Multi-channel Autobidding with Budget and ROI Constraints

Yuan Deng
Google Research, dengyuan@google.com

Negin Golrezaei
Sloan School of Management, Massachusetts Institute of Technology, golrezae@mit.edu

Patrick Jaillet
Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, jaillet@mit.edu

Jason Cheuk Nam Liang
Operations Research Center, Massachusetts Institute of Technology, jcnliang@mit.edu

Vahab Mirrokni
Google Research, mirrokni@google.com

In digital online advertising, advertisers procure ad impressions simultaneously on multiple platforms, or so-called *channels*, such as Google Ads, Meta Ads Manager, etc., each of which consists of numerous ad auctions. We study how an advertiser maximizes total conversion (e.g. ad clicks) while satisfying aggregate *return-on-investment (ROI)* and budget constraints across all channels. In practice, an advertiser does not have control over, and thus cannot globally optimize, which individual ad auctions she participates in for each channel, and instead authorizes a channel to procure impressions on her behalf: the advertiser can only utilize two levers on each channel, namely setting a per-channel budget and per-channel target ROI. In this work, we first analyze the effectiveness of each of these levers for solving the advertiser's global multi-channel problem. We show that when an advertiser only optimizes over per-channel ROIs, her total conversion can be arbitrarily worse than what she could have obtained in the global problem. Further, we show that the advertiser can achieve the global optimal conversion when she only optimizes over per-channel budgets. In light of this finding, under a bandit feedback setting that mimics real-world scenarios where advertisers have limited information on ad auctions in each channels and how channels procure ads, we present an efficient learning algorithm that produces per-channel budgets whose resulting conversion approximates that of the global optimal problem. Finally, we argue that all our results hold for both single-item and multi-item auctions from which channels procure impressions on advertisers' behalf.

Key words: Online advertising, autobidding, bandit feedback, online optimization

1. Introduction

In today's online advertisers world, advertisers (including but not limited to small businesses, marketing practitioners, non-profits, etc) have been embracing an expanding array of advertising platforms such as search engines, social media platforms, web publisher display etc. which present a plenitude of channels for advertisers to procure ad impressions and obtain traffic. In this growing multi-channel environment, the booming online advertising activities have fueled extensive research and technological advancements in *attribution analytics* to answer questions like which channels are more effective in targeting certain users? Or, which channels produce more user conversion (e.g. ad clicks) or *return-on-investment* (ROI) with the same amount of investments? (see Kannan et al. (2016) for a comprehensive survey on attribution analytics). Yet, this area of research has largely left out a crucial phase in the workflow of advertisers' creation of a digital ad campaign, namely how advertisers interact with advertising channels, which is the physical starting point of a campaign.

To illustrate the significance of advertiser-channel interactions, consider for example a small business who is relatively well-informed through attribution research that Google Ads and Meta ads are the two most effective channels for its products. The business instantiates its ad campaigns through interacting with the platforms' ad management interfaces (see Figure 1), on which the business utilizes levers such as specifying budget and a target ROI[1] to control campaigns. Channels then input these specified parameters into their *autobidding* procedures, where they procure impressions on the advertiser's behalf through automated blackbox algorithms. Eventually, channels report performance metrics such as expenditure and conversion back to the advertiser once the campaign ends. Therefore, one of the most important decisions advertisers need to make involves how to optimize over these levers provided by channels. Unfortunately, this has rarely been addressed in attribution analytics and relevant literature. Hence, this works contributes to filling this vacancy by addressing two themes of practical significance:

How effective are these channel levers for advertisers to achieve their conversion goals? And how should advertisers optimize decisions for such levers?

To answer these questions, we study a setting where an advertiser simultaneously procures ads on multiple channels, each of which consists of multiple ad auctions that sell ad impressions. The advertiser's *global optimization problem* is to maximize total conversion over all channels, while respecting a global budget constraint that limits total spend, and a global ROI constraint that ensures total conversion is at least the target ROI times total spend. However, channels operate as independent entities and conduct autobidding procurement on behalf of advertisers, thereby there are no realistic means for an advertiser to implement the global optimization problem via optimizing over individual auctions. Instead, advertisers can only use two levers, namely a per-channel ROI and per-channel budget, to influence how channels should autobid for impressions. Our goal is to understand how effective are these levers by comparing the total conversion via optimizing levers versus the globally optimal conversion, and also present methodologies to help advertisers optimize over the usage of these levers. We summarize our contributions as followed:

[1] Target ROI is the numerical inverse of CPA or cost per action on Google Ads, and cost per result goal in Meta Ads.

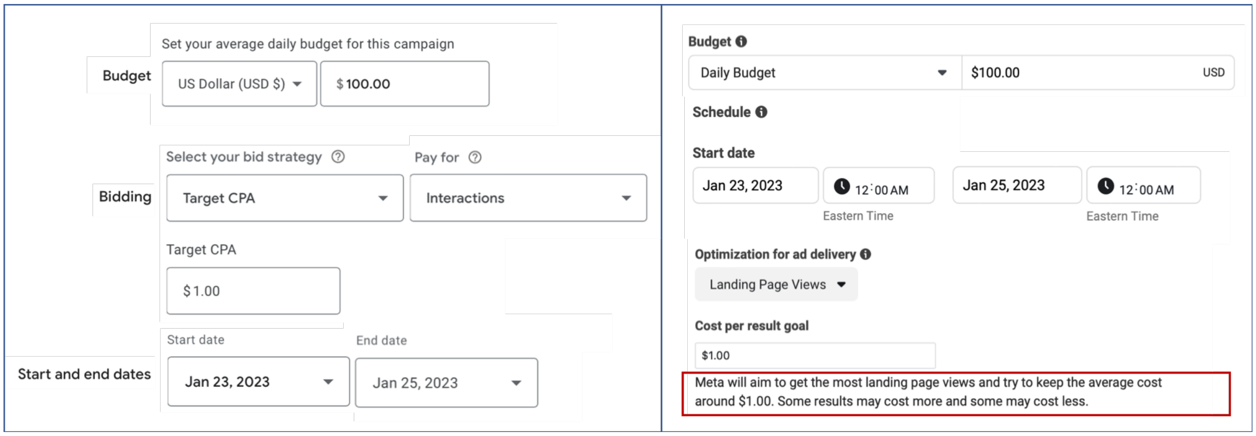

Figure 1 **Interfaces on Google Ads (left) and Meta Ads Manager (right) for creating advertising campaigns that allow advertisers to set budgets, target ROIs, and campaign duration. CPA, or cost per action on Google Ads, as well as cost per result goal on Meta Ads Manager, is effectively the inverse value for an advertiser's per-channel target ROI. Meta Ads Manager specifically highlights that the impression procurement methodology via autobidding maximizes total conversion while respecting advertisers' per-channel target ROI (see red box highlighted), providing evidence that supports the GL-OPT and CH-OPT models in Eq.** (1), (3)**, respectively.**

1.1. Main contributions

1. Modelling ad procurement through per-channel ROI and budget levers. In Section 2 we develop a novel model for online advertisers to optimize over the per-channel ROI and budget levers to maximize total conversion over channels while respecting a global ROI and budget constraint. This multi-channel optimization model closely imitates real-world practices (see Figure 1 for evidence), and to the best of our knowledge is the first of its kind to characterize advertisers' interactions with channels to run ad campaigns.

2. Solely optimizing per-channel budgets are sufficient to maximize conversion. In Theorem 3.2 of Section 3, we show that solely optimizing for per-channel ROIs is inadequate to optimize conversion across all channels, possibly resulting in arbitrary worse total conversions compared to the hypothetical global optimal where advertisers can optimize over individual auctions. In contrast, in Theorem 3.3 and Corollary 3.4 we show that solely optimizing for per-channel budgets allows an advertiser to achieve the global optimal.

3. Algorithm to optimize per-channel budget levers. Under a realistic bandit feedback structure where advertisers can only observe the total conversion and spend in each channel after making a per-channel budget decision, in Section 4 we develop an algorithm that augments stochastic gradient descent (SGD) with the upper-confidence bound (UCB) algorithm, and eventually outputs within T iterations a per-channel budget profile with which advertisers can achieve $\mathcal{O}(T^{-1/3})$ approximation accuracy in total conversion to that of the optimal per-channel budget profile, and a $\mathcal{O}(T^{-1/2})$ violation in both global budget and ROI constraints. Our algorithm relates to constrained convex optimization with uncertain constraints and bandit feedback under a "one point estimation" regime, and to the best of our knowledge, our proposed algorithm is the first to handle such a setting; see more discussions in Section 1.2 and Remark 4.2 of Section 4. Finally, we also present an extended version of our algorithm that achieves the same $\mathcal{O}(T^{-1/3})$ conversion accuracy, while respecting both constraints exactly.

4. Extensions to general advertiser objectives and mutli-impression auctions. In Sections 5 and 6, we shed light on the applicability of our results in Section 3 and 4 to more general settings when auctions correspond to the sale of multiple auctions, or when advertisers aim to optimize a private cost model instead of conversion.

1.2. Related works.

Generally speaking, our work focuses on advertisers' impression procurement process or the interactions between advertisers and impression sellers, which has been addressed in a vast amount of literature in mechanism design and online learning; see e.g. Braverman et al. (2018), Deng et al. (2019), Golrezaei et al. (2019b,a), Balseiro et al. (2019b), Golrezaei et al. (2021a) to name a few. Here, we review literature that relate to key themes of this work, namely autobidding, budget and ROI management, and constrained optimization with bandit feedback.

Autobidding. There has been a rich line of research that model the autobidding setup as well as budget and ROI management strategies. The autobidding model has been formally developed in Aggarwal et al. (2019), and has been analyzed through the lens of welfare efficiency or price of anarchy in Deng et al. (2021), Balseiro et al. (2021a), Deng et al. (2022b), Mehta (2022), as well as individual advertiser fairness in Deng et al. (2022a). The autobidding model has also been compared to classic quasi-linear utility models in Balseiro et al. (2021b). The autobidding model considered in these papers assume advertisers can directly optimize over individual auctions, whereas in this work we address a more realistic setting that mimics practice where advertisers can only use levers provided by channels, and let channels procure ads on their behalf.

Budget and ROI management. Budget and ROI management strategies have been widely studied in the context of mechanism design and online learning. Balseiro et al. (2017) studies the "system equilibria" of a range of budget management strategies in terms of the platforms' profits and advertisers' utility; Balseiro and Gur (2019), Balseiro et al. (2022) study online bidding algorithms (called pacing) that help advertisers achieve high utility in repeated second-price auctions while maintaining a budget constraint, whereas Feng et al. (2022) studies similar algorithms but considers respecting a long term ROI constraint in addition to a fixed budget. All of these works on budget and ROI management focus on bidding strategies in a single repeated auction where advertisers' decisions are bid values submitted directly to the auctions. In contrast, this work focuses on the setting where advertisers procure ads from multiple auctions through channels, and make decisions on how to adjust the per-channel ROI and budget levers while leaving the bidding to channels' blackbox algorithms.

Online optimization. Section 4 where we develop an algorithm to optimize over per-channel target ROI and budgets relates to the area of convex constrained optimization with bandit feedback (also referred to as zero-order or gradient-less feedback) since in light of Lemma 4.3 in Section 4 our problem of interest is also constrained and convex. First, there has been a plenitude of algorithms developed for deterministic constrained convex optimization under a bandit feedback structures where function evaluations for the objective and constraints are non-stochastic. Such algorithms include filter methods Audet and Dennis Jr (2004), Pourmohamad and Lee (2020), barrier-type methods Fasano et al. (2014), Dzahini et al. (2022), as well as Nelder-Mead type algorithms Bűrmen et al. (2006), Audet and Tribes (2018); see Nguyen and Balasubramanian (2022) and references therein for a comprehensive survey. In contrast to these works, our optimization algorithm developed in Section 4 handles noisy bandit feedback.

Regarding works that also address stochastic settings, Flaxman et al. (2004) presents online optimization algorithms under the *known constraint* regime, which assumes the optimizer can evaluate whether all constraints are satisfied, i.e. constraints are analytically available. Further, the algorithm achieves a $\mathcal{O}(T^{-1/4})$ accuracy. In this work, our setting is more complex as the optimizer (i.e. the advertiser) cannot tell whether the ROI constrained is satisfied (due to unknown value and cost distributions in each channels' auctions). Yet our proposed algorithm can still achieve a more superior $\mathcal{O}(T^{-1/3})$ accuracy. Most relevant to this paper is the very recent works Usmanova et al. (2019), Nguyen and Balasubramanian (2022), which considers a similar setting to ours that optimizes for a constrained optimization problem where the objective and constraints are only available through noisy function value evaluations (i.e. unkownn constraints). Usmanova et al. (2019) focuses on a special (unknown) linear constraint setting, and Nguyen and Balasubramanian (2022) extends to general convex constraints. Although Usmanova et al. (2019) and Nguyen and Balasubramanian (2022) achieve $\mathcal{O}(T^{-1})$ and $\mathcal{O}(T^{-1/2})$ approximation accuracy to the optimal solution which contrasts our $\mathcal{O}(T^{-1/3})$ accuracy, these works imposes several assumptions that are stronger than the ones that we consider. First, the objective and constraint functions are strongly smooth (i.e. the gradients are Lipschitz continuous) and Nguyen and Balasubramanian (2022) further assume strong convexity. But in our work, our objectives and constraints are piece-wise linear and do not satisfy such salient properties. Second, and most importantly, both works consider a setting with "two point estimations" that allows the optimizer to access the objective and constraint function values twice in each iteration, enabling more efficient estimations. This work, however, lies in the one-point setting where we can only access function values once per iteration. Finally, we remark that the optimal accuracy/oracle complexity for the one-point setting for constrained (non-smooth) convex optimization with bandit feedback and unknown constraints remains an open question; see Remark 4.2 in Section 4 for more details. We refer readers to Table 4.1 in Larson et al. (2019) for a survey on best known bounds under different one-point bandit feedback settings.

2. Preliminaries

Advertisers' global optimization problem. Consider an advertiser running a digital ad campaign to procure ad impressions on $M \in \mathbb{N}$ platforms such as Google Ads, Meta Ads Manager etc., each of which we call a *channel*. Each channel j consists of $m_j \in \mathbb{N}$ parallel ad auctions, each of which corresponds to the sale of an ad impression.[2] An ad auction $n \in [m_j]$ is associated with a value $v_{j,n} \geq 0$ that represents the expected conversion (e.g. number of clicks) of the impression on sale, and a cost $d_{j,n} \geq 0$ that is required for the purchase of the impression. For example, the cost in a single slot second-price auction is the highest competing bid of competitors in the market, and in a posted price auction the cost is simply the posted price by the seller of the impression. Writing $\boldsymbol{v}_j = (v_{j,n})_{n\in[m_j]}$ and $\boldsymbol{d}_j = (d_{j,n})_{n\in[m_j]}$, we assume that $\boldsymbol{z}_j := (\boldsymbol{v}_j, \boldsymbol{d}_j)$ is sampled from some discrete distribution $\boldsymbol{p}_j$ supported on some finite set $F_j \subseteq \mathbb{R}_+^{m_j} \times \mathbb{R}_+^{m_j}$.

The advertiser's goal is to maximize total conversion of procured ad impressions, while subject to a *return-on-investment (ROI)* constraint that states total conversion across all channels is no less than γ times total spend for some pre-specified target ROI $0 < \gamma < \infty$, as well as a budget constraint that states total spend over all channels is no greater than the total budget $\rho \geq 0$. Mathematically, the advertiser's *global optimization problem* across all M channels can be written as:

$$
\begin{aligned}
\text{GL-OPT} \;=\; \max_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_M} \quad & \sum_{j\in[M]} \mathbb{E}\left[\boldsymbol{v}_j^\top \boldsymbol{x}_j\right] \\
\text{s.t.} \quad & \sum_{j\in[M]} \mathbb{E}\left[\boldsymbol{v}_j^\top \boldsymbol{x}_j\right] \;\geq\; \gamma \sum_{j\in[M]} \mathbb{E}\left[\boldsymbol{d}_j^\top \boldsymbol{x}_j\right] \\
& \sum_{j\in[M]} \mathbb{E}\left[\boldsymbol{d}_j^\top \boldsymbol{x}_j\right] \;\leq\; \rho \\
& \boldsymbol{x}_j \in [0,1]^{m_j} \quad j \in [M]\,.
\end{aligned} \tag{1}
$$

Here, the decision variable $\boldsymbol{x}_j \in [0,1]^{m_j}$ is a vector where $x_{j,n}$ denotes whether impression in auction n for channel j is procured. We remark that $\boldsymbol{x}$ depends on the realization of $\boldsymbol{z} = (\boldsymbol{v}_j, \boldsymbol{d}_j)_{j\in[M]}$ and is also random. We note that the ROI and budget constraints are taken in expectation because an advertiser procures impressions from a very large number of auctions (since the number of auctions in each platform is typically very large) and thus the advertiser only demands to satisfy constraints in an average sense. We note that GL-OPT is a widely adopted formulation for autobidding practices

[2] Ad auctions for each channel may be run by the channel itself or other external ad inventory suppliers such as web publishers.

in modern online advertising, which represents advertisers' conversion maximizing behavior while respecting certain financial targets for ROIs and budgets; see e.g. Aggarwal et al. (2019), Balseiro et al. (2021a), Deng et al. (2021, 2022b). In Section 6.1 we discuss more general advertiser objectives, e.g. maximizing quasi-linear utility.

Our overarching goal of this work is to develop methodologies that enable an advertiser to achieve total campaign conversion that match GL-OPT while respecting her global ROI γ and budget ρ. However, directly optimizing GL-OPT may not be plausible as we discuss in the following.

Advertisers' levers to solve their global problems. To solve the global optimization problem GL-OPT, ideally advertisers would like to optimize over individual auctions across all channels. However, in reality channels operate as independent entities, and typically do not provide means for general advertisers to participate in specific individual auctions at their discretion. Instead, channels provide advertisers with specific *levers* to express their ad campaign goals on spend and conversion. In this work, we focus on two of the most widely used levers, namely the per-channel ROI target and per-channel budget (see illustration in Fig. 1). After an advertiser inputs these parameters to a channel, the channel then procures on behalf of the advertiser through autonomous programs (we call this programmatic process *autobidding*) to help advertiser achieve procurement results that match with the inputs. We will elaborate on this process later.

Formally, we consider the setting where for each channel $j \in [M]$, an advertiser is allowed to input a per-channel target ROI $0 \leq \gamma_j < \infty$, and a per-channel budget $\rho_j \in [0, \rho]$ where we recall $\rho > 0$ is the total advertiser budget for a certain campaign. Then, the channel uses these inputs in its autobidding algorithm to procure ads, and returns the total conversion $V_j(\gamma_j, \rho_j; \boldsymbol{z}_j) \geq 0$, as well as total spend $D_j(\gamma_j, \rho_j; \boldsymbol{z}_j) \geq 0$ to the advertiser, where we recall $\boldsymbol{z}_j = (\boldsymbol{v}_j, \boldsymbol{d}_j) \in \mathbb{R}^{m_j} \times \mathbb{R}^{m_j}$ is the vector of value-cost pairs in channel j sampled from discrete support F_j according to distribution $\boldsymbol{p}_j$; V_j and D_j will be further specified later.

As the advertiser has the freedom of choice to input either per-channel target ROI's, budgets, or both, we consider three options for the advertiser: 1. input only a per-channel target ROI for each channel; 2. input only a per-channel budget for each channel; 3. input both per-channel target ROI and budgets for each channel. Such options correspond to the following decision sets for $(\gamma_j, \rho_j)_{j \in [M]}$:

Per-channel budget only option: $\mathcal{I}_B = \{(\gamma_j, \rho_j)_{j \in [M]} \in \mathbb{R}_+^{2 \times M} : \gamma_j = 0, \rho_j \in [0, \rho] \text{ for } \forall j\}$.

Per-channel target ROI only option: $\mathcal{I}_R = \{(\gamma_j, \rho_j)_{j \in [M]} \in \mathbb{R}_+^{2 \times M} : \gamma_j \geq 0, \rho_j = \infty \text{ for } \forall j\}$. (2)

General option: $\mathcal{I}_G = \{(\gamma_j, \rho_j)_{j \in [M]} : \gamma_j \geq 0, \rho_j \in [0, \rho] \text{ for } \forall j\}$.

The advertiser's goal in practice is to maximize their total conversion of procured ad impressions through optimizing over per-channel budgets and target ROIs, while subject to the global ROI and budget constraint similar to those in GL-OPT. Mathematically, for any option $\mathcal{I} \in \{\mathcal{I}_B, \mathcal{I}_R, \mathcal{I}_G\}$, the advertiser's optimization problem through channels can be written as

$$\begin{aligned} \text{CH-OPT}(\mathcal{I}) = \max_{(\gamma_j, \rho_j)_{j\in[M]}\in\mathcal{I}} \quad & \sum_{j\in M} \mathbb{E}\left[V_j(\gamma_j, \rho_j; \boldsymbol{z}_j)\right] \\ \text{s.t.} \quad & \sum_{j\in M} \mathbb{E}\left[V_j(\gamma_j, \rho_j; \boldsymbol{z}_j)\right] \geq \gamma \sum_{j\in M} \mathbb{E}\left[D_j(\gamma_j, \rho_j; \boldsymbol{z}_j)\right] \\ & \sum_{j\in[M]} \mathbb{E}\left[D_j(\gamma_j, \rho_j; \boldsymbol{z}_j)\right] \leq \rho\,, \end{aligned} \tag{3}$$

where the expectation is taken w.r.t. randomness in $\boldsymbol{z}_j$. We remark that for any channel $j \in [M]$, the number of auctions m_j as well as the distribution $\boldsymbol{p}_j$ are fixed and not a function of the input parameters γ_j, ρ_j.

The functions (V_j, D_j) that map per-channel target ROI and budgets γ_j, ρ_j to the total conversion and expenditure are specified by various factors including but not limited to channel j's autobidding algorithms deployed to procure ads on advertisers' behalf as well as the auctions mechanisms that sell impressions. In this work, we study a general setup that closely mimics industry practices. We assume that on the behalf of the advertiser, each channel aims to optimize their conversion over all m_j auctions while respecting the advertiser's input (i.e., per-channel target ROI and budgets). (See e.g. Meta Ads Manager in Figure 1 specifically highlights the channel's autobidding procurement methodology provides evidence to support the aforementioned setup). Hence, each channel j's optimization problem can be written as

$$\boldsymbol{x}_j^*(\gamma_j, \rho_j; \boldsymbol{z}_j) \; = \; \arg\max_{\boldsymbol{x}\in[0,1]^{m_j}} \quad \boldsymbol{v}_j^\top \boldsymbol{x} \quad \text{s.t.} \quad \boldsymbol{v}_j^\top \boldsymbol{x} \geq \gamma_j \boldsymbol{d}_j^\top \boldsymbol{x}, \quad \boldsymbol{d}_j^\top \boldsymbol{x} \leq \rho_j\,, \tag{4}$$

where $\boldsymbol{x} = (x_n)_{n\in[m_j]} \in [0,1]^{m_j}$ denotes the vector of probabilities to win each of the parallel auctions, i.e. $x_n \in [0,1]$ is the probability to win auction $n \in [m_j]$ in channel j. In light of this representation, the corresponding conversion and spend functions are given by

$$\begin{aligned} V_j(\gamma_j, \rho_j; \boldsymbol{z}_j) = \boldsymbol{v}_j^\top \boldsymbol{x}_j^*(\gamma_j, \rho_j; \boldsymbol{z}_j) \quad & \text{and} \quad V_j(\gamma_j, \rho_j) = \mathbb{E}[V_j(\gamma_j, \rho_j; \boldsymbol{z}_j)] \\ D_j(\gamma_j, \rho_j; \boldsymbol{z}_j) = \boldsymbol{d}_j^\top \boldsymbol{x}_j^*(\gamma_j, \rho_j; \boldsymbol{z}_j) \quad & \text{and} \quad D_j(\gamma_j, \rho_j) = \mathbb{E}[D_j(\gamma_j, \rho_j; \boldsymbol{z}_j)]\,. \end{aligned} \tag{5}$$

Here, the expectation is taken w.r.t. randomness in $\boldsymbol{z}_j = (\boldsymbol{v}_j, \boldsymbol{d}_j) \in \mathbb{R}_+^{m_j} \times \mathbb{R}_+^{m_j}$. We assume that for any (γ_j, ρ_j) and realization $\boldsymbol{z}_j$, $V_j(\gamma_j, \rho_j; \boldsymbol{z}_j)$ is bounded above by some absolute constant $\bar{V} \in (0, \infty)$ almost surely. We remark that Eq.(5) assumes channels are able to achieve optimal procurement performance. Later in Section 6.2, we will briefly discuss setups where channels does not optimally solve for Eq.(4).

Key focuses and organization of this work. In this paper, we address two key topics:

1. How effective are the per-channel ROI and budget levers to help advertisers achieve the globally optimal conversion GL-OPT while respecting the global ROI and budget constraints? In particular, for each of the advertiser options $\mathcal{I} \in \{\mathcal{I}_B, \mathcal{I}_R, \mathcal{I}_G\}$ defined in Eq. (2), what is the discrepancy between CH-OPT($\mathcal{I}$), i.e. the optimal conversion an advertiser can achieve in practice, versus the optimal GL-OPT?
2. Since in reality advertisers can only utilize the two per-channel levers offered by channels, how can advertisers optimize per-channel target ROIs and budgets to solve for CH-OPT($\mathcal{I}$)?

In Section 3, we address the first question to determine the gap between CH-OPT($\mathcal{I}$) and GL-OPT for different advertiser options. In Section 4, we develop an efficient algorithm to solve for per-channel levers that optimize CH-OPT($\mathcal{I}$).

