

# Risk-Aware Bid Optimization for Online Display Advertisement \*

Rui Fan<sup>†1</sup> and Erick Delage<sup>‡1</sup>

<sup>1</sup>GERAD & Department of Decision Sciences, HEC Montréal, Montréal, QC H3T 2A7, Canada.

## Abstract

This research focuses on the bid optimization problem in the real-time bidding setting for online display advertisements, where an advertiser, or the advertiser’s agent, has access to the features of the website visitor and the type of ad slots, to decide the optimal bid prices given a predetermined total advertisement budget. We propose a risk-aware data-driven bid optimization model that maximizes the expected profit for the advertiser by exploiting historical data to design upfront a bidding policy, mapping the type of advertisement opportunity to a bid price, and accounting for the risk of violating the budget constraint during a given period of time. After employing a Lagrangian relaxation, we derive a parametrized closed-form expression for the optimal bidding strategy. Using a real-world dataset, we demonstrate that our risk-averse method can effectively control the risk of overspending the budget while achieving a competitive level of profit compared with the risk-neutral model and a state-of-the-art data-driven risk-aware bidding approach.

**Keywords:** Risk-aware Bidding policy, Risk Aversion, Display Advertising, Machine Learning, Entropic Risk Measure, Lagrangian Relaxation, Real-Time Bidding

## 1 Introduction

In a real-time-bidding (RTB) online display advertisement setting, the advertiser is given the website traffic of clients and ad slots with many different features, including logs, regions, ad slots format, etc., in order to decide a bidding policy for ads potentially displayed to a wide range of future users. The bidding process in digital advertising [Ostrovsky et al., 2007] is based on the second-price auction [Vickrey, 1961] model, where advertisers or their agents win the auction of a given opportunity if they offer the highest bid price among the competitors and pays the second highest bidder’s bid price. The bidders in the auction can be the advertiser themselves or their agents, which we will refer to as the decision makers for the bidding policy.

These bidding policies typically need to be made with incomplete information about the potential click-through rate (CTR)<sup>1</sup>, winning prices<sup>2</sup>, the net value of each customer to the

---

\*This article was reviewed and accepted in Proceedings of the 31st ACM International Conference on Information and Knowledge Management (CIKM’22) DOI .

<sup>†</sup>Corresponding author, email addresses: *rui.fan@hec.ca*

<sup>‡</sup>Email addresses: *erick.delage@hec.ca*

<sup>1</sup>The click-through rate of an advertisement is the number of times a click is made on the advertisement, divided by the number of times the ad is shown, which is also called the number of impressions [Wikipedia, 2022].

<sup>2</sup>The winning price (market price) is the price that is paid for the ad spot.

company, etc. Instead, they rely on historical data that could be shared by the advertisement platform or accumulated internally by advertisers themselves. The decision makers in this problem need to use this historical data to estimate these critical values and develop the optimal bidding policy for the forthcoming bidding opportunities.

Many companies around the world exploit online marketing, and in 2020, the total spending on digital ads worldwide reached 378.16 billion US\$ [Statista, 2022]. Yet, effective use of online marketing budgets requires robust statistical methods that can identify key interactions among a possibly biased historical dataset and algorithms that efficiently converge to optimal risk-aware bidding policies. Moreover, the currently available methods seldom consider the risk in the bidding activity and, to the best of our knowledge, have not explicitly addressed the risk of incurring excessive expenses.

In reality, the marketing budget is determined ahead and depends on a certain period of time. The budget is given at the beginning of the period when the decision makers have no knowledge about what opportunities will precisely realize. If they do not account for the risk of going over budget, valuable opportunities that appear at a later time will likely be missed. Therefore, being able to identify bidding policies that control the risk of excessive expense during a given period of time is of practical concern.

This paper considers an advertiser who needs to design a bidding policy, mapping the type of advertisement opportunity to a bid price, that will be deployed over a period of time while effectively controlling the risk of violating a specified budget constraint. For this purpose, we propose a static risk-aware bid optimization model that can exploit historical data to prescribe an optimal bidding policy. The approach is hybrid in how it considers a stochastic model that mixes both an empirical distribution to model the type of opportunities, and a parametric distribution to model the winning price and click potential when formulating the problem. This allows us to derive novel closed-form expressions for the optimal risk-neutral and risk-averse bidding policies. The optimal bidding policies are easy to implement and interpret given that they involve simple functions of the estimated conditional CTR and conditional mean and variance of the anticipated winning price. To control the risk of running out of budget, we employ an expected utility model that can account for the advertiser’s risk aversion. Finally, the proposed bidding strategies are implemented and evaluated on a real-world dataset. These experiments<sup>3</sup> show that our bidding policy effectively controls the risk of expense going beyond the budget while outperforming state-of-the-art risk-aware bidding strategies in terms of the average profit that is achieved.

The rest of the paper is divided as follows. Section 2 presents our review of the literature. Section 3 proposes a stochastic model for describing a random bid opportunity offered to an advertiser. Sections 4 and 5 respectively present our so-called risk-neutral and risk-averse bid optimization problems. Finally, sections 6 and 7 respectively describe our experiment design and numerical results, while Section 8 concludes.

## 2 Literature Review

Our literature review covers the three main topics of this paper.

### 2.1 CTR and winning price prediction

The CTR prediction is a binary classification problem, commonly used estimators in computational advertising are linear-based [Richardson et al., 2007], tree-based [Friedman, 2002], and many models have been developed based on the Factorization Machines (FM) [Rendle, 2010, Ta, 2015, Pan et al., 2016, Juan et al., 2017, Pan et al., 2018] which can better fit the feature

---

<sup>3</sup>The code of experiments is available at [https://github.com/ReneeRuiFAN/risk-aware\\_bid\\_optimization](https://github.com/ReneeRuiFAN/risk-aware_bid_optimization)

combination and sparse data often found in the display advertising datasets. In recent years, with the development of Deep Learning (DL) research and recommendation systems, many researchers have applied Deep Neural Network (DNN) based models [Liu et al., 2015, Qu et al., 2018, Cheng et al., 2016]. The idea of combining the DNN and FM is widely accepted in the research and represents the leading performance in real-world usage [Guo et al., 2017, Huang et al., 2019, Pande, 2021].

The winning price prediction is also called the bid landscape problem. [Cui et al., 2011] uses the gradient boosting decision trees to model the winning price. [Wu et al., 2015] proposed the censored regression model to deal with the problem when some historical winning prices are unknown to the advertiser. The recent DL advancement also applies to the bid landscape problem. [Wu et al., 2018b] proposed the Deep Censored Learning model that uses a DL model for CTR prediction to boost the prediction quality on the winning price and considers its distribution into the learning. Deep Landscape Forecasting (DLF) model [Kan et al., 2019] combines DL for probability distribution forecasting and survival analysis for censorship handling based on a recurrent neural network (RNN) to model the conditional winning probability with respect to each bid price.

## 2.2 Real-Time Bidding Strategies

Linear-based strategies are commonly used in developing bidding strategies. [Ostrovsky et al., 2007, Chen et al., 2011] proposed bidding policies linearly-related with the estimation of the value of click, so-called truthful bidding, while [Perlich et al., 2012] constructed bid prices that depend on the predicted CTR. Besides the profit maximization objective, a dual-based bidding framework [Liu et al., 2017] derived from a strict second-price auction assumption is generally applicable to the multiple ads scenario with various objectives and constraints. [Yang et al., 2019] studied the common case where advertisers aim to maximize the number of conversions, and set cost-per-click (CPC) as a constraint.

In [Zhang et al., 2014], the authors introduced a non-linear bidding strategy model (called ORTB) with the estimated CTR as the input of the bidding function that tries to bid on more impressions rather than focus on a small set of high-value impressions. The paper [Fernandez-Tapia et al., 2016] solves the bid problem in cases where impressions are generated by homogeneous Poisson processes and winning prices are independent and identically distributed (i.i.d). [Ren et al., 2018] model CTR learning and winning price estimation as part of bid optimization for campaign profit maximization as a whole and perform a joint optimization.

Researchers also consider bid optimization in a multi-stage setting as a sequential decision process, where Reinforcement Learning (RL) can play an important role. Indeed, the bid optimization problem can be formulated by Markov Decision Process (MDP), where the bid prices are the actions and the realized clicks provide rewards to the RL agent. In this RL model, the left-over budget can be integrated into the state space, such as in [Du et al., 2017, Cai et al., 2017], or used to influence the reward in [Wu et al., 2018a]. Researchers have also formulated the problem using a multi-agent RL framework [Jin et al., 2018, Zhao et al., 2018]. However, all these RL approaches are generally more computationally expensive to solve compared with static models. Moreover, their solutions usually lack interpretability.

## 2.3 Risk-Aware Bidding Strategies

Based on a multi-stage problem setting, many researchers looked into the feedback control problem during the bidding process [Karlsson, 2020], which controls the risk of unstable performance and keeps the optimization process along the stages using a dynamic system. [Grislain et al., 2019] add a penalty to the cost, if the bidding policy falls short of its key performance indicators to improve the robustness of performance under uncertainty. [Karlsson, 2016] introduced a bid

randomization mechanism to help exploration in a partially observed market and control the uncertainty in the auction-based bidding process.

