# Automated Data-driven Algorithm and Mechanism Design in Online Advertising Markets

by

Jason Cheuk Nam Liang

B.S., Columbia University in the City of New York (2018)

Submitted to the Sloan School of Management
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Operations Research

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2023



Author ........................................................

Sloan School of Management

May 5, 2023

Certified by........................................................

Patrick Jaillet

Dugald C. Jackson Professor, Department of Electrical Engineering and Computer Science

Co-Director, Operations Research Center

Thesis Supervisor

Certified by........................................................

Negin Golrezaei

KDD Career Development Professor in Communications and Technology and Assistant Professor of Operations Management

Thesis Supervisor

Accepted by ........................................................

Georgia Perakis

William F. Pounds Professor of Management Science

Co-Director, Operations Research Center

# Automated Data-driven Algorithm and Mechanism Design in Online Advertising Markets

by

Jason Cheuk Nam Liang

Submitted to the Sloan School of Management
on May 5, 2023, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Operations Research

## Abstract

Modern automated-bidding (or autobidding) ecosystems have fueled the prevalence of programmatic advertising, which accounts for 90% of total digital ad transaction volume and more than \$120 billion dollar ad spend in 2022. In an autobidding ecosystem, online advertisers convey high-level goals to ad platforms using levers presented by the platforms, and subsequently platforms run automated algorithms to procure ads on advertisers' behalf. While autobidding significantly simplifies and scales up ad procurement processes, it also brings about new challenges: for advertisers, the simplification to ad procurement comes at the cost of information dilution as advertisers no longer have access to granular procurement details; for ad platforms, booming advertiser activities have incentivized growing sophistication in advertising campaign and objectives.

My thesis is dedicated to address two themes from a data-centric perspective: how should advertisers effectively interact with ad platforms in limited information environments? And how should ad platforms design procurement mechanisms to achieve revenue and advertiser welfare goals under complex advertiser objectives and behaviors?

The thesis first addresses the advertisers' problem by exploring how advertisers can utilize levers presented by ad platforms such as budgets or target return-on-investments (ROI) to optimize ad procurement objectives. We analyze the effectiveness of standard platform levers, and then present efficient data-driven algorithms for an advertiser to optimize over lever decisions under limited information, where the procurement algorithm and selling mechanisms in the autobidding ecosystem are treated as a blackbox. Then, the thesis explores an ad platform's problem of designing selling mechanisms against advertisers with complex objectives. In particular, this part of the thesis concerns strategic advertisers who may manipulate selling mechanisms via submitting corrupted information to achieve better long-term rewards, as well as constrained advertisers who are subject to financial restrictions. We first design data-driven pricing algorithms to maximize long-term platform revenue in the presence of different advertiser types. Then, we also investigate how ad platforms can augment standard ad auctions with machine-learned advice to improve worst-case welfare guarantees on the individual advertiser level when advertisers are financially constrained and adopt arbitrary strategies.

Thesis Supervisor: Patrick Jaillet
Title: Dugald C. Jackson Professor, Department of Electrical Engineering and Computer Science
Co-Director, Operations Research Center

Thesis Supervisor: Negin Golrezaei
Title: KDD Career Development Professor in Communications and Technology and Assistant Professor of Operations Management

# Acknowledgments

First and foremost, I would like to express my sincere gratitude to my PhD advisors, Patrick Jaillet and Negin Golrezaei, for their invaluable support, mentorship, and encouragement throughout my graduate studies. As I reflect on my journey towards completing this thesis, I am reminded of the pivotal moment when I first met with Patrick and Negin during the PhD admission open house. At that time, I was struggling with anxiety and uncertainty, feeling unsure of what direction to take my research in for the next five years. In spite of my anxiety, Patrick and Negin offered unwavering support and kindness, encouraging me to pursue my passions and explore new areas of research. They also introduced me to the fascinating domain of learning and mechanism design, which sparked my interest and opened up a world of possibilities. As I delved deeper into research throughout the years, Patrick and Negin continued to offer their support and guidance, always challenging me to think beyond academic publications and to focus on making impact in real-world applications. Their mentorship has shaped my research mindset which will undoubtedly benefit me in the future as an independent researcher. In the past five years, Patrick and Negin are not only esteemed advisors, but have also become dear friends and invaluable mentors.

This thesis greatly benefited from several internship experiences through my PhD studies. I am grateful to Vahab Mirrokni and Yuan Deng for kindly hosting me at the Market Algorithms Team in Google Research NYC during the end of 2022, and for introducing me to many fascinating research topics on autobidding ecosystems, which eventually led to many results of this thesis. Throughout our collaborations, I have benefited greatly from Vahab and Yuan's insightful discussions and feedback. They challenged me to think critically beyond my own perspectives and to develop novel solutions to problems of both theoretical and practical significance. Their support has been essential to the completion of this thesis. I also thank Amine Allouah, Okke Schrijvers, and Vahid Liaghat for their mentorship and guidance during my first PhD internship in the Core Data Science Team at Facebook Research, which helped me think more broadly about research in a practical context.

I would also like to express my appreciation to Georgia Perakis and Thodoris Lykouris for serving on my thesis committee, and for offering invaluable advice and feedback during my thesis proposal meetings. I also want to thank Georgia for serving on my general exam committee. Special thanks to Adam Elmachtoub, Garud Iyengar, and Paul Glasserman for advising me on various research projects during my undergraduate studies at Columbia. Without their encouragement and support, I think I would not have pursued a PhD.

I was fortunate to have the opportunity to collaborate (either in academic research or classes), or discuss interesting problems, with many extraordinary researchers throughout my PhD journey: Haihao, Baoyu, Qinyi, Wenyu, Tiancheng, Zhengyuan. Outside of academics, I want to thank my friends Wenyu, Xiaoqi, Buxuan, Zhechao, and Luyao for teaming up with me to participate in all sorts of data science competitions/projects during our free time. These were fun and rewarding experiences that I will always cherish. I also want to thank Davide for being a great friend ever since the first semester in college. Whether it was sharing a meal together, or just having a good chat, his presence has made my journey more enjoyable.

Finally, I am deeply grateful for the unconditioned love from my parents Yongning and Fanny throughout the years. Growing up, my parents always taught me being a good person was more far important than being a smart person. They instilled in me the values of hard work, determination, and compassion, which have been the foundation of my character. I am forever grateful for the role that they have played in shaping me as the person I am today. I also want to express my appreciation for the love, support, and companionship from Clara throughout the past six years. From undergrad to my PhD, Clara has been by my side, contributing to my emotional well-being and making my life full of love and joy. Through the ups and downs during my academic journey, Clara has been a constant source of strength and motivation. Her kindness and positive energy has lifted my spirits and provided me with the courage to pursue my passions and to navigate important life decisions.

Thank you all from the bottom of my heart, and to you, I dedicate this thesis.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 Background on autobidding ecosystems

Over the past few years, programmatic advertising–the automation of purchase and sales of online digital inventories–has become the prevalent mode for digital ad transactions. In 2022, it is estimated that 90% of digital display ads are traded programmatically, resulting in more than \$120 billion dollar ad spend, and these figures are expected to continue their steady increases in following years [2, 1]. Programmatic advertising practices are primarily powered by autobidding ecosystems, in which advertisers only need to convey high-level ad campaign goals to ad platforms, and subsequently ad platforms deploy automated algorithms to procure ads on advertisers' behalf. Although specifics of autobidding ecosystems may vary across different market segments (e.g. display ads versus key word search), the vast majority shares three data-centric components: 1. advertiser-platform interactions; 2. automated proxy procurement; and 3. ad selling mechanism.

**Advertiser-platform interactions.** An advertiser's digital ad campaign starts from the advertiser setting high-level ad procurement goals for their ad campaign using pre-defined levers provided by the platform. These levers include, but are not limited to, procurement objective functions (e.g. optimize quasi-linear utility which is the net monetary gain from conversions after cost, or simply maximize total conversion), total budget, target return-on-investment ROI, maximum cost per conversion, campaign duration, etc.; see e.g. [16, 111] that study setups where advertisers utilize the budget and target ROI levers to express procurement goals. As advertisers generally run sequential campaigns, they make lever decisions based on feedback or conversion outcomes from previous interactions with platforms.

**Automated proxy procurement.** After an advertiser sets their lever decisions, the ad platform deploys automated algorithms to make procurement decisions on behalf of the advertiser w.r.t. ad selling mechanisms (e.g. determining bid values in ad auctions). The automated algorithm may dynamically analyze data on market demand, competitor behavior, and audience demographics to determine the optimal procurement decision. The goal of the deployed algorithms is to meet the specified goals of advertisers conveyed through their lever decisions. Specifically, in the case where selling mechanisms are ad auctions, procurement decisions take the form of submitted bids, and there has been a vast line of research that develops bidding algorithms for different kinds of auctions under a spectrum of advertiser objectives; see e.g. [109, 14, 66, 65, 18].

**Ad selling mechanism.** Independent of the design and deployment of automated procurement algorithms, the ad platform also determines the selling mechanism, which can include auction-based systems [92, 106, 46, 61], fixed-price models [77, 25, 38], or other pricing strategies [41, 42, 75].[1] The selling mechanism determines which ad slots are sold to which advertisers and at what price, with the goal of maximizing revenue for the ad platform and/or optimizing advertiser welfare. In modern ad platforms, the design of selling mechanisms are largely driven by data generated from past advertiser-platform interactions, e.g. using outcomes from past ad auctions to learn optimal auction formats [96, 60, 43], or augmenting standard auction formats with predictions from machine learning models on advertisers' perceived monetary values on ads [35, 36].

These three components constitute key functionalities of data-centric autobidding ecosystems, and their harmonious interplay has been largely beneficial for both advertisers and ad platforms. For advertisers, autobidding has simplified the ad procurement process by removing much of the manual work involved in purchasing ads. Traditionally, advertisers would manually negotiate with ad platforms for ad space and make bids based on their discretionary views on market demand and competitor behavior. However, with autobidding services offered by ad platforms, advertisers only need to convey high-level procurement goals such as budget and ROI targets, and the rest is taken care of by ad platforms that analyze large amounts of data in real-time and make procurement decisions on behalf of advertisers. This makes it

[1] In many modern ad platforms, the design of selling mechanisms is isolated from the design and deployment of automated procurement algorithms to prevent market manipulation and enforce fair prices for ads.

possible for advertisers to participate in millions of digital ad sales per day and allow them to focus on developing effective ad campaigns rather than on the intricacies of the procurement process. On the other hand from the perspective of ad platforms, autobidding has greatly increased online ad transaction volume and revenue: with automated algorithms making procurement decisions on behalf of advertisers, the number of ad sales that can be processed in real-time has increased dramatically. Further, increased advertiser activity due to autobidding generates a large amount of data, with which ad platforms can refine their selling mechanisms to improve revenue and welfare outcomes.

Despite such benefits, data-driven autobidding also brings about many challenges. For advertisers, the simplification of the ad procurement process has led to a lack of transparency not only in procurement behavior, but also in selling mechanics and procurement outcomes; e.g. ad platforms typically do not reveal details of procurement algorithms, making the procurement process a complete black box from advertisers' perspective. Further, advertisers in general do not have access to information on competitors' actions or behavior, nor do they know the procurement outcomes of others. This lack of transparency creates opaque environments for advertisers, making it difficult for them to determine ad campaign lever decisions such as setting budgets or ROI targets. On the other hand, for ad platforms, the vast amount of repeated interactions between automated procurement algorithms and platform selling mechanisms generates abundant data that incentivizes the adoption of even more complex behavior and algorithms to express long and short term goals. This complexity makes it difficult for platforms to design effective selling mechanisms that can achieve revenue and welfare goals.

As autobidding continues to be the dominant mode for advertisers to engage in online adverising markets, finding ways to overcome the aforementioned challenges that advertisers and ad platforms face becomes essential to ensuring the efficiency of modern online advertising systems.

## 1.2 Thesis goal and methodology

This thesis is dedicated to navigate through the aforementioned challenges by presenting theoretical frameworks to study both the advertiser and ad platform's problems, and further present data-driven approaches to help both parties achieve their respective goals in complex autobidding ecosystems. In particular, we present novel data-driven methods to address the following questions.

1. *(For advertisers) Facing opaque ad platform procurement mechanics and procedures, how can advertisers design their ad campaigns by effectively setting lever decisions to achieve their business objectives?*
2. *(For ad platforms) Facing complex advertiser objectives and automated procurement algorithms, how should ad platforms design selling mechanisms to achieve revenue goals and/or optimize overall social welfare among advertisers?*

Answering these two questions is challenging because the three data-centric components in the autobidding ecosystem, namely advertiser-platform interaction, automated proxy procurement, and ad selling mechanism, create a "three-body problem"–the components are tightly interlinked such that any perturbations or changes in one may cause intractable rippling effects on the others. For instance, as an advertiser varies her input levers presented by platforms, the subsequent automated procurement algorithm would adjust accordingly, shifting competition dynamics between the advertiser and competitors, and potentially leading to drastically different outcomes or equilibria of the selling mechanism. Additionally, modifications to the selling mechanism motivates new designs of procurement algorithms, and also prompts advertisers to rethink their lever decisions with which they utilize to achieve certain campaign goals.

In this thesis, our methodology to address both the advertiser's and the ad platform's questions in face of this "three-body problem" is to isolate a single component, and abstract the other two as a whole while retaining key features that interact with the single isolated component. In particular, in Section 1.2.1, we describe our approach to address the advertiser's problem regarding effective decision making for platform levers by isolating the advertiser-platform interaction component, while modelling automated procurement algorithms and the ad mechanism together as a blackbox. In Section 1.2.2, we study ad platforms' revenue maximization problem in the context of the selling mechanism component, while facing automated procurement algorithms that are governed by pre-specified advertiser levers. Finally in Section 1.2.3, we shed light on ad platforms' welfare maximization problem for individual advertisers through the lens of the selling mechanism component, while only considering satisfaction of advertiser levers without specifications of actual automated procurement algorithms.

### 1.2.1 An advertiser's view: multi-channel ad procurement with limited information (Chapter 2)

We address an advertiser's problem to make effective lever decisions under limited information via viewing the automated procurement algorithms and selling mechanism together as a blackbox that inputs lever decisions, and outputs conversion and spend. In particular, we focus on the advertiser's problem in Chapter 2 under a multi-channel ad procurement setup, where 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 ROI and budget constraints across all channels.