3. On the efficacy of the per-channel target ROIs and budgets as levers in solving the global problem

In this section, we examine the effectiveness of the per-channel target ROI and per-channel budget levers in achieving the global optimal GL-OPT. In particular, we study if the optimal solution to the channel problem CH-OPT($\mathcal{I}$) defined in Eq. (3) for $\mathcal{I} \in \{\mathcal{I}_B, \mathcal{I}_R, \mathcal{I}_G\}$ is equal to the global optimal GL-OPT. As a summary of our results, we show that the per-channel budget only option, and the general option achieves GL-OPT, but the per-channel ROI only option can yield conversion arbitrarily worse than GL-OPT for certain instance, even when there is no global budget constraint (i.e., $\rho = \infty$). This implies that the per-channel ROI lever is inadequate to help advertisers achieve the globally optimal conversion, whereas the per-channel budget lever is effective to attain optimal conversion even when the advertiser solely uses this lever.

Our first result in this section is the following Lemma 3.1 which shows that GL-OPT serves as a theoretical upper bound for an advertiser's conversion through optimizing CH-OPT($\mathcal{I}$) with any option $\mathcal{I}$.

Lemma 3.1 (GL-OPT is the theoretical upper bound for conversion) *For any option* $\mathcal{I} \in \{\mathcal{I}_B, \mathcal{I}_R, \mathcal{I}_G\}$ *defined in Eq.* (2)*, we have* $\text{GL-OPT} \geq \text{CH-OPT}(\mathcal{I})$*, where we recall the definitions of* GL-OPT *and* CH-OPT *in Eq.* (1) *and* (3)*, respectively.*

The proof of Lemma 3.1 is deferred to Appendix A.1. In light of the theoretical upper bound GL-OPT, we are now interested in the gap between GL-OPT and CH-OPT($\mathcal{I}$) for option $\mathcal{I} \in \{\mathcal{I}_B, \mathcal{I}_R, \mathcal{I}_G\}$. In the following Theorem 3.2, we show that there exists a problem instance under which the ratio $\frac{\text{CH-OPT}(\mathcal{I}_R)}{\text{GL-OPT}}$ nears 0, implying the per-channel ROIs alone fail to help advertisers optimize conversion.

Theorem 3.2 (Per-channel ROI only option fails to optimize conversion) *Consider an advertiser with a (global) target ROI of $\gamma = 1$ procuring impressions from $M = 2$ channels, where channel 1 consists of a single auction and channel 2 consists of two auctions. The advertiser has unlimited budget $\rho = \infty$, and chooses the per-channel target ROI only option $\mathcal{I}_R$ defined in Eq. (2). Assume there is only one realization of value-cost pairs $\boldsymbol{z} = (\boldsymbol{v}_j, \boldsymbol{d}_j)_{j\in[M]}$ (i.e. the support $F = F_1 \times F_2$ is a singleton), and the realization is presented in the following table, where $X > 0$ is some arbitrary parameter. Then, for this problem instance we have $\lim_{X\to\infty} \frac{\text{CH-OPT}(\mathcal{I}_R)}{\text{GL-OPT}} = 0$.*

	Channel 1	*Channel 2*	
	Auction 1	*Auction 2*	*Auction 3*
Value $v_{j,n}$	*1*	X	$2X$
Spend $d_{j,n}$	*0*	$1+X$	$2(1+X)$

Proof. Let $\widetilde{\boldsymbol{\gamma}} = (\widetilde{\gamma}_1, \widetilde{\gamma}_2)$ be the optimal solution to CH-OPT$(\mathcal{I}_R)$ and recall under the option $\mathcal{I}_R$, we let per-channel budgets to be infinity. It is easy to see that $\widetilde{\gamma}_1$ can be any arbitrary nonnegative number because the advertiser always wins auction 1, and $\widetilde{\gamma}_2 > \frac{X}{1+X}$: if otherwise $\widetilde{\gamma}_2 \le \frac{X}{1+X}$, then the optimal outcome of channel 2 is to win both auctions 2 and 3. However, in this case, the advertiser wins all auctions and acquires total value $1 + X + 2X = 1 + 3X$, and incurs total spend $0 + (1 + X) + 2(1 + X) = 3 + 3X$, which violates the ROI constraint in CH-OPT$(\mathcal{I}_R)$ because $\frac{1+3X}{3+3X} < 1$. Therefore the advertiser can only win auction 1, or in other words $\widetilde{\gamma}_2 > \frac{X}{1+X}$. This implies that the optimal objective to CH-OPT$(\mathcal{I}_R)$ is 1. On the other hand, it is easy to see that the optimal solution to GL-OPT is to only win auctions 1 and 2, yielding an optimal value of $1 + X$. Therefore $\frac{\text{CH-OPT}(\mathcal{I}_R)}{\text{GL-OPT}} = \frac{1}{1+X}$. Taking $X \to \infty$ yeilds the desired result. $\square$

In contrast to the per-channel ROI only option, the budget only option in fact allows an advertiser's conversion to reach the theoretical upper bound GL-OPT through solely optimizing for per-channel budgets. This is formalized in the following theorem whose proof we present in Appendix A.2.

Theorem 3.3 (Per-channel budget only option suffices to achieve optimal conversion) *For the budget only option $\mathcal{I}_B$ defined in Eq.(2), we have GL-OPT = CH-OPT$(\mathcal{I}_B)$ for any global target ROI $\gamma > 0$ and total budget $\rho > 0$, even for $\rho = \infty$.*

As an immediate extension of Theorem 3.3, the following Corollary 3.4 states per-channel ROIs in fact become redundant once advertisers optimize for per-channel budgets.

Corollary 3.4 (Redundancy of per-channel ROIs) *For the general option $\mathcal{I}_G$ defined in Eq.(2) where an advertiser sets both per-channel ROI and budgets, we have GL-OPT = CH-OPT$(\mathcal{I}_G)$ for* *any aggregate ROI* $\gamma > 0$ *and total budget* $\rho > 0$*, even for* $\rho = \infty$*. Further, there exists and optimal solution* $(\gamma_j, \rho_j)_{j\in[M]}$ *to* CH-OPT$(\mathcal{I}_G)$*, s.t.* $\gamma_j = 0$ *for all* $j \in [M]$.

In light of the redundancy of per-channel ROIs as illustrated in Corollary 3.4, in the rest of the paper we will fix $\gamma_j = 0$ for any channel $j \in [M]$, and omit γ_j in all relevant notations; e.g. we will write $D_j(\rho_j; \boldsymbol{z}_j)$ and $D_j(\rho_j)$, instead of $D_j(\gamma_j, \rho_j; \boldsymbol{z}_j)$ and $D_j(\gamma_j, \rho_j)$. Equivalently, we will only consider the per-channel budget only option $\mathcal{I}_B$.

4. Optimization algorithm for per-channel budgets under bandit feedback

In this section, we develop an efficient algorithm to solve for per-channel budgets that optimize CH-OPT$(\mathcal{I}_B)$ defined in Eq. (3), which achieves the theoretical optimal conversion, namely GL-OPT, as illustrated in Theorem 3.3. In particular, we consider algorithms that run over $T > 0$ periods, where each period for example corresponds to the duration of 1 hour or 1 day. At the end of T periods, the algorithm produces some per-channel budget profile $(\rho_j)_{j\in[M]} \in [0, \rho]^M$ that approximates CH-OPT$(\mathcal{I}_B)$, and satisfies aggregate budget and ROI constraints, namely

$$\text{ROI:} \quad \sum_{j\in M} V_j(\rho_j) \geq \gamma \sum_{j\in M} D_j(\rho_j), \quad \text{and} \quad \text{Budget:} \quad \sum_{j\in[M]} D_j(\rho_j) \leq \rho\,, \tag{6}$$

where we recall the expected conversion and spend functions $(V_j(\rho_j), D_j(\rho_j))$ defined in Eq. (5).

The algorithm proceeds as follows: at the beginning of period $t \in [T]$, the advertiser sets per-channel budgets $(\rho_{j,t})_{j\in[M]}$, while simultaneously values and costs $\boldsymbol{z}_t = (\boldsymbol{v}_{j,t}, \boldsymbol{d}_{j,t}) \in \mathbb{R}_+^{M_j} \times \mathbb{R}_+^{M_j}$ are sampled (independently in each period) from finite support $F = F_1 \times \ldots F_M$ according to discrete distributions $(\boldsymbol{p}_j)_{j\in[M]}$. Each channel j then takes as input $\rho_{j,t} \in [0, \rho]$ and procures ads on behalf of the advertiser, and reports the total realized conversion $V_j(\rho_{j,t}; \boldsymbol{z}_t)$ as well as total spend $D_j(\rho_{j,t}; \boldsymbol{z}_t)$ to the advertiser, where $V_j(\rho_{j,t}; \boldsymbol{z}_t)$ and $D_j(\rho_{j,t}; \boldsymbol{z}_t)$ are defined in Eq. (5). For simplicity we also assume for any realization $\boldsymbol{z} = (\boldsymbol{v}_j, \boldsymbol{d}_j) \in F$ we have the ordering $\frac{v_{j,1}}{d_{j,1}} > \frac{v_{j,2}}{d_{j,2}} > \cdots > \frac{v_{j,m_j}}{d_{j,m_j}}$ for all channels $j \in [M]$.

Here, we highlight that the advertiser receives *bandit feedback* from channels, i.e. the advertiser only observes the numerical values $V_j(\rho_{j,t}; \boldsymbol{z}_t)$ and $D_j(\rho_{j,t}; \boldsymbol{z}_t)$, but does not get to observe $V_j(\rho_j'; \boldsymbol{z}')$ and $D_j(\rho_j'; \boldsymbol{z}')$ evaluated at any other per-channel budget $\rho_j' \neq \rho_{j,t}$ and realized value-cost pairs $\boldsymbol{z}' \neq \boldsymbol{z}_t$. More discussions on challenges that arise from this bandit feedback structure can be found in Section 4.1.

We also make the following Assumption 4.1 that states that if the advertiser allocates any feasible per-channel budget to a channel $j \in [M]$, the channel will almost surely deplete the entire budget in the impression procurement process. This is a natural assumption that mimics practical scenarios, e.g. small businesses who have moderate-sized budgets.

Assumption 4.1 (Moderate budgets) *We assume the total budget is finite, i.e.* $\rho < \infty$*, and for any channel* $j \in [M]$*, value-cost realization* $\boldsymbol{z} = (\boldsymbol{v}, \boldsymbol{d}) \in F_j$*, and per-channel budget* $\rho_j \in [0, \rho]$*, the optimal solution* $\boldsymbol{x}_j^*(\rho; \boldsymbol{z})$ *defined in Eq. (4) is budget binding, i.e.* $D_j(\rho_j; \boldsymbol{z}) = \boldsymbol{d}_j^\top \boldsymbol{x}_j^*(\rho_j; \boldsymbol{z}) = \rho_j$*.*

4.1. The SGD-UCB algorithm to optimize per-channel budgets

Here, we describe our algorithm to solve for optimal per-channel budgets w.r.t. CH-OPT$(\mathcal{I}_B)$. Similar to most algorithms for constrained optimization, we take a dual stochastic gradient descent (SGD) approach; see a comprehensive survey on dual descent methods in Bertsekas (2014). First, we consider the Lagrangian functions w.r.t. CH-OPT$(\mathcal{I}_B)$ where we let $\boldsymbol{c} = (\lambda, \mu) \in \mathbb{R}_+^2$ be the dual variables corresponding to the ROI and budget constraints, respectively:

$$\mathcal{L}_j(\rho_j, \boldsymbol{c}; \boldsymbol{z}_j) = (1+\lambda) V_j(\rho_j; \boldsymbol{z}_j) - (\lambda\gamma + \mu)\rho_j \quad \text{and} \quad \mathcal{L}_j(\rho_j, \boldsymbol{c}) = \mathbb{E}\left[\mathcal{L}_j(\rho_j, \boldsymbol{c}; \boldsymbol{z}_j)\right]. \tag{7}$$

Then, in each period $t \in [T]$ given dual variables $\boldsymbol{c}_t = (\lambda_t, \gamma_t)$, SGD decides on a primal decision, i.e. per-channel budget $(\rho_{j,t})_{j \in [M]}$ by optimizing the following:

$$\rho_{j,t} \;=\; \arg\max_{\rho_j \in [0,\rho]} \mathcal{L}_j(\rho_j, \boldsymbol{c}_t; \boldsymbol{z}_t). \tag{8}$$

Having observed the realized values $\left(V_j(\rho_{j,t}; \boldsymbol{z}_t)\right)_{j \in [M]}$ (note that spend is $(\rho_{j,t})_{j \in [M]}$ in light of Assumption 4.1), we calculate the current period violation in budget and ROI constraints, namely $g_{1,t} := \sum_{j \in M} \left(V_j(\rho_{j,t}; \boldsymbol{z}_t) - \gamma \rho_{j,t}\right)$ and $g_{2,t} = \rho - \sum_{j \in [M]} \rho_{j,t}$. Next, we update dual variables via $\lambda_{t+1} = (\lambda_t - \eta g_{1,t})_+$ and $\mu_{t+1} = (\mu_t - \eta g_{2,t})_+$, where η is some pre-specified step size.[3]

However, the above SGD approach faces a fatal drawback, namely we cannot realistically find the primal decisions by solving Eq. (8) since the function $\mathcal{L}_j(\cdot, \boldsymbol{c}_t; \boldsymbol{z}_t)$ is unknown due to the bandit feedback structure. Therefore, we provide a modification to SGD to handle this issue. Before we present our approach, we briefly note that although bandit feedback prevents the naive application of SGD for our problem of interest, this may not be the case in other online advertising scenarios that involve relevant learning tasks, underlining the challenges of our problem; see following Remark 4.1 for details.

Remark 4.1 *Our problem of interest under bandit feedback is more difficult than similar problems in related works that study online bidding strategies under budget and ROI constraints; see e.g. Balseiro et al. (2017, 2022), Feng et al. (2022). To illustrate, consider for instance Balseiro et al. (2017) in which a budget constrained advertiser's primal decision at period* t *is to submit a bid value* b_t *after observing her value* v_t*. The advertiser competes with some unknown highest competing bid* d_t *in the*

[3] Here, the dual updates follow the vanilla gradient descent approach, and one can also employ more general mirror descent updates; see e.g. Balseiro et al. (2022).

market, and after submitting bid b_t, does not observe d_t if she does not win the competition, which involves a semi-bandit feedback structure. Nevertheless, the corresponding Lagrangian under SGD takes the special form $\mathcal{L}_j(b,\mu_t;\boldsymbol{z}_t) = (v_t - (1+\mu_t)d_t)\,\mathbb{I}\{b_t \geq d_t\}$ where μ_t is the dual variable w.r.t. the budget constraint. This simply allows an advertiser to optimize for her primal decision by bidding $\arg\max_{b\geq 0}\mathcal{L}_j(b,\boldsymbol{c}_t;\boldsymbol{z}_t) = \frac{v_t}{1+\mu_t}$*. So even though Balseiro et al. (2017, 2022), Feng et al. (2022) study dual SGD under bandit feedback, the special structures of their problem instances permits SGD to effectively optimize for primal decisions in each period, as opposed to Eq.* (8) *in our setting which can not be solved.*

To resolve challenges that arise with bandit feedback in our model, we take a natural approach to augment SGD with the *upper-confidence bound (UCB)* algorithm, which is well celebrated for solving learning problems under bandit feedback such as multi-arm bandits; see an introduction to bandits in Slivkins et al. (2019). In particular, we first discretize our per-channel buudget decision set $[0,\rho]$ into granular "arms" that are separated by some distance $\delta > 0$, so that the discretized per-channel budget decisions become

$$\mathcal{A}(\delta) = \{a_k\}_{k\in[K]} \text{ where } a_k = (k-1)\delta \text{ and } K := \lceil \rho/\delta \rceil + 1\,. \tag{9}$$

In the following we will use the terms "per-channel budget" and "arm" interchangeably. In the spirit of UCB, in each period t we maintain some estimate $\left(\bar{V}_j(a_k)\right)_{j\in[M]}$ of the conversions $\left(V_j(a_k)\right)_{j\in[M]}$ as well as an upper confidence bound $\texttt{UCB}_{j,t}(a_k)$ for each arm a_k using historical payoffs from periods in which arm a_k is pulled. Finally, we update primal decisions for each channel $j \in [M]$ using the "best arm" $\rho_{j,t} = \arg\max_{a_k\in\mathcal{A}(\delta)}\ (1+\lambda_t)\left(\bar{V}_{j,t}(a_k) + \texttt{UCB}_{j,t}(a_k)\right) - (\lambda_t\gamma + \mu_t)\,a_k$. We summarize our algorithm, called SGD-UCB, in the following Algorithm 1.

We remark that there has been very recent works that combine SGD with Thompson sampling which is another well-known algortithm for solving bandit problems (e.g. Ding et al. (2021) and references therein), and works that employ SGD in bandit problems (e.g. Han et al. (2021)). Yet to the best of our knowledge, approach to augment SGD with UCB is novel.

4.2. Analyzing the SGD-UCB algorithm

In this subsection, we analyze the performance of SGD-UCB in Algorithm 1, and present accuracy guarantees on the final output $\overline{\boldsymbol{\rho}}_T = \left(\frac{1}{T}\sum_{t\in[T]}\rho_{j,t}\right)_{j\in[M]}$ of the algorithm. The backbone of our analysis strategy is to show the cumulative conversion loss over T periods, namely $T\cdot\text{GL-OPT} - \mathbb{E}\left[\sum_{t\in[T]}\sum_{j\in[M]}V_j(\rho_{j,t})\right]$ consists of two main parts, namely the error induced by the UCB in our algorithm, and the error due to SGD (or what is typically viewed as the deviations from complementary slackness), as shows in the following Proposition 4.1. Then we further bound each part, respectively.

Algorithm 1 SGD-UCB

Input: Budget discretization decision set $\mathcal{A}(\delta)$ defined in Eq.(9). Step size $\eta > 0$. Initialize $N_{j,1}(a_k) = \bar{V}_{j,1}(a_k) = 0$ for all $j \in [M]$ and $k \in [K]$, and dual variables $\lambda_1 = \mu_1 = 0$.

1: **Output:** Per channel budget.

2: **for** $t = 1 \ldots T$ **do**

3: **Update (primal) per-channel budget.** For each channel $j \in [M]$ set (primal) per-channel budget:

- If $t \leq K$, set $\rho_{j,t} = a_t$.
- If $t > K$, set

$$\rho_{j,t} = \arg\max_{a_k \in \mathcal{A}(\delta)} \ (1+\lambda_t)\left(\bar{V}_{j,t}(a_k) + \texttt{UCB}_{j,t}(a_k)\right) - (\lambda_t \gamma + \mu_t)\, a_k, \quad \text{where } \texttt{UCB}_{j,t}(a_k) = \sqrt{\frac{2\log(T)}{N_{j,t}(a_k)}} \tag{10}$$

4: Observe realized values $\{V_j(\rho_{j,t}; \boldsymbol{z}_t)\}_{j \in [M]}$, and update for each arm $k \in [K]$:

$$\begin{aligned} &N_{j,t+1}(a_k) = N_{j,t}(a_k) + \mathbb{I}\{\rho_{j,t} = a_k\} \\ &\bar{V}_{j,t+1}(a_k) = \frac{1}{N_{j,t+1}(a_k)}\left(N_{j,t}(a_k)\bar{V}_{j,t}(a_k) + V_j(\rho_{j,t}; \boldsymbol{z}_t)\mathbb{I}\{\rho_{j,t} = a_k\}\right) \quad , \text{ for } j = 1 \ldots M \end{aligned} \tag{11}$$

5: **Update dual variables.** Update dual variables with $g_{1,t} := \sum_{j \in M}\left(V_j(\rho_{j,t}; \boldsymbol{z}_t) - \gamma \rho_{j,t}\right)$ and $g_{2,t} = \rho - \sum_{j \in [M]} \rho_{j,t}$:

$$\lambda_{t+1} = \left(\lambda_t - \eta g_{1,t}\right)_+ \ \text{ and } \ \mu_{t+1} = \left(\mu_t - \eta g_{2,t}\right)_+ \tag{12}$$

6: **end for**

7: Output $\overline{\boldsymbol{\rho}}_T = \left(\frac{1}{T}\sum_{t \in [T]} \rho_{j,t}\right)_{j \in [M]}$.

Proposition 4.1 *For any channel $j \in [M]$ define $\rho_j^*(t) = \arg\max_{\rho_{j \in [M]}} \mathcal{L}_j(\rho_j; \boldsymbol{c}_t)$ to be the optimal per-channel budget w.r.t. dual variables $\boldsymbol{c}_t = (\lambda_t, \mu_t)_{t \in [T]}$ during period $t \in [T]$. Then we have*

$$\begin{aligned} &T \cdot \text{GL-OPT} - \sum_{t \in [T]} \sum_{j \in [M]} \mathbb{E}\left[V_j(\rho_{j,t})\right] \\ \leq\ & M\bar{V}K + \sum_{j \in [M]} \underbrace{\sum_{t > K} \mathbb{E}\left[\mathcal{L}_j(\rho_j^*(t), \lambda_t, \mu_t) - \mathcal{L}_j(\rho_{j,t}, \lambda_t, \mu_t)\right]}_{\textit{UCB error}} + \underbrace{\sum_{t > K} \mathbb{E}\left[\lambda_t g_{1,t} + \mu_t g_{2,t}\right]}_{\textit{SGD complementary slackness deviations}} \quad , \end{aligned} \tag{13}$$

where we recall the definitions of $g_{1,t}$ and $g_{2,t}$ in step 4 of Algorithm 1, and the fact that the conversion $V_j(\rho_j; \boldsymbol{z}_j)$ is bounded above by absolute constant $\bar{V} \in (0, \infty)$ almost surely for any channel $j \in [M]$, (γ_j, ρ_j) and realization $\boldsymbol{z}_j$.

The bound on SGD complementary slackness violation is presented in the following Lemma 4.2, and follows a standard analyses for SGD; we refer readers to the proof in Appendix B.2.

Lemma 4.2 (Bounding complementary slackness deviations) *Recall $g_{1,t} = \sum_{j \in [M]}\left(V_j(\rho_{j,t}; \boldsymbol{z}_t) - \gamma \rho_{j,t}\right)$, $g_{2,t} = \rho - \sum_{j \in [M]} \rho_{j,t}$ and $\eta > 0$ the step size defined in Algorithm 1. Then we have*

$$\sum_{t \in [T]} \mathbb{E}\left[\lambda_t g_{1,t} + \mu_t g_{2,t}\right] \ \leq\ \mathcal{O}\left(\eta T + \frac{1}{\eta}\right) \tag{14}$$

Challenges in bounding UCB error due to adversarial contexts and continuum-arm dicretization. Bounding our UCB error is much more challenging than doing so in classic stochastic multi-arm bandit settings: first, our setup involves discretizing a continuum of arms i.e. our discretization in Eq.(9) for $[0,\rho]$; second, and more importantly, the dual variables $\{\boldsymbol{c}_t\}_{t\in[T]}$ are effectively *adversarial contexts* since they are updated via SGD instead of being stochastically sampled from some nice distribution, and correspondingly the Lagrangian function $\mathcal{L}_j(a_k,\boldsymbol{c}_t;\boldsymbol{z}_t)$ can be viewed as a reward function that maps any arm-context pair $(a_k,\boldsymbol{c}_t)$ to (stochastic) payoffs. Both continuum-arms and adversarial contexts have been notorious in making reward function estimations highly inefficient; see e.g. discussions in Agrawal (1995), Agarwal et al. (2014). We further elaborate on specific challenges that adversarial contexts bring about:

- **Boundedness of rewards.** In classic stochastic multi-arm bandtis and UCB, losses in total rewards grow linearly with the magnitude of rewards. In our setting, the reward function, i.e. the Lagrangian function $\mathcal{L}_j(a_k,\boldsymbol{c}_t;\boldsymbol{z}_t)$, scales linearly with the magnitude of contexts (see Eq. (7), so large contexts (i.e. large dual variables) may lead to large losses.
- **Context-dependent exploration-exploitation tradeoffs.** The typical trade-off for arm exploration and exploitation in our setting depends on the particular values of the contexts (i.e. the dual variables), which means there may exist "bad" contexts that lead to poor tradeoffs that require significantly more explorations to achieve accurate estimates of arm rewards than other "good" contexts. We elaborate more in Lemma 4.5 and discussions thereof.

In the following, we first handle continuum arm discretization via analyzing structural properties of the reward (i.e. Lagrangian) functions. Fortunately, the specific form of conversion functions $V(\rho_j;\boldsymbol{z})$ defined in Eq. (4) imposes a salient structure on the Lagrangian for pulling an arm. Specifically, the following lemma shows that the Lagrangian is continuous, piecewise linear, concave, and unimodal[4]; we present the proof in Appendix B.3

Lemma 4.3 (Structural properties of conversion and Lagrangian functions)

- *For any channel $j\in[M]$ and per-channel budget ρ_j the conversion function $V_j(\rho_j)$ is continuous, piece-wise linear, strictly increasing, and concave. In particular, $V_j(\rho_j)$ takes the form*

$$V_j(\rho_j)=\sum_{n\in[S_j]}(s_{j,n}\rho_j+b_{j,n})\,\mathbb{I}\{r_{j,n-1}\le\rho_j\le r_{j,n}\}\,, \tag{15}$$

where the parameters $S_j\in\mathbb{N}$ and $\{(s_{j,n},b_{j,n},r_{j,n})\}_{n\in[S_j]}$ only depend on the support F_j and distribution $\boldsymbol{p}_j$ from which value-to-cost pairs are sampled. These parameters satisfy $s_{j,1}>s_{j,2}>\cdots>s_{j,S_j}>0$ and $0=r_{j,0}<r_{j,1}<r_{j,2}<\cdots<r_{j,S_j}=\rho$, as well as $b_{j,n}>0$ s.t. $s_{j,n}r_{j,n}+b_{j,n}=s_{j,n+1}r_{j,n}+b_{j,n+1}$ for all $n\in[S_j-1]$, implying $V_j(\rho_j)$ is continous in ρ_j.

