Toward Controlling Discrimination in Online Ad Auctions

L. Elisa Celis [1] **Anay Mehrotra** [2] **Nisheeth K. Vishnoi** [3]

Abstract

Online advertising platforms are thriving due to the customizable audiences they offer advertisers. However, recent studies show that advertisements can be discriminatory with respect to the gender or race of the audience that sees the ad, and may inadvertently cross ethical and/or legal boundaries. To prevent this, we propose a constrained ad auction framework that maximizes the platform's revenue conditioned on ensuring that the audience seeing an advertiser's ad is distributed appropriately across sensitive types such as gender or race. Building upon Myerson's classic work, we first present an optimal auction mechanism for a large class of fairness constraints. Finding the parameters of this optimal auction, however, turns out to be a non-convex problem. We show that this non-convex problem can be reformulated as a more structured non-convex problem with no saddle points or local-maxima; this allows us to develop a gradient-descent-based algorithm to solve it. Our empirical results on the A1 Yahoo! dataset demonstrate that our algorithm can obtain uniform coverage across different user types for each advertiser at a minor loss to the revenue of the platform, and a small change to the size of the audience each advertiser reaches.

1. Introduction

Online advertisements are the main source of revenue for social-networking sites and search engines such as Google (Alphabet Inc.). Ad exchange platforms allow advertisers to select the target audience for their ad by specifying desired user demographics, interests and browsing histories (Facebook, Ad Targeting). Every time a user loads a webpage or enters a search term, bids are collected from relevant advertisers (Google, Ad Rank), and an auction is conducted to determine which ad is shown, and how much the advertiser is charged (Muthukrishnan, 2009; Yuan et al., 2012; Varian, 2007). As it is not practical for advertisers to place individual bids for every user, the advertiser instead gives some high-level preferences about their budget and target audience, and the platform places bids on their behalf (Google, Automated Bidding).

More formally, let there be n advertisers, and m types of users. Each advertiser i specifies their target demographic, average bid, and budget to the platform, which then decides a distribution, $\mathcal{P}_{ij}$, of bids of advertiser $i \in [n]$ for user type $j \in [m]$. These distributions represent the value of the user to the advertiser, and ensure that the advertiser only bids for users in their target demographic, with the expected bid not exceeding the amount specified by the advertiser (Facebook, Bid Strategies). At each time step, a user visits a web page (e.g., Facebook or Twitter), the user's type j is observed, and a bid v_i is drawn from $\mathcal{P}_{ij}$, for each advertiser i. Receiving these bids as input, the mechanism $\mathcal{M}$ decides an allocation $x(v)$ and price $p(v)$ for the advertisement slot. Several Ad Exchanges including Google Ads (Ad targeting) and Facebook Ads (About Ad Auctions), use variants of second price auction mechanism (Ostrovsky & Schwarz, 2011)[1].

Overall, such targeted advertising leads to higher utilities for the advertisers who show content to relevant audiences, for the users who view related advertisements, and for the platform which can benefit from selling targeted advertisements (Farahat, 2013; Yan et al., 2009; Fox-Brewster, 2017; Goldfarb & Tucker, 2011). However, targeted advertising can also lead to discriminatory practices. For instance, searches with "black-sounding" names were much more likely to be shown ads suggestive of an arrest record (Sweeney, 2013). Another study found that women were shown fewer advertisements for high paying jobs than men with similar profiles (Datta et al., 2015). In fact, recent experiments demonstrate that ads can be *inadvertently* discriminatory; (Lambrecht & Tucker, 2018) found that STEM job ads, specifically designed to be unbiased by the advertisers, were shown to more men than women across all major platforms (Facebook Ads, Google Ads, Instagram and Twitter). On Facebook, a platform with 52% women (Vermeren,

[1]Yale University, USA [2]Indian Institute of Technology Kanpur, India [3]Yale University, USA. Correspondence to: L. Elisa Celis <elisa.celis@yale.edu>.

Proceedings of the 36^{th} International Conference on Machine Learning, Long Beach, California, PMLR 97, 2019.

[1]If the auction sells a single item, then Myerson's mechanism (Myerson, 1981) reduces to a second price auction mechanism with a reserve price (Hartline, 2017).

2015) the advertisement was shown to 20% more men than women. (Ali et al., 2019) find that this could be a result of *competitive spillovers* among advertisers, and is neither a pure reflection of pre-existing cultural bias, nor a result of user input to the algorithm. Such (likely inadvertent) discrimination has led to two recent cases filed against Facebook, which will potentially lead to civil lawsuits alleging employment and housing discrimination (Guynn, 2018; Timberg & Jan, 2018; NFHA, 2018; Angwin & Parris Jr., 2016).

To gain intuition on how inadvertent discrimination could happen, consider the setting in which there are two advertisers with similar bids/budgets, but one advertiser specifically targets women (which is allowed for certain types of ads, e.g., related to clothing), while the second advertiser does not target based on gender (e.g., because they are advertising a job). The first advertiser creates an imbalance on the platform by taking up ad slots for women and, as a consequence, the second advertiser ends up advertising to disproportionately fewer women and is inadvertently discriminatory. Currently, online advertising platforms have no mechanism to check this type of discrimination. In fact, the only way around this would be for the advertiser to set up separate campaigns for different user types and ensure that each campaign reached a similar number of the sub-target audience. However, online platforms often reject such campaigns in the apprehension of discriminatory practices (Lambrecht & Tucker, 2018; Discriminatory practices).

Our Contributions

Our main contribution is an optimization-based framework which maximizes the revenue of the platform subject to satisfying constraints that prevent the emergence of inadvertent discrimination as described above. The constraints can be formulated as any one of a wide class of "group fairness" constraints as presented in (Celis et al., 2019c), which constrains the distribution of an ad's audience across the sensitive types to ensure proportionality across types as defined by the platform. The framework allows for intersectionality, allowing constraints across multiple sensitive attributes (e.g., gender, race, geography and economic class) and allows for restricting different advertisers to different constraints.

Formally, building on Myerson's seminal work (Myerson, 1981), we characterize the truthful revenue-optimal mechanism which satisfies the given constraints (Theorem 1). The user types, as defined by their sensitive attributes, are taken as input along with the type-specific bid distributions for each advertiser, and we assume that bids are drawn from these distributions independently. Our mechanism is parameterized by constant "shifts" which it applies to bids for each advertiser-type pair. Finding the parameters of this optimal mechanism, however, is a non-convex optimization problem, both in the objective and the constraints. Towards solving this, we first propose a novel reformulation of the objective as a composition of a convex function constrained on a polytope, and an unconstrained non-convex function (Theorem 2). Interestingly, the non-convex function is reasonably well behaved, with no saddle-points or local-maxima. This allows us to develop a gradient descent based scheme (Algorithm 1) to solve the reformulated program, which under mild assumptions has a fast convergence rate of $\widetilde{O}(1/\varepsilon^2)$ (Theorem 3).

We evaluate our approach empirically by studying the effect of the constraints on the revenue of the platform and the advertisers using the *Yahoo! Search Marketing Advertising Bidding Data* (Yahoo). We find that our mechanism can obtain uniform coverage across different user types for each advertiser while losing less than 5% of the revenue (Figure 1(b)). Further, we observe that the total-variation distance between the fair and unconstrained distributions of total advertisements an advertiser shows on the platform is less than 0.05 (Figure 1(c)).

To the best of our knowledge, we are the first to give a framework to prevent inadvertent discrimination in online ad auctions.

2. Our model

Preliminaries

We refer the reader to the excellent treatise (Hartline, 2017) on Mechanism design for a detailed discussion of the preliminaries. In addition, we provide some key definitions in Section B in the Supplementary File.

