Reserve Price Optimization for First Price Auctions in Display Advertising

Zhe Feng [1] **Sébastien Lahaie** [2] **Jon Schneider** [2] **Jinchao Ye** [2]

Abstract

The display advertising industry has recently transitioned from second- to first-price auctions as its primary mechanism for ad allocation and pricing. In light of this, publishers need to re-evaluate and optimize their auction parameters, notably reserve prices. In this paper, we propose a gradient-based algorithm to adaptively update and optimize reserve prices based on estimates of bidders' responsiveness to experimental shocks in reserves. Our key innovation is to draw on the inherent structure of the revenue objective in order to reduce the variance of gradient estimates and improve convergence rates in both theory and practice. We show that revenue in a first-price auction can be usefully decomposed into a *demand* component and a *bidding* component, and introduce techniques to reduce the variance of each component. We characterize the bias-variance trade-offs of these techniques and validate the performance of our proposed algorithm through experiments on synthetic data and real display ad auctions data from a major ad exchange.

1. Introduction

A reserve price in an auction specifies a minimum acceptable winning bid, below which the item remains with the seller. The reserve price may correspond to some outside offer, or the value of the item to the seller itself, and more generally may be set to maximize expected revenue (Myerson, 1981). In a data-rich environment like online advertising auctions it becomes possible to learn a revenue-optimal reserve price over time, and there is a substantial literature on optimizing reserve prices for *second-price* auctions, which have been commonly used to allocate ad space (Paes Leme et al., 2016; Mohri & Medina, 2016; Munoz & Vassilvitskii, 2017).

In this work we examine the problem of reserve price optimization in *first-price* (i.e., pay-your-bid) auctions, motivated by the fact that all the major ad exchanges have recently transitioned to this auction format as their main ad allocation mechanism (Chen, 2017; Bigler, 2019). First-price auctions have grown in favor because they are considered more transparent, in the sense that there is no uncertainty in the final price upon winning (Benes, 2017).[1] Unless restrictive assumptions are met, there is in theory no revenue ranking between first- and second-price auctions (Krishna, 2009), and there is no guarantee that reserve prices optimized for second-price auctions will continue to be effective in a first-price setting.

From a learning standpoint the shift from second- to first-price auctions introduces several new challenges. In a second-price auction, truthful bidding is a dominant strategy no matter what the reserve. The bidders' value distributions are therefore readily available, and bids stay static (in principle) as the reserve is varied. In a first-price auction, in contrast, bidders have an incentive to shade their values when placing their bids, and bid-shading strategies can vary by bidder. The gain from setting a reserve price now comes if (and only if) it induces higher bidding, so an understanding of bidder responsiveness becomes crucial to setting effective reserves.

Bid adjustments in response to a reserve price can occur at different timescales. If a bidder observes that it wins too few auctions because of the reserve price, it may increase its bid in the long-term (in a matter of hours up to weeks). Our focus here is on setting reserves prices by taking into account *immediate* bidder responses to reserves. We assume that each bidder has a fixed, unknown bidding function $b(r, v)$ that depends on its private value v and the observed auction reserve r. This agrees with practice in display ad auctions because the reserve r is normally sent out in the 'bid request' message to potential bidders (IAB, 2016). To the extent that the bid function responds to r, first-price reserves can show an immediate positive effect on revenue.

Our Results. We propose a gradient-based approach to adaptively improve and optimize reserve prices, where we

[1]Harvard University, this work was done when the first author was an intern in Google Inc, NYC. [2]Google Inc, NYC. Correspondence to: Zhe Feng <zhe_feng@g.harvard.edu>.

Proceedings of the 38^{th} International Conference on Machine Learning, PMLR 139, 2021.

[1]The full reasons for the transition are complex, and include the rise of "header bidding" (Weiss, 2019). A header bidding auction is a first-price auction usually triggered by code in a webpage header (hence the name).

perturb current reserves upwards and downwards (e.g., by 10%) on random slices of traffic to obtain gradient estimates.

Our key innovation is to draw on the inherent structure of the revenue objective in order to reduce the variance of gradient estimates and improve convergence rates in both theory and practice. We show that revenue in a first-price auction can be usefully decomposed into two terms: a *demand curve* component which depends only on the bidder's value distribution; and a *bidding* component whose variance can be reduced based on natural assumptions on bidding functions.

A demand curve is a simpler, more structured object than the original revenue objective (e.g., it is downward-sloping), so the demand component lends itself to parametric modeling to reduce the variance. We offer two variance reduction techniques for the bidding component,[2] referred to as *bid truncation* and *quantile truncation*. Bid truncation can strictly decrease variance with no additional bias assuming the right bidding model (perfect response model), whereas quantile truncation may introduce bias but is less sensitive to assumptions on the bidding model. Indeed, the quantile truncation variance reduction method works for a very general class of bidding strategies that satisfy a *diminishing sensitivity property*; this class includes best-response bidding for m i.i.d bidders with uniform value distribution and a *near* perfect response model defined in Section 2.

We evaluate our approach over synthetic data where bidder values are drawn uniformly, and also over real bid distributions collected from the logs of the major ad exchange. Our experimental results confirm that the combination of variance reduction on both objective components leads to the fastest convergence rate. For the demand component, a simple logistic model works well over the synthetic (i.e., uniform) data, but a flexible neural net is needed over the semi-synthetic data. For the bidding component, we find that quantile truncation is much more robust to assumptions on the bidding model.

Related Work. This paper connects with the rich literature on *reserve price optimization for auctions*, e.g., (Myerson, 1981; Riley et al., 1981). How to set optimal reserve prices in *second price auctions* based on access to bidders' historical bid data has been an increasingly popular research direction in machine learning, e.g., (Ostrovsky & Schwarz, 2011; Mohri & Medina, 2016; Munoz & Vassilvitskii, 2017). Another related line of work uses no-regret learning in second price auctions with partial information feedback to optimize reserve prices, e.g., (Blum et al., 2003; Cesa-Bianchi et al., 2015). All of the works cited so far rely on the fact that the seller can directly learn the valuation distribution from historical bid data, since the second price auction is truthful.