Most closely related to our work, [Haifeng et al., 2017] proposed the risk management on profit (RMP) model that also models the bidding process as a static problem. The model focuses on controlling the risk of the generated profit, while, in sharp contrast with our model, it does not address the risk of expense going over the budget. The authors also assume that profit risk is solely caused by CTR estimation error, which they model using Bayesian logistic regression, and end up over-simplifying the problem by assuming that the winning price is independent of the type of advertisement opportunity.

In the display advertisement field, the utility of bidding is often defined as the profit of clicks. The papers [Chapelle, 2015, Vasile et al., 2017, Haifeng et al., 2017] define the value of click  $v$  as the value of the sum of winning prices divided by total clicks and measure utility as the profit of bidding which is the difference between the total value of clicks and the expense paid. In these papers, the authors use expected utility theory [Bernoulli, 1954, John von Neumann, 1944]. The exponential utility function, which accounts for a constant absolute risk aversion, is probably the most commonly used utility function. It can also be interpreted as employing an entropic risk measure [Rudloff and Wunderlich, 2008], which satisfies the axioms of convex risk measures [Foellmer and Schied, 2010].

### 3 Stochastic Model and Training

For each bidding opportunity with an observable feature vector  $X$  that represents both user and ad information, the bidding optimization problem will account for three dependent random variables: the realized click  $C$  represents if the fact is that the ad gets clicked, the winning price  $W$ , and the net value of the customer to the company  $V$ . Also, we make the following assumption to facilitate the modeling of  $C, W, V$  given  $X$ . These dependency assumptions are commonly used in other literature [Zhang et al., 2014, Haifeng et al., 2017, Cui et al., 2011, Ren et al., 2018].

**Assumption 1.** *The winning price  $W$ , realized  $C$ , and the net value of the customer  $V$  are mutually independent given  $X$ .*

Additionally, our approach will optimize a bidding policy over a batch of  $M$  opportunities, which are assumed i.i.d. This will simulate a one-shot decision situation where a bidding policy needs to be defined in order to run for a given period of time under a fixed budget. Table 1 summarizes the definitions of the variables used in our formulation, which are discussed next.

**Remark 1.** *The idea of considering a batch of  $M$  opportunities is a distinctive feature of our decision models. Previous static approaches (see [Haifeng et al., 2017, Zhang et al., 2014]) usually assume that  $M$  is large enough for the law of large numbers to apply and impose that the expected expense be smaller than the average budget. In [Haifeng et al., 2017], for example, the proposed bidding policy is simply scaled to the same extent that gives the maximum overall profit under the predetermined budget during the hyperparameters tuning. Our models instead account for the fact that  $M$  can in practice be too small for the average expense to have converged to its expected value. We will however exploit a property of the entropic risk measure that enables us to reduce a batch problem of size  $M > 1$  to an equivalent instantaneous  $M = 1$  problem, which greatly simplifies the analysis.*

In the rest of this section, we propose conditional models for CTR, winning price, winning probability, and value of the customer.

### 3.1 Modeling Conditional CTR

We assume that the CTR depends on the opportunity's features  $X$ , and formally denote:  $\theta(X) := \mathbb{P}(C|X)$ . In the context of this work, we will employ a DeepFM model [Guo et al., 2017] to estimate  $\theta(X)$ . We note that this choice is not limiting and that other CTR prediction models could also be employed if one can improve the accuracy.

### 3.2 Modeling Conditional Winning Price Distribution

We assume that conditional on observing  $X$ , the winning price  $W$  follows the normal distribution  $W \sim N(\hat{w}(X), \sigma(X))$ , where the conditional standard deviation  $\sigma(X)$  also depends on the given opportunity's features  $X$ , so that we have the parametrized probability distribution function of the winning price  $W$  modeled as follows:

$$f_{W|X}(w) = \frac{1}{\sigma(X)\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{w-\hat{w}(X)}{\sigma(X)}\right)^2}.$$

Therefore, to model the distribution of the winning price, we need estimators for the conditional mean  $\hat{w}(X)$  and the conditional standard deviation  $\sigma(X)$ .

Given a dataset containing observed  $(X, W)$  pairs, we can obtain an estimator of the expected winning price, conditioned on  $X$ , by running the regression model:

$$\hat{w}(X) := \arg \min_{w \in \mathcal{W}} \mathbb{E}[(W - w(X))^2],$$

where  $\mathcal{W}$  is the set of estimation functions modeled by a certain DNN architecture, and where  $\mathbb{E}$  refers to the expected value under the empirical distribution observed in the dataset.

The same dataset can also be used to train an estimator of  $\sigma(X)$  following a method introduced in [Fan and Yao, 1998]. Conceptually, we consider the residual  $Z$  defined as:  $Z := (W - \hat{w}(X))^2$ , which depends on the estimator  $\hat{w}(X)$ , as well as the observed  $W$ . We can approximate the conditional variance of the winning price  $\sigma^2(X)$  as the expected value of the residual  $Z$ .

$$\hat{z}(X) := \arg \min_{z \in \mathcal{Z}} \mathbb{E}[(Z - z(X))^2],$$

where again  $\mathcal{Z}$  is the set of estimation functions modeled by a certain DNN architecture. Finally, an estimator of the conditional standard deviation of winning price  $\sigma(X)$  can be obtained by:  $\sigma(X) := \sqrt{\max(\hat{z}(X), \epsilon)}$ , for some small  $\epsilon > 0$ , which ensures that the variance estimate is always positive.

In our experiments, similarly to the CTR estimator  $\theta(X)$ , the estimators of mean of winning price  $\hat{w}(X)$  and standard deviation of winning price  $\sigma(X)$  are DeepFM models.

### 3.3 Modeling Conditional Winning Probability

Following the second price auction [Ostrovsky et al., 2007] process, we assume that the advertisers or their agents can win the bid if they offer a bid price that is larger than the winning price. Since the expense only happens when the advertiser wins the bid, to model uncertainty about the expense, we need to model the probability of winning the bid which depends on the bid price and the winning price. For this purpose, we define a function that indicates whether the bid price wins the auction:

$$s(b, W) := 1\{b \geq W\} = \begin{cases} 1 & b \geq W \\ 0 & \text{otherwise} \end{cases}$$

where  $b$  is the bid price and  $W$  is the winning price. Hence the conditional winning probability given  $X$  can be obtained from  $\mathbb{E}[s(b, W)|X]$ , which depends on both the bid price  $b$  and the conditional winning price distribution given  $X$ .

Table 1: Variable definitions

Variables	Description
$V$	net value of customer for opportunity
$\hat{V}(X)$	conditional net value of customer estimator given $X$
$X$	observed features of opportunity
$C$	indicates of click for opportunity
$\theta(X)$	conditional CTR estimator given $X$
$b(X)$	bid price for opportunity given $X$
$W$	winning price for opportunity
$\hat{w}(X)$	conditional expected winning price for opportunity given $X$
$s(b, W)$	indicates of winning the auction given $b$ and $W$
$\sigma(X)$	conditional standard deviation of winning price for opportunity given $X$
$M$	number of opportunities in a batch

**Remark 2.** *It is important to note that both the ORTB approach [Zhang et al., 2014] and the RMP approach [Haifeng et al., 2017] assume that the winning price is independent of the opportunity’s features  $X$ . The ORTB approach obtains the winning probability using a certain parametrized reciprocal function and tunes its parameter to best fit the winning probability curve as a function of the bid price. In the case of the RMP approach, the winning probability is simplistically estimated by assuming that the winning price is independent and using its empirical distribution. In contrast, we assume that the winning price’s distribution depends on the given opportunity  $X$ , which is a more natural and logical assumption as advertisers are willing to bid more for the more valuable opportunities.*

### 3.4 Modeling Conditional Value of Customer

Ideally, if the decision maker identifies a customer who could bring a higher value to the company from the observable features vector  $X$ , they will be willing to bid at a higher price to improve the probability of winning the auction. Therefore, we model the value of a customer  $V$  conditionally on the type of opportunity  $X$ :  $\hat{V}(X) := \mathbb{E}[V|X]$ . However, we note that, in our experiments,  $\hat{V}(X)$  is considered as a known constant (the same value from [Haifeng et al., 2017, Zhang et al., 2014]) for simplicity.

## 4 Risk-neutral Problem

In this section, we first introduce a static bid optimization model that captures a risk-neutral attitude regarding the possibility of exceeding the predefined budget. While our original model considers the performance of the bidding policy over a batch, we show how this problem can be equivalently reduced to its “instantaneous” version (i.e.  $M = 1$ ). We then establish that by employing a classical Lagrangian relaxation, our optimal bidding policy has a closed-form expression involving the conditional statistics of CTR, winning price, and customer value.

