

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

Accepted, *Management Science*, October 2025

In digital online advertising, advertisers procure ad impressions simultaneously on multiple platforms, or so-called *channels*, such as Google Ads, Meta Ads Manager, etc. We study how an advertiser maximizes their 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 they participate in for each channel, and instead authorizes a channel to procure impressions on their behalf. The advertiser can only utilize two levers on each channel, namely setting a per-channel budget and a 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, their total conversion can be arbitrarily worse than what they could have obtained in the global problem. Further, we show that the advertiser can achieve the global optimal conversion when they only optimize 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 channel 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 conduct numerical studies to demonstrate that our proposed algorithm accurately approximates optimal per-channel budgets in practical setups.

Key words: Online advertising, autobidding, multi-channel ad procurement, return-on-investment, budget management, ad campaign management, bandit learning

1. Introduction

In today’s world of online advertising, 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, and web publisher displays, which present a multitude 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 conversions (e.g., ad clicks) or *return-on-investment* (ROI) with the same amount of investment? (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 that is relatively well-informed through attribution research, indicating that Google Ads and Meta ads are the two most effective channels for its products. The business instantiates its ad campaigns by interacting with the platforms’ ad management interfaces (see Figure 1), on which the business utilizes levers such as specifying budget and a target ROI¹ to control campaigns. Channels then input these specified parameters into their *autobidding* procedures, where they procure impressions on the advertiser’s behalf through automated black-box 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 work contributes to filling this gap by addressing two themes of practical significance:

How effective are these channel levers in helping advertisers 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 conversions across 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. Therefore, there are no realistic means for an advertiser to implement the global optimization problem by optimizing over individual auctions. Instead, advertisers can only use two levers: a per-channel ROI and a per-channel budget, to influence how channels should autobid for impressions. Our goal is to understand how effective these levers are by comparing the total conversion via optimizing levers with the globally optimal conversion, and also to present methodologies that help advertisers optimize the usage of these levers. We summarize our contributions as follows:

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

The figure displays two side-by-side screenshots of advertising campaign creation interfaces. The left interface is for Google Ads, featuring a 'Budget' section with a dropdown for 'US Dollar (USD \$)' and a text input for '\$100.00'. Below this is a 'Bidding' section with a 'Select your bid strategy' dropdown set to 'Target CPA' and a 'Pay for' dropdown set to 'Interactions'. A 'Target CPA' text input is set to '\$1.00'. The 'Start and end dates' section shows a 'Start date' dropdown for 'Jan 23, 2023' and an 'End date' dropdown for 'Jan 25, 2023'. The right interface is for Meta Ads Manager, showing a 'Budget' section with a 'Daily Budget' dropdown set to '\$100.00' and a 'USD' currency selector. The 'Schedule' section includes 'Start date' and 'End date' dropdowns for 'Jan 23, 2023' and 'Jan 25, 2023', both with '12:00 AM' and 'Eastern Time' specified. The 'Optimization for ad delivery' section has a dropdown set to 'Landing Page Views'. The 'Cost per result goal' section has a text input set to '\$1.00'. A red box highlights a note: 'Meta will aim to get the most landing page views and try to keep the average cost around \$1.00. Some results may cost more and some may cost less.'

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 Eqs. (1) and (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 the per-channel ROI and budget levers to maximize total conversions across channels while respecting global ROI and budget constraints. 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 is 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 arbitrarily 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.* In Section 4, we develop an algorithm for learning the optimal per-channel budget profile 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. The algorithm augments [dual subgradient descent \(DSD\)](#) with the upper-confidence bound (UCB) design and outputs, within T iterations, a per-channel budget profile that enables advertisers to achieve $\mathcal{O}(T^{-1/3})$ approximation accuracy in total conversion, relative to the optimal

per-channel budget profile. Our approach connects to constrained convex optimization with uncertain constraints and bandit feedback under a "one-point estimation" regime. For further discussions, see Section 1.2 and Remark 4.2 in Section 4.

4. Extensions to multi-impression auctions. In Sections 5, we shed light on the applicability of our results to more general settings when auctions correspond to the sale of multiple items (ad spots).

5. Numerical studies. In Section 6, we conduct numerical studies to demonstrate that our proposed algorithm DSD-UCB accurately approximates optimal per-channel budgets in a stochastic setting, even with a relatively small number of data points. Additionally, we show that its performance degrades gracefully when channels do not optimally procure ads on advertisers' behalf. Reflecting on practical settings, we extend this analysis to a multi-agent setup, where multiple buyers compete with each other while channels implement pacing strategies on their behalf. Our empirical results demonstrate that DSD-UCB allows buyers to retain a large fraction of the hindsight global optimal value, showcasing its effectiveness in dynamic multi-agent environments.

1.2. Related works.

Here, we review literature related 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 models 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 the price of anarchy in Deng et al. (2021), Balseiro et al. (2021a), Deng et al. (2022b), Mehta (2022), as well as individual advertiser welfare in Deng et al. (2022a), and the impact of consumer data collection and sharing on welfare in Bergemann and Bonatti (2023). The autobidding model, under which buyers are usually value-maximizers, has also been compared to classic quasi-linear utility models in Balseiro et al. (2021b).

However, the autobidding model examined in these works assumes advertisers can directly optimize over individual auctions. In contrast, our work addresses a more realistic scenario, reflecting industry practices where advertisers must navigate using levers provided by channels, delegating the procurement of ads to these channels. Moreover, recent research, such as Alimohammadi et al. (2023), Feng et al. (2023), delves into whether advertisers have an incentive to misreport their target ROIs or budgets to a single autobidding platform. In contrast, our paper focuses on optimizing per-channel budget decisions across multiple channels, presenting a distinct perspective on this complex issue.

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. (2022b) 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. See also Ai et al. (2022), Conitzer et al. (2022), Susan et al. (2023), Wang et al. (2023), Liao et al. (2024) for some recent works that study pacing strategies under non-truthful auctions, such as first-price and generalized price auctions.

There has been a recent line of work that studies the setting where multiple budget- or ROI-constrained bidders run pacing-type algorithms and analyze time-average welfare guarantees among all bidders Gaitonde et al. (2022), Lucier et al. (2023), Fikioris and Tardos (2023). All of these works on budget and ROI management focus on bidding strategies in a single repeated auction, where advertisers’ decisions are bids 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 the channels’ black-box 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). In light of Lemma 4.3 in Section 4, our problem of interest is also constrained and convex. First, there has been a plethora of algorithms developed for deterministic constrained convex optimization under 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 Burmen 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 $\mathcal{O}(T^{-1/4})$ accuracy. In this work, our setting is more complex, as the optimizer (i.e., the advertiser) cannot tell whether the ROI constraint is satisfied (due to unknown value and cost distributions in each channel’s auctions). Yet our proposed algorithm can still achieve a superior $\mathcal{O}(T^{-1/3})$ accuracy.

Most relevant to this paper are the very recent works Usmanova et al. (2019), Nguyen and Balasubramanian (2022), which consider 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., unknown 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 with our $\mathcal{O}(T^{-1/3})$ accuracy, these works impose several assumptions that are stronger than the ones we consider. First, the objective and constraint functions are strongly smooth (i.e., the gradients are Lipschitz continuous), and Nguyen and Balasubramanian (2022) further assumes strong convexity. But in our work, our objectives and constraints are piecewise 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. We refer readers to Table 4.1 in Larson et al. (2019) for a survey on the best-known bounds under different one-point bandit feedback settings.

Bandits with Knapsack (BwK). Our work is also related to the literature on the BwK framework, introduced by Badanidiyuru et al. (2013, 2018), for online resource allocation under constraints. Regret-optimal algorithms for Stochastic BwK have been developed by Agrawal and Devanur (2014), Immorlica et al. (2022), with extensions to general resource settings Agrawal and Devanur (2019), contextual bandits Agrawal et al. (2016), Badanidiyuru et al. (2014), and frameworks addressing both stochastic and adversarial settings Castiglioni et al. (2022a).

In our problem, budgets can be viewed as resources, and similar to works on BwK, we also use a primal-dual framework to learn how to optimally allocate them. While such frameworks have been employed in prior work, our setting presents unique challenges. Specifically, the value obtained from a channel is governed by an optimization process, and the primal variables (budget parameters) are continuous. Although discretization is a common approach for handling continuous decision spaces, bounding the resulting error requires proving that the Lagrangian function—dependent on the optimization process—is well-behaved, specifically unimodal and piecewise linear, as we establish.

Additionally, our work addresses the challenge of handling continuous dual variables in a primal-dual framework, which could be of independent interest for analyzing such algorithms. Our algorithm establishes an optimal mapping between two-dimensional continuous dual variables and per-channel budget parameters. While discretizing the budget parameter is feasible, the same is not true for dual variables. To address this, we employ a UCB (Upper Confidence Bound) design in the primal step,

treating the two-dimensional continuous dual variables as contextual information. These contexts are not drawn from a stochastic process but arise from the interaction between the primal and dual algorithms. Typically, the regret of a UCB algorithm scales with $\sqrt{C}$, where C is the number of contexts. For continuous contexts, this would render the regret bound vacuous. To overcome this, we leverage a concept we term *complete cross-learning*. Once the primal reward (computed as the Lagrangian function evaluated at the current dual variables) is known for a specific action in one context, the reward can be inferred for the same action in any other context. This enables complete cross-learning, as defined in Balseiro et al. (2022a), ensuring that the regret of our UCB algorithm remains independent of the number of contexts. This critical property forms the cornerstone of our analysis.

2. Preliminaries

Advertisers' global optimization problem. Consider an advertiser running a digital ad campaign across $M \in \mathbb{N}$ 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 determined by the highest competing bid in the market, and in a posted price auction, the cost is simply the posted price by the seller of the impression. Let $\mathbf{v}_j = (v_{j,n})_{n \in [m_j]}$ and $\mathbf{d}_j = (d_{j,n})_{n \in [m_j]}$. We assume that $\mathbf{z}_j := (\mathbf{v}_j, \mathbf{d}_j)$ is sampled from some discrete distribution $\mathbf{p}_j$ 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:

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

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

Here, the decision variable $\mathbf{x}_j \in [0, 1]^{m_j}$ is a vector where $x_{j,n}$ denotes whether the impression in auction n for channel j is procured. We note that $\mathbf{x}$ depends on the realization of $\mathbf{z} = (\mathbf{v}_j, \mathbf{d}_j)_{j \in [M]}$ and is thus random. The ROI and budget constraints are taken in expectation because an advertiser procures impressions from a large number of auctions (as the number of auctions in each platform is typically very large), and thus only needs to satisfy the constraints on average. We also note that GL-OPT is a widely adopted formulation for autobidding practices in modern online advertising, representing 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 A, we discuss more general advertiser objectives, such as maximizing quasi-linear utility.

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

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: 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 the advertiser achieve procurement results that match 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 that $\rho > 0$ 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 $V_j(\gamma_j, \rho_j; \mathbf{z}_j) \geq 0$, as well as the total spend $D_j(\gamma_j, \rho_j; \mathbf{z}_j) \geq 0$ to the advertiser, where we recall that $\mathbf{z}_j = (\mathbf{v}_j, \mathbf{d}_j) \in \mathbb{R}^{m_j} \times \mathbb{R}^{m_j}$ is the vector of value-cost pairs in channel j , sampled from the discrete support F_j according to distribution $\mathbf{p}_j$; V_j and D_j will be further specified later.

As the advertiser has the freedom to choose whether to input per-channel target ROIs, 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

ROIs and budgets for each channel. These 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 is to maximize the total conversion of procured ad impressions by optimizing per-channel budgets and target ROIs, while adhering to the global ROI and budget constraints 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 across channels can be formulated as:

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

where the expectation is taken with respect to (w.r.t.) the randomness in $\mathbf{z}_j$. We note that for any channel $j \in [M]$, the number of auctions m_j and the distribution $\mathbf{p}_j$ are fixed and do not depend on the input parameters γ_j and ρ_j .

The functions (V_j, D_j) , which map per-channel target ROI and budgets γ_j, ρ_j to total conversion and expenditure, are determined by various factors, including but not limited to channel j 's autobidding algorithms deployed to procure ads on the advertiser's behalf, as well as the auction mechanisms that sell impressions. In this work, we study a general setup that closely mirrors industry practices. We assume that, on behalf of the advertiser, each channel aims to optimize conversion across all m_j auctions while adhering to the advertiser's inputs (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). Consequently, each channel j 's optimization problem can be written as:

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

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

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

Here, the expectation is taken with respect to the randomness in $\mathbf{z}_j = (\mathbf{v}_j, \mathbf{d}_j) \in \mathbb{R}_+^{m_j} \times \mathbb{R}_+^{m_j}$. We assume that for any (γ_j, ρ_j) and realization of $\mathbf{z}_j$, the total conversion $V_j(\gamma_j, \rho_j; \mathbf{z}_j)$ is bounded above by some constant $\bar{V} \in (0, \infty)$ almost surely. It is important to note that Eq. (5) assumes that channels are able to achieve optimal procurement performance. In Section 6, we conduct numerical studies to explore scenarios where channels do not optimally solve Eq. (4).

Additionally, in Eq. (5), it is assumed that $\mathbf{z}_j = (\mathbf{v}_j, \mathbf{d}_j)$ are drawn from a fixed distribution. In our numerical studies in Section 5.1, we relax this assumption by considering a multi-agent non-stochastic setting, where buyers' bids are modeled using standard learning algorithms. We evaluate our algorithm in such real-world-inspired ad procurement processes, where the channels use a pacing algorithm (or DSD from Balseiro and Gur (2019), Balseiro et al. (2022b)) to set the bids and participate in numerous auctions on the advertiser's behalf.

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 in helping advertisers achieve the globally optimal conversion GL-OPT while respecting global ROI and budget constraints? Specifically, 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, and the optimal GL-OPT?
2. Since, in reality, advertisers can only utilize the two per-channel levers offered by channels, how can they optimize per-channel target ROIs and budgets to solve for CH-OPT($\mathcal{I}$)?

In Section 3, we address the first question by determining the gap between CH-OPT($\mathcal{I}$) and GL-OPT for different advertiser options. In Section 4, we develop an efficient algorithm to learn the optimal 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 investigate whether 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 achieve GL-OPT, but the per-channel ROI-only option can yield conversions arbitrarily worse than GL-OPT for certain instances, even when there is no global budget constraint (i.e., $\rho = \infty$). This implies that the per-channel ROI lever is inadequate for helping advertisers achieve the globally optimal conversion, whereas the per-channel budget lever is effective in attaining 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 when 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 $GL-OPT \geq CH-OPT(\mathcal{I})$, where we recall the definitions of $GL-OPT$ and $CH-OPT$ in Eqs. (1) and (3), respectively.*

The proof of Lemma 3.1 is deferred to Appendix B.1. Given the theoretical upper bound $GL-OPT$, we are now interested in the gap between $GL-OPT$ and $CH-OPT(\mathcal{I})$ for the 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 $\frac{CH-OPT(\mathcal{I}_R)}{GL-OPT}$ approaches 0, implying that 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 each channel consists of a single auction. The advertiser has an 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 $\mathbf{z} = (\mathbf{v}_j, \mathbf{d}_j)_{j \in [M]}$ (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: $\lim_{X \rightarrow \infty} \frac{CH-OPT(\mathcal{I}_R)}{GL-OPT} = 0$.*

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

Proof of Theorem 3.2 Let $\tilde{\gamma} = (\tilde{\gamma}_1, \tilde{\gamma}_2)$ be the optimal solution to $CH-OPT(\mathcal{I}_R)$, and recall that under the option $\mathcal{I}_R$, we assume per-channel budgets are infinite. It is easy to see that $\tilde{\gamma}_1$ can be any arbitrary nonnegative number because the advertiser always wins auction 1. Additionally, we must have $\tilde{\gamma}_2 > \frac{X}{1+X}$.

To see this, assume $\tilde{\gamma}_2 \leq \frac{X}{1+X}$. In this case, the advertiser would win the auction in channel 2. However, under this assumption, the advertiser would win all auctions across the two channels and acquire a total value of $1 + 2X$, while incurring a total spend of $0 + 2(1 + X) = 2 + 2X$. This would violate the ROI constraint in $CH-OPT(\mathcal{I}_R)$ because $\frac{1+2X}{2+2X} < 1$. Therefore, the advertiser can only win auction 1, which implies $\tilde{\gamma}_2 > \frac{X}{1+X}$. Consequently, the optimal objective for $CH-OPT(\mathcal{I}_R)$ is 1.

On the other hand, it is straightforward to see that the optimal solution to $GL-OPT$ is to win auction 1 with probability 1 and auction 2 with probability $\frac{1}{2}$, yielding an optimal value of $1 + X$. Thus, $\frac{CH-OPT(\mathcal{I}_R)}{GL-OPT} = \frac{1}{1+X}$. Taking the limit as $X \rightarrow \infty$ yields the desired result.

□

We now provide intuition behind the inadequacy of the ROI-only option to maximize conversions. In the proof of Theorem 3.2, as observed, there are only two possible outcomes:

1. By setting $\tilde{\gamma}_2 > \frac{X}{1+X}$ and solving problem (4), the advertiser wins auction 2 with $\rho_2 = \infty$. However, as shown in the proof, this leads to a violation of the global ROI constraint.
2. By setting $\tilde{\gamma}_2 < \frac{X}{1+X}$ and solving problem (4), the advertiser avoids violating the global ROI constraint but is forced into a suboptimal solution—completely losing auction 2.

This demonstrates that under the option $\mathcal{I}_R$, it is impossible to achieve all possible outcomes, including the globally optimal solution to GL-OPT. Specifically, the constraint on the per-channel ROI forces the advertiser into a limited set of solutions, which can lead to arbitrarily poor performance. As a result, the ROI-only option does not provide the flexibility needed to reach the global optimum.

In contrast to the per-channel ROI only option, the budget-only option 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 is presented in Appendix B.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 shows that 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 an optimal solution $(\gamma_j, \rho_j)_{j \in [M]}$ to $CH-OPT(\mathcal{I}_G)$, such that $\gamma_j = 0$ for all $j \in [M]$.*

In light of Corollary 3.4, we observe that introducing per-channel ROI targets $(\gamma_j)_{j \in [M]}$ does not improve the optimal value of the problem: simply optimizing over per-channel budgets $(\rho_j)_{j \in [M]}$ suffices to satisfy both the ROI and budget constraints. Importantly, this does *not* mean the global ROI constraint is irrelevant; rather, the constraint must be satisfied, but there is no additional benefit from optimizing over ROI targets at the channel level. Therefore, for the rest of the paper, we fix $\gamma_j = 0$ for all $j \in [M]$, and omit γ_j from the notation. For example, we write $D_j(\rho_j; \mathbf{z}_j)$ and $D_j(\rho_j)$ instead of $D_j(\gamma_j, \rho_j; \mathbf{z}_j)$ and $D_j(\gamma_j, \rho_j)$. Equivalently, we focus on the per-channel budget-only option $\mathcal{I}_B$.

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

In this section, we develop an efficient algorithm to learn optimal per-channel budgets that optimize $\text{CH-OPT}(\mathcal{I}_B)$ 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 $\text{CH-OPT}(\mathcal{I}_B)$, and satisfies aggregate ROI and budget constraints, namely

$$\sum_{j \in M} V_j(\rho_j) \geq \gamma \sum_{j \in M} D_j(\rho_j), \quad \sum_{j \in [M]} D_j(\rho_j) \leq \rho,$$

where we recall $(V_j(\rho_j), D_j(\rho_j))$ are defined in Eq. (5).

The algorithm proceeds as follows: at the beginning of period $t \in [T]$, the advertiser sets per-channel budgets $(\rho_{j,t})_{j \in [M]}$, without observing values and costs $\mathbf{z}_t = (\mathbf{z}_{j,t})_{j \in [M]} = (\mathbf{v}_{j,t}, \mathbf{d}_{j,t})_{j \in [M]}$, where $(\mathbf{v}_{j,t}, \mathbf{d}_{j,t}) \in \mathbb{R}_+^{m_j} \times \mathbb{R}_+^{m_j}$ are sampled (independently in each period) from a finite support $F = F_1 \times \dots \times F_M$ according to discrete distributions $(\mathbf{p}_j)_{j \in [M]}$. Each channel j then takes as input $\rho_{j,t} \in [0, \rho]$ and procures ads on behalf of the advertiser, reporting the total realized conversion $V_j(\rho_{j,t}; \mathbf{z}_{j,t})$ as well as total spend $D_j(\rho_{j,t}; \mathbf{z}_{j,t})$ to the advertiser (see definitions in Eq. (5)). For simplicity, we assume that any realization $\mathbf{z}_j = (\mathbf{v}_j, \mathbf{d}_j) \in F_j$ admits an ordering $\frac{v_{j,1}}{d_{j,1}} > \dots > \frac{v_{j,m_j}}{d_{j,m_j}}$ for all channels $j \in [M]$.

Bandit feedback: We highlight that the advertiser receives *bandit feedback* from the channels, i.e., the advertiser only observes the numerical values $V_j(\rho_{j,t}; \mathbf{z}_{j,t})$ and $D_j(\rho_{j,t}; \mathbf{z}_{j,t})$, but does not observe $V_j(\rho'_j; \mathbf{z}'_j)$ and $D_j(\rho'_j; \mathbf{z}'_j)$ evaluated at any other per-channel budget $\rho'_j \neq \rho_{j,t}$ or realized value-cost pairs $\mathbf{z}'_j \neq \mathbf{z}_{j,t}$.

We also make two mild assumptions: In Assumption 4.1, we assume that each channel will deplete its input per-channel budgets. In Assumption 4.2, we assume that for any realization of value-cost pairs $\mathbf{z}_j = (\mathbf{v}_j, \mathbf{d}_j)$ in a channel $j \in [M]$, there always exists an auction $n \in [m_j]$ in this channel whose value-to-cost ratio is at least γ , i.e., $v_{j,n} \geq \gamma d_{j,n}$.

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

Assumption 4.1 holds in large online ad markets, where advertisers have moderate budgets and fully utilize them due to the abundance of ad impressions. In practice, bidders can estimate an upper bound on the budget for each channel, ensuring that if the budget is set below this bound, it will be fully depleted. This is not an issue for large channels with many auctions, but for smaller ones with fewer auctions, budget depletion may not always occur. However, by incorporating this upper bound in our framework, the algorithm and results remain valid even for smaller channels.