For first-price auctions, we have found little work on setting optimal reserves for asymmetric bidders, since there are no characterizations of equilibrium strategies for this case. Results are only available for limited environments, such as bidders with uniform valuation distributions (Krishna, 2009; Matthews, 1995). Recently, there has been a line of work regarding *revenue optimization against strategic bidders in repeated auctions*, e.g., (Amin et al., 2013; Huang et al., 2018; Drutsa, 2020). In this paper, instead of assuming that bidders act strategically, we assume each bidder has a fixed bidding function in response to reserves. This is a common assumption in large market settings and in the dynamic pricing literature (Mao et al., 2018).

The algorithms developed in this paper are related to the literature on *online convex optimization with bandit feedback* (Flaxman et al., 2005; Hazan & Levy, 2014; Agarwal et al., 2010; 2011). However, there are two key differences with our work: (1) the revenue function in a first price auction is non-convex, and (2) the seller cannot obtain perfect revenue feedback under perturbed reserves with just a single query (i.e., auction)—the seller needs multiple queries to achieve accurate estimates with high confidence. Our algorithm is also related to *zeroth-order stochastic gradient methods* (Ghadimi & Lan, 2013; Balasubramanian & Ghadimi, 2018; Ghadimi, 2019; Liu et al., 2018), which we discuss in detail later in Section 3.

2. Preliminaries

We consider a setting where a seller sells a single item to a set of m bidders via a first price auction. In such an auction, the seller first sends out a reserve price r to all bidders. Each bidder i then submits a bid b_i. The bidder with the highest bid larger than r wins the item and pays their bid; if no bidder bids above r, the item goes unallocated. Note that the type of reserve price we consider in this work is *anonymous* in the sense that each bidder sees the same reserve price.

Each bidder i has a private valuation $v_i \in [0, 1]$ for the item, where each value v_i is drawn independently (but not necessarily identically) from some unknown distribution F_i.[3] With a slight abuse of notation, we write $b_i(r, v_i)$ to denote the bid function of bidder i when the reserve price is r and her value is v_i. In a first-price auction, only the highest bid matters for both allocation and pricing. Given this property, we have the following reduction from multiple bidders to a single "meta-bidder" in a first price auction.

Theorem 2.1. *Let $\mathcal{F}$ be the distribution of $\max(v_1, v_2, \ldots, v_m)$, where each v_i is independently*

[2] Variance reduction of the bidding component relies on the insight that bids far above the reserves are little affected by them (under natural bidding models), so these bids can be filtered out when computing gradient estimates—changes in such bids are likely due to noise rather than any effect of reserves.

[3] The normalization on the valuation domain is without loss of generality: our analysis easily extends to any bounded valuation setting.

drawn from F_i, and let $\mathcal{B}(r)$ be the distribution of $\max(b_1(r, v_1), b_2(r, v_2), \ldots, b_m(r, v_m))$. Then there exists a bid function $b(r, v)$ such that the distribution of $b(r, v)$ when $v \sim \mathcal{F}$ is equal to the distribution $\mathcal{B}(r)$.

Proof. Let $B_r(b)$ be the CDF of $\mathcal{B}(r)$ and $F(v)$ be the CDF of $\mathcal{F}$. Denote $B_r^{-1}(x) = \inf\{b : u \leq B_r(x)\}$ and $F^{-1}(u) = \inf\{x : u \leq F(x)\}$ for any $u \in [0, 1]$. We construct bidding strategy $b(r, v)$,

$$b(r, v) = B_r^{-1}(F(v)). \tag{1}$$

Then we have

$$\begin{aligned} \mathbb{P}_{v\sim\mathcal{F}}\left(b(r, v) \leq b\right) &= \mathbb{P}_{v\sim\mathcal{F}}\left(B_r^{-1}(F(v)) \leq b\right) \\ &= \mathbb{P}_{v\sim\mathcal{F}}\left(v \leq F^{-1}(B_r(b))\right) \\ &= B_r(b) \end{aligned}$$

This guarantees that if $v \sim \mathcal{F}$, then $b(r, v) \sim \mathcal{B}(r)$. □

The above proposition implies that we can without loss of generality focus on a single bidder setting, defined as follows. Let $v = \max_i v_i$ denote the maximum value; v is drawn i.i.d. from an unknown distribution $\mathcal{F}$ across each auction. We write $b(r, v)$ to denote the *maximum* bid when the reserve price is r and the maximum value is v, and $\mathcal{B}(r)$ to denote the distribution of $b(r, v)$ for a fixed r when v is drawn according to $\mathcal{F}$. (Note that the "meta-bidder" may not be one of the n *real* bidders, since the maximum bid can be from a bidder whose value is not maximum.)

The main goal of the seller considered in this work is to learn the optimal reserve price $r \in [0, 1]$ that maximizes expected revenue:

$$\mathbb{E}_{v\sim\mathcal{F}}\left[b(r, v) \cdot \mathbb{I}\{b(r, v) \geq r\}\right]. \tag{2}$$

Note that there is no reason for a bidder to bid a positive value less than the reserve r: such a bid is guaranteed to lose. Therefore, without loss of generality we can assume that if $b(r, v) < r$, then $b(r, v) = 0$. This allows us to write the revenue simply as:

$$\mu(r) = \mathbb{E}_{b\sim\mathcal{B}(r)}\left[b\right] = \mathbb{E}_{v\sim\mathcal{F}}\left[b(r, v)\right].$$

In this paper, we focus on maximizing the revenue function $\mu(r)$ in the steady state, where $b(\cdot, \cdot)$ and $\mathcal{F}$ are unknown but fixed. This assumption is reasonable in real display ads system where changes to the valuation of an advertiser happen much more gradually than shocks to the market (changes in reserve price, new bidders entering, etc.) and the number of auctions is very large.

Response Models

We begin by describing some general properties of bidding functions that hold for any utility-maximizing bidders; see (Matthews, 1995) for further discussion.

Definition 2.2. *A* bidding function $b(r, v)$ *satisfies the following properties: I.* $b(r, v) \leq v$ *for all* v*; II.* $b(r, v) \geq r$ *for* $v \geq r$*; III.* $b(r, v) = 0$ *for* $v < r$*; IV.* $b(r, v)$ *is non-decreasing in* v *for all* r.