### 4.1 Risk-neutral Problem Formulation

Considering a random batch of  $M$  i.i.d. opportunities denoted by  $\{(X_i, W_i, C_i, V_i)\}_{i=1}^M$ , with each  $V_i$ ,  $C_i$ , and  $W_i$  mutually independent given  $X_i$  (as per Assumption 1), the Risk-neutral

Problem (RNP) seeks a bidding policy that maximizes the expected profit generated over this batch while ensuring that the total budget will be satisfied in expectation. This problem takes the following form:

$$\begin{aligned} b^{\text{rnp}}(\cdot) &:= \underset{b: \mathcal{X} \rightarrow \mathbb{R}^+}{\operatorname{argmax}} && \mathbb{E}[\text{Batch profit}] \\ &\text{subject to} && \mathbb{E}[\text{Batch expense}] \leq BM, \end{aligned}$$

where  $b(\cdot)$  is a bid price policy that will be employed over a batch of  $M$ , while  $B$  captures the average budget per opportunity. Furthermore, we have that:

$$\begin{aligned} \text{Batch profit} &= \sum_{i=1}^M V_i C_i s(b(X_i), W_i) - \sum_{i=1}^M W_i s(b(X_i), W_i), \\ \text{Batch expense} &= \sum_{i=1}^M W_i s(b(X_i), W_i). \end{aligned}$$

Based on the linearity of expectation, batch expressions can be simplified. For example, we can simplify the batch profit to expected instantaneous profit per opportunity format by:

$$\begin{aligned} \mathbb{E} \left[ \sum_{i=1}^M V_i C_i s(b(X_i), W_i) \right] &- \mathbb{E} \left[ \sum_{i=1}^M W_i s(b(X_i), W_i) \right] \\ &= M (\mathbb{E}[V C s(b(X), W)] - \mathbb{E}[W s(b(X), W)]), \end{aligned}$$

and similarly for the batch expense formula.

In the rest of this subsection, we exploit Assumption 1 and the normality assumption for  $W$  to obtain closed-form expressions for both the objective and constraint.

#### 4.1.1 Reducing expected instantaneous revenue expression

We start by rewriting the expected instantaneous revenue expression as:

$$\mathbb{E}[V C s(b(X), W)] = \mathbb{E}[\hat{V}(X) \theta(X) \mathbb{E}[s(b(X), W) | X]],$$

where we exploited (see Assumption 1) the fact that  $V$ ,  $C$ , and  $W$  are mutually independent of each other given  $X$ .

The conditional winning probability can be further reduced since we assume  $W$  is normally distributed conditionally on  $X$ :

$$\mathbb{E}[s(b, W) | X] = \int_{W \leq b} f_{W|X}(W) dW = \Phi \left( \frac{b - \hat{w}(X)}{\sigma(X)} \right),$$

where  $f_{W|X}(w)$  is the probability density function of the winning price  $W$  given  $X$ , and  $\Phi(\cdot)$  is the cumulative distribution function of a standard normal distribution.

Therefore, we can calculate the expected revenue given  $X$  using:

$$\mathbb{E}[V C s(b(X), W) | X] = \hat{V}(X) \theta(X) \Phi \left( \frac{b - \hat{w}(X)}{\sigma(X)} \right). \quad (1)$$

### 4.1.2 Reducing expected instantaneous expense expression

In the case of the budget constraint, we have that:

$$\begin{aligned}
\mathbb{E}[Ws(b, W)|X] &= \int_{W \leq b} W f_{W|X}(W) dW \\
&= \int_{w_s \leq \frac{b - \hat{w}}{\sigma(X)}} (\sigma(X)w_s + \hat{w}(X)) \phi(w_s) dw_s \\
&= \sigma(X) \int_{w_s \leq \frac{b - \hat{w}(X)}{\sigma(X)}} w_s \phi(w_s) dw_s + \hat{w}(X) \int_{w_s \leq \frac{b - \hat{w}}{\sigma(X)}} \phi(w_s) dw_s \\
&= g(b, X) := \hat{w}(X) \Phi\left(\frac{b - \hat{w}(X)}{\sigma(X)}\right) - \sigma(X) \phi\left(\frac{b - \hat{w}(X)}{\sigma(X)}\right), \tag{2}
\end{aligned}$$

where  $w_s$  follows the standard normal distribution, and  $\phi(\cdot)$ ,  $\Phi(\cdot)$  are respectively the density function and cumulative distribution function of a standard normal distribution.

## 4.2 Optimal Solution for RNP Model

After we derive the reduced revenue (1) and expense (2) expressions, we can reformulate the model for  $b^{\text{rnp}}$  using the reduced forms:

$$\begin{aligned}
b^{\text{rnp}}(\cdot) &:= \operatorname{argmax}_{b: \mathcal{X} \rightarrow \mathbb{R}^+} \mathbb{E} \left[ \hat{V}(X) \theta(X) \Phi\left(\frac{b(X) - \hat{w}(X)}{\sigma(X)}\right) - g(b(X), X) \right] \\
&\text{subject to } \mathbb{E}[g(b(X), X)] \leq B.
\end{aligned}$$

In an attempt to solve this problem, one can introduce the Lagrangian coefficient  $\lambda \geq 0$  to obtain a relaxation of this risk-neutral profit maximizing problem:

$$\begin{aligned}
\tilde{b}_\lambda^{\text{rnp}}(\cdot) &:= \operatorname{argmax}_{b: \mathcal{X} \rightarrow \mathbb{R}^+} \mathbb{E} \left[ \hat{V}(X) \theta(X) \Phi\left(\frac{b(X) - \hat{w}(X)}{\sigma(X)}\right) - g(b(X), X) \right] \\
&\quad - \lambda \mathbb{E}[g(b(X), X) - B] \\
&= \operatorname{argmax}_{b: \mathcal{X} \rightarrow \mathbb{R}^+} \mathbb{E}[\mathcal{G}_\lambda(b(X), X)],
\end{aligned}$$

where

$$\mathcal{G}_\lambda(b, X) := \hat{V}(X) \theta(X) \Phi\left(\frac{b - \hat{w}(X)}{\sigma(X)}\right) - g(b, X) - \lambda(g(b, X) - B).$$

The optimal bid price  $b^{\text{rnp}}$  can be approximated using  $\tilde{b}_{\lambda^*}^{\text{rnp}}$  with  $\lambda^*$  as the smallest value of  $\lambda \geq 0$  such that  $\mathbb{E}[g(\tilde{b}_{\lambda^*}^{\text{rnp}}(X), X)] \leq B$ . Furthermore,  $\lambda^*$  can be found using the bisection method.

We next provide a closed-form solution for  $\tilde{b}_\lambda^{\text{rnp}}$  in the form of Lemma 1 (see complete proof in Appendix A.1).

**Theorem 1.** *For any  $\lambda \geq 0$ , a maximizer of the Lagrangian relaxation takes the form:*

$$\tilde{b}_\lambda^{\text{rnp}}(X) := \arg \max_{b \in \left\{0, \frac{\hat{V}(X)\theta(X)}{\lambda+1}, \infty\right\}} \mathcal{G}_\lambda(b, X), \quad \forall X \in \mathcal{X}.$$

Therefore, we can that the optimal bid price  $\tilde{b}_{\lambda^*}^{\text{rnp}}$  is proportional to the estimated value of the customer and CTR given  $\lambda^*$ .



## 5 Risk-averse Problem

The distinguishing point of the Risk-averse Problem (RAP) is that we take the risk of going over budget into consideration. We develop the risk-averse bid optimization model  $b^{\text{rap}}$ , which maximizes the return of bidding while controlling the risk of violating the total budget for the batch of opportunities.

### 5.1 The Risk-averse Budget Constraint

We introduce the exponential utility function to model risk aversion in the budget constraint. Namely, we replace the expected expense constraint with:

$$\mathbb{E}[u_\alpha((1/M)\text{Batch expense})] \geq u_\alpha(B),$$

where  $u_\alpha(y) := -\exp(\alpha y)$  is a concave utility function that allows the decision maker to control risk exposure using the parameter  $\alpha$ . Based on [Rudloff and Wunderlich, 2008], this risk-averse budget constraint can be interpreted as imposing an upper bound of  $B$  on the entropic risk of the average expense in the batch:

$$\rho((1/M)\text{Batch expense}) \leq B, \quad (3)$$

where the entropic risk measure  $\rho$  is a well-known convex risk measure. Furthermore, this risk-averse constraint is known to reduce to the risk-neutral one when  $\alpha \rightarrow 0$  and for a batch of  $M$  opportunities, the constraint takes the form:

$$\mathbb{E} \left[ u_\alpha \left( \frac{1}{M} \sum_{i=1}^M W_i s(b(X_i), W_i) \right) \right] \geq u_\alpha(B). \quad (4)$$

One can simplify this constraint based on the fact that the winning price  $W_i$  and features  $X_i$  are i.i.d. variables, and that  $W$  is normally distributed given  $X$ . We refer the reader to Appendix A.2 for detailed proof of the following lemma.

