

# Yield Optimization of Display Advertising with Ad Exchange

Santiago Balseiro <sup>\*</sup>   Jon Feldman <sup>†</sup>   Vahab Mirrokni <sup>‡</sup>   S. Muthukrishnan <sup>§</sup>

February 15, 2011

**Abstract.** In light of the growing market of Ad Exchanges for the real-time sale of advertising slots, publishers face new challenges in choosing between the allocation of contract-based reservation ads and spot market ads. In this setting, the publisher should take into account the tradeoff between short-term revenue from an Ad Exchange and the long-term benefits of delivering good quality spots to the reservation ads. In this paper, we formalize this combined optimization problem as a stochastic control problem and derive an efficient policy for online ad allocation in settings with general joint distribution over placement quality and exchange bids. We prove asymptotic optimality of this policy in terms of any trade-off between quality of delivered reservation ads and revenue from the exchange, and provide a rigorous bound for its convergence rate to the optimal policy. We also give experimental results on data derived from real publisher inventory, showing that our policy can achieve any Pareto-optimal point on the quality vs. revenue curve. Finally, we study a parametric training-based algorithm in which instead of learning the dual variables from a data sample (as is done in non-parametric training-based algorithms), we learn the parameters of the distribution and construct those dual variables from the learned parameter values. We compare parametric and non-parametric ways to estimate from data both analytically and experimentally in the special case without the ad exchange, and show that though both methods converge to the optimal policy as the sample size grows, our parametric method converges faster, and thus performs better on smaller samples.

## 1 Introduction

Ad Exchanges like RightMedia, AdECN or DoubleClick are an emerging market for the real-time sale of online ad slots on the Internet. While exchanges differ in their implementations, in a generic Ad Exchange (AdX) [17], publishers post an ad slot with a reservation price, advertisers post bids, and an auction is run; this happens between the time a user visits a page and the ad is displayed. Thus AdX implements a spot market for display ads.

From the publisher’s point of view, the goal is to maximize revenue gained from sale of the ad inventory. In isolation, this problem is similar to optimal auction design, and Myerson [18] shows that the optimal reserve price can be calculated from the bid distribution. However, AdX represents only one possible sales outlet for the publisher. In fact, even in the presence of this efficient real-time market, many web publishers still get a significant portion of their revenue from making individual deals with specific advertisers or agencies. These deals usually take the form of a specific number of ad impressions reserved over a particular time horizon (e.g., one million impressions over a month). A publisher can make many such deals with different advertisers, with potentially sophisticated relationships between the advertisers’ targeting criteria. So for a given piece of ad inventory, the publisher must quickly decide whether to send the inventory to AdX (and at what price), or to assign it to an advertiser with a reservation.

In this paper, we study the problem faced by the publisher, jointly optimizing over AdX and the reservations. We bring to bear techniques from revenue management and stochastic optimization, and derive an efficient policy for making real-time ad allocation decisions. We prove that our policy is asymptotically optimal in terms of an arbitrary (i.e., publisher-defined) trade-off between quality delivered to reservation ads and revenue from the exchange. Our policy and analysis is quite general, and works for any joint distribution over placement quality and exchange bids, even allowing correlation between advertisers, or between quality and exchange bids. In particular, we provide a rigorous bound on the convergence rate of our policy to the optimal policy (Theorem 2). Typically ad allocation research compares to the optimal offline policy in hindsight; instead, we compare our policy with an optimal online policy, obtaining a bound on regret, as in online machine learning. We also give experimental results on data derived from real publisher inventory, showing that our policy can achieve any Pareto-optimal point on the quality vs. revenue curve.

A special case of our model is one without the ad exchange, where the publisher needs only to choose be-

<sup>\*</sup>Columbia University, New York, NY.

<sup>†</sup>Google, Inc., New York, NY.

<sup>‡</sup>Google, Inc., New York, NY.

<sup>§</sup>Google, Inc., New York, NY.

tween reservation ads, maximizing quality while respecting reservations. A well-studied technique for this problem in the algorithms literature is to solve an optimization problem on a data sample, then use the dual variables online for the remaining inventory. Interestingly, when one specializes our policy to the case without an exchange, a similar dual program emerges; however in our parametric model, the dual variables are derived algebraically from the parameters of the distribution rather than from the data. Of course in practice, parameters must be learned from data; thus a fair comparison to the known sampling-based methods is to allow our policy to learn its parameters from the same-sized sample. We perform this comparison both analytically and experimentally, and show that though both methods converge to the optimal policy as the sample size grows, our parametric method converges faster, and thus performs better on smaller samples. Our analysis may be of independent interest as a general comparison of parametric vs. non-parametric estimators of distribution quantiles.

## 2 Model

Consider a publisher displaying ads in a web page. The publisher signs contracts with a set  $A$  of advertisers guaranteeing them a certain number of targeted impressions within a given time horizon. We denote by  $\mathcal{A} = \{1, \dots, A\}$  the set of advertisers. We assume that a total of  $N$  users arrive in the time horizon. The web page has a single slot for ads, and each user is shown at most one impression per page.

Users have different characteristics, and depending on their profile they may be more or less attractive for different advertisers. We assume that the  $n$ -th impression is endowed with a vector of qualities  $Q_n = \{Q_{n,a}\}_{a \in \mathcal{A}}$ , where  $Q_{n,a}$  is the predicted quality advertiser  $a$  would perceive if the impression is assigned to her. Furthermore, qualities  $\{Q_n\}_{n=1, \dots, N}$  are random, independent, and identically distributed; however, we do allow qualities to be jointly distributed across advertisers. This captures the fact that advertisers might have similar target criteria, and hence the qualities perceived might be correlated. Qualities are, hence, drawn independently from the same joint c.d.f  $G(\cdot)$ . We do not impose any further restrictions on the qualities, other than finite second moments.

The publisher has agreed to deliver exactly  $C_a$  impressions to advertiser  $a \in \mathcal{A}$ ; neither over-delivery nor under-delivery is allowed. We denote by  $\rho = \{\rho_a\}_{a \in \mathcal{A}}$  with  $\rho_a = \frac{C_a}{N}$ , the capacity to impression ratio of each advertiser. Note that a necessary condition for the feasibility of the operation is that the number of arriving impressions suffices to satisfy the contracts, or  $\sum_{a \in \mathcal{A}} \rho_a \leq 1$ . An

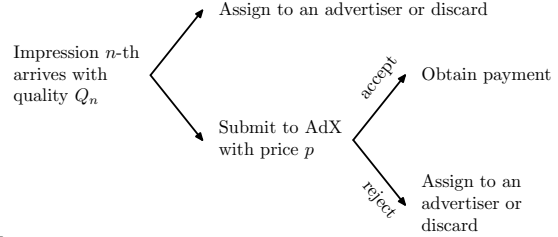


Figure 1: Publisher’s decision tree for a new impression.

assumption of this general model is that any user can be potentially assigned to any advertiser. In practice each advertiser may be interested in a particular group of user types. It is important to note that this is not a limitation of our results, but rather a modeling choice; in §F.4 we show how to handle targeting criteria by setting  $Q_{n,a} = -\tau_a$  for impressions not matching the targeting criteria of an advertiser. This can also be interpreted as forcing the publisher to pay a penalty to the advertisers each time an undesired impression is incorrectly assigned.

Arriving impressions may either be assigned to the advertisers, discarded or auctioned in the Ad Exchange (AdX) for profit. In AdX [17], the publisher contacts the exchange with a minimum price she is willing to take for the slot. Additionally, the publisher may submit some partial information of the user visiting the website. For simplicity, we first assume that no information about the user is revealed. However, in section F.3 we relax this assumption. Internally the exchange contacts different ad networks, and in turn they return bids for the slot. The exchange determines the winning bid among those that exceed the reserve price via an auction, and returns a payment to the publisher. In this case we say that the impressions is *accepted*, and the publisher is contractually obligated to display the winning impression. In the case that no bid attains the reserve price, no payment is made and the impression is *rejected*. We present the formal model of the exchange in Section 2.2. The entire operation above is executed before the page is rendered in the user’s screen. Thus, in the event that the impression is rejected by the exchange, the publisher may still be able to assign it to some advertiser. Figure 1 summarizes the decisions involved.

For notational simplicity we extend the set of advertisers to  $\mathcal{A}_0 = \{0\} \cup \mathcal{A}$  by including an outside option 0 that represents discarding an impression. We set the quality of the outside option identically to zero, i.e.  $Q_{n,0} = 0$  for all impressions  $n = 1, \dots, N$ . In the following, the terms discarding an impression or assigning it to advertiser 0 are used interchangeably. We set  $\rho_0 = 1 - \sum_{a \in \mathcal{A}} \rho_a$  to be the fraction of impressions that are not assigned to any advertiser. To wit, a fraction of the  $\rho_0$  impressions will be assigned to the winning impression of AdX, and the

remainder effectively discarded.

**2.1 Objective** The publisher’s problem is to maximize the overall placement quality of the impressions assigned to the advertisers together with the total revenue obtained with AdX, while complying with the contractual obligations. Note that the objectives are potentially conflicting; in the short-term, the publisher might boost the revenue stream from AdX at the expense of assigning lower quality impressions to the advertisers. In the long term, however, it may be convenient for the publisher to prioritize her advertisers, in view of attracting future contracts.

We attack the multi-objective problem by taking a weighted sum of both objectives. The publisher has at her disposal a parameter  $\gamma$ , which allows her to trade-off between these conflicting objectives. The aggregated objective is given by

$$\text{yield} = \text{revenue}(\text{AdX}) + \gamma \cdot \text{quality}(\text{advertisers}),$$

Hence, by choosing a suitable large  $\gamma$  the advertisers may focus on assigning high quality impressions to the advertisers; while a small  $\gamma$  would prioritize the revenue from AdX. Without loss of generality, we set  $\gamma = 1$  for the remainder of this paper, except when noted otherwise.

**2.2 AdX Model** The publisher submits an impression to AdX with the minimum price it is willing to take, denoted by  $p \geq 0$ . The impression is accepted if there is a bid of value  $p$  or more. We denote by  $B$  the winning bid random variable. In the following we assume that bids are independent of the quality of the impression, and identically distributed according to a c.d.f.  $F(\cdot)$ . Hence, the impression is accepted with probability  $1 - F(p) = \bar{F}(p)$ . In this first model, when the impression is accepted, the publisher is paid the minimum price  $p$ . In Sections F.2 and F.3 we drop this assumption and consider a more general second-price auction with side information.

Suppose the publisher has computed an *opportunity cost*  $c$  for selling this inventory in the exchange; that is, the publisher stands to gain  $c$  if the impression is given to a reservation advertiser.

Given opportunity cost  $c \geq 0$  the publisher picks the price that maximizes its expected revenue. Hence, the publisher solves the optimization problem  $R(c) = \max_{p \geq 0} \bar{F}(p)p + F(p)c$ . Changing variables, we can define  $r(s) = s\bar{F}^{-1}(s)$  to be the expected revenue under acceptance probability  $s$ , and rewrite this as

$$R(c) = \max_{s \in [0,1]} r(s) + (1 - s)c. \quad (1)$$

Also, let  $s^*(c) = \min \{\arg \max_{0 \leq s \leq 1} r(s) + (1 - s)c\}$

be the least maximizer of (1), and  $p^*(c) = \bar{F}^{-1}(s^*(c))$  be the price that verifies the maximum.

**Assumption 1** *The expected revenue under survival probability  $s$  is continuous, concave, non-negative, bounded, and satisfies  $\lim_{s \rightarrow 0} r(s) = 0$ . We call a function  $r(s)$  that satisfies all of the assumptions above a regular revenue function.*

These assumptions are common in RM literature (see, e.g., [9]). A sufficient condition for the concavity of the revenue is that  $B$  has increasing generalized failure rates [13]. Regularity implies, among other things, the existence of a *null price*  $p_\infty$  such that  $\lim_{p \rightarrow p_\infty} \bar{F}(p)p = 0$ .

**Proposition 1** *Suppose that  $r(s)$  is regular revenue function. Then,  $R(c)$  is non-decreasing, convex, continuous, and  $R(c) \geq c$ . Additionally,  $R(c) - c$  is non-increasing,  $s^*(c)$  is non-increasing, and  $p^*(c)$  is non-decreasing.*

An important consequence of above is that the maximum revenue expected from submitting an impression to AdX is always greater than the opportunity cost. This should not be surprising, since the publisher can pick a price high enough to compensate for the revenue loss of not assigning the impression. Hence, assigning an impression directly to an advertiser (rather than first testing the exchange) is never the right decision, and so in Figure 1 the upper branch is never taken.

In §F.2 we show that  $r(s)$  remains regular in the presence of multiple bidders in the AdX by considering the joint density of the highest and second-highest bids (given by [12]).

### 3 Our Bid-Price Control Policy

We can formulate an optimal control policy for yield maximization based on dynamic programming (DP), where the state of the system is represented by the number of impressions yet to arrive, and a vector of the number of impressions needed to comply with each advertiser’s contract. We present the details of this DP in Section A. Unfortunately, the state space of the DP has size  $O(N^{A+1})$ , and in most real-world problems the number of impressions in a single horizon can be in the order of millions, and so the DP is not efficiently solvable. Thus, here we give an approximation algorithm in which stochastic quantities are replaced by their expected values, and are assumed to be continuous. Such “deterministic approximation problems (DAP)” are popular in RM (see, e.g., [22]). In our setting, the approximation we make is to enforce contracts to be satisfied only in expectation. In §3.1, we formulate and solve an optimization problem

based on this assumption, then in §3.2 wrap a full stochastic policy around it (one that always meets the contracts, not just in expectation).

### 3.1 Deterministic Approximation Problem (DAP)

When an impression arrives, the publisher controls the reserve price submitted to AdX, and the advertiser to whom the impression is assigned, if rejected by AdX. Alternatively, in this formulation we state the controls in terms of total probabilities, where each control is a function from the quality domain to  $[0, 1]$ . Let  $\vec{s} = \{s_n(\cdot)\}_{n=1, \dots, N}$  and  $\vec{i} = \{i_n(\cdot)\}_{n=1, \dots, N}$  be vectors of controls, such that when the  $n^{\text{th}}$  impression arrives with quality  $Q$  the impression is accepted by AdX with probability  $s_n(Q)$ , and with probability  $i_{n,a}(Q)$  it is assigned to advertiser  $a$ . The conditional probability of an impression being assigned to advertiser  $a$  given that it has been rejected by AdX is given by  $I_{n,a}(Q) = i_{n,a}(Q)/(1 - s_n(Q))$ .

We denote by  $\mathcal{P}$  the set of feasible controls. A control is feasible if (i) the individual controls are non-negative, (ii) it satisfies the contractual constraint in expectation, and (iii) for every realization of the qualities the probabilities sum up to at most one. The objective of the DAP is to find a sequence of real-valued functions that maximize the expected yield:

$$\begin{aligned} J_N^D &= \max_{\vec{s}, \vec{i} \in \mathcal{P}} \sum_{n=1}^N \mathbb{E} \left[ r(s_n(Q_n)) + \sum_{a \in \mathcal{A}} i_{n,a}(Q_n) Q_{n,a} \right] \\ \text{s.t. } &\sum_{n=1}^N \mathbb{E}[i_{n,a}(Q_n)] = N\rho_a \quad \forall a \in \mathcal{A}, \\ &\sum_{a \in \mathcal{A}} i_{n,a}(Q) + s_n(Q) \leq 1, \text{ (a.s.) } \forall n, \\ &s_n(Q) \geq 0, i_{n,a}(Q) \geq 0, \text{ (a.s.) } \forall n, a \in \mathcal{A} \end{aligned}$$

The first term of the objective accounts for the revenue from AdX, while the second accounts for the quality perceived by the advertisers. Notice that in the DAP we wrote the total capacity as  $N\rho_a$  instead of  $C_a$  to allow the problem to be scaled.

Alas, the problem is still hard to solve since the number of functions is linear in  $N$ . However, exploiting the regularity of the revenue function, we can show that in the optimal solution to DAP, we can drop the dependence on  $n$  in the controls. This follows from the linearity of the constraints together with the concavity of the objective. We formalize this discussion in the following proposition.

**Proposition 2** *Suppose that the revenue function is regular. Then, there exists a time-homogenous optimal solution to the DAP, i.e. where  $s_n(Q_n) = s(Q_n)$  for all  $n = 1, \dots, N$  and  $i_n(Q_n) = i(Q_n)$  for all  $n = 1, \dots, N$ .*

The previous proposition allows us scale the problem so that  $N = 1$ , and consider the maximum expected revenue of one impression, denoted by  $J_1^D$ . In order to compute the DAP's optimal solution, we consider its dual problem, which is given by (see §B for derivation):

$$\psi(v) = \min_v \left\{ \mathbb{E}R\left(\max_{a \in \mathcal{A}_0} \{Q_a - v_a\}\right) + \sum_{a \in \mathcal{A}} v_a \rho_a \right\} \quad (3)$$

Note that  $R$  is convex and non-decreasing and the maximum is convex w.r.t  $v$ , hence the composite function within the expectation is convex. Using the fact that expectation preserves convexity, we obtain that the objective  $\psi(v)$  is convex in  $v$ .

When the distribution  $Q$  is known, the dual problem in (3) can be solved using a Subgradient Descent Method. It is worth noting that in many applications the distribution of  $Q$  is unknown and should be learned as impressions arrive. We postpone the analysis of that problem until §E.2. Once the optimal dual variables  $v$  are known, the primal solution can be constructed as given in the following theorem.

**Theorem 1** *Suppose that the revenue function is regular, and there is zero probability of a tie occurring, i.e.  $\mathbb{P}\{Q_a - v_a = Q_{a'} - v_{a'}\} = 0$  for all distinct  $a, a' \in \mathcal{A}_0$ . Then, the optimal controls for the DAP are  $s(Q) = s^*(\max_{a \in \mathcal{A}_0} \{Q_a - v_a\})$ , and  $I_a(Q) = \mathbf{1}\{Q_a - v_a > Q_{a'} - v_{a'} \forall a' \in \mathcal{A}_0\}$ , that is, the impression is assigned to the unique advertiser maximizing discounted quality. Furthermore, the optimal dual variables solve the equations*

$$\mathbb{E}[(1 - s^*(Q_a - v_a))I_a(Q)] = \rho_a, \quad \forall a \in \mathcal{A}.$$

**3.2 Our Stochastic Policy** The solution of the DAP suggests a policy for the stochastic control problem, but we must deal with two technical issues: (i) when more than one advertiser maximizes  $Q_a - v_a$  we need to decide how to break the tie, and (ii) we are only guaranteed to meet the contracts in expectation, whereas we must meet them exactly. We defer the first issue until §B.1, where we give an algorithm for generalizing the controls  $I_a(Q)$  to the case where ties are possible.

In the following let  $x_{n,a}$  be the total number of impressions left to assign to advertiser  $a$  to comply with the contract, and  $m = N - n$  the total number of impressions remaining to arrive.

### Stochastic Policy $\mu$

1. Observe state  $(m, X)$  and the realization  $Q_n$ .
2. Let  $\mathcal{A}_n = \{a \in \mathcal{A} : x_{n,a} > 0\}$ .
3. Let  $a_n^* = \arg \max_{a \in \mathcal{A}_n \cup \{0\}} \{Q_{n,a} - v_a\}$ .
4. If  $\sum_{a \in \mathcal{A}} x_{n,a} < m$ , set  $p = p^*(Q_{n,a_n^*} - v_{a_n^*})$ ; else,  $p = p_\infty$ .
5. Submit to AdX with price  $p$ . If rejected, and  $a_n^* \neq 0$ , assign to advertiser  $a_n^*$ .