For the "meta-bidder", properties I, II and III in Definition 2.2 hold trivially assuming that all individual bidders are utility maximizing, and property IV holds based on our construction given in Eq. (1). In this paper, we also investigate additional constraints on the response model which, while not a consequence of utility-maximizing behavior, are likely to hold in practice. One such constraint is the *diminishing sensitivity in value of bid to reserve*. This says that *meta-bidder* with a larger value will change its bid less in response to a change in reserves.

Definition 2.3 (Diminishing Sensitivity Property)**.** *If $v_H > v_L$, then for $\delta > 0$ and $v_L \geq r + \delta$ we have $b(r + \delta, v_H) - b(r, v_H) \leq b(r + \delta, v_L) - b(r, v_L)$.*

Indeed, this diminishing sensitivity property of bid to reserve holds in many scenarios. For example, if there are m i.i.d bidders with uniform value distribution, the Bayesian Nash Equilibrium (BNE) bidding strategy satisfies the diminishing sensitivity property. Other response models, such as, no response model (see Appendix B), perfect response model (Definition 2.5) or a mixture of these two models, all satisfy the diminishing sensitivity property. Based on our construction of $b(r, v)$ in Eq. (1), we have the following sufficient condition for the diminishing sensitivity property.

Proposition 2.4. *Let $B_r(b)$ be the CDF of $\mathcal{B}(r)$. If $B_{r+\delta}^{-1}(\cdot) - B_r^{-1}(\cdot)$ is a non-increasing function for any reserve r and $\delta > 0$, then the bid function satisfies diminishing sensitivity property.*

In practice, one natural and concrete example of a response model is a bidder that increases its bid to the reserve as long as the reserve is below its value. We refer to this as the *perfect response model*, formally defined as follows.

Definition 2.5. *A* perfect response *bidding function takes the form:*

$$b(r, v) = \begin{cases} b(0, v) & \text{if } b(0, v) \geq r \\ r & \text{if } b(0, v) < r \leq v \\ 0 & \text{if } v < r \end{cases}$$

Note that the perfect response model is based on the *original bid* of the bidder under reserve price 0, namely $b(0, v)$. If $b(0, v)$ is already above the reserve, then this bidder is unaffected by the reserve. If the value v is larger than r but the original bid is smaller than r, the bidder increases its bid just enough to meet the reserve r. Finally, if the value v is smaller than r, then the bidder submits a bid of 0 in accordance with Definition 2.2 (equivalently, the bidder places an irrelevant bid below the reserve, or simply declines to bid). Note that the perfect response model satisfies the *diminishing sensitivity property*.

In practice, bidders are unlikely to exactly follow the perfect response model; for example, bidders will often increase their bid to some amount strictly above the reserve r so as to remain competitive with other bidders. For this reason, we propose a relaxation of the perfect response model which we call the *ε-bounded response model*: the bid is at most ε greater than what it would have been under the perfect response model if $b(0, v) < r \le v$ (see also Definition A.6). Note that the ε-bounded response model becomes the perfect response model when $\varepsilon = 0$.

Remark. We believe the perfect response bidding strategy is a reasonable response model in practice, however, this paper is not specific to this model. Indeed, our gradient-based algorithm works for all response models, and the *quantile truncation* variance reduction method proposed in section 4 holds for any bidding strategy satisfying diminishing sensitivity property. For the special case of the perfect response model, we design a *bid truncation* variance reduction (in section 4) tailored to this model which significantly improves convergence theoretically and empirically.

3. Gradient Descent Framework

The first-price auction setting introduces several challenges for setting reserve prices. First, the seller cannot observe true bidder values because truthful bidding is not a dominant strategy in a first-price auction. Second, how the bidders will react to different reserves is unknown to the seller—the only information that the seller receives is bids drawn from distribution $\mathcal{B}(r)$ when the seller sets a reserve price r.

One natural idea, and the approach we take in this paper, is to optimize the reserve price via gradient descent. Gradient descent is only guaranteed to converge to the optimal reserve when our objective is convex (or at least, unimodal), which is not necessarily true for an arbitrary revenue function. However, gradient descent has a number of practical advantages for reserve price optimization, including:

1. Gradient descent allows us to incorporate prior information we may have about the location of a good reserve price (possibly significantly reducing the overall search cost).

2. The adaptivity of gradient descent allows us to quickly converge to a local optimum and follow this optimum if it changes over time, significantly saving on search cost (over global methods such as grid search).

3. In practice, many revenue curves have a unique local optimum (see Section 5), so gradient descent is likely to converge to the optimal reserve.

Algorithm 1 Zeroth-order stochastic projected gradient framework for reserve optimization.

Input: Initial reserve $r_1 \in (0, 1)$, and variables to be fixed later: total number of iterations T, perturbation size β_t, learning rate α_t.
Output: Reserve prices $r_2, r_3, \dots, r_{T+1}$.
for $t = 1, 2, \dots, T$ **do**
  Set a reserve price of $r_t^+ = (1 + \beta_t)r_t$ in n_t auctions.
  Set a reserve price of $r_t^- = (1 - \beta_t)r_t$ in n_t auctions.
  Construct an estimate $\hat{G}_t$ of the gradient of revenue at r_t, based on the feedback of experiments.
  Update reserve: $r_{t+1} = \Pi(r_t + \alpha_t \hat{G}_t)$, where

$$\Pi(x) = \underset{z \in (0,1)}{\arg\min} |z - x|.$$

end for

More specifically, since the seller has no direct access to the gradients (i.e, first-order information) of $\mu(r)$, we consider approaches that fit in the framework of *zeroth-order stochastic optimization*. Our framework, summarized in Algorithm 1, proceeds in rounds. In round t where the current reserve is r_t, the seller selects a perturbation size β_t and randomly sets the reserve price to either $(1 + \beta_t)r_t$ or $(1 - \beta_t)r_t$ on separate slices of experiment traffic, until it has received n_t samples from both $\mathcal{B}((1 + \beta_t)r_t)$ and $\mathcal{B}((1 - \beta_t)r_t)$. The seller then uses these $2n_t$ samples to estimate the gradient $\hat{G}_t$ of the revenue curve $\mu(r)$ at r_t and updates the reserve price based on this gradient estimate using learning rate (step size) α_t.