Assumption 4.2 (Strictly feasible global ROI constraints) *Fix any channel $j \in [M]$ and any realization of value-cost pairs $\mathbf{z}_j = (\mathbf{v}_j, \mathbf{d}_j) \in F_j$. Then, the channel's optimization problem in Eq. (4) is strictly feasible, i.e. the set $\{\mathbf{x}_j \in [0, 1]^{m_j} : \mathbf{v}_j^\top \mathbf{x}_j > \gamma \mathbf{d}_j^\top \mathbf{x}_j\}$ is nonempty.*

Assumption 4.2 ensures that the per-channel ROI constraint is strictly satisfied for every realization of value-cost pairs $\mathbf{z}_j = (\mathbf{v}_j, \mathbf{d}_j) \in F_j$. This assumption is not overly restrictive for several reasons. First, if ROI constraints are frequently violated, it poses both technical and practical challenges. Technically, violations impede the ability to accurately assess feedback from the channel, preventing us from establishing any meaningful results. Practically, while occasional violations are acceptable, frequent ones signal that the target ROI is likely too ambitious. Advertisers can adjust their target ROI in response to this feedback. Second, in real-world scenarios, the number of auctions m_j within a channel j is typically large. Given this, we expect Assumption 4.2 to hold with high probability, as the set $\{\mathbf{x}_j \in [0, 1]^{m_j} : \mathbf{v}_j^\top \mathbf{x}_j > \gamma \mathbf{d}_j^\top \mathbf{x}_j\}$ will generally be non-empty when m_j is large enough. In particular, Assumption 4.2 is satisfied whenever there exists $n \in [m_j]$ such that $v_{j,n} > \gamma d_{j,n}$.

4.1. Optimize per-channel budgets with DSD-UCB

Here, we describe our algorithm to solve for optimal per-channel budgets with respect to $\text{CH-OPT}(\mathcal{I}_B)$. Similar to most algorithms for constrained optimization, including those in the BwK framework discussed in Section 1.2, we take a dual subgradient descent (DSD) approach; see a comprehensive survey on dual descent methods in Bertsekas (2014). First, we consider the Lagrangian functions with respect to $\text{CH-OPT}(\mathcal{I}_B)$, where we let $\mathbf{c} = (\lambda, \mu) \in \mathbb{R}_+^2$ be the dual variables corresponding to the ROI and budget constraints, respectively:

$$\begin{aligned}\mathcal{L}_j(\rho_j, \mathbf{c}; \mathbf{z}_j) &= (1 + \lambda)V_j(\rho_j; \mathbf{z}_j) - (\lambda\gamma + \mu)\rho_j \\ \mathcal{L}_j(\rho_j, \mathbf{c}) &= \mathbb{E}[\mathcal{L}_j(\rho_j, \mathbf{c}; \mathbf{z}_j)] .\end{aligned}\tag{6}$$

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

$$\rho_{j,t} = \arg \max_{\rho_j \in [0, \rho]} \mathcal{L}_j(\rho_j, \mathbf{c}_t; \mathbf{z}_{j,t}).\tag{7}$$

Having observed the realized values $(V_j(\rho_{j,t}; \mathbf{z}_{j,t}))_{j \in [M]}$ (where spend is $(\rho_{j,t})_{j \in [M]}$ under Assumption 4.1), we calculate the current period violation in budget and ROI constraints, namely $g_{1,t} := \sum_{j \in M} (V_j(\rho_{j,t}; \mathbf{z}_{j,t}) - \gamma \rho_{j,t})$ and $g_{2,t} = \rho - \sum_{j \in [M]} \rho_{j,t}$. Next, we update the dual variables³ via

³ The dual variable λ reflects how tightly the ROI constraint is satisfied: it remains small when the ROI surplus (value minus γ -scaled cost) is large, and increases as this surplus shrinks. The dual variable μ captures the pressure from the total budget constraint. Though not directly observable, both variables are dynamically updated by the algorithm and can be monitored over time to understand which constraint is binding. In practice, these trajectories could help calibrate or adjust campaign-level parameters such as per-channel ROI targets or budget parameters based on historical trends.

$\Pi_{[0, C_F]}(\lambda_t - \eta g_{1,t})$ and $\mu_{t+1} = \Pi_{[0, C_F]}(\mu_t - \eta g_{2,t})$, where Π is the projection operator, η is some pre-specified step size, and C_F is some dual variable upper bound specified in Eq. (9).⁴

However, we cannot realistically determine the primal decisions by solving Eq. (7) since the function $\mathcal{L}_j(\cdot, \mathbf{c}_t; \mathbf{z}_{j,t})$ is unknown due to the bandit feedback structure. Therefore, we provide a modification to DSD to handle this issue. We briefly note that although bandit feedback prevents naively applying DSD to our problem, this may not be the case in other online advertising scenarios that involve relevant learning tasks, underlining the challenges of our problem; see the 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, 2022b), 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 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 DSD takes the special form $\mathcal{L}_j(b, \mu_t; \mathbf{z}_t) = (v_t - (1 + \mu_t)d_t)\mathbb{I}\{b_t \geq d_t\}$ 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 $\arg \max_{b \geq 0} \mathcal{L}_j(b, \mathbf{c}_t; \mathbf{z}_t) = \frac{v_t}{1 + \mu_t}$. So, even though Balseiro et al. (2017, 2022b), Feng et al. (2022) study DSD under bandit feedback, the special structures of their problem instances permit DSD to effectively optimize for primal decisions in each period. In contrast, Eq. (7) in our setting cannot be solved directly.*

To handle bandit feedback, we take a natural approach to augment DSD with the celebrated *upper-confidence bound (UCB)* algorithm; see the introduction to UCB and multi-arm bandits in Slivkins et al. (2019). In particular, we first discretize our per-channel budget decision set $[0, \rho]$ into granular "arms" separated by distance $\delta > 0$:

$$\mathcal{A}(\delta) = \{a_k\}_{k \in [K]} \text{ where } a_k = (k-1)\delta. \quad (8)$$

for $K := \lceil \rho/\delta \rceil + 1$. 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 $(\bar{V}_j(a_k))_{j \in [M]}$ of the conversions $(V_j(a_k))_{j \in [M]}$ as well as an upper confidence bound $\text{UCB}_{j,t}(a_k)$ 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 $j \in [M]$: $\rho_{j,t} = \arg \max_{a_k \in \mathcal{A}(\delta)} (1 + \lambda_t) (\bar{V}_{j,t}(a_k) + \text{UCB}_{j,t}(a_k)) - (\lambda_t \gamma + \mu_t) a_k$.

⁴ One can also employ more general mirror descent dual variable updates; see e.g. Balseiro et al. (2022b).

Finally, to ensure aggregate ROI and budget constraint satisfaction, we maintain variables that check ROI and budget balances, namely $S_{1,t}$ and $S_{2,t}$, to record the cumulative ROI and spend across all channels up until period t . When the ROI balance check $S_{1,t}$ is too negative, or the budget balance check is too large, we "stop" the algorithm and naively set some pre-defined small per-channel budget $\underline{\rho} \in (0, \rho)$ (later chosen in Theorem 4.6) during all periods after the "stopping time" denoted as τ_A . We remark that similar approaches to ensure constraint satisfaction have been introduced in, e.g., Balseiro et al. (2022b), Feng et al. (2022).

We summarize our algorithm, called DSD-UCB, in Algorithm 1.⁵

Algorithm 1 DSD-UCB

- 1: **Input:** Budget discretization set of arms $\mathcal{A}(\delta)$ defined in Eq.(8). 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 $\lambda_1 = \mu_1 = 0$. Set $\underline{\rho} \in (0, \rho/M)$, $\beta > 0$ and dual variable upper bound

$$C_F = M\bar{V} \max \left\{ \frac{1}{\beta \underline{\rho}}, \frac{1}{\rho - M\underline{\rho}} \right\}, \quad (9)$$

where $\bar{V} \geq \max_{j \in [M]} \max_{\rho_j \in [0, \rho]} \max_{\mathbf{z}_j \in F_j} V_j(\rho_j, \mathbf{z}_j)$ is the conversion upper bound.

- 2: Set initial constraint balance checks: $S_{1,t} = S_{2,t} = 0$ for $t = 1$, and start period counter $t = 1$.
3: **while** $t \leq T$ and $S_{1,t} - \gamma M \rho + \beta \underline{\rho}(T - t) \geq 0$ and $S_{2,t} + M \rho + M \underline{\rho}(T - t) \leq \rho T$ **do**
4: **Set per-channel budget.** For each channel $j \in [M]$: If $t \leq K$, set $\rho_{j,t} = a_t$. Else if $t > K$, set $\rho_{j,t} =$

$$\arg \max_{a_k \in \mathcal{A}(\delta)} \bar{V}_{j,t}(a_k) + \text{UCB}_{j,t}(a_k) - \frac{(\lambda_t \gamma + \mu_t) a_k}{1 + \lambda_t},$$

where $\text{UCB}_{j,t}(a_k) = \sqrt{\frac{2 \log(T)}{N_{j,t}(a_k)}}$.

- 5: Observe realized conversion $\{V_j(\rho_{j,t}; \mathbf{z}_{j,t})\}_{j \in [M]}$, and update for each arm $k \in [K]$ and channel $j \in [M]$

$$N_{j,t+1}(a_k) = N_{j,t}(a_k) + \mathbb{I}\{\rho_{j,t} = a_k\}, \quad \bar{V}_{j,t+1}(a_k) = \frac{N_{j,t}(a_k) \bar{V}_{j,t}(a_k) + V_j(\rho_{j,t}; \mathbf{z}_{j,t}) \mathbb{I}\{\rho_{j,t} = a_k\}}{N_{j,t+1}(a_k)}$$

- 6: **Update dual variables.** Calculate $g_{1,t} = \sum_{j \in [M]} (V_j(\rho_{j,t}; \mathbf{z}_{j,t}) - \gamma \rho_{j,t})$ and $g_{2,t} = \rho - \sum_{j \in [M]} \rho_{j,t}$. Then, set

$$\lambda_{t+1} = \Pi_{[0, C_F]}(\lambda_t - \eta g_{1,t}) \quad \text{and} \quad \mu_{t+1} = \Pi_{[0, C_F]}(\mu_t - \eta g_{2,t}). \quad (10)$$

- 7: Update balance check: $S_{1,t+1} = S_{1,t} + g_{1,t}$ and $S_{2,t+1} = S_{2,t} + \sum_{j \in [M]} \rho_{j,t}$.

- 8: Increment period counter $t \leftarrow t + 1$.

- 9: **end while**

- 10: Record $\tau_A = t - 1$ and for all $t = \tau_A + 1 \dots T$ set $\rho_{j,t} = \underline{\rho}$ for all $j \in [M]$.

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

4.2. Analyzing the DSD-UCB algorithm

In this subsection, we analyze the performance of DSD-UCB in Algorithm 1 and present accuracy guarantees on the final output $\bar{\rho}_T = \left(\frac{1}{T} \sum_{t \in [T]} \rho_{j,t} \right)_{j \in [M]}$.

⁵ There has been very recent works that combine DSD with adversarial bandit type algorithms such as EXP3 Castiglioni et al. (2022b, 2023), or with Thompson sampling which is another well-known algorithm for stochastic bandit problems (e.g. Ding et al. (2021)), and works that employ DSD in bandit problems (e.g. Han et al. (2021)). Yet to the best of our knowledge, our approach to integrate DSD with UCB is novel.

To show the result, we first decompose the cumulative loss over T periods, defined as

$$T \cdot \text{GL-OPT} - \mathbb{E} \left[\sum_{t \in [T]} \sum_{j \in [M]} V_j(\rho_{j,t}) \right],$$

into three main components: (1) the "stopping error," which arises when the while-loop condition is violated and a small per-channel budget $\underline{\rho}$ is naively set after the stopping time τ_A (see step 10); (2) the error induced by the UCB component of our algorithm; and (3) the error resulting from DSD, often interpreted as deviations from complementary slackness (see Proposition 4.1). This type of regret decomposition is standard in the bandit literature.

We then proceed to bound each component individually.

Proposition 4.1 (Regret decomposition) *For any channel $j \in [M]$ define $\rho_j^*(t) = \arg \max_{\rho_j \in [M]} \mathcal{L}_j(\rho_j; \mathbf{c}_t)$ to be the optimal per-channel budget w.r.t. dual variables $\mathbf{c}_t = (\lambda_t, \mu_t)_{t \in [T]}$. Then $T \cdot \text{GL-OPT} - \sum_{t \in [T]} \sum_{j \in [M]} V_j(\rho_{j,t})$ is bounded by*

$$\underbrace{M\bar{V}(T - \tau_A)}_{\text{Stopping error}} + \underbrace{\sum_{t \in [\tau_A]} (\lambda_t g_{1,t} + \mu_t g_{2,t})}_{\text{DSD complementary slackness deviations}} + \underbrace{\sum_{j \in [M]} \sum_{t \in [\tau_A]} \mathcal{L}_j(\rho_j^*(t), \mathbf{c}_t) - \mathcal{L}_j(\rho_{j,t}, \mathbf{c}_t)}_{\text{UCB error}}.$$

where $\tau_A \in [T]$ is defined in step 10 of Algorithm 1.

Recall the definitions of $g_{1,t}$ and $g_{2,t}$ from step 5 of Algorithm 1, and note that the conversion $V_j(\rho_j; \mathbf{z}_j)$ is almost surely bounded above by a constant $\bar{V} \in (0, \infty)$. We provide a bound on the stopping error, along with the DSD complementary slackness violations, in Lemma 4.2. This result follows standard analyses for DSD, and the proof can be found in Appendix C.2.

Lemma 4.2 (Bounding stopping error and complementary slackness deviations)

Assume Assumptions 4.1 and 4.2 hold. Recall $\eta > 0$ is the step size. Then we have $M\bar{V}(T - \tau_A) + \sum_{t \in [\tau_A]} (\lambda_t g_{1,t} + \mu_t g_{2,t}) \leq \mathcal{O}\left(\eta T + \frac{1}{\eta}\right)$.

Challenges in bounding UCB error due to adversarial contexts and continuum-arm discretization. Bounding our UCB error is significantly more challenging than in classic stochastic multi-arm bandit settings. First, our setup involves discretizing a continuum of arms, i.e., our discretization in Eq.(8) for $[0, \rho]$. Second, and more importantly, the dual variables $\{\mathbf{c}_t\}_{t \in [T]}$ represent *adversarial contexts* because they are updated via DSD, rather than being stochastically sampled from some well-behaved distribution. Consequently, the Lagrangian function $\mathcal{L}_j(a_k, \mathbf{c}_t; \mathbf{z}_t)$ can be viewed as a reward function that maps any arm-context pair $(a_k, \mathbf{c}_t)$ to (stochastic) payoffs. Both the continuum of arms and adversarial contexts have been well-known for making reward function estimation highly inefficient; see, e.g., discussions in Agrawal (1995), Agarwal et al. (2014). We

further elaborate on the specific challenges that adversarial contexts introduce:

1. Boundedness of rewards. In classic stochastic multi-arm bandits and UCB, losses in total rewards grow linearly with the magnitude of rewards. In our setting, the reward function, i.e., the Lagrangian function $\mathcal{L}_j(a_k, \mathbf{c}_t; \mathbf{z}_t)$, scales linearly with the magnitude of contexts (see Eq. (6)). Therefore, large contexts (i.e., large dual variables) may lead to large losses.

2. Context-dependent exploration-exploitation tradeoffs. The typical trade-off between arm exploration and exploitation in our setting depends on the particular values of the contexts (i.e., the dual variables). This means there may exist "bad" contexts that lead to poor trade-offs, requiring significantly more exploration to achieve accurate estimates of arm rewards than other "good" contexts. We elaborate further in Lemma 4.5 and the discussions thereof.

We first handle continuum arm discretization by showing that the specific form of the conversion functions $V(\rho_j; \mathbf{z})$ in Eq. (4) induces salient structures for the Lagrangian function. Namely, it is continuous, piecewise linear, concave, and unimodal⁶. We present the proof in Appendix C.3.

Lemma 4.3 (Structural properties) *For any channel $j \in [M]$:*

- *The conversion function $V_j(\rho_j)$ is continuous, piecewise linear, strictly increasing, and concave. In particular, $V_j(\rho_j)$ takes the form*

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

where $S_j \in \mathbb{N}$ and $(s_{j,n}, b_{j,n}, r_{j,n})_{n \in [S_j]}$ depend only on the support F_j and distribution $\mathbf{p}_j$ from which values and costs are sampled. These parameters satisfy $s_{j,1} > \dots > s_{j,S_j} > 0$ and $0 = r_{j,0} < r_{j,1} < \dots < r_{j,S_j} = \rho$, as well as $b_{j,n} \geq 0$ such that $s_{j,n} r_{j,n} + b_{j,n} = s_{j,n+1} r_{j,n} + b_{j,n+1}$ for all $n \in [S_j - 1]$, with $b_{j,1} = 0$. This implies that $V_j(\rho_j)$ is continuous in ρ_j .

- *For any dual variables $\mathbf{c} = (\lambda, \mu) \in \mathbb{R}_+^2$, $\mathcal{L}_j(\rho_j, \mathbf{c})$ defined in Eq. (6) is continuous, piece-wise linear, concave, and unimodal in ρ_j . In particular,*

$$\mathcal{L}_j(\rho_j, \mathbf{c}) = \sum_{n=1}^{S_j} (\sigma_{j,n}(\mathbf{c}) \rho_j + b'_{j,n}) \mathbb{I}\{\rho_j \in [r_{j,n-1}, r_{j,n}]\},$$

where the slope $\sigma_{j,n}(\mathbf{c}) = (1 + \lambda)s_{j,n} - (\mu + \gamma\lambda)$ and $b'_{j,n} = (1 + \lambda)b_{j,n}$. This implies $\arg \max_{\rho_j \geq 0} \mathcal{L}_j(\rho_j, \mathbf{c}) = \max\{r_{j,n} : n = 0, 1, \dots, S_j, \sigma_{j,n}(\mathbf{c}) \geq 0\}$.

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

⁶ A function $f: \mathbb{R} \rightarrow \mathbb{R}$ is unimodal if there exists a point y such that $f(y)$ strictly increases when $y \leq y^*$ and strictly decreases when $y \geq y^*$.

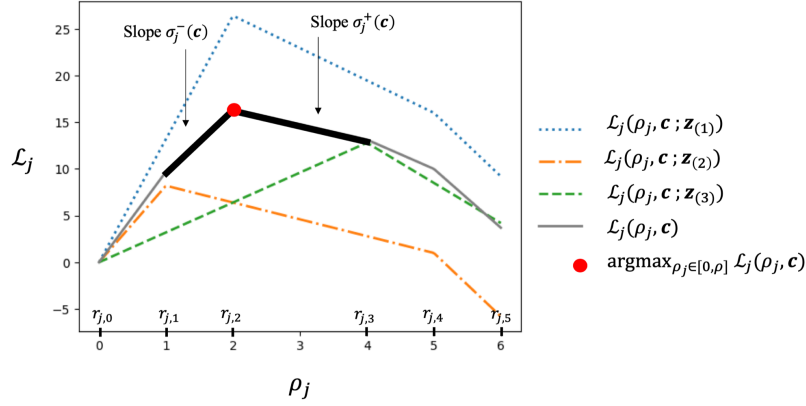


Figure 2 Illustration of Lagrangian functions defined in Eq. (6) with $M_j = 2$ auctions in channel j , and support F_j that contains 3 elements, $\mathbf{z}_{(1)} = ((8, 2), (2, 3))$, $\mathbf{z}_{(2)} = ((3, 4), (1, 4))$, $\mathbf{z}_{(3)} = ((8, 1), (4, 2))$, and context $\mathbf{c} = (\lambda, \mu) = (4, 2)$. Under Lemma 4.3, $S_j = 5$, where the “turning points” $r_{j,0} \dots r_{j,S_j}$ are indicated on the x-axis, and the optimal budget w.r.t. $\mathbf{c}$ is $\arg \max_{\rho_j \in [0, \rho]} \mathcal{L}_j(\rho_j; \mathbf{c}) = r_{j,2}$. The adjacent slopes in Eq. (11) are $\sigma_j^-(\mathbf{c}) = \sigma_{j,2}(\mathbf{c})$, and $\sigma_j^+(\mathbf{c}) = \sigma_{j,3}(\mathbf{c})$.

We now address the reward boundedness issue for the Lagrangian functions that arise from adversarial contexts. In Lemma 4.4 (proof in Appendix C.4), we show that the Lagrangian functions are bounded by some absolute constants:

Lemma 4.4 (Bounding Lagrangian functions) *For any $t \in [T]$, $j \in [M]$ and $\rho_j \in [0, \rho]$, we have $-(1 + \gamma)\rho C_F \leq \mathcal{L}_j(\rho_j, \mathbf{c}_t) \leq (1 + C_F)\bar{V}$, where the dual variables $\mathbf{c}_t = (\lambda_t, \mu_t)_{t \in [T]}$ are generated from Algorithm 1.*

We now address the context-dependent exploration-exploitation tradeoff. To illustrate (e.g., Figure 2), define the slopes adjacent to the optimal per-channel budget with respect to $\mathbf{c} = (\lambda, \mu)$ as follows: assume the n -th “turning point” $r_{j,n} = \arg \max_{\rho_j \in [0, \rho]} \mathcal{L}_j(\rho_j, \mathbf{c})$, then

$$\sigma_j^-(\mathbf{c}) = \sigma_{j,n}(\mathbf{c}) \quad \text{and} \quad \sigma_j^+(\mathbf{c}) = \sigma_{j,n+1}(\mathbf{c}). \quad (11)$$

Similar to standard exploration-exploitation tradeoffs in bandits, the flatter the slope (e.g., $\sigma_j^-(\mathbf{c})$ is close to 0), the more pulls are 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, or equivalently, the adjacent slopes $\sigma_j^-(\mathbf{c})$ and $\sigma_j^+(\mathbf{c})$, depends on the adversarial contexts. In Lemma 4.5, we bound the UCB error by handling this context-dependent tradeoff through separately analyzing periods when the adjacent slopes $\sigma_j^-(\mathbf{c})$ and $\sigma_j^+(\mathbf{c})$ are less than or greater than some parameter $\underline{\sigma} > 0$ chosen later, and characterize the context-dependent tradeoff using $\underline{\sigma}$.

Lemma 4.5 (Bounding UCB error) *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 the UCB error in Proposition 4.1 is upper-bounded by $\mathcal{O}\left(\delta T + \underline{\sigma} T + \frac{1}{\underline{\sigma}\delta}\right)$, where $\underline{\sigma} > 0$ is any positive number.*

See Appendix C.5 for the proof. Finally, we put together Proposition 4.1, Lemma 4.2, and Lemma 4.5, and obtain the main result, Theorem 4.6, whose proof we detail in Appendix C.6.

