

Dynamic Budget Throttling in Repeated Second-Price Auctions

Zhaohua Chen^{*1}, Chang Wang^{*2}, Qian Wang^{*1}, Yuqi Pan³, Zhuming Shi⁴, Zheng Cai⁵, Yukun Ren⁵, Zhihua Zhu⁵, Xiaotie Deng^{1,6}

¹ CFCS, School of Computer Science, Peking University

² Northwestern University

³ School of Electronics Engineering and Computer Science, Peking University

⁴ Stony Brook University

⁵ Tencent Technology (Shenzhen) Co., Ltd.

⁶ CMAR, Institute for Artificial Intelligence, Peking University

chenzhaohua@pku.edu.cn, wc@u.northwestern.edu, charlie@pku.edu.cn, pyq0419@stu.pku.edu.cn, zhuming.shi@stonybrook.edu, {zhengcai, rickyren, zhihuazhu}@tencent.com, xiaotie@pku.edu.cn

Abstract

In today's online advertising markets, a crucial requirement for an advertiser is to control her total expenditure within a time horizon under some budget. Among various budget control methods, throttling has emerged as a popular choice, managing an advertiser's total expenditure by selecting only a subset of auctions to participate in. This paper provides a theoretical panorama of a single advertiser's dynamic budget throttling process in repeated second-price auctions. We first establish a lower bound on the regret and an upper bound on the asymptotic competitive ratio for any throttling algorithm, respectively, when the advertiser's values are stochastic and adversarial. Regarding the algorithmic side, we propose the OGD-CB algorithm, which guarantees a near-optimal expected regret with stochastic values. On the other hand, when values are adversarial, we prove that this algorithm also reaches the upper bound on the asymptotic competitive ratio. We further compare throttling with pacing, another widely adopted budget control method, in repeated second-price auctions. In the stochastic case, we demonstrate that pacing is generally superior to throttling for the advertiser, supporting the well-known result that pacing is asymptotically optimal in this scenario. However, in the adversarial case, we give an exciting result indicating that throttling is also an asymptotically optimal dynamic bidding strategy. Our results bridge the gaps in theoretical research of throttling in repeated auctions and comprehensively reveal the ability of this popular budget-smoothing strategy.

Introduction

In recent years, the online advertising market has experienced significant growth, driven by the rise of new social media platforms such as short videos. When a user submits an ad query to the market, an auction is held among all interested advertisers, and the winner is awarded the opportunity to display their ad. Owing to the vast market volume, it is common for advertisers to set a budget to regulate their expenditure over a specified period. Correspondingly, advertising platforms

provide advertisers with various budget control methods to choose from.

This work studies one of these methods, called *throttling* (a.k.a. probabilistic pacing), which is widely adopted by major advertising platforms including Facebook (Facebook 2017), Google (Karande, Mehta, and Srikant 2013), LinkedIn (Agarwal et al. 2014) and Yahoo! (Xu et al. 2015). Under this method, an advertiser's accumulated payment is controlled by being excluded from a fraction of auctions throughout the entire period (e.g., a day, a week, or a month). Compared to other budget management strategies (Balseiro et al. 2021), such as pacing (a.k.a. bid-shading), an essential feature of throttling is that an advertiser's bid is never altered in any auction instance. As revealed in the literature (Karande, Mehta, and Srikant 2013; Chen, Kroer, and Kumar 2021), this feature attracts a large number of advertisers to use the throttling strategy due to two primary reasons. (a) Some advertisers do not permit the platform to modify their bids, forcing the platform to exclude them from some auctions to control their budget, i.e., adopting the throttling strategy. (b) Strategies that modify bids, such as pacing, may allocate an advertiser to those auctions where she is superior to other advertisers, and could be detrimental for the advertiser to exploit other parts of the market. In contrast, with unmodified bids competing, the throttling strategy provides advertisers with a more straightforward and unbiased observation of various users, enabling them to gain a clearer understanding of their competitiveness in different market sectors. For instance, under pacing, a small budget on some market sectors to be explored would lead to a small pacing multiplier on the advertiser's bid, which would cause the advertiser to win nothing in this sector, considering other advertisers who are superior. In comparison, under throttling, even with a small budget, the advertiser has a chance to have a competitive bid (as the bid is not modified) and win in some auctions, thus helping the advertiser to explore the new market sector with a small cost. These phenomena illustrate the importance of throttling as a popular budget control method.

The throttling strategy has been extensively explored in the literature, primarily from an empirical perspective evalua-

^{*}These authors contributed equally.

Copyright © 2024, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

ating the performance of specific algorithms (Agarwal et al. 2014; Karande, Mehta, and Srikant 2013; Xu et al. 2015). Some work has also considered the theoretical equilibrium problem when multiple buyers simultaneously adopt the throttling strategy (Balseiro et al. 2021; Chen, Kroer, and Kumar 2021). However, there is currently no literature that takes the perspective of the advertiser and concentrates on how to *theoretically* optimize their accumulated revenue through a throttling strategy in repeated auctions. The aim of this paper is to address this problem and design dynamic throttling algorithms that achieve good performance in various input models. In practice, such a throttling strategy is implemented through an auto-bidding service that receives the advertiser’s values and makes binary choices on behalf of the advertiser in each auction. However, to simplify the terminology and description, this work shifts the throttling control to the advertiser as a subjective strategy that takes objective values as inputs, which is equivalent to the real-world scenario.

Our Contributions

This work gives a theoretical panorama of an advertiser’s dynamic throttling strategy in repeated second-price auctions. Specifically, our contributions are three-wise along the following lines.