We assume that we have access to a fixed total number of samples $N = \sum_{t=1}^{T} n_t$ (the number of iterations T is a variable that will be fixed later). There is then a trade-off between n_t (i.e, the number of samples per iteration) and T (the number of iterations available to optimize the reserve price).

Zeroth-order stochastic gradient descent is a well-studied problem (Ghadimi & Lan, 2013; Balasubramanian & Ghadimi, 2018; Ghadimi, 2019; Liu et al., 2018). In this paper, we focus on taking advantage of the structure of $b(r, v)$ to construct good *discrete gradient* estimates $\hat{G}_t$, as this aspect is specific to the problem of reserve price optimization. Specifically, we tackle the following problem which we term the *discrete gradient problem*:

- **Input:** n samples $X_1^+, \cdots, X_n^+$ drawn i.i.d from $\mathcal{B}(r^+)$ and n samples $X_1^- \cdots, X_n^-$ drawn i.i.d from $\mathcal{B}(r^-)$, for known $r^+ > r^-$.

- **Output:** An estimator $\hat{G}$ for the discrete derivative $(\mu(r^+) - \mu(r^-))/(r^+ - r^-)$. This estimator has bias $\mathrm{Bias}(\hat{G})$ and variance $\mathrm{Var}(\hat{G})$, where $\mathrm{Bias}(\hat{G}) = \left|\mathbb{E}[\hat{G}] - \frac{\mu(r^+) - \mu(r^-)}{r^+ - r^-}\right|$.

Solutions to the discrete gradient problem with small bias and variance directly translate into faster convergence rates for our gradient descent. We provide a detailed convergence result in Theorem A.2 in Appendix A.1. We summarize this result informally as follows.

Theorem 3.1 (Informal Restatement of Theorem A.2)**.** *If for all t,* $\mathrm{Bias}(\hat{G}_t) \leq B$ *and* $\mathrm{Var}(\hat{G}_t) \leq V$ *then for appropriate choices of* α_t *and* n_t *(and fixing* $\beta_t = \delta/2r_t$*), Algorithm 1 satisfies*

$$\min_{t\in[T]} |\mathcal{P}_{\mathcal{C}}^t|^2 = \tilde{\mathcal{O}}\left(T^{-1/2} + \delta^2 + B^2 + V + (T/N)^2\right).$$

Here $\mathcal{P}_{\mathcal{C}}^t$ *can be thought of as the true gradient at round* t *(see Definition A.1 in Appendix).*

Intuitively, we want to design an estimator and choose our parameters α_t, β_t, n_t, so as to trade off between δ, B, and V. In the following sections, we show how to do this for a variety of bidder response models. The choices of α_t, β_t, n_t are summarized in the full version of the main theorem (Theorem A.2), in Appendix A.1. In this paper, we focus on convergence rate results since the revenue curve may not be concave (w.r.t. reserve). Analyzing the revenue guarantee of the algorithm is an interesting future direction.

Naive Gradient Estimation

The simplest method for estimating the discrete gradient is to take the difference between the average revenue from bids from $\mathcal{B}(r^+)$ and the average revenue from bids from $\mathcal{B}(r^-)$. More formally, we compute discrete gradient as,

$$\hat{G} = \frac{\sum_{i=1}^n X_i^+ - \sum_{i=1}^n X_i^-}{n(r^+ - r^-)}. \tag{3}$$

We show that $\hat{G}$ has the following properties.

Theorem 3.2. *Assume that* $r^+ - r^- = \delta$*, then* $\mathrm{Bias}(\hat{G}) = 0, \mathrm{Var}(\hat{G}) \leq \frac{1}{2\delta^2 n}$.

This leads to the following convergence rate via Theorem 3.1.

Corollary 3.1. *Using this estimator* $\hat{G}$*, and setting* $T = N^{1/2}$ *and* $\delta = \Theta(N^{-1/8})$*, Algorithm 1 achieves convergence,* $\min_{t\in[T]} |\mathcal{P}_{\mathcal{C}}^t|^2 \leq \widetilde{\mathcal{O}}\left(N^{-1/4}\right)$.

Although there are no matching lower bounds, this is the best known asymptotic convergence rate for zeroth-order optimization over a non-convex objective (Ghadimi & Lan, 2013; Balasubramanian & Ghadimi, 2018). The naive gradient estimation approach has the advantage that it works regardless of response model, is simple to compute (it uses only revenue information and not individual bids), and leads to an *unbiased* estimator for the discrete derivative. The disadvantage is that the variance of this estimator can be large (especially as we take δ small). In the following section, we show how to address this by taking into account the inherent structure of the revenue objective based on an underlying bidder response model.

4. Variance Reduced Gradient Estimation

In this section, we first introduce another representation of the revenue formula by decomposing it into a *demand* component and a *bidding* component. We then propose techniques to reduce the variance of the discrete gradient of each component.

4.1. Revenue Decomposition

We can decompose the revenue $\mu(r)$ in the following way.

Theorem 4.1. *We have that*

$$\mu(r) = \mathbb{E}_{v\sim\mathcal{F}}[\max(b(r,v) - r, 0)] + r \Pr_{v\sim\mathcal{F}}[v \geq r]. \tag{4}$$

Define $E(r) = \mathbb{E}_{v\sim\mathcal{F}}[\max(b(r,v) - r, 0)]$ and $D(r) = \Pr_{v\sim\mathcal{F}}[v \geq r]$, so that $\mu(r) = E(r) + rD(r)$. These two terms capture two different aspects of bidder behavior which contribute to revenue. The function $D(r)$ amounts to a "demand curve" which gives the proportion of values that clear the reserve r, and therefore the proportion of auctions that are bid on at r. If the auction were just a simple posted-price auction (i.e., the winner is charged the quoted price r), then the *demand* component $rD(r)$ would be the associated revenue. However, in a first-price auction the winning bidder pays its bid, not the reserve. Therefore the *bidding* component $E(r)$ captures the excess contribution from bids greater than the reserve.

