} \quad \sum_{i=1}^{n} |w_i| x_i \tag{21a}$$

$$\text{subject to} \quad p_i \geq c_i(\epsilon_i) = v_i \epsilon_i(\mathbf{x}), \quad \forall i \in [n] \tag{21b}$$

$$\sum_{i=1}^{n} p_i \leq B \tag{21c}$$

$$0 \leq x_i \leq 1, \quad \forall i \in [n] \tag{21d}$$

where $\epsilon_i(\mathbf{x}) = \frac{|w_i| x_i}{\sum_i |w_i|(1-x_i)}$. A budget feasible, individually rational, (but not necessarily discrete or truthful) canonical Laplace mechanism for the inner product has a minimal distortion among all such mechanisms if given input $(\mathbf{v}, \mathbf{w}, B)$ it outputs $(\mathbf{x}^*, \mathbf{p}^*)$, where the latter constitute an optimal solution to the above problem. This characterization will yield the approximation guarantee of the DCLEF mechanism[1].

Lemma 6. *Recall that $v_1 \leq v_2 \leq \ldots \leq v_n$. For $0 \leq k \leq n$, define $p(k) := \sum_{i=k+1}^{n} |w_i|$, if $0 \leq k \leq n-1$, and $p(n) := 0$. For $0 \leq k \leq n$, define $q(0) := 0$, and $q(k) := \sum_{i=1}^{k} v_i |w_i|$, if $1 \leq k \leq n$. Define $\ell := \min\{k : \ \forall i > k, q(i) - Bp(i) > 0\}$ and let*

$$x_i^* := \begin{cases} 1, & \text{if } i \leq \ell \\ \frac{Bp(\ell) - q(\ell)}{(v_{\ell+1} + B)|w_{\ell+1}|}, & \text{if } i = \ell + 1 \\ 0, & \text{if } i > \ell + 1 \end{cases}, \text{ and } p_i^* = v_i |w_i| x_i^* / \sigma(\mathbf{x}^*) \quad i \in [n].$$

Then $(\mathbf{x}^*, \mathbf{p}^*)$ is an optimal solution to (21).

Proof. We show first that the quantities ℓ and x_i^* are well defined. For $p(i)$, $q(i)$, $i \in \{0, \ldots, n\}$, as defined in the statement of the theorem, observe that $g(i) = q(i) - Bp(i)$ is strictly increasing and that $g(0) < 0$ while $g(n) > 0$. Hence, ℓ is well defined; in particular, $\ell \leq n-1$. The monotonicity of g implies that $g(i) \leq 0$ for all $0 \leq i \leq \ell$ and $g(i) > 0$ for $i > \ell$. For $a \in [0,1]$, let $h(a) = q(\ell) + v_{\ell+1}|w_{\ell+1}|a - B(p(\ell+1) + |w_{\ell+1}|(1-a))$. Then $h(0) = g(\ell) \leq 0$ and $h(1) = g(\ell+1) > 0$. As $h(a)$ is continuous and strictly increasing in the reals, there exists a unique $a^* \in [0,1]$ s.t. $h(a) = 0$; since h is linear, it is easy to verify that $a^* = q(\ell) - Bp(\ell)/(v_\ell + B)|w_{\ell+1}| = x_{\ell+1}^*$ and, hence, $x_{\ell+1}^* \in [0,1]$. To solve (21), we need only consider cases for which constraint (21b) is tight, *i.e.*, $p_i = v_i \epsilon_i(\mathbf{x})$. Any solution for which (21b) is not tight can be converted to a solution where it is; this will only strengthen constraint (21c), and will not affect the objective. Thus, (21) is equivalent to:

$$\text{Max.} \quad F(\mathbf{x}) = \sum_{i=1}^{n} |w_i| x_i \tag{22a}$$

$$\text{subj. to} \quad \sum_{i=1}^{n} v_i |w_i| x_i - B \sum_{i=1}^{n} w_i (1 - x_i) \leq 0, \quad \mathbf{x} \in [0,1]^n \tag{22b}$$

[1] An analogous characterization of the budget-limited knapsack mechanism in [Sin10] can be used to show that the mechanism is 5-approximate instead of 6-approximate.