A mechanism $\mathcal{M}$ is defined by its allocation rule $x\colon \mathbb{R}^n \to [0,1]^n$, and its payment rule $p\colon \mathbb{R}^n \to \mathbb{R}^n_{\geq 0}$. It is a well known fact (Nisan et al., 2007) that for any truthful mechanism, p is uniquely defined by x. Hence, for any truthful mechanism our only concern is the allocation rule x.

Let $\mathcal{P}$ be the distribution of valuation of a bidder, $\mathrm{pdf}\colon \mathbb{R} \to \mathbb{R}_{>0}$ be its probability density function, and $\mathrm{cdf}\colon \mathbb{R} \to [0,1]$ be its cumulative density function, then we define the *virtual valuation* $\phi\colon \mathrm{supp}(\mathcal{P}) \to \mathbb{R}$, as $\phi(v) \coloneqq v - (1 - \mathrm{cdf}(v))(\mathrm{pdf}(v))^{-1}$. We say $\mathcal{P}$ is *regular* if $\phi(v)$ is non-decreasing in v. Likewise, we say $\mathcal{P}$ is *strictly regular* if $\phi(v)$ is strictly increasing in v.

Let $\phi_{ij} \in \mathbb{R}$ be the virtual valuation of advertiser $i \in [n]$ for type $j \in [m]$, $f_{ij}\colon \mathbb{R} \to \mathbb{R}_{\geq 0}$ be its probability density function, and $F_{ij}\colon \mathbb{R} \to [0,1]$ be its cumulative density function. We denote the joint virtual valuation of all advertisers for type j by $\phi_j \in \mathbb{R}^n$, and its joint probability density function by $f_j\colon \mathbb{R}^n \to \mathbb{R}_{\geq 0}$. The types $j \in [m]$ are distributed according to a known distribution $\mathcal{U}$. Finally, given a user of type j, let a mechanism's allocation rule be $x_j\colon \mathbb{R}^n \to [0,1]^n$.

2.1. Fairness Constraints

We would like to guarantee that advertisers have a fair coverage across user types. We do so by placing constraints on the coverage of an advertiser. Formally, we define advertiser i's coverage of type j, q_{ij}, as the joint probability that advertiser i wins the auction and the user is of type j

$$q_{ij}(x_j) := \Pr_{\mathcal{U}}[j] \int_{\mathrm{supp}(\phi_j)} x_{ij}(\phi_j) df_j(\phi_j), \qquad \text{(Coverage, 1)}$$

where $x_{ij}(\phi_j)$ is the i-th component of $x_j(\phi_j)$. Then, we consider the proportional coverage of the advertiser on each type. Given vectors $\ell_j,\, u_j \in [0,1]^n\ \forall\, j \in [m]$, we define (ℓ, u)-fairness constraints for each advertiser i and type j, as a lower bound ℓ_{ij}, and an upper bound u_{ij}, on the proportion of users of type j the advertiser shows ads to, i.e., we impose the following constraints for all $i \in [n]$ and $j \in [m]$

$$\ell_{ij} \leq \frac{q_{ij}}{\sum_{t=1}^{m} q_{it}} \leq u_{ij}. \quad \text{(}(\ell, u)\text{-fairness constraints, 2)}$$

2.2. Discussion of Fairness constraints

Returning to the example presented in the introduction, we can ensure that the advertiser shows $x\%$ of total ads to women, by choosing a lower bound of x for this advertiser on women. More generally, for m user types, moderately placed lower bounds and upper bounds ($\ell_{ij} \sim {}^1\!/\!_m$ and $u_{ij} \sim {}^1\!/\!_m$), for some subset of advertisers, ensure this subset has a uniform coverage across all types, while allowing other advertisers to target specific types.

Importantly, while ensuring fairness across multiple types our constraints allow for targeting within any single type. This is vital as the advertiser may not derive the same utility from each user, and could be willing to pay a higher amount for more relevant users in the same type. For example, if the advertiser is displaying job ads, then a user already looking for job opportunities may be of a higher value to the advertiser than one who is not.

For a detailed discussion on how such constraints can encapsulate other popular metrics, such as satistical parity, we refer the reader to (Celis et al., 2019a).

2.3. Optimization Problem

We would like to develop a mechanism which maximizes the revenue while satisfying the upper and lower bound constraints in Eq. (2). Towards formally stating our problem, we define the revenue of mechanism $\mathcal{M}$, with an allocation rule $x_j \colon \mathbb{R}^n \to [0,1]^n$ for type j as

$$\mathrm{rev}_{\mathcal{M}} := \sum_{i\in[n],\, j\in[m]} \Pr_{\mathcal{U}}[j] \int_{\mathrm{supp}(\phi_j)} \phi_{ij} x_{ij}(\phi_j) df_j(\phi_j), \quad \text{(Revenue, 3)}$$

where $x_{ij}(\phi_j)$ and ϕ_{ij} are the i-th component of $x_j(\phi_j)$ and ϕ_j respectively. Thus, we can express our optimization problem with respect to *functions* $x(\cdot)$, or as an infinite dimensional optimization problem as follows.

(**Infinite-dimensional fair advertising problem).** For all user types $j \in [m]$, find the optimal allocation rule $x_j(\cdot) \colon \mathbb{R}^n \to [0,1]^n$ for

$$\max_{x_{ij}(\cdot)\geq 0} \ \mathrm{rev}_{\mathcal{M}}(x_1, x_2, \ldots, x_m) \tag{4}$$

$$\text{s.t.,} \quad q_{ij}(x_j) \geq \ell_{ij} \sum\nolimits_{t=1}^{m} q_{it}(x_t) \ \ \forall\, i \in [n], j \in [m] \tag{5}$$

$$q_{ij}(x_j) \leq u_{ij} \sum\nolimits_{t=1}^{m} q_{it}(x_t) \ \ \forall\, i \in [n], j \in [m] \tag{6}$$

$$\sum\nolimits_{i=1}^{n} x_{ij}(\phi_j) \leq 1 \qquad \forall\, j \in [m], \phi_j, \tag{7}$$

where (5) and (6) encode the lower bound and upper bound constraints, and (7) ensures that only one ad is allocated. In the above problem, we are looking for a collection of optimal continuous function $x^\star$. To be able to solve this problem, we need – in the least – a finite dimensional formulation of the fair online advertisement problem.

3. Other Related Work

(Dwork & Ilvento, 2019) consider a framework which selects an ad category (e.g., job or housing) every time a user visits the platform. Given fair mechanisms for each category, they construct a *fair composition* of these mechanisms. However, they do not show how to design fair mechanisms for each category, or study how the composition affects the platform's ad revenue. Another related problem is to design optimal mechanisms which satisfy contract constraints (Ghosh et al., 2009; Balseiro et al., 2014; Pai & Vohra, 2012); these constraints allocate a minimum number of ad spots to advertisers with a contract, and are different from our constraints which control the fraction of each sensitive type the ads are shown to.

Several prior works address the problems of polarization and algorithmic bias, including (Garimella et al., 2018; Celis et al., 2019b) who control polarization in social-networks and personalized feeds, (Panigrahi et al., 2012) who diversify personal feeds, and (Celis et al., 2018b) who create a diverse and balanced summary of a set of results. In addition, (Radlinski et al., 2008; Asudeh et al., 2017; Celis et al., 2018c) study fair ranking algorithms; these could be used to generate a balanced list of results on job platforms and other search engines. While these works are related to our broad goal of controlling algorithmic bias, their formulation is different since they do not involve a bidding mechanism. Therefore, their solutions cannot be applied to our problem.

Finally, a framework approach to fairness constraints has shown to be effective in various other applications such as classification (Celis et al., 2019a; Huang & Vishnoi, 2019; Zafar et al., 2017), selection of representatives (Celis et al., 2018a), and personalization (Celis & Vishnoi, 2017).

4. Theoretical Results

Our first result is structural, and gives a characterization of the optimal solution $x^\star$, to the infinite-dimensional fair advertising problem, in terms of a matrix $\alpha \in \mathbb{R}^{n\times m}$, making it a finite dimensional optimization problem in α.