Formalization, bounds, and impossibility results. We formalize the problem of dynamic throttling in repeated second-price auctions from the perspective of a single advertiser in two value input models; namely, where private values v are stochastic and adversarial, respectively. To model other bidders’ bids, we assume the highest competing bids p to be stochastic and following an unknown i.i.d. distribution. When v are stochastic, we measure the performance of a throttling strategy by considering its regret and establish an $\Omega(\sqrt{T})$ lower bound on the expected regret (Theorem 1). On the other hand, when v are adversarial, we measure the performance of a throttling strategy by considering its asymptotic competitive ratio and demonstrate that any throttling strategy’s asymptotic competitive ratio cannot exceed the advertiser’s regularized average budget (Theorem 2), i.e., the average budget divided by the maximum value. Note that the i.i.d. assumption on p is a common practice in the literature. In fact, effective learning could be impossible for the bidder without this assumption, as we show in the full version of this work (Chen et al. 2022b) that any throttling strategy can behave arbitrarily poorly when p are adversarial.

Our OGD-CB algorithm and analysis. Following the above bound results, we combine the online optimization method and the distribution estimation technique and propose an algorithm that is oblivious to the value input model. More technically, it is based on online gradient descent and confidence intervals, referred to as the OGD-CB algorithm (Algorithm 1). Practically, the ability to observe a sample of the highest competing bid affects the algorithm’s estimation accuracy of the distribution of p , and we consider our algorithm’s performance under two information models in order: (a) full information feedback, where the advertiser can observe the highest competing bid in each round; (b) partial information feedback, where the advertiser can only acquire

the highest competing bid when she participates in an auction. The latter information structure presents greater difficulty in distribution estimation but is more practical for auto-bidding services in the industry, especially if the platform announces the highest bid to all participants after an auction (Google Ad Exchange 2022).

Theoretically, we prove that with either full or partial information feedback, our OGD-CB algorithm achieves an $O(\sqrt{T} \log T)$ regret with probability $1 - O(1/T)$ for the stochastic value input model (Theorems 3 and 5). This result implies a near-optimal $O(\sqrt{T} \log T)$ *expected* regret. For the adversarial value input model, we show that it also possesses an optimal asymptotic competitive ratio with either full or partial information feedback (Theorems 4 and 6). We summarize these performance guarantees together with the bound results in Table 1.

In addition to its high performance, our algorithm has two additional advantages. Firstly, it is computation-friendly, as the decision and update process only takes constant time in each round. Second, compared with other algorithms, our solution does not rely on any particular structure of the distribution of the highest competing bid. Notably, our solution does not require the (interim) reward/cost function to be linear, convex/concave, or even continuous. In particular, our solution even works for discrete distributions.

Comparison between throttling and pacing. Subsequently, we compare the throttling strategy with the celebrated pacing strategy (Balseiro and Gur 2019) in repeated second-price auctions. It is worth noting that the latter is known to be asymptotically optimal when v and p are simultaneously stochastic or adversarial. When v and p are stochastic, we show that, in general cases, throttling results in a $\Theta(T)$ loss compared to pacing on the advertiser’s expected revenue (Theorem 7). We also give special conditions under which these two strategies exhibit only an $\tilde{O}(\sqrt{T})$ difference under full/partial information feedback (Theorem 8), for completeness. Excitingly, when v is adversarial and p is stochastic, we demonstrate that throttling is an *asymptotically optimal* bidding strategy under full/partial information feedback. Furthermore, our OGD-CB algorithm is also optimal in this case. This result reveals the importance of throttling as a budget-smoothing method in advertising, and fills the gaps in research on dynamic bidding strategies with adversarial values and stochastic competing bids under repeated second-price auctions.

Related Work

In this part, we will review two popular budget management strategies in repeated auctions: throttling and pacing, further discussion on technically related works.

Previous work on dynamic throttling mainly centers on experimental investigations. Among them, Karande, Mehta, and Srikant (2013) explore the concept of fair allocation in generalized second-price (GSP) auctions, wherein they present an optimal throttling algorithm for diverse objectives. Agarwal et al. (2014) also focus on GSP auctions from an advertiser’s side and implements their algorithm in LinkedIn’s ad serving system. Xu et al. (2015) evaluate a practical online throttling

Value Input Model		Stochastic		Adversarial	
Information Structure		Full	Partial	Full	Partial
Bounds		$\Omega(\sqrt{T})$ regret (Theorem 1)		(ρ/\bar{v}) -asymptotic competitive ratio (Theorem 2)	
	OGD-CB	$O(\sqrt{T \log T})$ regret (Theorems 3 and 5)		(ρ/\bar{v}) -asymptotic competitive ratio (Theorems 4 and 6)	

T : Number of auctions. ρ : Average budget. \bar{v} : Maximum value.

Table 1: The bounds and the performance of OGD-CB under different value input models and information structures.

algorithm on the demand-side platform. Recently, Gui, Nair, and Niu (2022) show how to conduct causal inference of online advertising effects in the budget throttling market. On the theoretical side, some work (Balseiro et al. 2021; Chen, Kroer, and Kumar 2021) focuses on the market equilibrium when all advertisers simultaneously follow a random throttling strategy from either a continuous or discrete view. In contrast, our work examines the dynamics of throttling in the repeated advertising market. Additionally, Charles et al. (2013) study the regret-free allocation for advertisers’ ROI, and shows that such a heuristic outperforms the random throttling strategy for advertisers. Meanwhile, other work looks into a similar problem for Internet keyword search, known as the AdWords problem (Mehta et al. 2007, 2013) in the framework of online matching. However, this line of work concentrates on the platform’s side rather than the advertiser’s side.

Pacing is another well-studied budget control strategy, in which an advertiser shades her value by a constant factor on her bid. Existing work studies this strategy from both dynamic view (Balseiro and Gur 2019; Borgs et al. 2007; Gaitonde et al. 2023; Celli et al. 2022; Lee, Jalali, and Dasdan 2013) and equilibrium perspectives (Balseiro, Besbes, and Weintraub 2015; Conitzer et al. 2022a,b). Among these, the result of Balseiro and Gur (2019) is highly correlated with our solution, which also considers the dual space. Nevertheless, the analysis of their algorithm depends on the continuity of the distribution function, which is not necessary for our algorithm. Some papers compare various budget control methods and explore their relationships from the equilibrium view (Balseiro et al. 2021; Chen et al. 2022a; Balseiro, Kroer, and Kumar 2023). In particular, Balseiro et al. (2021) show that in the symmetric system equilibrium, throttling yields a higher profit for the platform than pacing under certain assumptions. A part of our results extends the comparison between throttling and pacing from an advertiser’s viewpoint from a dynamic view.