It thus suffices to show that $\mathbf{x}^*$ is an optimal solution to (22). The latter is a linear program and its Lagrangian is

$$L(\mathbf{x}, \lambda, \mu, \nu) = -F(\mathbf{x}) + \lambda\big(\sum_{i=1}^{n} v_i|w_i|x_i - B\sum_{i=1}^{n}|w_i|(1-x_i)\big) + \sum_{i=1}^{n}\mu_i(x_i - 1) - \sum_{i=1}^{n}\nu_i x_i.$$

It is easy to verify that $\mathbf{x}^*$ satisfies the KKT conditions of (22) with $\lambda^* = \frac{1}{v_{\ell+1}+B}$, $\mu_i^* = \mathbb{1}_{(i\le\ell)} \cdot \frac{v_{\ell+1}-v_i}{v_{\ell+1}+B}|w_i|$, and $\nu_i^* = \mathbb{1}_{(i>\ell+1)} \cdot \frac{v_i - v_{\ell+1}}{v_{\ell+1}+B}|w_i|$. □

A canonical Laplace mechanism that outputs $(\mathbf{x}^*, \mathbf{p}^*)$ given by Lemma 6 would be optimal. Moreover, the objective value $S(\mathbf{x}^*; \mathbf{w}) \ge OPT$.

Proposition 6. *Let ℓ be as is defined in Lemma 6, and k as defined in FairInnerProduct. Then, $\ell \ge k$.*

Proof. Assume that $\ell < k$. Then

$$B(W - w([k])) \le B(W - w([\ell+1])) < \sum_{i=1}^{\ell+1}|w_i|v_i \le \sum_{i=1}^{k}|w_i|v_i \le v_k\sum_{i\le k}w_i = v_k w([k]).$$

However, this contradicts the fact that $B/w(k) \ge v_k/(W - w([k]))$. □

Proposition 7. *Let $\{x_i^*\}$ and ℓ be as defined in Lemma 6, and k as defined in FairInnerProduct. Then, $w([k+1]) > \sum_{i=k+1}^{\ell+1}|w_i|x_i^*$.*

Proof. If $\ell = k$, the statement is trivially true. Consider thus the case $\ell > k$. Assume that $\sum_{i=1}^{k+1} w_i \le \sum_{i=k+1}^{\ell+1}|w_i|x_i^*$. Then,

$$\begin{aligned}\frac{B(W - w([k+1]))}{w([k+1])} &\ge \frac{B(W - \sum_{i=k+1}^{\ell+1}|w_i|x_i^*)}{\sum_{i=k+1}^{\ell+1}|w_i|x_i^*} \ge \frac{B(W - \sum_{i=1}^{\ell+1}|w_i|x_i^*)}{\sum_{i=k+1}^{\ell+1}|w_i|x_i^*} \\ &\ge \frac{\sum_{i=k+1}^{\ell+1}|w_i|v_i x_i^*}{\sum_{i=k+1}^{\ell+1}|w_i|x_i^*} \ge v_{k+1}\end{aligned}$$

since $v_{k+1} \le v_i$ for all $(k+1) \le i \le \ell$. However, this contradicts the fact that $B/w([k+1]) < v_{k+1}/(W - w([k+1]))$. □

Now we will show that $S(\mathbf{x}; \mathbf{w}) \ge \frac{1}{5}OPT$ using Proposition 7. First notice that since (21) is a relaxation of (7), $OPT \le S(\mathbf{x}^*; \mathbf{w})$, where $\{x_i^*\}$ are defined in Lemma 6. Therefore, we have that $OPT \le S(\mathbf{x}^*; \mathbf{w}) = \sum_{i\le k}|w_i| + \sum_{i=k+1}^{\ell+1}|w_i|x_i^* \overset{\text{Prop. 7}}{<} w([k]) + w([k+1]) \le 2w([k]) + |w_{i^*}|$ It follows that if $O = [k]$, it implies $w([k]) \ge |w_{i^*}|$ and therefore $w([k]) = S(\mathbf{x}; \mathbf{w}) \ge \frac{1}{3}OPT$. On the other hand, if $O = \{i^*\}$, then $|w_{i^*}| > \sum_{j\in[k+1]\setminus\{i^*\}}|w_j|$, which implies $2w_{i^*} > w([k])$. Therefore, $OPT \le 2w([k]) + |w_{i^*}| < 5|w_{i^*}| = 5S(\mathbf{x}; \mathbf{w})$.