As illustrated in Section 1.1, the autobidding ecosystem does not allow advertisers to control over, and thus globally optimize, her procurement decisions in individual ad sales for each channel, as such granular procurement decisions are handled completely by channels: the advertiser can only utilize levers on each channel, such as setting a per-channel budget and per-channel target ROI, to convey high-level procurement goals. In Chapter 2, we specifically focus on two widely used levers, namely per-channel budget and per-channel ROI, and analyze the effectiveness of each of these levers for solving the advertiser's global multi-channel procurement problem. We show that when an advertiser only optimizes over per-channel ROI levers, her total conversion can be arbitrarily worse than what she could have obtained in the global problem which assumes the advertiser can optimize procurement decisions in individual sales. Further, we show that the advertiser can achieve the global optimal conversion when she only optimizes over per-channel budgets. In light of this finding, under a bandit feedback setting that mimics real-world scenarios where advertisers have limited information on ad auctions in channels and how channels procure ads, we present an efficient learning algorithm that produces per-channel budgets whose resulting conversion approximates that of the global optimal problem. Finally, we argue that all our results hold for selling mechanisms that take standard single-item or multi-item auction formats.

### 1.2.2 An ad platform's view I: revenue-maximization against strategic and financially constrained advertisers (Chapters 3 & 4)

We address an ad platform's revenue maximization problem in the context of the selling mechanism component while viewing the advertiser-platform interaction and automated procurement algorithm components as a whole–we perceive advertisers directly running general automated procurement algorithms with the aim to satisfy pre-determined ad procurement goals. In particular, we focus on two advertiser types, those who maximize long term cumulative quasi-linear utility, and those who maximize cumulative conversion subject to financial constraints, respectively.

In Chapter 3, we first study selling to quasi-linear utility maximizing advertisers (or buyers for short) under a dynamic reserve-price optimization setup for repeated contextual second-price auctions. The ad platform (or the seller) has limited information on buyers' overall demand curves which depends on a non-parametric market-noise distribution, whereas buyers run automated bidding algorithms that aim to maximize their long-term time discounted utility. Buyer algorithms may take advantage of the seller's lack of information, and thereby may potentially submit corrupted bids (relative to true valuations) to manipulate the seller's reserve price policies for more favorable reserve prices in the long run. We focus on designing the seller's learning policy to set contextual reserve prices where the seller's goal is to minimize regret compared to the revenue of a benchmark clairvoyant policy that has full information of buyers' demand. We propose a policy with a phased-structure that incorporates randomized "isolation" periods, during which a buyer is randomly chosen to solely participate in the auction. We show that this design allows the seller to control the number of periods in which buyers significantly corrupt their bids. We then prove that our policy enjoys a $T$-period regret of $\widetilde{\mathcal{O}}(\sqrt{T})$ facing strategic buyers.

In Chapter 4, we focus on selling to a financially constrained value-maximizing buyer who is subject to long-term ROI and budget constraints that help ensure efficient utilization of limited monetary resources. In particular, we study from a seller's perspective how to learn and dynamically price ads through repeated posted price mechanisms to maximize revenue, while the constrained buyer runs an data-driven algorithm to learn her optimal strategy to acquire impressions. For this two-sided learning setup, we first show that under full information, the seller's revenue function admits bell-shaped structure when the buyer best responds to prices under budget and ROI constraints. Motivated by this structural property, we propose a seller pricing algorithm that utilizes an episodic binary-search procedure to identify a revenue-optimal selling price. We show that such a simple learning algorithm enjoys low seller regret if within each episode the budget and ROI constrained buyer approximately best responds to the posted price. We present simple yet natural buyer's bidding algorithms under which the buyer approximately best responds while satisfying budget and ROI constraints, leading to a low regret for our proposed seller pricing algorithm.

### 1.2.3 An ad platform's view II: Improving individual advertiser welfare with ML advice (Chapter 5)

In parallel to Section 1.2.2 (and correspondingly Chapters 3 and 4), in this section we investigate an ad platform's goal to optimize advertiser welfare, rather than platform revenue, in the context of the seller mechanism component. We again combine the advertiser-platform interaction and automated procurement components by assuming advertisers directly make procurement decision according to their campaign goals, and consider how an ad platform can improve welfare for individual advertisers in face of worst-case advertiser decision profiles and outcomes under which every advertiser's procurement goals are satisfied.

In Chapter 5, we examine a setup where an ad platform sells ads over parallel auctions, in which multiple advertisers aim to maximize value (e.g. conversion) subject to a return-on-ad spent (ROAS) constraint that limits total spending to the value acquired. The ad platform aims to utilize machine-learned advice (i.e. predictions for advertiser values) to enforce welfare guarantees for individual advertisers while maintaining overall constraint satisfaction. This setup is motivated by the lack of understanding in the literature for advertiser welfare on the individual level: although the literature has studied auction design by incorporating ML advice through various forms to improve total welfare among advertisers, such improvements could come at the cost of individual bidders' welfare and do not shed light on how particular advertiser bidding strategies impact individual fairness.

To address this gap, Chapter 5 demonstrates how ad platforms can utilize ML advice to improve welfare guarantee on both the aggregate and individual bidder level by setting ML advice as personalized reserve prices. Under parallel VCG auctions with such ML advice-based reserve prices, we present a worst-case welfare lower-bound guarantee for an individual advertiser, and show that the lower-bound guarantee is positively correlated with ML advice quality as well the scale of bids induced by an advertiser's bidding strategies. Further, we prove an impossibility result showing that no truthful, and possibly randomized mechanism with anonymous allocations can achieve universally better individual welfare guarantees than VCG, in presence of personalized reserves based on ML-advice of equal quality. Finally, we extend our analysis to generalized first price (GFP) and generalized second price (GSP) auctions.

# Chapter 2

# Multi-channel ad procurement with limited information

*This chapter is based on [34], which is joint work with Yuan Deng, Negin Golrezaei, Patrick Jaillet, and Vahab Mirrokni.*

## 2.1 Introduction

In this chapter, we focus on advertiser-ad platform interactions, and study how advertiser's can effectively communicate with platforms to achieve their ad procurement goals. In today's online advertisers world, advertisers (including but not limited to small businesses, marketing practitioners, non-profits, etc) have been embracing an expanding array of advertising platforms such as search engines, social media platforms, web publisher display etc. which present a plenitude of channels for advertisers to procure ad impressions and obtain traffic. In this growing multi-channel environment, the booming online advertising activities have fueled extensive research and technological advancements in *attribution analytics* to answer questions like which channels are more effective in targeting certain users? Or, which channels produce more user conversion (e.g. ad clicks) or *return-on-investment* (ROI) with the same amount of investments? (see [73] for a comprehensive survey on attribution analytics). Yet, this area of research has largely left out a crucial phase in the workflow of advertisers' creation of a digital ad campaign, namely how advertisers interact with advertising channels, which is the physical starting point of a campaign.

To illustrate the significance of advertiser-channel interactions, consider for example a small business who is relatively well-informed through attribution research that Google Ads and Meta ads are the two most effective channels for its products. The business instantiates its ad campaigns through interacting with the platforms' ad management interfaces (see Figure 2-1), on which the business utilizes levers such as specifying budget and a target ROI[1] to control campaigns. Channels then input these specified parameters into their *autobidding* procedures, where they procure impressions on the advertiser's behalf through automated blackbox algorithms. Eventually, channels report performance metrics such as expenditure and conversion back to the advertiser once the campaign ends. Therefore, one of the most important decisions advertisers need to make involves how to optimize over these levers provided by channels. Unfortunately, this has rarely been addressed in attribution analytics

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

and relevant literature. Hence, this chapters contributes to filling this vacancy by addressing two themes of practical significance:

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

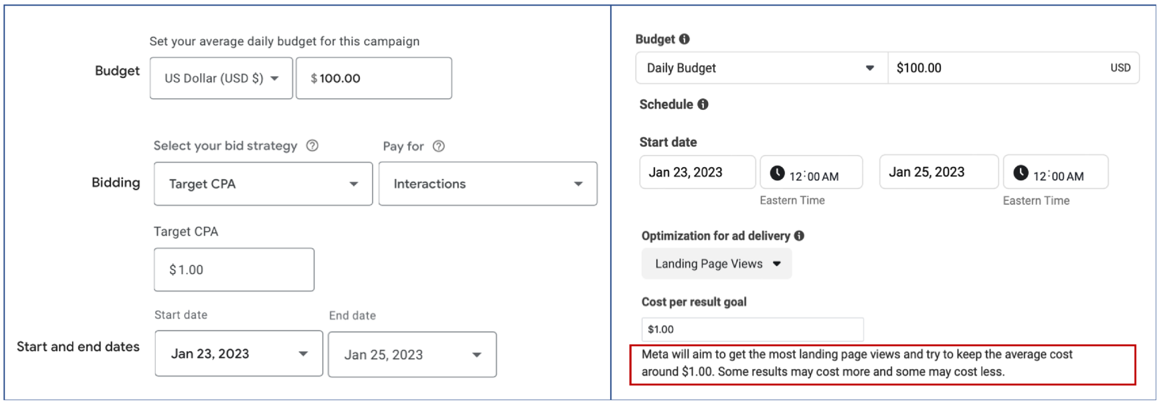


Figure 2-1: Interfaces on Google Ads (left) and Meta Ads Manager (right) for creating advertising campaigns that allow advertisers to set budgets, target ROIs, and campaign duration. CPA, or cost per action on Google Ads, as well as cost per result goal on Meta Ads Manager, is effectively the inverse value for an advertiser's per-channel target ROI.

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

**Modelling ad procurement through per-channel ROI and budget levers.** In Section 2.2 we develop a novel model for online advertisers to optimize over the per-channel ROI and budget levers to maximize total conversion over channels while respecting a global ROI and budget constraint. This multi-channel optimization model closely imitates real-world practices (see Figure 2-1 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.

**Solely optimizing per-channel budgets are sufficient to maximize conversion.** In Theorem 2.3.2 of Section 2.3, we show that solely optimizing for per-channel ROIs is inadequate to optimize conversion across all channels, possibly resulting in arbitrary worse total conversions compared to the hypothetical global optimal where advertisers can optimize over individual auctions. In contrast, in Theorem 2.3.4 and Corollary 2.3.5 we show that solely optimizing for per-channel budgets allows an advertiser to achieve the global optimal.

**Algorithm to optimize per-channel budget levers.** Under a realistic bandit feedback structure where advertisers can only observe the total conversion and spend in each channel after making a per-channel budget decision, in Section 2.4 we develop an algorithm that augments stochastic gradient descent (SGD) with the upper-confidence bound (UCB) algorithm, and eventually outputs within $T$ iterations a per-channel budget profile with which advertisers can achieve $\mathcal{O}(T^{-1/3})$ approximation accuracy in total conversion to that of the optimal per-channel budget profile, and a $\mathcal{O}(T^{-1/2})$ violation in both global budget and ROI constraints. Our algorithm relates to constrained convex optimization with uncertain constraints and bandit feedback under a "one point estimation" regime, and to the best of our knowledge, our proposed algorithm is the first to handle such a setting; see more discussions in Section 2.1 and Remark 2.4.2 of Section 2.4. Finally, we also present an extended version of our algorithm that achieves the same $\mathcal{O}(T^{-1/3})$ conversion accuracy, while respecting both constraints exactly.

**Extensions to general advertiser objectives and mutli-impression auctions.** In Sections 2.5 and 2.6, we shed light on the applicability of our results in Section 2.3 and 2.4 to more general settings when auctions correspond to the sale of multiple auctions, or when advertisers aim to optimize a private cost model instead of conversion.

## Related works

Generally speaking, this chapter focuses on advertisers' impression procurement process or the interactions between advertisers and impression sellers, which has been addressed in a vast amount of literature in mechanism design and online learning; see e.g. [25, 38, 58, 54, 13, 55] to name a few. Here, we review literature that relate to key themes of this chapter, namely autobidding, budget and ROI management, and constrained optimization with bandit feedback.

**Autobidding.** There has been a rich line of research that model the autobidding setup as well as budget and ROI management strategies. The autobidding model has been formally developed in [4], and has been analyzed through the lens of welfare efficiency or price of anarchy in [37, 11, 36, 89], as well as individual advertiser fairness in [33]. The autobidding model has also been compared to classic quasi-linear utility models in [16]. The autobidding model considered in these papers assume advertisers can directly optimize over individual auctions, whereas in this chapter we address a more realistic setting that mimics practice where advertisers can only use levers provided by channels, and let channels procure ads on their behalf.

**Budget and ROI management.** Budget and ROI management strategies have been widely studied in the context of mechanism design and online learning. [14] studies the "system equilibria" of a range of budget management strategies in terms of the platforms' profits and advertisers' utility; [17, 18] study online bidding algorithms (called pacing) that help advertisers achieve high utility in repeated second-price auctions while maintaining a budget constraint, whereas [50] studies similar algorithms but considers respecting a long term ROI constraint in addition to a fixed budget. All of these works on budget and ROI management focus on bidding strategies in a single repeated auction where advertisers' decisions are bid values submitted directly to the auctions. In contrast, this chapter focuses on the setting where advertisers procure ads from multiple auctions through channels, and make decisions on how to adjust the per-channel ROI and budget levers while leaving the bidding to channels' blackbox algorithms.

**Online optimization.** Section 2.4 where we develop an algorithm to optimize over per-channel target ROI and budgets relates to the area of convex constrained optimization with bandit feedback (also referred to as zero-order or gradient-less feedback) since in light of Lemma 2.4.3 in Section 2.4 our problem of interest is also constrained and convex. First, there has been a plenitude of algorithms developed for deterministic constrained convex optimization under a bandit feedback structures where function evaluations for the objective and constraints are non-stochastic. Such algorithms include filter methods [8, 97], barrier-type methods [48, 45], as well as Nelder-Mead type algorithms [26, 9]; see [93] and references therein for a comprehensive survey. In contrast to these works, our optimization algorithm developed in Section 2.4 handles noisy bandit feedback.