[4] We say a real-valued function $f:\mathbb{R}\to\mathbb{R}$ is unimodal if there exists some y^* such that $f(y)$ strictly increases when $y\le y^*$ and strictly decreases when $y\ge y^*$.

- *For any dual variables* $\boldsymbol{c} = (\lambda, \mu) \in \mathbb{R}_+^2$*, the Lagrangian function* $\mathcal{L}_j(\rho_j, \boldsymbol{c})$ *defined in Eq. (7) is continuous, piece-wise linear, concave, and unimodal in* ρ_j*. In particular,*

$$\mathcal{L}_j(\rho_j, \boldsymbol{c}) = \sum_{n \in [S_j]} \left(\sigma_{j,n}(\boldsymbol{c})\rho_j + (1+\lambda) b_{j,n}\right) \mathbb{I}\{r_{j,n-1} \le \rho_j \le r_{j,n}\}, \tag{16}$$

where the slope $\sigma_{j,n}(\boldsymbol{c}) = (1+\lambda)s_{j,n} - (\mu + \gamma\lambda)$*. Then, this also implies* $\arg\max_{\rho_j \ge 0} \mathcal{L}_j(\rho_j, \boldsymbol{c}) = \max\{r_{j,n} : n = 0, 1 \ldots, S_j\ , \sigma_{j,n}(\boldsymbol{c}) \ge 0\}$.

In fact, for any realized value-cost pairs $\boldsymbol{z}$, the "realization versions" of the conversion and Lagrangians functions, namely $V_j(\rho_j; \boldsymbol{z})$ and $\mathcal{L}_j(\rho_j, \boldsymbol{c}; \boldsymbol{z})$, also satisfy the same properties as those of $V_j(\rho_j)$ and $\mathcal{L}_j(\rho_j, \boldsymbol{c})$, respectively. We provide a visual illustration for these structural properties in Figure 2.

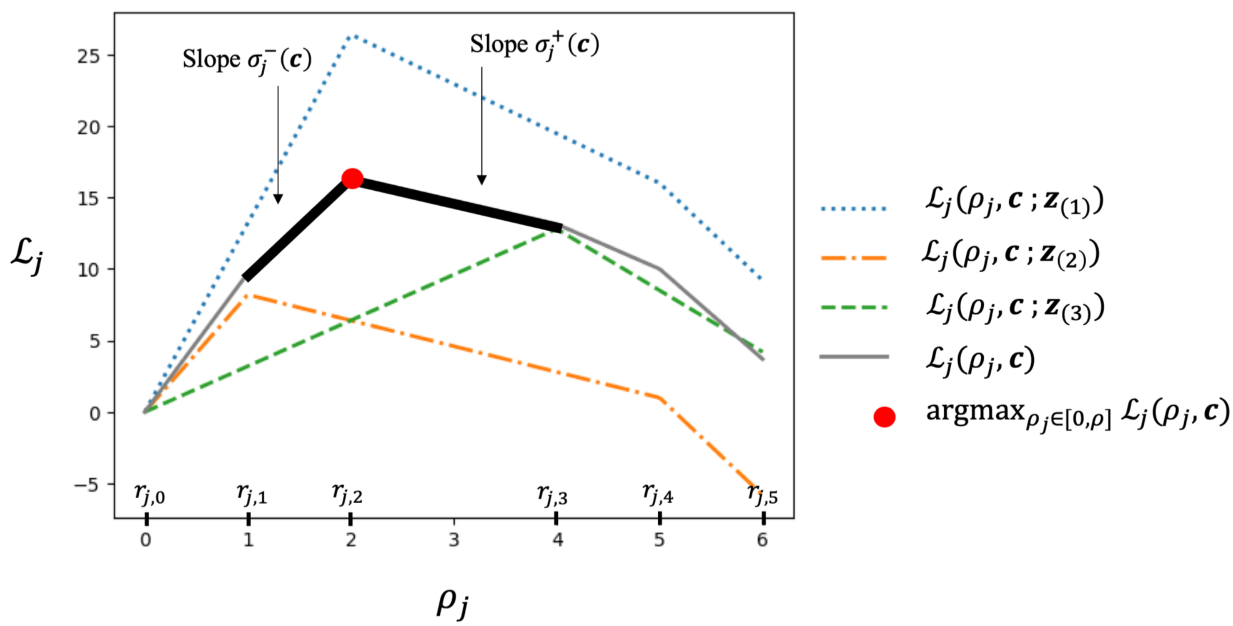

Figure 2 **Illustration of Lagrangian functions defined in Eq.** (7) **with** $M_j = 2$ **auctions in channel** j**, and support** F_j **that contains 3 elements,** $\boldsymbol{z}_{(1)} = (\boldsymbol{v}_{(1)}, \boldsymbol{d}_{(1)}) = ((8,2),(2,3))$**,** $\boldsymbol{z}_{(2)} = ((3,4),(1,4))$**,** $\boldsymbol{z}_{(3)} = ((8,1),(4,2))$**, and context** $\boldsymbol{c} = (\lambda, \mu) = (4,2)$**. In light of Lemma 4.3,** $S_j = 5$**, where the "turning points"** $r_{j,0} \ldots r_{j,S_j}$ **are indicated on the x-axis, and the optimal budget w.r.t.** $\boldsymbol{c}$ **is** $\arg\max_{\rho_j \in [0,\rho]} \mathcal{L}_j(\rho_j; \boldsymbol{c}) = r_{j,2}$**. The adjacent slopes in Eq.** (18) **are** $\sigma_j^-(\boldsymbol{c}) = \sigma_{j,2}(\boldsymbol{c})$**, and** $\sigma_j^+(\boldsymbol{c}) = \sigma_{j,3}(\boldsymbol{c})$**, respectively.**

We now handle the reward boundedness issue in the Lagrangian functions defined in Eq. (7) that arise from adversarial contexts. In the following Lemma 4.4, we show that the Lagrangian functions, as well as dual variables, are indeed bounded by some absolute constants under a mild feasibility Assumption 4.2 stated below:

Assumption 4.2 (Strictly feasible global ROI constraints) *For any realization of value-cost pairs* $\boldsymbol{z} = (\boldsymbol{v}_j, \boldsymbol{d}_j)_{j\in[M]} \in F_1 \times \dots F_M$*, the realized version of the ROI constraint in* GL-OPT *defined in Eq. 1 is strictly feasible, i.e. the set* $\left\{\boldsymbol{x} = (\boldsymbol{x}_j)_{j\in[M]} : \boldsymbol{x}_j \in [0,1]^{m_j} \text{ for } \forall j \in [M],\ \sum_{j\in[M]} \boldsymbol{v}_j^\top \boldsymbol{x}_j > \gamma \sum_{j\in[M]} \boldsymbol{d}_j^\top \boldsymbol{x}_j\right\}$ *is nonempty.*[5]

Lemma 4.4 (Bounding dual variables and Lagrangian functions) *Let* $(\lambda_t, \mu_t)_{t\in[T]}$ *be the variables generated from Algorithm 1. Under Assumption 4.2, and assuming the step size* $\eta > 0$ *satisfies* $\eta < \frac{1}{M\cdot\max\{\bar{V},\rho,\bar{V}^2\}}$*, there exists some absolute constant* $C_F > 0$ *that depends only on the support of value-cost pairs* $F = F_1 \times \cdots \times F_M$ *as well as aggregate target ROI and budget* (γ, ρ) *such that* $\lambda_t, \mu_t \le C_F$ *for all* $t \in [T]$*. Moreover, for any for any* $t \in [T]$*,* $j \in [M]$ *and* $\rho_j \in [0, \rho]$ *we have*

$$-(1+\gamma)\,\rho C_F \ \le\ \mathcal{L}_j(\rho_j, \lambda_t, \mu_t) \ \le\ (1+C_F)\bar{V}\,. \tag{17}$$

See proof in Appendix B.4.

Finally, we address the context-dependent exploration-exploitation tradeoff. We remark that this tradeoff is embodied in the "flatness" of the reward function that depends on adversarial contexts. To illustrate (see e.g. Figure 2), we define the slopes that are adjacent to the optimal budget $\arg\max_{\rho_j\in[0,\rho]} \mathcal{L}_j(\rho_j, \boldsymbol{c})$ for any $\boldsymbol{c} = (\lambda, \mu)$ as followed: assuming the optimal budget $\arg\max_{\rho_j\in[0,\rho]} \mathcal{L}_j(\rho_j, \boldsymbol{c})$ is located at the nth "turning point" $r_{j,n}$ we have

$$\sigma_j^-(\boldsymbol{c}) = \sigma_{j,n}(\boldsymbol{c}) \ \text{ and } \ \sigma_j^+(\boldsymbol{c}) = \sigma_{j,n+1}(\boldsymbol{c}) \tag{18}$$

Similar to standard multi-arm bandits exploration-exploitation tradeoffs, the flatter the slope (e.g. $\sigma_j^-(\boldsymbol{c})$ is close to 0), the more pulls required to accurately estimate rewards for sub-optimal arms on the slope, but the lower the loss in conversion for pulling sub-optimal arms. Our setting is challenging because the magnitude of this tradeoff depends on the adversarial contexts $\boldsymbol{c}_t$, i.e., the dual variables, which requires delicate treatments. In the following Lemma 4.5 where we bound the UCB error, we handle this context-dependent tradeoff by separately analyzing periods during which the adjacent slopes $\sigma_j^-(\boldsymbol{c})$ and $\sigma_j^+(\boldsymbol{c})$ are less or greater than some parameter $\underline{\sigma}$, and characterize the context-dependent tradeoff w.r.t. flatness of adjacent slopes using $\underline{\sigma}$.

Lemma 4.5 (Bounding UCB error in primal per-channel budgets) *Assume the discretization width* δ *satisfies* $\delta < \underline{r}_j := \min_{n\in[S_j]} r_{j,n} - r_{j,n-1}$*, where* S_j *and* $\{r_{j,n}\}_{n=0}^{S_j}$ *are defined in Lemma 4.3. Then we have*

$$\sum_{t\in[T]} \mathbb{E}\left[\mathcal{L}_j(\rho_{j,t}^*, \lambda_t, \mu_t) - \mathcal{L}_j(\rho_{j,t}, \lambda_t, \mu_t)\right] \ \le\ \mathcal{O}\left(\delta T + \underline{\sigma} T + \frac{1}{\underline{\sigma}\delta}\right) \tag{19}$$

where $\underline{\sigma} > 0$ *is any small positive number.*

[5] Equivalently, for any realization of value-cost pairs $\boldsymbol{z} = (\boldsymbol{v}_j, \boldsymbol{d}_j)_{j\in[M]}$ there always exists a channel $j \in [M]$ and an auction $n \in [m_j]$ in this channel whose value-to-cost ratio is at least γ, i.e. $v_{j,n} > \gamma d_{j,n}$.

We refer the readers to Appendix B.5 for the proof. Note that the parameter $\underline{\sigma}$ will be chosen later. Finally, returning to bounding the UCB error in Proposition 4.1, we put together Lemmas 4.2 and 4.5, and obtain the main result of this section in the following Theorem 4.6 whose proof we detail in Appendix B.6

Theorem 4.6 (Putting everything together) *Assume assumptions 4.1 and 4.2 hold. Let* $(\rho_{j,t})_{j\in[M],t\in[T]}$ *be the per-channel budgets generated from Algorithm 1 and assume we take step size* $\eta = \Theta(1/\sqrt{T})$*, discretization width* $\delta = \Theta(T^{-1/3})$*, and* $\underline{\sigma} = \Theta(T^{-1/3})$ *in Lemma 4.5. Then for large enough* T *we have* $T \cdot \text{GL-OPT} - \mathbb{E}\left[\sum_{t\in[T]}\sum_{j\in[M]} V_j(\rho_{j,t})\right] \leq \mathcal{O}(T^{2/3})$*. Recalling* $\overline{\boldsymbol{\rho}}_T = \left(\frac{1}{T}\sum_{t\in[T]}\rho_{j,t}\right)_{j\in[M]}$ *is the vector of time-averaged per-channel budgets, this implies*

$$\text{GL-OPT} - \sum_{j\in[M]} \mathbb{E}\left[V_j(\overline{\rho}_{T,j})\right] \leq \mathcal{O}(T^{-1/3}),$$

as well as approximate constraint satisfaction

$$\sum_{j\in[M]} \mathbb{E}\left[V_j(\overline{\rho}_{T,j}) - \gamma\overline{\rho}_{T,j}\right] \geq -\mathcal{O}(T^{-1/2}), \quad \textit{and } \rho - \sum_{j\in[M]} \mathbb{E}[\overline{\rho}_{T,j}] \geq -\mathcal{O}(T^{-1/2})$$

Here, we note that the above approxamate constraint satisfaction is in expectation, similar to our definition of $\text{CH-OPT}(\mathcal{I}_B)$ defined in Eq. (3). To conclude, we make an important remark that distinguishes our result in Theorem 4.6 with related literature on convex optimization.

Remark 4.2 *In light of Lemma 4.3, the advertiser's optimization problem* $\text{CH-OPT}(\mathcal{I}_B)$ *in Eq.* (3) *effectively becomes a convex problem (see Proposition B.4 in Appendix B.9). Hence it may be tempting for one to directly employ off-the-shelf convex optimization algorithms. However, our problem involves stochastic bandit feedback, and more importantly, uncertain constraints, meaning that we cannot analytically determine whether a primal decision satisfies the constraints of the problem. For example, in* $\text{CH-OPT}(\mathcal{I}_B)$*, for some primal decision* $(\rho_j)_{j\in[M]}$*, we cannot determine whether the ROI constraint* $\sum_{j\in M}\mathbb{E}\left[V_j(\gamma_j,\rho_j;\boldsymbol{z}_j) - \gamma D_j(\gamma_j,\rho_j;\boldsymbol{z}_j)\right] \geq 0$ *holds because the distribution* $(\boldsymbol{p}_j)_{j\in[M]}$ *from which* $\boldsymbol{z}$ *is sampled is unknown. To the best of our knowledge, there are only two recent works that handle a similar stochastic bandit feedback, and uncertain constraint setting, namely Usmanova et al. (2019) and Nguyen and Balasubramanian (2022). Nevertheless, our setting is more challenging because these works consider a "two-point estimation" regime where one can make function evaluations to the objective and constraints twice each period, whereas our setting involves "one-point estimation" such that we can only make function calls once per period. We note the optimal oracle complexities for unknown constraint convex optimization with one-point bandit feedback, remains an open problem.*[6]

[6] See Table 4.1 in Larson et al. (2019) for best known complexity bounds for one-point bandit feedback setups.

4.3. Extension to strict constraint satisfaction: UCB-SGD-II

In Theorem 4.6, we showed that the final output of the UCB-SGD Algorithm 1 outputs a per-channel budget profile $\overline{\boldsymbol{\rho}}_T = \left(\frac{1}{T}\sum_{t\in[T]}\rho_{j,t}\right)_{j\in[M]}$ that satisfies both ROI and budget constraints in CH-OPT$(\mathcal{I}_B)$ approximately, i.e. there can be at most violations in the magnitude of $\mathcal{O}(T^{-1/2})$ for both constraints. In this subsection, we present a modification to UCB-SGD that enables us to achieve no-constraint violations, while still retaining the $\mathcal{O}(T^{-1/3})$ accuracy in total conversion. Similar modification techniques have been introduced in Balseiro et al. (2022), Feng et al. (2022).

Our modification strategy handles ROI and budget constraint satisfactions differently. For budget constraints, we simply maintain a spend balance B_t in each period starting from $B_1 = 0$, and increase the balance by the expenditure in each period. When the balance nears ρT, i.e. total spend comes close to ρT, we simply terminate the algorithm. Regarding the ROI constraint, we develop two phases. Phase 1 is a "safety buffer phase" where we conservatively set per-channel budgets to accumulate a positive "ROI balance", i.e. in this phase (assume ending in period T_1) we hope to achieve $\sum_{t\in[T_1]} g_{1,t} \geq \Theta(\sqrt{T})$, where we recall $-g_{1,t}$ defined in step 4 of Algorithm 1 can be viewed as the ROI constraint violations in period t. For phase 2, we then naively run SGD-UCB. The motivation for this two-phase design is that we aim to have a buffer, i.e. positive ROI balance, in phase 1 that can compensate for possible constraint violations in phase 2 when we run SGD-UCB (see Theorem 4.6). We call our algorithm SGD-UCB-II which we present in Algorithm 2.

We remark that in order to implement the buffer phase 1 to attain a positive ROI balance, we rely on the following Assumption 4.3 which is a strengthened version of Assumption 4.2 that states in each channel there is always an auction that has a value-to-cost ratio above the global target ROI γ. Then by setting a small budget we denote as β, the channel will only procure impressions with high value-to-cost ratios (due to the structure of conversion functions in Lemma 4.3), and thus ensuring that the ROI balance increases.

Assumption 4.3 (Strictly feasible per-channel ROI constraints) *Fix any channel $j \in [M]$ and any realization of value-cost pairs $\boldsymbol{z}_j = (\boldsymbol{v}_j, \boldsymbol{d}_j) \in F_j$, the channel's optimization problem in Eq. 4 is strictly feasible, i.e. the set $\left\{\boldsymbol{x}_j \in [0,1]^{m_j} : \boldsymbol{v}_j^\top \boldsymbol{x}_j > \gamma \boldsymbol{d}_j^\top \boldsymbol{x}_j\right\}$ is nonempty.*[7]

Our strategy to bound the performance of SGD-UCB-II is as followed: In the first phase, we show that we acquire sufficient ROI balance buffers to compensate for global ROI constraint violation in the second phase. On the other hand, conversion loss in the second phase is solely due to SGD-UCB (Algorithm 1), and thus our proof to bound such loss follows from similar ideas in Theorem 4.6 (but here we have to additionally handle "early stopping" of UCB-SGD-II due to the spend balance check.

[7] Equivalently, for any realization of value-cost pairs $\boldsymbol{z}_j = (\boldsymbol{v}_j, \boldsymbol{d}_j)$ there always exists an auction $n \in [m_j]$ in this channel whose value-to-cost ratio is at least γ, i.e. $v_{j,n} > \gamma d_{j,n}$.

Algorithm 2 SGD-UCB-II

Input: Set spend balance $B_1 = 0$, and $g_{1,0} = 0$, $\beta > 0$

Phase 1 – Accumulate ROI balance buffer

1: **while** $\sum_{t' \in [t-1]} g_{1,t'} \le \sqrt{T}\log(T)$ **do**
2:     **if** $B_t + M\rho > \rho T$ **then**
3:         Terminate algorithm and output $\overline{\boldsymbol{\rho}} = \left(\frac{1}{T}\sum_{t' \in [t]} \rho_{j,t'}\right)_{j \in [M]}$.
4:     **end if**
5:     Set $\rho_{j,t} = \beta$ for all $j \in [M]$
6:     Observe conversion $V_j(\beta; \boldsymbol{z}_t)$ for all channels $j \in [M]$. Calculate

$$g_{1,t} = \sum_{j \in M} \left(V_j(\rho_{j,t}; \boldsymbol{z}_t) - \gamma \rho_{j,t}\right).$$

7:     Calculate $B_{t+1} = B_t + \sum_{j \in [M]} \rho_{j,t}$.
8:     Increment $t \leftarrow t + 1$
9: **end while**
10: Denote end period of Phase 1 as $T_1 = t - 1$.

Phase 2 – Run SGD-UCB

11: For remaining $T - T_1$ periods, run SGD-UCB in Algorithm 1 with a spend balance check during each period t:
12: **if** $B_t + M\rho > \rho T$ **then**
13:     Terminate and output $\overline{\boldsymbol{\rho}} = \left(\frac{1}{T}\sum_{t' \in [t]} \rho_{j,t'}\right)_{j \in [M]}$.
14: **else**
15:     Set per-channel budgets $\{\rho_{j,t}\}_{j \in [M]}$ according to SGD-UCB for all channels.
16:     Update spend balance $B_{t+1} = B_t + \sum_{j \in [M]} \rho_{j,t}$.
17: **end if**

Note that in the first ROI balance buffer phase, we are not optimizing for per-channel budgets which may lead to significant per-period conversion loss. Nevertheless, in the following lemma, we first show that the first ROI balance buffer phase does not last too long; We refer readers to Appendix B.7 for the proof.

Lemma 4.7 (Bounding length of Phase 1) *Recall $T_1 \in [T]$ is the end period of Phase 1 in the SGD-UCB-II algorithm (see step 10). Denote the event $\mathcal{E} = \{T_1 \ge 2\sqrt{T}\log^3(T)\}$, and take small budget $\beta = \frac{1}{\log(T)}$ in the SGD-UCB-II algorithm. Then, under Assumption 4.2, for large enough T we have $\mathbb{P}(\mathcal{E}) \le \frac{1}{T}$.*

Our main result in this subsection is the following Theorem 4.8 which bounds the conversion loss of SGD-UCB-II. The proof is detailed in Appendix B.8.

Theorem 4.8 *Let* $T_2 \in [T]$ *be the termination period of SGD-UCB-II, and recall* $\overline{\boldsymbol{\rho}} = \left(\frac{1}{T}\sum_{t\in[T_2]}\rho_{j,t}\right)_{j\in[M]}$ *is the final outputs of SGD-UCB-II (Algorithm 2). Then under Assumptions 4.1 and 4.3 and the same SGD-UCB (Algorithm 1) parameter choices in Theorem 4.6, for large enough* T *we have* $\text{GL-OPT} - \sum_{j\in[M]}\mathbb{E}\left[V_j(\overline{\rho}_j)\right] \leq \mathcal{O}(T^{-1/3})$*, and further* $\sum_{j\in[M]}\mathbb{E}\left[V_j(\overline{\rho}_j) - \gamma\overline{\rho}_j\right] \geq 0$ *and* $\sum_{j\in[M]}\mathbb{E}\left[\overline{\rho}_j\right] \leq \rho$.

5. Generalizing to autobidding in multi-item auctions

In previous sections, we assumed that each channel consists of multiple auctions, each of which is associated with the sale of a single ad impression (see Eq. (4) and discussions thereof). Yet, in practice, there are many scenarios in which ad platforms sell multiple impressions in each auction (see e.g. Varian (2007), Edelman et al. (2007)). Thereby in this section, we extend all our results for the single-item auction setting in previous sections to the multi-item auction setup. In Section 5.1, we formally describe the multi-item setup; in Section 5.2 we show that in the multi-item setting, the per-budget ROI lever is again redundant (similar to what is shown in Theorem 3.3 and Corollary 3.4), and an advertiser can solely optimize over per-channel budgets to achieve the global optimal conversion; in Section 5.3, we show our proposed UCB-SGD algorithm is directly applicable to the multi-item auction setup for a broad class of auctions, and similar to Theorem 4.6, our algorithm produces accurate lever estimates with which the advertiser can approximate the globally optimal lever decisions.

5.1. Multi-item autobidding setup

We first formalize our multi-item setup as followed. For each auction $n \in [m_j]$ of channel $j \in [M]$, assume $L_{j,n} \in \mathbb{N}$ impressions are sold, and channel j is only allowed to procure at most 1 impression in auction n on the advertiser's behalf. The value acquired and cost incurred by the advertiser when procuring impression $\ell \in [L_{j,n}]$ are $v_{j,n}(\ell)$ and $d_{j,n}(\ell)$, respectively. With a slight abuse of notation from previous sections, we write $\boldsymbol{v}_{j,n} = (v_{j,n}(1), \dots v_{j,n}(L_{j,n})) \in \mathbb{R}_+^{L_{j,n}}$ as the $L_{j,n}$-dimensional vector that includes all impression values of auction n in channel j, and further write $\boldsymbol{v}_j = (\boldsymbol{v}_{j,1} \dots \boldsymbol{v}_{j,m_j}) \in \mathbb{R}_+^{\sum_{n\in[m_j]} L_{j,n}}$ as the vector that concatenates all value vectors across auctions in channel j. We also define $\boldsymbol{d}_{j,n} \in \mathbb{R}_+^{L_{j,n}}$ and $\boldsymbol{d}_j \in \mathbb{R}_+^{\sum_{n\in[m_j]} L_{j,n}}$ accordingly for costs. Similar to Section 2, we assume $\boldsymbol{z}_j = (\boldsymbol{v}_j, \boldsymbol{d}_j)$ is sampled from finite support F_j according to discrete distribution $\boldsymbol{p}_j$ for any channel $j \in [M]$, and further we assume that for any element $\boldsymbol{z}_j \in F_j$, individual impressions in each auction are labelled such that $v_{j,n}(1) > \ldots > v_{j,n}(L_{j,n})$ for any $n \in [m_j]$.

Under the above multi-item setup, an advertiser's global optimization problem (analogous to GL-OPT in Eq. (1) for the single-item auction setup in previous sections), can be written as the following problem called GL-OPT$^+$:

$$\begin{aligned}
\text{GL-OPT}^+ \ = \\
\max_{\left(\boldsymbol{x}_j=(\boldsymbol{x}_{j,1}\ldots,\boldsymbol{x}_{j,m_j})\right)_{j\in[M]}} & \sum_{j\in[M]} \mathbb{E}\left[\boldsymbol{v}_j^\top \boldsymbol{x}_j\right] \\
\text{s.t.} & \sum_{j\in[M]} \mathbb{E}\left[\boldsymbol{v}_j^\top \boldsymbol{x}_j\right] \ \geq \ \gamma \sum_{j\in[M]} \mathbb{E}\left[\boldsymbol{d}_j^\top \boldsymbol{x}_j\right] \\
& \sum_{j\in[M]} \mathbb{E}\left[\boldsymbol{d}_j^\top \boldsymbol{x}_j\right] \ \leq \ \rho \\
& \boldsymbol{x}_{j,n} \in [0,1]^{\sum_{n\in[m_j]} L_{j,n}} \text{ and } \sum_{\ell\in[L_{j,n}]} x_{j,n}(\ell) \leq 1, \quad \forall j\in[M],\ n\in[m_j]
\end{aligned} \tag{20}$$

Here $\boldsymbol{x}_{j,n} = (x_{j,n}(\ell))_{\ell\in[L_{j,n}]}$ denotes the indicator vector for procuring impressions $\ell \in L_{j,n}$ in auction $n \in [m_j]$ of channel $j \in [M]$. Compared to GL-OPT, the key difference for GL-OPT$^+$ is that we introduced additional constraints which states "at most 1 impression is procured in every multi-item auction".