**Theorem 2.** *Constraint (4) is equivalent to  $\mathbb{E}[h(b(X), X)] \geq -1$ , where*

$$\begin{aligned} h(b, X) := & \\ & -e^{\gamma_1(X)} \Phi \left( \frac{b - \hat{w}(X) - \alpha' \sigma(X)^2}{\sigma(X)} \right) - e^{\gamma_2} + e^{\gamma_2} \Phi \left( \frac{b - \hat{w}(X)}{\sigma(X)} \right) \end{aligned} \quad (5)$$

with  $\alpha' := \alpha/M$ ,  $\gamma_1(X) := (1/2)(\alpha')^2 \sigma(X)^2 + \alpha' \hat{w}(X) - \alpha' B$ , and  $\gamma_2 := -\alpha' B$ .

**Remark 3.** *While the RMP approach [Haifeng et al., 2017] controls the risk of low profit by trading off between the mean and standard deviation of profit, which randomness is caused by CTR uncertainty, our method controls the risk of going over budget by measuring the expected utility of expenses, which randomness is caused by both the uncertainty of winning price distribution and the probability of winning the auction.*

### 5.2 Optimal Solution for RAP Model

Following the reductions presented in Sections 5.1, we can reduce the problem to the following risk-averse expected instantaneous profit maximization problem:

$$\begin{aligned} b^{\text{rap}}(\cdot) := \argmax_{b: \mathcal{X} \rightarrow \mathbb{R}^+} & \quad \mathbb{E}[VCs(b(X), W)] - \mathbb{E}[Ws(b(X), W)] \\ \text{subject to} & \quad \mathbb{E}[h(b(X), X)] \geq -1. \end{aligned} \quad (6)$$

In an attempt to solve this problem, one can again introduce the Lagrangian coefficient  $\lambda \geq 0$  to obtain a relaxation of this risk-averse profit maximizing problem:

$$\begin{aligned}\tilde{b}_\lambda^{\text{rap}}(\cdot) &:= \arg\max_{b: \mathcal{X} \rightarrow \mathbb{R}^+} \mathbb{E}[VCs(b(X), W)] - \mathbb{E}[Ws(b(X), W)] \\ &\quad - \lambda(-1 - \mathbb{E}[h(b(X), X)]) \\ &= \arg\max_{b: \mathcal{X} \rightarrow \mathbb{R}^+} \mathbb{E}[\mathcal{G}_\lambda(b(X), X)],\end{aligned}$$

where

$$\mathcal{G}_\lambda(b, X) := \hat{V}(X)\theta(X)\Phi\left(\frac{b - \hat{w}(X)}{\sigma(X)}\right) - g(b, X) + \lambda(1 - h(b, X)), \quad (7)$$

following our definitions of  $g(b, X)$  and  $h(b, X)$  in (2) and (5) respectively.

The optimal bid price  $b^{\text{rap}}$  can be approximated using  $\tilde{b}_{\lambda^*}^{\text{rap}}$  with  $\lambda^*$  as the smallest value of  $\lambda \geq 0$  such that  $\mathbb{E}[h(\tilde{b}_{\lambda^*}^{\text{rap}}(X), X)] \geq -1$ , where  $\lambda^*$  can again be found using the bisection method.

We next provide a closed-form solution for  $\tilde{b}_\lambda^{\text{rap}}$  in the form of Lemma 3 (see complete proof in Appendix A.3).

**Theorem 3.** *For any  $\lambda \geq 0$ , a maximizer of the Lagrangian relaxation takes the form:*

$$\begin{aligned}\forall X \in \mathcal{X}, \tilde{b}_\lambda^{\text{rap}}(X) &:= \\ \arg \max_{b \in \left\{0, -\frac{\mathbf{W}(\lambda \alpha' e^{(\hat{V}(X)\theta(X) + \lambda e^{\gamma_2} - B)\alpha'})}{\alpha'} + \hat{V}(X)\theta(X) + \lambda e^{\gamma_2}, \infty\right\}} &\mathcal{G}_\lambda(b, X),\end{aligned}$$

where  $\mathbf{W}$  is the Lambert  $W$ -function [Corless et al., 1996], i.e. the inverse of  $f(x) := xe^x$ .

We note that given its analytical form, Lemma 3 explicitly characterizes the influence of the customer value, budget, estimated CTR, and risk aversion level on the bid price, thus making the prescribed bid price highly interpretable.

## 6 Experimental Set-up

To test the effectiveness of our bidding policies, we design our experiments using the real-life iPinYou dataset<sup>4</sup>. This dataset includes logs of ad biddings, impressions, clicks, and final conversions, while data are collected from different industries. Researchers in [Zhang et al., 2015] have analyzed the distributions of data from different industries in this dataset and observed high variations between industries compared to within the same industry. Since the bidding models that we designed are based on the assumptions that each opportunity is i.i.d., we focus on one industry data: the Chinese vertical e-commerce industry collected from *Advertiser ID* 1458.

The estimators  $\theta(X), \hat{w}(X), \sigma(X)$  are trained and Lagrangian parameter  $\lambda$  optimized using the **Training** set (3,083,056 observations), while other hyperparameters are tuned using the **Validation** set (307,319 observations). The **Test** set (307,319 observations) is used for out-of-sample (OOS) performance evaluation. During validation and test, the simulated batches ( $M = 10,000$  opportunities) are interrupted whenever the budget is fully utilized to mimic realistic practices. For OOS experiments, the risk aversion parameter  $\alpha$  was fixed to the value that achieved, during validation, the highest Sharpe ratio of batch profit, i.e.

$$\text{Sharpe ratio of profit} = \frac{\mathbb{E}[\text{Batch profit}]}{\sigma(\text{Batch profit})}$$

---

<sup>4</sup><https://contest.ipinyou.com/>

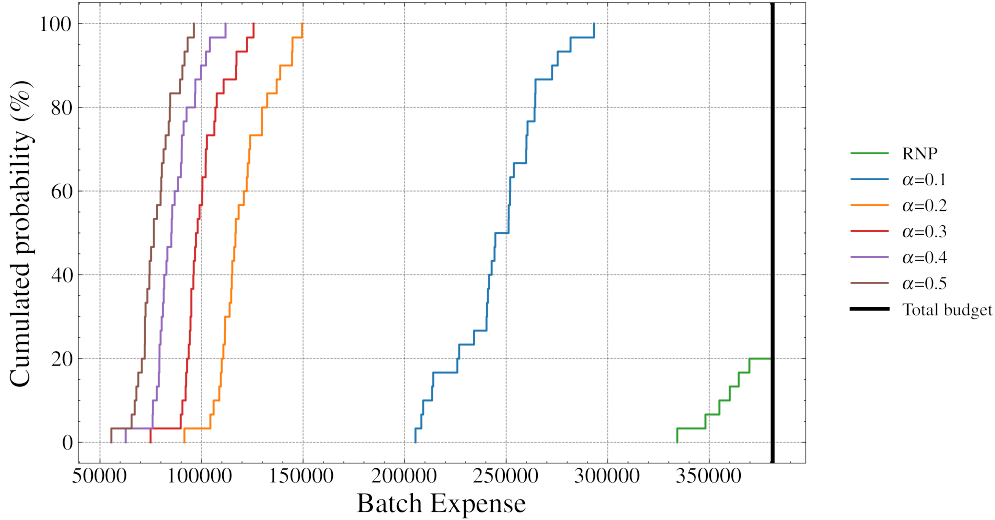


Figure 1: Empirical distribution (in validation) of batch expense for RAP, under different risk aversion levels, and RNP when  $B=1/2 \bar{B}$

while preserving the “early stop frequency” lower than 5%:

$$\text{Early stop frequency} = \mathbb{P} \left( \frac{\text{Batch expense}}{M} \geq B \right).$$

Finally, we investigate the role of the budget parameter  $B$  on performance by considering  $B \in \{\bar{B}, \bar{B}/2, \bar{B}/4, \bar{B}/8, \bar{B}/16, \bar{B}/32, \bar{B}/64\}$  with  $\bar{B}$  as the historical average winning price in the dataset. This is in line with the ranges investigated in [Zhang et al., 2014, Haifeng et al., 2017].

## 7 Numerical Results

In this section, we present our numerical results. We first investigate the effect of the risk aversion parameter on the capacity of the bidding policy to stay within budget. We then more formally compare the OOS performance of our proposed models to the model proposed in [Haifeng et al., 2017].

### 7.1 RAP’s Control of Budget Risk

First, we investigated RAP’s ability to control the budget risk in the validation set, through manipulation of the risk aversion parameter  $\alpha$ , and compared the performance to RNP. Figure 1 presents the empirical cumulative distribution functions (CDF) of the batch expense obtained, for the different problem formulations when  $B = \bar{B}/2$ . Compared with the RNP model, the risk-averse approach RAP demonstrates effective control of budget risk. Indeed, one can remark that all of the RAP models remain within total budget (black vertical line) for 100% of the runs. This is not the case for RNP, which exceeds the budget 80% of the time.