Regarding works that also address stochastic settings, [52] presents online optimization algorithms under the *known constraint* regime, which assumes the optimizer can evaluate whether all constraints are satisfied, i.e. constraints are analytically available. Further, the algorithm achieves a $\mathcal{O}(T^{-1/4})$ accuracy. In this chapter, our setting is more complex as the optimizer (i.e. the advertiser) cannot tell whether the ROI constrained is satisfied (due to unknown value and cost distributions in each channels' auctions). Yet our proposed algorithm can still achieve a more superior $\mathcal{O}(T^{-1/3})$ accuracy. Most relevant to this chapter is the very recent works [105, 93], which considers a similar setting to ours that optimizes for a constrained optimization problem where the objective and constraints are only available through noisy function value evaluations (i.e. unkownn constraints). [105] focuses on a special (unknown) linear constraint setting, and [93] extends to general convex constraints. Although [105] and [93] achieve $\mathcal{O}(T^{-1})$ and $\mathcal{O}(T^{-1/2})$ approximation accuracy to the optimal solution which contrasts our $\mathcal{O}(T^{-1/3})$ accuracy, these works imposes several assumptions that are stronger than the ones that we consider. First, the objective and constraint functions are strongly smooth (i.e. the gradients are Lipschitz continuous) and [93] further assume strong convexity. But in this chapter, our objectives and constraints are piece-wise linear and do not satisfy such salient properties. Second, and most importantly, both works consider a setting with "two point estimations" that allows the optimizer to access the objective and constraint function values twice in each iteration, enabling more efficient estimations. this chapter, however, lies in the one-point setting where we can only access function values once per iteration. Finally, we remark that the optimal accuracy/oracle complexity for the one-point setting for constrained (non-smooth) convex optimization with bandit feedback and unknown constraints remains an open question; see Remark 2.4.2 in Section 2.4 for more details. We refer readers to Table 4.1 in [81] for a survey on best known bounds under different one-point bandit feedback settings.

## 2.2 Preliminaries

**Advertisers' global optimization problem.** Consider an advertiser running a digital ad campaign to procure ad impressions on $M \in \mathbb{N}$ platforms such as Google Ads, Meta Ads Manager etc., each of which we call a *channel*. Each channel $j$ consists of $m_j \in \mathbb{N}$ parallel ad auctions, each of which corresponds to the sale of an ad impression.[2] An ad auction $n \in [m_j]$ is associated with a value $v_{j,n} \geq 0$ that represents the expected conversion (e.g. number of clicks) of the impression on sale, and a cost $d_{j,n} \geq 0$ that is required for the purchase of the impression. For example, the cost

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

in a single slot second-price auction is the highest competing bid of competitors in the market, and in a posted price auction the cost is simply the posted price by the seller of the impression. Writing $\boldsymbol{v}_j = (v_{j,n})_{n\in[m_j]}$ and $\boldsymbol{d}_j = (d_{j,n})_{n\in[m_j]}$, we assume that $\boldsymbol{z}_j := (\boldsymbol{v}_j, \boldsymbol{d}_j)$ is sampled from some discrete distribution $\boldsymbol{p}_j$ supported on some finite set $F_j \subseteq \mathbb{R}_+^{m_j} \times \mathbb{R}_+^{m_j}$.

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

$$\begin{aligned} \text{GL-OPT} \;=\; \max_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_M} \quad & \sum_{j\in[M]} \mathbb{E}\left[\boldsymbol{v}_j^\top \boldsymbol{x}_j\right] \\ \text{s.t.} \quad & \sum_{j\in[M]} \mathbb{E}\left[\boldsymbol{v}_j^\top \boldsymbol{x}_j\right] \;\geq\; \gamma \sum_{j\in[M]} \mathbb{E}\left[\boldsymbol{d}_j^\top \boldsymbol{x}_j\right] \\ & \sum_{j\in[M]} \mathbb{E}\left[\boldsymbol{d}_j^\top \boldsymbol{x}_j\right] \;\leq\; \rho \\ & \boldsymbol{x}_j \in [0,1]^{m_j} \quad j \in [M]\,. \end{aligned} \tag{2.1}$$

Here, the decision variable $\boldsymbol{x}_j \in [0,1]^{m_j}$ is a vector where $x_{j,n}$ denotes whether impression in auction $n$ for channel $j$ is procured. We remark that $\boldsymbol{x}$ depends on the realization of $\boldsymbol{z} = (\boldsymbol{v}_j, \boldsymbol{d}_j)_{j\in[M]}$ and is also random. We note that the ROI and budget constraints are taken in expectation because an advertiser procures impressions from a very large number of auctions (since the number of auctions in each platform is typically very large) and thus the advertiser only demands to satisfy constraints in an average sense. We note that GL-OPT is a widely adopted formulation for autobidding practices in modern online advertising, which represents advertisers' conversion maximizing behavior while respecting certain financial targets for ROIs and budgets; see e.g. [4, 11, 37, 36]. In Section 2.6 we discuss more general advertiser objectives, e.g. maximizing quasi-linear utility.

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

**Advertisers' levers to solve their global problems.** To solve the global optimization problem GL-OPT, ideally advertisers would like to optimize over individual auctions across all channels. However, in reality channels operate as independent entities, and typically do not provide means for general advertisers to participate in specific individual auctions at their discretion. Instead, channels provide advertisers with specific *levers* to express their ad campaign goals on spend and conversion. In this work, we focus on two of the most widely used levers, namely the per-channel ROI target and per-channel budget (see illustration in Fig. 2-1). After an advertiser inputs these parameters to a channel, the channel then procures on behalf of the advertiser through autonomous programs (we call this programmatic process *autobidding*) to help advertiser achieve procurement results that match with the inputs. We will elaborate on this process later.

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

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

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

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

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

(2.2)

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

$$\begin{aligned} \text{CH-OPT}(\mathcal{I}) = \max_{(\gamma_j,\rho_j)_{j\in[M]}\in\mathcal{I}} \quad & \sum_{j\in M} \mathbb{E}\left[V_j(\gamma_j, \rho_j; \boldsymbol{z}_j)\right] \\ \text{s.t.} \quad & \sum_{j\in M} \mathbb{E}\left[V_j(\gamma_j, \rho_j; \boldsymbol{z}_j)\right] \geq \gamma \sum_{j\in M} \mathbb{E}\left[D_j(\gamma_j, \rho_j; \boldsymbol{z}_j)\right] \\ & \sum_{j\in[M]} \mathbb{E}\left[D_j(\gamma_j, \rho_j; \boldsymbol{z}_j)\right] \leq \rho\,, \end{aligned} \tag{2.3}$$

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

The functions $(V_j, D_j)$ that map per-channel target ROI and budgets $\gamma_j, \rho_j$ to the total conversion and expenditure are specified by various factors including but not limited to channel $j$'s autobidding algorithms deployed to procure ads on advertisers' behalf as well as the auctions mechanisms that sell impressions. In this work, we study a general setup that closely mimics industry practices. We assume that on the behalf of the advertiser, each channel aims to optimize their conversion over all $m_j$ auctions while respecting the advertiser's input (i.e., per-channel target ROI and budgets); see e.g. Meta Ads Manager in Figure 2-1 specifically highlights the channel's autobidding procurement methodology provides evidence to support the aforementioned setup.

Hence, each channel $j$'s optimization problem can be written as

$$\boldsymbol{x}_j^*(\gamma_j, \rho_j; \boldsymbol{z}_j) \;=\; \arg\max_{\boldsymbol{x}\in[0,1]^{m_j}} \;\; \boldsymbol{v}_j^\top \boldsymbol{x} \quad \text{s.t.} \quad \boldsymbol{v}_j^\top \boldsymbol{x} \geq \gamma_j \boldsymbol{d}_j^\top \boldsymbol{x}, \quad \boldsymbol{d}_j^\top \boldsymbol{x} \leq \rho_j \,, \tag{2.4}$$

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

$$\begin{aligned} V_j(\gamma_j, \rho_j; \boldsymbol{z}_j) &= \boldsymbol{v}_j^\top \boldsymbol{x}_j^*(\gamma_j, \rho_j; \boldsymbol{z}_j) \quad \text{and} \quad V_j(\gamma_j, \rho_j) = \mathbb{E}[V_j(\gamma_j, \rho_j; \boldsymbol{z}_j)] \\ D_j(\gamma_j, \rho_j; \boldsymbol{z}_j) &= \boldsymbol{d}_j^\top \boldsymbol{x}_j^*(\gamma_j, \rho_j; \boldsymbol{z}_j) \quad \text{and} \quad D_j(\gamma_j, \rho_j) = \mathbb{E}[D_j(\gamma_j, \rho_j; \boldsymbol{z}_j)]\,. \end{aligned} \tag{2.5}$$

Here, the expectation is taken w.r.t. randomness in $\boldsymbol{z}_j = (\boldsymbol{v}_j, \boldsymbol{d}_j) \in \mathbb{R}_+^{m_j} \times \mathbb{R}_+^{m_j}$. We assume that for any $(\gamma_j, \rho_j)$ and realization $\boldsymbol{z}_j$, $V_j(\gamma_j, \rho_j; \boldsymbol{z}_j)$ is bounded above by some absolute constant $\bar{V} \in (0, \infty)$ almost surely. We remark that Eq.(2.5) assumes channels are able to achieve optimal procurement performance. Later in Section 2.6, we will briefly discuss setups where channels does not optimally solve for Eq.(2.4).

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

1. How effective are the per-channel ROI and budget levers to help advertisers achieve the globally optimal conversion GL-OPT while respecting the global ROI and budget constraints? In particular, for each of the advertiser options $\mathcal{I} \in \{\mathcal{I}_B, \mathcal{I}_R, \mathcal{I}_G\}$ defined in Eq. (2.2), what is the discrepancy between CH-OPT($\mathcal{I}$), i.e. the optimal conversion an advertiser can achieve in practice, versus the optimal GL-OPT?

2. Since in reality advertisers can only utilize the two per-channel levers offered by channels, how can advertisers optimize per-channel target ROIs and budgets to solve for CH-OPT($\mathcal{I}$)?

In Section 2.3, we address the first question to determine the gap between CH-OPT($\mathcal{I}$) and GL-OPT for different advertiser options. In Section 2.4, we develop an efficient algorithm to solve for per-channel levers that optimize CH-OPT($\mathcal{I}$).

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

In this section, we examine the effectiveness of the per-channel target ROI and per-channel budget levers in achieving the global optimal GL-OPT. In particular, we study if the optimal solution to the channel problem CH-OPT($\mathcal{I}$) defined in Eq. (2.3) for $\mathcal{I} \in \{\mathcal{I}_B, \mathcal{I}_R, \mathcal{I}_G\}$ is equal to the global optimal GL-OPT. As a summary of our results, we show that the per-channel budget only option, and the general option achieves GL-OPT, but the per-channel ROI only option can yield conversion arbitrarily worse than GL-OPT for certain instance, even when there is no global budget constraint (i.e., $\rho = \infty$). This implies that the per-channel ROI lever is inadequate to help advertisers achieve the globally optimal conversion, whereas the per-channel budget lever is effective to attain optimal conversion even when the advertiser solely uses this lever.

Our first result in this section is the following Lemma 2.3.1 which shows that GL-OPT serves as a theoretical upper bound for an advertiser's conversion through optimizing CH-OPT($\mathcal{I}$) with any option $\mathcal{I}$.

**Lemma 2.3.1** (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.* (2.2)*, we have* GL-OPT $\geq$ CH-OPT($\mathcal{I}$)*, where we recall the definitions of* GL-OPT *and* CH-OPT *in Eq.* (2.1) *and* (2.3)*, respectively.*

The proof of Lemma 2.3.1 is deferred to Appendix A.1.1. In light of the theoretical upper bound GL-OPT, we are now interested in the gap between GL-OPT and CH-OPT($\mathcal{I}$) for option $\mathcal{I} \in \{\mathcal{I}_B, \mathcal{I}_R, \mathcal{I}_G\}$. In the following Theorem 2.3.2, we show that there exists a problem instance under which the ratio $\frac{\text{CH-OPT}(\mathcal{I}_R)}{\text{GL-OPT}}$ nears 0, implying the per-channel ROIs alone fail to help advertisers optimize conversion.

**Theorem 2.3.2** (Per-channel ROI only option fails to optimize conversion)**.** *Consider an advertiser with a (global) target ROI of $\gamma = 1$ procuring impressions from $M = 2$ channels, where channel 1 consists of a single auction and channel 2 consists of two auctions. The advertiser has unlimited budget $\rho = \infty$, and chooses the per-channel target ROI only option $\mathcal{I}_R$ defined in Eq.* (2.2)*. Assume there is only one realization of value-cost pairs $\boldsymbol{z} = (\boldsymbol{v}_j, \boldsymbol{d}_j)_{j\in[M]}$ (i.e. the support $F = F_1 \times F_2$ is a singleton), and the realization is presented in the following table, where $X > 0$ is some arbitrary parameter. Then, for this problem instance we have $\lim_{X\to\infty} \frac{\text{CH-OPT}(\mathcal{I}_R)}{\text{GL-OPT}} = 0$.*

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

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

**Definition 2.3.1** (Strictly quasi-concave value-cost landscape)**.** *We say that a value-cost pair vector $\boldsymbol{z} = \{(v_1, d_1) \ldots (v_n, d_n)\}$ has a strictly quasi-concave landscape if for* *any subset $\mathcal{S} \subset [n]$, and $n' \notin \mathcal{S}$, we have*

$$\frac{v_{n'} + \sum_{k\in\mathcal{S}} v_k}{d_{n'} + \sum_{k\in\mathcal{S}} d_k} < \max\left\{\frac{v_{n'}}{d_{n'}}, \frac{\sum_{k\in\mathcal{S}} v_k}{\sum_{k\in\mathcal{S}} d_k}\right\}$$

**Lemma 2.3.3** (Conditions when per ROI-only option is sufficient)**.** *We have* GL-OPT $=$ CH-OPT$(\mathcal{I}_R)$ *if and only if for each channel $j$, any realization of $\boldsymbol{z}_j \in F_j$ has a strictly quasi-concave landscape (see Definition 2.3.1).*

In fact, we show that the optimal per-channel ROIs under strictly quasi-concave value-cost landscapes is the realized expected ROI in that channel for auctions selected by GL-OPT. The intuition is that for GL-OPT, we rank individual auctions according to their value-to-cost ratios. And for each channel, winning any other auction (apart from those selected by GL-OPT) will bring down the per-channel realized ROI under the strictly quasi-concave value-cost landscape condition, and hence violating the per-channel ROI constraint.

In contrast to the per-channel ROI only option, the budget only option in fact allows an advertiser's conversion to reach the theoretical upper bound GL-OPT through solely optimizing for per-channel budgets. This is formalized in the following theorem whose proof we present in Appendix A.1.2.

**Theorem 2.3.4** (Per-channel budget only option suffices to achieve optimal conversion)**.** *For the budget only option $\mathcal{I}_B$ defined in Eq.*(2.2)*, 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 2.3.4, the following Corollary 2.3.5 states per-channel ROIs in fact become redundant once advertisers optimize for per-channel budgets.