Technically, our problem is closely related to the network revenue management (NRM) problem (Gallego and Van Ryzin 1994, 1997) and the contextual bandits with knapsacks (CBwK) problem. However, there are significant differ-

ences between our work and the existing literature. In contrast to NRM, our agent’s reward and cost are random with the highest competing bid, while such randomness is not present in NRM problems. Moreover, our setting diverges from the literature on CBwK in multiple ways. Firstly, early work in CBwK (Agrawal, Devanur, and Li 2016; Badanidiyuru, Langford, and Slivkins 2014) assumes that the context set is finite, whereas we do not make such assumptions. Secondly, some work in CBwK (Agrawal and Devanur 2016; Sivakumar, Zuo, and Banerjee 2022; Han et al. 2023; Slivkins, Sankararaman, and Foster 2023) assumes a specific relationship between the expectation of reward/cost and the context, i.e., requires a specific distribution of the highest competing bid. However, our work sets no such limitation on this distribution. Finally, other work (Wu et al. 2015; Balseiro, Besbes, and Pizarro 2023; Liu and Grigas 2022; Ai et al. 2022) supposes that the context is drawn stochastically i.i.d. from some distribution. In contrast, our solution is oblivious to the context input model, allowing it to deal with adversarial context inputs. This is particularly important for auto-bidding services, as the value for an ad slot can be affected by multiple features and could vary over time. Therefore, it is crucial to design algorithms that can handle different value input models simultaneously. At last, our problem is also correlated with the bandits with knapsacks (BwK) problem (e.g., (Castiglioni, Celli, and Kroer 2022)). However, in that problem, the action in each round is chosen without any context, and the optimal action mode is universal. In contrast, in our problem, the optimal action is relevant to the value observed at the start of each auction.

Model

Basic settings. In this work, we consider the repeated second-price auction market, where an advertiser with a budget constraint competes against other advertisers. The market comprises T rounds of auctions, and in each round $t \in [T] := \{1, 2, \dots, T\}$, an item is to be sold to a buyer via a second-price auction. Here, we suppose that $T \gg 1$. To match the more prevalent terminology in the literature, we refer to the “advertiser” as the “buyer” in the remaining parts.

This work adopts the perspective of a fixed buyer. In each

round $t \in [T]$, she obtains a personal value for the item, represented by $v_t \in [0, \bar{v}]$, where \bar{v} is a constant upper bound on v_t . We denote the highest competing bid against the buyer as p_t , which is assumed to be i.i.d. sampled from an unknown distribution G with a support of $[0, \bar{v}]$. This assumption comes from the mean-field approximation (Balseiro, Besbes, and Weintraub 2015) and is commonly used in the literature. We use \mathbf{v} to represent the buyer's value vector across T rounds, i.e., $\mathbf{v} := (v_t)_{t \in [T]}$; similarly, $\mathbf{p} := (p_t)_{t \in [T]}$.

We assume the buyer has a total budget of B across all T rounds, with the maximum average expenditure being $\rho := B/T$ per round. In this paper, we suppose that $\rho \leq \bar{v}$ is a constant. This assumption comes from the practice in which the buyer is always asked to set a budget for a fixed period, with relatively fixed rounds of auctions. In each round t , the buyer makes a decision $x_t \in \{0, 1\}$ based on the value v_t . The binary selection of the buyer reflects the nature of throttling, where $x_t = 1$ denotes participation in the auction, and $x_t = 0$ means saving the budget and not participating in the auction¹. After the decision, the buyer receives a reward of $x_t r_t$ and incurs a cost of $x_t c_t$ in this round, where r_t and c_t are defined as:

$$r_t = (v_t - p_t)^+, \quad c_t = p_t \mathbf{1}[v_t \geq p_t].$$

In the above, $(v_t - p_t)^+$ in the expression for r_t stands for the positive part of $(v_t - p_t)$. Thus, the buyer can obtain a positive reward and cost in round t only by opting to "enter" (i.e., $x_t = 1$) and wins (i.e., $v_t \geq p_t$). In this case, her cost is p_t , and her reward is $v_t - p_t$ for a second-price auction. Once again, we mention here that in the literature on throttling, it is commonly assumed that the buyer bids truthfully as long as she enters an auction.

Information structure. We consider two different information models in this work:

1. **[Full information feedback.]** The buyer observes p_t at the end of any round t .
2. **[Partial information feedback.]** The buyer observes p_t at the end of round t only if she chooses to participate in the auction in this round, i.e., if $x_t = 1$.

In comparison to the full information feedback model, it is evident that the partial information feedback model is more challenging to manage since the buyer can access less information. Additionally, a natural model with even less information than the two we discuss is the bandit feedback model, in which the buyer only sees r_t and c_t in round t instead of p_t . Recently, some research (Slivkins, Sankararaman, and Foster 2023; Han et al. 2023) has investigated this model in the problem of contextual bandits with knapsacks (CBwK), which is a generalization of our problem. Nevertheless, this line uses online regression techniques and has specific requirements on $\mathbb{E}[r_t, c_t | v_t]$ as a function of v_t , e.g., being linear. In other words, their method has strong assumptions on the distribution G of p_t in our problem. As far as we know, these assumptions are *necessary* in the literature for bandits information feedback. In contrast, in this work, we do not

¹It is required in throttling that the buyer truthfully bids. The motivation here is discussed in the Introduction.

impose any restriction on the distribution G . Our solution is *oblivious* to the distribution G , and thus model-free. In this respect, our information feedback model is comparable to existing works.