When reducing the marginal budget to  $B = \bar{B}/32$ , the risk aversion parameter starts playing a more important role on budget risk. Indeed, Figure 2 presents the empirical early stop frequency as a function of  $\alpha$ , where  $\alpha$  needs to be greater than 0.28 for the bidding policy to have the guarantee to stay within budget. This confirms that RAP successfully handles the budget risk through the parametrization of  $\alpha$ .

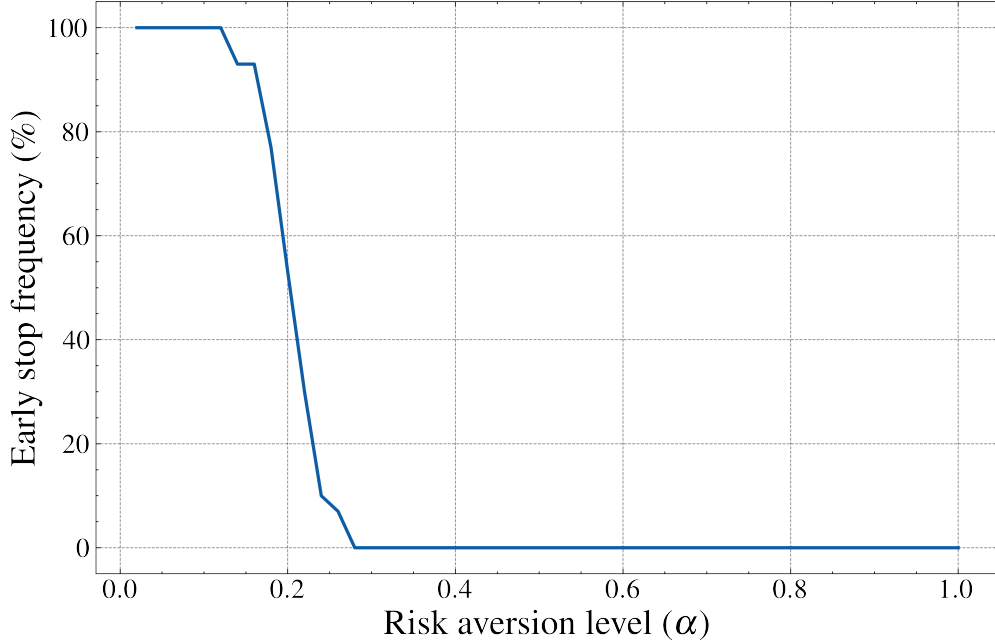


Figure 2: Empirical early stop frequency (in validation) under different risk levels for the profit model with  $B = 1/32\bar{B}$

Table 2: RAP and RNP performance in test set

Metrics	$B/2$		$B/4$		$B/8$		$B/16$		$B/32$		$B/64$	
	RAP	RNP	RAP	RNP	RAP	RNP	RAP	RNP	RAP	RNP	RAP	RNP
Avg. batch clicks	5.600	6.367	5.067	3.300	2.633	1.700	0.967	1.433	0.333	1.267	0.333	1.133
Avg. batch profit	192574	169873	260767	95035	156302	51845	49833	76413	20844	85811	22367	86183
Avg. batch expense	292121	381178	177766	190589	71620	95294	33835	47645	8006	23822	6484	11910
Avg. impression rate	64.5%	69.6%	47.1%	34.8%	28.9%	18.4%	19.4%	11.1%	7.7%	6.9%	6.8%	4.4%
Sharpe ratio of profit	1.083	0.847	1.381	0.595	1.077	0.430	0.497	0.702	0.450	0.802	0.480	0.732
Early stop frequency	0%	100%	13.3%	100%	0%	100%	0%	100%	0%	100%	3.3%	100%

We close this subsection with Table 2, which presents OOS performance metrics, on test data, for RNP and RAP, with risk aversion level selected based on validation data using Sharpe ratio and early stop frequency, under different levels of budget. Based on this table, we first conclude by looking at the early stop frequency that the RAP controls better the risk of violating the budget constraint, at all budget levels, compared with the risk-neutral model. When the budget is relatively large, i.e.  $B \geq \bar{B}/8$ , we also observe that the RAP outperforms the risk-neutral model RNP in Sharpe ratio of profit. When the budget is relatively small  $B \leq \bar{B}/16$ , the RNP model generally outperforms the RAP models in terms of Sharpe ratio of profit and average batch clicks, whereas the RAP model has a better batch impression rate. This is because the bid price tends to be lower (i.e. more conservative) under the RAP. RAP is, therefore, able to invest in more ads thus getting more impressions. On the other hand, it is less competitive for the costly opportunities that end up generating more profit.

## 7.2 Comparison to RMP from [Haifeng et al., 2017]

We now compare the performance of RAP and RNP on the test set to the risk-neutral (called RMP-N) and risk-averse (called RMP-A) bidding policy obtained from the RMP proposed in

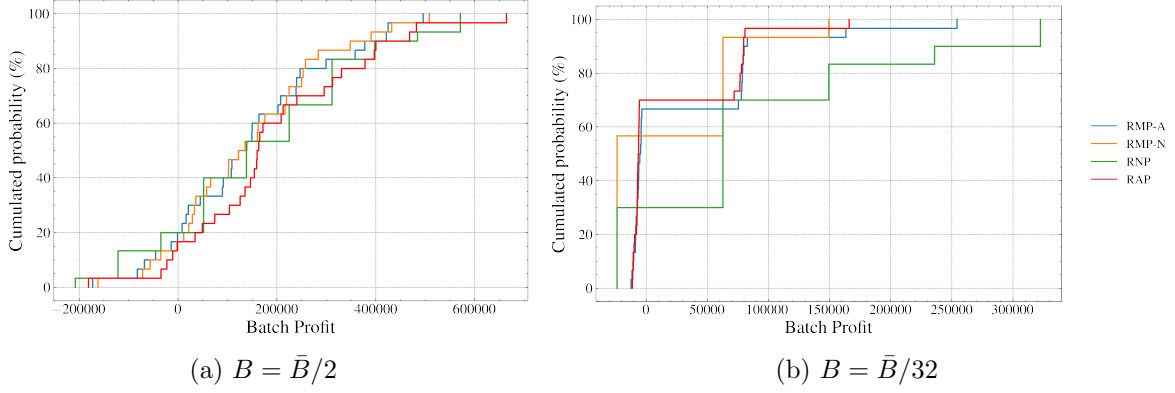


Figure 3: Out-of-sample empirical distribution of batch profit under two different budget levels.

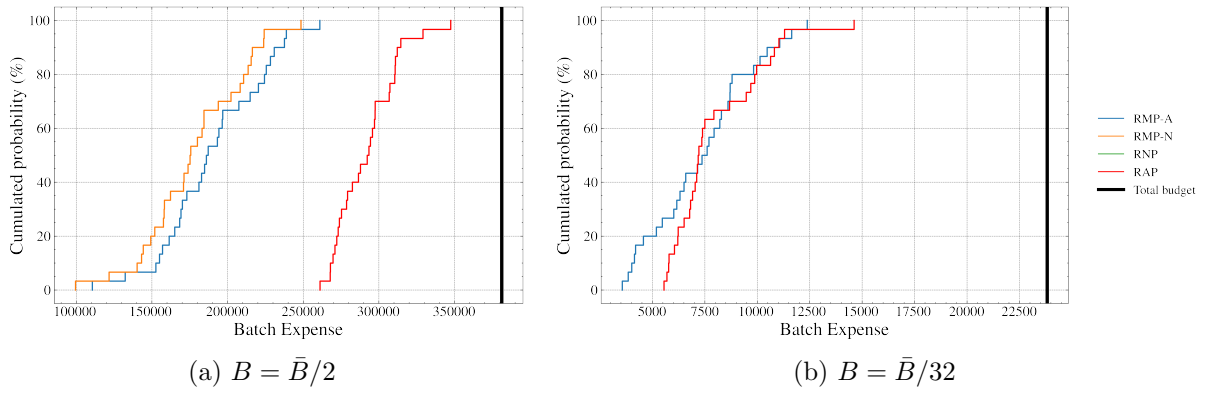


Figure 4: Out-of-sample empirical distribution of batch expense under two different budget levels. Note that curves are not presented when they overlap with the total budget bar.

Table 3: Performance comparison between our RAP and RNP, and the RMP-N and RMP-A from [Haifeng et al., 2017]

Metrics	$\bar{B}/2$				$\bar{B}/32$			
	RAP	RNP	RMP-A	RMP-N	RAP	RNP	RMP-A	RMP-N
Avg. batch clicks	5.600	<b>6.367</b>	3.867	3.700	0.333	<b>1.267</b>	0.433	0.500
Avg. batch profit	<b>192574</b>	169873	144471	142193	20844	<b>85811</b>	30136	19454
Avg. batch expense	292121	381178	190200	<b>178052</b>	8006	23822	<b>7370</b>	23822
Avg. impression rate	64.5%	<b>69.6%</b>	41.5%	38.8%	<b>7.7%</b>	6.9%	5.0%	7.3%
Sharpe ratio of profit	<b>1.083</b>	0.847	0.892	0.922	0.450	<b>0.802</b>	0.487	0.363
Early stop frequency	0	100%	0	0	0	100%	0	100%

[Haifeng et al., 2017]. To offer a fair comparison, we select hyperparameters in RMP using the same procedure as for RAP. Table 3 presents the different metrics for the OOS performance of the different approaches under two different levels of budget (i.e. high and low).