Notice that impressions are only assigned to advertisers with contracts that have yet to be fulfilled. When all contracts are fulfilled, impressions are sent to AdX with the revenue maximizing price  $p^*(0)$ . Moreover, when the total number of impressions left is equal to the number of impressions necessary to fulfill the contracts, the price is set to  $p_\infty$ , and thus all incoming impressions are directly assigned to advertisers. Hence, the stochastic policy  $\mu$  satisfies the contracts with probability 1.

## 4 Asymptotic Analysis

In this section we give our main result:

**Theorem 2** Let  $J_N^\mu$  be the expected yield under the stochastic policy  $\mu$ , and  $J_N^*$  the maximum expected yield of any feasible stochastic policy. Then,

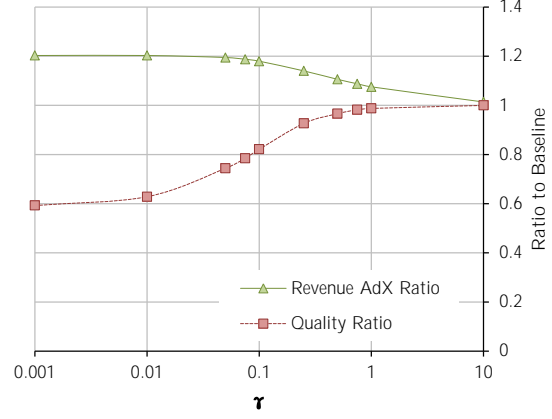
$$\frac{J_N^\mu}{J_N^*} \geq \frac{J_N^\mu}{J_N^D} \geq 1 - \frac{1}{\sqrt{N}} K(\rho),$$

where  $K(\rho) = \sqrt{\frac{A}{A+1} \sum_{a \in \mathcal{A}_0} \frac{1-\rho_a}{\rho_a}}$ .

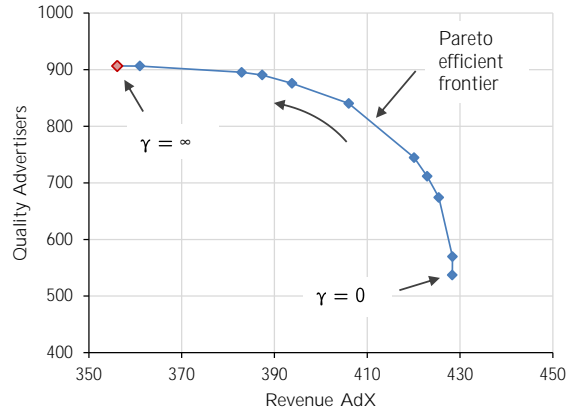
In terms of regret, our previous bound can be written as  $J_N^* - J_N^\mu \leq \sqrt{N} K(\rho) J_1^D$ , achieving an  $O(\sqrt{N})$  regret w.r.t the optimal online policy. In particular, we may fix the capacity to impression ratio of each advertiser, and consider a sequence of problems in which capacity and impressions are scaled up proportionally according to  $\rho$ . Then, the yield under policy  $\mu$  converges to the yield of the optimal online policy as  $N$  goes to infinity.

Theorem 2 is proven in two steps. First, we show that the optimal objective value of the DAP provides an upper bound on the objective value of the optimal policy, i.e.  $J_N^* \leq J_N^D$ . For this first result, we proceed by taking the optimal stochastic control policy, and construct a feasible solution for the DAP by taking expectations over the history. Later, we exploit the concavity of the objective and apply Jensen's inequality to show that this new solution attains a greater revenue in the DAP.

Second, we lower bound the yield of the stochastic policy in terms of the DAP objective. In proving that bound, we look at  $N^*$ , the first time that any advertisers contract



(a) Ratios to baseline



(b) Pareto Frontier

**Figure 2:** Figure 2a plots the quality and revenue relative to the case  $\gamma = \infty$ , as a function of  $\gamma$ . Figure 2b plots, in a quality vs. revenue graph, the objective values of the optimal solutions for the different choices of  $\gamma$ , together with the Pareto frontier. Both plots are for Instance 2, given in §E.

is fulfilled or the point is reached where all arriving impressions need to be assigned to the advertisers. We refer to the time after  $N^*$  as the *left-over regime*. The first key observation in the proof was that before time  $N^*$ , the controls of the stochastic policy behave exactly as the optimal deterministic controls. The second key observation was that the expected number of impressions in the left-over regime is  $O(\sqrt{N})$ , and the left-over regime has a small impact on the objective. See §C for details.

## 5 Experiments

Experimental results on real publisher data confirm that, as we increase the trade-off parameter  $\gamma$ , the quality of the impressions assigned to the advertisers increases, while the revenue from AdX subsides (Figure 2a). Starting from

the baseline case that disregards the AdX ( $\gamma = \infty$ ), we observe that the revenue from AdX can be substantially increased by sacrificing a small fraction of the overall quality of the impressions assigned. More generally, the Pareto frontier of efficiency vs. revenue (Figure 2b) is highly concave, relatively horizontally flat around  $\gamma = \infty$ , and vertically flat around  $\gamma = 0$ ; thus there are huge marginal improvements at the extremes, and it is highly beneficial to choose intermediate values of  $\gamma$ . We detail the data model in §D, and give the full experimental results in §E.

## 6 Comparison to the Primal-Dual Method

Consider the allocation problem faced by a publisher in display advertising in which arriving impressions need to be assigned to advertisers, and there is no option of sending to an exchange. This problem is a particular case of our model where the winning bid random variable is identically zero, i.e.  $B = 0$ . The following proposition shows that the optimal controls admit simple analytical expressions.

**Proposition 3** *Suppose that ties have zero probability. Then, in the case without AdX the optimal controls are  $I_a(Q) = \mathbf{1}\{Q_a - v_a \geq Q_{a'} - v_{a'} \forall a' \in \mathcal{A}_0\}$  where  $v = \{v_a\}_{a \in \mathcal{A}_0}$  satisfies  $v_0 = 0$  and*

$$\mathbb{P}\{Q_a - v_a \geq Q_{a'} - v_{a'} \forall a' \in \mathcal{A}_0\} = \rho_a \quad \forall a \in \mathcal{A}.$$

The resulting decision rule  $\arg \max_a \{Q_a - v_a\}$  is identical to the rule studied in previous work (e.g., [6]), where  $v_a$  is an optimal dual variable resulting from solving an assignment problem on a sample of the data, where the distribution is unknown. Roughly speaking, in [6] (and similarly in other work [7, 25, 1]), it is shown that as long as the sample is of size  $\approx \epsilon n$ , the overall assignment will be  $\approx \epsilon$  close to the optimal offline solution.

In our model, the parameters of the quality distribution are known, so we do not need to use a sample. Of course in practice, the parameters need to be learned, and so we would need to use a sample of the data in order to learn them; but in many settings (including online advertising) it is reasonable to assume that we at least know the *form* of the distribution (e.g., normal, exponential, Zipf), albeit not the specific parameters (mean, variance, covariance, etc.). The techniques in [6] are powerful *because* they don't need to assume anything about the distribution, but it is important to ask what can be gained from knowing the *form* of the distribution, which is what we do in the remainder of this section, both analytically and experimentally.

More formally, suppose the distribution of quality is not known with certainty, but we have at our disposal a sample of  $M$  quality vectors  $\{q_m\}_{m=1}^M$  that may be used to pin-down the distribution. Additionally, it is known that the qualities are drawn independently from a population with continuous density  $g(x|\theta)$ , where  $\theta$  is an unknown parameter to be estimated. Let  $G(x|\theta)$  be the c.d.f. which we assume to be strictly monotonic. For simplicity, we deal with the case of one advertiser with capacity to impression ratio of  $\rho$ . From Proposition 3 the optimal DAP control is the  $(1 - \rho)$ -quantile of  $Q$ , that is,  $v = \bar{G}^{-1}(\rho|\theta)$ . We compare the asymptotic efficiency of a parametric and a non-parametric estimation of the model.

**Parametric estimation method.** Let  $\hat{\theta}_M^{\text{mle}}$  be the maximum likelihood estimator (MLE) of the unknown parameter  $\theta$ . That is,  $\hat{\theta}_M^{\text{mle}}$  is an optimal solution of the program  $\max_{\theta} \sum_{m=1}^M \log g(q_m|\theta)$ . Once we have our estimator, we plug-in the estimated distribution in the dual problem, and solve for the optimal dual variable  $\hat{v}_M^{\text{mle}}$ . Again, from Proposition 3 we have that the optimal dual variable, given our maximum likelihood estimation, is given by  $\hat{v}_M^{\text{mle}} = \bar{G}^{-1}(\rho|\hat{\theta}_M^{\text{mle}})$ .

In turn, by the invariance property of the MLE, it is the case that  $\hat{v}_M^{\text{mle}}$  is the maximum likelihood estimator of the true optimally dual variable  $v$  (see, e.g., [4]). As a consequence, under some regularity conditions, we have that our new estimator is consistent, asymptotically efficient, and asymptotically normal

$$\sqrt{M}(\hat{v}_M^{\text{mle}} - v) \Rightarrow \mathcal{N}(0, u(\theta)),$$

where the  $u(\theta)$  is the Cramér-Rao lower bound on the variance of any unbiased estimator. The Cramér-Rao lower bound is  $u(\theta) = \left(\frac{\partial \bar{G}^{-1}(\rho|\theta)}{\partial \theta}\right)^2 I(\theta)^{-1}$ , where  $I(\theta) = \mathbb{E} \left[ \left(\frac{\partial}{\partial \theta} \ln g(Q|\theta)\right)^2 \right]$  is the Fisher information of parameter  $\theta$ .

**Non-parametric estimation method.** Considering the Sample Average Approximation (SAA) of the dual problem (3) we can obtain a non-parametric estimator of the truly optimal dual variable (see Chapter 5 from [20] for a review of the topic). In a SAA the expected value of the stochastic program is approximated by the sample average function over the observations  $\{q_m\}_{m=1}^M$ . In our case, we have that

$$\hat{v}_M = \arg \min_v \frac{1}{M} \sum_{m=1}^M \max\{q_m - v, 0\} + \rho v.$$

Equivalently, the previous problem can be stated as a linear program, and in this case one obtains the training-based Primal-Dual method as described in [6]. It can be shown that, under some conditions, the non-parametric estimator  $\hat{v}_M$  is consistent and asymptotically normal [20]. As we shall see, this estimator is not necessarily efficient. It is not hard to prove that the sample  $(1 - \rho)$ -quantile is an optimal solution to the SAA problem. Hence, from the asymptotic distribution of the  $(1 - \rho)$ -quantile we have that

$$\sqrt{m}(\hat{v}_M - v) \Rightarrow \mathcal{N}(0, u'(\theta)),$$

where the variance is  $u'(\theta) = \frac{\rho(1-\rho)}{g(v|\theta)^2}$ .

**Analysis.** Both the parametric and the non-parametric estimators converge, as the number of samples increases, to the true optimal solution. However, the non-parametric estimator is not as efficient as the parametric counterpart. Indeed, this is expected since the maximum likelihood estimator is known to be asymptotically efficient. We measure the *relative efficiency* as the ratio of asymptotic variance of the non-parametric estimator to parametric one, i.e.  $\varepsilon(\theta) = \frac{u'(\theta)}{u(\theta)}$ .

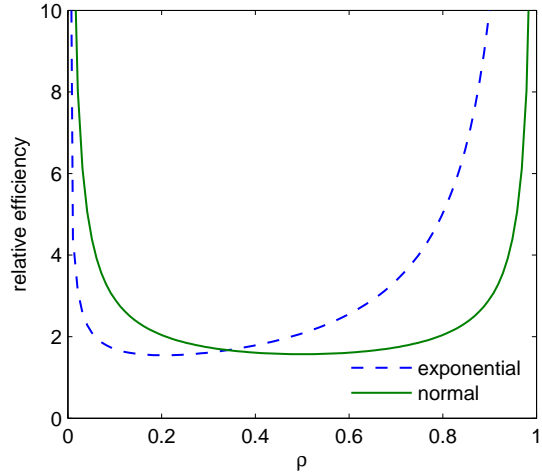
Until now, our analysis has been in terms of the optimal dual solution. The rationale is that the closer the dual variable is to the true value  $v$ , the better the performance of the policy should be. Next, we quantify analytically how does a deviation from the optimal solution impacts the performance of the policy. To assess the performance of the policy we look at the fluid limit as described in Section I. The next proposition shows that the relative efficiency in terms of the performance is exactly equal to  $\varepsilon(\theta)$ . Hence, there is no loss in looking at the relative efficiency of the estimators instead.

**Proposition 4** *The relative efficiency of the non-parametric estimator is*

$$\varepsilon(\theta) = \rho(1 - \rho)I(\theta) \left( \frac{\partial \bar{G}}{\partial \theta}(v|\theta) \right)^{-2} \geq 1, \quad (4)$$

*which is exactly equal to the relative efficiency in terms of the policies' performance.*

**Examples.** To fix ideas we consider two simple examples. First, suppose that  $Q \sim \exp(\theta)$ . The maximum likelihood estimator is given by  $\hat{\theta}_M^{\text{mle}} = \left( \frac{1}{M} \sum_{m=1}^M q_m \right)^{-1}$ , and the Fisher information is  $I(\theta) = \theta^{-2}$ . The optimal dual variable is  $v = -\theta^{-1} \ln \rho$ . Hence, the relative efficiency is  $\varepsilon(\theta) = (1 - \rho)/(\rho \ln^2 \rho)$ . In this case, the relative efficiency is lower bounded by  $\varepsilon(\theta) \geq 1.544$ . The



**Figure 3:** Relative efficiency as a function of the capacity to impression ratio  $\rho$  for the exponential distribution, and the normal distribution with known variance.

lower bound is tight, and attained at  $\rho \approx 0.2032$ . The relative efficiency as a function of the capacity to impression ratio is plotted in Figure 3. As shown in the figure, the relative efficiency may be arbitrarily bad as the capacity to impression ratio gets close to zero or one.

For the next example we assume that qualities are normal with known variance  $\sigma^2$  and unknown mean, that is,  $Q \sim \mathcal{N}(\theta, \sigma^2)$ . The maximum likelihood estimator is the sample mean,  $\hat{\theta}_M^{\text{mle}} = \frac{1}{M} \sum_{m=1}^M q_m$ , and the Fisher Information is  $I(\theta) = \sigma^{-2}$ . In this case the relative efficiency is given by  $\varepsilon(\theta) = 2\pi\rho(1 - \rho) \exp(\Phi^{-1}(1 - \rho)^2)$ , with  $\Phi^{-1}$  being the inverse of the standard normal c.d.f. Here the relative efficiency is lower bounded by  $\varepsilon(\theta) \geq \pi/2$ , with the minimum attained at  $\rho = 1/2$ . Interestingly, the relative efficiency is invariant under monotonic transformations of any random variable. Hence, the previous result holds too for the log-normal distribution.

**Experiments.** The previous analysis was in terms of a single advertiser; we show experimentally that the advantage of parametric estimation extends to multiple advertisers as well. The computational experiment is conducted as follows. First, a *training* data set of  $M$  impressions is generated. We denote the sampled quality vectors by  $\{q_m\}_{m=1}^M$ . Then, we estimate the parameters of the model on the training set as follows. For each type we estimate the type probabilities  $\hat{\pi}_T$ ; and mean  $\hat{\mu}_T$ , and covariance matrix  $\hat{\Sigma}_T$  of the logarithm of the qualities. Next, the dual problem (3) is solved on the estimated parametric model using a Gradient Descent Method, which we

describe in §H. Note that, since no AdX is considered, the maximum expected revenue function  $R(\cdot)$  is the identity. Using the optimal solution  $v^{\text{EST}}$  we construct a policy  $\mu^{\text{EST}}$  as in §3.2.

Simultaneously, we employ the primal-dual (PD) method on the training data, as in [6]. The PD method amounts to solving a sample average approximation of problem (3), which results in a linear program (see (15)). The linear program is solved using CPLEX 12. Again, using the dual optimal solution  $v^{\text{PD}}$  we construct a policy  $\mu^{\text{PD}}$ .

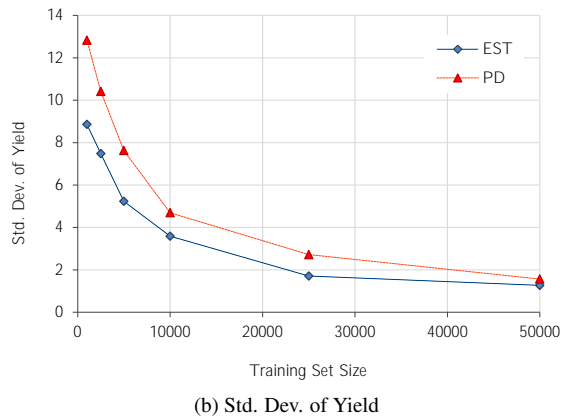
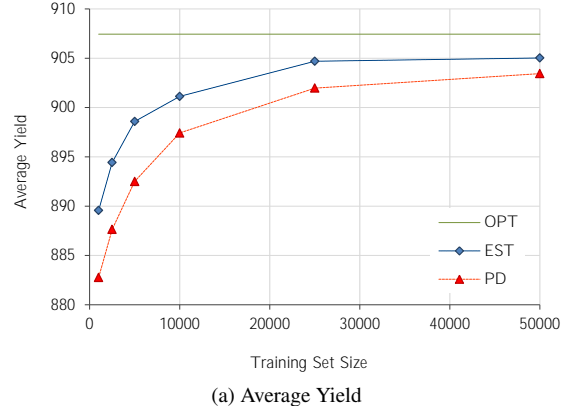
Afterwards, we assess the performance of both policies using a fluid limit (i.e., we compute the performance numerically, not experimentally; see §I). These steps are replicated on 50 different training sample sets. Plots of the results for one instance are shown in Figure 4. Details of the data model are given in §D, and Table 2 in §E reports the average results over the training sets for different sizes of training sets, and instances.

Results show that for both algorithms, as the size of the training set increases, the optimality gap decreases at a rate of  $O(M^{\frac{1}{2}})$ . However, the parametric method performs uniformly better than the non-parametric PD method. Also, the variability across different training sets diminishes as the size of the training set increases. Indeed, we observe that the standard deviation over training sets converges to zero for both methods, but the convergence is faster for the parametric one. In some sense this is expected, since the true data model follows exactly the distributional assumptions. However, the PD method is expected to be more robust to model misspecification.

Another experiment, though results are not reported, was conducted to test the strength of the parametric method on real data. We observed that, when the training set is small (around thousands), the parametric method performs better than the non-parametric one. However, as the sample size increases the non-parametric method outperforms the other. The rationale for this behavior is that, when data is scarce, the parametric method can exploit the distributional assumptions to reconstruct a fair representation of the data. However when data is larger, the fit of our model to real data is not perfect, and the non-parametric method can withstand deviations more robustly.

## 7 Related Work

**Online Allocation.** Our work is closely related to online ad allocation problems, including the *Display Ads Allocation (DA)* problem [8, 7, 1, 25], and the *AdWords (AW)* problem [16, 6]. In both of these problems, the publisher must assign online impressions to an inventory of ads, optimizing efficiency or revenue of the allocation while re-



**Figure 4:** Average (a) and standard deviation (b) of yield as a function of training set sample size; results are shown for the parametric method (EST) based on our policy  $\mu$  and the primal-dual method (PD) as in [6]. Both policies converge to the optimal yield, but EST converges faster, and with less variance.

specting pre-specified contracts.