Theorem 4.6 (Putting everything together) *Assume that assumptions 4.1 and 4.2 hold. Take step size $\eta = \Theta(1/\sqrt{T})$, discretization width $\delta = \Theta(T^{-1/3})$, and $\beta = \underline{\rho} = \frac{1}{\log(T)}$ in Algorithm 1, as well as $\underline{\sigma} = \Theta(T^{-1/3})$ in Lemma 4.5. Then, for large enough T , we have*

$$T \cdot \text{GL-OPT} - \mathbb{E} \left[\sum_{t \in [T]} \sum_{j \in [M]} V_j(\rho_{j,t}) \right] \leq \mathcal{O}(T^{2/3}).$$

Further, recalling that $\bar{\mathbf{p}}_T$ is the final output of Algorithm 1, we have

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

and the constraint satisfaction:

$$\rho - \sum_{j \in [M]} \mathbb{E} [\bar{\rho}_{j,T}] \geq 0,$$

as well as

$$\sum_{j \in [M]} \mathbb{E} [V_j(\bar{\rho}_{j,T}) - \gamma \bar{\rho}_{j,T}] \geq 0.$$

We make an important remark that distinguishes our result in Theorem 4.6 with related literature on convex optimization. We refer the readers to Section 1.2 for a discussion on

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 C.4 in Appendix C.7). 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 $\sum_{j \in [M]} \mathbb{E} [V_j(\gamma_j, \rho_j; \mathbf{z}_j) - \gamma D_j(\gamma_j, \rho_j; \mathbf{z}_j)] \geq 0$ holds because the distribution $(\mathbf{p}_j)_{j \in [M]}$ from which $\mathbf{z}$ is sampled is unknown. To the best of our knowledge, within the online convex optimization literature, two recent works address a similar stochastic bandit feedback and uncertain constraint setting, namely Usmanova et al. (2019) and Nguyen and Balasubramanian (2022). However, our setting is more challenging because these works consider a two-point estimation regime, where function evaluations to the objective and constraints can be made twice per period, whereas we deal with a*

*one-point estimation regime, where function calls are limited to once per period. The optimal oracle complexities for unknown constraint convex optimization with one-point bandit feedback remain an open problem.*⁷

In light of Remark 4.2, we would like to point out that stochastic bandit feedback has been extensively studied in the BwK literature, including works such as Immorlica et al. (2022), Castiglioni et al. (2022a), and Slivkins et al. (2023). As discussed earlier in Section 1.2, we use a primal-dual framework for optimal budget allocation, similar to these works. However, our problem introduces unique challenges due to the continuous nature of the primal variables and the optimization process governing the value from a channel. We address these challenges by discretizing the continuous decision space and proving that the resulting Lagrangian function is well-behaved (unimodal and piecewise linear), as established in Lemma 4.3. Additionally, we propose a novel approach for handling continuous dual variables by employing UCB (Upper Confidence Bound) in the primal step. To avoid vacuous regret bounds that typically arise with continuous contexts (i.e., dual variables), we introduce the concept of complete cross-learning Balseiro et al. (2022a), which enables rewards from one context to be inferred for others. This ensures that the regret remains independent of the number of contexts, a key property in our analysis, as formalized in Lemma 4.5.

5. Generalizing to autobidding in multi-item auctions

In the previous sections, we assumed that each channel consists of multiple auctions, each associated with the sale of a single ad impression (see Eq. (4) and the accompanying discussion). However, in practice, there are many scenarios where ad platforms sell multiple impressions in each auction (see, e.g., Varian (2007), Edelman et al. (2007)). In this section, we extend all the results from the single-item auction setting 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 remains redundant (as shown in Theorem 3.3 and Corollary 3.4), and an advertiser can optimize solely over per-channel budgets to achieve globally optimal conversion. In Section 5.3, we demonstrate that our proposed DSD-UCB algorithm is directly applicable to the multi-item auction setup for a broad class of auctions. Similar to Theorem 4.6, our algorithm produces accurate lever estimates, enabling the advertiser to approximate the globally optimal lever decisions.

5.1. Multi-item autobidding setup

We first formalize our multi-item setup as follows. 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 the cost incurred by the advertiser when

⁷ See Table 4.1 in Larson et al. (2019) for the best-known complexity bounds for one-point bandit feedback setups.

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 $\mathbf{v}_{j,n} = (v_{j,n}(1), \dots, v_{j,n}(L_{j,n})) \in \mathbb{R}_+^{L_{j,n}}$ as the $L_{j,n}$ -dimensional vector that includes all impression values of auction n in channel j , and further write $\mathbf{v}_j = (\mathbf{v}_{j,1}, \dots, \mathbf{v}_{j,m_j}) \in \mathbb{R}_+^{\sum_{n \in [m_j]} L_{j,n}}$ as the vector that concatenates all value vectors across auctions in channel j . We also define $\mathbf{d}_{j,n} \in \mathbb{R}_+^{L_{j,n}}$ and $\mathbf{d}_j \in \mathbb{R}_+^{\sum_{n \in [m_j]} L_{j,n}}$ accordingly for costs. Similar to Section 2, we assume $\mathbf{z}_j = (\mathbf{v}_j, \mathbf{d}_j)$ is sampled from finite support F_j according to the discrete distribution $\mathbf{p}_j$ for any channel $j \in [M]$, and without loss of generality, we assume that for any element $\mathbf{z}_j \in F_j$, the values and costs for individual impressions in any auction $n \in [m_j]$ satisfy $v_{j,n}(1) > \dots > v_{j,n}(L_{j,n}) > 0$ and $d_{j,n}(1) > \dots > d_{j,n}(L_{j,n}) > 0$.

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⁺:

$$\begin{aligned}
\text{GL-OPT}^+ = & \max_{(\mathbf{x}_j = (\mathbf{x}_{j,1}, \dots, \mathbf{x}_{j,m_j}))_{j \in [M]}} \sum_{j \in [M]} \mathbb{E} [\mathbf{v}_j^\top \mathbf{x}_j] \\
& \text{s.t.} \quad \sum_{j \in [M]} \mathbb{E} [\mathbf{v}_j^\top \mathbf{x}_j] \geq \gamma \sum_{j \in [M]} \mathbb{E} [\mathbf{d}_j^\top \mathbf{x}_j] \\
& \quad \sum_{j \in [M]} \mathbb{E} [\mathbf{d}_j^\top \mathbf{x}_j] \leq \rho \\
& \quad \mathbf{x}_{j,n} \in [0, 1]^{\sum_{n \in [m_j]} L_{j,n}} \text{ and } \sum_{\ell \in [L_{j,n}]} x_{j,n}(\ell) \leq 1, \quad \forall j \in [M], n \in [m_j].
\end{aligned} \tag{12}$$

Here, $\mathbf{x}_{j,n} = (x_{j,n}(\ell))_{\ell \in [L_{j,n}]}$ 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 introduce additional constraints which state: "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:

$$\begin{aligned}
\mathbf{x}_j^{*,+}(\gamma_j, \rho_j; \mathbf{z}_j) = & \arg \max_{\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_{m_j})} \mathbf{v}_j^\top \mathbf{x} \\
& \text{s.t.} \quad \mathbf{v}_j^\top \mathbf{x} \geq \gamma_j \mathbf{d}_j^\top \mathbf{x}, \text{ and } \mathbf{d}_j^\top \mathbf{x} \leq \rho_j \\
& \quad \mathbf{x}_n \in [0, 1]^{L_{j,n}} \text{ and } \sum_{\ell \in [L_{j,n}]} x_n(\ell) \leq 1, \quad \forall n \in [m_j],
\end{aligned} \tag{13}$$

where $\mathbf{x}_n = (x_n(\ell))_{\ell \in [L_{j,n}]} \in [0, 1]^{m_j}$ denotes the (possibly random) vector of indicators for winning each impression in auction n of channel j . With respect to this per-channel multi-item auction optimization problem in Eq. (13), we further define $V_j^+(\gamma_j, \rho_j; \mathbf{z}_j)$, $D_j^+(\gamma_j, \rho_j; \mathbf{z}_j)$, $V_j^+(\gamma_j, \rho_j)$, and $D_j^+(\gamma_j, \rho_j)$ 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 Theorem 5.1, which demonstrates that an advertiser can achieve the global optimal conversion $GL-OPT^+$ by 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 when $\rho = \infty$. Further, there exists an optimal solution $(\gamma_j, \rho_j)_{j \in [M]}$ to $CH-OPT^+(\mathcal{I}_G)$, such that $\gamma_j = 0$ for all $j \in [M]$.*

It is easy to see that the proof of Lemma 3.1, Theorem 3.3, and Corollary 3.4 with respect to 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 $CH-OPT^+$). Therefore, 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 we omit the per-channel ROI γ_j when the context is clear.

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

We now turn to our second main focus in the multi-item setting: understanding whether our proposed DSD-UCB algorithm can achieve accurate approximations of the optimal per-channel budgets, similar to Theorem 4.6 in the single-item setting. A key observation is that the only difference between bounding the error of DSD-UCB in the single and multi-item settings lies in the structure of the conversion and corresponding Lagrangian functions (see Lemma 4.3). This is because 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. This will allow us to directly apply the proof for bounding the error of DSD-UCB (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 of conversion and Lagrangian functions from the single-item setting (as shown in Lemma 4.3), and is demonstrated 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 $\mathbf{z}_j = (\mathbf{v}_j, \mathbf{d}_j)$, we have

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

where we recall $v_{j,n}(1) > \dots > v_{j,n}(L_{j,n}) > 0$ and $d_{j,n}(1) > \dots > d_{j,n}(L_{j,n}) > 0$.

Increasing marginal values intuitively means that in a multi-item auction, the marginal value per cost increases as higher-value impressions are procured. Many classic position auction formats satisfy the increasing marginal gains property, such as the Vickrey–Clarke–Groves (VCG) auction. For more details on position auctions, see Varian (2007), Edelman et al. (2007).

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 $(\mathbf{v}_{j,n}, \mathbf{d}_{j,n}) = (v_{j,n}(\ell), d_{j,n}(\ell))_{\ell \in [L_{j,n}]}$, there exists some $\tilde{v}_{n,j} > 0$, position discounts $1 \geq \theta_{n,j}(1) > \theta_{n,j}(2) > \dots > \theta_{n,j}(L_{n,j}) > 0$, and the $L_{n,j}$ -highest competing bids from competitors in the market $\tilde{d}_{n,j}(1) > \tilde{d}_{n,j}(2) > \dots > \tilde{d}_{n,j}(L_{n,j}) > 0$, such that the acquired value for procuring impression $\ell \in L_{n,j}$ is $v_{n,j}(\ell) = \theta_{n,j}(\ell) \cdot \tilde{v}_{n,j}$, and the corresponding cost is

$$d_{j,n}(\ell) = \sum_{\ell'=\ell}^{L_{j,n}} (\theta_{n,j}(\ell') - \theta_{n,j}(\ell' + 1)) \tilde{d}_{n,j}(\ell'),$$

where we denote $\theta_{n,j}(L_{j,n} + 1) = 0$.⁸ Under VCG, the marginal values are given by

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

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

We remark that the generalized second-price auction (GSP) does not necessarily have increasing marginal values. However, if all auctions in a channel have increasing marginal values, then we can show that the conversion function $V_j^+(\rho_j)$ and the corresponding Lagrangian function for multi-item auctions admit 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 $V_j^+(\rho_j) = \mathbb{E}[\mathbf{v}^T \mathbf{x}_j^{*,+}(\rho_j; \mathbf{z}_j)]$ is continuous, piecewise linear, strictly increasing, and concave. Here, recall that

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

$\mathbf{x}_j^{*,+}(\rho_j; \mathbf{z}_j)$ is the optimal solution to the channel's optimization problem in Eq. (13). Further, for any dual variables $\mathbf{c} = (\lambda, \theta) \in \mathbb{R}_+^2$, the Lagrangian function $\mathcal{L}_j^+(\rho_j, \mathbf{c}) := (1 + \lambda)V_j^+(\rho_j) - (\theta + \gamma\lambda)\rho_j$ is continuous, piecewise linear, concave, and unimodal in ρ_j .

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

Theorem 5.3 (DSD-UCB applied to channel procurement for multi-item auctions)

Assume multi-item auctions in any channel $j \in [M]$ have 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 $\bar{\rho}_T = \left(\frac{1}{T} \sum_{t \in [T]} \rho_{j,t}\right)_{j \in [M]}$ is the vector of time-averaged per-channel budgets produced by DSD-UCB, we have

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

as well as constraint satisfaction

$$\sum_{j \in [M]} \mathbb{E}[V_j^+(\bar{\rho}_{T,j}) - \gamma \bar{\rho}_{T,j}] \geq 0, \quad \text{and} \quad \rho - \sum_{j \in [M]} \mathbb{E}[\bar{\rho}_{T,j}] \geq 0,$$

where we recall $GL-OPT^+$ is defined in Eq. (12), $V_j^+(\rho_j) = \mathbb{E}[\mathbf{v}_j^\top \mathbf{x}_j^{*,+}(\rho_j; \mathbf{z}_j)]$ and $\mathbf{x}_j^{*,+}(\rho_j; \mathbf{z}_j)$ is defined in Eq. (13).

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.

6. Numerical studies

In this section, we conduct numerical studies using synthetic data to showcase the performance of our proposed DSD-UCB algorithm in both a single-agent stochastic and multi-agent environment. In particular, in Section 6.1, we first consider the setting where we take the perspective of a single buyer who runs our proposed DSD-UCB algorithm in a hypothetical setting where both values and costs are stochastically generated for auctions in numerous channels. We demonstrate that under this setup, the final per-channel budget profile output of our proposed algorithm accurately approximates the optimal per-channel budget profile, even when the total number of periods T is moderately small. Further, we show that when channels do not optimally autobid on the advertisers' behalf (i.e.,

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 $\mathbf{z} = (\mathbf{v}_j, \mathbf{d}_j)_{j \in [M]} \in F_1 \times \dots \times F_M$, the realized version of the ROI constraint in $GL-OPT^+$ defined in Eq. (12) is strictly feasible.

channels do not optimally solve for Eq. (4)), the performance of our proposed algorithm deteriorates gracefully.

Next, in Section 6.2, we consider a multi-agent setup where multiple buyers compete with each other across multiple channels, while channels deploy standard budget management algorithms on the buyers' behalf to acquire ads (in particular, pacing strategies, which we will describe later in the section). We show that in this multi-agent environment, our proposed DSD-UCB algorithm allows buyers to retain a large fraction of the hindsight global optimal value.

6.1. Stochastic environment

Setup and data generation. We take the perspective of an advertiser with an aggregate budget $\rho = 10$ as well as an aggregate ROI $\gamma = 1.3$, and consider $M = 10$ channels, each of which consists of $m_j = 100$ auctions for $j \in [M]$. We conduct experiments over $L = 100$ independent trials, where each trial corresponds to a different support F for value-cost pairs, as well as a different distribution $\mathbf{p} \in \Delta(F)$ from which value-cost pairs are sampled. Here, $\Delta(F)$ is the probability simplex over F . In particular, we fix a value-cost pair support size of $f = 5000$. For each trial $\ell \in [L]$ and channel $j \in 1, \dots, 5$, we sample f values from $\text{Uniform}([0, 1]^{m_j})$ as well as f costs from $\text{Uniform}([0, 1]^{m_j})$, with which we construct the support $F_j^{(\ell)} \subset \mathbb{R}^{m_j} \times \mathbb{R}^{m_j}$ for channel j (note that the support size $|F_j^{(\ell)}| = f$). On the other hand, for any channel $j \in 6, \dots, 10$, we sample f values from $\text{Uniform}([0, 2]^{m_j})$ as well as f costs from $\text{Uniform}([0, 1]^{m_j})$ for channel j , with which we construct the support $F_j^{(\ell)} \subset \mathbb{R}^{m_j} \times \mathbb{R}^{m_j}$. Correspondingly, for any channel $j \in [M]$, we generate uniformly at random a probability distribution $\mathbf{p}^{(\ell)} \in \Delta(F^{(\ell)})$, where $F^{(\ell)} = F_1^{(\ell)} \times \dots \times F_M^{(\ell)}$. Finally, for trial $\ell \in [L]$, with respect to the support $F^{(\ell)}$ and distribution $\mathbf{p}^{(\ell)}$, we denote the associated optimal conversion defined in Eq. (1) as $\text{GL-OPT}^{(\ell)}$, as well as the expected per-channel conversion defined in Eq. (5) as $V_j^{(\ell)}(\rho_j)$.

We remark that it is not difficult to see that auctions in channels 6–10 have a higher value-to-cost ratio on average than those in channels 1–5. Therefore, we refer to channels 6–10 as the “lucrative” channels, and channels 1–5 as the “non-lucrative” channels. Further, for any trial, under the hindsight optimal per-channel budget profiles, lucrative channels consume approximately 80% of our total budget, whereas non-lucrative channels demand the remaining 20%.

Non-optimal autobidding. In previous sections, we assumed that each channel adopts “optimal autobidding” that solves Eq. (4) to optimality. This raises the natural question of whether our findings still hold when channels do not procure ads optimally, perhaps due to non-stationary environments (Besbes et al. 2014, Luo et al. 2018, Cheung et al. 2019, Chen et al. 2022), or the presence of strategic market participants aiming to manipulate the market (Golrezaei et al. 2023a, Drutsa 2020, Golrezaei et al. 2021, 2023b). To model non-optimal autobidding behavior (i.e., when a channel j does not optimally solve for $V_j(\rho_j; \mathbf{z}_j)$ in Eq. (4)), we introduce corruption factors

$(\alpha_1, \alpha_2) \in [0, 1]^2$ for non-lucrative and lucrative channels, respectively.¹⁰ These factors capture each channel's degree of sub-optimal autobidding: after submitting a per-channel budget ρ_j to channel $j \in [M]$, the channel will report back a conversion value $\tilde{V}_j(\rho_j; \mathbf{z}_j) = \alpha_1 V_j(\rho_j; \mathbf{z}_j)$ for $j \in \{1, \dots, 5\}$, or $\tilde{V}_j(\rho_j; \mathbf{z}_j) = \alpha_2 V_j(\rho_j; \mathbf{z}_j)$ for $j \in \{6, \dots, 10\}$, given any realized value-cost pairs $\mathbf{z}_j$. Consequently, in step 4 of the DSD-UCB algorithm, we observe $\tilde{V}_j(\rho_j; \mathbf{z}_t)$ instead of the optimal conversion $V_j(\rho_j; \mathbf{z}_t)$. In other words, we assume that channel j can only achieve a fraction of the optimal conversion, where $\alpha_1 = 1$ or $\alpha_2 = 1$ corresponds to optimal autobidding.

Experiment procedure. We repeat the following procedure for trial $\ell \in [L]$, corruption factors $\alpha_1, \alpha_2 \in \{0.2, 0.4, 0.6, 0.8, 1\}$, and $T \in \{100, 200, 500, 1000\}$. We run the DSD-UCB algorithm over T periods, where in each period $t \in [T]$, value-cost pairs for all $\sum_{j \in [M]} m_j = 1000$ auctions, namely $\mathbf{z}_t = (\mathbf{v}_t, \mathbf{d}_t) \in \mathbb{R}^{\sum_{j \in [M]} m_j} \times \mathbb{R}^{\sum_{j \in [M]} m_j}$, are sampled from support $F^{(\ell)}$ according to distribution $\mathbf{p}^{(\ell)}$. After submitting $(\rho_{j,t})_{j \in [M]}$ to the channels, we observe $\tilde{V}_j(\rho_j, \mathbf{z}_j) = \alpha_1 V_j(\rho_j, \mathbf{z}_j)$ for $j \in [M]$. We denote the final output per-channel budget profile as $\bar{\rho}_T^{(\ell, \alpha_1, \alpha_2)}$.

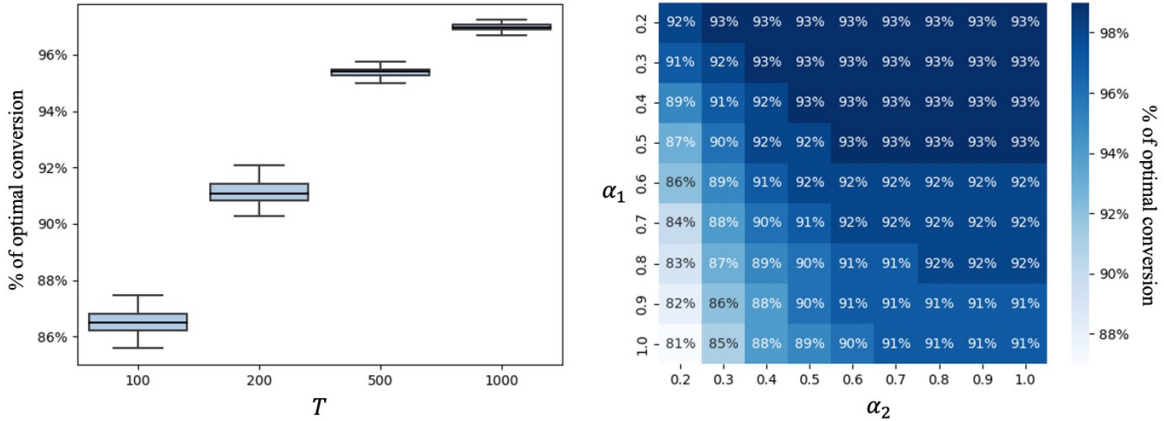


Figure 3 Left: percentage of global optimal conversion under different total periods T and optimal autobidding (i.e. $\alpha_1 = \alpha_2 = 1$). Each point in box plot corresponds to the percentage conversion of a single trial $\ell \in [L]$, namely, $(\sum_{j \in [M]} V_j(\bar{\rho}_T^{(\ell, \alpha_1, \alpha_2)})) / \text{GL-OPT}^{(\ell)}$ for some $\ell \in [L]$; Right: percentage of global optimal conversion under $T = 200$, and non-optimal autobidding as we vary $\alpha_1, \alpha_2 \in \{0.2, 0.4, 0.6, 0.8, 1\}$. Each point in the heatmap is the average conversion percentage over all L trials, namely, $\frac{1}{L} \sum_{\ell \in [L]} (\sum_{j \in [M]} V_j(\bar{\rho}_T^{(\ell, \alpha_1, \alpha_2)})) / \text{GL-OPT}^{(\ell)}$.

Results. To analyze the influence of the total number of periods on our DSD-UCB algorithm under optimal autobidding, or in other words, to assess how data-hungry our algorithm is, we vary T

¹⁰ To model non-optimal autobidding behavior, in Section 6.2, we conduct another robustness check by considering a dynamic, non-stationary multi-agent setting, where autobidding is done with the help of a dynamic pacing algorithm in Balseiro et al. (2017), Balseiro and Gur (2019).

and display the corresponding percentage of conversion our algorithm’s output achieves, compared to the hindsight global optimal conversion, namely

$$\frac{\sum_{j \in [M]} V_j(\bar{\rho}_T^{(\ell, \alpha_1, \alpha_2)})}{\text{GL-OPT}^{(\ell)}}$$

for $\alpha_1 = \alpha_2 = 1$, in the left subgraph of Figure 3. Note that each box plot corresponds to the variation over different trials $\ell \in [L]$. We observe that increasing the number of periods T in our algorithm yields improved performance by allowing for more opportunities to collect data and learn per-channel budgets, while simultaneously reducing the variance associated with its performance. More interestingly, we note that for $T = 200$ (i.e., 8-9 days if one period corresponds to running an ad campaign for an hour), our algorithm yields a per-channel budget profile that achieves more than 91% of the global optimal conversion.