To construct a good estimator $\hat{G}$ for the discrete gradient of $\mu(r)$, it suffices to construct good estimators $\hat{G}_E$ and $\hat{G}_D$ for the discrete gradients of $E(r)$ and $rD(r)$ respectively, and then output $\hat{G} = \hat{G}_E + \hat{G}_D$. Note that $\mathrm{Bias}(\hat{G}) \leq \mathrm{Bias}(\hat{G}_E) + \mathrm{Bias}(\hat{G}_D)$ and $\mathrm{Var}(\hat{G}) \leq 2(\mathrm{Var}(\hat{G}_D) + \mathrm{Var}(\hat{G}_E))$, so it suffices to bound the bias and variance of each component separately.

4.2. Estimating the Demand Component Gradient

We begin by discussing how to estimate the gradient $\hat{G}_D$ of the demand component of revenue. As in Section 3, it is possible to form a naive unbiased estimate of the demand component via the estimator $\hat{D}(r) = \frac{1}{n}\sum_{i=1}^n \mathbb{I}(X_i \geq r)$. The variance of the resulting unbiased estimator $\hat{G}_D$ is then bounded by (see Theorem A.5), $\mathrm{Var}(\hat{G}_D) \leq \frac{(r^+)^2}{2\delta^2 n}$. Note that for small r, the variance guarantee here is significantly better than the variance guarantee in Theorem 3.2. Thus, in instances where the optimal reserve is small (and hence we mostly test small r^+), combining this naive estimator with better estimators for $\hat{G}_E$ (like the ones we explore in the next section) can already lead to better convergence rates overall.

To obtain even better estimators, we can leverage the following two facts about the demand function. First, the demand function only depends on the value distribution $\mathcal{F}$ of the bidders, and not their specific bidding behavior. Since we expect value to be relatively stable in comparison to bidding behavior, this means that we can reasonably use data from previous rounds to learn the demand function and inform calculation of $\hat{G}_D$ (whereas the naive gradient update only uses data from the current round). Second, we expect the demand function $D(r)$ to be simpler and more nicely structured than the full revenue function $\mu(r)$—for example, $D(r)$ is weakly decreasing in r—and therefore more amenable to parametric modeling. Formally speaking, we can estimate $D(r)$ with a parametric function $f_\theta(r)$, and using this approximation to estimate the gradient $\hat{G}_D$. Suppose that we have access to additional historical data $\mathcal{S}$ with which we can fit our parametric class to $D(r)$; let $\hat{\theta}$ be the resulting learned parameter. This learned demand function gives rise to the following estimator $\hat{G}_D$:

$$\hat{G}_D = \frac{r^+ f_{\hat{\theta}}(r^+) - r^- f_{\hat{\theta}}(r^-)}{r^+ - r^-} \tag{5}$$

Note that this decreases overall variance, the variance of $\hat{G}_D$ is 0 because the randomness of $\hat{G}_D$ only comes from historical samples $\mathcal{S}$, which are independent of the samples obtained in the current round, at the cost of a possible increase in bias (due to inaccuracy in estimating $D(r)$).

4.3. Estimating the Bidding Component Gradient

In this section we propose a variance reduction method to achieve a better estimator for $\hat{G}_E$ for a variety of bidder models.

Variance reduction via bid truncation. We first consider the special case of the perfect response (and more generally, the ε-bounded response) bidding model. In the perfect response model, if you were going to bid $b > r^+$ when the reserve was r^+, you will bid the same bid b when the reserve is r^-. This means that large bids (bids larger than r^+) do not contribute in expectation to $\mu(r^+) - \mu(r^-)$, but they do add noise to our gradient estimation. By filtering these out, we can reduce the variance of our estimator while keeping our estimator unbiased.

Since we only apply this filtering when estimating the bidding component $E(r)$ but not the demand component $rD(r)$, we must be careful when implementing this. Note that a large bid $b > r^+$ contributes $b - r^+$ to $E(r^+)$ and $b - r^-$ to $E(r^-)$, and therefore $r^+ - r^-$ to $E(r^+) - E(r^-)$. We can therefore construct an unbiased estimator for $E(r^+) - E(r^-)$ by computing the contribution of unfiltered bids ($b < r^+$) from both $\mathcal{B}(r^+)$ or $\mathcal{B}(r^-)$ and then adding $r^+ - r^-$ for each filtered bid in $\mathcal{B}(r^-)$ (or equivalently, each filtered bid in $\mathcal{B}(r^+)$; under perfect response, the fraction of filtered bids is equal in both models in expectation). Note that every bid from $\mathcal{B}(r^+)$ is either filtered or has excess 0, so we can write this gradient $\hat{G}_E$ entirely in terms of bids from $\mathcal{B}(r^-)$. Formally, we define truncated bid Y_i^- as

$$Y_i^- = \begin{cases} \max(X_i^- - r^-, 0) & \text{if } X_i^- \leq r^+ \\ (r^+ - r^-) & \text{otherwise} \end{cases}$$

Our estimate for the gradient of $E(r)$ is then given by

$$\hat{G}_E = -\frac{\sum_{i=1}^n Y_i^-}{n(r^+ - r^-)} \tag{6}$$

Since any bid in an ε-bounded model only differs from one in the perfect response model by at most ε, we can apply this same estimator to an ε-bounded response model. The following theorem characterizes the bias and variance of the estimator for the ε-bounded response model.

Theorem 4.2. *Assume that $r^+ - r^- = \delta$, then the estimator $\hat{G}_E$ in Eq. (6) for ε-bounded response model, satisfies:* $\mathrm{Bias}(\hat{G}_E) = \frac{2\varepsilon}{\delta}, \mathrm{Var}(\hat{G}_E) \leq \frac{1}{4n}$.

Note that the bias of estimator $\hat{G}_E$ is 0 for the perfect response model. The complete proof is given in Appendix A.5. Combining the above results for $\hat{G}_E$ and $\hat{G}_D$, we have the following improved convergence result for the ε-bounded response model.

Corollary 4.1. *Suppose* $\mathrm{Bias}(\hat{G}_D) \leq \varepsilon_D/\delta$. *Using the estimator $\hat{G}_E$ proposed in Eq. (6) for the ε-bounded response model, setting $T = N^{2/3}$ and $\delta = \Theta(\sqrt{\varepsilon + \varepsilon_D})$, Algorithm 1 achieves convergence,* $\min_{t\in[T]} |\mathcal{P}_{\mathcal{C}}^t|^2 \leq \widetilde{\mathcal{O}}\left(\varepsilon + \varepsilon_D + N^{-1/3}\right)$.