In the DA problem, given a set of  $m$  advertisers with a set  $S_j$  of eligible impressions and demand of at most  $n(j)$  impressions, the publisher must allocate a set of  $n$  impressions that arrive online. Each impression  $i$  has value  $w_{ij} \geq 0$  for advertiser  $j$ . The goal of the publisher is to assign each impression to one advertiser maximizing the value of all the assigned impressions. The adversarial online DA problem was considered in [8], which showed that the problem is inapproximable without exploiting *free disposal*; using this property (that advertisers are at worst indifferent to receiving more impressions than required by their contract), a simple greedy algorithm is  $\frac{1}{2}$ -competitive, which is optimal. When the demand of each advertiser is large, a  $(1 - \frac{1}{e})$ -competitive algorithm exists [8], and it is tight. The stochastic model of the DA problem is more related to our problem. Following a



training-based dual algorithm by Devanur and Hayes [6], training-based  $(1 - \epsilon)$ -competitive algorithms have been developed for the DA problem and its generalization to various packing linear programs [7, 25, 1]. These papers develop a  $(1 - \epsilon)$ -competitive algorithm for online stochastic packing problems in which  $\frac{\text{OPT}}{w_{ij}} \geq O(\frac{m \log n}{\epsilon^3})$  (or  $\frac{\text{OPT}}{w_{ij}} \geq O(\frac{m \log n}{\epsilon^2})$ ) applying the technique of [1]) and the demand of each advertiser is large, in the random-order and the i.i.d model.

In the AdWords (AW) problem, the publisher allocates impressions resulting from search queries. Advertiser  $j$  has a budget  $b(j)$  on the total spend instead of a bound  $n(j)$  on the number of impressions. Assigning impression  $i$  to advertiser  $j$  consumes  $w_{ij}$  units of  $j$ 's budget instead of 1 of the  $n(j)$  slots, as in the DA problem. Other than training-based dual algorithms and primal-dual algorithms that get similar bounds as in the DA problem [6], online adaptive optimization techniques have been applied to online stochastic ad allocation [24]. Such control-based adaptive algorithms achieve asymptotic optimality following an updating rule inspired by the primal-dual algorithms. Our work differ from all the above in three main aspects: (i) We study both the parametric and non-parametric models, and compare their effectiveness in terms of the size of the sample sizes—both analytically for various distributions and experimentally on real data sets. (ii) Instead of using the framework of competitive analysis and comparing the solution with the optimum solution in hindsight, we compare the performance of our algorithm with the optimal online policy, and present a rate of convergence bound under this model. This is akin to regret bounds found in online Machine Learning; and (iii) None of the above work considers the simultaneous allocation of reservation ads and ads from AdX. In particular, these previous works do not consider the trade-off between the revenue from a spot market based on real-time bidding and the efficiency of reservation-based allocation. It is tempting to simply reduce to online stochastic packing by considering the AdX as just another “advertiser.” The problem with this is that it does not allow adjusting the reserve price, or allocating a reservation advertiser if the AdX rejects. In fact one can make such a reduction go through by considering online decisions on pairs of (reserve price, advertiser)s, and formalize the problem as an online allocation problem with general packing constraints. After a couple more steps, one can apply the techniques of [6, 7, 25, 1] to derive an online algorithm for the combined problem. However this approach discretizes the price space into multiples of  $\delta$ ; thus (a) we lose  $1 + \delta$  in the yield, (b) we increase the running time by a factor of  $\frac{1}{\delta}$ , and (c) for the the competitiveness proof to hold, the OPT

be larger than  $O(\frac{1}{\delta} \frac{m \log n}{\epsilon^3})$ . In addition, there is no clear way to apply the parametric technique. The method presented in this work not only avoids this dependence, it is a much more natural, extensible solution to the problem.

There has been recent work on display ad allocation with both contract-based advertisers and spot market advertisers [11, 10]. These focus on “fair” representative bidding strategies on behalf of the contract-based advertisers competing with the spot market bidders [11]. This line of work is mainly concerned with computing such fair representative bidding strategies for contract-based advertisers. Instead we focus on combined yield optimization and present a model and an algorithm taking into account any trade-off between quality delivered to reservation ads and revenue from the spot market.

**Revenue Management.** Another stream of relevant work is that of Revenue Management (RM). Even though RM is typically applied to airlines, car rentals, hotels and retailing [23], our problem formulation and analysis is inspired by RM techniques. As in the prototypical RM problem, we look for a policy maximizing the ex-ante expected revenue, which can be obtained using dynamic programming (DP). Since the resulting DP is intractable, we aim for a deterministic version in which stochastic quantities are replaced by their expect values and quantities assumed to be continuous. These are common in the literature [9, 14], and provide policies with provably good performance.

The case without AdX can be thought of as a Network RM problem (see, e.g., [22]) in which users’ click probabilities are requests for itineraries, and advertisers are edges in the network. The only caveat is that we aim to satisfy all contracts, or completely deplete all resources by the end of the horizon. A popular method for controlling the sale of inventory in revenue management applications is the use of bid-price controls. These were originally introduced by [21], and thoroughly analyzed by [22]. In this setting, a bid-price control sets a threshold or bid price for each advertiser, which may be interpreted as the opportunity cost of assigning one additional impression to the advertiser. When a user arrives, an impression is assigned to the advertiser whose click probability exceeds this opportunity cost in the largest amount.

From this perspective, our contribution is the inclusion of a spot market, the exchange, as an new sales channel. In this case, our policy is a suitable modification of a bid-price control that takes into account AdX by incorporating a pricing function. This pricing function quotes a reserve price to submit to the exchange that depends on the opportunity cost of assigning the impression to an advertiser.

## 8 Extensions and Conclusions

Our approach is general and applies to solve many extensions of the basic problems we have described thus far:

**Target Quality Constraint.** Previously, we considered the objective of a publisher as a suitable combination of quality and AdX revenue. An alternate formulation is one in which the publisher imposes a minimum overall quality for the impressions assigned to the advertisers. This might be more natural for some publishers; they might feel more comfortable specifying target quality constraint than picking a Lagrange multiplier to weight the impact of quality in the objective. Additionally, in some settings the advertisers themselves might demand that certain level of quality is guaranteed. In Section F.1, we extend our algorithm to this alternative objective.

**AdX with Multiple Bidders.** We generalize our results to the case where multiple buyers participate in the Ad Exchange. We model AdX as an auction with  $K$  risk neutral buyers. In Section F.2 we show that all our results hold when individual valuations are drawn independently from the same distribution.

**AdX with User Information.** In most systems, the publisher shares some user information with the exchange. In turn, the exchange may partially disclose the user information to their advertisers. The advertisers may react to this information, and bid strategically [17]. We extend our model to the case when the bids from AdX are correlated with the quality of the impression (a surrogate for user information). This is presented in Section F.3.

**User Types.** We have thus far assumed that any user could be potentially assigned to any advertiser. In practice, however, advertisers have specific targeting criteria. For instance, a guaranteed contract may demand for females with certain age range living in New York, while other contract may demand for males in California. We show in Section F.4 how to handle this by forcing the publisher to pay a good-will penalty to the advertisers each time an undesired impression is incorrectly assigned.

Our approach helps publishers determine when and how to access AdX to complement their contract sales of impressions. We also hope our insights here—in particular the comparison between parametric and non-parametric approaches—will help us understand ad allocation problems more deeply.

**Acknowledgments.** We thank Omar Besbes, Ciamac Moallemi, Daryl Pregibon, Ana Radovanovic, Nemo Semret, and Nicolás Stier for helpful discussions.

## References

[1] S. Agrawal, Z. Wang, and Y. Ye. A dynamic near-optimal algorithm for online linear programming. Working paper posted at

- <http://www.stanford.edu/yyye/>.
- [2] D. P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, 2nd edition, 2000.
- [3] P. Billingsley. *Convergence of Probability Measures*. John Wiley & Sons Inc, January 1968.
- [4] G. Casella and R. L. Berger. *Statistical Inference; 2nd ed.* Duxbury advanced series. Wadsworth, New Delhi, 2002.
- [5] J. Dai and B. Prabhakar. The throughput of data switches with and without speedup. In *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 2, pages 556–564, 2000.
- [6] N. Devanur and T. Hayes. The adwords problem: Online keyword matching with budgeted bidders under random permutations. In *ACM EC*, 2009.
- [7] J. Feldman, M. Henzinger, N. Korula, V. Mirrokni, and C. Stein. Online stochastic packing applied to display ad allocation. In *ESA*, 2010.
- [8] J. Feldman, N. Korula, V. Mirrokni, S. Muthukrishnan, and M. Pal. Online ad assignment with free disposal. In *WINE*, 2009.
- [9] G. Gallego and G. v. Ryzin. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science*, 40(8):999–1020, 1994.
- [10] A. Ghosh, P. McAfee, K. Papineni, and S. Vassilvitskii. Bidding for representative allocations for display advertising. In *WINE*, pages 208–219, 2009.
- [11] A. Ghosh, B. I. P. Rubinstein, S. Vassilvitskii, and M. Zinkevich. Adaptive bidding for display advertising. In *WWW*, pages 251–260, 2009.
- [12] J.-J. Laffont and E. Maskin. Optimal reservation price in the vickrey auction. *Economics Letters*, 6(4):309 – 313, 1980.
- [13] M. A. Lariviere. A note on probability distributions with increasing generalized failure rates. *Oper. Res.*, 54:602–604, May 2006.
- [14] Q. Liu and G. van Ryzin. On the choice-based linear programming model for network revenue management. *Manufacturing Service Operations Management*, 10(2):288–310, 2008.
- [15] D. G. Luenberger. *Optimization by Vector Space Methods*. John Wiley & Sons, Inc., New York, NY, USA, 1st edition, 1997.
- [16] A. Mehta, A. Saberi, U. Vazirani, and V. Vazirani. Adwords and generalized online matching. In *FOCS*, 2005.
- [17] S. Muthukrishnan. Ad exchanges: Research issues. In *Internet and Network Economics, Lecture Notes in Computer Science*, pages 1–12, 2009.
- [18] R. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981.
- [19] A. Shapiro. Asymptotic analysis of stochastic programs. *Annals of Operations Research*, 30:169–186, 1991.
- [20] A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on Stochastic Programming: Modeling and Theory*. MOS-SIAM Series on Optimization 9. Society for Industrial and Applied Mathematics (SIAM), 2009.
- [21] R. Simpson. Using network flow techniques to find shadow prices for market demands and seat. Flight Transportation Laboratory Memorandum M89-1, 1989.
- [22] K. Talluri and G. van Ryzin. An analysis of bid-price controls for network revenue management. *Management Science*, 44(11):1577–1593, 1998.

- [23] K. T. Talluri and G. J. van Ryzin. *The Theory and Practice of Revenue Management*. International Series in Operations Research & Management Science, Vol. 68. Springer, 2004.
- [24] B. Tan and R. Srikant. Online advertisement, optimization and stochastic networks. *CoRR*, abs/1009.0870, 2010.
- [25] E. Vee, S. Vassilvitskii, and J. Shanmugasundaram. Optimal online assignment with forecasts. In *ACM EC*, 2010.

## A Dynamic Programming Formulation

In this section we derive an optimal policy based on dynamic programming, which will not be efficient (due to the size of the state space). Interestingly, however, the resulting Bellman equation closely resembles in form the decision rule we use in our policy.

Let  $(m, X)$  be the state of the system, where we denote by  $m$  the total number of impressions remaining to arrive, and by  $X = \{x_a\}_{a \in \mathcal{A}}$  the number of impressions needed to comply with each advertiser's contract. Let the value function, denoted by  $J_m(X)$ , be defined as the optimal expected revenue obtainable under state  $(m, X)$ . Using the fact that is optimal to first test the exchange, we obtain the following Bellman equation

$$\begin{aligned} J_m(X) &= \mathbb{E}_{Q_n} \left[ \max_{p \geq 0} \left\{ \bar{F}(p)(p + J_{m-1}(X)) + (1 - \bar{F}(p)) \max_{a \in \mathcal{A}_0} \{Q_{n,a} + J_{m-1}(X - \mathbf{1}_a)\} \right\} \right] \\ &= J_{m-1}(X) + \mathbb{E}_{Q_n} \left[ R \left( \max_{a \in \mathcal{A}_0} \{Q_{n,a} - \Delta_a J_{m-1}(X)\} \right) \right], \end{aligned} \quad (5)$$

where we defined  $\mathbf{1}_a$  as a vector with a one in entry  $a$  and zero elsewhere,  $\mathbf{1}_0 = 0$ , and  $\Delta_a J_m(X) = J_m(X) - J_m(x - \mathbf{1}_a)$  as the expected marginal revenue of one extra impression for advertiser  $a$ . In (5) the objective accounts for the revenue obtained from attempting to send the impression to AdX. In the revenue has two terms that depend on whether the impression is accepted or not by AdX. In the latter case the maximum accounts for the decision of assigning the impression directly to the advertiser or discarding the impression (when  $a = 0$ ). In (5) we used the fact that assigning an impression directly to an advertiser is never the right decision (except in boundary conditions, see below). The publisher, however, may choose to discard impressions with low quality after being rejected by AdX.

Our objective is to compute  $J_N^* = J_N(C)$ . The boundary conditions are

$$\begin{aligned} J_m(x) &= -\infty, & \forall x \text{ s.t. } x_a < 0 \text{ for some } a \in \mathcal{A}, \\ J_m(x) &= -\infty, & \forall m < \sum_{a \in \mathcal{A}} x_a. \end{aligned}$$

Recall that when the contract with an advertiser is fulfilled, no extra revenue is obtained from assigning to her more impressions. This is the case of the first boundary condition, which guarantees that advertisers whose contract is fulfilled are excluded from the assignment. In particular, when  $X = 0$  all remaining impressions are sent to AdX with the revenue maximum price  $p^*(0)$  when  $x = 0$ . The second boundary condition guarantees that the contracts with the advertisers are always fulfilled. When  $\sum_{a \in \mathcal{A}} x_a = m$  AdX must be bypassed, and impressions should be assigned directly to the advertisers.

Summarizing, the optimal policy is:

---

### Algorithm 1 Optimal policy for the case with one advertiser.

---

```

Observe state  $(m, X)$  and the realization  $Q_n$ .
Let  $a^* = \arg \max_{a \in \mathcal{A}_0} \{Q_{n,a} - \Delta_a J_{n-1}(X)\}$ .
Submit to AdX with price  $p^*(Q_{i,a^*} - \Delta_{a^*} J_{n-1}(X))$ .
if impression rejected by AdX and  $a^* > 0$  then
    Assign to advertiser  $a^*$ .
end if

```

---

In the above policy, when the impression is submitted to AdX, the optimal price ponders an opportunity cost of  $Q_{i,a^*} - \Delta_{a^*} J_{n-1}(X)$ . This opportunity cost, when positive, is just the value of the impression discounted by the loss of potential revenue from assigning the impression right now. Note that the two boundary conditions are implicit in the optimal policy. By the first boundary condition, when the contract of an advertiser is fulfilled or  $x_a = 0$  for some advertiser  $a \in \mathcal{A}$ , the marginal expected revenue of the advertiser becomes infinity, i.e.  $\Delta_a J_{n-1}(X) = \infty$ , and that advertiser is excluded from future decisions (since it is dominated at least by the outside option). By the second boundary condition, when  $\sum_{a \in \mathcal{A}} x_a = m$  the expected marginal revenue of all advertisers whose contracts have already been fulfilled (including the outside option) becomes negative infinity. Hence, the set of available advertisers is restricted to those whose contract is yet to be fulfilled. This guarantees that the policy complies with the contracts.

## B Derivation of the Dual to DAP

The DAP is given by:

$$J_N^D = \max_{\bar{s}, \bar{i} \in \mathcal{P}} \sum_{n=1}^N \mathbb{E} \left[ r(s_n(Q_n)) + \sum_{a \in \mathcal{A}} i_{n,a}(Q_n) Q_{n,a} \right]$$

$$\text{s.t. } \sum_{n=1}^N \mathbb{E} [i_{n,a}(Q_n)] = N \rho_a \quad \forall a \in \mathcal{A}, \quad (6a)$$

$$\sum_{a \in \mathcal{A}} i_{n,a}(Q) + s_n(Q) \leq 1, \text{ (a.s.) } \forall n, \quad (6b)$$

$$s_n(Q) \geq 0, i_{n,a}(Q) \geq 0, \text{ (a.s.) } \forall n, a \in \mathcal{A}$$

To find the dual, we introduce Lagrange multipliers  $v = \{v_a\}_{a \in \mathcal{A}}$  for the capacity constraints (6a), and multiplier  $\lambda(Q) \geq 0$  for constraint (6b). Additionally, non-negativity constraints are kept in the feasible set. The Lagrangian, denoted by  $\mathcal{L}(s, i; v, \lambda)$  is

$$\mathcal{L}(s, i; v, \lambda) = \mathbb{E} \left[ r(s(Q)) + \sum_{a \in \mathcal{A}} i_a(Q) Q_a \right]$$

$$- \mathbb{E} \left[ \sum_{a \in \mathcal{A}} v_a (i_a(Q) - \rho_a) + \lambda(Q) \left( \sum_{a \in \mathcal{A}} i_a(Q) + s(Q) - 1 \right) \right].$$

The dual function, denoted by  $\phi(v, \lambda)$ , is the supremum of the Lagrangian over the feasible set. Thus, we have that

$$\phi(v, \lambda) = \sup_{s(\cdot) \geq 0, i_a(\cdot) \geq 0} \mathcal{L}(s, i; v, \lambda)$$

$$= \sup_{s(\cdot) \geq 0} \mathbb{E} [r(s(Q)) + (1 - s(Q))\lambda(Q)] + \sum_{a \in \mathcal{A}} v_a \rho_a$$

$$+ \sum_{a \in \mathcal{A}} \sup_{i_a(\cdot) \geq 0} \mathbb{E} [i_a(Q) (Q_a - v_a - \lambda(Q))]$$

$$= \mathbb{E} R(\lambda(Q)) + \sum_{a \in \mathcal{A}} v_a \rho_a + \sum_{a \in \mathcal{A}} \sup_{i_a(\cdot) \geq 0} \mathbb{E} [i_a(Q) (Q_a - v_a - \lambda(Q))],$$

where the first equation follows because the objective is separable, and the last equation from solving the AdX variational problem. In order for the dual function to be finite we require that  $Q_a - v_a \leq \lambda(Q)$  almost surely in  $Q$ . Otherwise, the dual function is unbounded.

Next, the dual problem is  $\min_{v, \lambda(\cdot) \geq 0} \phi(v, \lambda)$ . When the revenue function is regular the DAP's objective is concave and bounded from above. Moreover, the constraints of the primal problem are linear, the feasible set convex, and 0 is an interior solution w.r.t constraints (6a), and (6b). Hence, by the Strong Duality Theorem (p.224 in [15]) the dual problem attains the primal objective value. So, we have that

$$J_1^D = \min_{v, \lambda(\cdot)} \mathbb{E} R(\lambda(Q)) + \sum_{a \in \mathcal{A}} v_a \rho_a$$

$$\text{s.t. } \lambda(Q) \geq Q_a - v_a, \quad \forall Q, a \quad (7)$$

$$\lambda(Q) \geq 0 \quad \forall Q.$$

In the dual problem,  $\lambda(Q)$  is bounded from below by  $Q_a - v_a$  for all  $a \in \mathcal{A}_0$ , with  $v_0 = 0$ . Because the optimal AdX revenue function  $R$  is non-decreasing and the problem is to minimize the objective, it should be the case that the bound is tight in the optimal solution. Thus, we let  $\lambda(Q) = \max_{a \in \mathcal{A}_0} \{Q_a - v_a\}$ , and obtain the equivalent minimization problem

$$\min_v \left\{ \mathbb{E}R \left( \max_{a \in \mathcal{A}_0} \{Q_a - v_a\} \right) + \sum_{a \in \mathcal{A}} v_a \rho_a \right\}. \quad (8)$$

**B.1 Handling ties** Theorem 1 had an assumption that there would be no ties between advertisers verifying the maximum  $Q_a - v_a$ . In this section we address this situation. Devanur and Hayes [6] proposed introducing small random and independent perturbations to the qualities, or smoothing the dual problem to break ties. We provide an alternate method that directly attacks ties, and provides a randomized tie-breaking rule. Computing the parameters of the tie-breaking rule requires solving a flow problem on a graph of size  $2^{|\mathcal{A}|}$ ; thus in some settings it may not be possible. In section H, we show that in practice ties do not occur frequently. However, for completeness we provide a full characterization of the problem.