Regarding the control of budget risk, we can observe, based on a 0% early stop frequencies, that both the RAP and RMP-A approaches produce bid policies that control well this risk. This is not the case for RNP and RMP-N, which are by design risk-neutral. In fact, RMP-N appears to be slightly over-conservative at higher budget levels with an early stop frequency of 0%. It is confirmed in Figure 4, which presents the CDFs of batch expense, and where RMP-N’s curve lands a significant margin away from the budget mark for  $B = \bar{B}/2$ .

Next, we observe that RNP always outperforms both RMP-A and RMP-N in terms of average batch profit. This is especially noticeable at the low budget level where the average batch profit is 2.8 times larger than what is achieved by the best competing approach. Figure 3 presents the CDFs of batch profit. In the case of the high budget (see Figure 3(a)), the CDFs appear to mostly overlap with a slight trend to the right for both RNP and RAP, which explains the better average performances. We note that the RNP also appears to have a heavier left tail, which would indicate that it exposes the decision maker to a slightly larger risk of a net loss. In contrast, RAP appears protected from these losses and achieves a larger average batch profit than RNP. We explain the better performance of RAP over RNP here by the fact that when the budget is large, the optimal risk aversion parameter can be reduced in favor of improved expected profit. Moreover, it appears that a small level of risk aversion might actually play the role of making the policies performance more robust out-of-sample, i.e. reducing generalization error. Note that the observations made about average profit can also be made about the Sharpe ratio, identifying our RAP as a clear winner for a higher budget and our RNP as a clear winner for a lower budget.

Finally, in terms of impression rates, we observe that while all methods appear to perform similarly at the low budget level, both RAP and RNP clearly outperform RMP-A/N models at the higher budget level. We interpret this as evidence that our approach is more successful at predicting the price at which an opportunity will be won due to our modeling of the conditional influence of  $X$  on the distribution of  $W$ .

## 8 Conclusion

In this paper, we proposed an approach for optimizing bid policies in a context where one wishes to control the risk of spending more than a given budget in a certain period of time (captured by a total number of opportunities  $M$ ). We proposed both a risk-averse and a risk-neutral problem that let the decision maker control how much protection is needed against overspending. We further derived closed-form expressions for the optimal bid policies when

using a mixed stochastic model, which employed the dataset’s empirical distribution for the type of opportunities  $X$  and trained CTR and conditional Gaussian distribution models for the clicks and winning prices respectively. The closed-form expressions benefit from being interpretable and easy to implement in production. Our two approaches were then compared to analogous risk-neutral and risk-averse bid policies from the RMP approach in [Haifeng et al., 2017] using the iPinyou dataset. The results provide empirical evidence that RAP and RNP provide significant improvement in generating high profit while controlling the risk of going over budget.

In terms of future work, one could consider accounting for risk aversion with respect to profit generation, i.e.  $\mathbb{E}[u(\text{Batch profit})]$ . Alternatively, one could investigate how the choice of alternate convex risk measures, such as conditional value-at-risk, might affect the closed-form solutions and performance. Finally, many other aspects of bid optimization could be handled using similar risk measures: e.g. risk of not reaching a batch revenue, impression rate, or total click target. Finally, one could explore other types of implementations for our bidding policy. A shrinking horizon scheme could be used where the bidding policy continuously accounts for the budget that is left to invest in new ads. Alternatively, a policy optimization approach could be used to optimize the parameters of our policy in an RL environment.

## Acknowledgement

This research was supported financially by the NSERC (RGPIN-2016-05208), the Fin-ML CRE-ATE program, MITACS’ RTA program, and enabled in part by support provided by the Digital Research Alliance of Canada.

## A Proofs

### A.1 Proof of Lemma 1

We start by exploiting the interchangeability property of expected value (see [Shapiro, 2017]), which implies that the optimal bid price for the Lagrangian relaxation can be obtained as the price, for each  $X$ , that maximizes the Lagrangian relaxation function  $\mathcal{G}_\lambda(b, X)$  with

$$\begin{aligned}\mathcal{G}_\lambda(b, X) &:= \hat{V}\theta\Phi\left(\frac{b-\hat{w}}{\sigma}\right) - (\lambda+1)\left[\hat{w}\Phi\left(\frac{b-\hat{w}}{\sigma}\right) - \sigma\phi\left(\frac{b-\hat{w}}{\sigma}\right)\right] + \lambda B \\ &= -(\hat{w} + \lambda\hat{w} - \hat{V}\theta)\Phi\left(\frac{b-\hat{w}}{\sigma}\right) + (1+\lambda)\sigma\phi\left(\frac{b-\hat{w}}{\sigma}\right) + \lambda B,\end{aligned}$$

where we dropped the relation to  $X$  for simplicity of presentation.

Since  $\mathcal{G}_\lambda(b, X)$  is twice differentiable with respect to  $b$ , the maximizer for  $b$  is either 0,  $\infty$  or at a value where the derivative is 0. For the latter case, we get that:

$$\begin{aligned}\frac{d\mathcal{G}_\lambda(b, X)}{db} = 0 &\Leftrightarrow \\ \frac{\hat{w} + \lambda\hat{w} - \hat{V}\theta}{\sigma}\phi\left(\frac{b-\hat{w}}{\sigma}\right) + (1+\lambda)\frac{b-\hat{w}}{\sigma}\phi\left(\frac{b-\hat{w}}{\sigma}\right) &= 0.\end{aligned}$$

Hence, we can conclude that the value of  $b$  where  $\mathcal{G}_\lambda(b, X)$  has a derivative of zero is  $b = \hat{V}(X)\theta(X)/(\lambda+1)$ .  $\square$

## A.2 Proof of Lemma 2

We can simplify constraint (4) based on the fact that the winning price  $W_i$  and features  $X_i$  are i.i.d. variables:

$$\begin{aligned} \mathbb{E} \left[ u_\alpha \left( \frac{1}{M} \sum_{i=1}^M W_i s(b(X_i), W_i) \right) \right] &= -\mathbb{E} \left[ e^{\alpha \left( \frac{1}{M} \sum W_i s(b(X_i), W_i) \right)} \right] \\ &= -\mathbb{E} \left[ \prod_{i=1}^M e^{(\alpha/M) W_i s(b(X_i), W_i)} \right] = -\prod_{i=1}^M \mathbb{E} \left[ e^{(\alpha/M) W_i s(b(X_i), W_i)} \right] \\ &= -\mathbb{E} \left[ e^{\alpha/M W s(b(X), W)} \right]^M = \mathbb{E} \left[ u_{\alpha/M}(W s(b(X), W)) \right]^M, \end{aligned}$$

where the second equality is derived based on the independence assumption and the common distribution.

It means that constraint (4) can be rewritten as:

$$\mathbb{E}[u_{\alpha'}(W s(b(X), W))] \geq u_{\alpha'}(B), \quad (8)$$

where  $\alpha' := \frac{\alpha}{M}$ .

We further simplify the constraint to obtain a closed-form representation. In doing so, we start by dividing both side of constraint (8) by  $-u_{\alpha'}(B) > 0$  in order to normalize this constraint. Note that  $B \geq 0$  implies that  $-u_{\alpha'}(B) > 0$ . This leads us to

$$-\mathbb{E}[u_{\alpha'}(W s(b(X), W))/u_{\alpha'}(B)] \geq -1.$$

We then exploit the Gaussian nature of  $W$  when  $X$  is known to obtain the following reduction:

$$\begin{aligned} &-\mathbb{E}[u_{\alpha'}(W s(b(X), W))/u_{\alpha'}(B)|X] \\ &= -e^{-\alpha' B} \left( \int_{w \leq b(X)} \exp(\alpha' w) f_{W|X}(w) dw \right. \\ &\quad \left. + \int_{w > b(X)} \exp(\alpha' \cdot 0) f_{W|X}(w) dw \right) \\ &= -e^{-\alpha' B} \left( \int_{w \leq b} \frac{1}{\sigma \sqrt{2\pi}} \exp \left( -\frac{(w - \hat{w})^2}{2\sigma^2} + \alpha' w \right) dw + 1 - \Phi \left( \frac{b - \hat{w}}{\sigma} \right) \right) \\ &= -e^{-\alpha' B} \left( \int_{w \leq b} \frac{1}{\sigma \sqrt{2\pi}} \exp \left( -\frac{(w - (\hat{w} + \alpha' \sigma^2))^2}{2\sigma^2} \right. \right. \\ &\quad \left. \left. + \frac{(\alpha')^2 \sigma^2}{2} + \alpha' \hat{w} \right) dw + 1 - \Phi \left( \frac{b - \hat{w}}{\sigma} \right) \right) \\ &= -e^{\gamma_1} \int_{w \leq b} \frac{1}{\sigma \sqrt{2\pi}} \exp \left( -\frac{(w - \hat{w} - \alpha' \sigma^2)^2}{2\sigma^2} \right) dw - e^{\gamma_2} + e^{\gamma_2} \Phi \left( \frac{b - \hat{w}}{\sigma} \right) \\ &= -e^{\gamma_1(X)} \Phi \left( \frac{b(X) - \hat{w}(X) - \alpha' \sigma(X)^2}{\sigma(X)} \right) - e^{\gamma_2} + e^{\gamma_2} \Phi \left( \frac{b(X) - \hat{w}(X)}{\sigma(X)} \right) \end{aligned}$$

where we temporarily drop the relation to  $X$  for simplicity. This completes our proof.  $\square$