On the other hand, analogous to a channel's autobidding problem for the single-item auction setup in previous sections (Eq. (4)), in the multi-item setting each channel j's autobidding problem can be written as

$$\begin{aligned}
\boldsymbol{x}_j^{*,+}(\gamma_j,\rho_j;\boldsymbol{z}_j) \ = \ \arg\max_{\boldsymbol{x}=(\boldsymbol{x}_1\ldots\boldsymbol{x}_{m_j})} & \ \boldsymbol{v}_j^\top \boldsymbol{x} \\
\text{s.t.} & \ \boldsymbol{v}_j^\top \boldsymbol{x} \geq \gamma_j \boldsymbol{d}_j^\top \boldsymbol{x}, \text{ and } \quad \boldsymbol{d}_j^\top \boldsymbol{x} \leq \rho_j \\
& \ \boldsymbol{x}_n \in [0,1]^{L_{j,n}} \text{ and } \sum_{\ell\in[L_{j,n}]} x_n(\ell) \leq 1, \quad \forall n \in [m_j]
\end{aligned} \tag{21}$$

where $\boldsymbol{x}_n = (x_n(\ell))_{\ell\in[L_{j,n}]} \in [0,1]^{m_j}$ denotes the (possibly random) vector of indicators to win each impression of auction n in channel j. With respect to this per-channel multi-item auction optimization problem in Eq. (21), we can further define $V_j^+(\gamma_j,\rho_j;\boldsymbol{z}_j)$, $D_j^+(\gamma_j,\rho_j;\boldsymbol{z}_j)$, $V_j^+(\gamma_j,\rho_j)$, $D_j^+(\gamma_j,\rho_j)$ as in Eq.(5), and CH-OPT$^+(\mathcal{I})$ as in Eq.(3) for any advertiser lever option $\mathcal{I}$ in Eq.(2).

5.2. Optimizing per-channel budgets is sufficient to achieve global optimal

Our first main result for the multi-item setting is the following Theorem 5.1 which again shows an advertiser can achieve the global optimal conversion GL-OPT$^+$ via solely optimizing over per-channel budgets (analogous to Theorem 3.3 and Corollary 3.4).

Theorem 5.1 (Redundancy of per-channel ROIs in multi-slot auctions) *For the per-channel budget option $\mathcal{I}_B$ and general options $\mathcal{I}_G$ defined in Eq.(2), we have* GL-OPT$^+$ = CH-OPT$^+(\mathcal{I}_B)$ = CH-OPT$^+(\mathcal{I}_G)$ *for any aggregate ROI $\gamma > 0$ and total budget $\rho > 0$, even for*

$\rho = \infty$. *Further, there exists and optimal solution* $(\gamma_j, \rho_j)_{j\in[M]}$ *to* CH-OPT$^+(\mathcal{I}_G)$, *s.t.* $\gamma_j = 0$ *for all* $j \in [M]$.

It is easy to see that the proof of Lemma 3.1, Theorem 3.3, and Corollary 3.4 w.r.t. the single item setting in Section 3 can be directly applied to Theorem 5.1 since we did not rely on specific structures of the solutions to GL-OPT and CH-OPT other than the presence of the respective ROI and budget constraints (which are still present in GL-OPT$^+$ and CH-OPT$^+$). Thereby we will omit the proof of Theorem 5.1. In light of Theorem 5.1, we again conclude that the per-channel ROI lever is redundant, and hence omit per-channel ROI γ_j when the context is clear.

5.3. Applying UCB-SGD to the multi-item setting

We now turn to our second main focus of the multi-item setting, which is to understand whether our proposed UCB-SGD algorithm can achieve accurate approximations to the optimal per-channel budgets, similar to Theorem 4.6 for the single-item setting. A key observation is that the only difference between bounding the error of UCB-SGD in the single and multi-item settings is the structure of the conversion and corresponding Lagrangian functions (see Lemma 4.3), since the only change in the multi-item setting compared to the single-item setting is how a given per-channel budget translates into a certain conversion. Therefore, in this section we introduce a broad class of multi-item auction formats that induce the same conversion function structural properties as those illustrated in Lemma 4.3, which will allows us to directly apply the proof for bounding the error of UCB-SGD (Theorem 4.6) to the multi-item setting of interest.

To begin with, we introduce the following notion of increasing marginal values, which is a characteristic that preserves the structural properties for conversion and Lagrangian functions from the single-item setting (in Lemma 4.3), as shown later in Lemma 5.2.

Definition 5.1 (Multi-item auctions with increasing marginal values) *We say an auction* $n \in [m_j]$ *in channel* $j \in [M]$ *has increasing marginal values if for any realization* $\boldsymbol{z}_j = (\boldsymbol{v}_j, \boldsymbol{d}_j)$, *we have* $d_{j,n}(\ell) > d_{j,n}(\ell+1) > 0$ *for any* $\ell = 1 \ldots L_{j,n} - 1$ *and*

$$\frac{v_{j,n}(1) - v_{j,n}(2)}{d_{j,n}(1) - d_{j,n}(2)} > \ldots > \frac{v_{j,n}(L_{j,n}-1) - v_{j,n}(L_{j,n})}{d_{j,n}(L_{j,n}-1) - d_{j,n}(L_{j,n})} > \frac{v_{j,n}(L_{j,n})}{d_{j,n}(L_{j,n})} > 0\,.$$

Increasing marginal values intuitively means that in some multi-item auction, the marginal value per cost gained increases with procuring impressions of greater values. Many classic position auction formats satisfy increasing marginal gains, such as the Vickrey–Clarke–Groves (VCG) auction; see Varian (2007), Edelman et al. (2007) for more details on position auctions.

Example 5.1 (VCG auctions have increasing marginal values) *Let auction* $n \in [m_j]$ *in channel* $j \in [M]$ *be a VCG auction, where for any realization of* $(\boldsymbol{v}_{j,n}, \boldsymbol{d}_{j,n}) = (v_{j,n}(\ell), d_{j,n}(\ell))_{\ell\in[L_{j,n}]}$

there exists some $\widetilde{v}_{n,j} > 0$, *position discounts* $1 \geq \theta_{n,j}(1) > \theta_{n,j}(2) \ldots \theta_{n,j}(L_{n,j}) > 0$, *and* $L_{n,j}$*-highest competing bids from competitors in the market* $\widetilde{d}_{n,j}(1) > \widetilde{d}_{n,j}(2) \ldots > \widetilde{d}_{n,j}(L_{n,j}) > 0$, *such that the acquired value for procuring impression* $\ell \in L_{n,j}$ *is* $v_{n,j}(\ell) = \theta_{n,j}(\ell) \cdot \widetilde{v}_{n,j}$, *and the corresponding cost is* $d_{j,n}(\ell) = \sum_{\ell'=\ell}^{L_{j,n}} (\theta_{n,j}(\ell') - \theta_{n,j}(\ell'+1))\widetilde{d}_{n,j}(\ell')$ *where we denote* $\theta_{n,j}(L_{j,n}+1) = 0$.[8] *Thereby, under VCG the marginal values are*

$$\frac{v_{j,n}(\ell) - v_{j,n}(\ell+1)}{d_{j,n}(\ell) - d_{j,n}(\ell+1)} = \frac{(\theta_{n,j}(\ell) - \theta_{n,j}(\ell+1))\,\widetilde{v}_{j,n}}{(\theta_{n,j}(\ell) - \theta_{n,j}(\ell+1))\,\widetilde{d}_{j,n}(\ell)} = \frac{\widetilde{v}_{j,n}}{\widetilde{d}_{j,n}(\ell)}$$

which decreases in ℓ *since* $\widetilde{d}_{n,j}(1) > \widetilde{d}_{n,j}(2) \ldots > \widetilde{d}_{n,j}(L_{n,j}) > 0$. *Hence VCG auctions admit increasing marginal values.*

We remark that the generalized second price auction (GSP) does not necessarily have increasing marginal values. Now, if all auctions in a channel have increasing marginal values, then we can show the conversion function $V_j^+(\rho_j)$ and the corresponding Lagrangian function for multi-item auctions admits the same structural properties as those in Lemma 4.3:

Lemma 5.2 (Structural properties for multi-item auctions) *For any channel* $j \in [M]$ *whose auctions have increasing marginal values (see Definition 5.1), the conversion function* $V_j^+(\rho_j) = \mathbb{E}\left[\boldsymbol{v}_j^\top \boldsymbol{x}_j^{*,+}(\rho_j; \boldsymbol{z}_j)\right]$ *is continuous, piece-wise linear, strictly increasing, and concave. Here recall* $\boldsymbol{x}_j^{*,+}(\rho_j; \boldsymbol{z}_j)$ *is the optimal solution to the channel's optimization problem in Eq.* (21). *Further, for any dual variables* $\boldsymbol{c} = (\lambda, \theta) \in \mathbb{R}_+^2$, *the Lagrangian function* $\mathcal{L}_j^+(\rho_j, \boldsymbol{c}) := (1+\lambda)V_j^+(\rho_j) - (\theta + \gamma\lambda)\rho_j$ *is continuous, piece-wise linear, concave, and unimodal in* ρ_j.

See proof in Appendix C.1. In light of Lemma 5.2, we can argue that UCB-SGD produces per-channel budgets that yield the same accuracy as in Theorem 4.6 for the single-item setting,

Theorem 5.3 (UCB-SGD applied to channel procurement for multi-item auctions) *Assume multi-item auctions in any channel* $j \in [M]$ *has increasing marginal values (per Definition 5.1), and assume Assumptions 4.1 and 4.2 hold for the multi-item setting.*[9] *Then with the same parameter choices as in Theorem 4.6, and recalling* $\overline{\boldsymbol{\rho}}_T = \left(\frac{1}{T}\sum_{t \in [T]} \rho_{j,t}\right)_{j \in [M]}$ *is the vector of time-averaged per-channel budgets produced by UCB-SGD, we have*

$$\text{GL-OPT}^+ - \sum_{j \in [M]} \mathbb{E}\left[V_j^+(\overline{\rho}_{T,j})\right] \leq \mathcal{O}(T^{-1/3}),$$

[8] Here, the distribution over $(\boldsymbol{v}_{j,n}, \boldsymbol{d}_{j,n})$ can be viewed as the joint distribution over $\widetilde{v}_{n,j}$, $(\theta_{n,j}(\ell))_{\ell \in [L_{j,n}]}$ and $(\widetilde{d}_{n,j}(\ell))_{\ell \in [L_{j,n}]}$.

[9] Assumption 4.1 in the multi-item setting again implies the spend in any channel is exactly the input per-channel budget; Assumption 4.2 in the multi-item setting states that for any realization of value-cost pairs $\boldsymbol{z} = (\boldsymbol{v}_j, \boldsymbol{d}_j)_{j \in [M]} \in F_1 \times \ldots F_M$, the realized version of the ROI constraint in GL-OPT$^+$ defined in Eq. (20) is strictly feasible.

as well as approximate constraint satisfaction

$$\sum_{j\in[M]} \mathbb{E}\left[V_j^{+}(\overline{\rho}_{T,j}) - \gamma\overline{\rho}_{T,j}\right] \geq -\mathcal{O}(T^{-1/2}), \quad \textit{and } \rho - \sum_{j\in[M]} \mathbb{E}[\overline{\rho}_{T,j}] \geq -\mathcal{O}(T^{-1/2})$$

where we recall GL-OPT^{+} *is defined in Eq.* (20)*,* $V_j^{+}(\rho_j) = \mathbb{E}\left[\boldsymbol{v}_j^{\top}\boldsymbol{x}_j^{*,+}(\rho_j;\boldsymbol{z}_j)\right]$ *and* $\boldsymbol{x}_j^{*,+}(\rho_j;\boldsymbol{z}_j)$ *is defined in Eq.* (21)*.*

The proof for this theorem is identical to that of Theorem 4.6 given the same structural properties of the conversion and Lagrangian functions in Lemma 5.2 and Lemma 4.3. Hence we will omit the proof. Finally, we remark that UCB-SGD-II (Algorithm 2) can also be applied to the multi-item setting and yield per-channel budget estimates that achieve the same performance as illustrated in Theorem 4.8 while satisfying both global budget and ROI constraints exactly.

6. Additional discussions

In the following Section 6.1, we discuss extensions of our results to more general advertiser objectives; and in Section 6.2, we discuss future directions on non-optimal channel autobidding.

6.1. More general advertiser objectives

In GL-OPT and CH-OPT($\mathcal{I}$) defined Section 2 (or similarly GL-OPT^{+} and CH-OPT$^{+}(\mathcal{I})$ defined in the multi-item setting in Section 5), we can also consider more general objectives, namely $\max_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_M} \sum_{j\in[M]} \mathbb{E}\left[\boldsymbol{v}_j^{\top}\boldsymbol{x}_j - \alpha\boldsymbol{d}_j^{\top}\boldsymbol{x}_j\right]$ and $\max_{(\gamma_j,\rho_j)_{j\in[M]}\in\mathcal{I}} \sum_{j\in M} \mathbb{E}\left[V_j(\gamma_j,\rho_j;\boldsymbol{z}_j) - \alpha V_j(\gamma_j,\rho_j;\boldsymbol{z}_j)\right]$ for some private cost $\alpha \in [0,\gamma]$[10] in GL-OPT and CH-OPT($\mathcal{I}$), respectively. When $\alpha = 0$, we recover our considered models in the previous section, whereas in when $\alpha = 1$, we obtain the classic quasi-linear utility. We remark that this private cost model has been introduced and studied in related literature; see Balseiro et al. (2019b) and references therein. Nevertheless, when each channel's autobidding problem remains as is in Eq.(4), i.e. channels still aim to maximize conversion which causes a misalignment between advertiser objectives and channel behavior, it is not difficult to see in our proofs that all our results still hold in Section 3, and our UCB-SGD algorithm still produces estimates of the same order of accuracy via introducing α into the Lagrangian. In other words, even if channels aim to maximize total conversion for advertisers, advertisers can optimize for GL-OPT with a private cost α through optimizing CH-OPT($\mathcal{I}$) that also incorporates the same private cost.

[10] If $\alpha > \gamma$ the ROI constraints in GL-OPT as well as CH-OPT($\mathcal{I}$) become redundant.

6.2. Non-optimal autobidding in channels.

We recall in previous sections we assumed that each channel adopt "optimal autobidding" that solves Eq. (4) to optimality. This raises the natural question that whether our findings will still hold when channels do not procure ads optimally, perhaps because of non-stationary environments Besbes et al. (2014), Luo et al. (2018), Cheung et al. (2019), or the presence of strategic market participants who aim to manipulate the market Golrezaei et al. (2019a), Drutsa (2020), Golrezaei et al. (2021b,a). In such a scenario, an advertiser's (bandit) conversion feedback in a channel j would be $V(\gamma_j, \rho_j; \boldsymbol{z}_j) - \epsilon_j$ for some channel-specific and possibly adversarial loss $\epsilon_j > 0$. One potential resolution is to treat such ϵ_j as adversarial corruptions to bandit rewards, and instead of integrating vanilla UCB with SGD as in Algorithm 1, augment SGD with bandit algorithms that are robust to robust corruptions; see e.g. Lykouris et al. (2018), Gupta et al. (2019). Nevertheless, it remains an open question to prove how such augmentation would perform in our specific bandit-feedback constrained optimization setup. This leads to potential research directions of both practical and theoretical significance.

References

Agarwal A, Hsu D, Kale S, Langford J, Li L, Schapire R (2014) Taming the monster: A fast and simple algorithm for contextual bandits. *International Conference on Machine Learning*, 1638–1646 (PMLR).

Aggarwal G, Badanidiyuru A, Mehta A (2019) Autobidding with constraints. *International Conference on Web and Internet Economics*, 17–30 (Springer).

Agrawal R (1995) The continuum-armed bandit problem. *SIAM journal on control and optimization* 33(6):1926–1951.

Audet C, Dennis Jr JE (2004) A pattern search filter method for nonlinear programming without derivatives. *SIAM Journal on Optimization* 14(4):980–1010.

Audet C, Tribes C (2018) Mesh-based nelder–mead algorithm for inequality constrained optimization. *Computational Optimization and Applications* 71(2):331–352.

Balseiro S, Deng Y, Mao J, Mirrokni V, Zuo S (2021a) Robust auction design in the auto-bidding world. *Advances in Neural Information Processing Systems* 34:17777–17788.

Balseiro S, Golrezaei N, Mahdian M, Mirrokni V, Schneider J (2019a) Contextual bandits with cross-learning. *Advances in Neural Information Processing Systems* 32.

Balseiro S, Golrezaei N, Mirrokni V, Yazdanbod S (2019b) A black-box reduction in mechanism design with private cost of capital. *Available at SSRN 3341782* .

Balseiro S, Kim A, Mahdian M, Mirrokni V (2017) Budget management strategies in repeated auctions. *Proceedings of the 26th International Conference on World Wide Web*, 15–23.

Balseiro SR, Deng Y, Mao J, Mirrokni VS, Zuo S (2021b) The landscape of auto-bidding auctions: Value versus utility maximization. *Proceedings of the 22nd ACM Conference on Economics and Computation*, 132–133.

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

Balseiro SR, Lu H, Mirrokni V (2022) The best of many worlds: Dual mirror descent for online allocation problems. *Operations Research* .

Bertsekas DP (2014) *Constrained optimization and Lagrange multiplier methods* (Academic press).

Besbes O, Gur Y, Zeevi A (2014) Stochastic multi-armed-bandit problem with non-stationary rewards. *Advances in neural information processing systems* 27.

Braverman M, Mao J, Schneider J, Weinberg M (2018) Selling to a no-regret buyer. *Proceedings of the 2018 ACM Conference on Economics and Computation*, 523–538.

Bűrmen Á, Puhan J, Tuma T (2006) Grid restrained nelder-mead algorithm. *Computational optimization and applications* 34(3):359–375.

Cheung WC, Simchi-Levi D, Zhu R (2019) Learning to optimize under non-stationarity. *The 22nd International Conference on Artificial Intelligence and Statistics*, 1079–1087 (PMLR).

Deng Y, Golrezaei N, Jaillet P, Liang JCN, Mirrokni V (2022a) Fairness in the autobidding world with machine-learned advice. *arXiv preprint arXiv:2209.04748* .

Deng Y, Mao J, Mirrokni V, Zhang H, Zuo S (2022b) Efficiency of the first-price auction in the autobidding world. *arXiv preprint arXiv:2208.10650* .

Deng Y, Mao J, Mirrokni V, Zuo S (2021) Towards efficient auctions in an auto-bidding world. *Proceedings of the Web Conference 2021*, 3965–3973.

Deng Y, Schneider J, Sivan B (2019) Prior-free dynamic auctions with low regret buyers. *Advances in Neural Information Processing Systems* 32.

Ding Q, Hsieh CJ, Sharpnack J (2021) An efficient algorithm for generalized linear bandit: Online stochastic gradient descent and thompson sampling. *International Conference on Artificial Intelligence and Statistics*, 1585–1593 (PMLR).

Drutsa A (2020) Optimal non-parametric learning in repeated contextual auctions with strategic buyer. *International Conference on Machine Learning*, 2668–2677 (PMLR).

Dzahini KJ, Kokkolaras M, Le Digabel S (2022) Constrained stochastic blackbox optimization using a progressive barrier and probabilistic estimates. *Mathematical Programming* 1–58.

Edelman B, Ostrovsky M, Schwarz M (2007) Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *American economic review* 97(1):242–259.

Fasano G, Liuzzi G, Lucidi S, Rinaldi F (2014) A linesearch-based derivative-free approach for nonsmooth constrained optimization. *SIAM journal on optimization* 24(3):959–992.

Feng Z, Padmanabhan S, Wang D (2022) Online bidding algorithms for return-on-spend constrained advertisers. *arXiv preprint arXiv:2208.13713* .

Flaxman AD, Kalai AT, McMahan HB (2004) Online convex optimization in the bandit setting: gradient descent without a gradient. *arXiv preprint cs/0408007* .

Golrezaei N, Jaillet P, Liang JCN (2019a) Incentive-aware contextual pricing with non-parametric market noise. *arXiv preprint arXiv:1911.03508* .

Golrezaei N, Jaillet P, Liang JCN, Mirrokni V (2021a) Bidding and pricing in budget and roi constrained markets. *arXiv preprint arXiv:2107.07725* .

Golrezaei N, Javanmard A, Mirrokni V (2019b) Dynamic incentive-aware learning: Robust pricing in contextual auctions. *Advances in Neural Information Processing Systems* 32.

Golrezaei N, Manshadi V, Schneider J, Sekar S (2021b) Learning product rankings robust to fake users. *Proceedings of the 22nd ACM Conference on Economics and Computation*, 560–561.

Gupta A, Koren T, Talwar K (2019) Better algorithms for stochastic bandits with adversarial corruptions. *Conference on Learning Theory*, 1562–1578 (PMLR).

Han Y, Liang Z, Wang Y, Zhang J (2021) Generalized linear bandits with local differential privacy. *Advances in Neural Information Processing Systems* 34:26511–26522.

Kannan P, Reinartz W, Verhoef PC (2016) The path to purchase and attribution modeling: Introduction to special section.

Larson J, Menickelly M, Wild SM (2019) Derivative-free optimization methods. *Acta Numerica* 28:287–404.

Luo H, Wei CY, Agarwal A, Langford J (2018) Efficient contextual bandits in non-stationary worlds. *Conference On Learning Theory*, 1739–1776 (PMLR).

Lykouris T, Mirrokni V, Paes Leme R (2018) Stochastic bandits robust to adversarial corruptions. *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, 114–122.

Mehta A (2022) Auction design in an auto-bidding setting: Randomization improves efficiency beyond vcg. *Proceedings of the ACM Web Conference 2022*, 173–181.

Nguyen A, Balasubramanian K (2022) Stochastic zeroth-order functional constrained optimization: Oracle complexity and applications. *INFORMS Journal on Optimization* .

Pourmohamad T, Lee HK (2020) The statistical filter approach to constrained optimization. *Technometrics* 62(3):303–312.

Slivkins A, et al. (2019) Introduction to multi-armed bandits. *Foundations and Trends® in Machine Learning* 12(1-2):1–286.

Usmanova I, Krause A, Kamgarpour M (2019) Safe convex learning under uncertain constraints. *The 22nd International Conference on Artificial Intelligence and Statistics*, 2106–2114 (PMLR).

Varian HR (2007) Position auctions. *international Journal of industrial Organization* 25(6):1163–1178.

Appendices for

Multi-channel Autobidding with Budget and ROI constraints

Appendix A: Proofs for Section 3

A.1. Proof of Lemma 3.1

Fix any option $\mathcal{I} \in \{\mathcal{I}_B, \mathcal{I}_R, \mathcal{I}_G\}$ defined in Eq. (2), and let $(\widetilde{\boldsymbol{\gamma}}, \widetilde{\boldsymbol{\rho}}) \in \mathcal{I}$ be the optimal solution to CH-OPT($\mathcal{I}$). Note that for the per-channel ROI only option $\mathcal{I}_R$, we have $\widetilde{\rho}_j = \infty$ and for the per-channel budget only we have $\widetilde{\gamma}_j = 0$ for all $j \in [M]$. Further, for any realization of value-cost pairs over all auctions $\boldsymbol{z} = (\boldsymbol{v}_j, \boldsymbol{d}_j)_{j\in[M]}$, recall the optimal solution $\boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)$ to $V_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)$ for each channel $j \in [M]$ as defined in Eq. (4).

Due to feasibility of $(\widetilde{\boldsymbol{\gamma}}, \widetilde{\boldsymbol{\rho}}) \in \mathcal{I}$ for CH-OPT($\mathcal{I}$), we have

$$\sum_{j\in M} \mathbb{E}\left[V_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)\right] \geq \gamma \sum_{j\in M} \mathbb{E}\left[D_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)\right] \Longrightarrow \sum_{j\in[M]} \mathbb{E}\left[\boldsymbol{v}_j^\top \boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)\right] \geq \gamma \sum_{j\in[M]} \left[\boldsymbol{d}_j^\top \boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)\right]$$

where we used the definitions $V_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j) = \boldsymbol{v}_j^\top \boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)$ and $D_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j) = \boldsymbol{d}_j^\top \boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)$ in Eq. (5). This implies $\left(\boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)\right)_{j\in[M]}$ satisfies the ROI constraint in GL-OPT. A similar analysis implies $\left(\boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)\right)_{j\in[M]}$ also satisfies the budget constraint in GL-OPT. Therefore, $\left(\boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)\right)_{j\in[M]}$ is feasible to GL-OPT. So

$$\text{GL-OPT} \geq \sum_{j\in[M]} \mathbb{E}\left[\boldsymbol{v}_j^\top \boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)\right] = \sum_{j\in M} \left[V_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)\right] = \text{CH-OPT}(\mathcal{I})\,. \tag{22}$$

where the final equality follows from the assumption that $(\widetilde{\boldsymbol{\gamma}}, \widetilde{\boldsymbol{\rho}}) \in \mathcal{I}$ is the optimal solution to CH-OPT($\mathcal{I}$). □

A.2. Proof of Theorem 3.3

In light of Lemma 3.1, we only need to show CH-OPT($\mathcal{I}_B$) ≥ GL-OPT. Let $\widetilde{\boldsymbol{x}}(\boldsymbol{z}) = \{\widetilde{\boldsymbol{x}}_j(\boldsymbol{z}_j))\}_{j\in[N]}$ be the optimal solution to GL-OPT, and define $\widetilde{\gamma}_j = 0$ and $\widetilde{\rho}_j = \mathbb{E}\left[\boldsymbol{d}_j^\top \widetilde{\boldsymbol{x}}_j(\boldsymbol{z}_j))\right]$ to be the corresponding expected spend for each channel j under the optimal solution $\widetilde{\boldsymbol{x}}(\boldsymbol{z})$ to GL-OPT, respectively.