For perfect response bidding models, the above convergence rate is strictly faster than the convergence rate of naive estimator in Corollary 3.1 (state-of-the-art convergence rate for zeroth-order stochastic gradient descent), but with additional bias coming from demand estimation. However, we show this bias has practically negligible effect on the revenue in our experiments.

Variance reduction via quantile truncation. In Eq. (6), we reduced the variance of $\hat{G}_E$ by truncating all bids at the fixed threshold of $t = r^+$. In general, this does not quite work: for bidder response models that are far from perfect response, this truncation can introduce a very large bias. Here we demonstrate one technique for constructing good estimators $\hat{G}_E$ as long as the bidding function $b(r, v)$ possesses diminishing sensitivity in value to reserve.

Instead of truncating in *bid space*, we will instead want to truncate in *value space* to reduce the variance. Specifically, instead of throwing out all bids larger than some threshold t, we will instead throw out all bids whose *corresponding* *values* are larger than some threshold t. This has the nice property that we can write the bias of our resulting estimator $\hat{G}_E$ as $\mathrm{Bias}(\hat{G}_E) = \frac{1}{r^+ - r^-}\int_t^1 (b(r^+, v) - b(r^-, v))d\mathcal{F}(v)$.

The final key observation is that, even though we cannot directly truncate by values, since $b(r, v)$ is monotonically increasing in v, quantiles of bids (e.g., of $\mathcal{B}(r^+)$ and $\mathcal{B}(r^-)$) directly correspond to quantiles of values (of $\mathcal{F}$). Instead of setting a threshold t directly on the value, it is therefore equivalent to truncate at a *fixed quantile of the bid distribution*.

To achieve this, we first sort X_i^+ and X_i^- in ascending order. Then we compute $\hat{G}_E$ as

$$\hat{G}_E = \frac{\sum_{i=1}^{qn} \max(X_i^+ - r^+, 0) - \sum_{i=1}^{qn} \max(X_i^- - r^-, 0)}{n(r^+ - r^-)} - (1-q), \tag{7}$$

where $q \in [0, 1]$ is the *quantile threshold* used to truncate bids. The following theorem characterizes the bias and variance of the above $\hat{G}_E$,

Theorem 4.3. *Let* $r^+ - r^- = \delta$, $t = \mathcal{F}^{-1}(q)$, *and* $\tilde{t} = \mathcal{F}^{-1}(q + n^{-2/3})$. *Then the estimator* $\hat{G}_E$ *in Eq. (7) satisfies,* $\mathrm{Bias}(\hat{G}_E) \le \frac{(1-q)(b(r^+,t) - b(r^-,t))}{\delta} + O(n^{-2/3})$, $\mathrm{Var}(\hat{G}_E) \le \frac{2\tilde{t}^2}{n\delta^2} + O(n^{-5/3}\delta^{-2})$.

Unlike with bid truncation, with quantile truncation we have a clear bias-variance tradeoff as we change q: larger values of q decrease the bias (both by decreasing $(1-q)$ and $b(r^+, t) - b(r^-, t)$, which is decreasing due to diminishing sensitivity) but lead to larger variance. Since one can estimate this bound on the bias (by approximating $b(r^+, t) - b(r^-, t)$ via $Y_{qn}^+ - Y_{qn}^-$), it is possible to choose q to optimize this bias-variance tradeoff as one sees fit (for example, to minimize $B^2 + V$ in Theorem 3.1). We show a convergence rate result for this *quantile truncation* approach in the following Corollary,

Corollary 4.2. *Suppose* $\mathrm{Bias}(\hat{G}_D) \le \varepsilon_D/\delta$. *Using the estimator* $\hat{G}_E$ *proposed in Eq. (6) for the response model with diminishing sensitivity property, for any fixed quantile q, setting* $T = N^{2/3}$ *and* $\delta = \Theta(\sqrt{\varepsilon_D + 1 - q})$, *Algorithm 1 achieves convergence,*

$$\min_{t\in[T]} |\mathcal{P}_C^t|^2 \le \widetilde{\mathcal{O}}\left(\varepsilon_D + 1 - q + \left(1 + \frac{\mathcal{F}^{-1}(q + N^{-2/9})}{\varepsilon_D + 1 - q}\right)N^{-\frac{1}{3}}\right)$$

5. Experiments

We evaluate the performance of our algorithms on synthetic and semi-synthetic data sets. Due to space limitations, we present the complete experimental results in Appendix B.

5.1. Data Generation

The data generation process consists of two parts: a base bid distribution specifying the distribution of *maximum* bid when no reserve is set, and a response model describing how a "meta-bidder" (see Section 2) with bid b (under no reserve) would update its bid in response to a reserve of r.

Response models. We assume that in the absence of a reserve bidders bid a constant fraction γ of their value v (i.e., $b = \gamma v$), which we refer to as *linear shading*. We consider linear shading combined with perfect response and with ϵ-bounded response, which we implement by adding a uniform $[0, \epsilon]$ random variable to the bid. We also examine equilibrium bidding for m i.i.d. bidders with uniformly distributed valuation (Krishna, 2009): for each bidder $i \in [m]$, $b_i = \frac{r^m + (m-1)v_i^m}{mv_i^{m-1}}$.