For any non-empty subset  $S \subseteq \mathcal{A}_0$ , we define a  $S$ -tie as the event when the maximum is verified exactly by all the advertisers  $a \in S$ , and the impression is rejected by AdX. Note that the tie may be a singleton, in the case that exactly one advertiser verifies the maximum. Since the dual variables  $v$  are known, the probability of such event can be written as

$$\mathbb{P}(S\text{-tie}) = \mathbb{E} \left[ (1 - s^*(\lambda(Q))) \mathbf{1} \{ Q_a - v_a = \lambda(Q) \ \forall a \in S, \ Q_a - v_a < \lambda(Q) \ \forall a \notin S \} \right],$$

where  $\lambda(Q) = \max_{a \in \mathcal{A}_0} \{Q_a - v_a\}$ . With some abuse of notation we define the  $\emptyset$ -tie as the event when the impression is accepted by AdX, that is,  $\mathbb{P}(\emptyset\text{-tie}) = \mathbb{E}[s^*(\lambda(Q))]$ . Note that the tie events induce a partition of the quality space, and we have that  $\sum_{S \subseteq \mathcal{A}_0} \mathbb{P}(S\text{-tie}) = 1$ .

We look for a random tie-breaking rule that assigns an arriving impression to advertiser  $a \in S$  with conditional probability  $I_a(S)$  given that a  $S$ -tie occurs. Hence, the routing probabilities depend on which advertisers tie, and not on the particular realization of the qualities (they are independent of  $\lambda(Q)$ ). Therefore, under such policy the total probability, originating from  $S$ -ties, of an impression being assigned to advertiser  $a$  is  $y_a(S) = \mathbb{P}(S\text{-tie})I_a(S)$ . We can interpret  $y_a(S)$  as the normalized flow of impression assigned to the advertiser originating from  $S$ -ties. We will show that, in terms of  $y_a(S)$  as decision variables, finding the tie-breaking rule amounts to solving a transportation problem.

First, in order for the publisher to fulfill the contract with an advertiser  $a \in \mathcal{A}$  the incoming flow of impressions over all possible ties sums up to  $\rho_a$ . The previous constraint can be written as

$$\sum_{S \subseteq \mathcal{A}_0: a \in S} y_a(S) = \rho_a, \quad \forall a \in \mathcal{A}. \quad (9)$$

Notice that we impose no constraints for  $a = 0$  since any number of impressions can be discarded. Alternatively, we could set  $\rho_0^{\text{eff}} = 1 - \mathbb{P}(\emptyset\text{-tie}) - \sum_{a \in \mathcal{A}} \rho_a$  because the impressions effectively discarded are those that are rejected by AdX and not assigned to an advertiser. Second, the outgoing flow of impressions originating from a particular tie should sum up to the actual probability of that tie occurring. Then, we have that

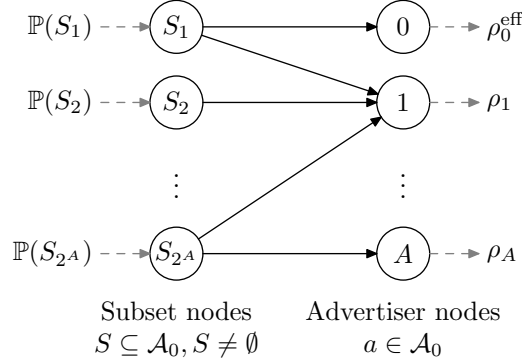
$$\sum_{a \in S} y_a(S) = \mathbb{P}(S\text{-tie}), \quad \forall S \subseteq \mathcal{A}_0. \quad (10)$$

Third, we require that  $y_a(S) \geq 0$  for all  $S \subseteq \mathcal{A}_0$  and  $a \in S$ . Finally, in order to obtain the tie-breaking rule we need a non-negative flow satisfying constraints (9) and (10). Once a such solution is found, the optimal controls can be computed as

$$I_a(Q) = \begin{cases} y_a(S)/\mathbb{P}(S\text{-tie}) & \text{if } a \in S, \text{ and } Q \text{ is an } S\text{-tie,} \\ 0 & \text{otherwise,} \end{cases}$$

$$s(Q) = s^* \left( \max_{a \in \mathcal{A}_0} \{Q_{n,a} - v_a\} \right).$$

It is not hard to see that the previous problem can be stated as a feasible flow problem in a bipartite graph. We briefly describe how to construct such graph next. On the left-hand side of the graph we include one node for each non-empty subset  $S \subseteq \mathcal{A}_0$ , and in the right-hand side we add one node for each advertiser  $a \in \mathcal{A}_0$ . In the following we refer to nodes in the left-hand side as *subset nodes*, and to those in the right-hand side as *advertiser nodes*. The supply for subset nodes is  $\mathbb{P}(S\text{-tie})$ , while the demand for advertiser nodes is  $\rho_a$ . Arcs in the graph represent the membership relation, i.e., the subset node  $S$  and advertiser node  $a$  are connected if and only if  $a \in S$ . Moreover, arc capacities are set to infinity. In Figure 5 the resulting bipartite graph is shown.



**Figure 5:** Bipartite flow problem solved to obtain the tie-breaking rule. On the left-hand side of the graph we include one node for each non-empty subset  $S \subseteq \mathcal{A}_0$  (subset nodes), and in the right-hand side we add one node for each advertiser  $a \in \mathcal{A}_0$  (advertiser nodes). The supply for subset nodes is  $\mathbb{P}(S\text{-tie})$ , while the demand for advertiser nodes is  $\rho_a$ . Arcs in the graph represent the membership relation, i.e., the subset node  $S$  and advertiser node  $a$  are connected if and only if  $a \in S$ . Arc capacities are set to infinity.

An important question is whether the flow problem admits a feasible solution. The next result proves that the answer is affirmative when the dual variables  $v$  are optimal for the dual problem (3). The proof proceeds by casting the feasible flow problem as a maximum flow problem, and then exploiting the optimality conditions of  $v$  to lower bound every cut in the bipartite graph.

**Proposition 5** *Suppose that  $v \in \mathbb{R}^A$  is an optimal solution for the dual problem (3). Then, there exists a non-negative flow satisfying constraints (9) and (10).*

We conclude this section by showing that the solution constructed is optimal for the primal problem. Notice that the solution is feasible because it satisfies constraints (9) and (10). In order to prove optimality it suffices to show that it attains the dual objective value, or that it satisfies the complementary slackness conditions. The latter follows trivially.

Once the optimal controls are calculated, we construct our stochastic policy as follows.

#### Stochastic Policy $\mu$

1. Observe state  $(m, X)$  and the realization  $Q_n$ .
2. Let  $\mathcal{A}_n = \{a \in \mathcal{A} : x_{n,a} > 0\}$ .
3. Let  $\mathcal{A}_n^* = \arg \max_{a \in \mathcal{A}_n \cup \{0\}} \{Q_{n,a} - v_a\}$ .
4. If  $\sum_{a \in \mathcal{A}} x_{n,a} < m$ , set  $p = p^*(\max_{a \in \mathcal{A}_n \cup \{0\}} \{Q_{n,a} - v_a\})$ ; else,  $p = p_\infty$ .
5. Submit to AdX with price  $p$ . If rejected, and  $\mathcal{A}_n^* \neq \{0\}$ , assign to advertiser  $a$  in  $\mathcal{A}_n^*$  with probability  $I_a(Q_n) / \sum_{a' \in \mathcal{A}_n^*} I_{a'}(Q_n)$ .

Notice that impressions are only assigned to advertisers with contracts that have yet to be fulfilled. Additionally, as the contracts of some advertisers are fulfilled, these are excluded of the assignment, and the routing probabilities  $I_a(\cdot)$  of the remaining advertisers are scaled-up and normalized.

## C Proof of the Main Result

In this section we prove Theorem 2. We proceed in the following way. First, we formulate the problem as a stochastic control problem (SCP). Though not practical, this abstract and equivalent formulation is useful from a theoretical point of view. Second, we show that the optimal objective value of the DAP provides an upper bound on the objective value of the SCP. Finally, we show that the upper bound is asymptotically tight.

A stochastic control policy maps states of the system to control actions (prices and target advertiser), and is adapted to the history up to the decision epoch. We restrict our attention to policies that always submit the impression to AdX, which were argued to be optimal. Recall that given the reserve price, the publisher knows the actual probability that the impression is accepted by AdX. As before, we recast the problem in terms of the survival probability control. Hence, the publisher picks the probability that the impression is accepted. Conversely, given a survival probability the reserve price can be easily computed using  $\bar{F}^{-1}(\cdot)$ . We denote by  $s_n^\mu(Q)$  the target survival probability under policy  $\mu$  at time  $n$  when an impression with quality  $Q$  arrives. Similarly, we let  $I_{n,a}^\mu(Q)$  indicate whether the  $n^{\text{th}}$  impression is assigned to advertiser  $a$  or not when policy  $\mu$  is used. In particular,  $I_{n,a}^\mu(Q) = 1$  indicates that the impression should be assigned to the advertiser if rejected by AdX.

We let the binary random variable  $X_n(s_n^\mu(Q))$  indicate whether the  $n^{\text{th}}$  impression is accepted by AdX or not when policy  $\mu$  is used. Specifically,  $X_n(s_n^\mu(Q)) = 1$  indicates that the impression is accepted by AdX, and when  $X_n(s_n^\mu(Q)) = 0$  the impression is rejected by AdX. Notice that, conditioning on the quality of the impression and the history,  $X_n(s_n^\mu(Q))$  is a Bernoulli random variable with success probability  $s_n^\mu(Q)$ .

We denote by  $\mathcal{M}$  the set of admissible policies, i.e. policies that are non-anticipating, adapting and feasible. A feasible policy should satisfy the contractual obligations with each advertiser, or equivalently the number of impressions assigned to advertiser  $a$  should be exactly equal to  $C_a$  in an almost sure sense. Additionally, the target advertiser controls should satisfy that  $\sum_{a \in \mathcal{A}} I_{n,a}^\mu(Q) \leq 1$ , since the impression should be assigned to at most one advertiser.

Finally, the equivalent stochastic optimal control problem is

$$\begin{aligned}
 J_N^* &= \max_{\mu \in \mathcal{M}} \mathbb{E} \left[ \sum_{n=1}^N r(s_n^\mu(Q_n)) + [1 - s_n^\mu(Q_n)] \sum_{a \in \mathcal{A}} I_{n,a}^\mu(Q_n) Q_{n,a} \right] \\
 \text{s.t. } &\sum_{n=1}^N [1 - X_n(s_n^\mu(Q_n))] I_{n,a}^\mu(Q_n) = \rho_a N \quad (\text{a.s.}) \quad \forall a \in \mathcal{A}, \\
 &\sum_{a \in \mathcal{A}} I_{n,a}^\mu(Q_n) \leq 1, \quad (\text{a.s.}) \quad \forall n, \\
 &0 \leq s_n^\mu(Q_n) \leq 1, I_{n,a}^\mu(Q_n) \in \{0, 1\}, \quad (\text{a.s.}) \quad \forall n, a
 \end{aligned} \tag{11}$$

where  $J_N^*$  denotes the optimal expected revenue over the set of admissible policies  $\mathcal{M}$ . The objective follows from the fact that conditioning on the quality of the impression and the history,  $X_n(s_n^\mu(Q))$  is a Bernoulli random variable. The first term of each summand in the objective accounts for the revenue obtained from AdX when the impression is accepted, while the second term accounts for the yield obtained when the impression is rejected and later assigned to an advertiser. By the Principle of Optimality it is the case that the dynamic program described in section A provides an optimal solution to the SCP [2].

We are now ready to prove the two theorems that imply Theorem 2. We first show the following:

**Proposition 6** *The optimal objective value of the DAP provides an upper bound on the objective value of the optimal policy, i.e.  $J_N^* \leq J_N^D$ .*

*Proof:* Let  $\mu^*$  be the optimal policy for the stochastic control problem. Let  $\hat{s} = \{\hat{s}_n(\cdot)\}_{n=1, \dots, N}$  and  $\hat{i} = \{\hat{i}_n(\cdot)\}_{n=1, \dots, N}$  be deterministic vectors of controls defined as

$$\begin{aligned}
 \hat{s}_n(Q) &= \mathbb{E}_{\mathcal{F}_n} \left[ s_n^{\mu^*}(Q) \mid Q \right] \quad \forall Q \text{ pointwise,} \\
 \hat{i}_{n,a}(Q) &= \mathbb{E}_{\mathcal{F}_n} \left[ (1 - s_n^{\mu^*}(Q)) I_{n,a}^{\mu^*}(Q) \mid Q \right] \quad \forall Q \text{ pointwise, } a \in \mathcal{A},
 \end{aligned}$$



where the expectation is taken over the history of the system until  $n$ , which is denoted by  $\mathcal{F}_n$ , and conditional on a particular realization of  $Q$ . The resulting controls are independent of the history, and dependent only on the realization of  $Q$  and the impression number  $n$ . Thus, they fulfill the first approximation and they are valid deterministic vectors of controls. We will show that  $(\hat{s}, \hat{i})$  is feasible for the DAP, and that its objective value (in the DAP) dominates the optimal objective value of the SCP. Then, we may conclude that  $J_N^* \leq J_N^D(\hat{s}, \hat{i}) \leq J_N^D$ , because no feasible solution is better than the optimal.

First, for the contract fulfillment constraint we have that for each advertiser  $a \in \mathcal{A}$

$$\begin{aligned} C_a &= \mathbb{E} \left[ \sum_{n=1}^N (1 - X_n(s_n^{\mu^*}(Q_n))) I_{n,a}^{\mu^*}(Q_n) \right] \\ &= \sum_{n=1}^N \mathbb{E} \left[ \mathbb{E}_{\mathcal{F}_n} \left[ (1 - s_n^{\mu^*}(Q_n)) I_{n,a}^{\mu^*}(Q_n) \mid Q_n \right] \right] = \sum_{n=1}^N \mathbb{E} [\hat{i}_{n,a}(Q)], \end{aligned}$$

where the first equality follows from taking expectations to the almost sure contract fulfillment constraint of  $\mu^*$ , the second from the tower rule, and the third from substituting  $\hat{s}$  and  $\hat{i}$  pointwise for all  $Q$  and the fact that impressions are i.i.d. Non-negativity of the controls follows trivially. Additionally, it is not hard to show that  $\sum_{a \in \mathcal{A}} \hat{i}_{n,a}(\cdot) + s_n(\cdot) \leq 1$  for all  $n$ . Thus,  $(\hat{s}, \hat{i})$  is a feasible deterministic control.

Second, the objective value of the optimal stochastic control is bounded by

$$\begin{aligned} J_N^* &= \mathbb{E} \left[ \sum_{n=1}^N \mathbb{E}_{\mathcal{F}_n} \left[ r \left( s_n^{\mu^*}(Q_n) \right) \mid Q_n \right] + \sum_{a \in \mathcal{A}} Q_{n,a} \mathbb{E}_{\mathcal{F}_n} \left[ (1 - s_n^{\mu^*}(Q_n)) I_{n,a}^{\mu^*}(Q_n) \mid Q_n \right] \right] \\ &\leq \mathbb{E} \left[ \sum_{n=1}^N r(\hat{s}_n(Q_n)) + \sum_{a \in \mathcal{A}} Q_{n,a} \hat{i}_{n,a}(Q) \right] = J_N^D(\hat{s}, \hat{i}), \end{aligned}$$

where the first equality follows from the tower rule and because  $Q_n$  is measurable w.r.t. the conditional expectation, and the inequality from applying Jensen's inequality to the concave revenue function.  $\blacksquare$

Now we complete the proof of Theorem 2 by bounding the yield of the stochastic policy in terms of the DAP objective.

*Proof of Theorem 2:* The first bound follows from Proposition 6.

Let  $S_{n,a}^{\mu} = \sum_{i=1}^n (1 - X_i(s_i^{\mu}(Q_i))) I_{i,a}^{\mu}(Q_i)$  be the total number of impressions assigned to advertiser  $a$  by time  $n$  when following the stochastic policy  $\mu$ . Additionally, we denote by  $S_n^{\mu} = \{S_{n,a}^{\mu}\}_{a \in \mathcal{A}}$  the random vector of impressions assigned to advertisers. Then,  $x_{n,a} = C_a - S_{n,a}^{\mu}$  is the total number of impressions left to assign to advertiser  $a$  to fulfill the contract, and  $m = N - n$  is the total number of impressions remaining to arrive.

To simplify the proof, we let  $C_0 = N - \sum_{a \in \mathcal{A}} C_a$  be the total number of impressions that are not assigned to any advertiser (accepted by AdX and discarded), and we refer to  $S_{n,0}^{\mu} = n - \sum_{a \in \mathcal{A}} S_{n,a}^{\mu}$  as total number of impressions not assigned to any advertiser by time  $n$  when following the stochastic policy  $\mu$ . Because  $C_0$  is the total number of impressions we can dispense of, when the point is reached that  $S_{n,0}^{\mu} = C_0$ , then all remaining impressions need to be assigned to the advertisers.

Let the random time  $N^* = \inf \{1 \leq n \leq N : x_{n,a} = 0 \text{ for some } a \in \mathcal{A} \text{ or } \sum_{a \in \mathcal{A}} x_{n,a} = m\}$  be the first time that any advertiser's contract is fulfilled or the point is reached where all arriving impressions need to be assigned to the advertisers. Clearly,  $N^*$  is a stopping time with respect to the stochastic process  $\{S_n^{\mu}\}_{n=1, \dots, N}$ .

In the following, let  $R_n^{\mu}$  be the revenue from time  $n$  under policy  $\mu$ . Similarly, we denote by  $R_n$  the revenue from time  $n$  when the deterministic control are used in an alternate system with no capacity constraints. Because the deterministic controls are time-homogeneous, and the underlying random variables are i.i.d., then the random variables  $\{R_n\}_{n=1, \dots, N}$  are i.i.d. too. Moreover, it is the case that  $\mathbb{E} R_n = J_1^D$ . Notice that when  $n < N^*$ , the controls of stochastic policy  $\mu$  behave exactly as the optimal deterministic controls. Thus,  $R_n = R_n^{\mu}$  for  $n < N^*$ . Using this fact together with the fact that  $N^*$  is a stopping time we get that

$$J_N^{\mu} = \mathbb{E} \left[ \sum_{n=1}^N R_n^{\mu} \right] = \mathbb{E} \left[ \sum_{n=1}^{N^*} R_n + \sum_{n=N^*+1}^N R_n^{\mu} \right] \geq \mathbb{E} \left[ \sum_{n=1}^{N^*} R_n \right] = \mathbb{E} N^* J_1^D, \quad (12)$$

where the inequality follows from the non-negativity of the revenues, and the last equality from Wald's equation. Then, we conclude that  $J_N^\mu/J_N^D \geq \mathbb{E}N^*/N$ .

Next, we turn to the problem of lower bounding  $\mathbb{E}N^*$ . Before proceeding we make some definitions. We define by  $S_{n,a}$  the number of impressions assigned to advertiser  $a$  by time  $n$  when following the deterministic controls in the alternate system with no capacity constraints. As for the revenues, it is the case that  $S_{n,a} = S_{n,a}^\mu$  for  $n < N^*$ . We define  $S_{n,0}$  in a similar fashion.