We first argue that $(\widetilde{\gamma}_j, \widetilde{\rho}_j)_{j\in[M]}$ is feasible to CH-OPT($\mathcal{I}_B$). Recall the optimal solution $\boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)$ to $V_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)$ for each channel $j \in [M]$ as defined in Eq. (4), as well as the definitions $V_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j) = \boldsymbol{v}_j^\top \boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)$ and $D_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j) = \boldsymbol{d}_j^\top \boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)$ in Eq. (5). Then, we have

$$\mathbb{E}\left[D_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)\right] \;=\; \mathbb{E}\left[\boldsymbol{d}_j^\top \boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j); \boldsymbol{z}_j\right] \;\stackrel{(i)}{\leq}\; \widetilde{\rho}_j \;=\; \mathbb{E}\left[\boldsymbol{d}_j^\top \widetilde{\boldsymbol{x}}_j(\boldsymbol{z}_j)\right]\,, \tag{23}$$

where (i) follows from feasibility of $\boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)$ to $V_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)$. Summing over $j \in [M]$ we conclude that $(\boldsymbol{\gamma}_j, \boldsymbol{\rho}_j)_{j\in[M]}$ satisfies the budget constraint in CH-OPT($\mathcal{I}_B$):

$$\sum_{j\in[M]} \mathbb{E}\left[D_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)\right] \leq \sum_{j\in[M]} \mathbb{E}\left[\boldsymbol{d}_j^\top \widetilde{\boldsymbol{x}}_j(\boldsymbol{z}_j)\right] \stackrel{(i)}{\leq} \rho\,. \tag{24}$$

Here (i) follows from feasibility of $\widetilde{\boldsymbol{x}}(\boldsymbol{z}) = \{\widetilde{\boldsymbol{x}}_j(\boldsymbol{z}_j))\}_{j\in[N]}$ to GL-OPT since it is the optimal solution.

On the other hand, we have

$$V_j(\widetilde{\gamma}_j,\widetilde{\rho}_j;\boldsymbol{z}_j) \;=\; \boldsymbol{v}_j^\top \boldsymbol{x}_j^*(\widetilde{\gamma}_j,\widetilde{\rho}_j;\boldsymbol{z}_j) \;\overset{(i)}{\geq}\; \boldsymbol{v}_j^\top \widetilde{\boldsymbol{x}}_j(\boldsymbol{z}_j) \tag{25}$$

where (i) follows from optimality of $\boldsymbol{x}_j^*(\widetilde{\gamma}_j,\widetilde{\rho}_j;\boldsymbol{z}_j)$ to $V_j(\widetilde{\gamma}_j,\widetilde{\rho}_j;\boldsymbol{z}_j)$. Hence, we have

$$\sum_{j\in M}\mathbb{E}\left[V_j(\widetilde{\gamma}_j,\widetilde{\rho}_j;\boldsymbol{z}_j)\right] \;\geq\; \sum_{j\in M}\mathbb{E}\left[\boldsymbol{v}_j^\top \widetilde{\boldsymbol{x}}_j(\boldsymbol{z}_j)\right] \;\overset{(i)}{\geq}\; \gamma\sum_{j\in M}\mathbb{E}\left[\boldsymbol{d}_j^\top \widetilde{\boldsymbol{x}}_j(\boldsymbol{z}_j)\right] \;\overset{(ii)}{\geq}\; \gamma\sum_{j\in[M]}\mathbb{E}\left[D_j(\widetilde{\gamma}_j,\widetilde{\rho}_j;\boldsymbol{z}_j)\right] \tag{26}$$

where (i) follows from feasibility of $\widetilde{\boldsymbol{x}}(\boldsymbol{z}) = \{\widetilde{\boldsymbol{x}}_j(\boldsymbol{z}_j))\}_{j\in[N]}$ to GL-OPT since it is the optimal solution; (ii) follows from Eq. (23). Hence combining Eq. (24) (26) we can conclude that $(\widetilde{\gamma}_j,\widetilde{\rho}_j)_{j\in[M]}$ is feasible to CH-OPT($\mathcal{I}_B$).

Finally, we have CH-OPT$(\mathcal{I}_B) \geq \sum_{j\in M}\mathbb{E}\left[V_j(\widetilde{\gamma}_j,\widetilde{\rho}_j;\boldsymbol{z}_j)\right] \geq \sum_{j\in M}\mathbb{E}\left[\boldsymbol{v}_j^\top \widetilde{\boldsymbol{x}}_j(\boldsymbol{z}_j)\right] =$ GL-OPT, where the last inequality follows from (26), and the final equality is because we assumed $\widetilde{\boldsymbol{x}}(\boldsymbol{z}) = \{\widetilde{\boldsymbol{x}}_j(\boldsymbol{z}_j))\}_{j\in[N]}$ is the optimal solution to GL-OPT. □

A.3. Proof of Corollary 3.4

In light of Lemma 3.1, we only need to show CH-OPT$(\mathcal{I}_G) \geq$ GL-OPT. Let $(\widetilde{\boldsymbol{\gamma}},\widetilde{\boldsymbol{\rho}}) \in \mathcal{I}_B$, and by definition of $\mathcal{I}_B$ in Eq. (2) we have $\widetilde{\gamma}_j = 0$ for all $j\in[M]$. Since $(\widetilde{\boldsymbol{\gamma}},\widetilde{\boldsymbol{\rho}})$ is feasible to CH-OPT($\mathcal{I}_B$), it is also feasible to CH-OPT($\mathcal{I}_G$) since these two problems share the same ROI and budget constraints. Because they also share the same objectives, we have

$$\text{CH-OPT}(\mathcal{I}_G) \geq \text{CH-OPT}(\mathcal{I}_B) = \text{GL-OPT} \tag{27}$$

where the final equality follows from Theorem 3.3. □

Appendix B: Proofs for Section 4

B.1. Proof of Proposition 4.1

Let $(\rho_j^*)_{j\in[M]}$ be the optimal per-channel budgets to CH-OPT($\mathcal{I}_B$), and define $\bar{\mu}_T = \frac{1}{T}\sum_{t\in[T]}\mu_t$ as well as $\bar{\lambda}_T = \frac{1}{T}\sum_{t\in[T]}\lambda_t$. Then

$$\begin{aligned}
& T\cdot\text{GL-OPT} - \mathbb{E}\left[\sum_{t\in[T]}\sum_{j\in[M]}V_j(\rho_{j,t})\right] \\
\overset{(i)}{\leq}\;& M\bar{V}K + (T-K)\cdot\text{CH-OPT}(\mathcal{I}_B) - \mathbb{E}\left[\sum_{t>K}\sum_{j\in[M]}V_j(\rho_{j,t})\right] \\
\leq\;& M\bar{V}K + (T-K)\cdot\mathbb{E}\left[\mathcal{L}_j(\rho_j^*,\bar{\lambda}_T,\bar{\mu}_T) + \rho\bar{\mu}_T\right] - \mathbb{E}\left[\sum_{t>K}\sum_{j\in[M]}V_j(\rho_{j,t})\right] \\
\overset{(ii)}{\leq}\;& M\bar{V}K + \rho\sum_{t>K}\mathbb{E}\left[\mu_t\right] + \sum_{t>K}\sum_{j\in[M]}\mathbb{E}\left[\mathcal{L}_j(\rho_j^*,\lambda_t,\mu_t)\right] - \sum_{t>K}\sum_{j\in[M]}\mathbb{E}\left[\mathcal{L}_j(\rho_{j,t},\boldsymbol{c}_t) - \lambda_t\left(V_j(\rho_{j,t}) - \gamma\rho_{j,t}\right) + \mu_t\rho_{j,t}\right] \\
\overset{(iii)}{\leq}\;& M\bar{V}K + \sum_{j\in[M]}\sum_{t>K}\mathbb{E}\left[\mathcal{L}_j(\rho_j^*(t),\boldsymbol{c}_t) - \mathcal{L}_j(\rho_{j,t},\boldsymbol{c}_t)\right] + \sum_{t>K}\left(\lambda_t g_{1,t} + \mu_t g_{2,t}\right)\,.
\end{aligned} \tag{28}$$

Here, (i) follows from Theorem 3.3 that states GL-OPT = CH-OPT($\mathcal{I}_B$); (ii) follows from the CH-OPT$(\mathcal{I}_B) = \sum_{j\in[M]}V_j(\rho_j^*)$ and the definition of the Lagrangian in Eq. (7); in (ii) we define $\rho_j^*(t) = \arg\max_{\rho_j\geq 0}\mathcal{L}_j(\rho_j,\boldsymbol{c}_t)$ to be the optimal budget that maximizes the Lagrangian w.r.t. the dual variables $\boldsymbol{c}_t = (\lambda_t,\mu_t)$. □

B.2. Proof for Lemma 4.2

Recall $g_{1,t} = \sum_{j\in[M]} (V_{j,t}(\rho_{j,t}) - \gamma\rho_{j,t})$ and $g_{2,t} = \rho - \sum_{j\in[M]} \rho_{j,t}$ defined in Algorithm 1.

Now from lemma B.3, we have for any $\lambda \geq 0$ and $\mu \geq 0$,

$$\sum_t (\lambda_t - \lambda)\, g_{1,t} \leq \frac{\eta M^2 \bar{V}^2}{2} \cdot T + \frac{1}{2\eta}(\lambda - \lambda_1)^2$$
$$\sum_t (\mu_t - \mu)\, g_{2,t} \leq \frac{\eta\rho^2}{2} \cdot T + \frac{1}{2\eta}(\mu - \mu_1)^2 .$$

By taking $\lambda = \mu = 0$ and recalling $\lambda_1 = \mu_1 = 0$, we get the desired bound in the statement of the lemma. □

B.3. Proof of Lemma 4.3

We first show for any realization $\boldsymbol{z} = (\boldsymbol{z}_j)_{j\in[M]} = (\boldsymbol{v}_j, \boldsymbol{d}_j)_{j\in[M]}$, the conversion function $V_j(\rho_j; \boldsymbol{z}_j)$ is piecewise linear, strictly inreasing, and concave for any $j \in [M]$.

Fix any channel j which consists of m_j parallel auctions, and recall that we assumed the orderding $\frac{v_{j,1}}{d_{j,1}} > \frac{v_{j,2}}{d_{j,2}} > \cdots > \frac{v_{j,m_j}}{d_{j,m_j}}$ for any realization $\boldsymbol{z}_j$. Then, with the option where the per-channel ROI is set to 0 (i.e. omitted) $V_j(\rho_j; \boldsymbol{z}_j)$ is exactly the LP relaxation of a 0-1 knapsack, whose optimal solution $\boldsymbol{x}_j^*(\rho_j; \boldsymbol{z}_j)$ is well known to be unique, and takes the form for any auction index $n \in [m_j]$:

$$x^*_{j,n}(\rho_j; \boldsymbol{z}_j) = \begin{cases} 1 & \text{if } \sum_{n'\in[n]} d_{j,n'} \leq \rho_j \\ \frac{\rho_j - \sum_{n'\in[n-1]} d_{j,n'}}{d_{j,n}} & \text{if } \sum_{n'\in[n]} d_{j,n'} > \rho_j \\ 0 & \text{otherwise} \end{cases} \tag{29}$$

where we denote $d_{j,0} = 0$. With this form, it is easy to see

$$V_j(\rho_j; \boldsymbol{z}_j) = \boldsymbol{v}_j^\top \boldsymbol{x}_j^*(\rho_j; \boldsymbol{z}_j) = \sum_{n\in[m_j]} \left(\frac{v_{j,n}}{d_{j,n}}\rho_j + b_{j,n}\right) \mathbb{I}\{d_{j,0} + \cdots + d_{j,n-1} \leq \rho_j \leq d_{j,0} + \cdots + d_{j,n}\} \tag{30}$$

where we denote $d_{j,0} = 0$ and also $b_{j,n} = \sum_{n'\in[n-1]} v_{j,n'} - \frac{v_{j,n}}{d_{j,n}} \cdot \left(\sum_{n'\in[n-1]} d_{j,n'}\right)$ and $v_{j,0} = 0$. It is easy to check that any two line segments, say $[X_{n-1}, X_n]$ and $[X_n, X_{n+1}]$ where we write $X_n = d_{j,0} + \cdots + d_{j,n}$, intersect at $\rho_j = X_n$, because $\frac{v_{j,n}}{d_{j,n}}\rho_j + b_{j,n} = \frac{v_{j,n+1}}{d_{j,n+1}}\rho_j + b_{j,n+1}$ at $\rho_j = X_n$. Hence, from Eq. (30) we can conclude $V_j(\rho_j; \boldsymbol{z}_j)$ is continuous, which further implies it is piecewise linear and strictly increasing. Further, the ordering $\frac{v_{j,1}}{d_{j,1}} > \frac{v_{j,2}}{d_{j,2}} > \cdots > \frac{v_{j,m_j}}{d_{j,m_j}}$ implies that the slopes on each segment $[X_n, X_{n+1}]$ decreases as n increases, which implies $V_j(\rho_j; \boldsymbol{z}_j)$ is concave.

Since $V_j(\rho_j) = \mathbb{E}[V_j(\rho_j; \boldsymbol{z}_j)]$, where the expectation is taken w.r.t. randomness in $\boldsymbol{z}_j$, and since the $\boldsymbol{z}_j$ is sampled from some discrete distribution $\boldsymbol{p}_j$ on finite support F_j, $V_j(\rho_j)$ is simply a weighted average over all $(V_j(\rho_j; \boldsymbol{z}_j))_{\boldsymbol{z}_j\in F_j}$ with weights in $\boldsymbol{p}_j$, so $V_j(\rho_j)$ is also continuous, piecewise linear, strictly increasing, and concave, and thus can be written as in Lemma 4.3 with parameters $\{(s_{j,n}, b_{j,n}, r_{j,n})\}_{n\in[S_j]}$ that only depend on the support F_j and distribution $\boldsymbol{p}_j$.

Finally, according to the definition of $\mathcal{L}_j(\rho_j, \boldsymbol{c}) = \mathbb{E}[\mathcal{L}_j(\rho_j, \boldsymbol{c}; \boldsymbol{z}_j)]$ and $\mathcal{L}_j(\rho_j, \boldsymbol{c}; \boldsymbol{z}_j) = (1+\lambda)V_j(\rho_j; \boldsymbol{z}_j) - (\lambda\gamma + \mu)\rho_j$ as defined in Eq. (7), we have

$$\mathcal{L}_j(\rho_j, \boldsymbol{c}) = (1+\lambda)V_j(\rho_j) - (\lambda\gamma + \mu)\rho_j \tag{31}$$

which implies $\mathcal{L}_j(\rho_j, \boldsymbol{c})$ is continuous, piecewise linear, and concave because $V_j(\rho_j)$ is continuous, piecewise linear, and concave as shown above. Combining Eq. (31) and the representation of $V_j(\rho_j)$ in Lemma (4.3), we have

$$\mathcal{L}_j(\rho_j, \boldsymbol{c}) = \sum_{n\in[S_j]} (\sigma_{j,n}(\boldsymbol{c})\rho_j + (1+\lambda)b_{j,n})\,\mathbb{I}\{r_{j,n-1} \le \rho_j \le r_{j,n}\}\,. \tag{32}$$

where the slope $\sigma_{j,n}(\boldsymbol{c}) = (1+\lambda)sj,n - (\mu+\gamma\lambda)$ decreases in n. Thus at the point $r_{j,n^*} = \max\{r_{j,n} : n = 0, 1 \ldots, S_j\ , \sigma_{j,n}(\boldsymbol{c}) \ge 0\}$ in which the slope to the right turns negative for the first time, $\mathcal{L}_j(\rho_j, \boldsymbol{c})$ takes its maximum value $\max_{\rho_j\ge 0} \mathcal{L}_j(\rho_j, \boldsymbol{c})$, because to the left of r_{j,n^*}, namely the region $[0, r_{j,n^*}]$, $\mathcal{L}_j(\rho_j, \boldsymbol{c})$ strictly increases because slopes are positive; and to the right of r_{j,n^*}, namely the region $[r_{j,n^*}, \rho]$, $\mathcal{L}_j(\rho_j, \boldsymbol{c})$ strictly decreases because slopes are negative. $\square$

B.4. Proof for Lemma 4.4

We first present some definitions for convenience: denote $\boldsymbol{c} = (\lambda, \mu)$, $\boldsymbol{c}_t = (\lambda_t, \mu_t)$, and $\boldsymbol{z}_t = (\boldsymbol{v}_t, \boldsymbol{d}_t)$. For any realization $\boldsymbol{z} = (\boldsymbol{v}, \boldsymbol{d}) = (\boldsymbol{v}_j, \boldsymbol{d}_j)_{j\in[M]} \in F_1 \times \ldots F_M$ of values and costs across channels, define the dual function

$$\begin{aligned}\mathcal{L}_d(\boldsymbol{c}; \boldsymbol{z}) \ &= \ \max_{\boldsymbol{\rho}\in[0,\rho]^M} \sum_{j\in[M]} V_j(\rho_j; \boldsymbol{z}) + \lambda \sum_{j\in[M]} (V_j(\rho_j; \boldsymbol{z}) - \gamma\rho_j) + \mu\left(\rho - \sum_{j\in[M]} \rho_j\right) \\ &= \ \mu\rho + \max_{\boldsymbol{\rho}\in[0,\rho]^M} \mathcal{L}(\boldsymbol{\rho}, \boldsymbol{c}; \boldsymbol{z})\,.\end{aligned} \tag{33}$$

From standard convex analysis we know that $\mathcal{L}_d$ is convex in (λ, μ), which implies for any $\boldsymbol{c}' = (\lambda', \mu')$

$$(\boldsymbol{c}' - \boldsymbol{c})^\top \nabla\mathcal{L}_d(\boldsymbol{c}; \boldsymbol{z}) \le \mathcal{L}_d(\boldsymbol{c}'; \boldsymbol{z}) - \mathcal{L}_d(\boldsymbol{c}; \boldsymbol{z}) \tag{34}$$

We prove the lemma by induction that

$$\lambda_t, \mu_t \le C_F := 1 + \frac{\max_{\boldsymbol{z}\in F_1\times\ldots F_M} \boldsymbol{e}^\top \boldsymbol{v} + 1}{\min_{\boldsymbol{z}\in F_1\times\ldots F_M}\left\{\sum_{j\in[M]}\left(V_j(\rho_{(\boldsymbol{z}),j}; \boldsymbol{z}) - \gamma\rho_j\right), \rho - \sum_{j\in[M]} \rho_{(\boldsymbol{z}),j}\right\}} \tag{35}$$

Here, in Proposition B.1 we defined $\boldsymbol{\rho}_{(\boldsymbol{z})}$ for any $\boldsymbol{z}$ such that Slater's condition holds, i.e. $\sum_{j\in[M]}\left(V_j(\rho_{(\boldsymbol{z}),j}; \boldsymbol{z}) - \gamma\rho_j\right) > 0$ and $\rho - \sum_{j\in[M]} \rho_{(\boldsymbol{z}),j} > 0$.

The base case for $t = 1$ is satisfied trivially since we take $\boldsymbol{c}_t = \boldsymbol{0}$. Now assume $\|\boldsymbol{c}_t\| \le C_F$, and we will show $\|\boldsymbol{c}_{t+1}\| \le C_F$ by considering two different case:

Case (A): If $\mathcal{L}_d(\boldsymbol{c}_t; \boldsymbol{z}_t) > \mathcal{L}_d(\boldsymbol{0}; \boldsymbol{z}_t) + \eta M^2 \bar{V}^2$.

Define $\boldsymbol{g}_t = (g_{1,t}, g_{2,t})$ where we recall $g_{1,t} = \sum_{j\in[M]} (V_{j,t}(\rho_{j,t}) - \gamma\rho_{j,t})$ and $g_{2,t} = \rho - \sum_{j\in[M]} \rho_{j,t}$. From Eq. (72) in the proof of Lemma B.3, we have for any $\lambda \ge 0$ and $\mu \ge 0$,

$$\begin{aligned}(\lambda - \lambda_{t+1})^2 \ &\le \ (\lambda - \lambda_t)^2 + 2\eta\,(\lambda - \lambda_t)\,g_{1,t} + \eta^2 M^2 \bar{V}^2 \\ (\mu - \mu_{t+1})^2 \ &\le \ (\mu - \mu_t)^2 + 2\eta\,(\mu - \mu_t)\,g_{1,t} + \eta^2 M^2 \bar{V}^2\,.\end{aligned}$$

By summing the two inequalities and taking $\lambda = \mu = 0$ we get

$$\begin{aligned}\|\boldsymbol{c}_{t+1}\|^2 \ &\le \ \|\boldsymbol{c}_t\|^2 + 2\eta\,(\boldsymbol{c}_t)^\top \boldsymbol{g}_t + 2\eta^2 M^2 \bar{V}^2 \\ &\overset{(i)}{\le} \ \|\boldsymbol{c}_t\|^2 + 2\eta\,(\mathcal{L}_d(\boldsymbol{0}; \boldsymbol{z}_t) - \mathcal{L}_d(\boldsymbol{c}_t; \boldsymbol{z}_t)) + 2\eta^2 M^2 \bar{V}^2 \quad ,\end{aligned} \tag{36}$$

where in (i) we applied Eq. (34) with $\boldsymbol{z} = \boldsymbol{z}_t$ and used the fact that $\boldsymbol{g}_t = (g_{1,t}, g_{r,t}) = \nabla\mathcal{L}_d(\boldsymbol{c}_t; \boldsymbol{z}_t)$.

Then from Eq. (36) and the assumption $\mathcal{L}_d(\boldsymbol{c}_t;\boldsymbol{z}_t) > \mathcal{L}_d(\boldsymbol{0};\boldsymbol{z}_t) + \eta M^2\bar{V}^2$ we have

$$\max\{\lambda_t,\mu_t\} \le \|\boldsymbol{c}_{t+1}\| < \|\boldsymbol{c}_t\| \le C_F$$

where the final inequality follows from the induction hypothesis.

Case (B): If $\mathcal{L}_d(\boldsymbol{c}_t;\boldsymbol{z}_t) \le \mathcal{L}_d(\boldsymbol{0};\boldsymbol{z}_t) + \eta M^2\bar{V}^2$, then

$$\begin{aligned}
&\mathcal{L}_d(\boldsymbol{0};\boldsymbol{z}_t) + \eta M^2\bar{V}^2 \\
\ge\ &\mathcal{L}_d(\boldsymbol{c}_t;\boldsymbol{z}_t) \\
\overset{(i)}{=}\ &\mu_t\rho + \max_{\boldsymbol{\rho}\in[0,\rho]^M} \mathcal{L}(\boldsymbol{\rho},\boldsymbol{c}_t;\boldsymbol{z}_t) \\
\overset{(ii)}{\ge}\ &\mu_t\rho + \mathcal{L}(\boldsymbol{\rho}_{(\boldsymbol{z}_t)},\boldsymbol{c}_t;\boldsymbol{z}_t) \\
=\ &\sum_{j\in[M]} V_j(\rho_{(\boldsymbol{z}_t),j};\boldsymbol{z}_t) + \lambda_t \sum_{j\in[M]} \big(V_j(\rho_{(\boldsymbol{z}_t),j};\boldsymbol{z}_t) - \gamma\rho_j\big) + \mu_t\left(\rho - \sum_{j\in[M]}\rho_{(\boldsymbol{z}_t),j}\right) \\
\ge\ &\lambda_t \sum_{j\in[M]} \big(V_j(\rho_{(\boldsymbol{z}_t),j};\boldsymbol{z}_t) - \gamma\rho_j\big) + \mu_t\left(\rho - \sum_{j\in[M]}\rho_{(\boldsymbol{z}_t),j}\right) \\
\overset{(iii)}{\ge}\ &(\lambda_t+\mu_t)\cdot \min_{\boldsymbol{z}\in F_1\times\ldots F_M} \left\{\sum_{j\in[M]} \big(V_j(\rho_{(\boldsymbol{z}),j};\boldsymbol{z}) - \gamma\rho_j\big), \rho - \sum_{j\in[M]}\rho_{(\boldsymbol{z}),j}\right\}.
\end{aligned} \tag{37}$$

Here, (i) follows from the definition of the dual function in Eq. (33); in (ii) we recall the definition of $\boldsymbol{\rho}_{(\boldsymbol{z})}$ that satisfies the Slater's condition in Lemma B.1; in (iii), we used the fact that $\lambda_t,\mu_t \ge 0$ and also under Slater's condition we have for any $\boldsymbol{z}$, $\sum_{j\in[M]}\big(V_j(\rho_{(\boldsymbol{z}),j};\boldsymbol{z}) - \gamma\rho_j\big) > 0$ and $\rho - \sum_{j\in[M]}\rho_{(\boldsymbol{z}),j} > 0$.