In the right subgraph of Figure 3, we further calculate the average percentage of conversion our algorithm’s output achieves compared to the global optimal conversion as we vary corruption factors α_1, α_2 . In particular, fixing $T = 200$, for any (α_1, α_2) -pair, we plot

$$\frac{1}{L} \sum_{\ell \in [L]} \frac{\sum_{j \in [M]} V_j(\bar{\rho}_T^{(\ell, \alpha_1, \alpha_2)})}{\text{GL-OPT}^{(\ell)}}$$

in the right subgraph of Figure 3. Note here that we are displaying the performance of the output per-channel budget profile of our algorithm $\bar{\rho}_T^{(\ell, \alpha_1, \alpha_2)}$ under optimal autobidding conversion $V_j(\cdot)$, instead of the observed non-optimal autobidding conversion $\tilde{V}_j(\cdot)$. This metric can be viewed as a normalized version of realized conversion by accounting for corruption factors α_1 and α_2 , which allows us to assess how much our algorithm’s output per-channel budget profile deviates from the optimal per-channel budget profile.

Fixing the corruption factor α_1 for the “non-lucrative” channels (i.e., channels with lower value-cost ratio auctions on average), the conversion performance of our algorithm improves as the other, more lucrative channels autobid more optimally. Nevertheless, even in the most corrupted case, where both channels have corruption factors of 0.2 (i.e., both achieve only 20% of optimal autobidding conversion), our algorithm still outputs a per-channel budget profile that attains more than 92% of the optimal (normalized) conversion. In contrast, when we fix the corruption factor α_2 for the “lucrative” channels, our algorithm achieves better performance as the non-lucrative channels become more sub-optimal in their autobidding. This is primarily because, with a decreasing corruption factor (indicating more sub-optimal autobidding) for the non-lucrative channels, our algorithm tends to allocate more budget to the lucrative channels, which inherently generate higher conversion for the same amount of spend.

6.2. Dynamic multi-agent environment

Setup. In this section, we consider a multi-agent non-stochastic setting where individual buyers' actions (i.e., bids) are modeled using standard learning algorithms. Our experimental process mimics real-world ad procurement processes, in which, after an advertiser sets a budget for a channel, the platform deploys a standard pacing algorithm for online constraint management—or equivalently, dual subgradient descent (DSD) from Balseiro et al. (2017)—to participate in a large number of auctions and acquire ads on the advertiser's behalf. The goal of this section is to analyze the behavior and performance of our proposed DSD-UCB algorithm when interacting with these standard pacing algorithms.

Experiment procedure. We consider the perspective of a single buyer acquiring ads across $M \geq 3$ channels, where each channel consists of N participant autobidders (including the buyer, indexed as 0), for some $N \in \mathbb{N}$. For simplicity, we assume all competitors are subject only to long-term budget constraints across multiple rounds of bidding. Each channel conducts repeated second-price auctions organized in phases. Specifically, consider $H = 100$ consecutive phases, where each phase consists of $T = 200$ periods, during which each channel independently runs a single auction in every period. At the start of each phase h , each buyer i determines the total budget $\rho_{j,h}^{(i)}T$ to allocate to each channel j . Budgets are fully replenished at the beginning of each new phase. Conceptually, a phase represents a single ad campaign deployed across multiple channels, where, within the campaign, each channel procures ads on behalf of the buyer over T periods. In each period, the valuations of all N buyers for the auctioned item in any channel are independently sampled from $\text{Uniform}([0, 1])$.¹¹

Most importantly, we assume that each channel runs DSD to submit bids on buyers' behalf for each period within a phase, based on a total budget set at the beginning of the phase. Specifically, for each phase $h \in [H]$, and for each channel $j \in [M]$ maintains a dual variable $\lambda_{j,t}^{(i)}$ for buyer i before bidding in period t . Once the buyer's valuation $v_{j,t}^{(i)}$ for the auctioned item in the channel is observed, the channel submits a bid value

$$b_{j,t}^{(i)} = \frac{v_{j,t}^{(i)}}{1 + \lambda_{j,t}^{(i)}}$$

on behalf of buyer i . Let $p_{j,t}^{(i)}$ denote the second-price auction payment for buyer i in channel j during period t , where $p_{j,t}^{(i)}$ equals the highest competing bid if buyer i wins the auction in channel j , and 0 otherwise. If buyer i wins the auction in channel j , they acquire a value of $v_{j,t}^{(i)}$; otherwise, the value is 0. The channel then updates the dual variables as follows:

$$\lambda_{j,t+1}^{(i)} = \lambda_{j,t}^{(i)} - \epsilon \left(\frac{\rho_{j,h}^{(i)}}{T} - p_{j,t}^{(i)} \right),$$

¹¹ The code and randomly generated data used for this experiment can be found at: https://anonymous.4open.science/r/multichannel_autobidding-6645/MultichannelSimulation.py.

where $\epsilon = 1/T$ is a fixed DSD step size that is identical across channels and buyers.

As stated earlier, we consider the perspective of a single buyer, indexed as 0, who uses our DSD-UCB algorithm to allocate their budget across multiple channels. We assume that competitors $i \in \{1, \dots, N\}$ have a per-period budget of $\rho_{j,h}^{(i)} = 0.1$. To implement DSD-UCB, at the end of each phase h , channel j reports back the total conversion value. Using the historical conversion data $(V_{j,1}^{(0)}, \dots, V_{j,h'}^{(0)})_{j \in [M]}$ for all $h' \leq h$, the buyer determines the per-channel budgets $(\rho_{j,h+1}^{(0)})_{j \in [M]}$ for the next phase. These budget parameters are updated according to the proposed DSD-UCB algorithm described in Algorithm 1.

For simplicity we fix our aggregate ROI $\gamma = 0.1$, and in the following section we analyze the performance of our DSD-UCB algorithm when we vary number of buyers $N \in \{2 \dots 10\}$ in each channel as well as our per-period budget $\rho \in \{0.1, 0.2 \dots 0.9\}$.¹² We run 50 independent trials for our experiment procedure (recall each consisting $H = 100$ with $T = 200$ periods per phase), and in each trial, we sample the number of channels $M \in \{3, 4 \dots 10\}$ uniformly at random.

Benchmarks. Denote our highest competing bid for each channel j as $d_{j,t} = \max_{i \in \{1, \dots, N\}} b_{j,t}^{(i)}$. Further, let

$$\tilde{d}_j = \frac{1}{HT} \sum_{t \in [HT]} d_{j,t} \quad \text{and} \quad \tilde{v}_j = \frac{1}{HT} \sum_{t \in [HT]} v_{j,t}^{(0)}$$

represent the average values of the highest competing bids and the buyer's valuations, respectively, over all phases and periods. Note that $\tilde{d}_j$ can be interpreted as the average cost of procuring ads in channel $j \in [M]$. We consider two benchmarks for comparing the performance of our proposed DSD-UCB algorithm:

1. **HINDSIGHT-OPT:** This measures the largest acquirable total value if all costs (i.e., highest competing bids) and realized values were observed prior to bidding.
2. **PHASE-GL-OPT:** This measures the total value obtained when fixing the single best budget for each channel across different phases. PHASE-GL-OPT serves as a reasonable benchmark since it computes the optimal solution that maximizes the expected value across channels while having access to their average performance, providing a competitive baseline for comparison.

Mathematically, these two benchmarks are defined as follows

$$\begin{aligned} \max_{x_t \in [0,1]} \quad & \frac{1}{HT} \sum_{j \in [M]} \sum_{t \in [HT]} v_{j,t} x_{j,t} \\ \text{s.t.} \quad & \sum_{j \in [M]} \sum_{t \in [HT]} v_{j,t} x_{j,t} \geq \gamma \sum_{j \in [M]} \sum_{t \in [HT]} d_{j,t} x_{j,t} \\ & \sum_{j \in [M]} \sum_{t \in [HT]} d_{j,t} x_{j,t} \leq \rho HT. \end{aligned} \quad \text{(HINDSIGHT-OPT)}$$

$$\begin{aligned} \max_{x_1, \dots, x_M \in [0,1]} \quad & \sum_{j \in [M]} \tilde{v}_j x_j \\ \text{s.t.} \quad & \sum_{j \in [M]} \tilde{v}_j x_j \geq \gamma \sum_{j \in [M]} \tilde{d}_j x_j \\ & \sum_{j \in [M]} \tilde{d}_j x_j \leq \rho. \end{aligned} \quad \text{(PHASE-GL-OPT)}$$

¹² We further investigate how convergence behavior scales with the amount of available budget. Additional results for small-budget campaigns, reported in Appendix E, show that, as expected, convergence slows when the per-round budget is smaller.

Results. In the following figures, we plot HINDSIGHT-OPT, PHASE-GL-OPT, and the realized acquired values (averaged over all HT periods) for varying budget ρ and total number of buyers N .

In Figure 4, we fix our budget parameter $\rho = 0.1$, equal to that of all competitors, and vary the per-channel number of buyers N . We observe that the per-period acquired value decreases as the number of competitors increases within each channel. This is due to increased competition, which raises the effective cost of acquiring items. More importantly, our proposed DSD-UCB algorithm achieves between 65% and 88% of the HINDSIGHT-OPT, which represents the largest achievable value given full knowledge of the highest competing bids for each period.¹³

Additionally, our algorithm achieves a 1.3x increase in total value compared to PHASE-GL-OPT when $N = 10$ and up to a 1.75x improvement when $N = 3$. This highlights the robustness of the DSD-UCB algorithm even under high competition.

In Figure 5, we fix the number of buyers at $N = 10$ and vary the budget parameter $\rho \in \{0.1, 0.2, \dots, 0.9\}$. We observe that the per-period acquired value increases as the budget parameter ρ grows, reflecting greater buying power allocated to each channel’s campaign. Our proposed DSD-UCB algorithm achieves between 60% and 90% of the HINDSIGHT-OPT, with the relative performance improving as ρ increases. For instance, when $\rho = 0.9$, the algorithm achieves 90% of the hindsight-optimal value — a significant increase compared to 65% at $\rho = 0.1$. Furthermore, the DSD-UCB algorithm achieves a 1.4x increase in total value compared to PHASE-GL-OPT when $\rho = 0.2$ and up to 1.8x when $\rho = 0.9$.

Managerial insights. Our numerical studies demonstrate that the DSD-UCB algorithm achieves a high fraction of the value obtained by the hindsight global optimal solution, even when autobidding algorithms do not explicitly solve an optimization problem and instead use pacing algorithms to update bidders’ bidding strategies. As shown in Balseiro and Gur (2019), when all bidders adopt adaptive pacing strategies, competing bids become nonstationary and endogenous. In such scenarios, under certain assumptions, these strategies converge to an approximate Nash equilibrium. We conjecture that the DSD-UCB algorithm implicitly leverages the convergence of the pacing strategies to a Nash equilibrium, enabling robust performance. Nonetheless, further investigation into this framework presents an intriguing avenue for future research.

7. Conclusion and Open Problems

This work addresses a critical yet overlooked aspect of online advertising: how advertisers interact with channels to optimize their campaigns. Our research focuses on how advertisers can effectively

¹³ Although we don’t show it here, the relative performance of the DSD-UCB algorithm improves as T (i.e., the number of periods within a phase) increases, which is currently set at 200. For example, increasing T from 200 to 400 improves the efficiency by approximately 5% due to reduced pacing inefficiency within a single phase.

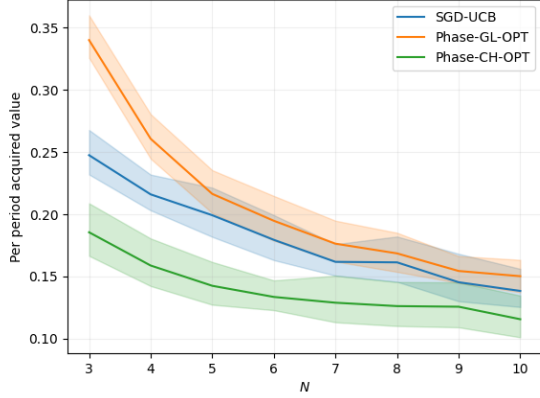


Figure 4 Fixed budget ratio $\rho = 0.1$. Shaded region represents distribution of average per-period value over 50 independent trials.

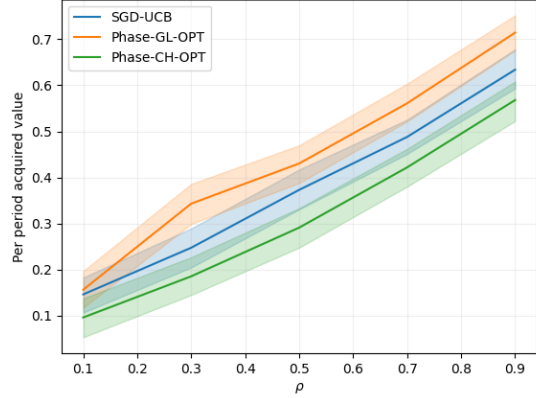


Figure 5 Fixed buyers $N = 10$ for each channel. Shaded region represents distribution of average per-period value over 50 independent trials.

utilize per-channel budget and ROI levers to maximize conversions under global budget and ROI constraints. We demonstrate that optimizing per-channel budgets alone is sufficient to achieve global optimality, while relying solely on per-channel ROI targets can lead to suboptimal outcomes. To tackle the challenge of optimizing budget decisions under bandit feedback, we propose an algorithm combining stochastic gradient descent with a UCB framework, achieving efficient sublinear regret guarantees. Our work extends to multi-impression auctions and general advertiser objectives, and numerical studies confirm the algorithm’s effectiveness.

Our work focuses on the stochastic setting, motivated by real-world scenarios. While advertisers may adapt their strategies over time, aggregated changes across a large market rarely create highly adversarial environments. Instead, the dynamics are typically stochastic and relatively stationary, making this setting more representative and applicable. Nevertheless, as in the BwK literature (e.g., Castiglioni et al. (2022a))—where approximate sublinear regret is achieved instead of exact sublinear regret in the stochastic setting—exploring adversarial settings could provide valuable theoretical insights and represent an interesting direction for future work. In adversarial settings, the values obtained from channels are not derived from an optimization problem assuming stochastic competing bids, but are instead generated adversarially. Under such conditions, to obtain meaningful results, one needs to bound the discretization error in the primal step, which is impossible without imposing certain restrictions on the adversary.

Another promising direction is to study market dynamics without assuming that the channel optimization problem is solved optimally. In this scenario, the outcome of budget allocation is derived from gradient descent dynamics, as modeled and simulated in Section 6.2. Our numerical results provide positive evidence, confirming that the global optimal solution can still be reached in this

setting. Nonetheless, further investigation into this framework presents an intriguing avenue for future research.

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.
- Agrawal S, Devanur NR (2014) Bandits with concave rewards and convex knapsacks. *Proceedings of the fifteenth ACM conference on Economics and computation*, 989–1006.
- Agrawal S, Devanur NR (2019) Bandits with global convex constraints and objective. *Operations Research* 67(5):1486–1502.
- Agrawal S, Devanur NR, Li L (2016) An efficient algorithm for contextual bandits with knapsacks, and an extension to concave objectives. *Conference on Learning Theory*, 4–18 (PMLR).
- Ai R, Wang C, Li C, Zhang J, Huang W, Deng X (2022) No-regret learning in repeated first-price auctions with budget constraints. *arXiv preprint arXiv:2205.14572* .
- Alimohammadi Y, Mehta A, Perloth A (2023) Incentive compatibility in the auto-bidding world. *arXiv preprint arXiv:2301.13414* .
- 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.
- Badanidiyuru A, Kleinberg R, Slivkins A (2013) Bandits with knapsacks. *2013 IEEE 54th Annual Symposium on Foundations of Computer Science (FOCS)*, 207–216 (IEEE Computer Society).
- Badanidiyuru A, Kleinberg R, Slivkins A (2018) Bandits with knapsacks. *Journal of the ACM (JACM)* 65(3):1–55.
- Badanidiyuru A, Langford J, Slivkins A (2014) Resourceful contextual bandits. *Conference on Learning Theory*, 1109–1134 (PMLR).
- 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 (2022a) Contextual bandits with cross-learning. *Mathematics of Operations Research* .

- Balseiro S, Golrezaei N, Mirrokni V, Yazdanbod S (2019) 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 (2022b) The best of many worlds: Dual mirror descent for online allocation problems. *Operations Research* .
- Bergemann D, Bonatti A (2023) Data, competition, and digital platforms. *arXiv preprint arXiv:2304.07653* .
- 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.
- Bürmen Á, Puhán J, Tuma T (2006) Grid restrained nelder-mead algorithm. *Computational optimization and applications* 34(3):359–375.
- Castiglioni M, Celli A, Kroer C (2022a) Online learning with knapsacks: the best of both worlds. *International Conference on Machine Learning*, 2767–2783 (PMLR).
- Castiglioni M, Celli A, Kroer C (2023) Online bidding in repeated non-truthful auctions under budget and roi constraints. *arXiv preprint arXiv:2302.01203* .
- Castiglioni M, Celli A, Marchesi A, Romano G, Gatti N (2022b) A unifying framework for online optimization with long-term constraints. *arXiv preprint arXiv:2209.07454* .
- Chen Q, Golrezaei N, Bouneffouf D (2022) Dynamic bandits with an auto-regressive temporal structure. *arXiv preprint arXiv:2210.16386* .
- 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).
- Conitzer V, Kroer C, Panigrahi D, Schrijvers O, Stier-Moses NE, Sodomka E, Wilkens CA (2022) Pacing equilibrium in first price auction markets. *Management Science* 68(12):8515–8535.
- 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.
- 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 Y, Lucier B, Slivkins A (2023) Strategic budget selection in a competitive autobidding world. *arXiv preprint arXiv:2307.07374* .
- Feng Z, Padmanabhan S, Wang D (2022) Online bidding algorithms for return-on-spend constrained advertisers. *arXiv preprint arXiv:2208.13713* .
- Fikioris G, Tardos É (2023) Liquid welfare guarantees for no-regret learning in sequential budgeted auctions. *Proceedings of the 24th ACM Conference on Economics and Computation*, 678–698.
- Flaxman AD, Kalai AT, McMahan HB (2004) Online convex optimization in the bandit setting: gradient descent without a gradient. *arXiv preprint cs/0408007* .
- Gaitonde J, Li Y, Light B, Lucier B, Slivkins A (2022) Budget pacing in repeated auctions: Regret and efficiency without convergence. *arXiv preprint arXiv:2205.08674* .
- Golrezaei N, Jaillet P, Liang JCN (2023a) Incentive-aware contextual pricing with non-parametric market noise. *International Conference on Artificial Intelligence and Statistics*, 9331–9361 (PMLR).
- Golrezaei N, Jaillet P, Liang JCN, Mirrokni V (2023b) Pricing against a budget and roi constrained buyer. *International Conference on Artificial Intelligence and Statistics*, 9282–9307 (PMLR).
- Golrezaei N, Manshadi V, Schneider J, Sekar S (2021) Learning product rankings robust to fake users. *Proceedings of the 22nd ACM Conference on Economics and Computation*, 560–561.
- 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.
- Immorlica N, Sankararaman K, Schapire R, Slivkins A (2022) Adversarial bandits with knapsacks. *Journal of the ACM* 69(6):1–47.

-
- 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.
- Liao L, Kroer C, Leonenkov S, Schrijvers O, Shi L, Stier-Moses N, Zhang C (2024) Interference among first-price pacing equilibria: A bias and variance analysis. *arXiv preprint arXiv:2402.07322* .
- Lucier B, Pattathil S, Slivkins A, Zhang M (2023) Autobidders with budget and roi constraints: Efficiency, regret, and pacing dynamics. *arXiv preprint arXiv:2301.13306* .
- Luo H, Wei CY, Agarwal A, Langford J (2018) Efficient contextual bandits in non-stationary worlds. *Conference On Learning Theory*, 1739–1776 (PMLR).
- 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, Sankararaman KA, Foster DJ (2023) Contextual bandits with packing and covering constraints: A modular lagrangian approach via regression. *The Thirty Sixth Annual Conference on Learning Theory*, 4633–4656 (PMLR).
- Slivkins A, et al. (2019) Introduction to multi-armed bandits. *Foundations and Trends® in Machine Learning* 12(1-2):1–286.
- Susan F, Golrezaei N, Schrijvers O (2023) Multi-platform budget management in ad markets with non-ic auctions. *arXiv preprint arXiv:2306.07352* .
- 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.
- Wang Q, Yang Z, Deng X, Kong Y (2023) Learning to bid in repeated first-price auctions with budgets. *arXiv e-prints* arXiv-2304.

Appendices for

Multi-channel Autobidding with Budget and ROI constraints

Appendix A: 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 $\max_{\mathbf{x}_1, \dots, \mathbf{x}_M} \sum_{j \in [M]} \mathbb{E} [\mathbf{v}_j^\top \mathbf{x}_j - \alpha \mathbf{d}_j^\top \mathbf{x}_j]$ and $\max_{(\gamma_j, \rho_j)_{j \in [M]} \in \mathcal{I}} \sum_{j \in M} \mathbb{E} [V_j(\gamma_j, \rho_j; \mathbf{z}_j) - \alpha V_j(\gamma_j, \rho_j; \mathbf{z}_j)]$ for some private cost $\alpha \in [0, \gamma]$ ¹⁴ in GL-OPT and CH-OPT($\mathcal{I}$), respectively. When $\alpha = 0$, we recover our considered models in the previous section, whereas 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. (2019) 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 DSD-UCB 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.

Appendix B: Proofs for Section 3

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

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

$$\sum_{j \in M} \mathbb{E} [V_j(\tilde{\gamma}_j, \tilde{\rho}_j; \mathbf{z}_j)] \geq \gamma \sum_{j \in M} \mathbb{E} [D_j(\tilde{\gamma}_j, \tilde{\rho}_j; \mathbf{z}_j)] \implies \sum_{j \in [M]} \mathbb{E} [\mathbf{v}_j^\top \mathbf{x}_j^*(\tilde{\gamma}_j, \tilde{\rho}_j; \mathbf{z}_j)] \geq \gamma \sum_{j \in [M]} \mathbb{E} [\mathbf{d}_j^\top \mathbf{x}_j^*(\tilde{\gamma}_j, \tilde{\rho}_j; \mathbf{z}_j)],$$

where we used the definitions $V_j(\tilde{\gamma}_j, \tilde{\rho}_j; \mathbf{z}_j) = \mathbf{v}_j^\top \mathbf{x}_j^*(\tilde{\gamma}_j, \tilde{\rho}_j; \mathbf{z}_j)$ and $D_j(\tilde{\gamma}_j, \tilde{\rho}_j; \mathbf{z}_j) = \mathbf{d}_j^\top \mathbf{x}_j^*(\tilde{\gamma}_j, \tilde{\rho}_j; \mathbf{z}_j)$ in Eq. (5).

This implies that $(\mathbf{x}_j^*(\tilde{\gamma}_j, \tilde{\rho}_j; \mathbf{z}_j))_{j \in [M]}$ satisfies the ROI constraint in GL-OPT. A similar analysis implies that $(\mathbf{x}_j^*(\tilde{\gamma}_j, \tilde{\rho}_j; \mathbf{z}_j))_{j \in [M]}$ also satisfies the budget constraint in GL-OPT. Therefore,

$$(\mathbf{x}_j^*(\tilde{\gamma}_j, \tilde{\rho}_j; \mathbf{z}_j))_{j \in [M]} \text{ is feasible to GL-OPT.}$$

Thus,

$$\text{GL-OPT} \geq \sum_{j \in [M]} \mathbb{E} [\mathbf{v}_j^\top \mathbf{x}_j^*(\tilde{\gamma}_j, \tilde{\rho}_j; \mathbf{z}_j)] = \sum_{j \in M} [V_j(\tilde{\gamma}_j, \tilde{\rho}_j; \mathbf{z}_j)] = \text{CH-OPT}(\mathcal{I}),$$

¹⁴ If $\alpha > \gamma$ the ROI constraints in GL-OPT as well as CH-OPT($\mathcal{I}$) become redundant.