Theorem 1. ***(Characterization of an optimal allocation rule).*** *There exists an $\alpha = \{\alpha_j\}_{j\in[m]} \in \mathbb{R}^{n\times m}$ such that if for all $j \in [m]$, $\mathcal{P}_j$ are strictly regular and independent, then the set of allocation rules $x_j(\cdot,\alpha_j)\colon \mathbb{R}^n \to [0,1]^n\ \forall\ j \in [m]$, defined below, is optimal for the infinite-dimensional fair advertising problem*

$$x_{ij}(v_j,\alpha_j) \coloneqq \mathbb{I}[\ i \in \mathrm{argmax}_{\ell\in[n]}(\phi_{\ell j}(v_{\ell j}) + \alpha_{\ell j})\ ]. \qquad (\alpha\text{-shifted mechanism, 8})$$

Where we randomly breaks ties if any (this is equivalent to the allocation rule of the VCG mechanism).

We present the proof of Theorem 1 in Section D.1 in the Supplementary File. In the proof, we analyze the dual of the infinite-dimensional fair advertising problem. We reduce the dual problem to one lagrangian variable, by fixing the lagrangian variables corresponding lower bound (5) and upper bound (6) constraints to their optimal values. The resulting problem turns out to be the dual of the unconstrained revenue maximizing problem, for which Myerson's mechanism is the optimal solution. We interpret the fixed lagrangian variables as shifting the original virtual valuations ϕ_{ij}. It then follows that for some shift $\alpha \in \mathbb{R}^{n\times m}$, the α-shifted mechanism (8) is the optimal solution to the infinite-dimensional fair advertising problem.

Now, our task is reduced from finding an optimal allocation rule, to finding an α characterizing the optimal allocation rule. Towards this, let us define the revenue, $\mathrm{rev}_{\mathrm{shift}}\colon \mathbb{R}^{n\times m} \to \mathbb{R}$ and coverage $q_{ij}\colon \mathbb{R}^{n\times m} \to [0,1]$ as functions of α

$$\mathrm{rev}_{\mathrm{shift}}(\alpha) \coloneqq \sum_{\substack{i\in[n]\\ j\in[m]}} \mathrm{Pr}_{\mathcal{U}}[j] \int_{\mathrm{supp}(f_{ij})} y f_{ij}(y) \prod_{k\in[n]\setminus\{i\}} F_{kj}(y+\alpha_{ij}-\alpha_{kj})dy \qquad (\text{Revenue } \alpha\text{-shifted mechanism, 9})$$

$$q_{ij}(\alpha) \coloneqq \mathrm{Pr}_{\mathcal{U}}[j] \int_{\mathrm{supp}(f_{ij})} f_{ij}(y) \prod_{k\in[n]\setminus\{i\}} F_{kj}(y+\alpha_{ij}-\alpha_{kj})dy. \qquad (\text{Coverage } \alpha\text{-shifted mechanism, 10})$$

These follow by observing that (8) selects the advertiser with the highest shifted virtual valuation, and then using this allocation rule in Eq. (3) and Eq. (1) respectively. Depending on the nature of the distribution, the gradients ${\partial \mathrm{rev}_{\mathrm{shift}}(\alpha)}/{\partial \alpha_i}$ and ${\partial q_{ij}(\alpha)}/{\partial \alpha_i}$ may not be monotone in α (e.g., consider the exponential distribution). Therefore, in general neither is $\mathrm{rev}_{\mathrm{shift}}(\cdot)$ a concave, nor is $q_{ij}(\cdot)$ a convex function of α (see Section D.2 in the Supplementary File for a concrete example). Hence, this optimization problem is non-convex both in its objective and in its constraints. We require further insights to solve the problem efficiently.

Towards this, we observe that revenue is a concave function of q. Consider two optimal allocation rules obtaining coverages $q_1,\ q_2 \in [0,1]^{n\times m}$ and revenues $R_1,\ R_2 \in \mathbb{R}$ respectively. If we use the first with probability $\gamma \in [0,1]$, we achieve a coverage $\gamma q_1 + (1-\gamma)q_2$ with revenue $\gamma R_1 + (1-\gamma)R_2$. Therefore, the optimal allocation rule achieving $\gamma q_1 + (1-\gamma)q_2$ has a revenue of at least $\gamma R_1 + (1-\gamma)R_2$. This shows that for optimal allocation rules revenue is a concave function of the coverage q.

Let $\mathrm{rev}\colon [0,1]^{(n-1)\times m} \to \mathbb{R}$, be the maximum revenue of the platform as a function of coverage q.[2] Consider the following two optimization problems.

(Optimal coverage problem). Find the optimal $q \in [0,1]^{n\times m}$ for,

$$\max_{q\in[0,1]^n}\ \mathrm{rev}(q) \qquad (11)$$

$$\text{s.t.,}\quad q_{ij} \geq \ell_{ij} \sum_{t=1}^{m} q_{it} \qquad \forall\ i\in[n],\ j\in[m] \quad (12)$$

$$q_{ij} \leq u_{ij} \sum_{t=1}^{m} q_{it} \qquad \forall\ i\in[n],\ j\in[m] \quad (13)$$

$$\sum_{i=1}^{n} q_{ij} \leq\ \mathrm{Pr}_{\mathcal{U}}[j] \qquad \forall\ j\in[m]. \quad (14)$$

(Optimal shift problem). Given the target coverage $\delta \in [0,1]^{n\times m}$, find the optimal $\alpha \in \mathbb{R}^{n\times m}$ for,

$$\min_{\alpha\in\mathbb{R}^{n\times m}} \mathcal{L}(\alpha) \coloneqq \|\delta - q(\alpha)\|_F^2. \qquad (15)$$

Our next result relates the solution of the above two problems with the infinite-dimensional fair advertising problem.

Theorem 2. *Given a solution $q^\star \in [0,1]^{n\times m}$ to the optimal coverage problem, the solution $\alpha^\star$ to the optimal shift problem with $\delta = q^\star$, defines an optimal α-shifted mechanism (8) for the infinite-dimensional fair advertising problem.*

Proof. For any $j \in [m]$ adding the all 1 vector, 1_n, to α_j does not change the allocation rule in (8). Thus, it suffices to show that for all $\delta \in [0,1]^{n\times m}$, there is a unique α with $\alpha_{1j} = 0\ \forall\ j \in [m]$, such that $q(\alpha) = \delta$.

We can show that for all $\delta \in [0,1]^{n\times m}$, there is at-least one $\alpha \in \mathbb{R}^{n\times m}$ such that $q(\alpha) = \delta$. In fact, the greedy algorithm which increases all α_{ij}, where $q_{ij}(\alpha) < \delta_{ij}$ and $i \neq 1$, will find the required α.

To prove it is unique consider distinct $\alpha, \beta \in \mathbb{R}^{n\times m}$ such that $\alpha_{1j} = \beta_{1j} = 0$. We can show that $q(\alpha) \neq q(\beta)$. In particular, that $q_{i'j'}(\alpha) \neq q_{i'j'}(\beta)$ for $(i',j') = \mathrm{argmax}_{i\in[n],\ j\in[m]} |\alpha_{ij} - \beta_{ij}|$. Now, the uniqueness of α follows by contradiction. ■

The above theorem allows us to find the optimal α by solving the *optimal coverage* and *optimal shift* problems. First, let us consider the *optimal coverage problem*. We already know that its objective is concave. We can further observe that its constraints are linear in q, and in particular, they

[2] We drop q_{ij} for one $i \in [n]$ and each $j \in [m]$. This is crucial to calculate $\nabla\mathrm{rev}$ (see Remark 5). By some abuse of notation we write $\mathrm{rev}(q)$ for $q \in \mathbb{R}^{n\times m}$ instead of using $q \in \mathbb{R}^{(n-1)\cdot m}$.

define a constraint-polytope $\mathcal{Q} \subseteq [0,1]^{n\times m}$. Therefore, it is a convex program, and one approach to solve it is to use gradient-based algorithms.