We generally use \mathcal{H}_t to denote the history that the buyer can access at the start of round t . For the full information feedback model, we use $\mathcal{H}_t^F := (v_\tau, x_\tau, p_\tau)_{1 \leq \tau < t}$ to denote the buyer's view at the start of round t . We should note that in this model, p_τ is always available to the buyer at the end of each round τ , and r_τ and c_τ can be inferred from v_τ , x_τ and p_τ . For the partial information feedback model, since p_τ is disclosed to the buyer if and only if $x_\tau = 1$, we accordingly define the history available at round t as $\mathcal{H}_t^P := (v_\tau, x_\tau, x_\tau p_\tau)_{1 \leq \tau < t}$. It is worth noting that p_τ cannot be deduced from x_τ and $x_\tau p_\tau$ when $x_\tau = 0$. Likewise, r_τ and c_τ can also be derived using v_τ , x_τ and $x_\tau p_\tau$ in this information model.

The throttling strategy. With the above notation, we now formally define the buyer's throttling strategy. Prior to making a decision, the buyer can see \mathcal{H}_t and v_t in each round t . Denoting the buyer's single-round strategy in round t by $\beta_t : [0, \bar{v}]^{2t-1} \times \{0, 1\}^{t-1} \times \Gamma \rightarrow \{0, 1\}$, we have

$$x_t = \beta_t(\mathcal{H}_t, v_t, \gamma_t).$$

Here, $\gamma_t \in \Gamma$ is sampled from a probability space to depict the potential randomness involved in calculating x_t . Let $\gamma = (\gamma_t)_{t \in [T]}$. As a result, $\beta := (\beta_t)_{t \in [T]}$ encompasses the buyer's single-round strategy for all T rounds and represents her overall throttling strategy.

With \mathbf{v} , \mathbf{p} and randomness γ given, we denote the stopping time of strategy β by $T_0^\beta(\mathbf{v}, \mathbf{p}, \gamma) \leq T$, i.e., the last round with $x_t = 1$. When the context is clear, we abbreviate this as T_0 . Consequently, the total revenue of β with inputs \mathbf{v} and \mathbf{p} is given by:

$$U^\beta(\mathbf{v}, \mathbf{p}, \gamma) := \sum_{t=1}^{T_0} x_t r_t.$$

Value Model and Benchmark

As previously suggested, this work examines two distinct input models for the values $(v_t)_{t \in [T]}$: the stochastic model and the adversarial model.

Stochastic values and regret. Concerning the stochastic value model, it is assumed that for each t , v_t is drawn i.i.d. from some *unknown* distribution F with a support on $[0, \bar{v}]$. We additionally suppose that F and G are independent. In this case, the performance of a throttling strategy is assessed by comparing its reward with the fluid adaptive throttling benchmark. The latter represents the optimal expected revenue of any random strategy given the value without exceeding the budget in expectation. Specifically,

$$\begin{aligned} \text{OPT} &:= T \cdot \max_{\pi: [0, \bar{v}] \rightarrow [0, 1]} \mathbb{E}_{v \sim F, p \sim G} [\pi(v) \cdot (v - p)^+], \\ \text{s.t.} \quad & \mathbb{E}_{v \sim F, p \sim G} [\pi(v) \cdot p \mathbf{1}[v \geq p]] \leq \rho. \end{aligned} \tag{1}$$

It is known that OPT provides an upper bound for the expected total reward of any throttling strategy with stochastic

values (Balseiro, Besbes, and Pizarro 2023). Consequently, we can define the regret of strategy β in relation to OPT given \mathbf{v} , \mathbf{p} and randomness γ . That is,

$$Reg^\beta(\mathbf{v}, \mathbf{p}, \gamma) := OPT - U^\beta(\mathbf{v}, \mathbf{p}, \gamma).$$

Clearly, under stochastic values, our goal is to design a throttling strategy β that results in a low expectation of $Reg^\beta(\mathbf{v}, \mathbf{p}, \gamma)$. A stronger requirement is to ensure a small $Reg^\beta(\mathbf{v}, \mathbf{p}, \gamma)$ with high probability on samples (\mathbf{v}, \mathbf{p}) and randomness γ .

Adversarial values and asymptotic competitive ratio.

We also consider the scenario that the inputs of $(v_t)_{t \in [T]}$ are adversarial on $[0, \bar{v}]$, which accounts for the situations where the buyer lacks confidence in the item distribution. In this regard, we first define the hindsight throttling benchmark, which is the optimal performance advertiser could attain with the benefit of hindsight on \mathbf{v} and \mathbf{p} . Specifically,

$$\begin{aligned} U_H(\mathbf{v}, \mathbf{p}) &:= \max_{\mathbf{x} \in \{0,1\}^T} \sum_{t=1}^T x_t (v_t - p_t)^+, \\ \text{s.t. } & \sum_{t=1}^T x_t p_t \mathbf{1}[v_t \geq p_t] \leq T\rho. \end{aligned} \quad (2)$$

Evidently, $U_H(\mathbf{v}, \mathbf{p})$ bounds $U^\beta(\mathbf{v}, \mathbf{p}, \gamma)$ for any strategy β and randomness γ . We then define a throttling strategy β to be asymptotically μ -competitive for $\mu \in (0, 1]$, if the following condition holds:

$$\liminf_{T \rightarrow \infty, B = \rho T} \inf_{\mathbf{v}, \mathbf{G}} \left(\frac{1}{T} \cdot \mathbb{E}_{\mathbf{p} \sim G^T, \gamma} [U^\beta(\mathbf{v}, \mathbf{p}, \gamma) - \mu \cdot U_H(\mathbf{v}, \mathbf{p})] \right) \geq 0.$$

Bounds on Dynamic Throttling

In this section, we explore the optimal performance attainable by any throttling strategy in the presence of either stochastic or adversarial values. In the case of stochastic values, we obtain a regret lower bound of $\Omega(\sqrt{T})$. On the other hand, with adversarial values, we give an asymptotic competitive ratio upper bound of ρ/\bar{v} . In the full version of this work (Chen et al. 2022b), we additionally derive an impossibility result indicating that any throttling algorithm could achieve arbitrarily small reward facing adversarial \mathbf{p} even when $\{v_t\}_{t \in T}$ are constant. This result underscores the necessity of presuming stationary highest competing bids in the context of throttling.