**Corollary 2.3.5** (Redundancy of per-channel ROIs)**.** *For the general option $\mathcal{I}_G$ defined in Eq.*(2.2) *where an advertiser sets both per-channel ROI and budgets, we have* GL-OPT $=$ CH-OPT$(\mathcal{I}_G)$ *for any aggregate ROI $\gamma > 0$ and total budget $\rho > 0$, even for $\rho = \infty$. Further, there exists and optimal solution $(\gamma_j, \rho_j)_{j\in[M]}$ to* CH-OPT$(\mathcal{I}_G)$*, s.t. $\gamma_j = 0$ for all $j \in [M]$.*

In light of the redundancy of per-channel ROIs as illustrated in Corollary 2.3.5, in the rest of the paper we will fix $\gamma_j = 0$ for any channel $j \in [M]$, and omit $\gamma_j$ in all relevant notations; e.g. we will write $D_j(\rho_j; \boldsymbol{z}_j)$ and $D_j(\rho_j)$, instead of $D_j(\gamma_j, \rho_j; \boldsymbol{z}_j)$ and $D_j(\gamma_j, \rho_j)$. Equivalently, we will only consider the per-channel budget only option $\mathcal{I}_B$.

## 2.4 Optimization algorithm for per-channel budgets under bandit feedback

In this section, we develop an efficient algorithm to solve for per-channel budgets that optimize CH-OPT($\mathcal{I}_B$) defined in Eq. (2.3), which achieves the theoretical optimal conversion, namely GL-OPT, as illustrated in Theorem 2.3.4. In particular, we consider algorithms that run over $T > 0$ periods, where each period for example corresponds to the duration of 1 hour or 1 day. At the end of $T$ periods, the algorithm produces some per-channel budget profile $(\rho_j)_{j\in[M]} \in [0, \rho]^M$ that approximates CH-OPT($\mathcal{I}_B$), and satisfies aggregate budget and ROI constraints, namely

$$\text{ROI:} \quad \sum_{j\in M} V_j(\rho_j) \geq \gamma \sum_{j\in M} D_j(\rho_j), \quad \text{and} \quad \text{Budget:} \quad \sum_{j\in[M]} D_j(\rho_j) \leq \rho\,, \tag{2.6}$$

where we recall the expected conversion and spend functions $(V_j(\rho_j), D_j(\rho_j))$ defined in Eq. (2.5).

The algorithm proceeds as follows: at the beginning of period $t \in [T]$, the advertiser sets per-channel budgets $(\rho_{j,t})_{j\in[M]}$, while simultaneously values and costs $\boldsymbol{z}_t = (\boldsymbol{v}_{j,t}, \boldsymbol{d}_{j,t}) \in \mathbb{R}_+^{M_j} \times \mathbb{R}_+^{M_j}$ are sampled (independently in each period) from finite support $F = F_1 \times \dots F_M$ according to discrete distributions $(\boldsymbol{p}_j)_{j\in[M]}$. Each channel $j$ then takes as input $\rho_{j,t} \in [0, \rho]$ and procures ads on behalf of the advertiser, and reports the total realized conversion $V_j(\rho_{j,t}; \boldsymbol{z}_t)$ as well as total spend $D_j(\rho_{j,t}; \boldsymbol{z}_t)$ to the advertiser, where $V_j(\rho_{j,t}; \boldsymbol{z}_t)$ and $D_j(\rho_{j,t}; \boldsymbol{z}_t)$ are defined in Eq. (2.5). For simplicity we also assume for any realization $\boldsymbol{z} = (\boldsymbol{v}_j, \boldsymbol{d}_j) \in F$ we have the ordering $\frac{v_{j,1}}{d_{j,1}} > \frac{v_{j,2}}{d_{j,2}} > \dots > \frac{v_{j,m_j}}{d_{j,m_j}}$ for all channels $j \in [M]$.

Here, we highlight that the advertiser receives *bandit feedback* from channels, i.e. the advertiser only observes the numerical values $V_j(\rho_{j,t}; \boldsymbol{z}_t)$ and $D_j(\rho_{j,t}; \boldsymbol{z}_t)$, but does not get to observe $V_j(\rho'_j; \boldsymbol{z}')$ and $D_j(\rho'_j; \boldsymbol{z}')$ evaluated at any other per-channel budget $\rho'_j \neq \rho_{j,t}$ and realized value-cost pairs $\boldsymbol{z}' \neq \boldsymbol{z}_t$. More discussions on challenges that arise from this bandit feedback structure can be found in Section 2.4.1.

We also make the following Assumption 2.4.1 that states that if the advertiser allocates any feasible per-channel budget to a channel $j \in [M]$, the channel will almost surely deplete the entire budget in the impression procurement process. This is a natural assumption that mimics practical scenarios, e.g. small businesses who have moderate-sized budgets.

**Assumption 2.4.1** (Moderate budgets)**.** *We assume the total budget is finite, i.e.* $\rho < \infty$, *and for any channel* $j \in [M]$, *value-cost realization* $\boldsymbol{z} = (\boldsymbol{v}, \boldsymbol{d}) \in F_j$, *and per-channel budget* $\rho_j \in [0, \rho]$, *the optimal solution* $\boldsymbol{x}^*_j(\rho; \boldsymbol{z})$ *defined in Eq.* (2.4) *is budget binding, i.e.* $D_j(\rho_j; \boldsymbol{z}) = \boldsymbol{d}_j^\top \boldsymbol{x}^*_j(\rho_j; \boldsymbol{z}) = \rho_j$.

### 2.4.1 The SGD-UCB algorithm to optimize per-channel budgets

Here, we describe our algorithm to solve for optimal per-channel budgets w.r.t. CH-OPT$(\mathcal{I}_B)$. Similar to most algorithms for constrained optimization, we take a dual stochastic gradient descent (SGD) approach; see a comprehensive survey on dual descent methods in [21]. First, we consider the Lagrangian functions w.r.t. CH-OPT$(\mathcal{I}_B)$ where we let $\boldsymbol{c} = (\lambda, \mu) \in \mathbb{R}^2_+$ be the dual variables corresponding to the ROI and budget constraints, respectively:

$$\mathcal{L}_j(\rho_j, \boldsymbol{c}; \boldsymbol{z}_j) = (1+\lambda)V_j(\rho_j; \boldsymbol{z}_j) - (\lambda\gamma + \mu)\rho_j \quad \text{and} \quad \mathcal{L}_j(\rho_j, \boldsymbol{c}) = \mathbb{E}\left[\mathcal{L}_j(\rho_j, \boldsymbol{c}; \boldsymbol{z}_j)\right] . \tag{2.7}$$

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

$$\rho_{j,t} \;=\; \arg\max_{\rho_j\in[0,\rho]} \mathcal{L}_j(\rho_j, \boldsymbol{c}_t; \boldsymbol{z}_t)\,. \tag{2.8}$$

Having observed the realized values $(V_j(\rho_{j,t}; \boldsymbol{z}_t))_{j\in[M]}$ (note that spend is $(\rho_{j,t})_{j\in[M]}$ in light of Assumption 2.4.1), we calculate the current period violation in budget and ROI constraints, namely $g_{1,t} := \sum_{j\in M}(V_j(\rho_{j,t}; \boldsymbol{z}_t) - \gamma\rho_{j,t})$ and $g_{2,t} = \rho - \sum_{j\in[M]}\rho_{j,t}$. Next, we update dual variables via $\lambda_{t+1} = (\lambda_t - \eta g_{1,t})_+$ and $\mu_{t+1} = (\mu_t - \eta g_{2,t})_+$, where $\eta$ is some pre-specified step size.[3]

However, the above SGD approach faces a fatal drawback, namely we cannot realistically find the primal decisions by solving Eq. (2.8) since the function $\mathcal{L}_j(\cdot, \boldsymbol{c}_t; \boldsymbol{z}_t)$ is unknown due to the bandit feedback structure. Therefore, we provide a modification to SGD to handle this issue. Before we present our approach, we briefly note that although bandit feedback prevents the naive application of SGD for our problem of interest, this may not be the case in other online advertising scenarios that involve relevant learning tasks, underlining the challenges of our problem; see following Remark 2.4.1 for details.

**Remark 2.4.1.** *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. [14, 18, 50]. To illustrate, consider for instance [14] 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 SGD takes the special form $\mathcal{L}_j(b, \mu_t; \boldsymbol{z}_t) = (v_t - (1+\mu_t)d_t)\,\mathbb{I}\{b_t \geq d_t\}$ where $\mu_t$ is the dual variable w.r.t. the budget constraint. This simply allows an advertiser to optimize for her primal decision by bidding $\arg\max_{b\geq 0}\mathcal{L}_j(b, \boldsymbol{c}_t; \boldsymbol{z}_t) = \frac{v_t}{1+\mu_t}$. So even though [14, 18, 50] study dual SGD*

[3]Here, the dual updates follow the vanilla gradient descent approach, and one can also employ more general mirror descent updates; see e.g. [18].

*under bandit feedback, the special structures of their problem instances permits SGD to effectively optimize for primal decisions in each period, as opposed to Eq.* (2.8) *in our setting which can not be solved.*

To resolve challenges that arise with bandit feedback in our model, we take a natural approach to augment SGD with the *upper-confidence bound (UCB)* algorithm, which is well celebrated for solving learning problems under bandit feedback such as multi-arm bandits; see an introduction to bandits in [102]. In particular, we first discretize our per-channel buudget decision set $[0, \rho]$ into granular "arms" that are separated by some distance $\delta > 0$, so that the discretized per-channel budget decisions become

$$\mathcal{A}(\delta) = \{a_k\}_{k \in [K]} \text{ where } a_k = (k-1)\delta \text{ and } K := \lceil \rho/\delta \rceil + 1\,. \tag{2.9}$$

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

We remark that there has been very recent works that combine SGD with Thompson sampling which is another well-known algortithm for solving bandit problems (e.g. [40] and references therein), and works that employ SGD in bandit problems (e.g. [67]). Yet to the best of our knowledge, approach to augment SGD with UCB is novel.

### 2.4.2 Analyzing the SGD-UCB algorithm

In this subsection, we analyze the performance of SGD-UCB in Algorithm 1, and present accuracy guarantees on the final output $\overline{\boldsymbol{\rho}}_T = \left(\frac{1}{T}\sum_{t \in [T]} \rho_{j,t}\right)_{j \in [M]}$ of the algorithm. The backbone of our analysis strategy is to show the cumulative conversion loss over $T$ periods, namely $T \cdot \text{GL-OPT} - \mathbb{E}\left[\sum_{t \in [T]} \sum_{j \in [M]} V_j(\rho_{j,t})\right]$ consists of two

**Algorithm 1** SGD-UCB

**Input:** Budget discretization decision set $\mathcal{A}(\delta)$ defined in Eq.(2.9). Step size $\eta > 0$. Initialize $N_{j,1}(a_k) = \hat{V}_{j,1}(a_k) = 0$ for all $j \in [M]$ and $k \in [K]$, and dual variables $\lambda_1 = \mu_1 = 0$.

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

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

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

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

$$\rho_{j,t} = \arg\max_{a_k \in \mathcal{A}(\delta)} \hat{V}_{j,t}(a_k) + \texttt{UCB}_{j,t}(a_k) - \frac{\lambda_t \gamma + \mu_t}{1 + \lambda_t} a_k, \quad \text{where } \texttt{UCB}_{j,t}(a_k) = \sqrt{\frac{2\log(T)}{N_{j,t}(a_k)}} \tag{2.10}$$

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

$$\begin{aligned} &N_{j,t+1}(a_k) = N_{j,t}(a_k) + \mathbb{I}\{\rho_{j,t} = a_k\} \\ &\hat{V}_{j,t+1}(a_k) = \frac{1}{N_{j,t+1}(a_k)} \Big( N_{j,t}(a_k)\hat{V}_{j,t}(a_k) + V_j(\rho_{j,t}; \boldsymbol{z}_t)\mathbb{I}\{\rho_{j,t} = a_k\} \Big) \quad , \text{ for } j = 1 \dots M \end{aligned} \tag{2.11}$$

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

$$\lambda_{t+1} = \left( \lambda_t - \eta g_{1,t} \right)_+ \text{ and } \mu_{t+1} = \left( \mu_t - \eta g_{2,t} \right)_+ \tag{2.12}$$

6: **end for**

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

main parts, namely the error induced by the UCB in our algorithm, and the error due to SGD (or what is typically viewed as the deviations from complementary slackness), as shows in the following Proposition 2.4.1. Then we further bound each part, respectively.

**Proposition 2.4.1.** *For any channel $j \in [M]$ define $\rho_j^*(t) = \arg\max_{\rho_{j \in [M]}} \mathcal{L}_j(\rho_j; \boldsymbol{c}_t)$ to be the optimal per-channel budget w.r.t. dual variables $\boldsymbol{c}_t = (\lambda_t, \mu_t)_{t \in [T]}$ during* *period $t \in [T]$. Then we have*

$$\begin{aligned} & T \cdot \text{GL-OPT} - \sum_{t\in[T]} \sum_{j\in[M]} \mathbb{E}\left[V_j(\rho_{j,t})\right] \\ \leq \;& M\bar{V}K + \sum_{j\in[M]} \underbrace{\sum_{t>K} \mathbb{E}\left[\mathcal{L}_j(\rho_j^*(t), \lambda_t, \mu_t) - \mathcal{L}_j(\rho_{j,t}, \lambda_t, \mu_t)\right]}_{\textit{UCB error}} \\ & + \underbrace{\sum_{t>K} \mathbb{E}\left[\lambda_t g_{1,t} + \mu_t g_{2,t}\right]}_{\textit{SGD complementary slackness deviations}}, \end{aligned} \tag{2.13}$$

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

The bound on SGD complementary slackness violation is presented in the following Lemma 2.4.2, and follows a standard analyses for SGD; we refer readers to the proof in Appendix A.2.2.