On the other hand, we have $\eta M^2\bar{V}^2 < 1$ and also $\mathcal{L}_d(\boldsymbol{0};\boldsymbol{z}_t) = \max_{\boldsymbol{\rho}\in[0,\rho]^M}\sum_{j\in[M]} V_j(\rho_j;\boldsymbol{z}_t) = \max_{\boldsymbol{z}\in F_1\times\ldots F_M} \boldsymbol{e}^\top\boldsymbol{v}$. Hence combining this with Eq. (37) we get

$$\lambda_t + \mu_t < \frac{\max_{\boldsymbol{z}\in F_1\times\ldots F_M}\boldsymbol{e}^\top\boldsymbol{v} + 1}{\min_{\boldsymbol{z}\in F_1\times\ldots F_M}\left\{\sum_{j\in[M]}\big(V_j(\rho_{(\boldsymbol{z}),j};\boldsymbol{z}) - \gamma\rho_j\big), \rho - \sum_{j\in[M]}\rho_{(\boldsymbol{z}),j}\right\}} := C_F - 1 \tag{38}$$

which further implies $\lambda_t,\mu_t \le C_F - 1$. Hence, we can finally conclude

$$\lambda_{t+1} = (\lambda_t - \eta g_{1,t})_+ \le \lambda_t + \eta|g_{1,t}| \le \|\boldsymbol{c}_t\| - 1 + \eta M\bar{V} < C_F \tag{39}$$

where in the final inequality we used the fact that $\eta M\bar{V} < 1$. A similar bound holds for μ_{t+1}. □

B.5. Proof for Lemma 4.5

For simplicity, for any context $\boldsymbol{c}$, assuming the optimal budget $\arg\max_{\rho_j\in[0,\rho]}\mathcal{L}_j(\rho_j,\boldsymbol{c})$ is located at the nth "turning point" $r_{j,n}$ we have

$$\begin{aligned}
&\sigma_j^-(\boldsymbol{c}) = \sigma_{j,n}(\boldsymbol{c}) \quad\text{and}\quad \sigma_j^+(\boldsymbol{c}) = \sigma_{j,n}(\boldsymbol{c}) \\
&\mathcal{U}_j^-(\boldsymbol{c}) = [r_{j,n-1}, r_{j,n}] \quad\text{and}\quad \mathcal{U}_j^+(\boldsymbol{c}) = [r_{j,n}, r_{j,n+1}].
\end{aligned} \tag{40}$$

Further, define the following

$$\begin{aligned}
&\Delta_k(\boldsymbol{c}) = \max_{\rho_j\in[0,\rho]}\mathcal{L}_j(\rho_j,\boldsymbol{c}) - \mathcal{L}_j(a_k,\boldsymbol{c}) \\
&\mathcal{C}(\underline{\sigma}) = \left\{\boldsymbol{c}\in\{\boldsymbol{c}_t\}_{t\in[T]} : \sigma_j^-(\boldsymbol{c}) > \underline{\sigma},\ |\sigma_j^+(\boldsymbol{c})| > \underline{\sigma}\right\} \text{ for } n = 0\ldots S_j \\
&\mathcal{C}_n = \left\{\boldsymbol{c}\in\{\boldsymbol{c}_t\}_{t\in[T]} : r_{j,n} = \arg\max_{\rho_j\ge 0}\mathcal{L}_j(\rho_j,\boldsymbol{c})\right\} \text{ for } n = 0\ldots S_j \\
&m_k(\boldsymbol{c}) = \frac{8\log(T)}{\Delta_k^2(\boldsymbol{c})} \text{ for } \forall(k,\boldsymbol{c}) \text{ s.t. } \Delta_k(\boldsymbol{c}) > 0.
\end{aligned} \tag{41}$$

where S_j and $\{r_{j,n}\}_{j\in[S_j]}$ are defined in Lemma 4.3. The set $\mathcal{C}_{k,n}$ represents all context $\boldsymbol{c}_t$ under which the optimal $\arg\max_{\rho_j\geq 0}\mathcal{L}_j(\rho_j,\boldsymbol{c}_t)$ is taken at the nth "turning point" $r_{j,n}$ of $\mathcal{L}_j(\rho_j,\boldsymbol{c}_t)$ which is piecewise linear in ρ_j and also the two slopes adjacent to $r_{j,n}$, namely $\sigma_{j,n}(\boldsymbol{c})$ and $\sigma_{j,n+1}(\boldsymbol{c})$ satisfy $\sigma_{j,n}(\boldsymbol{c})>\underline{\sigma}$ and $|\sigma_{j,n+1}(\boldsymbol{c})|>\underline{\sigma}$. Further it is easy to see that $\mathcal{C}=\cup_{n=0}^{S_j}\mathcal{C}_n$.

Further, define $N_{k,t}=\sum_{\tau\leq t-1}\mathbb{I}\{\rho_{j,\tau}=a_k\}$ to be the number of times arm k is pulled up to time t. we get

$$\begin{aligned}
&\sum_{t\in[T]}\mathcal{L}_j(\rho^*_{j,t},\boldsymbol{c}_t)-\mathcal{L}_j(\rho_{j,t},\boldsymbol{c}_t)=X_1+X_2+X_3 \quad \text{where}\\
&X_1=\sum_{t\in[T]:\boldsymbol{c}_t\notin\mathcal{C}(\underline{\sigma})}\sum_{k\in[K]}\Delta_k(\boldsymbol{c}_t)\mathbb{I}\{\rho_{j,t}=a_k,N_{k,t}\leq m_k(\boldsymbol{c}_t)\}\\
&X_2=\sum_{t\in[T]:\boldsymbol{c}_t\in\mathcal{C}}\sum_{k\in[K]}\Delta_k(\boldsymbol{c}_t)\mathbb{I}\{\rho_{j,t}=a_k,N_{k,t}\leq m_k(\boldsymbol{c}_t)\}\\
&X_3=\sum_{k\in[K]}\sum_{t\in[T]}\Delta_k(\boldsymbol{c}_t)\mathbb{I}\{\rho_{j,t}=a_k,N_{k,t}>m_k(\boldsymbol{c}_t)\}
\end{aligned}\tag{42}$$

In Section B.5.1, we show that $X_1\leq\mathcal{O}(\delta T+\underline{\sigma}T+\frac{1}{\delta})$; in Section B.5.2 we show that $X_2\leq\mathcal{O}(\delta T+\frac{1}{\delta\underline{\sigma}})$; in Section B.5.3 we show that $X_3\leq\mathcal{O}(1)$.

Here, regarding the final bound for$\sum_{t\in[T]}\mathcal{L}_j(\rho^*_{j,t},\boldsymbol{c}_t)-\mathcal{L}_j(\rho_{j,t},\boldsymbol{c}_t)$, without loss of generality we assume the optimal per-channel $\rho^*_j(t)=\arg\max_{\rho_j\in[0,\rho]}\mathcal{L}_j(\rho_j,\boldsymbol{c}_t)$ lies in the arm set $\mathcal{A}(\delta)$ for all t. This is because otherwise, we can decompose $\mathcal{L}_j(\rho^*_{j,t},\boldsymbol{c}_t)$ into $\mathcal{L}_j(\rho^*_{j,t},\boldsymbol{c}_t)-\max_{a_k\in\mathcal{A}(\delta)}\mathcal{L}_j(a_k,\boldsymbol{c}_t)$ plus $\max_{a_k\in\mathcal{A}(\delta)}(a_k,\boldsymbol{c}_t)-\mathcal{L}_j(\rho_{j,t},\boldsymbol{c}_t)$. The first term will yield an error in the order of $\mathcal{O}(\delta)$ in light of piecewise linearity of the Lagrangian function in Lemma 4.3 and also boundedness in the adjacent slopes $\sigma_j^-(\boldsymbol{c}_t)$ and $\sigma_j^+(\boldsymbol{c}_t)$ since $\boldsymbol{c}_t$ is bounded (Lemma 4.4). Hence, this "discretization error" will accumulate to a magnitude of $\mathcal{O}(\delta T)=\mathcal{O}(T^{2/3})$ when taking $\delta=\Theta(T^{-1/3})$, which will not impact the order of magnitude of our theorem's statement.

B.5.1. Bounding X_1. At each context $\boldsymbol{c}_t\notin\mathcal{C}(\underline{\sigma})$, we separate the set of arms $[K]$ into two groups, namely arms that lie at the left and right adjacent edges of optimal, namely $a_k\in\mathcal{U}_j^-(\boldsymbol{c}_t)\cup\mathcal{U}_j^+(\boldsymbol{c}_t)$ and all other arms (recall the sets $\mathcal{U}_j^-$ and $\mathcal{U}_j^+$ are defined in Eq. (40)).

(1) For arm k such that $a_k\in\mathcal{U}_j^-(\boldsymbol{c}_t)$, recall Lemma 4.3 that $\mathcal{L}_j(a,\boldsymbol{c}_t)$ is linear in a for $a\in\mathcal{U}_j^-(\boldsymbol{c}_t)$, so $\Delta_k(\boldsymbol{c}_t)=\sigma_j^-(\boldsymbol{c}_t)\cdot(\rho^*_{j,t}-a_k)\leq\underline{\sigma}(r_{j,n}-r_{j,n})\leq\underline{\sigma}\rho$, so summing over all such k we get

$$\begin{aligned}
&\sum_{t\in[T]:\boldsymbol{c}_t\notin\mathcal{C}(\underline{\sigma})}\;\sum_{k\in[K]:a_k\in\mathcal{U}_j^-(\boldsymbol{c}_t)}\Delta_k(\boldsymbol{c}_t)\mathbb{I}\{\rho_{j,t}=a_k,N_{k,t}\leq m_k(\boldsymbol{c}_t)\}\\
\leq&\sum_{t\in[T]:\boldsymbol{c}_t\notin\mathcal{C}(\underline{\sigma})}\;\sum_{k\in[K]:a_k\in\mathcal{U}_j^-(\boldsymbol{c}_t)}\underline{\sigma}\rho\cdot\mathbb{I}\{\rho_{j,t}=a_k,N_{k,t}\leq m_k(\boldsymbol{c}_t)\}\;\leq\;\underline{\sigma}\rho T=\mathcal{O}(\underline{\sigma}T)\,.
\end{aligned}\tag{43}$$

(2) For arm k such that $a_k<\min\mathcal{U}_j^-(\boldsymbol{c}_t)$, we further split contexts into groups $\mathcal{C}_n$ for $n=0\ldots S_j$ (defined in Eq. (41)) based on whether the corresponding optimal budget w.r.t. the Lagrangian at the context is

taken at the nth "turning point" (see Figure 2 of illustration). Then, for each context group n by defining $k' := \max\{k : a_k < r_{j,n-1}\}$ to be the arm closest to and less than $r_{j,n-1}$, we have

$$\begin{aligned}
&\sum_{t\in[T]:\boldsymbol{c}_t\in\mathcal{C}_n/\mathcal{C}(\underline{\sigma})} \sum_{k\in[K]:a_k<\min\mathcal{U}_j^-(\boldsymbol{c}_t)} \Delta_k(\boldsymbol{c}_t)\mathbb{I}\{\rho_{j,t}=a_k, N_{k,t}\le m_k(\boldsymbol{c}_t)\} \\
\overset{(i)}{=}& \sum_{t\in[T]:\boldsymbol{c}_t\in\mathcal{C}_n/\mathcal{C}(\underline{\sigma})} \sum_{k\in[K]:a_k<r_{j,n-1}} \Delta_k(\boldsymbol{c}_t)\mathbb{I}\{\rho_{j,t}=a_k, N_{k,t}\le m_k(\boldsymbol{c}_t)\} \\
=& \sum_{t\in[T]}\sum_{\boldsymbol{c}\in\mathcal{C}_n/\mathcal{C}(\underline{\sigma})} \sum_{k\in[K]:a_k<r_{j,n-1}} \Delta_k(\boldsymbol{c})\mathbb{I}\{\boldsymbol{c}_t=\boldsymbol{c},\rho_{j,t}=a_k, N_{k,t}\le m_k(\boldsymbol{c})\} \\
\overset{(ii)}{\le}& \sum_{t\in[T]}\sum_{\boldsymbol{c}\in\mathcal{C}_n/\mathcal{C}(\underline{\sigma})} \left(\Delta_{k'}(\boldsymbol{c})\mathbb{I}\{\boldsymbol{c}_t=\boldsymbol{c}\} + \sum_{k\in[K]:a_k<r_{j,n-1}-\delta} \Delta_k(\boldsymbol{c})\mathbb{I}\{\boldsymbol{c}_t=\boldsymbol{c},\rho_{j,t}=a_k, N_{k,t}\le m_k(\boldsymbol{c})\}\right) \\
\overset{(iii)}{\le}& \left((1+C_F)s_{j,n-1}\delta+\rho\underline{\sigma}\right)T + \sum_{k\in[K]:a_k<r_{j,n-1}-\delta}\ \sum_{\boldsymbol{c}\in\mathcal{C}_n/\mathcal{C}(\underline{\sigma})} \Delta_k(\boldsymbol{c})Y_k(\boldsymbol{c})
\end{aligned} \tag{44}$$

where in the final equality we defined $Y_k(\boldsymbol{c}) = \sum_{t\in[T]}\mathbb{I}\{\boldsymbol{c}_t=\boldsymbol{c},\rho_{j,t}=a_k,N_{k,t}\le m_k(\boldsymbol{c})\}$. In (i) we used the fact that the left end of the left adjacent region, i.e. $\min\mathcal{U}_j^-(\boldsymbol{c}_t)$ is exactly $r_{j,n-1}$ because for context $\boldsymbol{c}_t\in\mathcal{C}_n$ the optimal budget $\arg\max_{\rho_j\in[0,\rho]}\mathcal{L}_j(\rho_j,\boldsymbol{c}_t)$ is at the nth turning point; in (ii) we used the definition $k' := \max\{k : a_k < r_{j,n-1}\}$; (iii) follows the fact that under a context $\boldsymbol{c}\in\mathcal{C}_n/\mathcal{C}(\underline{\sigma})$, we have $\arg\max_{\rho_j\in[0,\rho]}\mathcal{L}_j(\rho_j,\boldsymbol{c}) = r_{j,n}$ so

$$\begin{aligned}
\Delta_{k'}(\boldsymbol{c}) &= \mathcal{L}_j(r_{j,n},\boldsymbol{c}) - \mathcal{L}_j(r_{j,n-1},\boldsymbol{c}) + \mathcal{L}_j(r_{j,n-1},\boldsymbol{c}) - \mathcal{L}_j(a_{k'},\boldsymbol{c}) \\
&= \sigma_j^-(\boldsymbol{c})(r_{j,n}-r_{j,n-1}) + \sigma_{j,n-1}(\boldsymbol{c})(r_{j,n-1}-a_{k'}) \\
&\overset{(iv)}{\le} \underline{\sigma}\rho + \sigma_{j,n-1}(\boldsymbol{c})\delta \\
&\overset{(v)}{\le} \underline{\sigma}\rho + (1+C_F)s_{j,n-1}\delta\,,
\end{aligned}$$

where in (iv) we used the fact that $\boldsymbol{c}\in\mathcal{C}_n/\mathcal{C}(\underline{\sigma})$ so $\sigma_j^-(\boldsymbol{c})\le\underline{\sigma}$, as well as k' lies on the line segment between points $r_{j,n-2}$ and $r_{j,n-1}$ since $\delta < \min_{n'\in[S_j]} r_{j,n'} - r_{j,n'-1}$; in (v) we recall $\sigma_{j,n-1}(\boldsymbol{c}) = (1+\lambda)s_{j,n-1} - (\mu+\gamma\lambda) \le (1+C_F)s_{j,n-1}$ where C_F is defined in Lemma 4.4.

We now bound $\sum_{\boldsymbol{c}\in\mathcal{C}_n/\mathcal{C}(\underline{\sigma})}\Delta_k(\boldsymbol{c})Y_k(\boldsymbol{c})$ in Eq. (44). It is easy to see the following inequality for any sequence of context $\boldsymbol{c}_{(1)},\dots,\boldsymbol{c}_{(\ell)} \in \{\boldsymbol{c}_t\}_{t\in[T]}$ (This is a more general inequality than a similar inequality shown in Balseiro et al. (2019a)):

$$Y_k(\boldsymbol{c}_{(1)}) + \cdots + Y_k(\boldsymbol{c}_{(\ell)}) \le \max_{\ell'=1\dots\ell} m_k(\boldsymbol{c}_{(\ell')})\,. \tag{45}$$

This is because

$$\begin{aligned}
\sum_{\ell'\in[\ell]} Y_k(\boldsymbol{c}_{(\ell')}) &= \sum_{t\in[T]}\sum_{\ell'\in[\ell]} \mathbb{I}\{\boldsymbol{c}_t=\boldsymbol{c}_{(\ell')},\rho_{j,t}=a_k,N_{k,t}\le m_k(\boldsymbol{c}_{(\ell')})\} \\
&\le \sum_{t\in[T]}\sum_{\ell'\in[\ell]} \mathbb{I}\{\boldsymbol{c}_t=\boldsymbol{c}_{(\ell')},\rho_{j,t}=a_k,N_{k,t}\le \max_{\ell'=1\dots\ell} m_k(\boldsymbol{c}_{(\ell')})\} \\
&= \sum_{t\in[T]} \mathbb{I}\{\boldsymbol{c}_t\in\{\boldsymbol{c}_{(\ell')}\}_{\ell'\in[\ell]},\rho_{j,t}=a_k,N_{k,t}\le \max_{\ell'=1\dots\ell} m_k(\boldsymbol{c}_{(\ell')})\} \\
&\le \max_{\ell'=1\dots\ell} m_k(\boldsymbol{c}_{(\ell')})\,.
\end{aligned}$$

For simplicity denote $L = |\mathcal{C}_n/\mathcal{C}(\underline{\sigma})|$, and order contexts in $\boldsymbol{c}\in\mathcal{C}_n/\mathcal{C}(\underline{\sigma})$ as $\{\boldsymbol{c}_{(\ell)}\}_{\ell\in[L]}$ s.t. $\Delta_k(\boldsymbol{c}_{(1)}) > \Delta_k(\boldsymbol{c}_{(2)}) > \cdots > \Delta_k(\boldsymbol{c}_{(L)})$, or equivalently $m_k(\boldsymbol{c}_{(1)}) < m_k(\boldsymbol{c}_{(2)}) < \cdots < m_k(\boldsymbol{c}_{(L)})$ according to Eq.(41). Then multiplying

Eq. (45) by by $\Delta_k(\boldsymbol{c}_{(\ell)}) - \Delta_k(\boldsymbol{c}_{(\ell+1)})$ (which is strictly positive due to the ordering of contexts), and summing $\ell = 1 \ldots L$ we get

$$\begin{aligned}
\sum_{\boldsymbol{c}\in\mathcal{C}_n/\mathcal{C}(\underline{\sigma})} \Delta_k(\boldsymbol{c})Y_k(\boldsymbol{c}) &= \sum_{\ell\in[L]} \Delta_k(\boldsymbol{c}_{(\ell)})Y_k(\boldsymbol{c}_{(\ell)}) \\
&\leq \sum_{\ell\in[L]} m_k(\boldsymbol{c}_{(\ell)})\left(\Delta_k(\boldsymbol{c}_{(\ell)}) - \Delta_k(\boldsymbol{c}_{(\ell+1)})\right) \\
&\stackrel{(i)}{=} 8\log(T) \sum_{\ell\in[L-1]} \frac{\Delta_k(\boldsymbol{c}_{(\ell)}) - \Delta_k(\boldsymbol{c}_{(\ell+1)})}{\Delta_k^2(\boldsymbol{c}_{(\ell)})} \\
&\stackrel{(ii)}{\leq} 8\log(T) \int_{z=\Delta_k(\boldsymbol{c}_{(L)})} \frac{dz}{z^2} \\
&= \frac{8\log(T)}{\Delta_k(\boldsymbol{c}_{(L)})} \stackrel{(iii)}{=} \frac{8\log(T)}{\min_{\boldsymbol{c}\in\mathcal{C}_n/\mathcal{C}(\underline{\sigma})} \Delta_k(\boldsymbol{c})}\,.
\end{aligned} \tag{46}$$

Here (i) follows from the definition in Eq. (41) where we defined $m_k(\boldsymbol{c}) = \frac{8\log(T)}{\Delta_k^2(\boldsymbol{c})}$; both (ii) and (iii) follow from the ordering of contexts so that $\Delta_k(\boldsymbol{c}_{(1)}) > \Delta_k(\boldsymbol{c}_{(2)}) > \cdots > \Delta_k(\boldsymbol{c}_{(L)})$. Note that for any $\boldsymbol{c} \in \mathcal{C}_n/\mathcal{C}(\underline{\sigma})$ and arm k such that $a_k < r_{j,n-1}$, we have

$$\begin{aligned}
\Delta_k(\boldsymbol{c}) &= \mathcal{L}_j(r_{j,n},\boldsymbol{c}) - \mathcal{L}_j(r_{j,n-1},\boldsymbol{c}) + \mathcal{L}_j(r_{j,n-1},\boldsymbol{c}) - \mathcal{L}_j(a_k,\boldsymbol{c}) \\
&> \mathcal{L}_j(r_{j,n-1},\boldsymbol{c}) - \mathcal{L}_j(a_k,\boldsymbol{c}) \\
&\stackrel{(i)}{\geq} \sigma_{j,n-1}(\boldsymbol{c})(r_{j,n-1} - a_k) \\
&\stackrel{(ii)}{\geq} (\sigma_{j,n-1}(\boldsymbol{c}) - \sigma_{j,n}(\boldsymbol{c}))\,(r_{j,n-1} - a_k) \\
&\stackrel{(iii)}{=} (1+\lambda)\,(s_{j,n-1} - s_{j,n})\,(r_{j,n-1} - a_k) \\
&> (s_{j,n-1} - s_{j,n})\,(r_{j,n-1} - a_k)
\end{aligned} \tag{47}$$

where in (i) we recall the slope $\sigma_{j,n-1}(\boldsymbol{c})$ is defined in Lemma 4.3 and further (i) follows from concavity of $\mathcal{L}_j(\rho j,\boldsymbol{c})$ in the first argument ρ_j; in (ii) we used the fact that $\sigma_{j,n}(\boldsymbol{c}) \geq 0$ since the optimal budget $\arg\max_{\rho_j\in[0,\rho]} \mathcal{L}_j(\rho_j,\boldsymbol{c})$ is taken at the nth turning point, and is the largest turning point whose left slope is non-negative from Lemma 4.3; (iii) follows from the definition $\sigma_{j,n'}(\boldsymbol{c}) = (1+\lambda)s_{j,n'} - (\mu + \gamma\lambda)$ for any n'.

Finally combining Eq. (44), (46) and (47), and summing over $n = 1 \ldots S_j$ we get

$$\begin{aligned}
&\sum_{t\in[T]:\boldsymbol{c}_t\notin\mathcal{C}(\underline{\sigma})} \sum_{k\in[K]:a_k<\min\mathcal{U}_j^-(\boldsymbol{c}_t)} \Delta_k(\boldsymbol{c}_t)\mathbb{I}\{\rho_{j,t} = a_k, N_{k,t} \leq m_k(\boldsymbol{c}_t)\} \\
&= \sum_{n\in[S_j]} \sum_{t\in[T]:\boldsymbol{c}_t\in\mathcal{C}_n/\mathcal{C}(\underline{\sigma})} \sum_{k\in[K]:a_k<\min\mathcal{U}_j^-(\boldsymbol{c}_t)} \Delta_k(\boldsymbol{c}_t)\mathbb{I}\{\rho_{j,t} = a_k, N_{k,t} \leq m_k(\boldsymbol{c}_t)\} \\
&\leq \sum_{n\in[S_j]} ((1+C_F)s_{j,n-1}\delta + \rho\underline{\sigma})\,T + \sum_{n\in[S_j]} \sum_{k\in[K]:a_k<r_{j,n-1}-\delta} \frac{8\log(T)}{(s_{j,n-1} - s_{j,n})\,(r_{j,n-1} - a_k)} \\
&\stackrel{(i)}{\leq} \sum_{n\in[S_j]} ((1+C_F)s_{j,n-1}\delta + \rho\underline{\sigma})\,T + \sum_{n\in[S_j]} \sum_{\ell=1}^{K} \frac{8\log(T)}{(s_{j,n-1} - s_{j,n})\,\ell\delta} \\
&\leq \sum_{n\in[S_j]} ((1+C_F)s_{j,n-1}\delta + \rho\underline{\sigma})\,T + \frac{8\log(T)\log(K)}{\delta} \sum_{n\in[S_j]} \frac{1}{(s_{j,n-1} - s_{j,n})} \\
&= \mathcal{O}(\delta T + \underline{\sigma}T + \frac{1}{\delta})\,.
\end{aligned} \tag{48}$$

Note that (i) follows because for all $a_k < r_{j,n-1} - \delta$, the a_k's distances from $r_{j,n-1}$ are at least $\delta, 3\delta, 2\delta \ldots$.

(3) The cases where arm $a_k \in \mathcal{U}_j^+(\boldsymbol{c}_t)$ and $a_k > \max\mathcal{U}_j^+(\boldsymbol{c}_t)$ are symmetric to $a_k \in \mathcal{U}_j^-(\boldsymbol{c}_t)$ and $a_k < \min\mathcal{U}_j^+(\boldsymbol{c}_t)$, respectively, and we omit from this paper.