Stochastic values. The following theorem gives our primary result for the regret lower bound with stochastic values:

Theorem 1. *There exists an instance tuple (F, G) and some constant $C_l > 0$, such that for any online throttling strategy β and $4|T$ (i.e., T is a multiple of 4), we have*

$$\mathbb{E}_{\mathbf{v} \sim F^T, \mathbf{p} \sim G^T, \gamma} [Reg^\beta(\mathbf{v}, \mathbf{p}, \gamma)] \geq C_l \sqrt{T}.$$

Similar results have been documented in Vera and Banerjee (2021); Bumpensanti and Wang (2020); Arlotto and Gurvich (2019) in the literature of network revenue management.

However, in that problem, there is no randomness in the reward and cost given the request (value), which is a simplification of our situation. The central idea of our proof is to give an appropriate problem instance with a *degenerate* fluid solution, where even the optimal throttling strategy with the hindsight of \mathbf{v} and \mathbf{p} cannot avoid a regret of $\Omega(\sqrt{T})$.

Note that the performance of an online throttling strategy may only be inferior with partial information. As a result, Theorem 1 implies that no online throttling strategy can always guarantee an $o(\sqrt{T})$ regret with partial information feedback.

Adversarial values. With adversarial values, we settle an upper bound on the asymptotic competitive ratio of any throttling algorithm. In fact, we have the following theorem stating that such an upper bound is ρ/\bar{v} .

Theorem 2. *For any $\mu > \rho/\bar{v}$,*

$$\liminf_{T \rightarrow \infty, B = \rho T} \inf_{\mathbf{v}, \mathbf{G}} \left(\frac{1}{T} \cdot \mathbb{E}_{\mathbf{p} \sim G^T, \gamma} [U^\beta(\mathbf{v}, \mathbf{p}, \gamma) - \mu \cdot U_H(\mathbf{v}, \mathbf{p})] \right) < 0.$$

Here, the \liminf notation stands for the limit inferior. The proof of Theorem 2 follows Balseiro and Gur (2019). More specifically, we first apply Yao’s principle (Yao 1977) and change our problem into constructing a “hard” problem instance with stochastic \mathbf{p} and half-stochastic \mathbf{v} that follows some vector distribution. Here, the “hard” instance should ensure that the asymptotic competitive ratio of any *deterministic* (rather than random) algorithm is no more than ρ/\bar{v} . We further let \mathbf{p} be fixed across time. Therefore we are further reduced to constructing a vector distribution of $(\mathbf{v} - \mathbf{p})^+$ that blocks any deterministic throttling strategy.

The OGD-CB Algorithm and Performance

In this section, we introduce an online throttling strategy called OGD-CB that is oblivious of the value model (stochastic or adversarial value) and works under either full or partial information feedback. For stochastic values, the regret of the algorithm is upper-bounded by $O(\sqrt{T} \log T)$, which is near optimal based on the lower bound (Theorem 1). For adversarial values, the OGD-CB algorithm is asymptotically (ρ/\bar{v}) -competitive regardless of the information model, matching the upper bound given in Theorem 2.

The algorithm. Our OGD-CB algorithm is presented in Algorithm 1. The algorithm starts with a one-round exploration to make an appropriate initialization (Line 3–Line 5). In each of the following rounds, after observing the value, the algorithm chooses the action based on a dynamically updated pricing parameter λ_t (Line 9), which is updated to control the rate of budget expenditure (Line 10). The update of λ_t follows an online gradient descent (OGD) procedure for a series of proper online reward functions, with step size η_t . Intuitively, a large λ_t indicates that the budget is being spent too quickly, and the algorithm reduces the frequency of entering the market. Conversely, an average expenditure below the ideal ρ in past rounds will result in a descent of λ_t and encourage the algorithm to participate in the auction. This intuition has been inspired by Balseiro, Lu, and Mirrokni

Algorithm 1. The OGD-CB Algorithm.

Input: ρ, T .
Initialization: $\mathcal{I}_1 \leftarrow \emptyset, B_1 \leftarrow B, \lambda_1 \leftarrow 0$.

```

1 for  $t \leftarrow 1$  to  $T$  do
2   Observe  $v_t$ ;
   /* A single round of
      exploration. */
3   if  $t = 1$  then
4      $x_t \leftarrow 1, \lambda_{t+1} \leftarrow \lambda_t$ ;
5   end
6   else
7     /* Estimate the revenue and
        cost with a confidence
        bound. */
8      $\epsilon_t \leftarrow \sqrt{(\log 2 + 2 \log T) / (2|\mathcal{I}_t|)}$ ;
9      $\tilde{r}_t(v_t) \leftarrow (\sum_{\tau \in \mathcal{I}_t} (v_t - p_\tau)^+) / |\mathcal{I}_t| + \epsilon_t v_t$ ,
10     $\tilde{c}_t(v_t) \leftarrow (\sum_{\tau \in \mathcal{I}_t} p_\tau \mathbf{1}[v_t \geq
11      p_\tau]) / |\mathcal{I}_t| - 2\epsilon_t v_t$ ;
12    /* Choose the action
       according to the
       estimates. */
13     $x_t \leftarrow \mathbf{1}[\tilde{r}_t(v_t) \geq \lambda_t \tilde{c}_t(v_t)]$ ;
14    /* Online gradient descent on
       the pricing variable. */
15     $\eta_t \leftarrow 1 / (\bar{v} \sqrt{t})$ ,
16     $\lambda_{t+1} \leftarrow (\lambda_t + \eta_t (x_t \tilde{c}_t(v_t) - \rho))^+$ ;
17  end
18  /* Observe the sample. */
19  if (FULL-INFO)  $\vee$  (PARTIAL-INFO  $\wedge x_t = 1$ )
19  then
20    Observe  $p_t, \mathcal{I}_{t+1} \leftarrow \mathcal{I}_t \cup \{t\}$ ;
21  end
22  else
23     $\mathcal{I}_{t+1} \leftarrow \mathcal{I}_t$ ;
24  end
25  /* Update the remaining budget.
26   */
27   $B_{t+1} \leftarrow B_t - x_t c_t$ ;
28  if  $B_{t+1} < \bar{v}$  then
29    break;
30  end
31 end
32 end

```

(2023). However, a crucial distinction between our setting and that work is that the buyer is unaware of the (expected) revenue and cost given the value. To address this issue, we employ the distribution estimation method and the confidence bound (CB) technique. Specifically, at the start of each round, the algorithm first provides estimates of the expected revenue and cost based on the history (Line 8) and incorporates a bias using the confidence bound, parameterized by ϵ_t , and then makes the decision and updates according to the biased

estimation. As the observation on the highest competing bid p accumulates, the estimation of the reward and the cost becomes more precise, and the bias value reduces to zero.