**Lemma 2.4.2** (Bounding complementary slackness deviations)**.** *Recall* $g_{1,t} = \sum_{j\in[M]} \left(V_j(\rho_{j,t}; \boldsymbol{z}_t) - \gamma\rho_{j,t}\right)$, $g_{2,t} = \rho - \sum_{j\in[M]} \rho_{j,t}$ *and* $\eta > 0$ *the step size defined in Algorithm 1. Then we have*

$$\sum_{t>K} \mathbb{E}\left[\lambda_t g_{1,t} + \mu_t g_{2,t}\right] \;\leq\; \mathcal{O}\left(\eta T + \frac{1}{\eta}\right). \tag{2.14}$$

**Challenges in bounding UCB error due to adversarial contexts and continuum-arm dicretization.** Bounding our UCB error is much more challenging than doing so in classic stochastic multi-arm bandit settings: first, our setup involves discretizing a continuum of arms i.e. our discretization in Eq.(2.9) for $[0, \rho]$; second, and more importantly, the dual variables $\{\boldsymbol{c}_t\}_{t\in[T]}$ are effectively *adversarial contexts* since they are updated via SGD instead of being stochastically sampled from some nice distribution, and correspondingly the Lagrangian function $\mathcal{L}_j(a_k, \boldsymbol{c}_t; \boldsymbol{z}_t)$ can be viewed as a reward function that maps any arm-context pair $(a_k, \boldsymbol{c}_t)$ to (stochastic) payoffs. Moreover, the UCB error in Eq.(2.13) concerns a dynamic benchmark $\mathcal{L}_j(\rho_j^*(t), \boldsymbol{c}_t; \boldsymbol{z}_t)$

instead of a single-arm benchmark across all contexts, namely $\max_{a\in\mathcal{A}(\delta)} \mathcal{L}_j(a, \boldsymbol{c}_t; \boldsymbol{z}_t)$ . Both continuum-arms and dynamic benchmarks under adversarial contexts have been notorious in making reward function estimations highly inefficient; see e.g. discussions in [5, 3]. We further elaborate on additional challenges that adversarial contexts bring about:

- **Boundedness of rewards.** In classic stochastic multi-arm bandtis and UCB, losses in total rewards grow linearly with the magnitude of rewards. In our setting, the reward function, i.e. the Lagrangian function $\mathcal{L}_j(a_k, \boldsymbol{c}_t; \boldsymbol{z}_t)$, scales linearly with the magnitude of contexts (see Eq. (2.7), so large contexts (i.e. large dual variables) may lead to large losses.

- **Context-dependent exploration-exploitation tradeoffs.** The typical trade-off for arm exploration and exploitation in our setting depends on the particular values of the contexts (i.e. the dual variables), which means there may exist "bad" contexts that lead to poor tradeoffs that require significantly more explorations to achieve accurate estimates of arm rewards than other "good" contexts. We elaborate more in Lemma 2.4.5 and discussions thereof.

In the following, we first handle continuum arm discretization and dynamic contextual benchmarks via analyzing structural properties of the reward (i.e. Lagrangian) functions. Fortunately, the specific form of conversion functions $V(\rho_j; \boldsymbol{z})$ defined in Eq. (2.4) imposes a salient structure on the Lagrangian for pulling an arm. Specifically, the following lemma shows that the Lagrangian is continuous, piece-wise linear, concave, and unimodal[4]; we present the proof in Appendix A.2.3

**Lemma 2.4.3** (Structural properties of conversion and Lagrangian functions)**.** •

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

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

*particular, $V_j(\rho_j)$ takes the form*

$$V_j(\rho_j) = \sum_{n\in[S_j]} \left(s_{j,n}\rho_j + b_{j,n}\right) \mathbb{I}\{r_{j,n-1} \le \rho_j \le r_{j,n}\}\,, \tag{2.15}$$

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

- *For any dual variables $\boldsymbol{c} = (\lambda, \mu) \in \mathbb{R}^2_+$, the Lagrangian function $\mathcal{L}_j(\rho_j, \boldsymbol{c})$ defined in Eq. (2.7) is continuous, piece-wise linear, concave, and unimodal in $\rho_j$. In particular,*

$$\mathcal{L}_j(\rho_j, \boldsymbol{c}) = \sum_{n\in[S_j]} \left(\sigma_{j,n}(\boldsymbol{c})\rho_j + (1+\lambda)b_{j,n}\right) \mathbb{I}\{r_{j,n-1} \le \rho_j \le r_{j,n}\}\,, \tag{2.16}$$

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

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

We now handle the reward boundedness issue in the Lagrangian functions defined in Eq. (2.7) that arise from adversarial contexts. In the following Lemma 2.4.4, we show that the Lagrangian functions, as well as dual variables, are indeed bounded by some absolute constants under a mild feasibility Assumption 2.4.2 stated below:

**Assumption 2.4.2** (Strictly feasible global ROI constraints)**.** *For any realization of value-cost pairs $\boldsymbol{z} = (\boldsymbol{v}_j, \boldsymbol{d}_j)_{j\in[M]} \in F_1 \times \ldots F_M$, the realized version of the ROI constraint in* GL-OPT *defined in Eq. 2.1 is strictly feasible, i.e. the set*

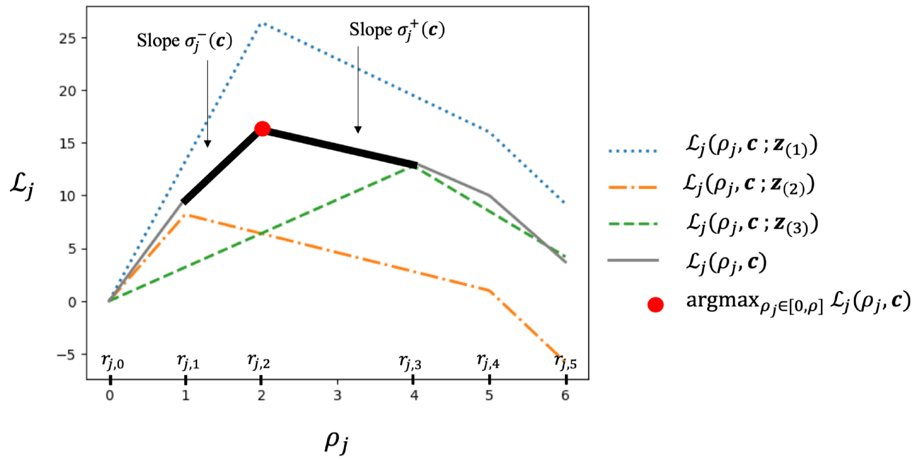


Figure 2-2: Illustration of Lagrangian functions defined in Eq. (2.7) with $M_j = 2$ auctions in channel $j$, and support $F_j$ that contains 3 elements, $\boldsymbol{z}_{(1)} = (\boldsymbol{v}_{(1)}, \boldsymbol{d}_{(1)}) = ((8,2),(2,3))$, $\boldsymbol{z}_{(2)} = ((3,4),(1,4))$, $\boldsymbol{z}_{(3)} = ((8,1),(4,2))$, and context $\boldsymbol{c} = (\lambda, \mu) = (4,2)$. In light of Lemma 2.4.3, $S_j = 5$, where the "turning points" $r_{j,0} \ldots r_{j,S_j}$ are indicated on the x-axis, and the optimal budget w.r.t. $\boldsymbol{c}$ is $\arg\max_{\rho_j \in [0,\rho]} \mathcal{L}_j(\rho_j; \boldsymbol{c}) = r_{j,2}$. The adjacent slopes in Eq. (2.18) are $\sigma_j^-(\boldsymbol{c}) = \sigma_{j,2}(\boldsymbol{c})$, and $\sigma_j^+(\boldsymbol{c}) = \sigma_{j,3}(\boldsymbol{c})$, respectively.

$\left\{\boldsymbol{x} = (\boldsymbol{x}_j)_{j\in[M]} : \boldsymbol{x}_j \in [0,1]^{m_j} \textit{ for } \forall j \in [M],\ \sum_{j\in[M]} \boldsymbol{v}_j^\top \boldsymbol{x}_j > \gamma \sum_{j\in[M]} \boldsymbol{d}_j^\top \boldsymbol{x}_j \right\}$ *is nonempty.*[5]

**Lemma 2.4.4** (Bounding dual variables and Lagrangian functions)**.** *Let $(\lambda_t, \mu_t)_{t\in[T]}$ be the variables generated from Algorithm 1. Under Assumption 2.4.2, and assuming the step size $\eta > 0$ satisfies $\eta < \frac{1}{M\sqrt{(\bar{V}+\gamma\rho)^2+\rho^2}}$, there exists some absolute constant $C_F > 0$ that depends only on the support of value-cost pairs $F = F_1 \times \cdots \times F_M$, as well as aggregate target ROI and budget $(\gamma, \rho)$, such that $\lambda_t, \mu_t \le C_F$ for all $t \in [T]$. Moreover, for any for any $t \in [T]$, $j \in [M]$ and $\rho_j \in [0, \rho]$ we have*

$$-(1+\gamma)\,\rho C_F \ \le\ \mathcal{L}_j(\rho_j, \boldsymbol{c}_t)\ \le\ (1 + C_F)\bar{V}\,. \tag{2.17}$$

See the proof in Appendix A.2.4.

Finally, we address the context-dependent exploration-exploitation tradeoff. We remark that this tradeoff is embodied in the "flatness" of the reward function that

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

depends on adversarial contexts. To illustrate (see e.g. Figure 2-2), we define the slopes that are adjacent to the optimal budget $\arg\max_{\rho_j \in [0,\rho]} \mathcal{L}_j(\rho_j, \boldsymbol{c})$ for any $\boldsymbol{c} = (\lambda, \mu)$ as followed: assuming the optimal budget $\arg\max_{\rho_j \in [0,\rho]} \mathcal{L}_j(\rho_j, \boldsymbol{c})$ is located at the $n$th "turning point" $r_{j,n}$ we have

$$\sigma_j^-(\boldsymbol{c}) = \sigma_{j,n}(\boldsymbol{c}) \text{ and } \sigma_j^+(\boldsymbol{c}) = \sigma_{j,n+1}(\boldsymbol{c}) \tag{2.18}$$

Similar to standard multi-arm bandits exploration-exploitation tradeoffs, the flatter the slope (e.g. $\sigma_j^-(\boldsymbol{c})$ is close to 0), the more pulls required to accurately estimate rewards for sub-optimal arms on the slope, but the lower the loss in conversion for pulling sub-optimal arms. Our setting is challenging because the magnitude of this tradeoff depends on the adversarial contexts $\boldsymbol{c}_t$, i.e., the dual variables, which requires delicate treatments. In the following Lemma 2.4.5 where we bound the UCB error, we handle this context-dependent tradeoff by separately analyzing periods during which the adjacent slopes $\sigma_j^-(\boldsymbol{c})$ and $\sigma_j^+(\boldsymbol{c})$ are less or greater than some parameter $\underline{\sigma}$, and characterize the context-dependent tradeoff w.r.t. flatness of adjacent slopes using $\underline{\sigma}$.

**Lemma 2.4.5** (Bounding the UCB error in per-channel budgets)**.** *Assume the discretization width $\delta$ 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 2.4.3. Then we have*

$$\sum_{t > K} \mathbb{E}\left[\mathcal{L}_j(\rho_j^*(t), \boldsymbol{c}_t) - \mathcal{L}_j(\rho_{j,t}, \boldsymbol{c}_t)\right] \;\leq\; \widetilde{\mathcal{O}}\left(\delta T + \underline{\sigma} T + \frac{1}{\underline{\sigma}\delta} + \frac{1}{\delta T}\right), \tag{2.19}$$

*where $\underline{\sigma} > 0$ is any small positive absolute constant that we can specify later, and recall $\rho_j^*(t) = \arg\max_{\rho_{j \in [M]}} \mathcal{L}_j(\rho_j; \boldsymbol{c}_t)$ is the optimal per-channel budget w.r.t. dual variables $\boldsymbol{c}_t$. $\widetilde{\mathcal{O}}$ hides all logarithmic factors.*

We refer the readers to Appendix A.2.5 for the proof. Note that the parameter $\underline{\sigma}$ will be chosen later. Finally, returning to bounding the UCB error in Proposition 2.4.1, we put together Lemmas 2.4.2 and 2.4.5, and obtain the main result of this section in the following Theorem 2.4.6 whose proof we detail in Appendix A.2.6.

**Theorem 2.4.6** (Putting everything together)**.** *Assume assumptions 2.4.1 and 2.4.2 hold. Let* $(\rho_{j,t})_{j\in[M],t\in[T]}$ *be the per-channel budgets generated from Algorithm 1 and assume we take step size* $\eta = \Theta(1/\sqrt{T})$*, discretization width* $\delta = \Theta(T^{-1/3})$*, and* $\underline{\sigma} = \Theta(T^{-1/3})$ *in Lemma 2.4.5. Then for large enough* $T$ *we have* $T \cdot \text{GL-OPT} - \mathbb{E}\left[\sum_{t\in[T]}\sum_{j\in[M]} V_j(\rho_{j,t})\right] \le \mathcal{O}(T^{2/3})$*. Recalling* $\overline{\boldsymbol{\rho}}_T = \left(\frac{1}{T}\sum_{t\in[T]}\rho_{j,t}\right)_{j\in[M]}$ *is the vector of time-averaged per-channel budgets, this implies*

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

*as well as approximate constraint satisfaction*

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

Here, we note that the above approxamate constraint satisfaction is in expectation, similar to our definition of CH-OPT($\mathcal{I}_B$) defined in Eq. (2.3). To conclude, we make an important remark that distinguishes our result in Theorem 2.4.6 with related literature on convex optimization.

**Remark 2.4.2.** *In light of Lemma 2.4.3, the advertiser's optimization problem* CH-OPT($\mathcal{I}_B$) *in Eq.* (2.3) *effectively becomes a convex problem (see Proposition A.2.4 in Appendix A.2.9). Hence it may be tempting for one to directly employ off-the-shelf convex optimization algorithms. However, our problem involves stochastic bandit feedback, and more importantly, uncertain constraints, meaning that we cannot analytically determine whether a primal decision satisfies the constraints of the problem. For example, in* CH-OPT($\mathcal{I}_B$)*, for some primal decision* $(\rho_j)_{j\in[M]}$*, we cannot determine whether the ROI constraint* $\sum_{j\in M}\mathbb{E}\left[V_j(\gamma_j,\rho_j;\boldsymbol{z}_j) - \gamma D_j(\gamma_j,\rho_j;\boldsymbol{z}_j)\right] \ge 0$ *holds because the distribution* $(\boldsymbol{p}_j)_{j\in[M]}$ *from which* $\boldsymbol{z}$ *is sampled is unknown. To the best of our knowledge, there are only two recent works that handle a similar stochastic bandit feedback, and uncertain constraint setting, namely [105] and [93]. Nevertheless, our setting is more challenging because these works consider a "two-point estimation" regime where one can make function evaluations to the objective and constraints twice* *each period, whereas our setting involves "one-point estimation" such that we can only make function calls once per period. We note the optimal oracle complexities for unknown constraint convex optimization with one-point bandit feedback, remains an open problem.*[6]

### 2.4.3 Extension to strict constraint satisfaction: UCB-SGD-II

In Theorem 2.4.6, we showed that the final output of the UCB-SGD Algorithm 1 outputs a per-channel budget profile $\overline{\boldsymbol{\rho}}_T = \left(\frac{1}{T}\sum_{t\in[T]}\rho_{j,t}\right)_{j\in[M]}$ that satisfies both ROI and budget constraints in CH-OPT($\mathcal{I}_B$) approximately, i.e. there can be at most violations in the magnitude of $\mathcal{O}(T^{-1/2})$ for both constraints. In this subsection, we present a modification to UCB-SGD that enables us to achieve no-constraint violations, while still retaining the $\mathcal{O}(T^{-1/3})$ accuracy in total conversion. Similar modification techniques have been introduced in [18, 50].