Let  $N_a = \inf\{n \geq 1 : S_{n,a} = C_a\}$  be the time when the contract of advertiser  $a \in \mathcal{A}$  is fulfilled, and  $N_0 = \inf\{n \geq 1 : S_{n,0} = C_0\}$  be the point in time where all arriving impressions need to be assigned to the advertisers. Even though these stopping times are defined with respect to the stochastic process that follows the deterministic controls, it is the case that  $N^* = \min_{a \in \mathcal{A}_0} \{N_a\}$ . In the remainder of the proof we study the mean and variance of each stopping time, and then conclude with a bound for  $\mathbb{E}N^*$  based on those central moments.

For the case of  $a \in \mathcal{A}$ , the summands of  $S_{n,a}$  are independent Bernoulli random variables with success probability  $\rho_a$ . The success probability follows from (6a). Hence,  $N_a$  is a negative binomial random variable with  $C_a$  successes and success probability  $\rho_a$ . The mean and variance are given by  $\mathbb{E}N_a = N$ , and  $\text{Var}[N_a] = N \frac{1-\rho_a}{\rho_a}$ , where we used that  $\rho_a = C_a/N$ . Similarly, for the case of  $a = 0$ , now the summands of  $S_{n,0}$  are Bernoulli random variables with success probability  $\rho_0$ . Hence,  $N_0$  is a negative binomial random variable with  $C_0$  successes and success probability  $\rho_0$ .

Finally, using the lower bound on the mean of the minimum of a number of random variables of Aven (1985) we get that

$$\begin{aligned} \mathbb{E}N^* &= \mathbb{E} \min_{a \in \mathcal{A}_0} \{N_a\} \geq \min_{a \in \mathcal{A}_0} \mathbb{E}N_a - \sqrt{\frac{A}{A+1} \sum_{a \in \mathcal{A}_0} \text{Var}[N_a]} \\ &= N - \sqrt{\frac{A}{A+1}} \sqrt{\sum_{a \in \mathcal{A}_0} N \frac{1-\rho_a}{\rho_a}} = N - \sqrt{N} K(\rho). \end{aligned} \quad (13)$$

The result follows from combining (12) and (13). ■

A key observation in proving the last theorem was that the number of impressions in the left-over regime is  $O(\sqrt{N})$ . In fact, using a Chernoff bound, we may show that the probability that the number of impressions in the left-over regime exceeds a fraction of the total impressions decays exponentially fast.

**Corollary 1** *The probability that the number of impressions in the left-over regime exceeds a fraction  $\epsilon > 0$  of the total impressions decays exponentially fast, as given by*

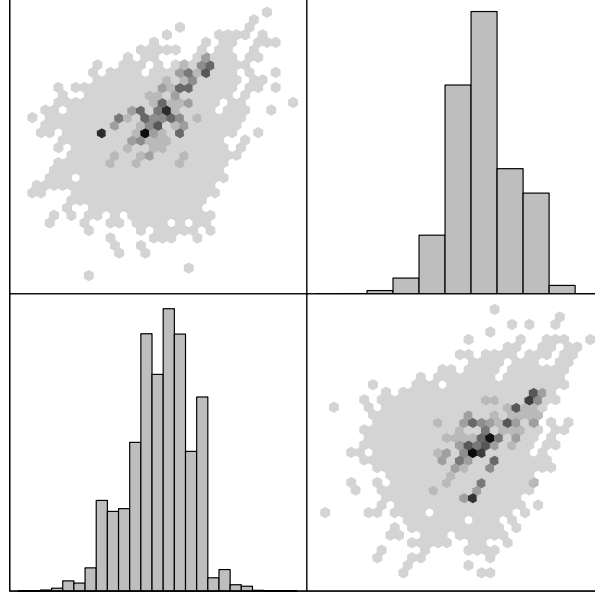
$$\mathbb{P}\{N - N^* \geq \epsilon N\} \leq \sum_{a \in \mathcal{A}_0} \exp(-2\epsilon^2 \rho_a N).$$

*Proof:* We prove the complement, that is, the probability that  $N^* \geq (1 - \epsilon)N$  converges exponentially fast to one. Notice that  $N^* \geq (1 - \epsilon)N$  if and only if by time  $(1 - \epsilon)N$  the contract of each advertiser is not yet fulfilled ( $S_{(1-\epsilon)N,a} < C_a$ ), and the point where all impressions need to be assigned to advertisers has not been reached ( $S_{(1-\epsilon)N,0} < C_0$ ). Combining De Morgan's law and Boole's inequality we get that

$$\mathbb{P}\{N^* \geq (1 - \epsilon)N\} = \mathbb{P}\{S_{(1-\epsilon)N,a} < C_a \ \forall a \in \mathcal{A}_0\} \geq 1 - \sum_{a \in \mathcal{A}_0} \mathbb{P}\{S_{(1-\epsilon)N,a} \geq C_a\}.$$

Recall that  $S_{(1-\epsilon)N,0}$  is the sum of  $(1 - \epsilon)N$  independent Bernoulli random variables with success probability  $\rho_a$ . Hence, we conclude by applying Chernoff's bound to the each summand to obtain  $\mathbb{P}\{S_{(1-\epsilon)N,a} \geq C_a\} \leq \exp(-2\epsilon^2 \rho_a N)$ . ■

*Proof of Proposition 3:* Because  $B = 0$ , then it is not hard to show that  $\bar{F}^{-1}(s) = 0$ , and that the revenue function is  $r(s) = 0$ . Hence, the revenue function is regular and satisfies Assumption 1. Moreover, the optimal survival probability is  $s^*(c) = 0$ , and  $R(c) = c$ . The result follows from substituting these functions in Theorem 1. ■



**Figure 6:** Graphical representation of the empirical distribution of log-quality for a type with two advertisers (data is log-transformed).

## D Data model

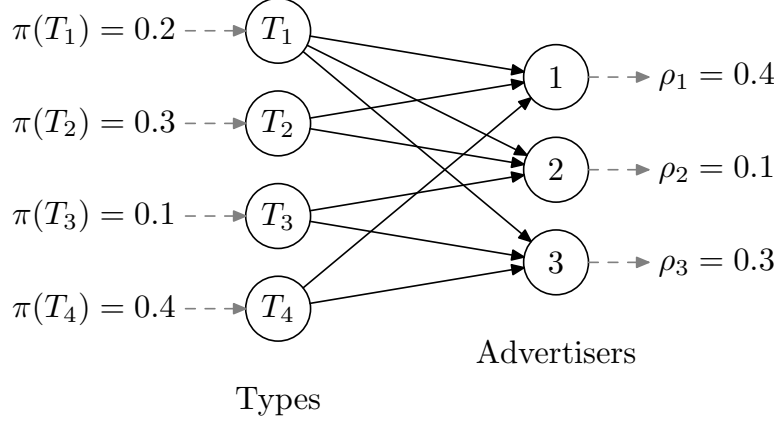
In this section we give a parametric model based on our observation of real data. Our model will be slightly more general than the discussion thus far. We characterize an impression type by the subset of advertisers  $T \subseteq \mathcal{A}$  that are interested on it. Although the number of types may be exponential in  $\mathcal{A}$ , in practice we observe that a linear number of them suffice to characterize 98% of the inventory. We observe that the predicted quality perceived by the advertisers within a type is approximately log-normal. This can be seen in Figure 6, where the empirical distribution of log-quality is graphically represented for a type with two advertisers (data is log-transformed). The histograms on the diagonal show the marginal log-quality of each advertiser, which approximately resemble a normal curve. On the off-diagonals, scatter plots show the correlation between advertisers, which is strongly positive. In some sense this is expected, since many advertisers have similar targeting criteria.

Given a particular type  $T$ , we assume that quality follows a multivariate log-normal with mean vector  $\mu_T$  and covariance matrix  $\Sigma_T$  for the advertisers in the type, and takes a value of  $-\tau_a$  for advertisers not in the type, that is, the advertiser pays a good-will penalty for assigned undesired impressions. The total distribution of quality is given by the mixture of these types distribution with mixing probabilities  $\pi(T)$ . Thus, we have that

$$Q \sim \begin{cases} \ln \mathcal{N}(\mu_T, \Sigma_T), & \text{for } a \in T, \\ -\tau_a, & \text{for } a \notin T, \end{cases} \quad \text{w.p. } \pi(T).$$

**D.1 Estimation** Data from four different publishers was analyzed for a consecutive period of seven days. First, the capacities of the reservations were used to compute the ratios  $\rho$ . Second, logs were analyzed to estimate the types frequencies, and the parameters of the underlying log-normal distributions. Figure 7 describes Instance 1, a publisher with 4 types and 3 advertisers. The parameters for the remaining publishers will be made available online in a later version.

With multiple bidders, AdX runs a sealed bid second-price auction. We analyze the first and second highest bids for the inventory submitted to AdX. In this first approach to the problem, we assume that bids are independent of the quality of the impressions. We denote by  $\{(B_1^m, B_2^m)\}_{m=1, \dots, M}$  the sampled highest and second highest bids from the exchange. Sample data is used to compute the two primitives of our model: (i) the complement of the quantile of the highest bid  $p(s)$ , and (ii) the revenue function of  $r(s)$ . Both functions are estimated on a uniform grid  $\{s_j\}_1^{100}$  of



Type	Ads	$\pi(T)$	$\mu_T$	$\Sigma_T$
$T_1$	{1, 2, 3}	0.2	$\begin{pmatrix} 7.8155 \\ 7.8155 \\ 7.8155 \end{pmatrix}$	$\begin{pmatrix} 0.3 & 0.1 & 0.1 \\ 0.1 & 0.3 & 0.1 \\ 0.1 & 0.3 & 0.1 \end{pmatrix}$
$T_2$	{1, 2}	0.3	$\begin{pmatrix} 6.6755 \\ 7.0655 \end{pmatrix}$	$\begin{pmatrix} 0.3180 & 0.1649 \\ 0.1649 & 0.3602 \end{pmatrix}$
$T_3$	{2, 3}	0.1	$\begin{pmatrix} 6.6355 \\ 7.8055 \end{pmatrix}$	$\begin{pmatrix} 0.4347 & 0.2357 \\ 0.2357 & 0.4367 \end{pmatrix}$
$T_4$	{1, 3}	0.4	$\begin{pmatrix} 7.2155 \\ 6.9155 \end{pmatrix}$	$\begin{pmatrix} 0.23 & 0.05 \\ 0.05 & 0.40 \end{pmatrix}$

**Figure 7:** Description of Instance 1.

survival probabilities in the  $[0, 1]$  range.

First, for each point in the grid  $j$ , the price  $p_j = p(s_j)$  is estimated as the  $(1 - s_j)$ -th population quantile of the highest bid. Then, using sampled bids, we estimate the revenue function w.r.t. to prices at the grid points as

$$r(p_j) = \frac{1}{M} \sum_{m=1}^M \mathbf{1}\{B_1^m \geq p_j\} \max\{B_2^m, p_j\} \quad (14)$$

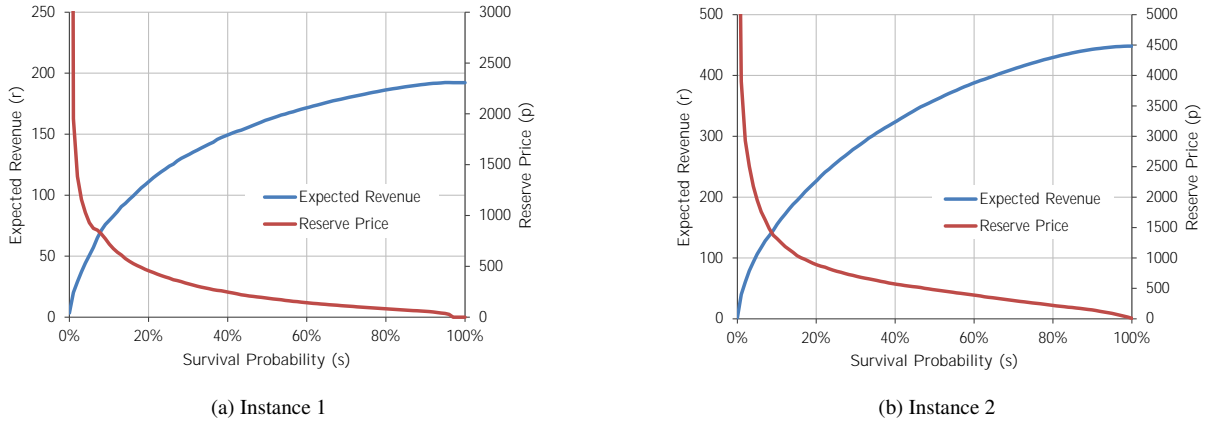
Finally, the revenue function is obtained by composing (14) and  $p(s)$ . The estimated survival probability and revenue function from two different publishers are shown in Figure 8.

## E Experimental Results

Two experiments were conducted to study our algorithm. First we study the impact of introducing an AdX on the publisher’s yield. Second, we compare the previously known primal-dual approach to ad allocation that is non-parametric to our approach here which is parametric.

**E.1 Impact of AdX** This first experiment explores the potential benefits of introducing an AdX, and how the publisher can take advantage of it. We study the impact of the trade-off parameter  $\gamma$  on both objectives, that is, the quality of the impressions assigned to the advertisers, and the revenue from AdX. The limiting choices of  $\gamma = 0$ , and  $\gamma = \infty$  are of particular interest. The first choice represents the case where the publisher disregards the quality of the impressions assigned to the advertisers, and strives to maximize the revenue extracted from AdX. Here the publisher strategically picks the reserve price so that just enough impressions are rejected to satisfy the contracts. In the second choice, the publishers prioritizes the quality of the impressions assigned, and submits the remanent inventory to AdX. We use this case as the baseline to which we compare our method.

The experiment was conducted as follows. First, we set up a grid on the trade-off parameter  $\gamma$ . Then, we solve the publisher’s problem as given in (3). The resulting policies are evaluated using a fluid limit (see I). Table 1 reports the expected quality and revenue for different choices of  $\gamma$ . Figure 2a plots the quality, and revenue relative to the baseline



**Figure 8:** Estimated survival probability and revenue function for AdX from two different publishers.

case; as a function of  $\gamma$ . In Figure 2b we plot, in a quality vs. revenue graph, the objective values of the optimal solutions for the different choices of  $\gamma$ , together with the Pareto frontier.

**Discussion.** Results confirm that, as we increase the trade-off parameter  $\gamma$ , the quality of the impressions assigned to the advertisers increases, while the revenue from AdX subsidizes. Interestingly, starting from the baseline case that disregards the AdX ( $\gamma = \infty$ ), we observe that the revenue from AdX can be substantially increased by sacrificing a small fraction of the overall quality of the impressions assigned. For instance, by exploiting strategically the AdX, the publisher can increase AdX’s revenue by 8% by giving up only 1% quality. Conversely, starting from the case that disregards the advertiser’s quality ( $\gamma = 0$ ), the publisher can raise the quality in a large amount at the expense of a small decrease in AdX’s revenue.

Alternatively, the previous analysis can be understood in terms of the Pareto frontier. Results show that the Pareto frontier is highly concave, relatively horizontally flat around  $\gamma = \infty$ , and vertically flat around  $\gamma = 0$ . This explains the huge marginal improvements at the extremes. There are several advantages to the quality vs. revenue representation. First, the Pareto frontier allows for quick grasp of the nature of the operation. When the publisher’s current operation is sub-optimal, its performance point should lie in the interior of the frontier. In this case, the Pareto frontier allows the publisher to measure its efficiency, and quantify the potential benefits an optimal policy may introduce. Second, when the choice of the trade-off parameter is not clear, the publisher may impose a lower bound on the overall quality of the impressions, and instead maximize the total revenue from AdX. The efficient frontier provides the maximum attainable revenue, and the proper  $\gamma$  to achieve the quality constraint.

**E.2 Comparison with the Primal-Dual Approach** In this second experiment we study the performance the our algorithm, and contrast it with a Primal-Dual (PD) method. This experiment was discussed in §6, but we give the full experiments and discussion here for the benefit of the reader. Since no existing PD method is known yet for the AdX problem, we consider instead the case with no AdX. The Primal-Dual approach [6], uses a sample from data to estimate the dual variables and uses it in a bid-price control policy. In contrast, our algorithm, as stated, assumes the parameters of the quality distribution are known, and uses that to estimate the dual variables. So, we do not need to use a sample. Of course in practice, the parameters need to be learned, and so we would need to use a sample of the data in order to learn them; but in many settings (including online advertising) it is reasonable to assume that we at least know the *form* of the distribution (e.g., normal, exponential, Zipf), albeit not the specific parameters (mean, variance, covariance, etc.). The techniques in [6] are powerful *because* they don’t need to assume anything about the distribution, but it is important to ask what can be gained from knowing the *form* of the distribution, which is what we do in the remainder of this section.

In order to objectively assess the performance of our algorithm we adopt the user type model described in §D as a generative model. The generative model is used to generate sample data on which both our algorithm and a PD method are tested. The advantages of adopting a generative model are twofold. First, it allows us to compute the truly optimal

Instance 1						
$\gamma$	0	0.001	0.01	0.05	0.075	0.1
Yield	110.94	112.78	128.81	202.46	249.11	296.95
Quality	1107.32	1779.07	1801.00	1864.95	1891.19	1913.78
Revenue	110.94	111.00	110.80	109.21	107.27	105.57
$\gamma$	0.25	0.5	0.75	1	10	$\infty$
Yield	590.54	1098.43	1608.80	2122.99	20764.82	$\infty$
Quality	1998.31	2044.89	2055.72	2061.33	2072.22	2075.52
Revenue	90.97	75.98	67.00	61.66	42.61	38.48

Instance 2						
$\gamma$	0	0.001	0.01	0.05	0.075	0.1
Yield	428.73	429.02	434.22	459.57	477.39	495.66
Quality	483.02	545.27	573.23	676.34	720.31	752.35
Revenue	428.73	428.47	428.49	425.75	423.37	420.42
$\gamma$	0.25	0.5	0.75	1	10	$\infty$
Yield	617.04	834.39	1056.05	1279.88	9425.63	$\infty$
Quality	843.47	880.69	891.49	896.89	906.46	907.05
Revenue	406.17	394.05	387.43	382.99	360.99	356.11

**Table 1:** Expected yield, advertisers' quality and revenue from AdX for two instances, and different choices of  $\gamma$ .

policy  $\mu^{\text{OPT}}$ . Second, the true performance of any policy can be evaluated efficiently using a fluid limit (see Section I).

The computational experiment is conducted as follows. First, a *training* data set of  $M$  impressions is generated. We denote the sampled quality vectors by  $\{q_m\}_{m=1}^M$ . Then, we estimate the parameters of the model on the training set as follows. For each type we estimate the type probabilities  $\hat{\pi}_T$ ; and mean  $\hat{\mu}_T$ , and covariance matrix  $\hat{\Sigma}_T$  of the logarithm of the qualities. Next, the dual problem (3) is solved on the estimated parametric model using a Gradient Descent Method as described in §H. Note that, since no AdX is considered, the maximum expected revenue function  $R(\cdot)$  is the identity. Using the optimal solution  $v^{\text{EST}}$  we construct a policy, which we refer as  $\mu^{\text{EST}}$ .

Simultaneously, we employ the PD method on the training data. The PD method amounts to solving a sample average approximation of problem (3), which results in the following linear program

$$\begin{aligned}
 \min_{v, \lambda} \quad & \frac{1}{M} \sum_{m=1}^M \lambda_m + \sum_{a \in \mathcal{A}} \rho_a v_a \\
 \text{s.t.} \quad & \lambda_m + v_a \geq q_{m,a}, \quad \forall m, a \\
 & \lambda_m \geq 0 \quad \forall m.
 \end{aligned} \tag{15}$$

The linear program is solved using CPLEX 12. Again, using the dual optimal solution  $v^{\text{PD}}$  we construct a policy  $\mu^{\text{PD}}$ .

Afterwards, we assess the performance of both policies using a fluid limit. These steps are replicated on 50 different training sets. Table 2 reports the average results over the training sets for different sizes of training sets, and instances. Plots of the results for a given instance are shown in Figure 4.