Stochastic values with full information feedback. We now analyze the performance of the OGD-CB algorithm with stochastic values and full information feedback. We show that in this scenario, OGD-CB achieves an $O(\sqrt{T} \log T)$ regret bound against OPT, as given in the following theorem.

Theorem 3. *With full information feedback, when $\{v_t\}_{t \in T}$ are sampled i.i.d. from some distribution F on $[0, \bar{v}]$, there is a constant C^F , such that with probability at least $1 - 4/T$, the OGD-CB algorithm guarantees*

$$\text{Reg}^\beta(\mathbf{v}, \mathbf{p}) \leq C^F \sqrt{T \log T}.$$

Adversarial values with full information feedback. We further investigate the scenario where the value input could be adversarial. In this case, we establish that the OGD-CB algorithm attains a ρ/\bar{v} asymptotic competitive ratio, which matches the upper bound provided in Theorem 2.

Theorem 4. *With full information feedback, the OGD-CB algorithm guarantees*

$$\liminf_{T \rightarrow \infty, B = \rho T} \inf_{\mathbf{v}, G} \left(\frac{1}{T} \cdot \mathbb{E}_{\mathbf{p} \sim GT} [U^\beta(\mathbf{v}, \mathbf{p}) - \frac{\rho}{\bar{v}} \cdot U_H(\mathbf{v}, \mathbf{p})] \right) \geq 0.$$

Partial information feedback. In the partial information setting, the buyer can only observe p_t if she chooses to enter the round t . As a result, the main obstacle here is that the algorithm may not gather sufficient historical data to provide accurate estimates of $r(v_t)$ and $c(v_t)$. In other words, we need to simultaneously limit the failure probability and the estimation error. We overcome this issue by bounding the entering frequency. More precisely, we use an induction method to show that $|\mathcal{I}_t|$, the entering frequency before round t , increases linearly with t . The result is expressed in the following important lemma.

Lemma 1. *Let $C_e := \min\{(1/2) \cdot (\rho/\bar{v})^2, (\sqrt{2}/4) \cdot (\rho/\bar{v})\}$. In the partial information setting, for any $t \geq 2$, the OGD-CB algorithm guarantees that $|\mathcal{I}_t| \geq C_e \cdot (t - 1)$.*

Utilizing Lemma 1, we can revisit Theorems 3 and 4 and obtain their corresponding versions under partial information. Specifically, we have the following results.

Theorem 5. *With partial information feedback, when $\{v_t\}_{t \in T}$ are sampled i.i.d. from some distribution F on $[0, \bar{v}]$, there is a constant C^P , such that with probability at least $1 - 4/T$, the OGD-CB algorithm guarantees*

$$\text{Reg}^\beta(\mathbf{v}, \mathbf{p}) \leq C^P \sqrt{T \log T}.$$

Theorem 6. *With partial information feedback, the OGD-CB algorithm guarantees*

$$\liminf_{T \rightarrow \infty, B = \rho T} \inf_{\mathbf{v}, G} \left(\frac{1}{T} \cdot \mathbb{E}_{\mathbf{p} \sim GT} [U^\beta(\mathbf{v}, \mathbf{p}) - \frac{\rho}{\bar{v}} \cdot U_H(\mathbf{v}, \mathbf{p})] \right) \geq 0.$$

We notice that in Theorems 3 and 5, the difference between C^P and C^F satisfies that

$$C^P - C^F = O \left(\bar{v} \cdot \sqrt{\max \left\{ \frac{2\bar{v}^2}{\rho^2}, \frac{4\bar{v}}{\sqrt{2}\rho} \right\}} \right).$$

In other words, the difficulty of partial information feedback highly correlates with the ratio \bar{v}/ρ . When this ratio increases, i.e., the buyer's average budget ρ becomes smaller, the partial information feedback becomes more challenging. Intuitively, with smaller ρ , the buyer has fewer chances to participate in auctions and thus learns less information on the distribution of the highest competing bid p with partial feedback, which leads to an increase in regret.

Comparison between Throttling and Pacing

In this section, we aim to compare two popular budget control approaches in the dynamic setting: throttling and pacing. Throttling, which is the focus of this work, involves limiting the frequency of entering the auction, while pacing (Balseiro and Gur 2019), involves shading the buyer's value by an adaptive multiplier and bid the shaded value in each round. The latter approach has been extensively studied in the literature and widely adopted in the industry, as it is known to be the asymptotically optimal bidding strategy when both v and p are stochastic or adversarial. Specifically, we are most interested in comparing the two dynamic strategies. This part extends the result in Balseiro et al. (2021), which compares these two strategies in system equilibrium, i.e., when the dynamic process converges.

We now compare the two strategies under stochastic and adversarial values, and restrict ourselves to full or partial information feedback model. We denote by $U^T(v, p, \gamma)$ the revenue of the optimal throttling strategy² given v , p , and γ . For comparison, we use $U^P(v, p)$ to represent the revenue of the adaptive pacing strategy given in Balseiro and Gur (2019).

Stochastic values. With stochastic values, we begin by giving the following two assumptions.

Assumption 1. G is a continuous distribution on $[0, \bar{v}]$ with density strictly no less than some constant $L > 0$.

Assumption 2. For any $\lambda > 0$, the measure that $r(v)/c(v) \neq \lambda$ is positive concerning distribution F .