Our modification strategy handles ROI and budget constraint satisfactions differently. For budget constraints, we simply maintain a spend balance $B_t$ in each period starting from $B_1 = 0$, and increase the balance by the expenditure in each period. When the balance nears $\rho T$, i.e. total spend comes close to $\rho T$, we simply terminate the algorithm. Regarding the ROI constraint, we develop two phases. Phase 1 is a "safety buffer phase" where we conservatively set per-channel budgets to accumulate a positive "ROI balance", i.e. in this phase (assume ending in period $T_1$) we hope to achieve $\sum_{t\in[T_1]} g_{1,t} \geq \Theta(\sqrt{T})$, where we recall $-g_{1,t}$ defined in step 4 of Algorithm 1 can be viewed as the ROI constraint violations in period $t$. For phase 2, we then naively run SGD-UCB. The motivation for this two-phase design is that we aim to have a buffer, i.e. positive ROI balance, in phase 1 that can compensate for possible constraint violations in phase 2 when we run SGD-UCB (see Theorem 2.4.6). We call our algorithm SGD-UCB-II which we present in Algorithm 2.

We remark that in order to implement the buffer phase 1 to attain a positive ROI balance, we rely on the following Assumption 2.4.3 which is a strengthened version of Assumption 2.4.2 that states in each channel there is always an auction that has a

[6] See Table 4.1 in [81] for best known complexity bounds for one-point bandit feedback setups.

value-to-cost ratio above the global target ROI $\gamma$. Then by setting a small budget we denote as $\beta$, the channel will only procure impressions with high value-to-cost ratios (due to the structure of conversion functions in Lemma 2.4.3), and thus ensuring that the ROI balance increases.

**Algorithm 2** SGD-UCB-II

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

*Phase 1 – Accumulate ROI balance buffer*

```
while $\sum_{t'\in[t-1]} g_{1,t'} \leq \sqrt{T}\log(T)$ do
    if $B_t + M\rho > \rho T$ then
        Terminate algorithm and output $\overline{\boldsymbol{\rho}} = \left(\frac{1}{T}\sum_{t'\in[t]}\rho_{j,t'}\right)_{j\in[M]}$.
    end if
    Set per-channel budget to be $\beta$ for all channels. Observe conversion $V_j(\beta;\boldsymbol{z}_t)$ for all $j \in [M]$,
    and calculate
        $$g_{1,t} = \sum_{j\in M}\left(V_j(\beta;\boldsymbol{z}_t) - \gamma\beta\right).$$
    Calculate $B_{t+1} = B_t + M\beta$.
    Increment $t \leftarrow t + 1$
end while
Denote end period of Phase 1 as $T_1 = t - 1$.
```

*Phase 2 – Run SGD-UCB*

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

**Assumption 2.4.3** (Strictly feasible per-channel ROI constraints)**.** *Fix any channel $j \in [M]$ and any realization of value-cost pairs $\boldsymbol{z}_j = (\boldsymbol{v}_j, \boldsymbol{d}_j) \in F_j$, the channel's optimization problem in Eq. 2.4 is strictly feasible, i.e. the set $\left\{\boldsymbol{x}_j \in [0,1]^{m_j} : \boldsymbol{v}_j^\top\boldsymbol{x}_j > \gamma\boldsymbol{d}_j^\top\boldsymbol{x}_j\right\}$ is nonempty.*[7]

Our strategy to bound the performance of SGD-UCB-II is as followed: In the first phase, we show that we acquire sufficient ROI balance buffers to compensate for

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

global ROI constraint violation in the second phase. On the other hand, conversion loss in the second phase is solely due to SGD-UCB (Algorithm 1), and thus our proof to bound such loss follows from similar ideas in Theorem 2.4.6 (but here we have to additionally handle "early stopping" of UCB-SGD-II due to the spend balance check.

Note that in the first ROI balance buffer phase, we are not optimizing for per-channel budgets which may lead to significant per-period conversion loss. Nevertheless, in the following lemma, we first show that the first ROI balance buffer phase does not last too long; We refer readers to Appendix A.2.7 for the proof.

**Lemma 2.4.7** (Bounding length of Phase 1)**.** *Recall $T_1 \in [T]$ is the end period of Phase 1 in the SGD-UCB-II algorithm (see step 10). Denote the event $\mathcal{E} = \{T_1 \geq 2\sqrt{T}\log^3(T)\}$, and take small budget $\beta = \frac{1}{\log(T)}$ in the SGD-UCB-II algorithm. Then, under Assumption 2.4.2, for large enough $T$ we have $\mathbb{P}(\mathcal{E}) \leq \frac{1}{T}$.*

Our main result in this subsection is the following Theorem 2.4.8 which bounds the conversion loss of SGD-UCB-II. The proof is detailed in Appendix A.2.8.

**Theorem 2.4.8.** *Suppose that Assumptions 2.4.1 and 2.4.3 hold. Let $T_2 \in [T]$ be the termination period of SGD-UCB-II (Algorithm 2), and recall $\overline{\boldsymbol{\rho}} = \left(\frac{1}{T}\sum_{t\in[T_2]}\rho_{j,t}\right)_{j\in[M]}$ are the final outputs of the algorithm. Further, we input $\beta = \frac{1}{\log(T)}$ to the algorithm that is used for Phase 1, and in Phase 2 where we run SGD-UCB, we use the same parameters as those in Theorem 2.4.6. Then, for large enough $T$ we have $\text{GL-OPT} - \sum_{j\in[M]}\mathbb{E}\left[V_j(\overline{\rho}_j)\right] \leq \mathcal{O}(T^{-1/3})$, and further $\sum_{j\in[M]}\mathbb{E}\left[V_j(\overline{\rho}_j) - \gamma\overline{\rho}_j\right] \geq 0$ and $\sum_{j\in[M]}\mathbb{E}\left[\overline{\rho}_j\right] \leq \rho$.*

## 2.5 Generalizing to autobidding in multi-item auctions

In previous sections, we assumed that each channel consists of multiple auctions, each of which is associated with the sale of a single ad impression (see Eq. (2.4) and discussions thereof). Yet, in practice, there are many scenarios in which ad platforms sell multiple impressions in each auction (see e.g. [106, 46]). Thereby in this section, we extend all our results for the single-item auction setting in previous sections to the multi-item auction setup. In Section 2.5.1, we formally describe the multi-item setup; in Section 2.5.2 we show that in the multi-item setting, the per-budget ROI lever is again redundant (similar to what is shown in Theorem 2.3.4 and Corollary 2.3.5), and an advertiser can solely optimize over per-channel budgets to achieve the global optimal conversion; in Section 2.5.3, we show our proposed UCB-SGD algorithm is directly applicable to the multi-item auction setup for a broad class of auctions, and similar to Theorem 2.4.6, our algorithm produces accurate lever estimates with which the advertiser can approximate the globally optimal lever decisions.

### 2.5.1 Multi-item autobidding setup

We first formalize our multi-item setup as followed. For each auction $n \in [m_j]$ of channel $j \in [M]$, assume $L_{j,n} \in \mathbb{N}$ impressions are sold, and channel $j$ is only allowed to procure at most 1 impression in auction $n$ on the advertiser's behalf. The value acquired and cost incurred by the advertiser when procuring impression $\ell \in [L_{j,n}]$ are $v_{j,n}(\ell)$ and $d_{j,n}(\ell)$, respectively. With a slight abuse of notation from previous sections, we write $\boldsymbol{v}_{j,n} = (v_{j,n}(1), \dots v_{j,n}(L_{j,n})) \in \mathbb{R}_+^{L_{j,n}}$ as the $L_{j,n}$-dimensional vector that includes all impression values of auction $n$ in channel $j$, and further write $\boldsymbol{v}_j = (\boldsymbol{v}_{j,1} \dots \boldsymbol{v}_{j,m_j}) \in \mathbb{R}_+^{\sum_{n\in[m_j]} L_{j,n}}$ as the vector that concatenates all value vectors across auctions in channel $j$. We also define $\boldsymbol{d}_{j,n} \in \mathbb{R}_+^{L_{j,n}}$ and $\boldsymbol{d}_j \in \mathbb{R}_+^{\sum_{n\in[m_j]} L_{j,n}}$ accordingly for costs. Similar to Section 2.2, we assume $\boldsymbol{z}_j = (\boldsymbol{v}_j, \boldsymbol{d}_j)$ is sampled from finite support $F_j$ according to discrete distribution $\boldsymbol{p}_j$ for any channel $j \in [M]$, and without loss of generality, we assume that for any element $\boldsymbol{z}_j \in F_j$, the value 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. (2.1) for the single-item auction setup in previous sections), can be written as the following problem called GL-OPT$^+$:

$$
\begin{aligned}
&\text{GL-OPT}^+ \;=\; \\
&\max_{\left(\boldsymbol{x}_j=(\boldsymbol{x}_{j,1}\ldots,\boldsymbol{x}_{j,m_j})\right)_{j\in[M]}} \quad \sum_{j\in[M]} \mathbb{E}\left[\boldsymbol{v}_j^\top \boldsymbol{x}_j\right] \\
&\qquad\qquad \text{s.t.} \quad \sum_{j\in[M]} \mathbb{E}\left[\boldsymbol{v}_j^\top \boldsymbol{x}_j\right] \;\geq\; \gamma \sum_{j\in[M]} \mathbb{E}\left[\boldsymbol{d}_j^\top \boldsymbol{x}_j\right] \\
&\qquad\qquad\qquad \sum_{j\in[M]} \mathbb{E}\left[\boldsymbol{d}_j^\top \boldsymbol{x}_j\right] \;\leq\; \rho \\
&\qquad\qquad\qquad \boldsymbol{x}_{j,n} \in [0,1]^{\sum_{n\in[m_j]} L_{j,n}} \text{ and } \sum_{\ell\in[L_{j,n}]} x_{j,n}(\ell) \leq 1, \quad \forall j \in [M],\ n \in [m_j]
\end{aligned}
\tag{2.20}
$$

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

On the other hand, analogous to a channel's autobidding problem for the single-item auction setup in previous sections (Eq. (2.4)), in the multi-item setting each channel $j$'s autobidding problem can be written as

$$
\begin{aligned}
\boldsymbol{x}_j^{*,+}(\gamma_j,\rho_j;\boldsymbol{z}_j) \;=\; \arg\max_{\boldsymbol{x}=(\boldsymbol{x}_1\ldots\boldsymbol{x}_{m_j})} \; & \boldsymbol{v}_j^\top \boldsymbol{x} \\
\text{s.t.} \quad & \boldsymbol{v}_j^\top \boldsymbol{x} \geq \gamma_j \boldsymbol{d}_j^\top \boldsymbol{x}, \text{ and } \quad \boldsymbol{d}_j^\top \boldsymbol{x} \leq \rho_j \\
& \boldsymbol{x}_n \in [0,1]^{L_{j,n}} \text{ and } \sum_{\ell\in[L_{j,n}]} x_n(\ell) \leq 1, \quad \forall n \in [m_j]
\end{aligned}
\tag{2.21}
$$

where $\boldsymbol{x}_n = (x_n(\ell))_{\ell\in[L_{j,n}]} \in [0,1]^{m_j}$ denotes the (possibly random) vector of indicators to win each impression of auction $n$ in channel $j$. With respect to this per-channel multi-item auction optimization problem in Eq. (2.21), we can further define $V_j^+(\gamma_j,\rho_j;\boldsymbol{z}_j)$, $D_j^+(\gamma_j,\rho_j;\boldsymbol{z}_j)$, $V_j^+(\gamma_j,\rho_j)$, $D_j^+(\gamma_j,\rho_j)$ as in Eq.(2.5), and CH-OPT$^+(\mathcal{I})$ as in Eq.(2.3) for any advertiser lever option $\mathcal{I}$ in Eq.(2.2).

### 2.5.2 Optimizing per-channel budgets is sufficient to achieve global optimal

Our first main result for the multi-item setting is the following Theorem 2.5.1 which again shows an advertiser can achieve the global optimal conversion GL-OPT$^+$ via solely optimizing over per-channel budgets (analogous to Theorem 2.3.4 and Corollary 2.3.5).

**Theorem 2.5.1** (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.(2.2), we have* GL-OPT$^+$ = CH-OPT$^+(\mathcal{I}_B)$ = CH-OPT$^+(\mathcal{I}_G)$ *for any aggregate ROI $\gamma > 0$ and total budget $\rho > 0$, even for $\rho = \infty$. Further, there exists and optimal solution $(\gamma_j, \rho_j)_{j\in[M]}$ to* CH-OPT$^+(\mathcal{I}_G)$*, s.t. $\gamma_j = 0$ for all $j \in [M]$.*

It is easy to see that the proof of Lemma 2.3.1, Theorem 2.3.4, and Corollary 2.3.5 w.r.t. the single item setting in Section 2.3 can be directly applied to Theorem 2.5.1 since we did not rely on specific structures of the solutions to GL-OPT and CH-OPT other than the presence of the respective ROI and budget constraints (which are still present in GL-OPT$^+$ and CH-OPT$^+$). Thereby we will omit the proof of Theorem 2.5.1. In light of Theorem 2.5.1, we again conclude that the per-channel ROI lever is redundant, and hence omit per-channel ROI $\gamma_j$ when the context is clear.

### 2.5.3 Applying UCB-SGD to the multi-item setting

We now turn to our second main focus of the multi-item setting, which is to understand whether our proposed UCB-SGD algorithm can achieve accurate approximations to the optimal per-channel budgets, similar to Theorem 2.4.6 for the single-item setting. A key observation is that the only difference between bounding the error of UCB-SGD in the single and multi-item settings is the structure of the conversion and corresponding Lagrangian functions (see Lemma 2.4.3), since the only change in the multi-item setting compared to the single-item setting is how a given per-channel budget translates into a certain conversion. Therefore, in this section we introduce a broad class of multi-item auction formats that induce the same conversion function structural properties as those illustrated in Lemma 2.4.3, which will allows us to directly apply the proof for bounding the error of UCB-SGD (Theorem 2.4.6) to the multi-item setting of interest.