The problem is that we do not have access to $\nabla\mathrm{rev}$. The key idea is that if we let $\alpha = q^{-1}(\delta)$, then we can calculate $\nabla\mathrm{rev}(\delta)$ by solving the following linear-system,

$$(J_q(\alpha))^\top \nabla\mathrm{rev}(\delta) = \nabla\mathrm{rev}_{\mathrm{shift}}(\alpha), \quad \text{(Gradient Oracle, 16)}$$

where $J_q(\alpha)$ is the Jacobian of $\mathrm{vec}(q(\alpha)) \in \mathbb{R}^{(n-1)m}$ [3], with respect to $\mathrm{vec}(\alpha) \in \mathbb{R}^{(n-1)m}$. It turns out that $J_q(\alpha)$ is invertible for all $\alpha \in \mathbb{R}^{n\times m}$(see Section 6.2), and therefore, the above linear-system has an exact solution.

Now, let us consider the *optimal shift problem*. Its objective is non-convex (see Figure 4 in the Supplementary File). $\nabla\mathcal{L}(\alpha)$ is a linear combination of $\nabla q_{ij}(\alpha)$ for all $i \in [n]$ and $j \in [m]$. Since $J_q(\alpha)$ is invertible, its rows $\{\nabla q_{ij}(\alpha)\}$, are linearly independent, and the gradient is never zero unless we are at the global minimum where $\alpha = q^{-1}(\delta)$. This guarantees that the objective does not have a saddle-point or local-maximum, and that any local-minimum is a global minimum. Using this we can develop an efficient algorithm to solve the *optimal coverage problem* (Lemma 3).

This brings us to our main algorithmic result, which is an algorithm to find the optimal allocation rule for the *infinite-dimensional fair advertising problem*.

Theorem 3. ***(An algorithm to solve the infinite-dimensional fair advertising problem).*** *There is an algorithm (Algorithm 1) which outputs $\alpha \in \mathbb{R}^{n\times m}$ such that if assumptions (17), (18), (19), and (20) are satisfied, the α-shifted mechanism (8) achieves a revenue ε-close to the optimal for the infinite-dimensional fair advertising problem in*

$$\widetilde{O}\left(\frac{n^7 \log m}{\varepsilon^2}\frac{(\mu_{\max}\rho)^2}{(\mu_{\min}\eta)^4}(L + n^2\mu_{\max}^2)\right) \text{ steps.}$$

Where the arithmetic calculations in each step are bounded by calculating $\nabla\mathrm{rev}$ *once and* $\widetilde{O}$ *hides* $\log$ *factors in* $n,\ \rho,\ \eta,\ \mu_{\max},\ {}^1\!/\!_\varepsilon$ *and* ${}^1\!/\!_{\mu_{\min}}$.

Roughly, the above algorithm has a convergence rate of $\widetilde{O}({}^1\!/\!_{\varepsilon^2})$, under the assumptions which we list below.

Assumptions

For all $i \in [n],\ j \in [m],$ and $y_1, y_2 \in \mathrm{supp}(f_{ij})$

$q_{ij} > \eta$	(η-coverage, 17)
$\mu_{\min} \le f_{ij}(y_1) \le \mu_{\max}$	(Distributed distribution, 18)
$\lvert f_{ij}(y_1) - f_{ij}(y_2)\rvert < L\lvert y_1 - y_2\rvert$	(Lipschitz distribution, 19)
$\lvert\mathbb{E}[\phi_{ij}]\rvert = \left\lvert \int_{\mathrm{supp}(f_{ij})} z f_{ij}(z)dz \right\rvert < \rho.$	(Bounded bid, 20)

Assumption (17) guarantees that all advertisers have at least an η probability of winning on every type, assumption (18) places lower and upper bounds on the probability density functions of the ϕ_{ij}, assumption (19) guarantees that the probability density functions of the ϕ_{ij} are L-Lipschitz continuous, and assumption (20) assumes that the expected ϕ_{ij} is bounded.

[3] $\mathrm{vec}(\cdot)$ represents the vectorization operator.

Algorithm 1 Algorithm1$(\mathcal{Q}, G, L, \eta, \mu_{\max}, \mu_{\min}, \varepsilon)$

Input: Constraint polytope $\mathcal{Q} \subseteq [0,1]^{n\times m}$, Lipschitz constant $G > 0$ of $\mathrm{rev}(\cdot)$, Lipschitz constant $L > 0$ of $f_{ij}(\cdot)$, minimum coverage $\eta > 0$, lower and upper bounds, $\mu_{\min}$ and $\mu_{\max}$ of $f_{ij}(\cdot)$, and a constant $\varepsilon > 0$.
Output: Shifts $\alpha \in \mathbb{R}^{n\times m}$ for the optimal mechanism.

```
Initialize γ:=ε/2G², ξ:=(Gγ)², T=(√2G/ε)²
Compute q_1 := proj_Q(q(0_{n×m}))
Compute α_1 := Algorithm2(q_t, α_t, ξ, L, η, μ_max, μ_min)
for t = 1,2,...,T do
    Compute J_q(α_t)
    Compute rev(q_t) from
        J_q(α_t)^T ∇rev(q_t) := ∇rev_shift(α_t)
    Update q_{t+1} := proj_Q(q_t + γ∇rev(q_t))
    Update
        α_{t+1} := Algorithm2(q_t, α_t, ξ, L, η, μ_max, μ_min)
end for
return α
```

We expect Assumptions (17) and (20) to hold in any real-world setting. We can drop the lower bound in Assumption (18) by introducing "jumps" in α to avoid ranges where the measure of bids is small. Removing assumption (19) would be an interesting direction for future work.

Remark 4. *We inherit the assumption of independent and regular distributions from Myerson. In addition, we require the the distributions of valuations are strictly regular to guarantee that ties between advertisers happen with* 0 *probability. We can drop this assumption by incorporating a randomized tie-breaking rule which retains fairness. The above allocation rule is monotone and allocates the ad spot to the bidder with the highest shifted valuation $\phi_{ij} + \alpha_{ij}$ for a given user. Thus, it defines a unique truthful mechanism and corresponding payment rule.*

5. Our Algorithm

Algorithm 1 performs a projected gradient descent to find the optimal $q^\star \in \mathcal{Q}$ (11). It starts with an initial coverage $q_1 \in \mathcal{Q}$, and the corresponding shift $\alpha_1 = q^{-1}(q_1)$. At step k, it calculates the gradient $\nabla\mathrm{rev}(q_k)$, by solving the linear-system in Eq. (16). To solve this linear-system, we need to calculate $J_q(\alpha_k)^\top$ and $\nabla\mathrm{rev}_{\mathrm{shift}}(\alpha_k)$. This can be done in $O(n^2m)$ steps if we have $\alpha_k = q^{-1}(q_k)$ (see Remark 6). Therefore, the algorithm requires a "good" approximation of α at each step, it maintains this by "updating" the previous approximation α_{k-1} using Algorithm 2 to approximately solve the optimal-shift problem (15).

After calculating $\nabla\mathrm{rev}(q_k)$, it takes a gradient step and projects the current iterate on $\mathcal{Q}$ in $O((nm)^\omega)$ time (Section 6.1), where ω is the fast matrix multiplication coefficient. It takes roughly $O({}^1\!/\!_{\varepsilon^2})$ steps to obtain an ε-accurate solution, and then returns its current shift $\alpha \approx \alpha^\star$. We can bound the error introduced by the approximation of α_k at each step by ensuring that Algorithm 2 has sufficient accuracy. In particular, if it is $O(\varepsilon^2)$ accurate we can prove that Algorithm 1 converges in $\widetilde{O}(1/\varepsilon^2)$ steps.