Under these two assumptions, we have the following theorem.

Theorem 7. Under Assumptions 1 and 2, when $\mathbb{E}_{v,p}[p \mathbf{1}[v \geq p]] > \rho$, we have

$$\mathbb{E}_{v,p}[U^P(v, p)] - \mathbb{E}_{v,p,\gamma}[U^T(v, p, \gamma)] = \Theta(T).$$

We should emphasize that Assumptions 1 and 2 are not particularly strong. In fact, Assumption 1 holds for most parameterized continuous distribution families, while Assumption 2 holds true unless G has a special form, e.g., being a uniform distribution on $[0, \bar{v}]$. As a result, we can conclude that under stochastic values, dynamic pacing (Balseiro and Gur 2019) would generally outperform even the optimal online throttling strategy by a linear term on the buyer's expected revenue.

²The “optimality” here concerns the expected total revenue. When v is stochastic, the expectation is taken on v , p and the algorithm randomness γ . When v is adversarial, the expectation is only on p and γ .

Nevertheless, for the completeness of this part, we also mention some special cases in which these two strategies have asymptotically similar performances.

Theorem 8. Let β be the OGD-CB algorithm. When (a) the highest competing bid p is fixed, or (b) $\mathbb{E}_{v,p}[p \mathbf{1}[v \geq p]] \leq \rho$, then under full or partial information feedback, we have

$$\begin{aligned} |\mathbb{E}_{v,p,\gamma}[U^T(v, p, \gamma)] - \mathbb{E}_{v,p}[U^P(v, p)]| &= O(\sqrt{T}), \\ |\mathbb{E}_{v,p}[U^\beta(v, p)] - \mathbb{E}_{v,p}[U^P(v, p)]| &= O(\sqrt{T}). \end{aligned}$$

These results conclude our discussions on the comparison of the two dynamic methods under stochastic values.

Adversarial values. Under adversarial values, we first recall the result of Balseiro and Gur (2019) on the performance of dynamic pacing in this scenario.

Proposition 2 (From Balseiro and Gur (2019)). *We have the following:*

$$\liminf_{T \rightarrow \infty, B = \rho T} \inf_{v,p} \left(\frac{1}{T} \cdot \left(U^P(v, p) - \frac{\rho}{\bar{v}} \cdot U_H(v, p) \right) \right) \geq 0.$$

A crucial point in Proposition 2 is that the adaptive pacing algorithm can handle the case when both v and p are adversarial and reaches optimality in this case (Balseiro and Gur 2019). However, by our further results in the full version (Chen et al. 2022b), this is impossible for any throttling strategy.

Nevertheless, when only v is adversarial and p is stochastic, notice that Theorem 2 can be extended to arbitrary bidding strategies (which certainly includes throttling and pacing) without any modification on the proof. Therefore, combining our positive result on the OGD-CB algorithm (Theorems 4 and 6), we conclude that OGD-CB is asymptotically optimal under this scenario.

Concluding Remarks

This work provides a comprehensive discussion on the dynamic throttling strategy in repeated second-price auctions from a buyer's viewpoint. More specifically, we investigate the optimal performance of any throttling algorithm when the buyer's values and the highest competing bids are stochastic or adversarial. On top of that, we propose an OGD-CB algorithm that achieves (near) optimality under both full and partial information structure, regardless of the value input model when the highest competing bid is stochastic. Furthermore, we compare the dynamic throttling strategy to dynamic pacing under different settings. When the values are stochastic, dynamic throttling strategy generally exhibits a linear gap in comparison to dynamic pacing concerning the buyer's revenue. However, with adversarial values, we demonstrate that OGD-CB is asymptotically the best bidding strategies with full or partial information feedback.

Acknowledgements

This work is supported by Tencent Marketing Solution RBFR202104. The authors also thank Santiago R. Balseiro, Chuyue Tang, Xiang Yan, and anonymous reviewers for their kind and useful suggestions.

References

Agarwal, D.; Ghosh, S.; Wei, K.; and You, S. 2014. Budget pacing for targeted online advertisements at linkedin. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 1613–1619.

Agrawal, S.; and Devanur, N. 2016. Linear contextual bandits with knapsacks. *Advances in Neural Information Processing Systems*, 29.

Agrawal, S.; Devanur, N. R.; and Li, L. 2016. An efficient algorithm for contextual bandits with knapsacks, and an extension to concave objectives. In *Conference on Learning Theory*, 4–18. PMLR.

Ai, R.; Chen, Z.; Deng, X.; Pan, Y.; Wang, C.; and Yang, M. 2022. On the Re-Solving Heuristic for (Binary) Contextual Bandits with Knapsacks. *arXiv preprint arXiv:2211.13952*.

Arlotto, A.; and Gurvich, I. 2019. Uniformly bounded regret in the multisecretary problem. *Stochastic Systems*, 9(3): 231–260.

Badanidiyuru, A.; Langford, J.; and Slivkins, A. 2014. Resourceful contextual bandits. In *Conference on Learning Theory*, 1109–1134. PMLR.

Balseiro, S.; Kim, A.; Mahdian, M.; and Mirrokni, V. 2021. Budget-Management Strategies in Repeated Auctions. *Operations Research*, 69(3): 859–876.

Balseiro, S.; Kroer, C.; and Kumar, R. 2023. Contextual standard auctions with budgets: Revenue equivalence and efficiency guarantees. *Management Science*.

Balseiro, S. R.; Besbes, O.; and Pizarro, D. 2023. Survey of Dynamic Resource-Constrained Reward Collection Problems: Unified Model and Analysis. *Operations Research*.

Balseiro, S. R.; Besbes, O.; and Weintraub, G. Y. 2015. Repeated auctions with budgets in ad exchanges: Approximations and design. *Management Science*, 61(4): 864–884.

Balseiro, S. R.; and Gur, Y. 2019. Learning in repeated auctions with budgets: Regret minimization and equilibrium. *Management Science*, 65(9): 3952–3968.