Therefore, combining Eq. (43) and (48) we can conclude

$$X_1 \leq \mathcal{O}(\delta T + \underline{\sigma} T + \frac{1}{\delta}) \tag{49}$$

B.5.2. Bounding X_2. We first rewrite X_2 as followed

$$\begin{aligned}
X_2 &= \sum_{t\in[T]:\boldsymbol{c}_t\in\mathcal{C}(\underline{\sigma})} \sum_{k\in[K]} \Delta_k(\boldsymbol{c}_t)\mathbb{I}\{\rho_{j,t} = a_k, N_{k,t} \leq m_k(\boldsymbol{c}_t)\} \\
&= \sum_{t\in[T]} \sum_{n\in[S_j]} \sum_{k\in[K]} \sum_{\boldsymbol{c}\in\mathcal{C}_n\cap\mathcal{C}(\underline{\sigma})} \Delta_k(\boldsymbol{c})\mathbb{I}\{\boldsymbol{c}_t = \boldsymbol{c}, \rho_{j,t} = a_k, N_{k,t} \leq m_k(\boldsymbol{c})\} \\
&\overset{(i)}{=} \sum_{n\in[S_j]} \sum_{k\in[K]} \sum_{\boldsymbol{c}\in\mathcal{C}_n\cap\mathcal{C}(\underline{\sigma})} \Delta_k(\boldsymbol{c}) Y_k(\boldsymbol{c}) \\
&\overset{(ii)}{=} T\delta(1+C_F) \sum_{n\in[S_j]} (s_{j,n} + s_{j,n+1}) + \sum_{n\in[S_j]} \sum_{k\in[K]/\{k_n^-,k_n^+\}} \sum_{\boldsymbol{c}\in\mathcal{C}_n\cap\mathcal{C}(\underline{\sigma})} \Delta_k(\boldsymbol{c}) Y_k(\boldsymbol{c})\,.
\end{aligned} \tag{50}$$

where in (i) we define $Y_k(\boldsymbol{c}) = \sum_{t\in[T]} \mathbb{I}\{\boldsymbol{c}_t = \boldsymbol{c}, \rho_{j,t} = a_k, N_{k,t} \leq m_k(\boldsymbol{c})\}$; in (ii) for any context $\boldsymbol{c} \in \mathcal{C}_n \cap \mathcal{C}(\underline{\sigma})$, the optimal budget $\arg\max_{\rho_j\in[0,\rho]} \mathcal{L}_j(\rho_j, \boldsymbol{c}) = r_{j,n}$ is taken at the nth turning point per the definition of $\mathcal{C}_n$ in Eq. (41), and further we defined $k_n^- := \max\{k \in [K] : a_k < r_{j,n}\}$ to be the arm closest to and no greater than $r_{j,n}$, whereas $k_n^+ := \min\{k \in [K] : a_k > r_{j,n}\}$ to be the arm closest to and no less than $r_{j,n}$. Thus for small enough $\delta < \min_{n'\in[S_j]} r_{j,n'} - r_{j,n'-1}$, k_n^- lies on the line segment between $r_{j,n-1}$ and $r_{j,n}$, so $\Delta_{k_n^-}(\boldsymbol{c}) = \sigma_j^-(\boldsymbol{c})(r_{j,n} - a_{k_n^-}) \leq \sigma_j^-(\boldsymbol{c})\delta \leq (1+C_F)s_{j,n-1}\delta$, where in the final inequality follows from the definition of $\sigma_j^-(\boldsymbol{c}) = \sigma_{j,n}(\boldsymbol{c}) = (1+\lambda)s_{j,n} - (\mu+\gamma\lambda) \leq (1+\lambda)s_{j,n} \leq (1+C_F)s_{j,n}$ where C_F is defined in Eq. (4.4). A similar bound holds for $\Delta_{k_n^+}(\boldsymbol{c})$.

Then, following the same logic as Eq. (45), (46), (47) in Section B.5.1 where we bound X_1, we can bound $\sum_{\boldsymbol{c}\in\mathcal{C}_n\cap\mathcal{C}(\underline{\sigma})} \Delta_k(\boldsymbol{c}) Y_k(\boldsymbol{c})$ as followed:

$$\sum_{\boldsymbol{c}\in\mathcal{C}_n\cap\mathcal{C}(\underline{\sigma})} \Delta_k(\boldsymbol{c}) Y_k(\boldsymbol{c}) \leq \frac{8\log(T)}{\min_{\boldsymbol{c}\in\mathcal{C}_n\cap\mathcal{C}(\underline{\sigma})} \Delta_k(\boldsymbol{c})}\,. \tag{51}$$

Now, the set $k \in [K]/\{k_n^-, k_n^+\}$ in Eq. (50) can be split into two sets, namely $\{k \in [K] : a_k < r_{j,n} - \delta\}$ and $\{k \in [K] : a_k > r_{j,n} + \delta\}$ due to the definitions $k_n^- := \max\{k \in [K] : a_k < r_{j,n}\}$ and $k_n^+ := \min\{k \in [K] : a_k > r_{j,n}\}$. Therefore, for any k s.t. $a_k < r_{j,n} - \delta$ and any $\boldsymbol{c} \in \mathcal{C}_n \cap \mathcal{C}(\underline{\sigma})$,

$$\Delta_k(\boldsymbol{c}) = \mathcal{L}_j(r_{j,n}, \boldsymbol{c}) - \mathcal{L}_j(a_k, \boldsymbol{c}) \geq \sigma_j^-(\boldsymbol{c})(r_{j,n} - a_k) \geq \underline{\sigma}(r_{j,n} - a_k),\,,$$

where the final inequality follows from the definition of $\mathcal{C}(\underline{\sigma})$ in Eq. (41) such that $\sigma_j^-(\boldsymbol{c}) \geq \underline{(}\sigma)$ for $\boldsymbol{c}\mathcal{C}(\underline{\sigma})$. Hence combining this with Eq. (51) we have

$$\sum_{k\in[K]:a_k<r_{j,n}-\delta} \sum_{\boldsymbol{c}\in\mathcal{C}_n\cap\mathcal{C}(\underline{\sigma})} \Delta_k(\boldsymbol{c}) Y_k(\boldsymbol{c}) \leq \sum_{k\in[K]:a_k<r_{j,n}-\delta} \frac{8\log(T)}{\underline{\sigma}(r_{j,n} - a_k)} \overset{(i)}{\leq} \sum_{\ell=1}^{K} \frac{8\log(T)}{\underline{\sigma}\ell\delta} \leq \frac{8\log(T)\log(K)}{\underline{\sigma}\delta} \tag{52}$$

where (i) follows because for all $a_k < r_{j,n} - \delta$, the a_k's distances from $r_{j,n-1}$ are at least $\delta, 3\delta, 2\delta \ldots$. Symmetrically, we can show an identical bound for the set $\{k \in [K] : a_k > r_{j,n} + \delta\}$. Hence, combining Eq. (50) and (52) we can conclude

$$X_2 \leq \mathcal{O}\left(\delta T + \frac{1}{\delta\underline{\sigma}}\right). \tag{53}$$

B.5.3. Bounding X_3. For any arm $k \in [K]$, denote $R_t(a_k) = \bar{V}_t(a_k) - \frac{\lambda_t\gamma+\mu_t}{1+\lambda_t}a_k$. Also, define $\bar{\mathcal{L}} = (1+\gamma)\,\rho C_F + (1+C_F)\bar{V}$, where C_F is specified in Lemma 4.4, and recall $\rho_j^*(t) = \arg\max_{a_{k'}\in\mathcal{A}(\delta)} \mathcal{L}_j(a_{k'}, \boldsymbol{c}_t)$ since we assumed WLOG $\rho_j^*(t) \in \mathcal{A}(\delta)$ for all t. Then we bound X_3 as followed

$$\begin{aligned} X_3 &= \sum_{t\in[T]} \mathbb{E}\left[\Delta_k(\boldsymbol{c})\mathbb{I}\{\rho_{j,t} = a_k, N_{k,t} > m_k(\boldsymbol{c})\}\right] \\ &\overset{(i)}{\leq} \bar{\mathcal{L}} \cdot \mathbb{P}\left(\rho_{j,t} = a_k, N_{k,t} > m_k(\boldsymbol{c}_t)\right) \\ &\leq \bar{\mathcal{L}} \cdot \sum_{t\in[T]} \mathbb{P}\left(R_t(a_k) + \texttt{UCB}_{j,t}(a_k) \geq R_t(\rho_j^*(t)) + \texttt{UCB}_{j,t}(\rho_j^*(t)), N_{k,t} > m_k(\boldsymbol{c}_t)\right)\,, \end{aligned} \tag{54}$$

where (i) follows from Lemma 4.4 such that $-(1+\gamma)\,\rho C_F \leq \mathcal{L}_j(\rho_j, \boldsymbol{c}_t) \leq (1+C_F)\bar{V}$ for all $j \in [M]$, $t\ in[T]$, $\rho_j \in [0, \rho]$ and $\boldsymbol{c}_t$.

Now let $\bar{R}_n(a_k)$ denote the average conversion of arm k over its first n pulls, i.e. letting $\tau = \min\{t \in [T] : N_{k,t} = n\}$ be the period during which arm k is pulled for the nth time, then $\bar{R}_n(a_k) = \bar{V}_{j,\tau}(a_k)$, where $\bar{V}_{j,\tau}$ is defined in Algorithm 1. Hence, we continue as follwoed

$$\begin{aligned} &\mathbb{P}\left(R_t(a_k) + \texttt{UCB}_{j,t}(a_k) \geq R_t(\rho_j^*(t)) + \texttt{UCB}_{j,t}(\rho_j^*(t)), N_{k,t} > m_k(\boldsymbol{c}_t)\right) \\ \leq\ & \mathbb{P}\left(\max_{m_k(\boldsymbol{c}_t)<n\leq t} \bar{R}_n(a_k) + \texttt{UCB}_t(a_k) - \frac{\lambda_t\gamma+\mu_t}{1+\lambda_t}a_k \geq \min_{1\leq n'\leq t} \bar{R}_n(\rho_j^*(t)) + \texttt{UCB}_t(\rho_j^*(t)) - \frac{\lambda_t\gamma+\mu_t}{1+\lambda_t}\rho_j^*(t)\right) \\ \leq\ & \sum_{n=\lceil m_k(\boldsymbol{c}_t)\rceil+1}^{t} \sum_{n'=1}^{t} \mathbb{P}\left(\bar{R}_n(a_k) + \texttt{UCB}_n(a_k) - \frac{\lambda_t\gamma+\mu_t}{1+\lambda_t}a_k > \bar{R}_{n'}(\rho_j^*(t)) + \texttt{UCB}_{n'}(\rho_j^*(t)) - \frac{\lambda_t\gamma+\mu_t}{1+\lambda_t}\rho_j^*(t)\right) \end{aligned} \tag{55}$$

Now, when the event $\left\{\bar{R}_n(a_k) + \texttt{UCB}_n(a_k) - \frac{\lambda_t\gamma+\mu_t}{1+\lambda_t}a_k > \bar{R}_n(\rho_j^*(t)) + \texttt{UCB}_{n'}(\rho_j^*(t)) - \frac{\lambda_t\gamma+\mu_t}{1+\lambda_t}\rho_j^*(t)\right\}$ occurs, it is easy to see that one of the following events must also occur:

$$\begin{aligned} &\mathcal{G}_{1,n} = \left\{\bar{R}_n(a_k) \geq V(a_k) + \texttt{UCB}_n(a_k)\right\} \\ &\mathcal{G}_{2,n'} = \left\{\bar{R}_{n'}(\rho_j^*(t)) \leq V(\rho_j^*(t)) - \texttt{UCB}_n(\rho_j^*(t))\right\} \\ &\mathcal{G}_3 = \left\{V(\rho_j^*(t)) - \frac{\lambda_t\gamma+\mu_t}{1+\lambda_t}\rho_j^*(t) \ < \ V(a_k) - \frac{\lambda_t\gamma+\mu_t}{1+\lambda_t}a_k + 2\cdot\texttt{UCB}_n(a_k)\right\} \end{aligned} \tag{56}$$

Note that for $n > m_k(\boldsymbol{c}_t)$, we have $\texttt{UCB}_n(a_k) = \sqrt{\frac{2\log(T)}{n}} < \sqrt{\frac{2\log(T)}{m_k(\boldsymbol{c}_t)}} = \frac{\Delta_k(\boldsymbol{c}_t)}{2}$ since we defined $m_k(\boldsymbol{c}) = \frac{8\log(T)}{\Delta_k^2(\boldsymbol{c})}$ in Eq. (41). Therefore

$$V(a_k) - \frac{\lambda_t\gamma+\mu_t}{1+\lambda_t}a_k + 2\cdot\texttt{UCB}_n(a_k) < \underbrace{V(a_k) - \frac{\lambda_t\gamma+\mu_t}{1+\lambda_t}a_k}_{=\mathcal{L}(a_k,\boldsymbol{c}_t)} + \Delta_k(\boldsymbol{c}_t) \overset{(i)}{=} \underbrace{V(\rho_j^*(t)) - \frac{\lambda_t\gamma+\mu_t}{1+\lambda_t}\rho_j^*(t)}_{=\mathcal{L}(\rho_j^*(t),\boldsymbol{c}_t)=\max_{a\in\mathcal{A}(\delta)}\mathcal{L}(a,\boldsymbol{c}_t)}$$

where (i) follows from the definition of $\Delta_k(\boldsymbol{c}) = \max_{a\in\mathcal{A}(\delta)} \mathcal{L}(a, \boldsymbol{c}) - \mathcal{L}(a_k, \boldsymbol{c})$ in Eq. (41) for any context $\boldsymbol{c}$. This implies that event $\mathcal{G}_3$ in Eq. (56) cannot hold for $n > m_k(\boldsymbol{c}_t)$. Therefore

$$\mathbb{P}\left(\bar{R}_n(a_k) + \texttt{UCB}_n(a_k) - \frac{\lambda_t\gamma+\mu_t}{1+\lambda_t}a_k > \bar{R}_{n'}(\rho_j^*(t)) + \texttt{UCB}_{n'}(\rho_j^*(t)) - \frac{\lambda_t\gamma+\mu_t}{1+\lambda_t}\rho_j^*(t)\right) \leq \mathbb{P}\left(\mathcal{G}_{1,n} \cup \mathcal{G}_{2,n'}\right) \tag{57}$$

From standard UCB analysis and the Azuma Hoeffding's inequality, we have $\mathbb{P}(\mathcal{G}_{1,n}) \leq \frac{\bar{V}}{T^4}$ and $\mathbb{P}(\mathcal{G}_{2,n'}) \leq \frac{\bar{V}}{T^4}$. Hence combining Eq. (54) (55), (57) we can conclude

$$X_3 \le \sum_{t\in[T]} \sum_{n=\lceil m_k(\boldsymbol{c}_t)\rceil+1}^{t} \sum_{n'=1}^{t} \left(\mathbb{P}\left(\mathcal{G}_{1,n}\right) + \mathbb{P}\left(\mathcal{G}_{2,n'}\right)\right)$$
$$\le \sum_{t\in[T]} \sum_{n=\lceil m_k(\boldsymbol{c}_t)\rceil+1}^{t} \sum_{n'=1}^{t} \frac{2\bar{V}}{T^4} \tag{58}$$
$$\le \frac{2\bar{V}}{T}$$

□

B.6. Proof for Theorem 4.6

Starting from Proposition 4.1, we get

$$T\cdot \text{GL-OPT} - \mathbb{E}\left[\sum_{t\in[T]}\sum_{j\in[M]} V_j(\rho_{j,t})\right] \le M\bar{V}K + \sum_{j\in[M]}\sum_{t>K}\mathbb{E}\left[\mathcal{L}_j(\rho_j^*(t),\boldsymbol{c}_t) - \mathcal{L}_j(\rho_{j,t},\boldsymbol{c}_t)\right] + \sum_{t>K}(\lambda_t g_{1,t} + \mu_t g_{2,t})$$
$$\overset{(i)}{\le} M\bar{V}K + \mathcal{O}\left(\underline{\sigma}T + \delta T + \frac{1}{\underline{\sigma}\delta}\right) + \mathcal{O}\left(\eta T + \frac{1}{\eta}\right) \tag{59}$$

where in (i) we applied Lemma 4.5 and 4.2. Taking $\eta = 1/\sqrt{T}$, $\delta = \underline{\sigma} = T^{-1/3}$ (i.e. $K = \mathcal{O}(T^{1/3})$ yields $T\cdot\text{GL-OPT} - \mathbb{E}\left[\sum_{t\in[T]}\sum_{j\in[M]} V_j(\rho_{j,t})\right] \le \mathcal{O}(T^{2/3})$. According to Lemma 4.3, $V_j(\rho_j)$ is concave for all $j\in[M]$, so

$$\mathcal{O}(T^{-1/3}) \ge \text{GL-OPT} - \frac{1}{T}\sum_{t\in[T]}\mathbb{E}\left[\sum_{j\in[M]} V_j(\rho_{j,t})\right]$$
$$\ge \text{GL-OPT} - \mathbb{E}\left[\sum_{j\in[M]} V_j\left(\frac{1}{T}\sum_{t\in[T]}\rho_{j,t}\right)\right]$$
$$\ge \text{GL-OPT} - \mathbb{E}\left[\sum_{j\in[M]} V_j(\bar{\rho}_{T,j})\right]$$

where in the final equality we used the definition $\bar{\boldsymbol{\rho}}_T$ as defined in Algorithm 1.

On the other hand, Lemma 4.2 shows

$$-\mathcal{O}(1/\sqrt{T}) \le \frac{1}{T}\sum_{t\in[T]}\mathbb{E}\left[g_{1,t}\right]$$
$$= \frac{1}{T}\sum_{t\in[T]}\sum_{j\in[M]}\mathbb{E}\left[(V_j(\rho_{j,t};\boldsymbol{z}_t) - \gamma\rho_{j,t})\right]$$
$$= \frac{1}{T}\sum_{t\in[T]}\sum_{j\in[M]}\mathbb{E}\left[(V_j(\rho_{j,t}) - \gamma\rho_{j,t})\right]$$
$$\overset{(i)}{\le} \sum_{j\in[M]}\left(V_j\left(\frac{1}{T}\sum_{t\in[T]}\rho_{j,t}\right) - \gamma\cdot\frac{1}{T}\sum_{t\in[T]}\rho_{j,t}\right)$$
$$= \sum_{j\in[M]}\left(V_j\left(\bar{\rho}_{T,j}\right) - \gamma\bar{\rho}_{T,j}\right).$$

where in (i) we again applied concavity of $V_j(\rho_j)$. We omit the analysis for the budget constraint as it is similar to the above. □

B.7. Proof of Lemma 4.7

Consider the hypothetical version of SGD-UCB-II where we ignore the budget balance condition $B_t < M\rho$ in step 2 of the algorithm, and terminate phase 1 only when the condition $\sum_{t'\in[t-1]} g_{1,t'} > \sqrt{T}\log(T)$ in Step 1 holds. Denote this hypothetical stopping time as $\widetilde{T}_1$, defined to be

$$\widetilde{T}_1 = \min\Big\{t\in[T]: \sum_{t'\in[t]} \underbrace{\sum_{j\in M}\left(V_j(\beta;\boldsymbol{z}_{t'})-\gamma\beta\right)}_{:=h_{t'}} > \sqrt{T}\log(T)\Big\} \tag{60}$$

It is easy to see that with probability 1, we have $\widetilde{T}_1 \geq T_1$ where we recall T_1 is the real stopping time of Phase 1 in SGD-UCB-II. Hence we have

$$\mathbb{P}\left(T_1 \geq R\right) \leq \mathbb{P}\left(\widetilde{T}_1 \geq R\right) \tag{61}$$

Now, it is not difficult to see $\{h_t\}_t$ where h_t defined in Eq. (60) are i.i.d. random variables, since the only randomness comes from the realization of value-cost pairs $\{\boldsymbol{z}_t\}_{t\in[T]}$ which are i.i.d.. Further, for any $t\in[T]$ we have

$$h_t := \sum_{j\in M}\left(V_j(\beta;\boldsymbol{z}_t)-\gamma\beta\right) \overset{(i)}{\geq} (\xi-\gamma)\beta \implies \bar{h} := \mathbb{E}[h_t] \geq (\xi-\gamma)\beta > 0 \tag{62}$$

where (i) follows from Claim B.1 where we also defined $\xi := \min_{j\in[M]}\min_{\boldsymbol{z}_j\in F_j}\frac{v_{j,1}}{d_{j,1}} > \gamma$ under Assumption 4.3; in the final inequality we let $\bar{h}$ be the mean of the i.i.d. random variables $\{h_t\}_{t\in[T]}$.

We note that $\widetilde{T}_1 > R$ implies that the sum of the first R h_t's do not exceed $\sqrt{T}\log(T)$ (see definition of $\widetilde{T}_1$ in Eq. (60)), hence we have

$$\begin{aligned}
\mathbb{P}\left(\widetilde{T}_1 > R\right) &\leq \mathbb{P}\left(\sum_{t=1}^{\lceil R\rceil} h_t \leq \sqrt{T}\log(T)\right)\\
&= \mathbb{P}\left(\sum_{t=1}^{\lceil R\rceil}(h_t-\bar{h}) \leq \sqrt{T}\log(T) - \lceil R\rceil\cdot\bar{h}\right)\\
&\overset{(i)}{\leq} \mathbb{P}\left(\sum_{t=1}^{\lceil R\rceil}(h_t-\bar{h}) \leq \sqrt{T}\log(T) - R(\xi-\gamma)\beta\right)\\
&\overset{(ii)}{=} \mathbb{P}\left(\sum_{t=1}^{\lceil R\rceil}(h_t-\bar{h}) \leq -\sqrt{T}\log(T)\right)\\
&\overset{(iii)}{\leq} \exp\left(-\frac{T\log^2(T)}{2\lceil R\rceil M^2\bar{V}^2}\right)\\
&\overset{(iv)}{\leq} \frac{1}{T}
\end{aligned} \tag{63}$$

Here (i) follows from $\bar{h} \geq (\xi-\gamma)\beta$ in Eq. (62); (ii) follows from the definition that $R = 2\sqrt{T}\log^3(T)$ and $\beta = \frac{1}{\log(T)}$ so $R(\xi-\gamma)\beta = 2(\xi-\gamma)\sqrt{T}\log^2(T) \geq 2\sqrt{T}\log(T)$ for large enough T such that $\log(T) > \frac{1}{\xi-\gamma}$; (iii) follows from Azuma Hoefding's inequality given that $h_t\in[0, M\bar{V}]$ for any $t\in[T]$; and finally (iv) follows from $T \geq \lceil R\rceil$, and $\log(T) > M^2\bar{V}^2$ for large enough T. Hence, combining Eq. (63) and (61) yields the desired statement of the lemma. □

B.8. Proof of Theorem 4.8

The proof of this theorem consists of 3 parts. In Part I, we bound the global budget constraint violation; in Part II, we bound the global ROI constraint violation; in Part III, we bound the conversion error.

Part I. Bounding global budget constraint violation. The design of the SGD-UCB-II algorithm ensures that the per-channel budget decisions never sum up to exceed $\rho T - M\rho$, so

$$\frac{1}{T}\sum_{t\in[T_2]}\sum_{j\in[M]}\rho_{j,t} \le \frac{1}{T}(\rho T - M\rho) < \rho\,.$$

Part II. Bounding global ROI constraint violation.

Recall the event $\mathcal{E} = \{T_1 \ge 2\sqrt{T}\log^3(T)\}$ defined in Lemma 4.7 where T_1 is the end period of Phase 1 in the SGD-UCB-II algorithm (see step 10). We consider two scenarios, namely when event $\mathcal{E}$ holds and doesn't hold. Recall $g_{1,t} = \sum_{j\in[M]}(V_j(\rho_{j,t};\boldsymbol{z}_t) - \gamma\rho_{j,t}) \in [-\gamma M\rho, MV]$

When event $\mathcal{E}$ holds, then

$$\mathbb{E}\left[\sum_{t\in[T_2]} g_{1,t} \;\middle|\; \mathcal{E}\right] \ge -T\gamma M\rho \tag{64}$$

When event $\mathcal{E}$ does not hold, Phase 1 terminates within the first $2\sqrt{T}\log^3(T)$ periods, i.e. $T_1 < 2\sqrt{T}\log^3(T)$, and the total spend balance in Phase 1 is at most

$$\sum_{t\in[T_1]}\sum_{j\in[M]}\beta = M\beta T_1 < 2M\sqrt{T}\log^2(T) < \rho T - \rho M$$

where the final inequality holds for large enough T. This implies that Phase 1 terminates because the ROI buffer condition in step 1 is met, i.e.

$$\sum_{t\in[T_1]} g_{1,t} > \sqrt{T}\log(T) \tag{65}$$

Now in periods $t = T_1 + 1 \dots T_2$, following the proof of Lemma B.2, we have

$$\sum_{t=T_1+1}^{T_2} g_{1,t} \ge \frac{1}{\eta}\left(\lambda_{T_1} - \lambda_{T_2}\right) \overset{(i)}{\ge} -\frac{C_F}{\eta} = -C_F\sqrt{T} \tag{66}$$

we (i) follows from Lemma 4.4 such that $0 \le \lambda_t \le C_F$ for any $t \in [T]$ for some absolute constant $C_F > 0$ that only depends on the support of value-cost pairs $F = F_1 \times \ldots \times F_M$, and recall the SGD step size $\eta = 1/\sqrt{T}$. Hence, combining Eq. (64), (65), (66), we have

$$\begin{aligned}
\mathbb{E}\left[\sum_{t\in[T_2]} g_{1,t}\right] &= \mathbb{E}\left[\sum_{t\in[T_2]} g_{1,t} \;\middle|\; \mathcal{E}\right]\mathbb{P}(\mathcal{E}) + \mathbb{E}\left[\sum_{t\in[T_2]} g_{1,t} \;\middle|\; \mathcal{E}^c\right]\mathbb{P}(\mathcal{E}^c) \\
&\ge -T\gamma M\rho\cdot\mathbb{P}(\mathcal{E}) + \mathbb{E}\left[\sum_{t\in[T_2]} g_{1,t} \;\middle|\; \mathcal{E}^c\right]\mathbb{P}(\mathcal{E}^c) \\
&\overset{(i)}{\ge} -\gamma M\rho + \left(\sqrt{T}\log(T) - C_F\sqrt{T}\right)\mathbb{P}(\mathcal{E}^c) \\
&\overset{(ii)}{\ge} 0
\end{aligned} \tag{67}$$

where (i) follows from Lemma 4.7 which states $\mathbb{P}(\mathcal{E}) \le \frac{1}{T}$; in (ii) we used the fact that for large enough T, we have $\log(T) > -C_F$, and $\mathbb{P}(\mathcal{E}^c) \ge 1 - 1/T \ge 1/2$ so $-\gamma M\rho + \frac{1}{2}\left(\sqrt{T}\log(T) - C_F\sqrt{T}\right) > 0$ since $\log(T) > C_F + \frac{2\gamma M\rho}{\sqrt{T}}$ for large T. Therefore Eq. (67) implies

$$\mathbb{E}\left[\sum_{t\in[T_2]}\sum_{j\in[M]}\left(V_j(\rho_{j,t}) - \gamma\rho_{j,t}\right)\right] \ge 0\,. \tag{68}$$

Finally, we have

$$\sum_{j\in[M]}\mathbb{E}\left[V_j(\overline{\rho}_j) - \gamma\overline{\rho}_j\right] \overset{(i)}{\ge} \frac{1}{T}\mathbb{E}\left[\sum_{t\in[T_2]}\sum_{j\in[M]}\left(V_j(\rho_{j,t}) - \gamma\rho_{j,t}\right)\right] \ge 0$$

where (i) follows from concavity of $V_j(\rho_j)$ according to Lemma 4.3.