6. Proof Overviews

6.1. Projection on Constraint Polytope ($\mathcal{Q}$)

Given any point $q \in [0,1]^{n\times m}$, by determining the constraints it violates, we can express the projection on the constraint polytope $\mathcal{Q}$, as a quadratic program with equality constraints. Using this we can construct a projection oracle $\mathrm{proj}_{\mathcal{Q}}$, which given a point $q \in [0,1]^{n\times m}$ projects it onto $\mathcal{Q}$ in $O((nm)^\omega)$ arithmetic operations, where ω is the fast matrix multiplication coefficient.

6.2. Calculating and Bounding $\nabla\mathrm{rev}(\cdot)$

We fix the shift of one advertiser $i \in [n]$ for each type $j \in [m]$. Then, to obtain $\nabla\mathrm{rev}(q)$ we use the fact that $J_q(\alpha)$ is always invertible (Lemma 1). Given $\alpha = q^{-1}(\delta)$ for some $\delta \in [0,1]^{n\times m}$, we can calculate $\nabla\mathrm{rev}(\delta)$ by solving the linear-system in Eq. (16).

Remark 5. *$J_q(\alpha)$ is invertible iff we fix the shift α_{ij} of one advertiser $i \in [n]$ for each type $j \in [m]$. Intuitively, if we increase the α_{ij} for all $i \in [n]$ and $j \in [m]$ by the same amount, then q remains invariant. This implies that each row of $J_q(\alpha)$ has 0 sum, or that $J_q(\alpha)$ is not invertible.*

Lemma 1. ***(Jacobian is invertible).*** *For all $\alpha \in \mathbb{R}^{(n-1)\times m}$, if all advertisers have non-zero coverage for all types $j \in [m]$, then $J_q(\alpha) \in \mathbb{R}^{(n-1)\cdot m\times(n-1)\cdot m}$ is invertible.*
See Section D.4 in the Supplementary File for the proof.

Remark 6. *For all $i, s \in [n]$ such that $i \neq s$, q_{ij} is independent of α_{st}. Therefore, that Jacobian $J_q(\alpha)$ is sparse. and the linear-system in Eq. (16) can be solved in $O(n^\omega m)$ steps, where ω is the fast matrix multiplication coefficient.*
In order to get a complexity bound for Algorithm 1, we show that the Frobenius norm of $\nabla\mathrm{rev}$ is bounded. The proof is presented in Section D.5 in the Supplementary File.

Lemma 2. ***(Revenue is Lipschitz).*** *For all coverages $q_1, q_2 \in \mathcal{Q}$, if assumptions (17), (18) and (20) are satisfied, then*

$$|\mathrm{rev}(q_1) - \mathrm{rev}(q_2)| \leq (\mu_{\max}\rho/\mu_{\min}\eta)n^2\|q_1 - q_2\|_F{}^{4}. \quad (21)$$

6.3. An Algorithm to Solve the Optimal Shift Problem

Lemma 3. ***(An algorithm to solve the optimal shift problem).*** *There is an algorithm (Algorithm 2) which outputs $\alpha \in \mathbb{R}^{n\times m}$ such that if assumptions (17), (18) and (19) are satisfied, then α is an ε-optimal solution for the optimal shift problem, i.e., $\mathcal{L}(\alpha) < \varepsilon$, in*

$$\log\left(m\mathcal{L}(\alpha_1)/\varepsilon\right)\cdot n^3(L + n^2\mu_{\max}^2)(\eta\mu_{\min})^{-2} \text{ steps.}$$

Where the arithmetic operations in each step are bounded by calculating $\nabla\mathcal{L}$ once.

[4] We use $\|\cdot\|_F$ to denote the Frobenius norm.

Algorithm 2 Algorithm2($\delta, \alpha_1, \xi, L, \eta, \mu_{\max}, \mu_{\min}$)

Input: Target coverage $\delta \in [0,1]^{n\times m}$, approximate shift α_1, a constant $\xi > 0$ that controls the accuracy, Lipschitz constant $L > 0$ of $f_{ij}(\cdot)$, minimum coverage $\eta > 0$, and lower and upper bounds, $\mu_{\min}$ and $\mu_{\max}$, of $f_{ij}(\cdot)$.
Output: Approximation $\alpha \in \mathbb{R}^{n\times m}$ of shifts for δ.

Initialize $\gamma := (4nL + 2n^3\mu_{\max}^2)^{-1}$
Initialize
$T := \log\left(mn^3\mathcal{L}(\alpha_1)/\varepsilon\right)\frac{(L+n^2\mu_{\max}^2)}{(\eta\mu_{\min})^2}$
for t = 1,2,...,T **do**
  Compute $\nabla L(\alpha_t) := \nabla[\mathcal{L}_1(\alpha_t), \ldots, \mathcal{L}_m(\alpha_t)]$
  Update $\alpha_{t+1} := \alpha_t - \gamma\nabla L(\alpha_t)$
end for
return α

We give an overview of the proof here, and defer the details to Section D.6 in the Supplementary File.

First, we observe that $\nabla\mathcal{L} = 2\sum_{i,j}(q_{ij}(\alpha) - \delta_{ij})\nabla q_{ij}(\alpha)$ is a linear combination of the rows, $\{\nabla q_{ij}\}$, of $J_q(\alpha)$. Since $\{\nabla q_{ij}(\alpha)\}$ are linearly independent, $\nabla\mathcal{L} \neq 0$ unless we are at the global minimum of $\mathcal{L}$, where $\delta = q(\alpha)$. This guarantees that $\mathcal{L}$ does not have any saddle-points or local-maxima, and that any local minimum is a global minimum.

Then, to get an efficient complexity with a gradient-based algorithm we want to avoid small gradients "far" from the optimal. We show that if $\mathcal{L}(\alpha) \geq \varepsilon$, then the Frobenius norm $\|\nabla\mathcal{L}(\alpha)\|_F \geq \sqrt{\varepsilon}$ (Lemma 6 in the Supplementary File).

Finally, we show that the gradient, $\nabla\mathcal{L}(\alpha)$ is $O(n(L + n^2\mu_{\max}^2))$-Lipschitz continuous (Lemma 7 in the Supplementary File). Therefore, at each step where $\mathcal{L}(\alpha) \geq \xi$, we improve the loss by a factor of $1 - \beta\xi$, where β does not depend on ξ. This gives us a complexity bound of $O(\log 1/\varepsilon)$.

6.4. Proof of Theorem 3

Starting from $q_0 \in \mathcal{Q}$, Algorithm 1 performs a projected gradient descent on $\mathcal{Q}$. Since $\mathcal{Q}$ is convex, the projection is contractive. In particular, for the optimal $q^\star \in \mathcal{Q}$

$$\forall\, q,\ \|\mathrm{proj}_{\mathcal{Q}}(q) - q^\star\|_2 \leq \|q - q^\star\|_2. \quad (22)$$

It queries the shift $\alpha_k \approx q^{-1}(q_k)$ from Algorithm 2 at each step. This introduces some error $\xi > 0$ at each step, which we fix later in the proof.