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

□

B.2. Proof of Theorem 3.3

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

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

$$\mathbb{E}[D_j(\tilde{\gamma}_j, \tilde{\rho}_j; \mathbf{z}_j)] = \mathbb{E}[\mathbf{d}_j^\top \mathbf{x}_j^*(\tilde{\gamma}_j, \tilde{\rho}_j; \mathbf{z}_j)] \stackrel{(i)}{\leq} \tilde{\rho}_j = \mathbb{E}[\mathbf{d}_j^\top \tilde{\mathbf{x}}_j(\mathbf{z}_j)], \quad (14)$$

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

$$\sum_{j \in [M]} \mathbb{E}[D_j(\tilde{\gamma}_j, \tilde{\rho}_j; \mathbf{z}_j)] \leq \sum_{j \in [M]} \mathbb{E}[\mathbf{d}_j^\top \tilde{\mathbf{x}}_j(\mathbf{z}_j)] \stackrel{(i)}{\leq} \rho. \quad (15)$$

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

On the other hand, we have

$$V_j(\tilde{\gamma}_j, \tilde{\rho}_j; \mathbf{z}_j) = \mathbf{v}_j^\top \mathbf{x}_j^*(\tilde{\gamma}_j, \tilde{\rho}_j; \mathbf{z}_j) \stackrel{(i)}{\geq} \mathbf{v}_j^\top \tilde{\mathbf{x}}_j(\mathbf{z}_j), \quad (16)$$

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

$$\sum_{j \in [M]} \mathbb{E}[V_j(\tilde{\gamma}_j, \tilde{\rho}_j; \mathbf{z}_j)] \geq \sum_{j \in [M]} \mathbb{E}[\mathbf{v}_j^\top \tilde{\mathbf{x}}_j(\mathbf{z}_j)] \stackrel{(i)}{\geq} \gamma \sum_{j \in [M]} \mathbb{E}[\mathbf{d}_j^\top \tilde{\mathbf{x}}_j(\mathbf{z}_j)] \stackrel{(ii)}{\geq} \gamma \sum_{j \in [M]} \mathbb{E}[D_j(\tilde{\gamma}_j, \tilde{\rho}_j; \mathbf{z}_j)], \quad (17)$$

where (i) follows from feasibility of $\tilde{\mathbf{x}}(\mathbf{z}) = \{\tilde{\mathbf{x}}_j(\mathbf{z}_j)\}_{j \in [N]}$ to GL-OPT since it is the optimal solution; (ii) follows from Eq. (14). Hence, combining Eq. (15) and Eq.(17), we can conclude that $(\tilde{\gamma}_j, \tilde{\rho}_j)_{j \in [M]}$ is feasible to CH-OPT($\mathcal{I}_B$).

Finally, we have CH-OPT($\mathcal{I}_B$) $\geq \sum_{j \in [M]} \mathbb{E}[V_j(\tilde{\gamma}_j, \tilde{\rho}_j; \mathbf{z}_j)] \geq \sum_{j \in [M]} \mathbb{E}[\mathbf{v}_j^\top \tilde{\mathbf{x}}_j(\mathbf{z}_j)] = \text{GL-OPT}$, where the last inequality follows from Eq. (17), and the final equality is because we assumed $\tilde{\mathbf{x}}(\mathbf{z}) = \{\tilde{\mathbf{x}}_j(\mathbf{z}_j)\}_{j \in [N]}$ is the optimal solution to GL-OPT.

□

B.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 $(\tilde{\gamma}, \tilde{\rho}) \in \mathcal{I}_B$, and by definition of $\mathcal{I}_B$ in Eq. (2) we have $\tilde{\gamma}_j = 0$ for all $j \in [M]$. Since $(\tilde{\gamma}, \tilde{\rho})$ 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

$$\text{CH-OPT}(\mathcal{I}_G) \geq \text{CH-OPT}(\mathcal{I}_B) = \text{GL-OPT}, \quad (18)$$

where the final equality follows from Theorem 3.3. □

Appendix C: Proofs for Section 4

C.1. Proof of Proposition 4.1

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

$$\begin{aligned}
& T \cdot \text{GL-OPT} - \sum_{t \in [T]} \sum_{j \in [M]} V_j(\rho_{j,t}) \\
& \stackrel{(i)}{\leq} M\bar{V}(T - \tau_A) + \tau_A \text{CH-OPT}(\mathcal{I}_B) - \sum_{t \in [\tau_A]} \sum_{j \in [M]} V_j(\rho_{j,t}) \\
& \stackrel{(ii)}{\leq} M\bar{V}(T - \tau_A) + \tau_A \cdot \left(\sum_{j \in [M]} \mathcal{L}_j(\rho_j^*, \bar{\lambda}_T, \bar{\mu}_T) + \rho \bar{\mu}_T \right) - \sum_{t \in [\tau_A]} \sum_{j \in [M]} V_j(\rho_{j,t}) \\
& \stackrel{(iii)}{\leq} M\bar{V}(T - \tau_A) + \rho \sum_{t \in [\tau_A]} \mu_t + \sum_{t \in [\tau_A]} \sum_{j \in [M]} \mathcal{L}_j(\rho_j^*, \lambda_t, \mu_t) - \left[\sum_{t \in [\tau_A]} \sum_{j \in [M]} \mathcal{L}_j(\rho_{j,t}, \mathbf{c}_t) - \lambda_t (V_j(\rho_{j,t}) - \gamma \rho_{j,t}) + \mu_t \rho_{j,t} \right] \\
& \stackrel{(iv)}{\leq} M\bar{V}(T - \tau_A) + \sum_{j \in [M]} \sum_{t \in [\tau_A]} \mathcal{L}_j(\rho_j^*(t), \mathbf{c}_t) - \mathcal{L}_j(\rho_{j,t}, \mathbf{c}_t) + \sum_{t \in [\tau_A]} (\lambda_t g_{1,t} + \mu_t g_{2,t}).
\end{aligned} \tag{19}$$

Here, (i) follows from Theorem 3.3 that states $\text{GL-OPT} = \text{CH-OPT}(\mathcal{I}_B)$ and $\text{CH-OPT}(\mathcal{I}_B)$ is apparently upper bounded by $M\bar{V}$; (ii) follows from the weak duality theorem that $\text{CH-OPT}(\mathcal{I}_B) \leq \sum_{j \in [M]} \mathcal{L}_j(\rho_j^*, \lambda, \mu) + \mu \rho$ for any $(\lambda, \mu) \in \mathbb{R}_+^2$; (iii) follows from that $\mathcal{L}_j(\rho_j^*, \lambda, \mu)$ is linear in both λ and μ and the definition of the Lagrangian in Eq. (6); in (iv) we define $\rho_j^*(t) = \arg \max_{\rho_j \geq 0} \mathcal{L}_j(\rho_j, \mathbf{c}_t)$ to be the optimal budget that maximizes the Lagrangian w.r.t. the dual variables $\mathbf{c}_t = (\lambda_t, \mu_t)$. Further, we have $g_{1,t} = \sum_{j \in [M]} (V_j(\rho_{j,t}) - \gamma \rho_{j,t})$ and $g_{2,t} = \rho - \sum_{j \in [M]} \rho_{j,t}$. Q.E.D.

C.2. Proof for Lemma 4.2

Recall $g_{1,t} = \sum_{j \in [M]} (V_{j,t}(\rho_{j,t}; \mathbf{z}_{j,t}) - \gamma \rho_{j,t})$ and $g_{2,t} = \rho - \sum_{j \in [M]} \rho_{j,t}$ defined in Algorithm 1. Also recall $\tau_A \in [T]$ defined in step 10 of Algorithm 1. In the following, we will show

$$\begin{aligned}
& M\bar{V}(T - \tau_A) + \sum_{t \in [\tau_A]} (\lambda_t g_{1,t} + \mu_t g_{2,t}) \\
& \leq C_F \max\{M\bar{V}, \rho\} + M^2 \bar{V} \rho \cdot \max\left\{\frac{1}{\beta_{\underline{\rho}}}, \frac{1}{\rho - M\underline{\rho}}\right\} + \frac{(\gamma M^2 \bar{V}^2 + \rho^2)}{2} \cdot \eta T + \frac{1}{2\eta} C_F^2 = \mathcal{O}\left(\eta T + \frac{1}{\eta}\right),
\end{aligned} \tag{20}$$

where we recall $C_F = M\bar{V} \max\left\{\frac{1}{\beta_{\underline{\rho}}}, \frac{1}{\rho - M\underline{\rho}}\right\}$ defined in Eq. (9).

From Lemma C.3, we have for any $t \in [T]$, and $\lambda, \mu \in [0, C_F]$,

$$\begin{aligned}
\sum_{\tau \in [t]} (\lambda_\tau - \lambda) g_{1,\tau} & \leq \frac{\eta M^2 \bar{V}^2}{2} \cdot t + \frac{1}{2\eta} \lambda^2 \\
\sum_{\tau \in [t]} (\mu_\tau - \mu) g_{2,\tau} & \leq \frac{\eta \rho^2}{2} \cdot t + \frac{1}{2\eta} \mu^2,
\end{aligned} \tag{21}$$

where we used the fact that $\lambda_1 = \mu_1 = 0$ in Algorithm 1.

Suppose that $\tau_A = T$ and thus $M\bar{V}(T - \tau_A) = 0$. Then, considering $\lambda = \mu = 0$ in Eq. (21), we have

$$\sum_{t \in [\tau_A]} \lambda_t g_{1,t} \leq \frac{\eta M^2 \bar{V}^2}{2} \cdot T \text{ and } \sum_{t \in [\tau_A]} \mu_t g_{2,t} \leq \frac{\eta \rho^2}{2} \cdot T. \tag{22}$$

Thus, Eq. (20) holds.

If $\tau_A < T$, then according to Algorithm 1, we either have $S_{1,\tau_A} - \gamma M\rho + \beta\rho(T - \tau_A) < 0$ or $S_{2,\tau_A} + M\rho + M\rho(T - \tau_A) > \rho T$, where we recall $S_{1,\tau_A} = \sum_{t \in [\tau_A-1]} g_{1,t}$ and $S_{2,\tau_A} = \sum_{t \in [\tau_A-1]} \sum_{j \in [M]} \rho_{j,t} = \sum_{t \in [\tau_A-1]} (\rho - g_{2,t})$:

- If $S_{1,\tau_A} - \gamma M\rho + \beta\rho(T - \tau_A) < 0$, then we have $\sum_{t \in [\tau_A-1]} g_{1,t} < \gamma M\rho - \beta\rho(T - \tau_A)$. Hence, considering $\lambda = \frac{M\bar{V}}{\beta\rho} \in [0, C_F]$ in Eq. (21), we have

$$\begin{aligned}
& M\bar{V}(T - \tau_A) + \sum_{t \in [\tau_A]} \lambda_t g_{1,t} \\
& \leq M\bar{V}(T - \tau_A) + \lambda_{\tau_A} g_{1,\tau_A} + \sum_{t \in [\tau_A-1]} \lambda g_{1,t} + \frac{\eta M^2 \bar{V}^2}{2} \cdot (\tau_A - 1) + \frac{1}{2\eta} \lambda^2 \\
& < M\bar{V}(T - \tau_A) + \lambda_{\tau_A} g_{1,\tau_A} - M\bar{V}(T - \tau_A) + \frac{\gamma M^2 \bar{V} \rho}{\beta\rho} + \frac{\eta M^2 \bar{V}^2}{2} \cdot (\tau_A - 1) + \frac{1}{2\eta} \lambda^2 \\
& \leq C_F M\bar{V} + \frac{\gamma M^2 \bar{V} \rho}{\beta\rho} + \frac{\eta M^2 \bar{V}^2}{2} \cdot T + \frac{1}{2\eta} C_F^2,
\end{aligned} \tag{23}$$

where the first inequality follows from Eq. (21), and the second inequality holds because as stated above, we have $\sum_{t \in [\tau_A-1]} g_{1,t} < \gamma M\rho - \beta\rho(T - \tau_A)$ in this case. The final inequality uses the fact that $\tau_A \leq T$, $\lambda \leq C_F$, and $g_{1,t} \leq M\bar{V}$ for any $t \in [T]$. Finally, by further taking $\mu = 0$, we have $\sum_{t \in [\tau_A]} \mu_t g_{2,t} \leq \frac{\eta \rho^2}{2} \cdot T$. This and Eq. (23) show that Eq. (20) holds.

- If $S_{2,\tau_A} + M\rho + M\rho(T - \tau_A) > \rho T$, then we have $\sum_{t \in [\tau_A-1]} (\rho - g_{2,t}) > \rho T - M\rho - M\rho(T - \tau_A)$, or equivalently $\sum_{t \in [\tau_A-1]} g_{2,t} < M\rho(T - \tau_A) + M\rho - \rho(T - \tau_A) \leq -(\rho - M\rho)(T - \tau_A) + M\rho$. Hence, considering $\mu = \frac{M\bar{V}}{\rho - M\rho} \in [0, C_F]$ in Eq.(21) we have

$$\begin{aligned}
M\bar{V}(T - \tau_A) + \sum_{t \in [\tau_A]} \mu_t g_{2,t} & \leq M\bar{V}(T - \tau_A) + \mu_{\tau_A} g_{2,\tau_A} + \sum_{t \in [\tau_A-1]} \mu g_{2,t} + \frac{\eta \rho^2}{2} \cdot \tau_A + \frac{1}{2\eta} \mu^2 \\
& < M\bar{V}(T - \tau_A) + \mu_{\tau_A} g_{2,\tau_A} - M\bar{V}(T - \tau_A) + \frac{M^2 \bar{V} \rho}{\rho - \rho} + \frac{\eta \rho^2}{2} \cdot \tau_A + \frac{1}{2\eta} \mu^2 \tag{24} \\
& \leq C_F \rho + \frac{M^2 \bar{V} \rho}{\rho - \rho} + \frac{\eta \rho^2}{2} \cdot T + \frac{1}{2\eta} C_F^2,
\end{aligned}$$

where the first inequality follows from Eq. (21), and the second inequality holds because as stated above, we have $\sum_{t \in [\tau_A-1]} g_{2,t} < M\rho(T - \tau_A) + M\rho - \rho(T - \tau_A) \leq -(\rho - M\rho)(T - \tau_A) + M\rho$ in this case. The final inequality uses the fact that $\tau_A \leq T$, $\mu \leq C_F$, and $g_{2,t} \leq \rho$ for any $t \in [T]$. Finally, by further taking $\lambda = 0$, we have $\sum_{t \in [\tau_A]} \lambda_t g_{1,t} \leq \frac{\eta M^2 \bar{V}^2}{2} \cdot T$. This and Eq. (24) show that Eq. (20) holds.

Q.E.D.

C.3. Proof of Lemma 4.3

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

Fix any channel j that consists of m_j parallel auctions, and recall that we assumed the order $\frac{v_{j,1}}{d_{j,1}} > \frac{v_{j,2}}{d_{j,2}} > \dots > \frac{v_{j,m_j}}{d_{j,m_j}}$ for any realization $\mathbf{z}_j$. Then, with the option where the per-channel ROI is set to 0 (i.e. omitted)

$V_j(\rho_j; \mathbf{z}_j)$ is exactly the LP relaxation of a 0-1 knapsack, whose optimal solution $\mathbf{x}_j^*(\rho_j; \mathbf{z}_j)$ is well known to be unique, and takes the form for any auction index $n \in [m_j]$:

$$x_{j,n}^*(\rho_j; \mathbf{z}_j) = \begin{cases} 1 & \text{if } \sum_{n' \in [n]} d_{j,n'} \leq \rho_j \\ \frac{(\rho_j - \sum_{n' \in [n-1]} d_{j,n'})^+}{d_{j,n}} & \text{otherwise} \end{cases}, \quad (25)$$

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

$$V_j(\rho_j; \mathbf{z}_j) = \mathbf{v}_j^\top \mathbf{x}_j^*(\rho_j; \mathbf{z}_j) = \sum_{n \in [m_j]} \left(\frac{v_{j,n}}{d_{j,n}} \rho_j + b_{j,n} \right) \mathbb{I}\{d_{j,0} + \dots + d_{j,n-1} \leq \rho_j < d_{j,0} + \dots + d_{j,n}\} \quad (26)$$

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 $X_n = d_{j,0} + \dots + d_{j,n}$, 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. (26) we can conclude $V_j(\rho_j; \mathbf{z}_j)$ is continuous, which further implies it is piecewise linear and strictly increasing. Further, the ordering $\frac{v_{j,1}}{d_{j,1}} > \frac{v_{j,2}}{d_{j,2}} > \dots > \frac{v_{j,m_j}}{d_{j,m_j}}$ implies that the slopes on each segment $[X_n, X_{n+1}]$ decreases as n increases, which implies $V_j(\rho_j; \mathbf{z}_j)$ is concave.

Since $V_j(\rho_j) = \mathbb{E}[V_j(\rho_j; \mathbf{z}_j)]$, where the expectation is taken w.r.t. randomness in $\mathbf{z}_j$, and since the $\mathbf{z}_j$ is sampled from some discrete distribution $\mathbf{p}_j$ on finite support F_j , $V_j(\rho_j)$ is simply a weighted average over all $(V_j(\rho_j; \mathbf{z}_j))_{\mathbf{z}_j \in F_j}$ with weights in $\mathbf{p}_j$, 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 $\mathbf{p}_j$.

Finally, according to the definition of $\mathcal{L}_j(\rho_j, \mathbf{c}) = \mathbb{E}[\mathcal{L}_j(\rho_j, \mathbf{c}; \mathbf{z}_j)]$ and $\mathcal{L}_j(\rho_j, \mathbf{c}; \mathbf{z}_j) = (1 + \lambda)V_j(\rho_j; \mathbf{z}_j) - (\lambda\gamma + \mu)\rho_j$ as defined in Eq. (6), we have

$$\mathcal{L}_j(\rho_j, \mathbf{c}) = (1 + \lambda)V_j(\rho_j) - (\lambda\gamma + \mu)\rho_j \quad (27)$$

which implies $\mathcal{L}_j(\rho_j, \mathbf{c})$ is continuous, piecewise linear, and concave because $V_j(\rho_j)$ is continuous, piecewise linear, and concave as shown above. Combining Eq. (27) and the representation of $V_j(\rho_j)$ in Lemma (4.3), we have

$$\mathcal{L}_j(\rho_j, \mathbf{c}) = \sum_{n \in [S_j]} (\sigma_{j,n}(\mathbf{c})\rho_j + (1 + \lambda)b_{j,n}) \mathbb{I}\{r_{j,n-1} \leq \rho_j \leq r_{j,n}\}. \quad (28)$$

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

C.4. Proof for Lemma 4.4

Recall the definition of the Lagrangian function $\mathcal{L}_j(\rho_j, \mathbf{c}; \mathbf{z}_j) = (1 + \lambda)V_j(\rho_j; \mathbf{z}_j) - (\lambda\gamma + \mu)\rho_j$ in Eq.(6). Then, since $V_j(\rho_j; \mathbf{z}_j) \leq \bar{V}$, and $\lambda_t, \mu_t \in [0, C_F]$ for any period $t \in [T]$ and per-channel budget $\rho_j \in [0, \rho]$, we can conclude $-(1 + \gamma)\rho C_F \leq \mathcal{L}_j(\rho_j, \lambda_t, \mu_t) \leq (1 + C_F)\bar{V}$. Q.E.D.

C.5. Proof for Lemma 4.5

In the following, instead of bounding $\sum_{t \in [\tau_A]} \mathcal{L}_j(\rho_{j,t}^*, \mathbf{c}_t) - \mathcal{L}_j(\rho_{j,t}, \mathbf{c}_t)$, we bound $\sum_{t \in [T]} \mathcal{L}_j(\rho_{j,t}^*, \mathbf{c}_t) - \mathcal{L}_j(\rho_{j,t}, \mathbf{c}_t)$ where we consider the hypothetical scenario in which we ignore the termination criteria for the while loop in Algorithm 1, and continue to set per-channel budgets based on steps 4-6 in the algorithm until the end of period T . This is due to the fact that $\sum_{t \in [T]} \mathcal{L}_j(\rho_{j,t}^*, \mathbf{c}_t) - \mathcal{L}_j(\rho_{j,t}, \mathbf{c}_t) \geq \sum_{t \in [\tau_A]} \mathcal{L}_j(\rho_{j,t}^*, \mathbf{c}_t) - \mathcal{L}_j(\rho_{j,t}, \mathbf{c}_t)$.

We fix some channel $j \in [M]$ and omit the subscript j when the context is clear. Also, we first introduce some definitions that will be used throughout our proof. Fix some positive constant $\underline{\sigma} > 0$ whose value we choose later, and recall a_k denotes the k th arm in the discretized budget set $\mathcal{A}(\delta)$ as we defined in Eq. (8). Then we define the following

$$\begin{aligned} \Delta_k(\mathbf{c}) &= \max_{\rho_j \in [0, \rho]} \mathcal{L}_j(\rho_j, \mathbf{c}) - \mathcal{L}_j(a_k, \mathbf{c}) \\ \mathcal{C}_n &= \left\{ \mathbf{c} \in \{\mathbf{c}_t\}_{t \in [T]} : r_{j,n} = \arg \max_{\rho_j \geq 0} \mathcal{L}_j(\rho_j, \mathbf{c}) \right\} \text{ for } n = 0 \dots S_j \\ \mathcal{C}(\underline{\sigma}) &= \left\{ \mathbf{c} \in \{\mathbf{c}_t\}_{t \in [T]} : \sigma_j^-(\mathbf{c}) > \underline{\sigma}, |\sigma_j^+(\mathbf{c})| > \underline{\sigma} \right\} \\ m_k(\mathbf{c}) &= \frac{8 \log(T)}{\Delta_k^2(\mathbf{c})} \text{ for } \forall(k, \mathbf{c}) \text{ s.t. } \Delta_k(\mathbf{c}) > 0. \end{aligned} \tag{29}$$

Here, the “adjacent slopes” $\sigma_j^-(\mathbf{c})$ and $\sigma_j^+(\mathbf{c})$, which are defined in Eq.(11), represent the slopes adjacent to the optimal budget $\arg \max_{\rho_j \in [0, \rho]} \mathcal{L}_j(\rho_j, \mathbf{c})$ for any context $\mathbf{c} = (\lambda, \mu)$. In addition, S_j and $\{r_{j,n}\}_{n \in [S_j]}$ are defined in Lemma 4.3. Here we state in words the meanings of $\Delta_k(\mathbf{c})$, $\mathcal{C}(\underline{\sigma})$ and $\mathcal{C}_n$, respectively.