Part III. Bounding conversion error. We first show that $T - T_2 \le M + \frac{C_F}{\rho}\sqrt{T}$ where C_F is an absolute constant independent of T defined in Lemma 4.4.

If $T_2 = T$, then there is nothing to show. Assume $T_2 < T$, the by the algorithm's termination criteria we have $\rho T \le B_{T_2} + M\rho = M\rho + \sum_{t\in[T_2]}\sum_{j\in[M]}\rho_{j,t}$, where we recall the definition of the spend balance $B_t = \sum_{t'\in[t]}\sum_{j\in[M]}\rho_{j,t'}$. Now, recalling $g_{2,t} = \rho - \sum_{j\in[M]}\rho_{j,t}$, we have

$$\begin{aligned}\rho T \;&\le\; T_2\rho - \sum_{t\in[T_2]} g_{2,t} + M\rho \;\overset{(i)}{\le}\; T_2\rho + \frac{C_F}{\eta} + M\rho\\ \Longrightarrow& T - T_2 \le M + \frac{C_F}{\rho}\sqrt{T}\,.\end{aligned} \tag{69}$$

where (i) follows from the proof of Lemma B.2 such that $\sum_{t=1}^{T_2} g_{2,t} \ge \frac{1}{\eta}\left(\lambda_1 - \lambda_{T_2}\right) \overset{(ii)}{\ge} -\frac{C_F}{\eta} = -C_F\sqrt{T}$, and (ii) follows from Lemma 4.4 such that $0 \le \lambda_t \le C_F$ for any $t\in[T]$ for some absolute constant $C_F > 0$ that only depends on the support of value-cost pairs $F = F_1 \times \ldots \times F_M$, and recall the SGD step size $\eta = 1/\sqrt{T}$.

Hence, recalling the event $\mathcal{E} = \{T_1 > 2\sqrt{T}\log^3(T)\}$ defined in Lemma 4.7, we have

$$\begin{aligned}&T\cdot\text{GL-OPT} - \mathbb{E}\left[\sum_{t\in[T_2]}\sum_{j\in[M]} V_j(\rho_{j,t})\right]\\ &\le\; \mathbb{E}\left[T\cdot\text{GL-OPT}\;\middle|\;\mathcal{E}\right]\mathbb{P}(\mathcal{E}) + \mathbb{E}\left[T\cdot\text{GL-OPT} - \sum_{t\in[T_2]}\sum_{j\in[M]} V_j(\rho_{j,t})\middle|\;\mathcal{E}^c\right]\\ &\overset{(i)}{\le}\; M\bar{V} + \mathbb{E}\left[(T_1 + T - T_2)\cdot\text{GL-OPT} + (T_2 - T_1)\cdot\text{GL-OPT} - \sum_{t=T_1+1}^{T_2}\sum_{j\in[M]} V_j(\rho_{j,t})\middle|\;\mathcal{E}^c\right]\\ &\overset{(ii)}{\le}\; M\bar{V} + M\bar{V}\left(2\sqrt{T}\log^3(T) + M + \frac{C_F}{\rho}\sqrt{T}\right) + \mathbb{E}\left[(T_2 - T_1)\cdot\text{GL-OPT} - \sum_{t=T_1+1}^{T_2}\sum_{j\in[M]} V_j(\rho_{j,t})\middle|\;\mathcal{E}^c\right]\\ &\overset{(iii)}{\le}\; \mathcal{O}(T^{2/3})\,.\end{aligned}$$

Here, (i) follows from Lemma 4.7 s.t. $\mathbb{P}(\mathcal{E}) \le \frac{1}{T}$; in (ii) we used the fact that under event $\mathcal{E}^c$ we have $T_1 \le 2\sqrt{T}\log^3(T)$ and also Eq. (69) that bounds $T - T_2$; in (iii), the term $(T_2 - T_1)\cdot\text{GL-OPT} - \sum_{t=T_1+1}^{T_2}\sum_{j\in[M]} V_j(\rho_{j,t})$ represents the convergence loss for UCB-SGD in Algorithm 1, which is in the order of $\mathcal{O}(T^{2/3})$ according to Theorem 4.6.

Finally, using concavity of $V_j(\rho_j)$ as illustrated in Lemma 4.3, we have $\text{GL-OPT} - \mathbb{E}\left[\sum_{j\in[M]} V_j(\bar{\rho}_j)\right] \le \text{GL-OPT} - \frac{1}{T}\mathbb{E}\left[\sum_{t\in[T_2]}\sum_{j\in[M]} V_j(\rho_{j,t})\right] \le \mathcal{O}(-T^{1/3})$ □

B.9. Additional Results for Section 4

Assumption 4.2 ensures that for any realization $\boldsymbol{z} = (\boldsymbol{v}, \boldsymbol{d})$ there must be some per-channel budget allocation that allows the advertiser to satisfy her ROI constraints as illustrated in the following proposition

Proposition B.1 (Slater's condition) *Assume Assumption 4.2 holds. Let* $\boldsymbol{z} = (\boldsymbol{v}, \boldsymbol{d}) = (\boldsymbol{v}_j, \boldsymbol{d}_j)_{j\in[M]} \in F_1 \times \dots F_M$ *be any realization of values and costs across all channels, then there exists some per-channel budget allocation* $\rho_{(\boldsymbol{z})} \in [0, \rho]^M$ *s.t.* $\sum_{j\in M} V_j(\rho_{(\boldsymbol{z}),j}; \boldsymbol{z}) > \gamma \sum_{j\in M} \rho_{(\boldsymbol{z}),j}$ *and* $\sum_{j\in[M]} \rho_{(\boldsymbol{z}),j} < \rho$.

Proof. Under Assumption 4.2, it is easy to see for any realization of value-cost pairs $\boldsymbol{z} = (\boldsymbol{v}_j, \boldsymbol{d}_j)_{j\in[M]}$ there always exists a channel $j \in [M]$ in which there is an auction $n \in [m_j]$ whose value-to-cost ration is at least γ, i.e. $v_{j,n} > \gamma d_{j,n}$. Since we assumed the ordering $\frac{v_{j,1}}{d_{j,1}} > \frac{v_{j,2}}{d_{j,2}} > \dots > \frac{v_{j,m_j}}{d_{j,m_j}}$, we know that $\frac{v_{j,1}}{d_{j,1}} \geq \frac{v_{j,n}}{d_{j,n}} > \gamma$. Now, in Eq. (30) within the proof of Lemma 4.3, we showed

$$V_j(\rho_j; \boldsymbol{z}_j) = \boldsymbol{v}_j^\top \boldsymbol{x}_j^*(\rho_j; \boldsymbol{z}_j) = \sum_{n\in[m_j]} \left(\frac{v_{j,n}}{d_{j,n}}\rho_j + b_{j,n}\right) \mathbb{I}\left\{d_{j,0} + \dots + d_{j,n-1} \leq \rho_j \leq d_{j,0} + \dots + d_{j,n}\right\}$$

where $b_{j,n} = \sum_{n'\in[n-1]} v_{j,n'} - \frac{v_{j,n}}{d_{j,n}} \cdot \left(\sum_{n'\in[n-1]} d_{j,n'}\right)$ and $d_{j,0} = v_{j,0} = 0$. Hence by taking $\rho_j = \underline{\rho}$ for some $\underline{\rho} < \min\{d_{j,1}, \rho\}$, we have $V_j(\underline{\rho}; \boldsymbol{z}_j) = \frac{v_{j,1}}{d_{j,1}}\underline{\rho} > \gamma\underline{\rho}$. Therefore, constructing $\rho_{(\boldsymbol{z}),j} = (0, \dots, 0, \rho_j, 0 \dots 0)$ for $\rho_j = \underline{\rho}$ satisfies $\sum_{j\in M} V_j(\rho_{(\boldsymbol{z}),j}; \boldsymbol{z}) > \gamma \sum_{j\in M} \rho_{(\boldsymbol{z}),j}$ and $\sum_{j\in[M]} \rho_{(\boldsymbol{z}),j} < \rho$. □

Lemma B.2 (Approximate constraint satisfaction) *Assume Assumption 4.2 holds. If Algorithm 1 is run with stepsize* $\eta = 1/\sqrt{T}$ *such that* $\eta < \frac{1}{M \max\{V, \rho, V^2\}}$, *then we have*

$$\frac{1}{T}\sum_{t\in[T]} g_{1,t} \geq -C_F/\sqrt{T} \text{ and } \frac{1}{T}\sum_{t\in[T]} g_{2,t} \geq -C_F/\sqrt{T}$$

where we recall $g_{1,t} = \sum_{j\in[M]} (V_j(\rho_{j,t}; \boldsymbol{z}_t) - \gamma\rho_{j,t})$, $g_{2,t} = \rho - \sum_{j\in[M]} \rho_{j,t}$, *and* $C_F > 0$ *is an absolute constant defined in Lemma 4.4 that depends only on the support* $F = F_1 \times \dots F_M$.

Proof. According to the dual update step in Algorithm 1 wee have

$$\lambda_{t+1} = (\lambda_t - \eta g_{1,t})_+ \geq \lambda_t - \eta g_{1,t}$$

Summing from $t = 1 \dots T-1$ and telescoping, we get

$$\lambda_T \geq \lambda_1 - \eta \sum_{t\in[T]} g_{1,t} = -\eta \sum_{t\in[T]} g_{1,t}$$

From Lemma 4.4, we have $\lambda_T \leq C_F$ for some absolute constant $C_F > 0$, therefore we have

$$\sum_{t\in[T]} g_{1,t} \geq -\frac{C_F}{\eta}$$

since we take $\eta = 1/\sqrt{T}$. A similar bound holds for $\sum_{t\in[T]} g_{2,t}$. □

Lemma B.3 *Let* $(\lambda_t, \mu_t)_{t\in[T]}$ *be the dual variables generated by Algorithm 1. Then for any* $\lambda, \mu \geq 0$ *we have*

$$\begin{aligned} \sum_t (\lambda_t - \lambda)\, g_{1,t} &\leq \frac{\eta M^2 \bar{V}^2}{2} \cdot T + \frac{1}{2\eta}(\lambda - \lambda_1)^2 \\ \sum_t (\mu_t - \mu)\, g_{2,t} &\leq \frac{\eta \rho^2}{2} \cdot T + \frac{1}{2\eta}(\mu - \mu_1)^2 . \end{aligned} \tag{70}$$

where we recall $g_{1,t} = \sum_{j\in[M]} (V_{j,t}(\rho_{j,t}) - \gamma\rho_{j,t})$ *and* $g_{2,t} = \rho - \sum_{j\in[M]} \rho_{j,t}$.

Proof. It only remains to prove Eq. (70). Starting with the first inequality w.r.t. λ_t's, we have

$$(\lambda_t - \lambda)\, g_{1,t} = (\lambda_{t+1} - \lambda)\, g_{1,t} + (\lambda_t - \lambda_{t+1})\, g_{1,t} \tag{71}$$

Since $\lambda_{t+1} = (\lambda_t - \eta g_{1,t})_+ = \arg\min_{\lambda \geq 0} (\lambda - (\lambda_t - \eta g_{1,t}))^2$, we have

$$(\lambda_{t+1} - (\lambda_t - \eta g_{1,t})) \cdot (\lambda - \lambda_{t+1}) \geq 0\,.$$

So we have

$$\begin{aligned}(\lambda_{t+1} - \lambda)\, g_{1,t} &\leq \frac{1}{\eta}(\lambda_{t+1} - \lambda_t) \cdot (\lambda - \lambda_{t+1}) \\ &= \frac{1}{2\eta}\left((\lambda - \lambda_t)^2 - (\lambda - \lambda_{t+1})^2 - (\lambda_{t+1} - \lambda_t)^2\right)\end{aligned}$$

Plugging the above back into Eq. (71) we get

$$\begin{aligned}(\lambda_t - \lambda)\, g_{1,t} &\leq (\lambda_t - \lambda_{t+1})\, g_{1,t} + \frac{1}{2\eta}\left((\lambda - \lambda_t)^2 - (\lambda - \lambda_{t+1})^2 - (\lambda_{t+1} - \lambda_t)^2\right) \\ &\leq \frac{\eta}{2} g_{1,t}^2 + \frac{1}{2\eta}\left((\lambda - \lambda_t)^2 - (\lambda - \lambda_{t+1})^2\right) \\ &\leq \frac{\eta M^2 \bar{V}^2}{2} + \frac{1}{2\eta}\left((\lambda - \lambda_t)^2 - (\lambda - \lambda_{t+1})^2\right)\,,\end{aligned} \tag{72}$$

where the final inequality follows from the fact that $V_{j,t}(\rho_{j,t}) \leq \bar{V}$ for any $j \in [M]$ and $t \in [T]$ so $g_{1,t} \leq M\bar{V}$. Summing the above over $t = 1 \ldots T$ and telescoping we get

$$\sum_t (\lambda_t - \lambda)\, g_{1,t} \leq \frac{\eta M^2 \bar{V}^2}{2} \cdot T + \frac{1}{2\eta}(\lambda - \lambda_1)^2$$

Following the same arguments above we can show

$$\sum_t (\mu_t - \mu)\, g_{2,t} \leq \frac{\eta \rho^2}{2} \cdot T + \frac{1}{2\eta}(\mu - \mu_1)^2$$

□

Proposition B.4 *Under Assumption 4.2, the advertiser's per-channel only budget optimization problem, namely* CH-OPT($\mathcal{I}_B$) *is a convex problem.*

Proof. Recalling the CH-OPT($\mathcal{I}_B$) in Eq. (3) and the definition of $\mathcal{I}_B$ in Eq. (2), we can write CH-OPT($\mathcal{I}_B$) as

$$\begin{aligned}\text{CH-OPT}(\mathcal{I}_B) = \max_{(\rho_j)_{j\in[M]} \in \mathcal{I}} \quad & \sum_{j \in M} V_j(\rho_j) \\ \text{s.t.} \quad & \sum_{j \in M} V_j(\rho_j) \geq \gamma \sum_{j \in M} \rho_j \\ & \sum_{j \in [M]} \rho_j \leq \rho\,,\end{aligned} \tag{73}$$

Here we used the definition $V_j(\rho_j) = \mathbb{E}\left[V_j(\rho_j; \boldsymbol{z}_j)\right]$ in Eq. (5), and $D_j(\rho_j; \boldsymbol{z}_j) = \rho_j$ for any $\boldsymbol{z}_j$ under Assumption 4.2. According to Lemma 4.3, $V_j(\rho_j)$ is concave in ρ_j for any j, so the objective of CH-OPT($\mathcal{I}_B$) maximizes a concave function. For the feasibility region, assume ρ_j and ρ_j' are feasible, then defining $\rho_j'' = \theta\rho_j + (1-\theta)\rho_j'$ for any $\theta \in [0,1]$, we know that

$$\begin{aligned}\sum_{j\in M}\left(V_j(\rho_j'') - \gamma\rho_j''\right) &\overset{(i)}{\geq} \sum_{j\in M}\left(\theta V_j(\rho_j) + (1-\theta)V_j(\rho_j') - \gamma\rho_j''\right)\\ &= \theta\sum_{j\in M}\left(V_j(\rho_j) - \gamma\rho_j\right) + (1-\theta)\sum_{j\in M}\left(V_j(\rho_j') - \gamma\rho_j'\right)\\ &\overset{(ii)}{\geq} 0\end{aligned}$$

where (i) follows from concavity of $V_j(\rho_j)$ and (ii) follows from feasiblity of ρ_j and ρ_j'. On the other hand it is apparent that $\sum_{j\in[M]}\rho_j'' \leq \rho$. Hence we conclude that for any ρ_j and ρ_j' feasible, $\rho_j'' = \theta\rho_j + (1-\theta)\rho_j'$ is also feasible, so the feasible region of CH-OPT($\mathcal{I}_B$) is convex. This concludes the statement of the proposition. □

Claim B.1 *Assume Assumption 4.3 holds, then for any channel $j \in [M]$ and value-cost realization $\boldsymbol{z}_j = (\boldsymbol{v}_j, \boldsymbol{d}_j) \in F_j$, we have $\frac{v_{j,1}}{d_{j,1}} > \gamma$. This further implies that $\xi := \min_{j\in[M]}\min_{\boldsymbol{z}_j\in F_j}\frac{v_{j,1}}{d_{j,1}} > \gamma$. Further, let $\beta = \frac{1}{\log(T)}$. Then, for large enough T we have $V_j(\beta;\boldsymbol{z}_j) = \frac{v_{j,1}}{d_{j,1}}\beta \geq \xi\beta$ for any realization $\boldsymbol{z}_j \in F_j$.*

Proof. Under Assumption 4.3, we know that for any realization of value-cost pairs $\boldsymbol{z}_j = (\boldsymbol{v}_j, \boldsymbol{d}_j)$ there always exists an auction $n \in [m_j]$ in this channel whose value-to-cost ratio is at least γ, i.e. $v_{j,n} > \gamma d_{j,n}$. Under the ordering $\frac{v_{j,1}}{d_{j,1}} > \frac{v_{j,2}}{d_{j,2}} > \cdots > \frac{v_{j,m_j}}{d_{j,m_j}}$, we have $\frac{v_{j,1}}{d_{j,1}} > \gamma$. In Eq. (30) within the proof of Lemma 4.3, we showed that for any realization $\boldsymbol{z}_j \in F_j$, $V_j(\rho_j;\boldsymbol{z}_j) = \frac{v_{j,1}}{d_{j,1}}\rho_j$ for all $\rho_j \leq d_{j,1}$. Hence we know that when T is large enough such that $\beta = \frac{1}{\log(T)} < \min_{j\in[M]}\min_{\boldsymbol{z}_j\in F_j} d_{j,1}$, we always have $V_j(\beta;\boldsymbol{z}_j) = \frac{v_{j,1}}{d_{j,1}}\beta \geq \xi\beta$. □

Appendix C: Proofs for Section 5

C.1. Proof of Lemma 5.2

Before we prove the lemma, we first show the following claim is true:

Claim C.1 *If auction n in channel j has increasing marginal values, i.e. for any realization $\boldsymbol{z}_j = (\boldsymbol{v}_j, \boldsymbol{d}_j)$, for any $n \in [m_j]$ we have $\frac{v_{j,n}(\ell-1) - v_{j,n}(\ell)}{d_{j,n}(\ell-1) - d_{j,n}(\ell)}$ decreases in ℓ, then $\frac{v_{j,n}(\ell)}{d_{j,n}(\ell)}$ also decreases in ℓ.*

Proof. We prove this claim by induction. The base case is $\ell = L_{j,n}$: it is easy to see

$$\frac{v_{j,n}(L_{j,n}-1) - v_{j,n}(L_{j,n})}{d_{j,n}(L_{j,n}-1) - d_{j,n}(L_{j,n})} > \frac{v_{j,n}(L_{j,n})}{d_{j,n}(L_{j,n})} \Longrightarrow \frac{v_{j,n}(L_{j,n}-1)}{d_{j,n}(L_{j,n}-1)} > \frac{v_{j,n}(L_{j,n})}{d_{j,n}(L_{j,n})}$$

Now assume the induction hypothesis $\frac{v_{j,n}(\ell)}{d_{j,n}(\ell)} > \frac{v_{j,n}(\ell+1)}{d_{j,n}(\ell+1)} > \cdots > \frac{v_{j,n}(L_{j,n})}{d_{j,n}(L_{j,n})}$. Then, we have

$$\begin{aligned}\frac{v_{j,n}(\ell)}{d_{j,n}(\ell)} > \frac{v_{j,n}(\ell+1)}{d_{j,n}(\ell+1)} &\Longrightarrow \frac{d_{j,n}(\ell+1) - d_{j,n}(\ell)}{d_{j,n}(\ell)} > \frac{v_{j,n}(\ell+1) - v_{j,n}(\ell)}{v_{j,n}(\ell)}\\ &\Longrightarrow \frac{d_{j,n}(\ell) - d_{j,n}(\ell+1)}{d_{j,n}(\ell)} < \frac{v_{j,n}(\ell) - v_{j,n}(\ell+1)}{v_{j,n}(\ell)}\\ &\Longrightarrow \frac{v_{j,n}(\ell)}{d_{j,n}(\ell)} < \frac{v_{j,n}(\ell) - v_{j,n}(\ell+1)}{d_{j,n}(\ell) - d_{j,n}(\ell+1)}\,.\end{aligned} \tag{74}$$

On the other since $\frac{v_{j,n}(\ell-1)-v_{j,n}(\ell)}{d_{j,n}(\ell-1)-d_{j,n}(\ell)}$ decreases in ℓ we have

$$\begin{aligned}&\frac{v_{j,n}(\ell-1)-v_{j,n}(\ell)}{d_{j,n}(\ell-1)-d_{j,n}(\ell)} > \frac{v_{j,n}(\ell)-v_{j,n}(\ell+1)}{d_{j,n}(\ell)-d_{j,n}(\ell+1)} \overset{(i)}{>} \frac{v_{j,n}(\ell)}{d_{j,n}(\ell)}\\ \Longrightarrow &\frac{v_{j,n}(\ell-1)}{d_{j,n}(\ell-1)} \overset{(ii)}{>} \frac{v_{j,n}(\ell)}{d_{j,n}(\ell)}\end{aligned}$$

where (i) follows from Eq. (74), and (ii) follows from the fact that $\frac{A}{B} > \frac{C}{D}$ for $A,B,C,D>0$ implies $\frac{A+C}{B+D} > \frac{C}{D}$ where we let $A = v_{j,n}(\ell-1) - v_{j,n}(\ell)$, $B = d_{j,n}(\ell-1) - d_{j,n}(\ell)$, $C = v_{j,n}(\ell)$ and $D = d_{j,n}(\ell)$. This concludes the proof. $\square$

Now we prove Lemma C.1. Similar to the proof of Lemma 4.3, we only need to show for any realization $\boldsymbol{z}_j = (\boldsymbol{v}_j, \boldsymbol{d}_j)_{j\in[M]}$, the conversion function $V_j^+(\rho_j;\boldsymbol{z}_j) = \boldsymbol{v}_j^\top \boldsymbol{x}_j^{*,+}(\rho_j;\boldsymbol{z}_j)$ where $\boldsymbol{x}_j^{*,+}(\rho_j;\boldsymbol{z}_j)$ is defined as Eq. (21) is piecewise linear, continuous, strictly increasing and concave.

For simplicity we use the shorthand notation $\boldsymbol{x}_j^* = \boldsymbol{x}_j^{*,+}(\rho_j;\boldsymbol{z}_j) \in [0,1]^{\sum_{n\in[m_j]} L_{j,n}}$ as the optimal solution to $V_j^+(\rho_j;\boldsymbol{z}_j)$, defined in Eq. (21). We first argue that in any auction, no impression ranked bellow the first would get procured, i.e. $x_{j,n}^*(\ell) = 0$ for any $\ell \in 2\dots L_{j,n}$ by contradiction. Assume there exists some auction $n\in[m_j]$ and impression slot $\ell' \in 2\dots L_{j,n}$ such that $x_{j,n}^*(\ell') > 0$, then by the constraint that at most 1 impression can be procured, i.e. $\sum_{\ell\in[L_{j,n}]} x_{j,n}^*(\ell) \le 1$ in Eq. (21), we know that $x_{j,n}^*(1) < 1$. Also, note that $x_{j,n}^*(\ell')$ incurs a cost of $d_{j,n}(\ell')\cdot x_{j,n}^*(\ell')$ amongst the total per-channel budget ρ_j. If we instead use this cost on the first impression, then we will obtain a value increase of

$$v_{j,n}(1)\cdot\frac{d_{j,n}(\ell')\cdot x_{j,n}^*(\ell')}{d_{j,n}(1)} - v_{j,n}(\ell')\cdot x_{j,n}^*(\ell') = d_{j,n}(\ell')\cdot x_{j,n}^*(\ell')\cdot\left(\frac{v_{j,n}(1)}{d_{j,n}(1)} - \frac{v_{j,n}(\ell')}{d_{j,n}(\ell')}\right) > 0,$$

where the final inequality follows from the assumption that $x_{j,n}^*(\ell') > 0$, and the multi-item auction has increasing marginal values (see Definition 5.1) so Claim C.1 holds. This contradicts the optimality of $\boldsymbol{x}_j^*$, and hence $x_{j,n}^*(\ell) = 0$ for any $\ell\in 2\dots L_{j,n}$, or in other words, a channel will only procure impressions ranked first. Hence, a channel's procurement problem in Eq. (21) can be restricted to the first impression in each auction, and thus similar to the proof of Lemma 4.3, is an LP-relaxation to the 0-1 knapsack with budget ρ_j, and m_j items whose values are $v_{j,1}(1)\dots v_{j,m_j}(1)$ with costs $d_{j,1}(1)\dots d_{j,m_j}(1)$. By re-labeling the auction indices in channel $j\in[M]$ such that $\frac{v_{j,1}(1)}{d_{j,1}(1)} > \frac{v_{j,2}(1)}{d_{j,2}(1)} > \dots > \frac{v_{j,m_j}(1)}{d_{j,m_j}(1)}$, the optimal solution takes the form

$$x_{j,n}^*(\ell) = \begin{cases} 1 & \text{if } \ell = 1 \text{ and } \sum_{n'\in[n]} d_{j,n'}(1) \le \rho_j \\ \frac{\rho_j - \sum_{n'\in[n-1]} d_{j,n'}(1)}{d_{j,n}(1)} & \text{if } \ell = 1 \text{ and } \sum_{n'\in[n]} d_{j,n'}(1) > \rho_j \\ 0 & \text{otherwise} \end{cases}$$

which is analogous to that of Eq. (29) in the proof of Lemma 4.3. The rest of the proof follows exactly from that for Lemma 4.3. $\square$