D.3 The Uniform-Weight Case

In the uniform-weight case, FairInnerProduct reduces to the Ghosh-Roth mechanism. Our benchmark OPT is stricter than that used by Ghosh and Roth. In particular, they show their mechanism is optimal among all truthful, individually rational, budget-feasible and envy-free mechanisms. In this section, we prove that when all weights are equal, FairInnerProduct is 2-approximate with respect to OPT.

Let $|w_i| = u$ for all $i \in [n]$. First, observe that in this case, FairInnerProduct always outputs $O = [k]$. Therefore, $S(\mathbf{x}; \mathbf{w}) = ku$. We use this observation to prove the result.

Lemma 7. *Assume that for all $i \in [n]$, $|w_i| = u$. Then, $S(\mathbf{x}; \mathbf{w}) \geq \frac{1}{2} OPT$.*

Proof. Observe that $OPT \leq S(\mathbf{x}^*; \mathbf{w}) = \sum_{i=1}^{\ell+1} |w_i| x_i^* = w([k]) + \sum_{i=k+1}^{\ell+1} |w_i| x_i^* < w([k]) + w([k+1])$ from Proposition 7, where $\{x_i^*\}$ and ℓ are defined in Lemma 6. Substituting $|w_i| = u$ for all i, we get $OPT < (2k+1)u$. Since OPT is the objective value attained by the optimal DCLEF mechanism, $OPT = mu$ for some $m \in [n]$. This implies $2k+1 > m$. Since k and m are integers, it follows that $2k \geq m$, or equivalently, $S(\mathbf{x}; \mathbf{w}) \geq \frac{1}{2} OPT$. □

E Proof of Theorem 4 (Impossibility of Approximation)

Consider the following example. Let $n = 4$. The private costs of the four individuals are given by $v_1 = a, v_2 = v_3 = v_4 = 2$, where $0 < a < 2$. The weights of the four individuals are given by $w_1 = w_2 = w_3 = w = d$, where $d > 0$. Let the budget $B = 1 + a/2 < 2$.

Observe that the optimal individually rational, budget-feasible, DCLEF mechanism would set $x_1^* = 1$ and exactly one of x_2^*, x_3^* and x_4^* to 1. Without loss of generality, assume that $x_1^* = x_2^* = 1$ and $x_3^* = x_4^* = 0$. Therefore, the optimal weight $OPT = 2d$. Consider a truthful, individually rational, budget-feasible, DCLEF mechanism that is $2 - \varepsilon$ approximate, for any $\varepsilon > 0$. Such a mechanism must set at least two of the x_i's to 1 (since it is $2 - \varepsilon$ approximate). Therefore, for such a mechanism $\sigma(\mathbf{x}) \leq 2d$. Moreover, since the mechanism is budget-feasible, it must set $x_i = 1$, as otherwise the sum of payments, $\frac{1}{\sigma(\mathbf{x})} \sum_i v_i w_i x_i \geq \frac{4d}{\sigma(\mathbf{x})} \geq 2 > B$. For such a mechanism, the cost of individual 1 is $c_1(\epsilon_1) = v_1 w_1 / \sigma(\mathbf{x}) \geq v_1 d / (2d) \geq v_1/2$. Since the mechanism is truthful, the payment p_1 cannot depend on v_1. Also, for this mechanism to be individually rational, p_1 must be at least 1 (since v_1 can be arbitrarily close to 2), which implies that the remaining budget is strictly less than 1. However, for this mechanism, for $i \in \{2, 3, 4\}$, $c_i(\epsilon_i) = 2d/\sigma(\mathbf{x}) \geq 1$. This means that this mechanism cannot be both individually rational and budget feasible. □