- $\Delta_k(\mathbf{c}) = \max_{\rho_j \in [0, \rho]} \mathcal{L}_j(\rho_j, \mathbf{c}) - \mathcal{L}_j(a_k, \mathbf{c})$ denotes the loss of contextual bandit rewards when pulling the arm a_k in context $\mathbf{c}$.
- $\mathcal{C}_n = \left\{ \mathbf{c} \in \{\mathbf{c}_t\}_{t \in [T]} : r_{j,n} = \arg \max_{\rho_j \geq 0} \mathcal{L}_j(\rho_j, \mathbf{c}) \right\}$ is the set including all context $\mathbf{c}_t$ under which the optimal per-channel budget $\arg \max_{\rho_j \geq 0} \mathcal{L}_j(\rho_j, \mathbf{c}_t)$ is taken at the n th “turning point” $r_{j,n}$ (see Lemma 4.3).
- $\mathcal{C}(\underline{\sigma}) = \left\{ \mathbf{c} \in \{\mathbf{c}_t\}_{t \in [T]} : \sigma_j^-(\mathbf{c}) > \underline{\sigma}, |\sigma_j^+(\mathbf{c})| > \underline{\sigma} \right\}$ is the set of all contexts, in which the adjacent slopes to the optimal point w.r.t. the context $\mathbf{c}$, namely $\arg \max_{\rho_j \geq 0} \mathcal{L}_j(\rho_j, \mathbf{c})$, have magnitude greater than $\underline{\sigma}$, or in other words, the adjacent slopes are steep.

On a related note, for any context $\mathbf{c}$, we define the following “adjacent regions” that sandwich the optimal budget w.r.t. $\mathbf{c}$

$$\mathcal{U}_j^-(\mathbf{c}) = [r_{j,n-1}, r_{j,n}] \text{ and } \mathcal{U}_j^+(\mathbf{c}) = [r_{j,n}, r_{j,n+1}] \text{ if } \mathbf{c} \in \mathcal{C}_n. \tag{30}$$

In other words, if $\mathbf{c} \in \mathcal{C}_n$, according to the definition of $\mathcal{C}_n$ above, $\arg \max_{\rho_j \in [0, \rho]} \mathcal{L}_j(\rho_j, \mathbf{c})$ is located at the n th “turning point” $r_{j,n}$, then $\mathcal{U}_j^-(\mathbf{c})$ and $\mathcal{U}_j^+(\mathbf{c})$ are, respectively, the left and right regions surrounding $r_{j,n}$.

With the above definitions, we demonstrate how to bound the UCB error. 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 , then we can decompose the UCB error as follows

$$\begin{aligned} \sum_{t > K} \mathcal{L}_j(\rho_j^*(t), \mathbf{c}_t) - \mathcal{L}_j(\rho_{j,t}, \mathbf{c}_t) &= X_1 + X_2 + X_3 \quad \text{where} \\ X_1 &= \sum_{t > K : \mathbf{c}_t \notin \mathcal{C}(\underline{\sigma})} \sum_{k \in [K]} \Delta_k(\mathbf{c}_t) \mathbb{I}\{\rho_{j,t} = a_k, N_{k,t} \leq m_k(\mathbf{c}_t)\} \\ X_2 &= \sum_{t > K : \mathbf{c}_t \in \mathcal{C}(\underline{\sigma})} \sum_{k \in [K]} \Delta_k(\mathbf{c}_t) \mathbb{I}\{\rho_{j,t} = a_k, N_{k,t} \leq m_k(\mathbf{c}_t)\} \\ X_3 &= \sum_{k \in [K]} \sum_{t > K} \Delta_k(\mathbf{c}_t) \mathbb{I}\{\rho_{j,t} = a_k, N_{k,t} > m_k(\mathbf{c}_t)\}. \end{aligned} \tag{31}$$

In Section C.5.1, we show that $X_1 \leq \tilde{\mathcal{O}}(\delta T + \underline{\sigma}T + \frac{1}{\delta})$; in Section C.5.2, we show that $X_2 \leq \tilde{\mathcal{O}}(\delta T + \frac{1}{\delta \underline{\sigma}})$; in Section C.5.3, we show that $X_3 \leq \tilde{\mathcal{O}}(\frac{1}{\delta T})$. Here, we provide some high-level discussion about the key challenges we face when bounding X_i , $i \in [3]$. To bound X_3 , which represents the regret when an arm a_k is played a sufficient number of times under context $\mathbf{c}_t$, our analysis, while being more involved, closely follows the classical proof to bound the regret of the UCB algorithm.

The main challenge lies in bounding X_1 and X_2 , where most of the complexity arises. To bound X_1 and X_2 , which correspond to the regret when we lack sufficient observations for budget a_k (i.e., $N_{k,t} \leq m_k(\mathbf{c}_t) = \frac{8 \log(T)}{\Delta_k^2(\mathbf{c}_t)}$), we must lower bound the Lagrangian differences $\mathcal{L}_j(\rho_j, \mathbf{c}_t) - \mathcal{L}_j(a_k, \mathbf{c}_t)$ over the entire sequence of time-varying dual variables $\mathbf{c}_t = (\lambda_t, \mu_t)$.

As seen in the definitions of X_1 and X_2 , we partition the rounds into two sets: the first set contains rounds t with $\mathbf{c}_t \notin \mathcal{C}(\underline{\sigma})$, and the second contains rounds t with $\mathbf{c}_t \in \mathcal{C}(\underline{\sigma})$. Here, $\mathcal{C}(\underline{\sigma}) = \{\mathbf{c} \in \{\mathbf{c}_t\}_{t \in [T]} : \sigma_j^-(\mathbf{c}) > \underline{\sigma}, |\sigma_j^+(\mathbf{c})| > \underline{\sigma}\}$, as formally defined in Eq. (29). This is the set of all contexts $\mathbf{c}$ where the adjacent slopes to the optimal point with respect to $\mathbf{c}$ (i.e., $\arg \max_{\rho_j \geq 0} \mathcal{L}_j(\rho_j, \mathbf{c})$) have magnitudes greater than $\underline{\sigma}$, meaning the adjacent slopes are steep. In other words, for any $\mathbf{c}_t \in \mathcal{C}(\underline{\sigma})$, the Lagrangian function is steep around its optimal budget, implying that $\Delta_k(\mathbf{c}_t)$ is large. Conversely, when $\mathbf{c}_t \notin \mathcal{C}(\underline{\sigma})$, the Lagrangian function is not steep around its optimal budget, implying that $\Delta_k(\mathbf{c}_t)$ is small.

This decomposition into steep and non-steep contexts is critical in the presence of continuously varying dual variables, because it allows us to tailor the regret analysis to the local geometry of the Lagrangian. Specifically, it enables us to isolate contexts where the Lagrangian is flat (making learning difficult and requiring more careful tracking) from those where the Lagrangian is sharp (where mistakes are costlier but learning is faster). By doing so, we ensure that the regret is controlled uniformly across time-varying contexts.

When $\Delta_k(\mathbf{c}_t)$ is small (i.e., to bound X_1 , $\mathbf{c}_t \notin \mathcal{C}(\underline{\sigma})$), distinguishing the optimal arm from a suboptimal one becomes challenging. In this case, the piecewise linear structure of the Lagrangian function is crucial in the analysis (see Eq. (37)) and in defining $\mathcal{C}(\underline{\sigma})$. More importantly, we leverage cross-learning between dual variables, as detailed in the proof between Eqs. (34) and (35). To illustrate cross-learning, consider two contexts $\mathbf{c}_t = (\lambda_t, \mu_t)$ and $\mathbf{c} = (\lambda, \mu)$, and fix an action ρ_j and a noise realization $\mathbf{z}_j$. Suppose that at round t , the learner selects ρ_j and observes the reward

$$\mathcal{L}_j(\rho_j, \mathbf{c}_t; \mathbf{z}_j) = (1 + \lambda_t)V_j(\rho_j; \mathbf{z}_j) - (\lambda_t\gamma + \mu_t)\rho_j.$$

Given this observation and the known form of the Lagrangian, the learner can recover $V_j(\rho_j; \mathbf{z}_j)$, and then use it to compute

$$\mathcal{L}_j(\rho_j, \mathbf{c}; \mathbf{z}_j) = (1 + \lambda)V_j(\rho_j; \mathbf{z}_j) - (\lambda\gamma + \mu)\rho_j$$

for any other $\mathbf{c} = (\lambda, \mu)$. This means a single observation at $\mathbf{c}_t$ allows the learner to evaluate the Lagrangian for the same action ρ_j across all other contexts. Hence, we say the learner can “cross-learn” the reward values across contexts without additional exploration.

When $\Delta_k(\mathbf{c}_t)$ is large (i.e., to bound X_2 , $\mathbf{c}_t \in \mathcal{C}(\underline{\sigma})$), distinguishing the optimal arm from suboptimal ones is easier, but each mistake incurs larger regret. Similar to the case of bounding X_1 , we use cross-learning between contexts and the piecewise linear structure of the Lagrangian function to bound X_2 .

Overall, the gap $\Delta_k(\mathbf{c}_t)$ depends on the time-varying dual variables $\mathbf{c}_t$. To address this challenge, we leverage cross-learning between contexts and exploit the structure of the Lagrangian function and the reward function. This approach enables us to effectively bound X_1 and X_2 , which are the primary contributors to the complexity of the regret analysis.

Remark C.1 *In the following Sections C.5.1, C.5.2 and C.5.3 where we bound X_1 , X_2 , and X_3 , respectively, we assume the optimal per-channel $\rho_j^*(t) = \arg \max_{\rho_j \in [0, \rho]} \mathcal{L}_j(\rho_j, \mathbf{c}_t)$ lies in the arm set $\mathcal{A}(\delta)$ for all t . This is because otherwise, we can consider the following decomposition of the UCB error in period t as follows:*

$$\mathcal{L}_j(\rho_j^*(t), \mathbf{c}_t) - \mathcal{L}_j(\rho_{j,t}, \mathbf{c}_t) = \mathcal{L}_j(\rho_j^*(t), \mathbf{c}_t) - \mathcal{L}_j(a_t^*, \mathbf{c}_t) + \mathcal{L}_j(a_t^*, \mathbf{c}_t) - \mathcal{L}_j(\rho_{j,t}, \mathbf{c}_t) \quad \text{where } a_t^* = \arg \max_{a_k \in \mathcal{A}(\delta)} \mathcal{L}_j(a_k, \mathbf{c}_t)$$

The first term will yield an error in the order of $\mathcal{O}(\delta)$ due to the Lagrangian function being unimodal, piecewise linear, which implies $|a_t^* - \rho_j^*(t)| \leq \delta$ so that $\mathcal{L}_j(\rho_j^*(t), \mathbf{c}_t) - \mathcal{L}_j(a_t^*, \mathbf{c}_t) = \mathcal{O}(\delta)$. Hence, this “discretization error” will accumulate to a magnitude of $\mathcal{O}(\delta T)$ over T periods, leading to an additional error that is already taken into account in the lemma statement.

C.5.1. Bounding X_1 . Our strategy to bound $X_1 = \sum_{t > K: \mathbf{c}_t \notin \mathcal{C}(\underline{\sigma})} \sum_{k \in [K]} \Delta_k(\mathbf{c}_t) \mathbb{I}\{\rho_{j,t} = a_k, N_{k,t} \leq m_k(\mathbf{c}_t)\}$ consists of 4 steps, namely bounding the loss of arm a_k at each context $\mathbf{c} \notin \mathcal{C}(\underline{\sigma}) = \{\mathbf{c} \in \{\mathbf{c}_t\}_{t \in [T]} : \sigma_j^-(\mathbf{c}) > \underline{\sigma}, |\sigma_j^+(\mathbf{c})| > \underline{\sigma}\}$, when (i) $a_k \in \mathcal{U}_j^-(\mathbf{c})$ lies on the left adjacent region of the optimal budget; (ii) $a_k < \min \mathcal{U}_j^-(\mathbf{c})$ lies to the left of the left adjacent region; (iii) $a_k \in \mathcal{U}_j^+(\mathbf{c})$ lies on the right adjacent region of the optimal budget; and (iv) $a_k > \max \mathcal{U}_j^+(\mathbf{c})$ lies to the right of the right adjacent region. Here, we recall that the adjacent regions are defined in Eq. (30). We note that while the analysis in the first and third steps is straightforward, the analysis in step 2 and step 4 is more involved, and this is where cross-learning between contexts is used.

Step 1: $a_k \in \mathcal{U}_j^-(\mathbf{c}_t)$. For arm k such that $a_k \in \mathcal{U}_j^-(\mathbf{c}_t)$, recall Lemma 4.3 that $\mathcal{L}_j(a, \mathbf{c}_t)$ is linear in a for $a \in \mathcal{U}_j^-(\mathbf{c}_t)$, so $\Delta_k(\mathbf{c}_t) = \sigma_j^-(\mathbf{c}_t) \cdot (\rho_j^*(t) - a_k) \leq \underline{\sigma} \rho$ where we used the condition that $\mathbf{c}_t \notin \mathcal{C}(\underline{\sigma})$ so the adjacent slopes have magnitude at most $\underline{\sigma}$, and $\rho_j^*(t) \leq \rho$. Thus, summing over all such k we get

$$\begin{aligned} & \sum_{t > K: \mathbf{c}_t \notin \mathcal{C}(\underline{\sigma})} \sum_{k \in [K]: a_k \in \mathcal{U}_j^-(\mathbf{c}_t)} \Delta_k(\mathbf{c}_t) \mathbb{I}\{\rho_{j,t} = a_k, N_{k,t} \leq m_k(\mathbf{c}_t)\} \\ & \leq \sum_{t > K: \mathbf{c}_t \notin \mathcal{C}(\underline{\sigma})} \sum_{k \in [K]: a_k \in \mathcal{U}_j^-(\mathbf{c}_t)} \underline{\sigma} \rho \cdot \mathbb{I}\{\rho_{j,t} = a_k, N_{k,t} \leq m_k(\mathbf{c}_t)\} \leq \underline{\sigma} \rho T = \mathcal{O}(\underline{\sigma} T). \end{aligned} \tag{32}$$

Step 2: $a_k < \min \mathcal{U}_j^-(\mathbf{c}_t)$. For arm k such that $a_k < \min \mathcal{U}_j^-(\mathbf{c}_t)$, we further split contexts into groups $\mathcal{C}_n = \{\mathbf{c} \in \{\mathbf{c}_t\}_{t \in [T]} : r_{j,n} = \arg \max_{\rho_j \geq 0} \mathcal{L}_j(\rho_j, \mathbf{c})\}$ for $n = 0 \dots S_j$ based on whether the corresponding optimal budget w.r.t. the Lagrangian at the context is taken at the n th “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

$$\begin{aligned}
& \sum_{t > K : \mathbf{c}_t \in \mathcal{C}_n / \mathcal{C}(\underline{\sigma})} \sum_{k \in [K] : a_k < \min \mathcal{U}_j^-(\mathbf{c}_t)} \Delta_k(\mathbf{c}_t) \mathbb{I}\{\rho_{j,t} = a_k, N_{k,t} \leq m_k(\mathbf{c}_t)\} \\
& \stackrel{(i)}{=} \sum_{t > K : \mathbf{c}_t \in \mathcal{C}_n / \mathcal{C}(\underline{\sigma})} \sum_{k \in [K] : a_k < r_{j,n-1}} \Delta_k(\mathbf{c}_t) \mathbb{I}\{\rho_{j,t} = a_k, N_{k,t} \leq m_k(\mathbf{c}_t)\} \\
& = \sum_{t > K} \sum_{\mathbf{c} \in \mathcal{C}_n / \mathcal{C}(\underline{\sigma})} \sum_{k \in [K] : a_k < r_{j,n-1}} \Delta_k(\mathbf{c}) \mathbb{I}\{\mathbf{c}_t = \mathbf{c}, \rho_{j,t} = a_k, N_{k,t} \leq m_k(\mathbf{c})\} \\
& \stackrel{(ii)}{\leq} \sum_{t > K} \sum_{\mathbf{c} \in \mathcal{C}_n / \mathcal{C}(\underline{\sigma})} \left(\Delta_{k'}(\mathbf{c}) \mathbb{I}\{\mathbf{c}_t = \mathbf{c}\} + \sum_{k \in [K] : a_k < r_{j,n-1} - \delta} \Delta_k(\mathbf{c}) \mathbb{I}\{\mathbf{c}_t = \mathbf{c}, \rho_{j,t} = a_k, N_{k,t} \leq m_k(\mathbf{c})\} \right) \\
& \stackrel{(iii)}{\leq} ((1 + C_F)s_{j,n-1}\delta + \rho\underline{\sigma})T + \sum_{k \in [K] : a_k < r_{j,n-1} - \delta} \sum_{\mathbf{c} \in \mathcal{C}_n / \mathcal{C}(\underline{\sigma})} \Delta_k(\mathbf{c}) Y_k(\mathbf{c})
\end{aligned} \tag{33}$$

where in the final equality we defined $Y_k(\mathbf{c}) = \sum_{t > K} \mathbb{I}\{\mathbf{c}_t = \mathbf{c}, \rho_{j,t} = a_k, N_{k,t} \leq m_k(\mathbf{c})\}$. In (i) we used the fact that the left end of the left adjacent region, i.e. $\min \mathcal{U}_j^-(\mathbf{c}_t)$ is exactly $r_{j,n-1}$ because for context $\mathbf{c}_t \in \mathcal{C}_n$ the optimal budget $\arg \max_{\rho_j \in [0, \rho]} \mathcal{L}_j(\rho_j, \mathbf{c}_t)$ is at the n th turning point; in (ii) we used the definition $k' := \max\{k : a_k < r_{j,n-1}\}$ where we recall arms are indexed such that $a_1 < a_2 < \dots < a_K$. Note that in (ii) we separate out the arm $a_{k'}$ because its distance to the optimal per-channel may be less than δ since it is the closest arm, and thus we ensure all other arms indexed by $k \in [K] : a_k < r_{j,n-1} - \delta$, are at least δ away from the optimal per-channel budget; (iii) follows from the fact that under a context $\mathbf{c} \in \mathcal{C}_n / \mathcal{C}(\underline{\sigma})$, we have $\arg \max_{\rho_j \in [0, \rho]} \mathcal{L}_j(\rho_j, \mathbf{c}) = r_{j,n}$ so

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

where in the second equality, we use that $\mathcal{L}(y, \mathbf{c})$ is piece-wise linear in y and by definition, the slope of $\mathcal{L}(y, \mathbf{c})$ when $y \in [r_{j,n-1}, r_{j,n}]$ is $\sigma_j^-(\mathbf{c})$ and the slope of $\mathcal{L}(y, \mathbf{c})$ when $y \in [a_{k'}, r_{j,n-1}]$ is $\sigma_{j,n-1}(\mathbf{c})$. In (iv) we used $\mathbf{c} \in \mathcal{C}_n / \mathcal{C}(\underline{\sigma})$ implies $\sigma_j^-(\mathbf{c}) \leq \underline{\sigma}$, as well as all $r_{j,n} \leq \rho$ for any n and the fact that 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 $\sigma_{j,n-1}(\mathbf{c}) = (1 + \lambda)s_{j,n-1} - (\mu + \gamma\lambda) \leq (1 + C_F)s_{j,n-1}$ where C_F is defined in Lemma 4.4.

We now bound $\sum_{\mathbf{c} \in \mathcal{C}_n / \mathcal{C}(\underline{\sigma})} \Delta_k(\mathbf{c}) Y_k(\mathbf{c})$ in Eq. (33). It is easy to see the following inequality for any sequence of context $\mathbf{c}_{(1)}, \dots, \mathbf{c}_{(\ell)} \in \{\mathbf{c}_t\}_{t \in [T]}$:

$$Y_k(\mathbf{c}_{(1)}) + \dots + Y_k(\mathbf{c}_{(\ell)}) \leq \max_{\ell'=1 \dots \ell} m_k(\mathbf{c}_{(\ell')}). \tag{34}$$

This is because

$$\sum_{\ell' \in [\ell]} Y_k(\mathbf{c}_{(\ell')}) = \sum_{t > K} \sum_{\ell' \in [\ell]} \mathbb{I}\{\mathbf{c}_t = \mathbf{c}_{(\ell')}, \rho_{j,t} = a_k, N_{k,t} \leq m_k(\mathbf{c}_{(\ell')})\}$$

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

An inequality similar to $Y_k(\mathbf{c}_{(1)}) + \dots + Y_k(\mathbf{c}_{(\ell)}) \leq \max_{\ell'=1 \dots \ell} m_k(\mathbf{c}_{(\ell')})$ that we just shown is proven in the proof of Lemma 3 of Balseiro et al. (2022a)).