**Discussion.** Results show that for both algorithms, as the size of the training set increases, the optimality gap decreases at a rate of  $O(M^{\frac{1}{2}})$ . However, the parametric method performs uniformly better than the non-parametric PD method. Additionally, the variability across different training sets diminishes as the size of the training set increases. Indeed, we observe that the standard deviation over training sets converges to zero for both methods, but the convergence is faster for the parametric one. In some sense this is expected, since the true data model follows exactly the distributional assumptions. However, the PD method is expected to be more robust to model misspecification.

**Instance 1** (A = 3, T = 4, OPT = 2075.09)

Training Set Size	EST		PD	
	mean	std.dev.	mean.	std.dev
100	2004.16 (3.42%)	33.978	1990.32 (4.08%)	37.552
1000	2053.41 (1.04%)	10.008	2047.92 (1.31%)	12.365
2500	2065.12 (0.48%)	4.956	2062.76 (0.59%)	5.838
5000	2068.44 (0.32%)	3.681	2066.99 (0.39%)	4.224

**Instance 1** (A = 6, T = 10, OPT = 907.44)

Training Set Size	EST		PD	
	mean	std.dev.	mean.	std.dev
1000	889.58 (1.97%)	8.861	882.77 (2.72%)	12.829
2500	894.43 (1.43%)	7.485	887.64 (2.18%)	10.418
5000	898.59 (0.98%)	5.231	892.51 (1.65%)	7.625
10000	901.13 (0.70%)	3.588	897.42 (1.10%)	4.692
25000	904.69 (0.30%)	1.712	901.97 (0.60%)	2.720
50000	905.03 (0.27%)	1.267	903.44 (0.44%)	1.567

**Instance 3** (A = 17, T = 15, OPT = 894.82)

Training Set Size	EST		PD	
	mean	std.dev.	mean.	std.dev
2500	859.83 (3.91%)	9.937	849.44 (5.07%)	14.615
5000	868.61 (2.93%)	5.870	861.06 (3.77%)	7.954
10000	877.59 (1.92%)	5.226	873.46 (2.39%)	6.577
25000	884.04 (1.20%)	2.585	881.13 (1.53%)	3.747
50000	887.34 (0.84%)	1.926	885.11 (1.08%)	2.728

**Instance 4** (A = 14, T = 10, OPT = 928.76)

Training Set Size	EST		PD	
	mean	std.dev.	mean.	std.dev
2500	892.55 (3.90%)	12.886	888.88 (4.29%)	13.427
5000	903.04 (2.77%)	8.537	901.79 (2.90%)	10.277
10000	911.25 (1.88%)	6.951	909.96 (2.02%)	6.935
25000	917.30 (1.23%)	3.353	915.81 (1.39%)	3.705
50000	921.36 (0.80%)	2.668	920.11 (0.93%)	2.716

**Table 2:** Experimental results comparing the performance of our parametric method (EST) with the non-parametric primal-dual method (PD). No AdX present in this experiment.

Another experiment, though results are not reported, was conducted to test the strength of the parametric method on real data. We observed that, when the training set is small (around thousands), the parametric method performs better than the non-parametric one. However, as the sample size increases the non-parametric method outperforms the other. The rationale for this behavior is that, when data is scarce, the parametric method can exploit the distributional assumptions to reconstruct a fair representation of the data. However when data is larger, the fit of our model to real data is not perfect, and the non-parametric method can withstand deviations more robustly.

## F Extensions

In this section we consider a number of extensions of the model and policy from the previous section.

**F.1 Target Quality Constraints** In section 2.1 we discussed an alternate formulation in which the publisher imposes a minimum overall quality for the impressions assigned to the advertisers. This might be more natural for some publishers; they might feel more comfortable specifying target quality constraint than picking a Lagrange multiplier to weight the impact of quality in the objective. Additionally, in some settings the advertisers themselves might demand that certain level of quality is guaranteed.

In the following, we impose that the average quality of the impressions assigned to advertiser  $a$  is larger or equal than a threshold value  $\ell_a$ . Now, the publisher would strive to maximize the revenue from AdX, while complying with the target quality constraints, and the contractual obligations. The one-impression DAP would be similar, except that the objective only accounts for AdX's revenue, and the inclusion of the constraints

$$\mathbb{E}[i_a(Q)Q_a] \geq \ell_a, \quad \forall a \in \mathcal{A}. \quad (16)$$

We attack the problem, as done before, by considering its dual. Let  $\gamma_a \geq 0$  be the Lagrange multiplier associated to (16). As a side note, problem (2) can be interpreted as the Lagrange relaxation of our new problem w.r.t. the target quality constraints, and the dual variables  $\gamma$  as the shadow price of the target quality constraints. The new constraints preserve the convexity of the program, and strong duality still holds. Following the same steps, we obtain the new dual problem

$$\min_{\gamma \geq 0, v} \left\{ \mathbb{E}R \left( \max_{a \in \mathcal{A}_0} \{ \gamma_a Q_a - v_a \} \right) + \sum_{a \in \mathcal{A}} v_a \rho_a - \gamma_a \ell_a \right\},$$

which still is a convex minimization problem. The publisher might now jointly optimize over  $v$ , and  $\gamma$  to construct a provably good policy. Additionally, in a similar fashion to Proposition 9, we may compute the directional derivative of the objective w.r.t. the dual variables  $\gamma$ .

Regarding the performance the bid-price control, Theorem 2 still holds, and the policy asymptotically attains the optimal revenue from AdX, while complying with the delivery targets. However, we still need to argue about the expected average quality assigned to the advertisers. Unfortunately, for those advertisers whose constraint (16) is binding, our algorithm might not attain the desired quality target. Nevertheless, from our asymptotic analysis we may show that the expected average quality is lower bounded by

$$\mathbb{E} \left[ \frac{1}{N} \sum_{n=1}^N [1 - s_n^\mu(Q_n)] I_{n,a}^\mu(Q_n) Q_{n,a} \right] \geq \frac{\mathbb{E}N^*}{N} \mathbb{E}[i_a^*(Q)Q_a] \geq \left( 1 - \frac{1}{\sqrt{N}} K(\rho) \right) \ell_a.$$

Hence, for advertisers with binding constraint (16), albeit not feasible, the expected average quality becomes arbitrary close to the threshold value as the number of impressions in the horizon increases. On the other hand, for the remaining advertisers whose target quality constraint is not binding, the expected average quality will surpass the threshold for suitably large  $N$ .

**F.2 AdX with Multiple bidders** Here we generalize our results to the case where multiple buyers participate in the Ad Exchange. We model AdX as an auction with  $K$  risk neutral buyers. The publisher believes that individual valuations are drawn independently from the same distribution with c.d.f  $F(\cdot)$ , density  $f(\cdot)$ , and support  $[p_0, p_\infty]$ .



Moreover, we assume that the distribution of the values have increasing failure rates, are absolutely continuous and strictly monotonic. The publisher must choose the reservation price  $p$  that maximizes her expected revenue given that her value for the impression is  $c \geq 0$ . As before, we denote by  $R(c)$  the optimal expected revenue of the publisher.

[18] argued that under our assumptions the optimal mechanism is a Vickrey or second-price sealed-bid auction. Moreover, it is known that in such auctions bidding the true valuation is a dominant strategy for the buyers, and that the optimal reservation price  $p^*(c)$  is independent of the number of buyers [12].

Let  $B_{1:K}$  and  $B_{2:K}$  be the order statistics which denote the highest and the second highest bid respectively. Given a reserve price  $p$ , the item is sold if  $B_{1:K} \geq p$ , i.e., there is some bid higher than the reserve price. The winning buyer pays the second highest bid, or alternatively  $\max\{B_{2:K}, p\}$ , since the seller should receive at least the reserve price  $p$ . Therefore, the publisher's maximization problem is

$$R(c) = \max_{p \geq 0} \mathbb{E} [\mathbf{1}\{B_{1:K} \geq p\} \max\{B_{2:K}, p\} + \mathbf{1}\{B_{1:K} < p\}c].$$

Recall that, instead of reserve prices, we casted our problem in terms of survival or winning probabilities. Then, letting  $s$  be the probability than the impression is sold, we have that  $s = \mathbb{P}\{B_{1:K} \geq p\} = 1 - F^K(p)$  since valuations are i.i.d. Conversely, the reserve price as a function of the survival probability is given by  $p(s) = \bar{F}^{-1}(1 - (1 - s)^{1/K})$ , which is well-defined due to the strict monotonicity of the c.d.f. In terms of survival probabilities, the problem is now

$$R(c) = \max_{0 \leq s \leq 1} r(s) + (1 - s)c,$$

where we defined the revenue function as  $r(s) = r(p(s))$ , and  $r(p) = \mathbb{E}[\mathbf{1}\{B_{1:K} \geq p\} \max\{B_{2:K}, p\}]$ .

The next proposition shows that the revenue function is regular, and as a consequence all previous results hold for the case with multiple bidders.

**Proposition 7** *Under the previous assumption, the revenue function  $r(s)$  is regular. Moreover, the optimal reserve price  $p^*(c)$  solves*

$$\frac{\bar{F}(p)}{f(p)} = p - c,$$

when  $c \in [p_0 - 1/f(p_0), p_\infty]$ . When the opportunity cost is higher than the null price ( $c > p_\infty$ ), the publisher bypasses the exchange ( $p^*(c) = p_\infty$ ). Finally, when the opportunity cost is low enough ( $c < p_0 - 1/f(p_0)$ ), the impression is kept by the highest bidder ( $p^*(c) = p_0$ ).

*Proof:* The joint distribution of  $B_{1:K}$  and  $B_{2:K}$  has a density function [12]

$$f(b_1, b_2) = \begin{cases} K(K-1)F(b_2)^{K-2}f(b_1)f(b_2) & \text{if } b_1 \geq b_2 \\ 0 & \text{otherwise} \end{cases}.$$

Then, we have that

$$\begin{aligned} r(p) &= \mathbb{E}[\mathbf{1}\{B_{2:K} \geq p\}B_{2:K} + p\mathbf{1}\{B_{1:K} \geq p, B_{2:K} < p\}] \\ &= \int_p^\infty \int_p^{b_1} b_2 f(b_1, b_2) db_2 db_1 + p \int_p^\infty \int_0^p f(b_1, b_2) db_2 db_1 \\ &= K(K-1) \int_p^\infty b_2 F(b_2)^{K-2} f(b_2) (1 - F(b_2)) db_2 + KpF(p)^{K-1}(1 - F(p)) \end{aligned}$$

Continuity of  $r(s)$  follows because the p.d.f. is continuous, and  $p(s)$  is continuous (if  $F$  not strictly monotone, the inverse may have jumps). Additionally, we may bound the revenue by

$$r(p) \leq \mathbb{E}[\mathbf{1}\{B_{1:K} \geq p\}B_{1:K}] \leq K\mathbb{E}[\mathbf{1}\{B \geq p\}B] \leq K\mathbb{E}B < \infty,$$

the first inequality follows because  $B_{1:K}$  is the maximum, the second because any order statistic is upper bounded by the sum of the bids, and the fourth because bids are integrable. Moreover, integrability of  $B$  implies that  $\lim_{p \rightarrow \infty} r(p) = 0$ .

Next, we turn to the concavity of  $r(s)$ . Differentiating w.r.t to  $p$  we get

$$\frac{dr}{dp} = KF(p)^{K-1}(\bar{F}(p) - pf(p)).$$

Then, using the fact that  $\frac{ds}{dp} = -KF(p)^{1-k}/f(p)$  we get from the composition rule that

$$\frac{dr}{ds} = \left. \frac{dr}{dp} \right|_{p(s)} \frac{dp}{ds} = p(s) - \frac{1}{h(p(s))},$$

where  $h(p) = f(p)/\bar{F}(p)$  is the hazard rate of the bidder's valuation. Because  $p(s)$  is non-increasing in  $s$  and the  $h(p)$  is non-decreasing in  $p$ , we conclude that  $\frac{dr}{ds}$  is non-increasing. Thus, the revenue function is concave.

Finally, notice that the that derivative of the objective w.r.t to  $s$  is

$$p(s) - \frac{1}{h(p(s))} - c, \quad (17)$$

which is non-increasing. When  $c > p_\infty$  we have that (17) is negative, so  $s^*(c) = 0$  and  $p^*(c) = p_\infty$ . Similarly, when  $c < p_0 - 1/h(p_0)$  we that (17) is positive, so  $s^*(c) = 1$  and  $p^*(c) = p_0$ . ■

Notice that by writing  $\max\{B_{2:K}, p\} = p + (B_{2:K} - p)^+$  we get that the setup of Section 2.2 is the particular case where  $K = 1$ .

**F.3 AdX with User Information** In most systems, the publisher shares some user information with the exchange. In turn, the exchange may partially disclose the user information to their advertisers. The advertisers may react to this information, and bid strategically [17]. In this section we extend our model to the case when the bids from AdX are correlated with the quality of the impression (a surrogate for user information). For simplicity we consider the case of one bidder. Nevertheless, our analysis can be easily extended to the general case.

Let  $\bar{F}(p|Q) = \mathbb{P}\{B \geq p \mid Q\}$  be the conditional probability that the bid from AdX is greater than  $p$  given that the impression quality vector is  $Q$ . Additionally, we define the *conditional revenue function* as  $r(s|Q) = s\bar{F}^{-1}(s|Q)$ . The publisher can exploit the correlation between user information and bids to update his prior on AdX bids. Conditioning on the impression quality, we obtain that the maximum expected revenue under opportunity cost  $c$ , denoted by  $R(c|Q)$ , is now

$$R(c|Q) = \max_{0 \leq s \leq 1} r(s|Q) + (1-s)c. \quad (18)$$

In order to apply the results from the previous sections we require that the conditional revenue function  $r(\cdot|Q)$  is regular for all qualities  $Q$  almost surely.

The rest of the analysis follows in a straightforward way by conditioning on the impression quality. The dual problem (3) now reads

$$\min_v \left\{ \mathbb{E}R \left( \max_{a \in \mathcal{A}_0} \{Q_a - v_a\} \mid Q \right) + \sum_{a \in \mathcal{A}} v_a \rho_a \right\},$$

where we replaced the maximum expected revenue by  $R(\cdot|Q)$ . It is worth noting that now the optimal reserve price to be submitted to AdX depends both on the maximum discounted quality, and the actual realization of the quality vector.

**F.4 User Types** Previously we assumed that any user could be potentially assigned to any advertiser. In practice, however, advertisers have specific targeting criteria. For instance, a guaranteed contract may demand for females with certain age range living in New York, while other contract may demand for males in California. Next, we show how

to handle this by forcing the publisher to pay a good-will penalty to the advertisers each time an undesired impression is incorrectly assigned.

Instead of grouping user types according to their attributes, we aggregate user types that match the criteria of the same subset of advertisers. This has the advantage of reducing the space of types to a function of the number of advertisers (which is typically small in practice) rather than the number of possible types (which is potentially large). Hence, a user type is characterized by the subset of advertisers  $T \subseteq \mathcal{A}$  that are interested in it. In the following, we let  $\mathcal{T}$  be the support of the type distribution, and  $\pi(T)$  the probability of an arriving impression being of type  $T$ . As before we assume that, across different impressions, types are independent and identically distributed. Given a particular type  $T$ , the predicted quality perceived by the advertisers within the type is modeled by the non-negative random vector  $Q(T) = \{Q_a(T)\}_{a \in T}$ .

Even if the total number of impressions suffices to satisfy the contracts, i.e.  $\sum_{a \in \mathcal{A}} \rho_a \leq 1$ , the inventory may not be enough to satisfy the contracts targeting criteria. Our algorithm guarantees that the total number of impressions  $C_a$  is always respected, yet some advertisers may be assigned impressions outside of their criteria. If an impression of type  $T$  happens to be assigned to an advertiser  $a \notin T$ , the publishers pays a nonnegative goodwill penalty  $\tau_a$ . These penalties allow the publisher to prioritize certain reservations, specially when the contracts are not feasible. Thus, the ex-ante distribution of quality is given by the mixture of the types distribution with mixing probabilities  $\pi(T)$ , and we have that

$$Q \sim \begin{cases} Q_a(T), & \text{for } a \in T, \\ -\tau_a, & \text{for } a \notin T, \end{cases} \quad \text{w.p. } \pi(T).$$

Notice that all our previous results hold if we apply the same analysis to the mixture distribution.

As we stated, the inventory may not be enough to satisfy the contracts targeting criteria. Even if penalties are high enough, it may be the case that advertisers are assigned impressions outside their criteria. It is straightforward to state the problem of determining whether contracts can be satisfied or not, as a feasible flow problem on a bipartite graph. The problem can be formulated on an graph with one node for each user type  $T$  with a supply of  $\pi(T)$ , on the left side; and one node for each advertisers  $a \in \mathcal{A}_0$  with a demand  $\rho_a$ , on the right side. Then, we say that the operation is *feasible* if the user type-advertiser graph admits a feasible flow.

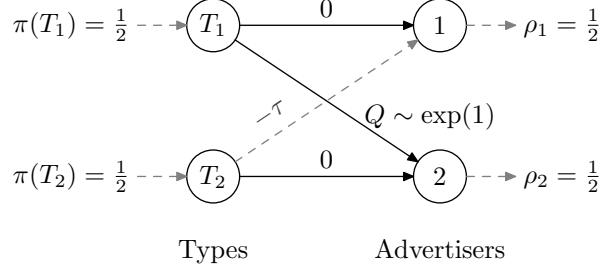
The feasibility of the operation, albeit necessary, does not suffice to guarantee that no impressions outside the targeting criteria are assigned to the advertisers. When advertisers compete for the same type, and one of them obtains a potentially unbounded reward for that type; it may be optimal to allow the latter advertiser to cannibalize the user type, and force the others advertisers to take types outside of their criteria. This may occur, surprisingly, for all conceivable penalties. However, if qualities are bounded, and penalties are set high enough, then the optimal policy would not recommend the assignment of impressions outside the targeting criteria. Even in this case some impressions may be incorrectly assigned in the left-over regime, but the probability of this event decays exponentially fast. We formalize this discussion in the following proposition.

**Proposition 8** *Suppose that the user type-advertiser graph admits a feasible flow, and that qualities and bids from AdX are bounded by  $\frac{1}{A} \min_a \tau_a$ . Then, the stochastic control policy does not assign any impressions outside of the targeting criteria, except perhaps for the left-over regime.*

Note that since the left-over regime is vanishingly small in proportion to the length of the horizon (Cor. 1) this implies that the number of unassigned impressions is small. Thus in practice, a publisher may set  $C'_a = C_a + \epsilon$ , discard any impressions assigned by the policy outside the targeting criteria, and ensure that contracts are filled properly.

Next, we prove by example that the requirement that qualities are bounded is necessary for the previous result to hold. Consider a publisher who contracts with two advertisers, and agrees to deliver one half of the arriving impressions to each one of them. Additionally, there are two impression types, denoted by  $T_1$  and  $T_2$ , each occurring 50% of the time. The first advertiser only cares about the first type. She obtains a reward of zero for  $T_1$ , and the advertisers pays a positive penalty  $\tau$  each time a  $T_2$  impression is assigned to her. The second advertisers admits both types, but only obtains a positive reward  $Q \sim \exp(1)$  for the first type. The setup is shown in Figure 9.

A feasible policy could assign all  $T_1$  impressions to the first advertiser, and  $T_2$  impressions to the second advertiser. However, such policy is not optimal. Notice that both advertisers compete for the  $T_1$  impressions, and the first



**Figure 9:** Example with two user types, and two advertisers.

advertiser could extract a potentially high quality from them. It is not hard to see that the optimal dual variables are  $v_1 = -\tau$ , and  $v_2 = 0$ ; and the optimal objective value is  $\frac{1}{2}\mathbb{E}[Q - \tau]^+ = \frac{1}{2}e^{-\tau}$ . Hence, it is optimal to assign those  $T_1$  impressions with quality greater than  $\tau$  to the second advertiser. Thus, no matter the value of the penalty, a fraction  $e^{-\tau}$  of the total impression assigned to the first advertiser are undesired.