Let $z_{k+1} = q_k + \gamma\nabla\mathrm{rev}(q_k)$ be the coverage at the $k+1$-th gradient-step, and $q_{k+1} = q(\alpha_{k+1})$ be the coverage obtained by querying $\alpha_k \approx q^{-1}(\mathrm{proj}_{\mathcal{Q}}(q_{k+1}))$. Then, we have the following bound on the error

$$\|\mathrm{proj}_{\mathcal{Q}}(z_{k+1}) - q_{k+1}\|_2^2 \leq \xi. \quad \text{(Error from Algorithm2, 23)}$$

We know that $\mathrm{rev}(\cdot)$ is a concave function of q. Using the first-order condition of concavity at $q^\star$ and q_k we have

$$\|z_{k+1} - q^\star\|_2^2 = \|q_k + \gamma\nabla\mathrm{rev}(q_k) - q^\star\|_2^2$$
$$\leq \|q_k - q^\star\|_2^2 + 2\gamma(\mathrm{rev}(q_k) - \mathrm{rev}(q^\star)) + \gamma^2\|\nabla\mathrm{rev}(q_k)\|_2^2 \quad (24)$$

Using the triangle inequality with Eq. (22) and (23) we get

$$\|q_{k+1}-q^\star\|_2^2 = \|q_{k+1}-\mathrm{proj}_{\mathcal{Q}}(z_{k+1})+\mathrm{proj}_{\mathcal{Q}}(z_{k+1})-q^\star\|_2^2$$
$$\leq \|z_{k+1}-q^\star\|_2^2+\xi \quad (25)$$
$$\overset{(24)}{\leq} \|q_k-q^\star\|_2^2+2\gamma(\mathrm{rev}(q_k)-\mathrm{rev}(q^\star))+\gamma^2\|\nabla \mathrm{rev}(q_k)\|_2^2+\xi \quad (26)$$

Expanding the above recurrence we get

$$\|q_{k+1}-q^\star\|_2^2 \overset{(26)}{\leq} k\xi+\|q_1-q^\star\|_2^2+\sum_{i=1}^k \gamma^2\|\nabla \mathrm{rev}(q_i)\|_2^2 + 2\sum_{i=1}^k \gamma(\mathrm{rev}(q_i)-\mathrm{rev}(q^\star)). \quad (27)$$

Substituting $\|q_{k+1}-q^\star\|_2^2 \geq 0$, and $\|q_1-q^\star\|_2^2 \leq 1$ we get

$$k\xi+1+2\sum_{i\in[k]}\gamma(\mathrm{rev}(q_i)-\mathrm{rev}(q^\star))+\sum_{i\in[k]}\gamma^2\|\nabla \mathrm{rev}(q_i)\|_2^2 \geq 0.$$

Replacing $\mathrm{rev}(q_i)$ by its maximum, choosing $\xi := G^2\gamma^2$, and using $\|\nabla \mathrm{rev}(q_i)\|_2 \leq G$ and $k := \left(\sqrt{2}G/\varepsilon\right)^2$ we get

$$\mathrm{rev}(q^\star) - \max_{i\in[k]}\left(\mathrm{rev}(x_i)\right) \leq \frac{1+k\xi+G^2\sum_{i\in[k]}\gamma^2}{2\sum_{i\in[k]}\gamma} \leq \varepsilon$$

At each step we perform a small update to q_k and query α_k, therefore, Algorithm 2 is always warm-started, i.e., $\|z_{k+1}-q_k\|_2^2 < G\gamma$. Now, from Lemma 3 the total steps required to update α are

$$\sum_{i=1}^k \log\left(mG\gamma/\xi\right)\cdot n^3(L+n^2\mu_{\max}^2)(\eta\mu_{\min})^{-2}$$
$$= (\sqrt{2}G/\varepsilon)^2 \log\left(2mG/\varepsilon\right)\cdot n^3(L+n^2\mu_{\max}^2)(\eta\mu_{\min})^{-2}$$

The sum of the total gradient steps by Algorithm 1, and the total gradient steps by all calls of Algorithm 2 is

$$O\left(G^2/\varepsilon^2 \log\left(2mG/\varepsilon\right)\cdot n^3(L+n^2\mu_{\max}^2)(\eta\mu_{\min})^{-2}\right).$$

Using $G = (\mu_{\max}\rho/\mu_{\min}\eta)\cdot n^2$ (from Lemma 2) we have that Algorithm 1 gets an ε-approximation of optimal revenue in

$$\widetilde{O}\left(\frac{n^7\log m}{\varepsilon^2}\frac{(\mu_{\max}\rho)^2}{(\mu_{\min}\eta)^4}(L+n^2\mu_{\max}^2)\right) \text{ steps.}$$

Where $\widetilde{O}$ hides log factors in $n, \rho, \eta, \mu_{\max}, 1/\varepsilon$ and $1/\mu_{\min}$. ■

7. Empirical Study

We evaluate our approach empirically on the Yahoo! A1 dataset (Yahoo). We vary the strength of the fairness constraint for all advertisers, find an optimal fair mechanism $\mathcal{F}$ using Algorithm 1 and compare it against the optimal unconstrained (and hence potentially unfair) mechanism $\mathcal{M}$, which is given by Myerson (Myerson, 1981). We first consider the impact of the fairness constraints on the revenue of the platform. Let $\mathrm{rev}_{\mathcal{N}}$ denote the revenue of mechanism $\mathcal{N}$. We report the revenue ratio $\kappa_{\mathcal{M},\mathcal{F}} := \mathrm{rev}_{\mathcal{F}}/\mathrm{rev}_{\mathcal{M}}$. Note that the revenue of $\mathcal{F}$ can be at most that of $\mathcal{M}$, as it solves a constrained version of the same problem; thus $\kappa_{\mathcal{M},\mathcal{F}} \in [0,1]$. We then consider the impact of the fairness constraints on the advertisers. Towards this, we consider the distribution of winners among advertisers in an auction given by $\mathcal{M}$ and an auction given by $\mathcal{F}$. We report the total variation distance $d_{TV}(\mathcal{M},\mathcal{F}) := 1/2\sum_{i=1}^n |\sum_{j=1}^m q_{ij}(\mathcal{M}) - q_{ij}(\mathcal{F})| \in [0,1]$ between the two distributions, as a measure of how much the winning distribution changes due to the fairness constraints. Lastly, we consider the fairness of the resultant mechanism $\mathcal{F}$. To this end, we measure selection lift (slift) achieved by $\mathcal{F}$, $\mathrm{slift}(\mathcal{F}) := \min_{i,j}(q_{ij}/1-q_{ij}) \in [0,1]$. Where $\mathrm{slift}(\mathcal{F}) = 1$, represents perfect fairness among the two user types.

7.1. Dataset

We use the Yahoo! A1 dataset (Yahoo), which contains bids placed by advertisers on the top 1000 keywords on *Yahoo! Online Auctions* between June 15, 2002 and June 14, 2003. The dataset has 10475 advertisers, and each advertiser places bids on a subset of keywords; there are approximately $2\cdot 10^7$ bids in the dataset.

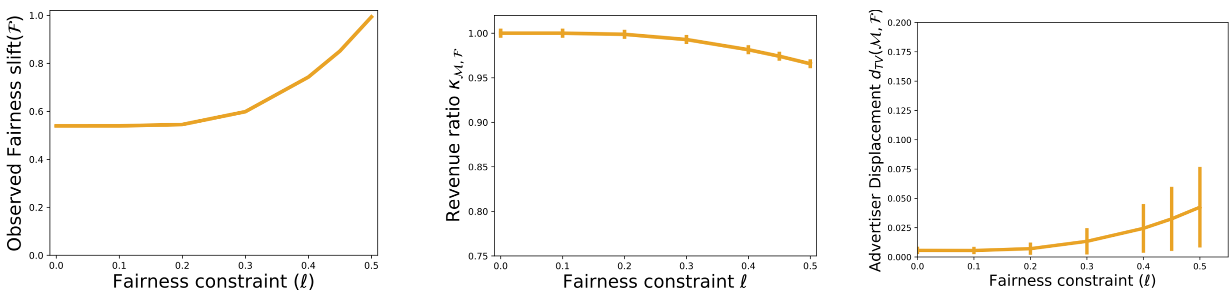
For each keyword k, let A_k be the set of advertisers that bid on it. We infer the distribution of valuation of an advertiser for a keyword by the bids they place on the keyword. In order to retain sufficiently rich valuation profiles for each advertiser, we remove advertisers who place less than 1000 bids on k or whose valuations have variance lower than $3\cdot 10^{-3}$ from A_k, and then those who win the auction less than 5% of the time. This retains more than $1.5\cdot 10^7$ bids.