For simplicity denote $L = |\mathcal{C}_n / \mathcal{C}(\underline{\sigma})|$, and order contexts in $\mathbf{c} \in \mathcal{C}_n / \mathcal{C}(\underline{\sigma})$ as $\{\mathbf{c}_{(\ell)}\}_{\ell \in [L]}$ s.t. $\Delta_k(\mathbf{c}_{(1)}) > \Delta_k(\mathbf{c}_{(2)}) > \dots > \Delta_k(\mathbf{c}_{(L)})$, or equivalently $m_k(\mathbf{c}_{(1)}) < m_k(\mathbf{c}_{(2)}) < \dots < m_k(\mathbf{c}_{(L)})$ according to Eq.(29). Then multiplying Eq. (34) by $\Delta_k(\mathbf{c}_{(\ell)}) - \Delta_k(\mathbf{c}_{(\ell+1)})$ (which is strictly positive due to the ordering of contexts), and summing $\ell = 1 \dots L$ we get

$$\begin{aligned}
\sum_{\mathbf{c} \in \mathcal{C}_n / \mathcal{C}(\underline{\sigma})} \Delta_k(\mathbf{c}) Y_k(\mathbf{c}) &= \sum_{\ell \in [L]} \Delta_k(\mathbf{c}_{(\ell)}) Y_k(\mathbf{c}_{(\ell)}) \\
&\leq \sum_{\ell \in [L]} m_k(\mathbf{c}_{(\ell)}) (\Delta_k(\mathbf{c}_{(\ell)}) - \Delta_k(\mathbf{c}_{(\ell+1)})) \\
&\stackrel{(i)}{=} 8 \log(T) \sum_{\ell \in [L-1]} \frac{\Delta_k(\mathbf{c}_{(\ell)}) - \Delta_k(\mathbf{c}_{(\ell+1)})}{\Delta_k^2(\mathbf{c}_{(\ell)})} \stackrel{(ii)}{\leq} 8 \log(T) \int_{\Delta_k(\mathbf{c}_{(L)})}^{\infty} \frac{dz}{z^2} \\
&= \frac{8 \log(T)}{\Delta_k(\mathbf{c}_{(L)})} \stackrel{(iii)}{=} \frac{8 \log(T)}{\min_{\mathbf{c} \in \mathcal{C}_n / \mathcal{C}(\underline{\sigma})} \Delta_k(\mathbf{c})}.
\end{aligned} \tag{35}$$

Here (i) follows from the definition of $m_k(\mathbf{c})$ in Eq. (29) where $m_k(\mathbf{c}) = \frac{8 \log(T)}{\Delta_k^2(\mathbf{c})}$; both (ii) and (iii) follow from the ordering of contexts so that $\Delta_k(\mathbf{c}_{(1)}) > \Delta_k(\mathbf{c}_{(2)}) > \dots > \Delta_k(\mathbf{c}_{(L)})$. Note that for any $\mathbf{c} \in \mathcal{C}_n / \mathcal{C}(\underline{\sigma})$ and arm k such that $a_k < r_{j,n-1}$, we have

$$\begin{aligned}
\Delta_k(\mathbf{c}) &= \mathcal{L}_j(r_{j,n}, \mathbf{c}) - \mathcal{L}_j(r_{j,n-1}, \mathbf{c}) + \mathcal{L}_j(r_{j,n-1}, \mathbf{c}) - \mathcal{L}_j(a_k, \mathbf{c}) \\
&> \mathcal{L}_j(r_{j,n-1}, \mathbf{c}) - \mathcal{L}_j(a_k, \mathbf{c}) \\
&\stackrel{(i)}{\geq} \sigma_{j,n-1}(\mathbf{c})(r_{j,n-1} - a_k) \\
&\stackrel{(ii)}{\geq} (\sigma_{j,n-1}(\mathbf{c}) - \sigma_{j,n}(\mathbf{c}))(r_{j,n-1} - a_k) \\
&\stackrel{(iii)}{=} (1 + \lambda)(s_{j,n-1} - s_{j,n})(r_{j,n-1} - a_k) \\
&> (s_{j,n-1} - s_{j,n})(r_{j,n-1} - a_k),
\end{aligned} \tag{36}$$

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

Finally combining Eqs. (33), (35) and (36), and summing over $n = 1 \dots S_j$ we get

$$\begin{aligned}
& \sum_{t > K: \mathbf{c}_t \notin \mathcal{C}(\underline{\sigma})} \sum_{k \in [K]: a_k < \min \mathcal{U}_j^-(\mathbf{c}_t)} \Delta_k(\mathbf{c}_t) \mathbb{I}\{\rho_{j,t} = a_k, N_{k,t} \leq m_k(\mathbf{c}_t)\} \\
&= \sum_{n \in [S_j]} \sum_{t > K: \mathbf{c}_t \in \mathcal{C}_n / \mathcal{C}(\underline{\sigma})} \sum_{k \in [K]: a_k < \min \mathcal{U}_j^-(\mathbf{c}_t)} \Delta_k(\mathbf{c}_t) \mathbb{I}\{\rho_{j,t} = a_k, N_{k,t} \leq m_k(\mathbf{c}_t)\} \\
&\leq \sum_{n \in [S_j]} ((1 + C_F) s_{j,n-1} \delta + \rho \underline{\sigma}) T + \sum_{n \in [S_j]} \sum_{k \in [K]: a_k < r_{j,n-1} - \delta} \frac{8 \log(T)}{(s_{j,n-1} - s_{j,n})(r_{j,n-1} - a_k)} \\
&\stackrel{(i)}{\leq} \sum_{n \in [S_j]} ((1 + C_F) s_{j,n-1} \delta + \rho \underline{\sigma}) T + \sum_{n \in [S_j]} \sum_{\ell=1}^K \frac{8 \log(T)}{(s_{j,n-1} - s_{j,n}) \ell \delta} \\
&\leq \sum_{n \in [S_j]} ((1 + C_F) s_{j,n-1} \delta + \rho \underline{\sigma}) T + \frac{8 \log(T) \log(K)}{\delta} \sum_{n \in [S_j]} \frac{1}{(s_{j,n-1} - s_{j,n})} \\
&= \tilde{\mathcal{O}}(\delta T + \underline{\sigma} T + \frac{1}{\delta}).
\end{aligned} \tag{37}$$

In the first inequality, the first term (i.e., $\sum_{n \in [S_j]} ((1 + C_F) s_{j,n-1} \delta + \rho \underline{\sigma}) T$) is the first term in the last inequality of Eq. (33) when we sum over all possible values for n . The second term is an upper bound on $\sum_{k \in [K]: a_k < r_{j,n-1} - \delta} \sum_{\mathbf{c} \in \mathcal{C}_n / \mathcal{C}(\underline{\sigma})} \Delta_k(\mathbf{c}) Y_k(\mathbf{c})$ (see the second term in the last inequality of Eq. (33)). We established this upper bound in Eqs. (35) and (36). Note that (i) follows because for all $a_k < r_{j,n-1} - \delta$, the distances a_k from $r_{j,n-1}$ are at least $\delta, 2\delta, 3\delta, \dots$. In the last equation, we hide all logarithmic factors using the notation $\tilde{\mathcal{O}}$, and note that the constants $C_F, (s_{j,n})_{n \in S_j}, S_j$ are all absolute constants that depend only on the support F_j and corresponding sampling distribution $\mathbf{p}_j$ for value-cost pairs; see definitions of these absolute constants in Lemmas 4.3 and 4.4.

Step 3 and 4: $a_k \in \mathcal{U}_j^+(\mathbf{c}_t)$ or $a_k > \max \mathcal{U}_j^+(\mathbf{c}_t)$. The cases where arm $a_k \in \mathcal{U}_j^+(\mathbf{c}_t)$ and $a_k > \max \mathcal{U}_j^+(\mathbf{c}_t)$ are symmetric to $a_k \in \mathcal{U}_j^-(\mathbf{c}_t)$ and $a_k < \min \mathcal{U}_j^-(\mathbf{c}_t)$, respectively, and we omit from this paper.

Therefore, combining Eqs. (32) and (37) we can conclude

$$X_1 \leq \tilde{\mathcal{O}}(\delta T + \underline{\sigma} T + \frac{1}{\delta}). \tag{38}$$

C.5.2. Bounding X_2 . We first rewrite $X_2 = \sum_{t > K: \mathbf{c}_t \in \mathcal{C}(\underline{\sigma})} \sum_{k \in [K]} \Delta_k(\mathbf{c}_t) \mathbb{I}\{\rho_{j,t} = a_k, N_{k,t} \leq m_k(\mathbf{c}_t)\}$ as follows

$$\begin{aligned}
X_2 &= \sum_{t > K: \mathbf{c}_t \in \mathcal{C}(\underline{\sigma})} \sum_{k \in [K]} \Delta_k(\mathbf{c}_t) \mathbb{I}\{\rho_{j,t} = a_k, N_{k,t} \leq m_k(\mathbf{c}_t)\} \\
&= \sum_{t > K} \sum_{n \in [S_j]} \sum_{k \in [K]} \sum_{\mathbf{c} \in \mathcal{C}_n \cap \mathcal{C}(\underline{\sigma})} \Delta_k(\mathbf{c}) \mathbb{I}\{\mathbf{c}_t = \mathbf{c}, \rho_{j,t} = a_k, N_{k,t} \leq m_k(\mathbf{c})\} \\
&\stackrel{(i)}{=} \sum_{n \in [S_j]} \sum_{k \in [K]} \sum_{\mathbf{c} \in \mathcal{C}_n \cap \mathcal{C}(\underline{\sigma})} \Delta_k(\mathbf{c}) Y_k(\mathbf{c}) \\
&\stackrel{(ii)}{=} \sum_{n \in [S_j]} \sum_{\mathbf{c} \in \mathcal{C}_n \cap \mathcal{C}(\underline{\sigma})} \sum_{k \in \{k_n^-, k_n^+\}} \Delta_k(\mathbf{c}) Y_k(\mathbf{c}) + \sum_{n \in [S_j]} \sum_{\mathbf{c} \in \mathcal{C}_n \cap \mathcal{C}(\underline{\sigma})} \sum_{k \in [K] / \{k_n^-, k_n^+\}} \Delta_k(\mathbf{c}) Y_k(\mathbf{c}) \\
&\stackrel{(iii)}{\leq} T \delta (1 + C_F) \sum_{n \in [S_j]} (s_{j,n} + s_{j,n+1}) + \sum_{n \in [S_j]} \sum_{\mathbf{c} \in \mathcal{C}_n \cap \mathcal{C}(\underline{\sigma})} \sum_{k \in [K] / \{k_n^-, k_n^+\}} \Delta_k(\mathbf{c}) Y_k(\mathbf{c}).
\end{aligned} \tag{39}$$

Here, we recall that $\mathcal{C}(\underline{\sigma}) = \{\mathbf{c} \in \{\mathbf{c}_t\}_{t \in [T]} : \sigma_j^-(\mathbf{c}) > \underline{\sigma}, |\sigma_j^+(\mathbf{c})| > \underline{\sigma}\}$. In (i) we define $Y_k(\mathbf{c}) = \sum_{t > K} \mathbb{I}\{\mathbf{c}_t = \mathbf{c}, \rho_{j,t} = a_k, N_{k,t} \leq m_k(\mathbf{c})\}$; in (ii) we separate out two arms k_n^- and k_n^+ defined as follows: recall for context

$\mathbf{c} \in \mathcal{C}_n \cap \mathcal{C}(\underline{\sigma})$, the optimal budget $\arg \max_{\rho_j \in [0, \rho]} \mathcal{L}_j(\rho_j, \mathbf{c}) = r_{j,n}$ is taken at the n th turning point per the definition of $\mathcal{C}_n$ in Eq. (29), and thereby 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 $r_{j,n}$; in (iii), for small enough $\delta < \min_{n' \in [S_j]} r_{j,n'} - r_{j,n'-1}$, we know that k_n^- lies on the line segment between $r_{j,n-1}$ and $r_{j,n}$, so $\Delta_{k_n^-}(\mathbf{c}) = \sigma_j^-(\mathbf{c})(r_{j,n} - a_{k_n^-}) \leq \sigma_j^-(\mathbf{c})\delta \leq (1 + C_F)s_{j,n-1}\delta$, where in the final inequality follows from the definition of $\sigma_j^-(\mathbf{c}) = \sigma_{j,n}(\mathbf{c}) = (1 + \lambda)s_{j,n} - (\mu + \gamma\lambda) \leq (1 + \lambda)s_{j,n} \leq (1 + C_F)s_{j,n}$ where C_F is defined in Lemma (4.4). A similar bound holds for $\Delta_{k_n^+}(\mathbf{c})$.

Then, following the same logic as in Eqs. (34), (35), (36) in Section C.5.1, where we bound X_1 , we can bound $\sum_{\mathbf{c} \in \mathcal{C}_n \cap \mathcal{C}(\underline{\sigma})} \Delta_k(\mathbf{c})Y_k(\mathbf{c})$ as follows for any arm $k \in [K] \setminus \{k_n^-, k_n^+\}$, i.e., arms that are at least δ away from the optimal per-channel budget with respect to $\mathbf{c}$:

$$\sum_{\mathbf{c} \in \mathcal{C}_n \cap \mathcal{C}(\underline{\sigma})} \Delta_k(\mathbf{c})Y_k(\mathbf{c}) \leq \frac{8 \log(T)}{\min_{\mathbf{c} \in \mathcal{C}_n \cap \mathcal{C}(\underline{\sigma})} \Delta_k(\mathbf{c})}. \quad (40)$$

Now, the set $k \in [K] \setminus \{k_n^-, k_n^+\}$ in Eq. (39) can be further split into two subsets, 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 $\mathbf{c} \in \mathcal{C}_n \cap \mathcal{C}(\underline{\sigma})$,

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

where the final inequality follows from the definition of $\mathcal{C}(\underline{\sigma})$ in Eq. (29) such that $\sigma_j^-(\mathbf{c}) \geq \underline{\sigma}$ for $\mathbf{c} \in \mathcal{C}(\underline{\sigma})$. Hence combining this with Eq. (40) we have

$$\sum_{k \in [K] : a_k < r_{j,n} - \delta} \sum_{\mathbf{c} \in \mathcal{C}_n \cap \mathcal{C}(\underline{\sigma})} \Delta_k(\mathbf{c})Y_k(\mathbf{c}) \leq \sum_{k \in [K] : a_k < r_{j,n} - \delta} \frac{8 \log(T)}{\underline{\sigma}(r_{j,n} - a_k)} \stackrel{(i)}{\leq} \sum_{\ell=1}^K \frac{8 \log(T)}{\underline{\sigma}\ell\delta} \leq \frac{8 \log(T) \log(K)}{\underline{\sigma}\delta}, \quad (41)$$

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 $\delta, 2\delta, 3\delta, \dots$. Symmetrically, we can show an identical bound for the set $\{k \in [K] : a_k > r_{j,n} + \delta\}$. Hence, combining Eqs. (39) and (41) we can conclude

$$X_2 \leq \tilde{O}\left(\delta T + \frac{1}{\delta \underline{\sigma}}\right). \quad (42)$$

Here, similar to our bound in Eq. (37) for bounding X_1 , we hide all logarithmic factors using the notation $\tilde{O}$, and note that the constants C_F , $(s_{j,n})_{n \in S_j}$, and S_j are all absolute constants that depend only on the support F_j and the corresponding sampling distribution $\mathbf{p}_j$ for value-cost pairs. See the definitions of these absolute constants in Lemmas 4.3 and 4.4.

C.5.3. Bounding X_3 . We first define

$$\bar{\mathcal{L}} = (1 + \gamma) \rho C_F + (1 + C_F) \bar{V}, \quad (43)$$

where C_F is specified in Lemma 4.4. Recalling the definition $\Delta_k(\mathbf{c}) = \max_{\rho_j \in [0, \rho]} \mathcal{L}_j(\rho_j, \mathbf{c}) - \mathcal{L}_j(a_k, \mathbf{c})$ in Eq. (29), and $-(1 + \gamma) \rho C_F \leq \mathcal{L}_j(\rho_j, \mathbf{c}) \leq (1 + C_F) \bar{V}$ for any $\rho_j \in [0, \rho]$ and context $\mathbf{c}$ (see Lemma 4.4), it is easy to see

$$\Delta_k(\mathbf{c}) \leq \bar{\mathcal{L}} \quad \forall k \in [K], \forall \mathbf{c}. \quad (44)$$

Then we bound X_3 as follows

$$\begin{aligned} X_3 &= \sum_{k \in [K]} \sum_{t > K} \mathbb{E} [\Delta_k(\mathbf{c}) \mathbb{I}\{\rho_{j,t} = a_k, N_{k,t} > m_k(\mathbf{c}_t)\}] \\ &\stackrel{(i)}{\leq} \bar{\mathcal{L}} \cdot \sum_{k \in [K]} \sum_{t > K} \mathbb{P}(\rho_{j,t} = a_k, N_{k,t} > m_k(\mathbf{c}_t)) \\ &\stackrel{(ii)}{\leq} \bar{\mathcal{L}} \cdot \sum_{k \in [K]} \sum_{t > K} \mathbb{P}\left(\hat{V}_{j,t}(a_k) - \frac{\lambda_t \gamma + \mu_t}{1 + \lambda_t} a_k + \text{UCB}_{j,t}(a_k) \geq \hat{V}_{j,t}(\rho_j^*(t)) - \frac{\lambda_t \gamma + \mu_t}{1 + \lambda_t} \rho_j^*(t) + \text{UCB}_{j,t}(\rho_j^*(t)), \right. \\ &\quad \left. N_{k,t} > m_k(\mathbf{c}_t)\right), \end{aligned} \quad (45)$$

where (i) follows from Eq. (44); in (ii), recall that we choose arm $\rho_{j,t} = a_k$ because the estimated UCB rewards of arm a_k are greater than those of any other arm, including $\rho_j^*(t)$, according to the DSD-UCB (Algorithm 1), or mathematically,

$$\hat{V}_{j,t}(a_k) - \frac{\lambda_t \gamma + \mu_t}{1 + \lambda_t} a_k + \text{UCB}_{j,t}(a_k) \geq \hat{V}_{j,t}(\rho_j^*(t)) - \frac{\lambda_t \gamma + \mu_t}{1 + \lambda_t} \rho_j^*(t) + \text{UCB}_{j,t}(\rho_j^*(t)).$$

Here, we also used the fact that $\rho_j^*(t)$ lies in the arm set $\mathcal{A}(\delta)$ for all t (see Remark C.1).

Now, let $\hat{R}_n(a_k)$ denote the average conversion of arm k over its first n pulls, i.e.,

$$\hat{R}_n(a_k) = \hat{V}_{j,\tau}(a_k) \text{ for } \tau = \min\{t \in [T] : N_{k,t} = n\}, \quad (46)$$

where we recall $\hat{V}_{j,\tau}(a_k)$ is the estimated conversion for arm a_k in channel j during period τ as defined in Algorithm 1. In other words, τ is the period during which arm a_k is pulled for the n th time so $\hat{R}_n(a_k) = \hat{V}_{j,\tau}(a_k)$.

Hence, we continue with Eq. (45) as follows:

$$\begin{aligned} &\mathbb{P}\left(\hat{V}_{j,t}(a_k) - \frac{\lambda_t \gamma + \mu_t}{1 + \lambda_t} a_k + \text{UCB}_{j,t}(a_k) \geq \hat{V}_{j,t}(\rho_j^*(t)) - \frac{\lambda_t \gamma + \mu_t}{1 + \lambda_t} \rho_j^*(t) + \text{UCB}_{j,t}(\rho_j^*(t)), N_{k,t} > m_k(\mathbf{c}_t)\right) \\ &\leq \mathbb{P}\left(\max_{n: m_k(\mathbf{c}_t) < n \leq t} \left\{ \hat{R}_n(a_k) + \text{UCB}_n(a_k) - \frac{\lambda_t \gamma + \mu_t}{1 + \lambda_t} a_k \right\} \right. \\ &\quad \left. \geq \min_{n': 1 \leq n' \leq t} \left\{ \hat{R}_{n'}(\rho_j^*(t)) + \text{UCB}_{n'}(\rho_j^*(t)) - \frac{\lambda_t \gamma + \mu_t}{1 + \lambda_t} \rho_j^*(t) \right\}\right) \\ &\leq \sum_{n = \lceil m_k(\mathbf{c}_t) \rceil + 1}^t \sum_{n' = 1}^t \mathbb{P}\left(\hat{R}_n(a_k) + \text{UCB}_n(a_k) - \frac{\lambda_t \gamma + \mu_t}{1 + \lambda_t} a_k > \hat{R}_{n'}(\rho_j^*(t)) + \text{UCB}_{n'}(\rho_j^*(t)) - \frac{\lambda_t \gamma + \mu_t}{1 + \lambda_t} \rho_j^*(t)\right) \end{aligned} \quad (47)$$

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

$$\begin{aligned} \mathcal{G}_{1,n} &= \left\{ \hat{R}_n(a_k) \geq V(a_k) + \text{UCB}_n(a_k) \right\} \quad \text{for } n \text{ s.t. } m_k(\mathbf{c}_t) < n \leq t \\ \mathcal{G}_{2,n'} &= \left\{ \hat{R}_{n'}(\rho_j^*(t)) \leq V(\rho_j^*(t)) - \text{UCB}_{n'}(\rho_j^*(t)) \right\} \quad \text{for } n' \text{ s.t. } 1 \leq n' \leq t \\ \mathcal{G}_3 &= \left\{ V_j(\rho_j^*(t)) - \frac{\lambda_t \gamma + \mu_t}{1 + \lambda_t} \rho_j^*(t) < V_j(a_k) - \frac{\lambda_t \gamma + \mu_t}{1 + \lambda_t} a_k + 2 \cdot \text{UCB}_n(a_k) \right\} \end{aligned} \quad (48)$$

Note that for $n > m_k(\mathbf{c}_t)$, we have $\text{UCB}_n(a_k) = \sqrt{\frac{2 \log(T)}{n}} < \sqrt{\frac{2 \log(T)}{m_k(\mathbf{c}_t)}} = \frac{\Delta_k(\mathbf{c}_t)}{2}$ since we defined $m_k(\mathbf{c}) = \frac{8 \log(T)}{\Delta_k^2(\mathbf{c})}$ in Eq. (29). Therefore

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

where (i) follows from the definition of $\Delta_k(\mathbf{c}) = \max_{a \in \mathcal{A}(\delta)} \mathcal{L}(a, \mathbf{c}) - \mathcal{L}(a_k, \mathbf{c})$ in Eq. (29) for any context $\mathbf{c}$. This implies that event $\mathcal{G}_3$ in Eq. (48) cannot hold for $n > m_k(\mathbf{c}_t)$. Therefore

$$\mathbb{P} \left(\hat{R}_n(a_k) + \text{UCB}_n(a_k) - \frac{\lambda_t \gamma + \mu_t}{1 + \lambda_t} a_k > \hat{R}_{n'}(\rho_j^*(t)) + \text{UCB}_{n'}(\rho_j^*(t)) - \frac{\lambda_t \gamma + \mu_t}{1 + \lambda_t} \rho_j^*(t) \right) \leq \mathbb{P}(\mathcal{G}_{1,n} \cup \mathcal{G}_{2,n'}). \quad (49)$$

From the 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 Eqs. (45) (47), (49) we can conclude

$$\begin{aligned} X_3 &\leq \sum_{k \in [K]} \sum_{t > K} \sum_{n = \lceil m_k(\mathbf{c}_t) \rceil + 1}^t \sum_{n' = 1}^t (\mathbb{P}(\mathcal{G}_{1,n}) + \mathbb{P}(\mathcal{G}_{2,n'})) \\ &\leq \sum_{k \in [K]} \sum_{t > K} \sum_{n = \lceil m_k(\mathbf{c}_t) \rceil + 1}^t \sum_{n' = 1}^t \frac{2\bar{V}}{T^4} \\ &\leq \frac{2K\bar{V}}{T} = \mathcal{O}\left(\frac{1}{\delta T}\right). \end{aligned} \quad (50)$$

Q.E.D.

C.6. Proof for Theorem 4.6

Starting from Proposition 4.1, we get

$$\begin{aligned} T \cdot \text{GL-OPT} - \mathbb{E} \left[\sum_{t \in [T]} \sum_{j \in [M]} V_j(\rho_{j,t}) \right] \\ \leq M\bar{V}(T - \tau_A) + \sum_{j \in [M]} \mathbb{E} \left[\sum_{t \in [\tau_A]} \mathcal{L}_j(\rho_j^*(t), \mathbf{c}_t) - \mathcal{L}_j(\rho_{j,t}, \mathbf{c}_t) \right] + \mathbb{E} \left[\sum_{t \in [\tau_A]} (\lambda_t g_{1,t} + \mu_t g_{2,t}) \right] \\ \stackrel{(i)}{\leq} M\bar{V}(T - \tau_A) + \mathcal{O} \left(\underline{\sigma} T + \delta T + \frac{1}{\underline{\sigma} \delta} \right) + \mathcal{O} \left(\eta T + \frac{1}{\eta} \right), \end{aligned} \quad (51)$$

where in (i) we applied Lemmas 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 $T \cdot \text{GL-OPT} - \mathbb{E} \left[\sum_{t \in [T]} \sum_{j \in [M]} V_j(\rho_{j,t}) \right] \leq \mathcal{O}(T^{2/3})$. According to Lemma 4.3, $V_j(\rho_j)$ is concave for all $j \in [M]$, so

$$\begin{aligned} \mathcal{O}(T^{-1/3}) &\geq \text{GL-OPT} - \frac{1}{T} \sum_{t \in [T]} \mathbb{E} \left[\sum_{j \in [M]} V_j(\rho_{j,t}) \right] \\ &\geq \text{GL-OPT} - \mathbb{E} \left[\sum_{j \in [M]} V_j \left(\frac{1}{T} \sum_{t \in [T]} \rho_{j,t} \right) \right] \\ &\geq \text{GL-OPT} - \mathbb{E} \left[\sum_{j \in [M]} V_j(\bar{\rho}_{j,T}) \right] \end{aligned} \quad (52)$$

where in the final equality we used the definition $\bar{\rho}_T = \left(\frac{1}{T} \sum_{t \in [T]} \rho_{j,t} \right)_{j \in [M]}$ as defined in Algorithm 1.