## G Proofs

In this section we give proofs for those results that were not included in the main body.

*Proof of Proposition 1:* First, observe that for all  $c$  the objective function of (1) is concave and continuous in  $s$ , and the feasible set is compact. Hence, by Weierstrass Theorem the set of optimal solutions is non-empty and compact. Thus, both  $R(c)$  and  $s^*(c)$  are well-defined.

Second,  $R(c) \geq c$  follows from letting  $s = 0$ . To see that  $R(c)$  is non-increasing, let  $c < c'$ , and  $s^*$  be the optimal solution under cost  $c$ . Then,  $R(c) = r(s^*) + (1 - s^*)c \leq r(s^*) + (1 - s^*)c' \leq R(c')$  where the first inequality follows because  $s^* \leq 1$ , and the second because no solution is better than the optimal. To see that  $R(c) - c$  is non-increasing, let  $c' < c$ , and  $s^*$  be the optimal solution under cost  $c$ . Then, by a similar argument we get that  $R(c') - c' \geq r(s^*) - s^*c' \geq r(s^*) - s^*c = R(c) - c$ . Convexity follows in a similar way (this is a standard result).

Third, observe that the objective function of (1) is jointly continuous in  $s$  and  $c$ . Thus, by the Maximum Theorem  $R(c)$  is continuous in  $c$ , and  $s^*(c)$  is upper-hemicontinuous.

Finally, because  $r(s) + (1 - s)c$  has decreasing differences in  $(s, c)$  and the feasible set is a lattice, by Topkis's Theorem  $s^*(c)$  is non-increasing in  $c$ . The result for  $p^*(c)$  follows from the fact that  $F^{-1}(s)$  is non-increasing in  $s$ . ■

*Proof of Theorem 1:* The optimality conditions of  $v$  for problem (3) imply that the directional derivative of  $\psi(v)$  along any direction is greater or equal to zero. In particular for each advertiser  $a \in \mathcal{A}$  it should be the case that  $\nabla_{\mathbf{1}_a} \psi(v) \geq 0$ , and  $\nabla_{-\mathbf{1}_a} \psi(v) \geq 0$ . Applying proposition 9 to both directions, together with the fact that there is zero probability of a tie occurring, we get that

$$\mathbb{E}[(1 - s^*(Q_a - v_a)) \mathbf{1}\{Q_a - v_a > Q_{a'} - v_{a'} \forall a' \in \mathcal{A}_0\}] = \rho_a,$$

and the result follows. ■

*Proof of Proposition 4:* For (4), we use that  $\frac{\partial \bar{G}^{-1}}{\partial \theta}(\rho|\theta) = \frac{\partial \bar{G}}{\partial \theta}(v|\theta)/g(v|\theta)$ , which follows from the implicit function theorem. In view of Cramér-Rao lower bound, we have that  $\varepsilon(\theta) \geq 1$ .

Next, we look at the average yield of the policy as the number of impressions grows to infinity when a bid price of  $u$  is employed, denoted by  $\bar{J}(u)$ . The limiting performance is given by

$$\bar{J}(u) = \begin{cases} \rho \mathbb{E}_\theta[Q|Q \geq u], & \text{if } u < v, \\ \mathbb{E}_\theta[Q] - (1 - \rho) \mathbb{E}_\theta[Q|Q \leq u], & \text{if } u \geq v. \end{cases}$$

Under our assumptions, the performance function is continuous  $u$ . One would be tempted to apply the Delta Method to derive the asymptotic distribution of the performance. Unfortunately,  $\bar{J}(\cdot)$  is not differentiable at  $v$ . However, it is the

case that the performance function is semi-differentiable at  $v$  with finite right-derivative  $\bar{J}'_+(v) \geq 0$  and left-derivative  $\bar{J}'_-(v) \leq 0$ . Thus, we can apply an extension of the Delta Method for directionally differentiable functions proved by [19], and obtain

$$\begin{aligned}\sqrt{m}(\bar{J}(\hat{v}_M) - \bar{J}(v)) &\Rightarrow d\bar{J}(v; \mathcal{N}(0, u'(\theta))), \\ \sqrt{m}(\bar{J}(\hat{v}_M^{\text{mle}}) - \bar{J}(v)) &\Rightarrow d\bar{J}(v; \mathcal{N}(0, u(\theta))),\end{aligned}$$

where  $d\bar{J}(v; \xi)$  is the Gâteaux derivative of  $\bar{J}$  at the point  $v$  along the direction  $\xi$ , which is given by  $d\bar{J}(v; \xi) = \bar{J}'_+(v)\xi$  when  $\xi \geq 0$ , and  $d\bar{J}(v; \xi) = \bar{J}'_-(v)\xi$  when  $\xi < 0$ . Note that the asymptotic variance of the performances are  $u'(\theta) \cdot K$ , and  $u(\theta) \cdot K$  respectively, where the performance scale factor is given by  $K = \frac{1}{2}(\bar{J}'_+(v)^2 + \bar{J}'_-(v)^2) - \frac{1}{2\pi}(\bar{J}'_+(v) - \bar{J}'_-(v))^2$ . Thus, the relative efficiency of the performance is identical to  $\varepsilon(\theta)$ . ■

*Proof:*[Proof of Proposition 2] Let  $\bar{s} = \{s_n(\cdot)\}_{n=1, \dots, N}$  and  $\bar{v} = \{i_n(\cdot)\}_{n=1, \dots, N}$  be any feasible vectors of controls. Let  $\bar{s}$  be the mean of the controls (in terms of prices  $\bar{p}$  would be the generalized  $\bar{F}$ -mean), which is defined as

$$\bar{s}(Q) = \frac{1}{N} \sum_{n=1}^N s_n(Q) \quad \forall Q \text{ pointwise,}$$

and let  $\bar{v}$  be such that

$$\bar{v}_a(Q) = \frac{1}{N} \sum_{n=1}^N i_{n,a}(Q) \quad \forall Q \text{ pointwise, } a \in \mathcal{A}.$$

We will show that solution in which  $(\bar{s}, \bar{v})$  are used for all impressions is a feasible control with greater or equal revenue than the original one.

First, for the feasibility of  $(\bar{s}, \bar{v})$  observe that for each advertiser  $a \in \mathcal{A}$

$$C_a = \mathbb{E}_Q \left[ \sum_{n=1}^N i_{n,a}(Q) \right] = N \mathbb{E}_Q [\bar{v}_a(Q)],$$

where the first equation follows from the feasibility of  $(\bar{p}, \bar{v})$  and the linearity of expectation, and the second from substituting  $\bar{v}_a$  pointwise for  $Q$ . Additionally, it is not hard to prove that  $\sum_{a \in \mathcal{A}} \bar{v}_a(\cdot) + \bar{s}(\cdot) \leq 1$ .

Second, we denote by  $J^D(\bar{s}, \bar{v})$  and  $J^D(\bar{s}, \bar{v})$  the objective value of the solutions  $(\bar{s}, \bar{v})$  and  $(\bar{s}, \bar{v})$  respectively. We have that

$$\begin{aligned}J^D(\bar{s}, \bar{v}) &= \mathbb{E} \left[ \sum_{n=1}^N r(s_n(Q)) + \sum_{a \in \mathcal{A}} Q_a \sum_{n=1}^N i_{n,a}(Q) \right] \\ &\leq \mathbb{E} \left[ Nr \left( \frac{1}{N} \sum_{n=1}^N s_n(Q) \right) + \sum_{a \in \mathcal{A}} Q_a \sum_{n=1}^N i_{n,a}(Q) \right] \\ &= N \mathbb{E} \left[ r(\bar{s}(Q)) + \sum_{a \in \mathcal{A}} Q_a \bar{v}_a(Q) \right] = J^D(\bar{s}, \bar{v}),\end{aligned}$$

where the inequality follows from the concavity of the revenue function, and the second equality from substituting  $\bar{s}$  and  $\bar{v}$  pointwise for all  $Q$ . ■

Given a subset of the quality space  $D \subseteq \Omega$ , we define the measure  $\mathbb{P}_R(D)$  as the probability that the quality vector belongs to that subset and the impression is rejected by the Ad Exchange when the optimal survival probability is used. More formally,

$$\mathbb{P}_R(D) = \mathbb{E} \left[ \left( 1 - s^* \left( \max_{a \in \mathcal{A}_0} \{Q_a - v_a\} \right) \right) \mathbf{1}\{Q \in D\} \right].$$

Notice that the latter is not a probability measure since  $\mathbb{P}_R(\Omega) \leq 1$ .

Proposition 9 characterizes the directional derivative of the objective function of the dual along some directions that, as we will show later, are of particular interest. Results are given in terms of the measure  $\mathbb{P}_R$ .

**Proposition 9** *Given a subset  $\alpha \in \mathcal{A}$ , the directional derivative of the objective function of the dual w.r.t. directions  $\mathbf{1}_\alpha$  and  $-\mathbf{1}_\alpha$  are respectively*

$$\begin{aligned}\nabla_{\mathbf{1}_\alpha} \psi(v) &= -\mathbb{P}_R \left\{ \max_{a \in \alpha} \{Q_a - v_a\} > \max_{a \in \mathcal{A}_0 \setminus \alpha} \{Q_a - v_a\} \right\} + \sum_{a \in \alpha} \rho_a, \\ \nabla_{-\mathbf{1}_\alpha} \psi(v) &= \mathbb{P}_R \left\{ \max_{a \in \alpha} \{Q_a - v_a\} \geq \max_{a \in \mathcal{A}_0 \setminus \alpha} \{Q_a - v_a\} \right\} - \sum_{a \in \alpha} \rho_a.\end{aligned}$$

*Proof:* We consider first the direction  $\mathbf{1}_\alpha$ . The difference quotient at  $v$  of  $\psi$  given a step  $\delta > 0$  is

$$\begin{aligned}\frac{\psi(v + \delta \mathbf{1}_\alpha) - \psi(v)}{\delta} &= \frac{1}{\delta} \mathbb{E} \left[ R \left( \max_{a \in \mathcal{A}_0} \{Q_a - v_a - \delta \mathbf{1}_{a \in \alpha}\} \right) - R \left( \max_{a \in \mathcal{A}_0} \{Q_a - v_a\} \right) \right] + \sum_{a \in \alpha} \rho_a, \\ &= \mathbb{E} \gamma_\delta(Q) + \sum_{a \in \alpha} \rho_a,\end{aligned}$$

where we defined  $\gamma_\delta(Q)$  as the difference quotient random variable. In order to obtain the directional derivative along vector  $\mathbf{1}_\alpha$  we need to look at the limit of the difference quotient as  $\delta$  goes to zero. We will first show that the  $\gamma_\delta(Q)$  converges to the desired expression, and then exchange limit and expectation using the dominated convergence theorem.

In the following we denote  $Q_\alpha - v_\alpha = \max_{a \in \alpha} \{Q_a - v_a\}$  and  $Q_{\alpha'} - v_{\alpha'} = \max_{a \in \mathcal{A}_0 \setminus \alpha} \{Q_a - v_a\}$ . Using this notation, the difference quotient random variable can be written as

$$\begin{aligned}\gamma_\delta(Q) &= \frac{R \left( \max(Q_\alpha - v_\alpha - \delta, Q_{\alpha'} - v_{\alpha'}) \right) - R \left( \max(Q_\alpha - v_\alpha, Q_{\alpha'} - v_{\alpha'}) \right)}{\delta} \\ &= -\frac{R(Q_\alpha - v_\alpha) - R(Q_\alpha - v_\alpha - \delta)}{\delta} \mathbf{1} \{Q_{\alpha'} - v_{\alpha'} + \delta \leq Q_\alpha - v_\alpha\} \\ &\quad + \frac{R(Q_{\alpha'} - v_{\alpha'}) - R(Q_\alpha - v_\alpha)}{\delta} \mathbf{1} \{Q_{\alpha'} - v_{\alpha'} < Q_\alpha - v_\alpha < Q_{\alpha'} - v_{\alpha'} + \delta\},\end{aligned}$$

where in the last equation we partitioned the difference based on which of the three terms verifies the maximum. Note that when  $Q_{\alpha'} - v_{\alpha'} \geq Q_\alpha - v_\alpha$  the difference is zero.

Because  $R(c) - c$  is non-increasing, we have that for  $c_1 \leq c_2$  then  $(R(c_2) - R(c_1))/(c_2 - c_1) \leq 1$ . So the difference quotients in each term are bounded by 1. As a result,  $|\gamma_\delta(Q)| \leq \mathbf{1} \{Q_{\alpha'} - v_{\alpha'} + \delta < Q_\alpha - v_\alpha\}$ , which is integrable. Moreover, the point-wise limit of  $\gamma_\delta(Q)$  as  $\delta \rightarrow 0^+$  is

$$\lim_{\delta \rightarrow 0^+} \gamma_\delta(Q) = -R'(Q_\alpha - v_\alpha) \mathbf{1} \{Q_{\alpha'} - v_{\alpha'} < Q_\alpha - v_\alpha\}$$

where we used the fact that first term converges to the derivative of  $R$  at  $Q_\alpha - v_\alpha$ , and the second term converges point-wise to zero.

Finally, by the dominated convergence theorem together with the fact that  $R'(c) = 1 - s^*(c)$  we get that

$$\begin{aligned}\nabla_{\mathbf{1}_\alpha} \psi(v) &= \lim_{\delta \rightarrow 0^+} \frac{\psi(v + \delta \mathbf{1}_\alpha) - \psi(v)}{\delta} = \lim_{\delta \rightarrow 0^+} \mathbb{E} \gamma_\delta(Q) + \sum_{a \in \alpha} \rho_a \\ &= \mathbb{E} \lim_{\delta \rightarrow 0^+} \gamma_\delta(Q) + \sum_{a \in \alpha} \rho_a = -\mathbb{P}_R \left\{ \max_{a \in \alpha} \{Q_a - v_a\} > \max_{a \in \mathcal{A}_0 \setminus \alpha} \{Q_a - v_a\} \right\} + \sum_{a \in \alpha} \rho_a.\end{aligned}$$

A similar result follows for the opposite direction  $-\mathbf{1}_\alpha$ . ■

*Proof:*[Proof of Proposition 5] The proof proceeds by contradiction, that is, we assume that there is no feasible flow. First, we cast the feasible flow problem as a maximum flow problem. Feasibility would imply the existence of a flow with value  $1 - \mathbb{P}(\emptyset\text{-tie})$ . But since we assume that no such feasible flow exists, by the max-flow min-cut theorem there should exist a cut with value strictly less than  $1 - \mathbb{P}(\emptyset\text{-tie})$ . The contradiction arises because the optimality conditions of  $v$  for the dual problem (3) imply that every cut is lower bounded by  $1 - \mathbb{P}(\emptyset\text{-tie})$ .

In order to write the feasible flow problem as a maximum flow problem, we first add a source  $s$  and a sink  $t$ . Second, we add one arc from  $s$  to each node associated to a non-empty subset  $S \subseteq \mathcal{A}_0$  (left-hand side nodes) with capacity  $\mathbb{P}(S\text{-tie})$ . Third, we add one arc from each advertiser  $a \in \mathcal{A}_0$  (right-hand side nodes) to  $t$  with capacity  $\rho_a$ . Lastly, we set the capacity of arcs from  $S$  to  $a \in S$  to infinity.

Now, since no feasible flow exists, by the max-flow min-cut theorem there should be a cut with value strictly less than  $1 - \mathbb{P}(\emptyset\text{-tie})$ . Let  $\alpha \subseteq \mathcal{A}_0$  be the advertiser nodes (right-hand) belonging to the  $t$  side of a *minimum* cut. Figure 10 shows the minimum cut. Next we argue that subset nodes in the  $s$  side verify that  $S \cap \alpha = \emptyset$ , while those in the  $t$  side verify that  $S \cap \alpha \neq \emptyset$ . First, because the cut has minimum value, there is no arc from a subset node to an advertiser node crossing the cut (those arcs have infinity capacity). Equivalently, within the  $s$  side of the cut, all subset nodes  $S \subseteq \mathcal{A}_0$  should verify that  $S \cap \alpha = \emptyset$ . Second, observe that any subset node with  $S \cap \alpha = \emptyset$  in the  $t$  side of the cut could be moved to the  $s$  side of the cut without increasing the value of the cut. Hence, with no loss of generality we can assume that all subset nodes in the  $t$  side of the cut verify that  $S \cap \alpha \neq \emptyset$ .

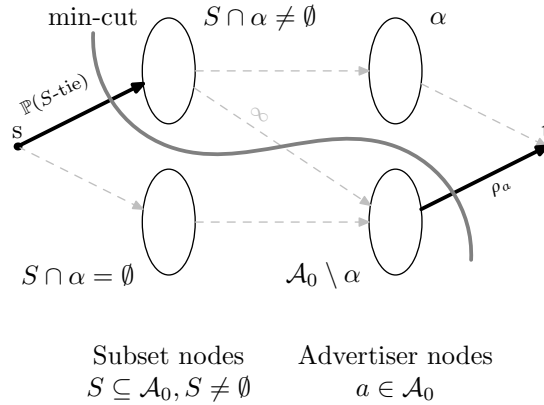
As a consequence, the only arcs crossing the cut are those from the source to the subsets  $S \cap \alpha \neq \emptyset$ , and those from advertisers  $\mathcal{A}_0 \setminus \alpha$  to the sink. The value of this cut is

$$\sum_{S \subseteq \mathcal{A}_0: S \cap \alpha \neq \emptyset} \mathbb{P}(S\text{-tie}) + \sum_{a \in \mathcal{A}_0 \setminus \alpha} \rho_a.$$

Because the value is strictly less than  $1 - \mathbb{P}(\emptyset\text{-tie})$  we get that

$$\sum_{S \subseteq \mathcal{A}_0: S \cap \alpha \neq \emptyset} \mathbb{P}(S\text{-tie}) < \sum_{a \in \alpha} \rho_a, \quad (19)$$

where we used that  $\sum_{a \in \mathcal{A}} \rho_a + \rho_0^{\text{eff}} = 1 - \mathbb{P}(\emptyset\text{-tie})$ .



**Figure 10:** The flow problem in the bipartite graph with a minimum cut. A source  $s$  connected to the subset nodes and a sink  $t$  connected to the advertiser nodes was included.  $\alpha \subseteq \mathcal{A}_0$  is the subset of advertiser nodes (right-hand) belonging to the  $t$  side. Note that there is no arc from a subset node to an advertiser node crossing the cut.

Next, we look at the optimality conditions of  $v$  for the dual problem (3). We distinguish between the case that  $0 \notin \alpha$  and  $0 \in \alpha$ . First suppose that  $0 \notin \alpha$ , and consider the direction  $-1_\alpha$  that has a  $-1$  if  $a \in \alpha$  and  $0$  elsewhere.

According to proposition 9 the directional derivative of the objective at  $v$  is

$$\begin{aligned}\nabla_{-\mathbf{1}_\alpha} \psi(v) &= \mathbb{P}_R \left\{ \max_{a \in \alpha} \{Q_a - v_a\} \geq \max_{a \in \mathcal{A}_0 \setminus \alpha} \{Q_a - v_a\} \right\} - \sum_{a \in \alpha} \rho_a \\ &= \sum_{S \subseteq \mathcal{A}_0: S \cap \alpha \neq \emptyset} \mathbb{P}(S\text{-tie}) - \sum_{a \in \alpha} \rho_a,\end{aligned}$$

where we have written the event that the maximum is verified non-exclusively by some advertiser  $a \in \alpha$  as all  $S$ -ties in which some advertiser  $a \in \alpha$  is involved. The optimality of  $v$  implies that the directional derivative along that direction is greater or equal to zero, contradicting equation (19).