The actual keywords in the dataset are anonymized; hence, in order to determine whether two keywords k_1 and k_2 are related, we consider whether they share more that one advertiser, i.e., $A_{k_1}\cap A_{k_2} > 1$. This allows us to identify keywords that are related (see Figure 3 in the Supplementary File), and hence for which spillover effects may be present as described in (Lambrecht & Tucker, 2018). Drawing that analogy, one can think of each keyword in the pair as a different type of user for which the same advertisers are competing, and the goal would be for the advertiser to win an equal proportion of each user.

There are 14,380 such pairs. However, we observe that spillover does not affect all keyword pairs (see Figure 2 in the Supplementary File). To test the effect of imposing fairness constrains in a challenging setting, we consider only the auctions which are not already fair; in particular there are 3282 keyword pairs which are less than $\ell = 0.3$ fair.

7.2. Experimental Setup

As we only consider pairs of keywords in this experiment, a lower bound constraint $\ell_{11} = \delta$ is equivalent to an upper bound constraint $u_{12} = 1-\delta$. Hence, it suffices to consider lower bound constraints. We set $\ell_{i1} = \ell_{i2} = \ell\ \forall\, i \in [2]$, and vary ℓ uniformly from 0 to 0.5 , i.e., from the completely unconstrained case (which is equivalent to Myerson's action) to completely constrained case (which requires each advertiser to win each keywords in the pair with exactly



(a) *Fairness.* We report the fairness $\mathrm{slift}(\mathcal{F})$ achieved by our fair $(\mathcal{F})$ mechanism for varying level of fairness.

(b) *Fairness and Revenue.* We report the revenue ratio $\mathrm{rev}_{\mathcal{M},\mathcal{F}}$ between the fair $(\mathcal{F})$ and the unconstrained $(\mathcal{M})$ mechanisms.

(c) *Effect of fairness on advertisers.* We report the $d_{TV}(\mathcal{M},\mathcal{F})$ of the distribution of winners in ads allocated by $\mathcal{F}$ and $\mathcal{M}$.

Figure 1. The x-axis represents fairness constraint ℓ (lower bound). Error bars represent the standard error of the mean.

the same probability). We report $\kappa_{\mathcal{N},\mathcal{M}}$, $d_{TV}(\mathcal{N},\mathcal{M})$, and $\mathrm{slift}(\mathcal{F})$ averaged over all auctions after 10^4 iterations in Figure 1; error bars represent the standard error of the mean over 10^4 iterations and 3282 auctions respectively.

Remark 7. *Computationally, we could consider more types (m). The bottleneck is empirical; whether the dataset contains enough keywords with m overlapping advertisers for the experiment to be meaningful. For $m < 7$ we get over 1000 such keywords sets, and observe results similar to $m = 2$ case, losing less than 5% of the revenue with a TV-distance smaller than 0.05 even for the setting with $\ell = 0.5$.*

7.3. Empirical Results

Fairness. Since the auctions are unbalanced to begin with, we expect the selection lift to increase with the fairness constraint. We observe a growing trend in the selection lift, eventually achieving perfect fairness for $\ell = 0.5$.

Revenue Ratio. We do not expect to outperform the optimal unconstrained mechanism. However, we observe that even in the perfectly balanced setting with $\ell = 0.5$ our mechanisms lose less than 5% of the revenue.

Advertiser Displacement. Since the auctions are unbalanced to begin with, we expect TV-distance to grow with the fairness constraint. We observe this growing trend in the TV-distance on lowering the risk-difference. Even for zero risk-difference ($\ell = 0.5$) our mechanisms obtain a TV-distance smaller than 0.05.

8. Limitations and Future Work

This work leaves several interesting directions open. On the technical side, it would be interesting to improve Theorem 3 by weakening the assumptions on the distributions, or by deriving better complexity bounds in terms of ε or n. Although our algorithm works for intersectional types, it considers a separate constraint for each intersection. Since there can be exponentially many intersections compared to the types, it would be important to improve the run-time in this setting. Exploring the utility lost from the advertiser's perspective, and potential ways of bounding it would also be of interest. Further, it would be relevant to extend our framework to the (non-truthful) general second price auction (Edelman et al., 2007; Varian, 2009), which is used to auction multiple ad slots together.

From a practical standpoint, a natural problem is that advertisers run their campaigns at different times; while an ad campaign is running on the platform, several other campaigns start and finish. Our framework does not account for this. Further, we do not ensure that users of different types derive similar value from an ad. An advertiser could intentionally design an ad to appeal to a specific type, and then, even though the ad receives a balanced coverage, it could generate biased value for users (Speicher et al., 2018).

Finally on the empirical side, testing our framework in the field and studying how the constraints affect user satisfaction, and the profile of ads they see would be important.

9. Conclusion

We initiate a formal study of designing mechanisms for online ad auctions that can ensure advertisements are not shown disproportionately to different populations. This is especially relevant for ads for employment opportunities, housing, and other regulated markets where biases in the recipient population can be illegal and/or unethical. As has been shown recently, existing platforms suffer from various spillover effects that result in such biased distributions. Our approach places constraints on the allocations achieved by an ad across different sub-populations in order to ensure balanced exposure of the content. It can be used flexibly placing constraints on some or all advertisers, across some or all sub-populations, and varying the tightness of the constraint depending on the level of fairness desired.

We present a truthful mechanism which attains the optimal revenue while satisfying the constraints necessary to attain such fairness, and present an efficient algorithm for finding this mechanism given the advertiser properties and fairness constraints. Empirically, we observe that our mechanisms can satisfy fairness constraints at a minor loss to the revenue of the platform, even when the constraints ensure it attains perfect fairness. Hence, fairness is not necessarily at odds with maximizing the platform's ad revenue. Furthermore, we show empirically that advertisers are not significantly impacted with respect to their winning percentages – the sub-populations their ads are shown to change to be fair, but overall they are still reach a similar number of users.

References

About Ad Auctions. Facebook Ads. http://bit.ly/2FLMLiT.

Ad targeting. Google Ads, About the Ad Auction. http://bit.ly/2L9IFFC.

Ali, M., Sapiezynski, P., Bogen, M., Korolova, A., Mislove, A., and Rieke, A. Discrimination Through Optimization: How Facebook's Ad Delivery Can Lead to Skewed Outcomes, 2019.

Alphabet Inc. Online Advertisements Contribute 85% to Alphabet INC.'s Total Revenue. Source: Alphabet Inc. Q3 2018 Results. http://bit.ly/2UKSjD8.

Angwin, J. and Parris Jr., T. Facebook Lets Advertisers Exclude Users by Race. *Propublica*, October 2016. http://bit.ly/2DueClN.

Asudeh, A., Jagadish, H. V., Stoyanovich, J., and Das, G. Designing Fair Ranking Schemes. *CoRR*, abs/1712.09752, 2017.

Balseiro, S. R., Feldman, J., Mirrokni, V. S., and Muthukrishnan, S. Yield Optimization of Display Advertising with Ad Exchange. *Management Science*, 60(12):2886–2907, 2014. doi: 10.1287/mnsc.2014.2017. URL https://doi.org/10.1287/mnsc.2014.2017.

Celis, L. E. and Vishnoi, N. K. Fair Personalization. *CoRR*, abs/1707.02260, 2017. URL http://arxiv.org/abs/1707.02260.

Celis, L. E., Huang, L., and Vishnoi, N. K. Multiwinner Voting with Fairness Constraints. In *IJCAI*, pp. 144–151. ijcai.org, 2018a.

Celis, L. E., Keswani, V., Straszak, D., Deshpande, A., Kathuria, T., and Vishnoi, N. K. Fair and Diverse DPP-Based Data Summarization. In *ICML*, volume 80 of *Proceedings of Machine Learning Research*, pp. 715–724. PMLR, 2018b.

Celis, L. E., Straszak, D., and Vishnoi, N. K. Ranking with Fairness Constraints. In *ICALP*, volume 107 of *LIPIcs*, pp. 28:1–28:15. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2018c.