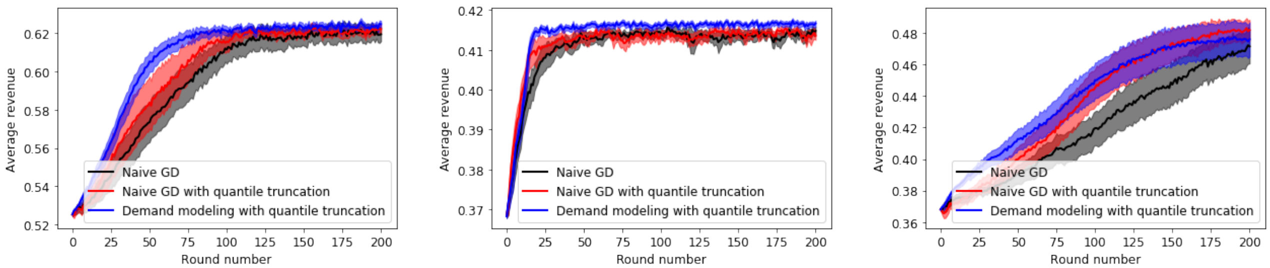
Synthetic data. In our synthetic data sets, the base (maximum) bid distribution is the uniform $[0, 1]$ distribution for perfect response model and ε-bounded response model. In the simulations, we apply a constant shading factor of 0.4 for the perfect response model and ε-bounded response model. For equilibrium bidding, we assume that each auction contains $m = 2$ bidders, thus the base (maximum) bid distribution is the distribution of $\max\{\frac{v_1}{2}, \frac{v_2}{2}\}$, where $v_1, v_2 \sim U[0, 1]$.

Semi-synthetic data. For our semi-synthetic data sets, we separately collected the empirical distributions of winning bids over one day for 20 large publishers on a major display ad exchange. Each distribution was filtered for outliers and normalized to the interval $[0, 1]$. For this semi-synthetic data we only test the perfect-response model and ε-bounded response model, since there is no closed-form solution for the equilibrium bidding strategy. We use 0.3 as the constant shading factor for semi-synthetic data.

5.2. Methodology

Gradient descent algorithms. We examine five different gradient descent algorithms: (I) *Naive GD*: naive gradient descent using the gradient estimator in Eq. (3); (II) *Naive GD with bid truncation*: gradient descent using the gradient estimator in Eq. (6) for the bidding component, and a naive estimate[4] of the demand component; (III) *Naive GD with quantile truncation*: gradient descent using the gradient estimator in Eq. (7) for the bidding component, and naive estimate of the demand component; (IV) *Demand modeling with bid truncation*: Same as the second variant, but with a parametric model of the demand curve to estimate demand component of gradient; (V) *Demand modeling with quantile truncation*: Same as the third variant, but with a parametric model of the demand curve to estimate demand component of gradient. The parameters used in these algorithms are specified in Appendix B.

[4]We can form a naive unbiased estimator $\hat{G}_D = \frac{\hat{D}(r^+) - \hat{D}(r^-)}{r^+ - r^-}$, where $\hat{D}(r^+) = \frac{1}{n}\sum_i \mathbb{I}\{x_i^+ \ge r^+\}$ and similarly for $\hat{D}(r^-)$.



(a) Synthetic data with perfect response. **(b)** Synthetic data with equilibrium response. **(c)** Semi-synthetic data with perfect response.

Figure 1: Revenue as a function of round t for (a) synthetic data with perfect response, (b) synthetic data with equilibrium response, and (c) semi-synthetic data with perfect response.

Demand curve estimation. To reduce variance following the ideas of Section 4, we need a model $\hat{G}_D$ for the demand component of the discrete gradient. Instead of estimating $\hat{G}_D$ from historical data, we adaptively learn the demand curve during the training process. Concretely, at each round t, we observe new (reserve, demand) pairs from $2n_t$ samples and retrain our demand curve using all the samples observed up to the current round. We use this trained demand curve to compute $\hat{G}_D$ based on (5). For the synthetic data, a simple logistic regression can effectively learn the demand curve. However, the semi-synthetic data required a more flexible model so for this case we model demand using a fully connected neural network with 1 hidden layer, 15 hidden nodes and ReLU activations.

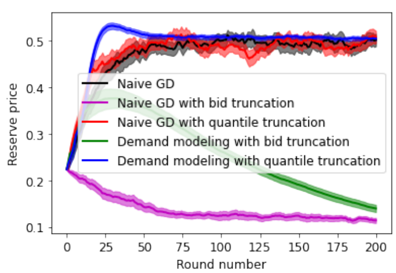


Figure 2: Reserve price as a function of round t for synthetic data with equilibrium response.

5.3. Evaluation

Effectiveness of gradient descent. First, we confirm that gradient descent can effectively find optimal reserves in our models. For each semi-synthetic model, we construct the revenue curve as a function of reserve with assumed response models. We find that 19 out of the 20 revenue curves have a clear single local maximum (the remaining curve has 2). In all cases (synthetic and semi-synthetic models), the revenue learned by the naive gradient descent algorithm is at least 95% of the revenue at the optimal reserve, which indicates that gradient descent can efficiently find the optimal reserve in these cases despite the lack of convexity.

Effectiveness of variance reduction methods. We first evaluate the performance of the quantile-based variance reduction method. We run the algorithm variants (I), (III) and (V) under synthetic data and semi-synthetic data with multiple bidder response models. Figures (1a) and (1c) show the revenue achieved by the three algorithms over time under the perfect response model. We find that quantile-based variance reduction leads to a more stable training process which converges faster than naive gradient descent. Figure (1b) evaluates the performance of the three algorithm variants under synthetic data and an equilibrium response model, with similar conclusions. Overall, quantile-based variance reduction outperforms naive gradient descent. Moreover, with the addition of demand curve estimation, algorithm variant (V) achieves better revenue and converges to an optimal reserve faster than the other two algorithms, in agreement with our theoretical guarantees.

We next consider variance reduction using bid truncation, which is used in algorithm variants (II) and (IV). Bid truncation is tailored to perfect response and performs very well for this response model, in accordance with the theoretical guarantees, but quantile truncation is competitive and often performs as well over the semi-synthetic data (see Appendix B for a detailed comparison). Under the equilibrium response model, bid truncation can in fact hinder the training process and lead to a substantially suboptimal reserve price (see Figure 2). In summary, quantile-based variance reduction coupled with a good demand-curve estimation is the method of choice to achieve good reserve prices under a range of different bid distributions and bidder response models.

6. Conclusions

In this paper, we propose a gradient-based algorithm to adaptively update and optimize reserve prices in first price auctions, based on estimates of bidders' responsiveness to experimental shocks in reserves. For a broad class of bidder response strategies that satisfy a natural diminishing sensitivity property, we obtain convergence rates that strictly improve over state-of-the-art algorithms for zeroth-order optimization (which do not take into account the specific structure of the revenue objective). For future work, we plan to investigate other techniques for variance reduction used in zeroth-order optimization, such as kernel methods, and to obtain results for more general settings where value distributions may drift over time.