## A.3 Proof of Lemma 3

Similar to the risk-neutral models, we also exploit the interchangeability property of expected value, which implies that the optimal bid price for the Lagrangian relaxation can be obtained



as the price, for each  $X$ , that maximizes the Lagrangian relaxation function  $\mathcal{G}_\lambda(b, X)$ :

$$\begin{aligned}\mathcal{G}_\lambda(b, X) &:= (\hat{V}\theta + \lambda e^{\gamma_2})\Phi\left(\frac{b - \hat{w}}{\sigma}\right) - \lambda e^{\gamma_1}\Phi\left(\frac{b - \hat{w} - \alpha'\sigma^2}{\sigma}\right) \\ &\quad + (1 - e^{\gamma_2})\lambda - \left[\hat{w}\Phi\left(\frac{b - \hat{w}}{\sigma}\right) - \sigma\phi\left(\frac{b - \hat{w}}{\sigma}\right)\right] \\ &= c_1\Phi\left(\frac{b - \hat{w}}{\sigma}\right) - c_2\Phi\left(\frac{b - c_3}{\sigma}\right) + (1 - e^{\gamma_2})\lambda + \sigma\phi\left(\frac{b - \hat{w}}{\sigma}\right),\end{aligned}$$

where we again drop the dependence on  $X$  for convenience and where we use  $c_1$ ,  $c_2$ , and  $c_3$  to refer to  $c_1(X) := \hat{V}(X)\theta(X) + \lambda e^{\gamma_2} - \hat{w}(X)$ ,  $c_2(X) := \lambda e^{\gamma_1(X)}$ , and  $c_3(X) := \hat{w}(X) + \alpha'\sigma(X)^2$ .

Since  $\mathcal{G}_\lambda(b, X)$  is differentiable with respect to  $b$ , the maximizer for  $b$  is either 0,  $\infty$  or at a value where the derivative is 0. For the latter case, we get that:

$$\begin{aligned}\frac{d\mathcal{G}_\lambda(b, X)}{db} &= 0 \\ \Leftrightarrow -\frac{c_1}{\sigma}\phi\left(\frac{b - \hat{w}}{\sigma}\right) + \frac{c_2}{\sigma}\phi\left(\frac{b - c_3}{\sigma}\right) + \frac{b - \hat{w}}{\sigma}\phi\left(\frac{b - \hat{w}}{\sigma}\right) &= 0 \\ \Leftrightarrow (c_1 - b + \hat{w})\phi\left(\frac{b - \hat{w}}{\sigma}\right) &= c_2\phi\left(\frac{b - c_3}{\sigma}\right) \\ \Leftrightarrow \ln\left((\hat{V}\theta + \lambda e^{\gamma_2} - b)\phi\left(\frac{b - \hat{w}}{\sigma}\right)\right) &= \ln\left(c_2\phi\left(\frac{b - c_3}{\sigma}\right)\right) \\ \Leftrightarrow \ln\left(\frac{\hat{V}\theta + \lambda e^{\gamma_2} - b}{\sqrt{2\pi}}\right) - \frac{(b - \hat{w})^2}{2\sigma^2} &= \ln\left(\frac{c_2}{\sqrt{2\pi}}\right) - \frac{(b - c_3)^2}{2\sigma^2} \\ \Leftrightarrow 2\sigma^2 \ln(\hat{V}\theta + \lambda e^{\gamma_2} - b) - (b - \hat{w})^2 &= \\ &\quad 2\sigma^2(\ln(\lambda) + \gamma_1) - c_3^2 + 2bc_3 - b^2 \\ \Leftrightarrow 2\sigma^2 \ln(\hat{V}\theta + \lambda e^{\gamma_2} - b) - (b - \hat{w})^2 &= \\ &\quad 2\sigma^2[\ln(\lambda) + \gamma_1] - (\hat{w} + \alpha'\sigma^2)^2 + 2b(\hat{w} + \alpha'\sigma) - b^2 \\ \Leftrightarrow 2\ln(\hat{V}\theta + \lambda e^{\gamma_2} - b) &= 2(\ln(\lambda) + 0.5(\alpha')^2\sigma^2 + \alpha'\hat{w} - \alpha'B) \\ &\quad - (\alpha')^2\sigma^2 - 2w\alpha' + 2b\alpha' \\ \Leftrightarrow \ln\left(\frac{\hat{V}\theta + \lambda e^{\gamma_2} - b}{\lambda}\right) &= \alpha'(b - B) \\ \Leftrightarrow \ln\left(\frac{\hat{V}\theta + \lambda e^{\gamma_2} - b}{\lambda}\right) &= (\hat{V}\theta + \lambda e^{\gamma_2} - B)\alpha' - (\hat{V}\theta + \lambda e^{\gamma_2} - b)\alpha' \\ \Leftrightarrow \frac{\hat{V}\theta + \lambda e^{\gamma_2} - b}{\lambda} e^{(\hat{V}\theta + \lambda e^{\gamma_2} - b)\alpha'} &= e^{(\hat{V}\theta + \lambda e^{\gamma_2} - B)\alpha'} \\ \Leftrightarrow (\hat{V}\theta + \lambda e^{\gamma_2} - b)e^{(\hat{V}\theta + \lambda e^{\gamma_2} - b)\alpha'} &= \lambda e^{(\hat{V}\theta + \lambda e^{\gamma_2} - B)\alpha'} \\ \Leftrightarrow (\hat{V}\theta + \lambda e^{\gamma_2} - b)\alpha' e^{(\hat{V}\theta + \lambda e^{\gamma_2} - b)\alpha'} &= \lambda\alpha' e^{(\hat{V}\theta + \lambda e^{\gamma_2} - B)\alpha'} \\ \Leftrightarrow (\hat{V}\theta + \lambda e^{\gamma_2} - b)\alpha' &= \mathbf{W}\left(\lambda\alpha' e^{(\hat{V}\theta + \lambda e^{\gamma_2} - B)\alpha'}\right) \\ \Leftrightarrow b = -\frac{\mathbf{W}\left(\lambda\alpha' e^{(\hat{V}\theta + \lambda e^{\gamma_2} - B)\alpha'}\right)}{\alpha'} &+ \hat{V}\theta + \lambda e^{\gamma_2},\end{aligned}$$

where the  $\mathbf{W}(\cdot)$  is the Lambert- $\mathbf{W}$  function. This completes our proof.  $\square$