Celis, L. E., Huang, L., Keswani, V., and Vishnoi, N. K. Classification with Fairness Constraints: A Meta-Algorithm with Provable Guarantees. In *FAT*, pp. 319–328. ACM, 2019a.

Celis, L. E., Kapoor, S., Salehi, F., Keswani, V., and Vishnoi, N. K. A Dashboard for Controlling Polarization in Personalization. *AI Commun.*, 32(1):77–89, 2019b.

Celis, L. E., Kapoor, S., Salehi, F., and Vishnoi, N. K. Controlling Polarization in Personalization: An Algorithmic Framework. In *FAT*, pp. 160–169. ACM, 2019c.

Clarke, E. H. Multipart Pricing of Public Goods. *Public Choice*, 11(1):17–33, Sep 1971. ISSN 1573-7101. doi: 10.1007/BF01726210. URL https://doi.org/10.1007/BF01726210.

Datta, A., Tschantz, M. C., and Datta, A. Automated Experiments on Ad Privacy Settings. *PoPETs*, 2015(1):92–112, 2015.

Discriminatory practices. Facebook: Advertising Policies, Discriminatory Practices. http://bit.ly/2MorDAM.

Dwork, C. and Ilvento, C. Fairness Under Composition. In *ITCS*, volume 124 of *LIPIcs*, pp. 33:1–33:20. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2019.

Edelman, B., Ostrovsky, M., and Schwarz, M. Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars Worth of Keywords. *American economic review*, 97(1): 242–259, 2007.

Facebook, Ad Targeting. Facebook Ads. http://bit.ly/2R28y7Y.

Facebook, Bid Strategies. Your Guide to Facebook Bid Strategies. http://bit.ly/2ZYDUC7.

Farahat, A. How Effective Is Targeted Advertising? In *ACC*, pp. 6014–6021. IEEE, 2013.

Fox-Brewster, T. Creepy or Cool? Twitter Is Tracking Where You've Been, What You like and Is Telling Advertisers. *Forbes Magazine*, May 2017. http://bit.ly/2vfCllp.

Garimella, K., Morales, G. D. F., Gionis, A., and Mathioudakis, M. Reducing Controversy by Connecting Opposing Views. In *IJCAI*, pp. 5249–5253. ijcai.org, 2018.

Ghosh, A., McAfee, R. P., Papineni, K., and Vassilvitskii, S. Bidding for Representative Allocations for Display Advertising. In *WINE*, volume 5929 of *Lecture Notes in Computer Science*, pp. 208–219. Springer, 2009.

Goldfarb, A. and Tucker, C. Online Display Advertising: Targeting and Obtrusiveness. *Marketing Science*, 30(3):389–404, May 2011. ISSN 1526-548X. doi: 10.1287/mksc.1100.0583. URL https://doi.org/10.1287/mksc.1100.0583.

Google, Ad Rank. Ad Rank Thresholds: Definition. http://bit.ly/2IzDY5N.

Google, Automated Bidding. About Automated Bidding. http://bit.ly/2Gu2rpE.

Groves, T. Incentives in Teams. *Econometrica*, 41(4):617–31, 1973.

Guynn, J. Facebook Accused of Bias in Job Ads That Let Companies Target Men, Exclude Women. *USA Today*, September 2018. http://bit.ly/2FPPxUm.

Hartline, J. D. *Mechanism Design and Approximation*. July 2017. Book draft. http://jasonhartline.com/MDnA/.

Huang, L. and Vishnoi, N. K. Stable and Fair Classification. 81, 2019.

Jones, E., Oliphant, T., Peterson, P., et al. SciPy: Open Source Scientific Tools for Python, 2001–. http://www.scipy.org/.

Lambrecht, A. and Tucker, C. E. Algorithmic bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of STEM Career Ads. March 2018.

Muthukrishnan, S. Ad Exchanges: Research Issues. In *WINE*, volume 5929 of *Lecture Notes in Computer Science*, pp. 1–12. Springer, 2009.

Myerson, R. B. Incentive Compatibility and the Bargaining Problem. *Econometrica*, 47(1):61–73, January 1979.

Myerson, R. B. Optimal Auction Design. *Mathematics of Operations Research*, 6(1):58–73, 1981.

NFHA. Facebook Sued by Civil Rights Groups for Discrimination in Online Housing Advertisements. *National Fair Housing Alliance*, March 2018. http://bit.ly/2Uab7XM.

Nisan, N., Roughgarden, T., Tardos, E., and Vazirani, V. V. *Algorithmic Game Theory*. Cambridge University Press, New York, NY, USA, 2007. ISBN 0521872820.

Ostrovsky, M. and Schwarz, M. Reserve Prices in Internet Advertising Auctions: A Field Experiment. *EC*, 11:59–60, 2011.

Pai, M. M. and Vohra, R. Auction Design with Fairness Concerns: Subsidies vs. Set-Asides. Technical report, Discussion Paper, Center for Mathematical Studies in Economics and and Management Science, 2012.

Panigrahi, D., Sarma, A. D., Aggarwal, G., and Tomkins, A. Online Selection of Diverse Results. In *WSDM*, pp. 263–272. ACM, 2012.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

Radlinski, F., Kleinberg, R., and Joachims, T. Learning Diverse Rankings with Multi-Armed Bandits. In *ICML*, volume 307 of *ACM International Conference Proceeding Series*, pp. 784–791. ACM, 2008.

Speicher, T., Ali, M., Venkatadri, G., Ribeiro, F. N., Arvanitakis, G., Benevenuto, F., Gummadi, K. P., Loiseau, P., and Mislove, A. Potential for Discrimination in Online Targeted Advertising. In *FAT*, volume 81 of *Proceedings of Machine Learning Research*, pp. 5–19. PMLR, 2018.

Sweeney, L. Discrimination in Online Ad Delivery. *Commun. ACM*, 56(5):44–54, 2013.

Timberg, C. and Jan, T. HUD Secretary Carson accuses Facebook of enabling housing discrimination. *The Washington Post*, August 2018. http://bit.ly/2sEB99V.

Varga, R. *Gervsgorin and His Circles*. Springer Series in Computational Mathematics. Springer Berlin Heidelberg, 2011. ISBN 9783540211006. URL https://books.google.ch/books?id=q4g0sMe3Ss8C.

Varian, H. R. Position Auctions. *international Journal of industrial Organization*, 25(6):1163–1178, 2007.

Varian, H. R. Online Ad Auctions. *American Economic Review*, 99(2):430–34, 2009.

Vermeren, I. Men vs. Women: Who Is More Active on Social Media? *Brandwatch*, January 2015. http://bit.ly/2Xvy6xY.

Vickrey, W. Counterspeculation, Auctions, and Competitive Sealed Tenders. *The Journal of Finance*, 16(1):8–37, 1961. ISSN 00221082, 15406261. URL http://www.jstor.org/stable/2977633.

Yahoo. A1 - Yahoo! Search Marketing Advertising Bidding Data, Version 1.0. http://bit.ly/2VX90Ib.

Yan, J., Liu, N., Wang, G., Zhang, W., Jiang, Y., and Chen, Z. How Much Can Behavioral Targeting Help Online Advertising? In *WWW*, pp. 261–270. ACM, 2009.

Yuan, S., Abidin, A. Z., Sloan, M., and Wang, J. Internet Advertising: An Interplay among Advertisers, Online Publishers, Ad Exchanges and Web Users. *CoRR*, abs/1206.1754, 2012. URL http://arxiv.org/abs/1206.1754.

Zafar, M. B., Valera, I., Gomez-Rodriguez, M., and Gummadi, K. P. Fairness Constraints: Mechanisms for Fair Classification. In *AISTATS*, volume 54 of *Proceedings of Machine Learning Research*, pp. 962–970. PMLR, 2017.