To begin with, we introduce the following notion of increasing marginal values, which is a characteristic that preserves the structural properties for conversion and Lagrangian functions from the single-item setting (in Lemma 2.4.3), as shown later in Lemma 2.5.2.

**Definition 2.5.1** (Multi-item auctions with increasing marginal values). *We say an auction $n \in [m_j]$ in channel $j \in [M]$ has increasing marginal values if for any realization $\boldsymbol{z}_j = (\boldsymbol{v}_j, \boldsymbol{d}_j)$, we have*

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

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

Increasing marginal values intuitively means that in some multi-item auction, the marginal value per cost gained increases with procuring impressions of greater values. Many classic position auction formats satisfy increasing marginal gains, such as the Vickrey–Clarke–Groves (VCG) auction; see [106, 46] for more details on position auctions.

**Example 2.5.1** (VCG auctions have increasing marginal values). *Let auction $n \in [m_j]$ in channel $j \in [M]$ be a VCG auction, where for any realization of $(\boldsymbol{v}_{j,n}, \boldsymbol{d}_{j,n}) = (v_{j,n}(\ell), d_{j,n}(\ell))_{\ell \in [L_{j,n}]}$ there exists some $\widetilde{v}_{n,j} > 0$, position discounts $1 \geq \theta_{n,j}(1) > \theta_{n,j}(2) \ldots \theta_{n,j}(L_{n,j}) > 0$, and $L_{n,j}$-highest competing bids from competitors in the market $\widetilde{d}_{n,j}(1) > \widetilde{d}_{n,j}(2) \ldots > \widetilde{d}_{n,j}(L_{n,j}) > 0$, such that the acquired value for procuring impression $\ell \in L_{n,j}$ is $v_{n,j}(\ell) = \theta_{n,j}(\ell) \cdot \widetilde{v}_{n,j}$, and the corresponding cost is $d_{j,n}(\ell) = \sum_{\ell'=\ell}^{L_{j,n}} (\theta_{n,j}(\ell') - \theta_{n,j}(\ell'+1)) \widetilde{d}_{n,j}(\ell')$ where we denote $\theta_{n,j}(L_{j,n}+1) = 0$.[8] Thereby, under*

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

*VCG the marginal values are*

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

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

We remark that the generalized second price auction (GSP) does not necessarily have increasing marginal values. Now, if all auctions in a channel have increasing marginal values, then we can show the conversion function $V_j^+(\rho_j)$ and the corresponding Lagrangian function for multi-item auctions admits the same structural properties as those in Lemma 2.4.3:

**Lemma 2.5.2** (Structural properties for multi-item auctions)**.** *For any channel $j \in [M]$ whose auctions have increasing marginal values (see Definition 2.5.1), the conversion function $V_j^+(\rho_j) = \mathbb{E}\left[\boldsymbol{v}_j^\top \boldsymbol{x}_j^{*,+}(\rho_j;\boldsymbol{z}_j)\right]$ is continuous, piece-wise linear, strictly increasing, and concave. Here recall $\boldsymbol{x}_j^{*,+}(\rho_j;\boldsymbol{z}_j)$ is the optimal solution to the channel's optimization problem in Eq.* (2.21)*. Further, for any dual variables $\boldsymbol{c} = (\lambda, \theta) \in \mathbb{R}_+^2$, the Lagrangian function $\mathcal{L}_j^+(\rho_j, \boldsymbol{c}) := (1+\lambda)V_j^+(\rho_j) - (\theta + \gamma\lambda)\rho_j$ is continuous, piece-wise linear, concave, and unimodal in $\rho_j$.*

See the proof in Appendix A.3.1. In light of Lemma 2.5.2, we can argue that UCB-SGD produces per-channel budgets that yield the same accuracy as in Theorem 2.4.6 for the single-item setting,

**Theorem 2.5.3** (UCB-SGD applied to channel procurement for multi-item auctions)**.** *Assume multi-item auctions in any channel $j \in [M]$ has increasing marginal values (per Definition 2.5.1), and assume Assumptions 2.4.1 and 2.4.2 hold for the multi-item setting.*[9] *Then with the same parameter choices as in Theorem 2.4.6, and recalling*

[9]Assumption 2.4.1 in the multi-item setting again implies the spend in any channel is exactly the input per-channel budget; Assumption 2.4.2 in the multi-item setting states that for any realization of value-cost pairs $\boldsymbol{z} = (\boldsymbol{v}_j, \boldsymbol{d}_j)_{j\in[M]} \in F_1 \times \ldots F_M$, the realized version of the ROI constraint in GL-OPT$^+$ defined in Eq. (2.20) is strictly feasible.

$\overline{\boldsymbol{\rho}}_T = \left(\frac{1}{T}\sum_{t\in[T]}\rho_{j,t}\right)_{j\in[M]}$ *is the vector of time-averaged per-channel budgets produced by UCB-SGD, we have*

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

*as well as approximate constraint satisfaction*

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

*where we recall* GL-OPT$^+$ *is defined in Eq.* (2.20), $V_j^+(\rho_j) = \mathbb{E}\left[\boldsymbol{v}_j^\top \boldsymbol{x}_j^{*,+}(\rho_j;\boldsymbol{z}_j)\right]$ *and* $\boldsymbol{x}_j^{*,+}(\rho_j;\boldsymbol{z}_j)$ *is defined in Eq.* (2.21).

The proof for this theorem is identical to that of Theorem 2.4.6 given the same structural properties of the conversion and Lagrangian functions in Lemma 2.5.2 and Lemma 2.4.3. Hence we will omit the proof. Finally, we remark that UCB-SGD-II (Algorithm 2) can also be applied to the multi-item setting and yield per-channel budget estimates that achieve the same performance as illustrated in Theorem 2.4.8 while satisfying both global budget and ROI constraints exactly.

## 2.6 Additional discussions and future research

### More general advertiser objectives

In GL-OPT and CH-OPT($\mathcal{I}$) defined Section 2.2 (or similarly GL-OPT$^+$ and CH-OPT$^+(\mathcal{I})$ defined in the multi-item setting in Section 2.5), we can also consider more general objectives, namely $\max_{\boldsymbol{x}_1,\dots,\boldsymbol{x}_M} \sum_{j\in[M]} \mathbb{E}\left[\boldsymbol{v}_j^\top \boldsymbol{x}_j - \alpha \boldsymbol{d}_j^\top \boldsymbol{x}_j\right]$ and $\max_{(\gamma_j,\rho_j)_{j\in[M]}\in\mathcal{I}} \sum_{j\in M} \mathbb{E}\left[V_j(\gamma_j,\rho_j;\boldsymbol{z}_j) - \alpha V_j(\gamma_j,\rho_j;\boldsymbol{z}_j)\right]$ for some private cost $\alpha \in [0,\gamma]$[10] in GL-OPT and CH-OPT($\mathcal{I}$), respectively. When $\alpha = 0$, we recover our considered models in the previous section, whereas in when $\alpha = 1$, we obtain the classic quasi-linear utility. We remark that this private cost model has been introduced

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

and studied in related literature; see [13] and references therein. Nevertheless, when each channel's autobidding problem remains as is in Eq.(2.4), 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 2.3, and our UCB-SGD algorithm still produces estimates of the same order of accuracy via introducing $\alpha$ 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 $\alpha$ through optimizing CH-OPT($\mathcal{I}$) that also incorporates the same private cost.

## Future research on non-optimal autobidding in channels

We recall in previous sections we assumed that each channel adopts "optimal autobidding" that solves Eq. (2.4) to optimality. This raises the natural question that whether our findings will still hold when channels do not procure ads optimally, perhaps because of non-stationary environments [22, 85, 29], or the presence of strategic market participants who aim to manipulate the market [54, 42, 63, 55]. In such a scenario, an advertiser's (bandit) conversion feedback in a channel $j$ would be $V(\gamma_j, \rho_j; \boldsymbol{z}_j) - \epsilon_j$ for some channel-specific and possibly adversarial loss $\epsilon_j > 0$. One potential resolution is to treat such $\epsilon_j$ as adversarial corruptions to bandit rewards, and instead of integrating vanilla UCB with SGD as in Algorithm 1, augment SGD with bandit algorithms that are robust to robust corruptions; see e.g. [86, 64]. Nevertheless, it remains an open question to prove how such augmentation would perform in our specific bandit-feedback constrained optimization setup. This leads to potential research directions of both practical and theoretical significance.

# Chapter 3

# Learning to price against strategic buyers

*This chapter is based on [54], which is joint work with Negin Golrezaei and Patrick Jaillet.*

## 3.1 Introduction

In the previous Chapter 2 we studied how advertisers can effectively interact with autobidding ecosystems to procure ads from ad platforms, and abstracted away from details on the procurement algorithms and selling mechanisms run by platforms. In this chapter, we turn to investigate an ad platform's problem to design ad selling mechanisms that specifically take the form of ad auctions with reserve prices, which is arguably one of the most widely adopted forms of selling mechanisms in online advertising; see e.g. [106, 46, 94, 39, 49].

To put the ad platform's auction design problem in more context, consider an autobidding ecosystem for display ads: when a user makes a request (e.g. viewing a webpage or mobile app), the ad platform receives the request and triggers an ad auction, in which automated procurement (i.e. bidding) algorithms compete on behalf of advertisers for the auctioned "item"-the opportunity to display an ad to the user. It is important to note that user requests are highly heterogeneous, meaning that users' preferences, or likelihood of clicking the ad, vary. Such variation in user preferences can be characterized by user features which are typically modelled as contexts of the auctioned item [54]. From an ad platform's perspective, both the advertiser-platform interaction component and the automated procurement component can be perceived as virtual buyers directly submitting bids according to some algorithm that aims to satisfy certain objectives (i.e. the corresponding advertisers' procurement goals). Under this view, this chapter focuses on the problem of designing dynamic reserve pricing policies for highly heterogeneous items against strategic buyers whose bidding behavior may aim to manipulate the platform's pricing policies. This involves learning buyers' demand, which is a mapping from item context and offered prices to the likelihood of the item being sold, under limited understanding of buyers' bidding behavior. The key goal of this chapter is to develop effective and robust dynamic pricing polices that facilitate such a complex learning process for very general non-parametric contextual demand curves facing strategic buyers.

Formally, we study the setting wherein any period $t$ over a finite time horizon

$T$, a seller (i.e. an ad platform) sells one item (i.e. an ad impression) to buyers via running a second price auction with a reserve price. The item is characterized by a $d$-dimensional context vector $x_t$, public to the seller and buyers. We consider an interdependent contextual valuation model in which a buyer's valuation for the item is the sum of common and private components. The common component determines the expected willingness-to-pay of buyers and is the inner product of the feature vector and a fixed "mean vector" $\beta$ that is homogeneous across buyers; the private component, which captures buyers' idiosyncratic preferences, is independently sampled from an unknown *non-parametric* noise distribution $F$. We note that such a linear valuation model is very common in the literature of dynamic pricing; e.g. see [57, 71, 74, 70].

Under this interdependent contextual valuation model, we study a *strategic setting* where buyers intend to maximize long-term discounted utility and may consequently submit *corrupted*, i.e., untruthful bids. The motivation of this strategic setting comes from the repeated buyer-seller interactions when the seller does not possess full information on buyers' demand and aims to learn it using buyers' submitted bids. In a single-shot second price auction, where there is no repeated interactions between the seller and buyers, bidding truthfully is a buyer's weakly dominant action. However, this is no longer the case in our repeated second price auction setting: repeated auctions may incentivize the buyers to submit corrupted bids, rather than their true valuations, in order to manipulate seller's future reserve prices; e.g. by underbidding, buyers may trick the seller to lower future reserve prices.

In this chapter, we would like to design a reserve price policy for the seller who does not know the mean vector $\beta$ and the noise distribution $F$. The policy dynamically learns/optimizes contextual reserve prices while being robust to corrupted data (bids), submitted by strategic buyers. In particular, our objective is to minimize our policy's regret computed against a clairvoyant benchmark policy that knows both $\beta$ and $F$. Designing low-regret policies in our setting involves overcoming the following challenges: (i) The demand curve is constantly shifting due to the change in contexts over time. (ii) The shape of the demand curve is unknown due to the lack of information on the market noise distribution $F$ which may not enjoy a parametric functional form.

Furthermore, we do not impose the *Monotone Hazard Rate* (MHR)[1] assumption on $F$. While the MHR assumption is common in the related literature and can significantly simplify reserve price optimization (see e.g. Remark 3.3.1), it has been shown to fail in practice (see [27, 60]). (iii) As stated earlier, in our strategic setting, buyers may take advantage of the seller's lack of knowledge about buyers' demand and submit corrupted bids to manipulate future reserve prices.

**Main contribution.** We develop a policy called *Non-Parametric Contextual Policy against Strategic Buyers* (NPAC-S) that enables the seller to efficiently learn the optimal contextual reserve prices while being robust against buyers' corrupted bids. Our policy design incorporates two simple yet effective features, namely a *phased structure* and *random isolation*. First, NPAC-S partitions the entire horizon into consecutive phases, and then estimates the mean vector and the distributions of the second-highest and highest valuations only using data from the previous phase. This reduces the buyers' manipulating power on future reserve prices as past corrupted bids prior to the previous phase will not affect future pricing decisions. Second, the NPAC-S policy incorporates randomized isolation periods, that is, in each period with some probability the seller chooses a particular buyer at random and let her be the single participant of the auction during this period. In these isolation periods, the isolated buyer faces no competition from other buyers, and hence may incur large utility loss if a significantly corrupted bid is submitted.[2]

For our main theoretical results, we show that in virtue of our isolation periods in our design of NPAC-S, the number of past periods with large corruptions is $\mathcal{O}(\log(t))$ for any period $t$ via leveraging the fact that buyers aim to maximize their long-term discounted utility. Furthermore, we present novel high probability bounds for our estimation errors in $\beta$ and $F$ which are estimated by ordinary least squares and empirical distributions, respectively, with the presence of corrupted bids. Finally, in

[1]Distribution $F$ is MHR if $\frac{f(z)}{1-F(z)}$ is non-decreasing in $z$, where $f$ is the corresponding pdf.

[2]In the isolation periods, when the valuation of the isolated buyer is greater than the reserve price, significantly underbidding may cause the item to not be allocated; when the valuation of the isolated buyer is lower than the reserve price, overbidding results in the buyer paying much higher prices (relative to valuation) to achieve the item. In either case, the isolated buyer will incur a significant utility loss compared to truthful bidding.

Theorem 3.4.2, we show that the NPAC-S policy achieves a regret of $\tilde{\mathcal{O}}(d\sqrt{T})$ for general non-parametric distributions $F$ against a clairvoyant benchmark policy.