Regarding ROI constraint satisfaction, consider

$$\begin{aligned}
0 &\stackrel{(i)}{\leq} \frac{1}{T} \sum_{t \in [T]} \mathbb{E}[g_{1,t}] \\
&= \frac{1}{T} \sum_{t \in [T]} \sum_{j \in [M]} \mathbb{E}[V_j(\rho_{j,t}; \mathbf{z}_{j,t}) - \gamma \rho_{j,t}] \\
&= \frac{1}{T} \sum_{t \in [T]} \sum_{j \in [M]} \mathbb{E}[V_j(\rho_{j,t}) - \gamma \rho_{j,t}] \\
&\stackrel{(ii)}{\leq} \sum_{j \in [M]} \mathbb{E} \left[V_j \left(\frac{1}{T} \sum_{t \in [T]} \rho_{j,t} \right) - \gamma \cdot \frac{1}{T} \sum_{t \in [T]} \rho_{j,t} \right] \\
&= \sum_{j \in [M]} \mathbb{E}[V_j(\bar{\rho}_{j,T}) - \gamma \bar{\rho}_{j,T}].
\end{aligned} \tag{53}$$

where (i) follows from Lemma C.2; in (ii), we again apply the concavity of $V_j(\rho_j)$. We omit the analysis of the budget constraint, as it is similar to the above.

Q.E.D.

C.7. Additional Results for Section 4

Proposition C.1 *Assume Assumption 4.2 holds, and recall $\mathbf{z}_j = (\mathbf{v}_j, \mathbf{d}_j) \in F_j$ is any realization of values and costs for channel $j \in [M]$. Then, for any channel $j \in [M]$, we have $\min_{\mathbf{z}_j \in F_j} \frac{v_{j,1}}{d_{j,1}} > \gamma$, where we recall the ordering $\frac{v_{j,1}}{d_{j,1}} > \frac{v_{j,2}}{d_{j,2}} > \dots > \frac{v_{j,m_j}}{d_{j,m_j}}$ for any element $\mathbf{z}_j = (\mathbf{v}_j, \mathbf{d}_j) \in F_j$ (see Section 4). Further, there exists some $\tilde{\rho} \in (0, \rho)$ s.t. for any per-channel budget $\rho_j \leq \tilde{\rho}$, we have $V_j(\rho_j; \mathbf{z}_j) = \frac{v_{j,1}}{d_{j,1}} \rho_j > \gamma \rho_j$ for any $j \in [M]$.*

Proof. Under Assumption 4.2, it is easy to see for any realization of value-cost pairs $\mathbf{z}_j = (\mathbf{v}_j, \mathbf{d}_j)$ there always exists an auction $n \in [m_j]$ whose value-to-cost ratio is at least γ , i.e. $v_{j,n} > \gamma d_{j,n}$. Hence we know that $\frac{v_{j,1}}{d_{j,1}} \geq \frac{v_{j,n}}{d_{j,n}} > \gamma$. Now, in Eq. (26) within the proof of Lemma 4.3, we showed

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

where $d_{j,0} = v_{j,0} = b_{j,1} = 0$. This implies that for any $\rho_j < d_{j,1}$, we have $V_j(\rho_j; \mathbf{z}_j) = \frac{v_{j,1}}{d_{j,1}} \rho_j > \gamma \rho_j$. Therefore, we can take $\tilde{\rho} = \min_{j \in [M]} \min_{\mathbf{z}_j \in F_j} d_{j,1}$, which ensures that for any $\rho_j \leq \tilde{\rho}$ and realization $\mathbf{z}_j \in F_j$ we have $V_j(\rho_j; \mathbf{z}_j) = \frac{v_{j,1}}{d_{j,1}} \rho_j > \gamma \rho_j$ for any channel $j \in [M]$. Q.E.D.

Lemma C.2 (Constraint satisfaction) *Assume Assumption 4.2 holds, and consider $\beta = \underline{\rho} = \frac{1}{\log(T)}$ in Algorithm 1. Then, for large enough T we have*

$$\frac{1}{T} \sum_{t \in [T]} g_{1,t} \geq 0 \quad \text{and} \quad \frac{1}{T} \sum_{t \in [T]} \sum_{j \in [M]} \rho_{j,t} \leq \rho,$$

where we recall $g_{1,t} = \sum_{j \in [M]} (V_j(\rho_{j,t}; \mathbf{z}_{j,t}) - \gamma \rho_{j,t})$.

Proof. Recall $\tau_A \in [T]$ defined in step 10 of Algorithm 1.

If $\tau_A = T$, then we know that Algorithm 1 does not exit the while loop, and therefore $S_{1,t} - \gamma M \rho + \beta \rho(T-t) \geq 0$ for $t = T$, or equivalently $S_{1,T} \geq \gamma M \rho > 0$. Since we recall $S_{1,T} = \sum_{t \in [T-1]} g_{1,t}$, we can conclude that

$\sum_{t \in [T]} g_{1,t} = S_{1,T} + g_{1,T} \geq M\rho + g_{1,T} \geq 0$ since $g_{1,T} \geq -\gamma M\rho$. Similarly, we also have $S_{2,t} + M\rho + \underline{\rho}(T-t) \leq \rho T$ for $t = T$, or equivalently $S_{2,T} \leq \rho T - M\rho$ where we used the fact that $\underline{\rho} = 1/\log(T) < \rho$ for large enough T and $M \geq 2$. Hence, recalling $S_{2,T} = \sum_{t \in [T-1]} \sum_{j \in [M]} \rho_{j,t}$, we can conclude that $\sum_{t \in [T]} \sum_{j \in [M]} \rho_{j,t} = S_{2,T} + \sum_{j \in [M]} \rho_{j,T} \leq \rho T - M\rho + \sum_{j \in [M]} \rho_{j,T} \leq \rho T$ since $\sum_{j \in [M]} \rho_{j,T} \leq M\rho$.

If $\tau_A < T$, then we know that at the stopping time τ_A , the while loop in Algorithm 1 has not yet exited, so we have

$$S_{1,\tau_A} - \gamma M\rho + \beta \underline{\rho}(T - \tau_A) \geq 0 \quad \text{and} \quad S_{2,\tau_A} + M\rho + M\underline{\rho}(T - \tau_A) \leq \rho T \quad (54)$$

Hence,

$$\begin{aligned} \sum_{t \in [T]} g_{1,t} &= \sum_{t \in [\tau_A-1]} g_{1,t} + g_{1,\tau_A} + \sum_{t=\tau_A+1}^T g_{1,t} \\ &\stackrel{(i)}{\geq} \gamma M\rho - \beta \underline{\rho}(T - \tau_A) + g_{1,\tau_A} + \sum_{t=\tau_A+1}^T g_{1,t} \\ &\geq \gamma M\rho - \beta \underline{\rho}(T - \tau_A) - \gamma M\rho + \sum_{t=\tau_A+1}^T g_{1,t} \\ &\stackrel{(ii)}{=} -\beta \underline{\rho}(T - \tau_A) + \sum_{t=\tau_A+1}^T \sum_{j \in [M]} (V_j(\underline{\rho}; \mathbf{z}_{j,t}) - \gamma \underline{\rho}) \\ &\stackrel{(iii)}{\geq} -\beta \underline{\rho}(T - \tau_A) + \sum_{t=\tau_A+1}^T \sum_{j \in [M]} \left(\underline{\rho} \cdot \min_{\mathbf{z}_j \in F_j} \frac{v_{j,1}}{d_{j,1}} - \gamma \underline{\rho} \right) \\ &= -\beta \underline{\rho}(T - \tau_A) + (T - \tau_A)M \left(\underline{\rho} \cdot \min_{\mathbf{z}_j \in F_j} \frac{v_{j,1}}{d_{j,1}} - \gamma \underline{\rho} \right) \\ &\stackrel{(iv)}{\geq} 0, \end{aligned} \quad (55)$$

where (i) follows from $S_{1,\tau_A} = \sum_{t \in [\tau_A-1]} g_{1,t}$ and Eq. (54) that gives $S_{1,\tau_A} \geq \gamma M\rho - \beta \underline{\rho}(T - \tau_A)$; (ii) follows from Algorithm 1 where we set $\rho_{j,t} = \underline{\rho}$ for all $j \in [M]$ and $t = \tau_A + 1 \dots T$, and that $g_{1,t} = \sum_{j \in [M]} (V_j(\rho_{j,t}; \mathbf{z}_{j,t}) - \gamma \rho_{j,t})$; for (iii), assuming the j th channel's realized value cost pairs $\mathbf{z}_{j,t}$ is the element $\mathbf{z}_j \in F_j$, then Proposition C.1 says $V_j(\underline{\rho}; \mathbf{z}_{j,t}) \geq \frac{v_{j,1}}{d_{j,1}} \underline{\rho}$ since $\underline{\rho} = \frac{1}{\log(T)} < \tilde{\rho}$ for large enough T . Hence $V_j(\underline{\rho}; \mathbf{z}_{j,t}) \geq \min_{\mathbf{z}_j \in F_j} \frac{v_{j,1}}{d_{j,1}} \underline{\rho}$; (iv) follows from the fact that $\min_{\mathbf{z}_j \in F_j} \frac{v_{j,1}}{d_{j,1}} > \gamma$ according to Proposition C.1, so $M \min_{\mathbf{z}_j \in F_j} \frac{v_{j,1}}{d_{j,1}} \geq M\gamma + \beta$ since $\beta = \frac{1}{\log(T)} \leq M \min_{\mathbf{z}_j \in F_j} \frac{v_{j,1}}{d_{j,1}} - M\gamma$ for large enough T .

Similarly, we have

$$\begin{aligned} \sum_{t \in [T]} \sum_{j \in [M]} \rho_{j,t} &= \sum_{t \in [\tau_A-1]} \sum_{j \in [M]} \rho_{j,t} + \sum_{j \in [M]} \rho_{j,\tau_A} + \sum_{t=\tau_A+1}^T \sum_{j \in [M]} \rho_{j,t} \\ &\stackrel{(i)}{\leq} \rho T - M\rho - M\underline{\rho}(T - \tau_A) + \sum_{j \in [M]} \rho_{j,\tau_A} + M(T - \tau_A)\underline{\rho} \\ &\leq \rho T - M\rho - M\underline{\rho}(T - \tau_A) + M\rho + M(T - \tau_A)\underline{\rho} \\ &= \rho T. \end{aligned} \quad (56)$$

where (i) follows from $S_{2,\tau_A} = \sum_{t \in [\tau_A-1]} \sum_{j \in [M]} \rho_{j,t}$ and Eq. (54), as well as in Algorithm 1 we set $\rho_{j,t} = \underline{\rho}$ for all $j \in [M]$ and $t = \tau_A, \tau_A + 1 \dots T$. Q.E.D.

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

$$\begin{aligned} \sum_{\tau \in [t]} (\lambda_\tau - \lambda) g_{1,\tau} &\leq \frac{\eta M^2 \bar{V}^2}{2} \cdot t + \frac{1}{2\eta} (\lambda - \lambda_1)^2 \\ \sum_{\tau \in [t]} (\mu_\tau - \mu) g_{2,\tau} &\leq \frac{\eta \rho^2}{2} \cdot t + \frac{1}{2\eta} (\mu - \mu_1)^2, \end{aligned} \quad (57)$$

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

Proof. We will show Eq. (57). Starting with the first inequality w.r.t. λ_τ 's, we have

$$(\lambda_\tau - \lambda) g_{1,\tau} = (\lambda_{\tau+1} - \lambda) g_{1,\tau} + (\lambda_\tau - \lambda_{\tau+1}) g_{1,\tau}. \quad (58)$$

Since $\lambda_{\tau+1} = \Pi_{[0, C_F]} (\lambda_\tau - \eta g_{1,\tau})_+ = \arg \min_{\lambda \in [0, C_F]} (\lambda - (\lambda_\tau - \eta g_{1,\tau}))^2$, we have

$$(\lambda_{\tau+1} - (\lambda_\tau - \eta g_{1,\tau})) \cdot (\lambda - \lambda_{\tau+1}) \geq 0 \quad \forall \lambda \in [0, C_F]. \quad (59)$$

So we have

$$\begin{aligned} (\lambda_{\tau+1} - \lambda) g_{1,\tau} &\leq \frac{1}{\eta} (\lambda_{\tau+1} - \lambda_\tau) \cdot (\lambda - \lambda_{\tau+1}) \\ &= \frac{1}{2\eta} ((\lambda - \lambda_\tau)^2 - (\lambda - \lambda_{\tau+1})^2 - (\lambda_{\tau+1} - \lambda_\tau)^2). \end{aligned} \quad (60)$$

Plugging the above back into Eq. (58), we get

$$\begin{aligned} (\lambda_\tau - \lambda) g_{1,\tau} &\leq (\lambda_\tau - \lambda_{\tau+1}) g_{1,\tau} + \frac{1}{2\eta} ((\lambda - \lambda_\tau)^2 - (\lambda - \lambda_{\tau+1})^2 - (\lambda_{\tau+1} - \lambda_\tau)^2) \\ &\leq \frac{\eta}{2} g_{1,\tau}^2 + \frac{1}{2\eta} ((\lambda - \lambda_\tau)^2 - (\lambda - \lambda_{\tau+1})^2) \\ &\leq \frac{\eta M^2 \bar{V}^2}{2} + \frac{1}{2\eta} ((\lambda - \lambda_\tau)^2 - (\lambda - \lambda_{\tau+1})^2) \end{aligned} \quad (61)$$

The second inequality holds because $(\lambda_\tau - \lambda_{\tau+1}) g_{1,\tau} - \frac{1}{2\eta} (\lambda_{\tau+1} - \lambda_\tau)^2 \leq \frac{\eta}{2} g_{1,\tau}^2$, following the inequality $2ab \leq a^2 + b^2$. The final inequality follows from $g_{1,\tau} = \sum_{j \in M} (V_j(\rho_{j,\tau}) - \gamma \rho_{j,\tau}) \leq M\bar{V}$, where the inequality holds because $V_{j,\tau}(\rho_{j,\tau}) \leq \bar{V}$ for all $j \in [M]$ and $\tau \in [t]$. Summing the above over $\tau = 1 \dots t$ and telescoping we get

$$\sum_{\tau \in [t]} (\lambda_\tau - \lambda) g_{1,\tau} \leq \frac{\eta M^2 \bar{V}^2}{2} \cdot t + \frac{1}{2\eta} (\lambda - \lambda_1)^2 \quad \text{for } \forall \lambda \in [0, C_F].$$

Following the same arguments above we can show

$$\sum_{\tau \in [t]} (\mu_\tau - \mu) g_{2,\tau} \leq \frac{\eta \rho^2}{2} \cdot T + \frac{1}{2\eta} (\mu - \mu_1)^2 \quad \text{for } \forall \mu \in [0, C_F].$$

Q.E.D.

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

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

$$\begin{aligned} \text{CH-OPT}(\mathcal{I}_B) = & \max_{(\gamma_j)_{j \in [M]} \in \mathcal{I}} \sum_{j \in M} V_j(\rho_j) \\ \text{s.t. } & \sum_{j \in M} V_j(\rho_j) \geq \gamma \sum_{j \in M} \rho_j \\ & \sum_{j \in [M]} \rho_j \leq \rho. \end{aligned} \quad (62)$$

Here, we used the definition $V_j(\rho_j) = \mathbb{E}[V_j(\rho_j; \mathbf{z}_j)]$ in Eq. (5), and $D_j(\rho_j; \mathbf{z}_j) = \rho_j$ for any $\mathbf{z}_j$ 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 $\text{CH-OPT}(\mathcal{I}_B)$ 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

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

where (i) follows from the concavity of $V_j(\rho_j)$ and (ii) follows from feasibility 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 $\text{CH-OPT}(\mathcal{I}_B)$ is convex. This concludes the statement of the proposition. Q.E.D.

Appendix D: Proofs for Section 5

D.1. Proof of Lemma 5.2

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

Claim D.1 Recall that $v_{j,n}(1) > \dots > v_{j,n}(L_{j,n}) > 0$ and $d_{j,n}(1) > \dots > d_{j,n}(L_{j,n}) > 0$ for any channel $j \in [M]$ and auction $n \in [m_j]$. If auction n in channel j has increasing marginal values, i.e., for any realization $\mathbf{z}_j = (\mathbf{v}_j, \mathbf{d}_j)$, 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

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

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

$$\begin{aligned} \frac{v_{j,n}(\ell)}{d_{j,n}(\ell)} > \frac{v_{j,n}(\ell+1)}{d_{j,n}(\ell+1)} & \implies \frac{d_{j,n}(\ell+1) - d_{j,n}(\ell)}{d_{j,n}(\ell)} > \frac{v_{j,n}(\ell+1) - v_{j,n}(\ell)}{v_{j,n}(\ell)} \\ & \implies \frac{d_{j,n}(\ell) - d_{j,n}(\ell+1)}{d_{j,n}(\ell)} < \frac{v_{j,n}(\ell) - v_{j,n}(\ell+1)}{v_{j,n}(\ell)} \\ & \implies \frac{v_{j,n}(\ell)}{d_{j,n}(\ell)} < \frac{v_{j,n}(\ell) - v_{j,n}(\ell+1)}{d_{j,n}(\ell) - d_{j,n}(\ell+1)}. \end{aligned} \quad (63)$$

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

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

where (i) follows from Eq. (63), 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 5.2. Similar to the proof of Lemma 4.3, we only need to show for any realization $\mathbf{z}_j = (\mathbf{v}_j, \mathbf{d}_j)_{j \in [M]}$, the conversion function $V_j^+(\rho_j; \mathbf{z}_j) = \mathbf{v}_j^\top \mathbf{x}_j^{*,+}(\rho_j; \mathbf{z}_j)$ where $\mathbf{x}_j^{*,+}(\rho_j; \mathbf{z}_j)$ is defined as Eq. (13) is piecewise linear, continuous, strictly increasing and concave.

For simplicity we use the shorthand notation $\mathbf{x}_j^* = \mathbf{x}_j^{*,+}(\rho_j; \mathbf{z}_j) \in [0, 1]^{\sum_{n \in [m_j]} L_{j,n}}$ as the optimal solution to $V_j^+(\rho_j; \mathbf{z}_j)$, defined in Eq. (13). By re-labeling the auction indices in channel $j \in [M]$ such that $\frac{v_{j,1}(1)}{d_{j,1}(1)} > \frac{v_{j,2}(1)}{d_{j,2}(1)} > \dots > \frac{v_{j,m_j}(1)}{d_{j,m_j}(1)}$, we claim that $\mathbf{x}_j^*$ takes the following form:

$$x_{j,n}^*(\ell) = \begin{cases} 1 & \text{if } \ell = 1 \text{ and } \sum_{n' \in [n]} d_{j,n'}(1) \leq \rho_j \\ \frac{\rho_j - \sum_{n' \in [n-1]} d_{j,n'}(1)}{d_{j,n}(1)} & \text{if } \ell = 1 \text{ and } \sum_{n' \in [n]} d_{j,n'}(1) > \rho_j \\ 0 & \text{otherwise} \end{cases} \quad (64)$$

which is analogous to that of Eq. (25) in the proof of Lemma 4.3. In other words, in the optimal solution, an advertiser would only procure impressions who are in the first position in each auction, and also those with high value-to-cost ratios. With the above representation of $\mathbf{x}_j^*$, the rest of the proof follows exactly from that for Lemma 4.3.

It remains to show that Eq. (64) holds. We proceed by contradiction: suppose that in some auction, an impression other than the first is procured, i.e., $x_{j,n}^*(\ell) > 0$ for some $\ell \in \{2, \dots, L_{j,n}\}$. Under this assumption, we know that by the constraint that at most one impression can be procured in each auction, i.e., $\sum_{\ell \in [L_{j,n}]} x_{j,n}^*(\ell) \leq 1$ (as in Eq. (13)), it must follow that $x_{j,n}^*(1) < 1$.

Furthermore, observe that $x_{j,n}^*(\ell')$ incurs a cost of $d_{j,n}(\ell') \cdot x_{j,n}^*(\ell')$ from the total per-channel budget ρ_j . If we instead allocate this cost to the first impression, we would observe an increase in the total value.

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

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 D.1 holds. This contradicts the optimality of $\mathbf{x}_j^*$, 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. (13) 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 $d_{j,1}(1), \dots, d_{j,m_j}(1)$.

$\square$

Appendix E: Additional Experiment: Convergence under Small Budgets

To study how per-period budget affects convergence, we replicate the experimental setup in Section 6.2 but consider smaller per-period budgets $\rho \in \{0.01, 0.03, 0.05, 0.1\}$. We run $H = 100$ phases, each consisting of $T = 200$ rounds, and update per-channel budgets only at the end of each phase. After each phase, we record the per-period value achieved by our algorithm and compute a 5-phase rolling standard deviation to measure stability over time.

As shown in Figure 6, the algorithm converges more slowly for smaller budgets. This occurs because smaller per-period budget produces smaller reward magnitudes and higher stochastic variance in feedback, which makes the dual updates noisier and increases the number of phases required for stabilization.

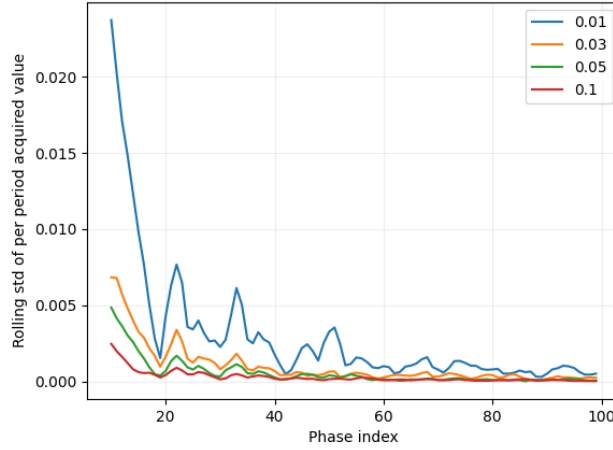


Figure 6 Convergence comparison under small per-period budgets $\rho \in \{0.01, 0.03, 0.05, 0.1\}$. After each of $H = 100$ phases (each with $T = 200$ rounds), we record the per-period value achieved by the proposed algorithm and plot the 5-phase rolling standard deviation across recent phases. A smaller rolling standard deviation indicates faster stabilization and convergence.