Acknowledgments

Zhe Feng is supported by a Google PhD fellowship. We thank Saeed Ghadimi for informing us on the state of the art in zeroth-order methods, and we thank Nitish Korula and Renato Paes Leme for insightful discussions on the paper. We also thank the anonymous reviewers for their constructive feedback.

References

Agarwal, A., Dekel, O., and Xiao, L. Optimal algorithms for online convex optimization with multi-point bandit feedback. In *Proceedings of the 23rd Annual Conference on Learning Theory (COLT)*, June 2010.

Agarwal, A., Foster, D. P., Hsu, D. J., Kakade, S. M., and Rakhlin, A. Stochastic convex optimization with bandit feedback. In Shawe-Taylor, J., Zemel, R. S., Bartlett, P. L., Pereira, F., and Weinberger, K. Q. (eds.), *Advances in Neural Information Processing Systems 24*, pp. 1035–1043, 2011.

Amin, K., Rostamizadeh, A., and Syed, U. Learning prices for repeated auctions with strategic buyers. In *Advances in Neural Information Processing Systems 26*, 2013.

Balasubramanian, K. and Ghadimi, S. Zeroth-order nonconvex stochastic optimization: Handling constraints, high-dimensionality and saddle-points, 2018.

Benes, R. How SSPs use deceptive price floors to squeeze ad buyers. https://digiday.com/marketing/ssps-use-deceptive-price-floors-squeeze-ad-buyers/, September 2017. Accessed: 2020-01-29.

Bigler, J. Rolling out first price auctions to google ad manager partners digiday. https://www.blog.google/products/admanager/rolling-out-first-price-auctions-google-ad-manager-partners, September 2019. Accessed: 2020-01-27.

Blum, A., Kumar, V., Rudra, A., and Wu, F. Online learning in online auctions. *Theor. Comput. Sci.*, 324:137–146, 2003.

Boucheron, S., Thomas, M., et al. Concentration inequalities for order statistics. *Electronic Communications in Probability*, 17, 2012.

Cesa-Bianchi, N., Gentile, C., and Mansour, Y. Regret minimization for reserve prices in second-price auctions. *IEEE Transactions on Information Theory*, 61(1):549–564, Jan 2015.

Chen, Y. Programmatic advertising is preparing for the first-price auction era. https://digiday.com/marketing/programmatic- advertising-readying-first-price-auction-era, October 2017. Accessed: 2020-01-29.

Drutsa, A. Reserve pricing in repeated second-price auctions with strategic bidders. *Proceedings of the 37th International Conference on Machine Learning*, 2020.

Flaxman, A. D., Kalai, A. T., Kalai, A. T., and McMahan, H. B. Online convex optimization in the bandit setting: Gradient descent without a gradient. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '05, pp. 385–394, Philadelphia, PA, USA, 2005.

Gentle, J. E. *Computational statistics*, volume 308. Springer, 2009.

Ghadimi, S. Conditional gradient type methods for composite nonlinear and stochastic optimization. *Mathematical Programming*, 173(1):431–464, Jan 2019.

Ghadimi, S. and Lan, G. Stochastic first- and zeroth-order methods for nonconvex stochastic programming. *SIAM Journal on Optimization*, 23(4):2341–2368, 2013.

Ghadimi, S., Lan, G., and Zhang, H. Mini-batch stochastic approximation methods for nonconvex stochastic composite optimization. *Mathematical Programming*, 155: 267–305, 2013.

Hazan, E. and Levy, K. Y. Bandit convex optimization: Towards tight bounds. In *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 1*, NIPS'14, pp. 784–792, 2014.

Huang, Z., Liu, J., and Wang, X. Learning optimal reserve price against non-myopic bidders. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, 2018.

IAB. OpenRTB API Specification, Version 2.5. www.iab.com/wp-content/uploads/2016/03/OpenRTB- API- Specification -Version-2-5-FINAL.pdf, December 2016. Accessed: 2020-01-26.

Krishna, V. *Auction theory*. Academic press, 2009.

Liu, S., Li, X., Chen, P., Haupt, J., and Amini, L. Zeroth-order stochastic projected gradient descent for nonconvex optimization. In *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 1179–1183, Nov 2018.

Mao, J., Leme, R., and Schneider, J. Contextual pricing for lipschitz buyers. In *Advances in Neural Information Processing Systems 31*, pp. 5643–5651, 2018.

Matthews, S. A. A Technical Primer on Auction Theory I: Independent Private Values. Discussion Papers 1096, Northwestern University, Center for Mathematical Studies in Economics and Management Science, May 1995.

Mohri, M. and Medina, A. M. n. Learning algorithms for second-price auctions with reserve. *J. Mach. Learn. Res.*, 17(1):2632–2656, January 2016.

Munoz, A. and Vassilvitskii, S. Revenue optimization with approximate bid predictions. In *Advances in Neural Information Processing Systems 30*, pp. 1858–1866, 2017.

Myerson, R. Optimal auction design. *Mathematics of Operations Research*, 6:58–73, 1981.

Ostrovsky, M. and Schwarz, M. Reserve prices in internet advertising auctions: A field experiment. In *Proceedings of the 12th ACM Conference on Electronic Commerce*, EC '11, pp. 59–60, 2011.

Paes Leme, R., Pal, M., and Vassilvitskii, S. A field guide to personalized reserve prices. In *Proceedings of the 25th international conference on world wide web*, pp. 1093–1102. International World Wide Web Conferences Steering Committee, 2016.

Reddi, S. J., Sra, S., Póczos, B., and Smola, A. J. Proximal stochastic methods for nonsmooth nonconvex finite-sum optimization. In *NIPS*, 2016.

Riley, J. G., Samuelson, W. F., et al. Optimal auctions. *American Economic Review*, 71(3):381–392, 1981.

Weiss, M. Digiday research: Header bidding and first-price auctions boost publisher revenues. https://digiday.com/media/digiday-research-header-bidding-and-first-price-auctions-boost-publisher-revenues, January 2019. Accessed: 2020-01-29.