Balseiro, S. R.; Lu, H.; and Mirrokni, V. 2023. The best of many worlds: Dual mirror descent for online allocation problems. *Operations Research*, 71(1): 101–119.

Borgs, C.; Chayes, J.; Immorlica, N.; Jain, K.; Etesami, O.; and Mahdian, M. 2007. Dynamics of bid optimization in online advertisement auctions. In *Proceedings of the 16th international conference on World Wide Web*, 531–540.

Bumpensanti, P.; and Wang, H. 2020. A re-solving heuristic with uniformly bounded loss for network revenue management. *Management Science*, 66(7): 2993–3009.

Castiglioni, M.; Celli, A.; and Kroer, C. 2022. Online learning with knapsacks: the best of both worlds. In *International Conference on Machine Learning*, 2767–2783. PMLR.

Celli, A.; Colini-Baldeschi, R.; Kroer, C.; and Sodomka, E. 2022. The parity ray regularizer for pacing in auction markets. In *Proceedings of the ACM Web Conference 2022*, 162–172.

Charles, D.; Chakrabarty, D.; Chickering, M.; Devanur, N. R.; and Wang, L. 2013. Budget smoothing for internet ad auctions: a game theoretic approach. In *Proceedings of the fourteenth ACM conference on Electronic commerce*, 163–180.

Chen, X.; Kroer, C.; and Kumar, R. 2021. Throttling equilibria in auction markets. *arXiv preprint arXiv:2107.10923*.

Chen, Z.; Deng, X.; Li, J.; Wang, C.; and Yang, M. 2022a. Budget-Constrained Auctions with Unassured Priors. *arXiv preprint arXiv:2203.16816*.

Chen, Z.; Wang, C.; Wang, Q.; Pan, Y.; Shi, Z.; Cai, Z.; Ren, Y.; Zhu, Z.; and Deng, X. 2022b. Dynamic Budget Throttling in Repeated Second-Price Auctions. *arXiv preprint arXiv:2207.04690*.

Conitzer, V.; Kroer, C.; Panigrahi, D.; Schrijvers, O.; Stier-Moses, N. E.; Sodomka, E.; and Wilkens, C. A. 2022a. Packing equilibrium in first price auction markets. *Management Science*, 68(12): 8515–8535.

Conitzer, V.; Kroer, C.; Sodomka, E.; and Stier-Moses, N. E. 2022b. Multiplicative pacing equilibria in auction markets. *Operations Research*, 70(2): 963–989.

Facebook. 2017. Your Guide to Facebook Bid Strategy. https://www.facebook.com/gms_hub/share/biddingstrategyguide_final.pdf?ref=advertising-principles. Accessed: 2022-06-28.

Gaitonde, J.; Li, Y.; Light, B.; Lucier, B.; and Slivkins, A. 2023. Budget Pacing in Repeated Auctions: Regret and Efficiency without Convergence. In *14th Innovations in Theoretical Computer Science Conference (ITCS 2023)*, volume 251, 52. Schloss Dagstuhl–Leibniz-Zentrum für Informatik.

Gallego, G.; and Van Ryzin, G. 1994. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management science*, 40(8): 999–1020.

Gallego, G.; and Van Ryzin, G. 1997. A multiproduct dynamic pricing problem and its applications to network yield management. *Operations research*, 45(1): 24–41.

Google Ad Exchange. 2022. Bid data sharing. <https://support.google.com/authorizedbuyers/answer/2696468>. Accessed: 2022-06-28.

Gui, G.; Nair, H.; and Niu, F. 2022. Auction Throttling and Causal Inference of Online Advertising Effects. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, 297–297.

Han, Y.; Zeng, J.; Wang, Y.; Xiang, Y.; and Zhang, J. 2023. Optimal Contextual Bandits with Knapsacks under Realizability via Regression Oracles. In *International Conference on Artificial Intelligence and Statistics*, 5011–5035. PMLR.

Karande, C.; Mehta, A.; and Srikant, R. 2013. Optimizing budget constrained spend in search advertising. In *Proceedings of the sixth ACM international conference on Web search and data mining*, 697–706.

Lee, K.-C.; Jalali, A.; and Dasdan, A. 2013. Real time bid optimization with smooth budget delivery in online advertising. In *Proceedings of the seventh international workshop on data mining for online advertising*, 1–9.

Liu, H.; and Grigas, P. 2022. Online Contextual Decision-Making with a Smart Predict-then-Optimize Method. *arXiv preprint arXiv:2206.07316*.

Mehta, A.; Saberi, A.; Vazirani, U.; and Vazirani, V. 2007. Adwords and generalized online matching. *Journal of the ACM (JACM)*, 54(5): 22–es.

Mehta, A.; et al. 2013. Online matching and ad allocation. *Foundations and Trends® in Theoretical Computer Science*, 8(4): 265–368.

Sivakumar, V.; Zuo, S.; and Banerjee, A. 2022. Smoothed adversarial linear contextual bandits with knapsacks. In *International Conference on Machine Learning*, 20253–20277. PMLR.

Slivkins, A.; Sankararaman, K. A.; and Foster, D. J. 2023. Contextual Bandits with Packing and Covering Constraints: A Modular Lagrangian Approach via Regression. In *The Thirty Sixth Annual Conference on Learning Theory*, 4633–4656. PMLR.

Vera, A.; and Banerjee, S. 2021. The bayesian prophet: A low-regret framework for online decision making. *Management Science*, 67(3): 1368–1391.

Wu, H.; Srikant, R.; Liu, X.; and Jiang, C. 2015. Algorithms with logarithmic or sublinear regret for constrained contextual bandits. *Advances in Neural Information Processing Systems*, 28.

Xu, J.; Lee, K.-c.; Li, W.; Qi, H.; and Lu, Q. 2015. Smart pacing for effective online ad campaign optimization. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2217–2226.

Yao, A. C.-C. 1977. Probabilistic computations: Toward a unified measure of complexity. In *18th Annual Symposium on Foundations of Computer Science (sfcs 1977)*, 222–227. IEEE Computer Society.