## References

- [Bernoulli, 1954] Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica*, 22(1):23–36.
- [Cai et al., 2017] Cai, H., Ren, K., Zhang, W., Malialis, K., Wang, J., Yu, Y., and Guo, D. (2017). Real-time bidding by reinforcement learning in display advertising. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*. ACM.
- [Chapelle, 2015] Chapelle, O. (2015). Offline evaluation of response prediction in online advertising auctions. In *Proceedings of the 24th International Conference on World Wide Web, WWW '15 Companion*, page 919–922, New York, NY, USA. Association for Computing Machinery.
- [Chen et al., 2011] Chen, Y., Berkhin, P., Anderson, B., and Devanur, N. R. (2011). Real-time bidding algorithms for performance-based display ad allocation. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '11*, page 1307–1315, New York, NY, USA. Association for Computing Machinery.
- [Cheng et al., 2016] Cheng, H.-T., Koc, L., Harmsen, J., Shaked, T., Chandra, T., Aradhye, H., Anderson, G., Corrado, G., Chai, W., Ispir, M., Anil, R., Haque, Z., Hong, L., Jain, V., Liu, X., and Shah, H. (2016). Wide & deep learning for recommender systems. In *Proceedings of the 1st Workshop on Deep Learning for Recommender Systems, DLRS 2016*, page 7–10, New York, NY, USA. Association for Computing Machinery.
- [Corless et al., 1996] Corless, R., Gonnet, G., Hare, D., Jeffrey, D., and Knuth, D. (1996). On the lambert w function. *Advances in Computational Mathematics*, 5:329–359.
- [Cui et al., 2011] Cui, Y., Zhang, R., Li, W., and Mao, J. (2011). Bid landscape forecasting in online ad exchange marketplace. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '11*, page 265–273, New York, NY, USA. Association for Computing Machinery.
- [Du et al., 2017] Du, M., Sassioui, R., Varisteas, G., State, R., Brorsson, M., and Cherkaoui, O. (2017). Improving real-time bidding using a constrained markov decision process.
- [Fan and Yao, 1998] Fan, J. and Yao, Q. (1998). Efficient estimation of conditional variance functions in stochastic regression. *Biometrika*, 85(3):645–660.
- [Fernandez-Tapia et al., 2016] Fernandez-Tapia, J., Guéant, O., and Lasry, J.-M. (2016). Optimal real-time bidding strategies.
- [Foellmer and Schied, 2010] Foellmer, H. and Schied, A. (2010). Convex and coherent risk measures. *Encyclopedia of Quantitative Finance*.
- [Friedman, 2002] Friedman, J. H. (2002). Stochastic gradient boosting. *Comput. Stat. Data Anal.*, 38(4):367–378.
- [Gislain et al., 2019] Gislain, N., Perrin, N., and Thabault, A. (2019). Recurrent neural networks for stochastic control in real-time bidding. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD '19*, page 2801–2809, New York, NY, USA. Association for Computing Machinery.
- [Guo et al., 2017] Guo, H., Tang, R., Ye, Y., Li, Z., and He, X. (2017). Deepfm: A factorization-machine based neural network for ctr prediction.
- [Haifeng et al., 2017] Haifeng, Z., Weinan, Z., Yifei, R., Kan, R., Wenxin, L., and Jun, W. (2017). Managing risk of bidding in display advertising. *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*.
- [Huang et al., 2019] Huang, T., Zhang, Z., and Zhang, J. (2019). Fibinet: Combining feature importance and bilinear feature interaction for click-through rate prediction. In *Proceedings of the 13th ACM Conference on Recommender Systems, RecSys '19*, page 169–177, New York, NY, USA. Association for Computing Machinery.

- [Jin et al., 2018] Jin, J., Song, C., Li, H., Gai, K., Wang, J., and Zhang, W. (2018). Real-time bidding with multi-agent reinforcement learning in display advertising. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, CIKM '18, page 2193–2201, New York, NY, USA. Association for Computing Machinery.
- [John von Neumann, 1944] John von Neumann, O. M. (1944). *Theory of Games and Economic Behavior*. Princeton University Press.
- [Juan et al., 2017] Juan, Y., Lefortier, D., and Chapelle, O. (2017). Field-aware factorization machines in a real-world online advertising system. In *Proceedings of the 26th International Conference on World Wide Web Companion*, WWW '17 Companion, page 680–688, Republic and Canton of Geneva, CHE. International World Wide Web Conferences Steering Committee.
- [Kan et al., 2019] Kan, R., Jiarui, Q., Lei, Z., Weinan, Z., and Yong, Y. (2019). Deep landscape forecasting for real-time bidding advertising. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM.
- [Karlsson, 2016] Karlsson, N. (2016). Control problems in online advertising and benefits of randomized bidding strategies. *European Journal of Control*, 30(C):31–49.
- [Karlsson, 2020] Karlsson, N. (2020). Feedback control in programmatic advertising: The frontier of optimization in real-time bidding. *IEEE Control Systems Magazine*, 40(5):40–77.
- [Liu et al., 2017] Liu, H., Zhu, M., Meng, X., Hu, Y., and Wang, H. (2017). Dual based dsp bidding strategy and its application.
- [Liu et al., 2015] Liu, Q., Yu, F., Wu, S., and Wang, L. (2015). A convolutional click prediction model. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, CIKM '15, page 1743–1746, New York, NY, USA. Association for Computing Machinery.
- [Ostrovsky et al., 2007] Ostrovsky, M., Edelman, B., and Schwarz, M. (2007). Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *American Economic Review*, 97:242–259.
- [Pan et al., 2018] Pan, J., Xu, J., Ruiz, A. L., Zhao, W., Pan, S., Sun, Y., and Lu, Q. (2018). Field-weighted factorization machines for click-through rate prediction in display advertising. In *Proceedings of the 2018 World Wide Web Conference*, WWW '18, page 1349–1357, Republic and Canton of Geneva, CHE. International World Wide Web Conferences Steering Committee.
- [Pan et al., 2016] Pan, Z., Chen, E., Liu, Q., Xu, T., Ma, H., and Lin, H. (2016). Sparse factorization machines for click-through rate prediction. In *2016 IEEE 16th International Conference on Data Mining (ICDM)*, pages 400–409.
- [Pande, 2021] Pande, H. (2021). Field-embedded factorization machines for click-through rate prediction.
- [Perlich et al., 2012] Perlich, C., Dalessandro, B., Hook, R., Stitelman, O., Raeder, T., and Provost, F. (2012). Bid optimizing and inventory scoring in targeted online advertising. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '12, page 804–812, New York, NY, USA. Association for Computing Machinery.
- [Qu et al., 2018] Qu, Y., Fang, B., Zhang, W., Tang, R., Niu, M., Guo, H., Yu, Y., and He, X. (2018). Product-based neural networks for user response prediction over multi-field categorical data. *ACM Trans. Inf. Syst.*, 37(1).
- [Ren et al., 2018] Ren, K., Zhang, W., Chang, K., Rong, Y., Yu, Y., and Wang, J. (2018). Bidding machine: Learning to bid for directly optimizing profits in display advertising. *IEEE Transactions on Knowledge and Data Engineering*, 30(4):645–659.

- [Rendle, 2010] Rendle, S. (2010). Factorization machines. In *2010 IEEE International Conference on Data Mining*, pages 995–1000. IEEE.
- [Richardson et al., 2007] Richardson, M., Dominowska, E., and Ragno, R. (2007). Predicting clicks: Estimating the click-through rate for new ads. In *Proceedings of the 16th International Conference on World Wide Web, WWW '07*, page 521–530, New York, NY, USA. Association for Computing Machinery.
- [Rudloff and Wunderlich, 2008] Rudloff, B. and Wunderlich, R. (2008). Entropic risk constraints for utility.
- [Shapiro, 2017] Shapiro, A. (2017). Interchangeability principle and dynamic equations in risk averse stochastic programming. *Operations Research Letters*, 45(4):377–381.
- [Statista, 2022] Statista (2022). Digital advertising spending worldwide from 2019 to 2024. <https://www.statista.com/statistics/237974/online-advertising-spending-worldwide/>.
- [Ta, 2015] Ta, A.-P. (2015). Factorization machines with follow-the-regularized-leader for ctr prediction in display advertising. In *2015 IEEE International Conference on Big Data (Big Data)*, pages 2889–2891.
- [Vasile et al., 2017] Vasile, F., Lefortier, D., and Chapelle, O. (2017). Cost-sensitive learning for utility optimization in online advertising auctions. In *Proceedings of the ADKDD'17, ADKDD'17*, New York, NY, USA. Association for Computing Machinery.
- [Vickrey, 1961] Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. *The Journal of Finance*, 16(1):8–37.
- [Wikipedia, 2022] Wikipedia (2022). Click-through rate. [https://en.wikipedia.org/wiki/Click-through\\_rate](https://en.wikipedia.org/wiki/Click-through_rate).
- [Wu et al., 2018a] Wu, D., Chen, X., Yang, X., Wang, H., Tan, Q., Zhang, X., Xu, J., and Gai, K. (2018a). Budget constrained bidding by model-free reinforcement learning in display advertising. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management, CIKM '18*, page 1443–1451, New York, NY, USA. Association for Computing Machinery.
- [Wu et al., 2018b] Wu, W., Yeh, M.-Y., and Chen, M.-S. (2018b). Deep censored learning of the winning price in the real time bidding. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD '18*, page 2526–2535, New York, NY, USA. Association for Computing Machinery.
- [Wu et al., 2015] Wu, W. C.-H., Yeh, M.-Y., and Chen, M.-S. (2015). Predicting winning price in real time bidding with censored data. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '15*, page 1305–1314, New York, NY, USA. Association for Computing Machinery.
- [Yang et al., 2019] Yang, X., Li, Y., Wang, H., Wu, D., Tan, Q., Xu, J., and Gai, K. (2019). Bid optimization by multivariable control in display advertising. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*.
- [Zhang et al., 2014] Zhang, W., Yuan, S., and Wang, J. (2014). Optimal real-time bidding for display advertising. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '14*, page 1077–1086, New York, NY, USA. Association for Computing Machinery.
- [Zhang et al., 2015] Zhang, W., Yuan, S., Wang, J., and Shen, X. (2015). Real-time bidding benchmarking with ipinyou dataset.
- [Zhao et al., 2018] Zhao, J., Qiu, G., Guan, Z., Zhao, W., and He, X. (2018). Deep reinforcement learning for sponsored search real-time bidding. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD '18*, page 1021–1030, New York, NY, USA. Association for Computing Machinery.