When  $0 \in \alpha$  we consider the direction  $\mathbf{1}_{\mathcal{A} \setminus \alpha}$  that has a 1 if  $a \notin \alpha$  and 0 elsewhere. The direction derivative is now

$$\begin{aligned}\nabla_{\mathbf{1}_{\mathcal{A} \setminus \alpha}} \psi(v) &= -\mathbb{P}_R \left\{ \max_{a \in \mathcal{A} \setminus \alpha} \{Q_a - v_a\} > \max_{a \in \alpha \cup \{0\}} \{Q_a - v_a\} \right\} + \sum_{a \in \mathcal{A} \setminus \alpha} \rho_a, \\ &= - \sum_{S \subseteq \mathcal{A}_0: S \subseteq \mathcal{A} \setminus \alpha} \mathbb{P}(S\text{-tie}) + \sum_{a \in \mathcal{A} \setminus \alpha} \rho_a = \sum_{S \subseteq \mathcal{A}_0: S \cap \alpha \neq \emptyset} \mathbb{P}(S\text{-tie}) - \sum_{a \in \alpha} \rho_a,\end{aligned}$$

where in the second equation we have written the event that the maximum is verified exclusively by some advertiser  $a \in \alpha$  as all  $S$ -ties in which only advertisers in  $\alpha$  are involved. Again, the optimality of  $v$  implies that the directional derivative along that direction is greater or equal to zero, contradicting equation (19). ■

*Proof:*[Proof Sketch of Proposition 8] For simplicity we consider the case without AdX. Let  $i$  be an optimal solution to the DAP, and suppose that some advertisers are assigned some types outside of their targeting criteria. We will construct another solution with greater or equal yield in which no incorrect assignments are made.

An optimal solution to the DAP is a vector of functions  $i_{T,a} : \Omega_T \rightarrow [0, 1]$  for  $a \in \mathcal{A}_0$  and  $T \in \mathcal{T}$  such that

$$\begin{aligned}\sum_{T \in \mathcal{T}} \mathbb{E} i_{T,a}(Q) &= \rho_a, \quad \forall a \in \mathcal{A}_0, \\ \sum_{a \in \mathcal{A}_0} i_{T,a}(Q) &= \pi(T), \quad (\text{a.s.}) \forall T \in \mathcal{T},\end{aligned}$$

We refer to each of the functions in a solution as components.

Next, we construct a feasible solution  $i^0$  to the DAP from a feasible flow of the user type-advertiser graph. Take the difference  $\Delta i = i^0 - i$ , which is a circulation in the user type-advertiser graph. The circulation  $\Delta$  may have components of mixed signs. Because  $i^0$  has no incorrect assignments, if advertiser  $a$  is assigned a type  $T \not\preceq a$  not in her criteria, then the circulation verifies that  $\Delta i_{T,a}(Q) = -i_{T,a}(Q)$ . Hence, the components with incorrect assignments are negative.

Let  $a$  be an advertiser that is assigned a type  $T \not\preceq a$  not in her criteria, that is,  $\mathbb{E}[i_{T,a}(Q)] > 0$ . We may find an augmenting cycle  $w$  containing the incorrect assignment, such that if we push some flow along this cycle, we construct another solution  $i + w$  with fewer incorrect assignments. The cycle  $w$  has at most  $A + 1$  positive components, and at most  $A + 1$  negative components. The cost associated to the negative components is at most  $A \frac{1}{A} \min_{a'} \tau_{a'} - \tau_a \leq 0$ . All positive components arcs have a cost of at least zero, and the total cost of this cycle is positive. Thus, the new solution  $i + w$  has greater or equal yield. Moreover,  $\mathbb{E}[(i+w)_{T,a}(Q)] < \mathbb{E}[i_{T,a}(Q)]$ , and no new incorrect assignments are introduced. Repeating this procedure, we may construct a solution with no incorrect assignments. ■

## H Computation

In this section we describe show to compute the optimal policy for our data model. The main problem resides in the computation of the dual objective in (3) and its gradient given a vector of dual variables.



**Objective.** The first term of the objective can be written as

$$\begin{aligned}
\mathbb{E}R\left(\max_{a \in \mathcal{A}_0}\{Q_a - v_a\}\right) &= \sum_{\forall T} \pi(T) \mathbb{E}\left[R\left(\max_{a \in \mathcal{A}_0}\{Q_a - v_a\}\right) \mid T\right] \\
&= \sum_{\forall T} \pi(T) \sum_{a \in T \cup T^c} \mathbb{E}[R(Q_a - v_a) \mathbf{1}\{Q_a - v_a \geq Q_{a'} - v_{a'} \forall a' \neq a\} \mid T] \\
&= \sum_{\forall T} \pi(T) \left( I_{T,0}(v) + \sum_{a \in T} I_{T,a}(v) \right)
\end{aligned}$$

where the first equation follows by conditioning on the type, and the second because the events are a partition of the sample space. Next, we show to compute the expectations  $I_{T,a}(v)$ .

Let  $M_T(v) = \max_{a \in \mathcal{A}_0} \{-\tau_a - v_a\}$  be the maximum discounted quality of the advertisers (including the outside option) that are not in the type, and  $\alpha_T(v)$  the set of advertisers that verify the maximum. Then, we have that

$$\begin{aligned}
I_{T,0}(v) &= R(M_T(v)) \mathbb{P}\{Q_a - v_a \leq M_T(v) \forall a' \in T\} \\
&= R(M_T(v)) G_T(M_T(v) + v_T),
\end{aligned}$$

where  $G_T(\cdot)$  is the c.d.f. of  $Q_T$ , and  $v_T$  is the vector of dual variables for the advertisers in the type.

For  $a \in T$ , we compute the expectation by conditioning on the continuous random variable  $Q_a$ . Further, suppose that we partition the mean vector and covariance matrix in a corresponding manner. That is,  $\mu_T = \begin{pmatrix} \mu_a \\ \mu_{-a} \end{pmatrix}$ , and  $\Sigma_T = \begin{pmatrix} \Sigma_{a,a} & \Sigma_{a,-a} \\ \Sigma_{-a,a} & \Sigma_{-a,-a} \end{pmatrix}$ . For instance,  $\mu_{-a}$  gives the means for the variables in  $T \setminus \{a\}$ , and  $\Sigma_{-a,-a}$  gives variances and covariances for the same variables. The matrix  $\Sigma_{-a,a}$  gives covariances between variables in  $T \setminus \{a\}$  and set  $a$  (as does matrix  $\Sigma_{a,-a}$ ). Because the marginal distribution of a multivariate normal is an univariate normal, we have that  $Q_a \sim \ln \mathcal{N}(\mu_a, \Sigma_{a,a})$ . We denote by  $g_{T,a}(\cdot)$  the p.d.f. of  $Q_a$ . Similarly, let  $Q_{-a}$  be the vector of qualities for advertisers in  $T \setminus \{a\}$ . Conditioning on  $Q_a = q_a$ , the distribution of  $Q_{-a}$  is log-normal with mean vector  $\mu_{-a} - \Sigma_{-a,a}(q_a - \mu_a)/(\Sigma_{a,a})$ , and covariance matrix  $\Sigma_{-a,-a} - (\Sigma_{-a,a}\Sigma_{a,-a})/(\Sigma_{a,a})$ . We denote its c.d.f. by  $G_{T,-a}(\cdot)$ . Putting all together, we have that

$$\begin{aligned}
I_{T,a}(v) &= \mathbb{E}[R(Q_a - v_a) \mathbb{P}\{Q_{a'} - v_{a'} \leq Q_a - v_a \forall a' \neq a \mid Q_a\} \mid T] \\
&= \int_{v_a + M_T(v)}^{\infty} R(q_a - v_a) G_{T,-a}(q_a - v_a + v_{-a}) g_{T,a}(q_a) dq_a,
\end{aligned}$$

where  $v_{-a}$  is the vector of dual variables for advertisers in  $T \setminus \{a\}$ .

**Gradient.** The forward derivative of the dual objective can be written as

$$\begin{aligned}
\nabla_a \psi(v) &= -\mathbb{P}_R \left\{ Q_a - v_a > \max_{a \in \mathcal{A}_0 \setminus a} \{Q_{a'} - v_{a'}\} \right\} + \rho_a \\
&= -\sum_{\forall T} \pi(T) \mathbb{E} \left[ (1 - s^*(Q_a - v_a)) \mathbf{1} \left\{ Q_a - v_a > \max_{a \in \mathcal{A}_0 \setminus a} \{Q_{a'} - v_{a'}\} \right\} \mid T \right] + \rho_a \\
&= -\sum_{T:a \in T} \pi(T) P_{T,a}(v) - \sum_{\substack{T:a \notin T \\ a \in \alpha_T(v), |\alpha_T(v)|=1}} \pi(T) P_{T,a}(v) + \rho_a,
\end{aligned}$$

where the contributing types for the forward derivative are those where  $a$  is in, and those where  $a$  is not in but verifies exclusively the maximum of the types not in ( $M_T(v)$ ). If two or more advertisers verify the maximum  $M_T(v)$ , then increasing  $v_a$  does not have an impact of the type's contribution to the objective. When  $a \notin T$ , the expectation is given by

$$P_{T,a}(v) = (1 - s^*(M_T(v))) G_T(M_T(v) + v_T).$$

Similarly to the objective, when  $a \in T$  we have that

$$P_{T,a}(v) = \int_{v_a + M_T(v)}^{\infty} (1 - s^*(q_a - v_a)) G_{T,-a}(q_a - v_a + v_{-a}) g_{T,a}(q_a) dq_a.$$

The backward derivative is computed in a similar fashion. The only exception is that, when  $a \notin T$ , and  $a$  verifies the maximum  $M_T(v)$ , the advertiser always contributes to the derivative regardless of the number of advertisers that attain the maximum. Hence,

$$\nabla_{-a} \psi(v) = \sum_{T:a \in T} \pi(T) P_{T,a}(v) + \sum_{\substack{T:a \notin T \\ a \in \alpha_T(v)}} \pi(T) P_{T,a}(v) - \rho_a.$$

**Optimization.** We solve the dual problem (3) using a Gradient Descent Method. At each step the objective and its objective are computed as described previously. Notice that, when multiple advertisers verify a tie, the objective is not differentiable. In this case a descent direction is constructed using the forward and backward derivatives (if possible).

**Ties.** For the following, we assume that the instance is not *degenerate*, that is, the variances within the types are positive, and no two advertisers are perfectly correlated. Then, within each type, non-trivial ties can only occur between the advertisers that are not in the type (we refer to the non-trivial ties as those in which multiple advertisers attain the same discounted quality). Moreover, there can be at most one tie within each type, and this happens when the maximum  $M_T(v)$  is verified by many advertisers, that is  $|\alpha_T(v)| > 1$ . With some abuse of notation, the probability of such a tie is given by  $\pi(T) P_{T,\alpha_T(v)}$  and it should be split among the advertisers  $\alpha_T(v)$ . Note that the number of non-trivial ties is  $O(T)$ , and the tie-breaking rule can be computed efficiently by solving a feasible flow problem.

## I Fluid Limit

Exploiting our generative model we can construct a fluid model, and obtain the limiting performance of an arbitrary bid-price policy as the number of impressions grows to infinity. We first describe the fluid equations governing the dynamics of the fluid model, then we construct a solution to such system, and then prove that the stochastic algorithm satisfies the fluid equation in the limit as done in [5]. For simplicity, we focus our analysis on the case with no AdX, and no ties; though, a similar analysis applies to the more general case.

In the following we analyze the performance of the stochastic control policy when implementing some (sub-optimal) bid-prices  $v$ . Let  $\omega$  be a sample path,  $J_n(\omega)$  be the cumulative yield collected up to impression  $n$ , and  $S_{n,a}(\omega)$  the number of impressions assigned to advertiser  $a$  up to time  $n$ . We extended the previous definitions for an arbitrary time, by taking their linear interpolations, so that they are continuous. The previous functions are random elements on  $\mathbb{C}[0, \infty)$ . We shall construct the fluid limit by scaling capacity and time proportionally to infinity, and considering a continuous flow of impressions arriving during an horizon of length 1. More formally, we define  $\bar{S}_a(t) = \lim_{N \rightarrow \infty} N^{-1} S_{tN,a}(\omega)$ , which can be interpreted as the fraction of impressions assigned to advertiser  $a$  by time  $t$ . Similarly, we define  $\bar{J}(t) = \lim_{N \rightarrow \infty} N^{-1} J_{tN}(\omega)$  as the cumulative yield up to time  $t$ . We are interested in computing  $\bar{J}(1)$ , the total limiting yield of the algorithm under bid-prices  $v$ .

When capacity is scaled, each advertiser has a capacity of  $\rho_a$ , and the fluid model should satisfy the following differential equations

$$J'(t) = \sum_{a \in \mathcal{A}(t)} \mathbb{E} \left[ Q_a \mathbf{1} \left\{ a = \arg \max_{a' \in \mathcal{A}(t)} \{ Q_{a'} - v_{a'} \} \right\} \right], \quad (20a)$$

$$S'_a(t) = \mathbb{P} \left\{ a = \arg \max_{a' \in \mathcal{A}(t)} \{ Q_{a'} - v_{a'} \} \right\}, \quad \forall a \in \mathcal{A}_0 \quad (20b)$$

$$\mathcal{A}(t) = \{ a \in \mathcal{A}_0 : \bar{S}_a(t) < \rho_a \}, \quad (20c)$$

with the initial conditions  $S_a(0) = 0$ , and  $J(0) = 0$ . In (20c),  $\mathcal{A}(t)$  is the set of advertisers that are yet to be fulfilled (including advertiser 0, which is fulfilled when the time comes all impressions should be assigned directly to the advertisers), and (20b) determines the rate at which impressions are assigned to each advertiser. When one advertiser is fulfilled and the fraction of impressions  $\bar{S}'_a(t)$  reaches its capacity  $\rho_a$ , it is excluded from  $\mathcal{A}(t)$ , and its rate is driven to zero. Finally, (20a) determines the rate at which yield is generated.

It is not hard to see that the solution to the fluid equations (20) is piecewise linear, and continuous. We construct a solution as follows. Let an *epoch*, denoted by  $t^k$ , be the time in which the contract of any advertiser is fulfilled (including advertiser 0). The horizon  $[0, 1]$  is partitioned in consecutive *pieces*, each culminating with an epoch. The  $k^{\text{th}}$  piece spans the interval  $[t^k, t^{k+1})$ , and has a length of  $\Delta^k = t^{k+1} - t^k$ . Since one advertiser is fulfilled at each epoch, there are at most  $A + 1$  pieces.

Let  $\mathcal{A}^k$  be set of advertisers yet to be satisfied at the beginning of stage  $k$  as given by (20c),  $r_a^k$  be the service rate for advertiser  $a$  during stage  $k$  as given by (20b), and  $y^k$  be the yield rate during stage  $k$  as given by (20a). The length of a stage is determined by the advertiser that is first fulfilled. Once determined, the solution is constructed recursively as follows

$$\begin{aligned} \Delta^k &= \min_{a \in \mathcal{A}^k} \left\{ \frac{\rho_a - S_a(t^k)}{r_a^k} \right\}, \\ t^{k+1} &= \Delta^k + t^k, \\ S_a(t^{k+1}) &= S_a(t^k) + r_a^k \Delta^k, \\ J(t^{k+1}) &= J(t^k) + y^k \Delta^k, \\ \mathcal{A}^{k+1} &= \mathcal{A}^k \setminus a^k, \end{aligned} \tag{21a}$$

where  $a^k$  is the advertiser that verifies the minimum in (21a), and initially  $t^0 = 0$ , and  $\mathcal{A}^0 = \mathcal{A}_0$ . The functions  $S_a$  and  $J$  are obtained as the linear interpolation of the values at the endpoints of the interval. Fortunately, the rates can be easily obtained by evaluating the dual objective and its gradient. Let  $\psi^k(v)$  be the objective in (3) when restricting to the set of advertisers  $\mathcal{A}^k$ . Then, we have that  $r^k = \rho - \nabla \psi^k(v)$ , and  $y^k = \psi^k(v) - v \cdot \nabla \psi^k(v)$ .

We conclude by showing that functions obtained are actually the fluid limit of the stochastic process induced by the algorithm.

**Proposition 10** *The fluid limits  $\bar{S}_a(t)$ , and  $\bar{J}(t)$  are a solution to the fluid equations (20).*

*Proof:* Consider the sequence of random elements  $\bar{S}_{N,a}(t, \omega) = N^{-1} S_{tN,a}(\omega)$ , and  $\bar{J}_N(t, \omega) = N^{-1} J_{tN}(\omega)$ . We would like to show that the previous sequences are tight. By Theorem 8.3 in [3], a sequence of random elements  $\{X_N\}$  in  $\mathbb{C}[0, \infty)$  is tight iff (i)  $\{X_N(0)\}$  is tight, and (ii) for all  $\epsilon > 0$  and  $\eta > 0$ , there is a  $\delta > 0$  and an integer  $N_0$  such that  $\mathbb{P}\{\sup_{t \leq t' \leq t+\delta} |X_N(t') - X_N(t)| \geq \epsilon\} \leq \eta$  for  $N \geq N_0$ .

The first condition is trivially satisfied for both sequences. Disregarding integrality issues, which are not important in the analysis, and using the fact that the processes are non-decreasing, we have that

$$\sup_{t \leq t' \leq t+\delta} |\bar{S}_{N,a}(t') - \bar{S}_{N,a}(t)| \leq \frac{1}{N} (S_{(t+\delta)N,a} - S_{tN,a}) \leq \delta.$$

Thus, by picking  $\delta < \epsilon$  the second condition is satisfied for the number of impressions assigned. For the yield processes, employing Markov's inequality, and the bound  $\mathbb{E}R_n^2 \leq A^2 \max_a \{\mathbb{E}Q_a^2\}$  for the yield in any single period, we obtain

$$\frac{1}{\delta} \mathbb{P}\{J_{(t+\delta)N} - J_{tN} \geq N\epsilon\} \leq \frac{1}{\delta N^2 \epsilon^2} \mathbb{E} \left[ \sum_{n=tN}^{(t+\delta)N} R_n \right]^2 \leq \delta \frac{A^2 \max_a \{\mathbb{E}Q_a^2\}}{\epsilon^2},$$

which can be bounded from above by  $\eta$  by picking a small enough  $\delta$ , and  $N > 1/\delta$ .

To see that the fluid limit satisfies the equation (20b), we write the difference quotient for the number of impressions assigned to an advertiser  $a$

$$\frac{\bar{S}_{N,a}(t + \delta) - \bar{S}_{N,a}(t)}{\delta} = \frac{S_{(t+\delta)N,a} - S_{tN,a}}{N\delta} = \frac{1}{N\delta} \sum_{n=tN}^{(t+\delta)N} I_{n,a},$$

with  $\delta_N = N^{-1/2}$ . The left-hand term converges in probability to  $\bar{S}'_a(t)$ . Because the summands in the right-hand term are binary random variables, by Kolmogorov's criterion, we have that  $\frac{1}{N\delta} \sum I_{n,a} - \frac{1}{N\delta} \sum \mathbb{E}I_{n,a}$  goes to zero. We

conclude with Dominated Convergence Theorem that  $\frac{1}{N^\delta} \sum \mathbb{E}I_{n,a}$  converges to  $\mathbb{P}\left\{a = \arg \max_{a' \in \mathcal{A}(t)} \{Q_{a'} - v_{a'}\}\right\}$  as  $N \rightarrow \infty$ , because each summand  $I_{n,a}$  converges to a Bernoulli random variable with aforementioned success probability. A similar argument holds for equation (20a). ■