## Related works

Here we discuss related works that study dynamic pricing against strategic buyers with stochastic valuations, [3].

Both [6, 7] study a dynamic pricing problem in a posted price auction against a single strategic buyer. [6] addresses the non-contextual stochastic valuation setting, where as [7] studies a linear contextual valuation model, but with no market noise disturbance. [7] proposes an algorithm that achieves $\widetilde{\mathcal{O}}(T^{2/3})$ regret in contrast with our regret of $\widetilde{\mathcal{O}}(\sqrt{T})$ using the NPAC-S policy. We point out that this is because the seller in their setting only observes the outcome of the auction (i.e. bandit feedback), while in our setting we assume that seller can examine all submitted bids. Our setting is more complex compared to [6, 7] as we handle the contextual pricing problem against multiple strategic buyers, and also deals with the issue of learning a non-parametric distribution function in the presence of strategic buyer behavior. [74] consider a contextual buyer valuation model similar to ours (but with the MHR assumption on the market noise distribution) and proposes a pricing algorithm that sets personalized reserve prices for individual buyers. They argue that the design of their algorithm induces an equilibrium where buyers always bid truthfully, and then further assume buyers act according to this equilibrium. This chapter distinguishes itself from two aspects. First, setting personalized reserve prices in [74] rely crucially on the MHR assumption, and in this chapter we relax this assumption such that our methodology works for a larger class of market noise distributions. Second, we consider more general buyers who do not necessarily play any equilibrium and are forward looking. [57] study a similar interdependent contextual valuation model to ours, but with

[3]The general theme of learning in the presence of strategic agents or corrupted information has also been studied in other applications; see, for example, [28, 23, 51]. There are also related works that study adversarial buyer valuations. For example, [41] studies the seller's pricing problem for repeated second-price auctions facing multiple strategic buyers with private valuations fixed overtime. In addition, buyers in this chapter also seek to maximize cumulative discounted utility. The paper proposes an algorithm that achieves $\mathcal{O}(\log\log(T))$ regret for worst-case (adversarial) valuations.

heterogeneous mean vector $\beta$ across agents. This chapter distinguishes itself from [57] in two major ways. First, they focus on optimizing contextual reserve w.r.t. the worst-case distribution among a known class of MHR market noise distributions. In contrast, this chapter relaxes this constraint and does not require the seller to have any prior knowledge on the possibly non-parametric distribution. Second, in their setting, the seller only utilizes the outcome of the auctions to learn buyer demand and results in a regret of $\widetilde{\mathcal{O}}(T^{2/3})$.[4] In this chapter, we exploit the information of all submitted bids by taking advantage of the fact that buyers' utility-maximizing behaviour constrains their degree of corruption on bids. This eventually allows us to achieve an improved regret of $\widetilde{\mathcal{O}}(\sqrt{T})$. Nevertheless, our proposed algorithm cannot not handle heterogeneous $\beta$'s, and hence this will be an interesting future research direction. [42] studies the posted price selling problem against a strategic agent with a non-linear (stochastic) contextual valuation model that satisfies some Lipschitz condition with no additive noise.

We summarize some key differences in the settings/results of the aforementioned works in Table 3.1.

## 3.2 Preliminaries

**Notation.** For $a \in \mathbb{N}^+$, denote $[a] = \{1, 2, \ldots, a\}$. For two vectors $x, y \in \mathbb{R}^d$, denote $\langle x, y \rangle$ as their inner product. Finally, $\mathbb{I}\{\cdot\}$ is the indicator function: $\mathbb{I}\{\mathcal{A}\} = 1$ if event $\mathcal{A}$ occurs and 0 otherwise.

We consider a seller who runs repeated second price auctions over a horizon with length $T$ that is unknown to the seller. In each auction $t \in [T]$, an item is sold to $N$ buyers, where the item is characterized by a $d$-dimensional feature vector $x_t \in \mathcal{X} \subset \{x \in \mathbb{R}^d : \|x\|_\infty \leq x_{\max}\}$ where $0 < x_{\max} < \infty$. We assume that $x_t$ is independently drawn from some distribution $\mathcal{D}$ unknown to the seller. We define $\Sigma$ as

[4] A recent work [35] builds on the result of [57] by considering a stronger benchmark that knows future buyer valuation distributions (noise distribution and all the future contextual information). They design robust pricing schemes whose regret is $\mathcal{O}(T^{5/6})$ against the aforementioned benchmark, confirming the generalizability of pricing schemes in [57].

| Algorithm | # buyers | Context | Noise/value dist. | Disc. util. | Regret |
|---|---|---|---|---|---|
| Phased [6] | 1 | False | Lipschitz | True | Sublinear[5] |
| LEAP [7] | 1 | True | No additive noise | True | $\mathcal{O}(T^{2/3})$ |
| PELS [42] | 1 | True | No additive noise | True | $\mathcal{O}(T^{d/(d+1)})$ |
| HO-SERP [75] | $\geq 2$ | True | MHR | False | $\mathcal{O}(\sqrt{T})$ |
| SCORP [57] | $\geq 2$ | True | MHR | True | $\mathcal{O}(T^{2/3})$ |
| NPAC-S | $\geq 2$ | True | Non-parametric | True | $\mathcal{O}(\sqrt{T})$ |

Table 3.1: Summary of settings and results for seller algorithms that sell against strategic agents with stochastic valuations. Note that the Discount util. column indicates whether the algorithm deals with buyers who discount their long-term utilities. Note that HO-SERP[75] and SCORP [57] set *personalized reserve prices* for each buyer, whereas NPAC-S sets a single reserve for all buyers. PELS in [42] learns a non-linear contextual valuation model and hence yields larger regret. Among all algorithms, only SCORP [57] handles heterogeneous $\beta$ across buyers.

the covariance matrix of distribution $\mathcal{D}$.[6] We assume that $\Sigma$ is positive definite and unknown to the seller, and define the smallest eigenvalue of $\Sigma$ to be $\lambda_0 > 0$.

**Buyer valuation model.** We focus on an interdependent valuation model where the valuation of buyer $i \in [N]$ at time $t \in [T]$is given by $v_{i,t} = \langle \beta, x_t \rangle + \epsilon_{i,t}$. Here, $\beta$ is called the *mean vector* and is fixed over time and unknown to the seller, while $\epsilon_{i,t}$ is idiosyncratic market noise sampled independently over time and across buyers from some time-invariant distribution $F$ with probability density function $f$, both unknown to the seller. Furthermore, $F$ has bounded support $(-\epsilon_{\max}, \epsilon_{\max})$, in which its probability density function is bounded by $c_f := \sup_{z \in [-\epsilon_{\max}, \epsilon_{\max}]} f(z) \geq \inf_{z \in [-\epsilon_{\max}, \epsilon_{\max}]} f(z) > 0$. The support boundary $\epsilon_{\max}$ is not necessarily known to the seller. We assume there exist $v_{\max} > 0$ so that $v_{i,t} \in [0, v_{\max}]$ for all $i \in [N]$, $t \in [T]$.

We highlight that our setting does not enforce distribution $F$ to be parametric nor to satisfy the MHR assumption. This is because via analyzing real auction data sets, it has been shown that the MHR assumption does not necessarily hold in online advertising markets [27, 60].

**Repeated contextual second price auctions with reserve.** The contextual

[6]The covariance matrix of a distribution $\mathcal{P}$ on $\mathbb{R}^d$ is defined as $\mathbb{E}_{x \sim \mathcal{P}}[xx^\top] - \mu\mu^\top$, where $\mu = \mathbb{E}_{x \sim \mathcal{P}}[x]$.

second price auction with reserve is described as followed for $N \geq 2$ buyers: In any period $t \geq 1$, a context vector $x_t \sim \mathcal{D}$ is revealed to the seller and buyers. The seller then computes reserve price $r_t$, while simultaneously each buyer $i \in [N]$ forms individual valuations $v_{i,t}$ and submits a bid $b_{i,t}$ to the seller. Let $i^\star = \arg\max_{i \in [N]} b_{i,t}$ be the buyer who submitted the highest bid.[7] If $b_{i^\star,t} \geq r_t$, the item is allocated to buyer $i^\star$ and he is charged the maximum between the reserve price and second highest bid, i.e. buyer $i^\star$ pays $p_{i^\star,t} = \max\{r_t, \max_{i \neq i^\star} b_{i,t}\}$. For any other buyer $i \neq i^\star$, the payment $p_{i,t} = 0$. In the case where $b_{i^\star,t} < r_t$, the item is not allocated and all payments are zero.

Here, the seller's reserve price $r_t$ can only depend on $x_t$ and the history set $\mathcal{H}_{t-1} := \{(r_1, \{b_{i,1}\}_{i \in [N]}, x_1), \ldots, (r_{t-1}, \{b_{i,t-1}\}_{i \in [N]}, x_{t-1})\}$ which includes all information available to the seller up to period $t-1$.

**Buyers' bidding behavior.** In the setting where buyers are strategic, we assume that in any period $t$, each buyer $i \in [N]$ aims at maximizing his long-term discounted utility $U_{i,t}$:

$$U_{i,t} \ := \ \sum_{\tau=t}^{T} \eta^{\tau} \mathbb{E}\left[v_{i,\tau} w_{i,\tau} - p_{i,\tau}\right] \ , \tag{3.1}$$

where $\eta \in (0,1)$ is the discount factor, $w_{i,t} \in \{0,1\}$ indicates whether buyer $i$ wins the item; the expectation is taken with respect to the randomness due to the noise distribution $F$, the context distribution $\mathcal{D}$, buyers' bidding behavior, and the seller's pricing policy. We point out that this discounted utility model illustrates the fact that buyers are less patient than the seller, and is a common framework in many dynamic pricing literature; see [6, 7, 57], and [84]. The motivation lies in many applications in online advertisement markets wherein the user traffic is usually very uncertain and as a result, advertisers (buyers) would not like to miss out an opportunity of showing their ads to targeted users. An alternative interpretation for the above discounted utility model is that each buyer has probability $\eta$ of leaving the repeated auctions, and thereby the expected cumulative utility of each bidder is exactly Eq. (3.1). It is

[7]No ties will occur since we assume that no valuations and bids are the same.

worth noting that [6] showed, in the case of a single buyer, it is not possible to obtain a no-regret policy when $\eta = 1$, that is, when the buyer is as patient as the seller.

Furthermore, we assume buyers corrupt their true valuations in an additive manner:

$$\forall i \in [N], t \in [T] \quad b_{i,t} = v_{i,t} - a_{i,t} \;\text{ where }\; |a_{i,t}| \leq a_{\max}\,.$$

Here, $a_{i,t}$ is called the degree of corruption, and we refer to the buyer behavior of submitting a bid $b_{i,t} \neq v_{i,t}$ (i.e., $a_{i,t} \neq 0$) as "corrupted bidding". Note that when $a_{i,t} > 0$, the buyer shades her bid, and when $a_{i,t} < 0$, the buyer overbids. Essentially, a buyer $i$'s strategic behavior is equivalent to deciding on a non-zero value of $a_{i,t}$. In this work, we impose no restrictions on the degree of corruption $a_{i,t}$ for a buyer $i$ in period $t$ other than it is bounded. [8]

## 3.3 Benchmark and seller's regret

The seller's revenue in period $t \in [T]$ is the sum of total payments from all buyers, and the expected revenue given context $x_t \in \mathcal{X}$ and reserve price $r_t$ is

$$\begin{aligned} \mathsf{rev}_t(r_t) &:= \mathbb{E}\Big[\sum_{i \in [N]} p_{i,t} \;\Big|\; x_t, r_t\Big], \\ \text{where} \;\; p_{i,t} \;&=\; \max\{b_t^-, r_t\}\mathbb{I}\{b_{i,t} \geq \max\{b_t^+, r_t\}\}\,. \end{aligned} \tag{3.2}$$

Here, $b_t^-$ and $b_t^+$ are the second-highest and highest bids in period $t$, respectively; the expectation is taken with respect to the noise distribution in period $t$ and any randomness in the reserve price $r_t$ as well as bid values submitted by buyers in period $t$ (as buyers' bidding strategies may be randomized).

The seller's objective is to maximize his expected revenue over a fixed time horizon $T$ through optimizing contextual reserve prices $r_t$ for any $t \in [T]$. To evaluate any seller pricing policy, we compare its total revenue against that of a benchmark policy run by a clairvoyant seller who knows the mean vector $\beta$ and the non-parametric

[8] A bound for the degree of corruption is natural as buyers always submit non negative bids and all bids are bounded by $v_{\max}$.

noise distribution $F$. This clairvoyant seller's benchmark policy sets the "optimal" contextual reserve price in each period to obtain the maximum achievable revenue $\max_r \mathsf{rev}_t(r)$ in each period, and hence facing such a seller there will be no incentive for buyers' to corrupt their bids. To provide a more formal definition for the revenue of the clairvoyant seller as well as "optimality" in contextual reserve prices, we rely on the following proposition that characterizes the seller's conditional expected revenue when buyers bid truthfully.

**Proposition 3.3.1** (Seller's Revenue with Truthful Buyers)**.** *Consider the case of $N \geq 2$ buyers who bid their true valuations, i.e., $v_{i,t} = b_{i,t}$ for any $i \in [N]$ and $t \in [T]$. Conditioned on the reserve price $r_t$ and the current context $x_t \in \mathcal{X}$, the seller's single period expected revenue in Equation (3.2) is*

$$\int_{-\infty}^{\infty} z dF^-(z) + \langle \beta, x_t \rangle + \int_0^{r_t} F^-(z - \langle \beta, x_t \rangle) dz \\ - r_t \left( F^+(r_t - \langle \beta, x_t \rangle) \right) , \tag{3.3}$$

*where for any $z \in \mathbb{R}$, $F^-(z) := NF^{N-1}(z) - (N-1)F^N(z)$ and $F^+(z) := F^N(z)$.*

The proof for this proposition is detailed in Appendix B.1.1. In Proposition 3.3.1, $F^+(\cdot)$ and $F^-(\cdot)$ are the cumulative distribution functions of $\epsilon_t^+ := v_t^+ - \langle \beta, x_t \rangle$ and $\epsilon_t^- := v_t^- - \langle \beta, x_t \rangle$ respectively, where $v_t^+$ and $v_t^-$ are the highest and second highest valuations in period $t \in [T]$.

In light of Proposition 3.3.1, we define the benchmark policy of the clairvoyant seller as followed,

**Definition 3.3.1** (Benchmark Policy)**.** *The benchmark policy knows the mean vector $\beta$ and noise distribution $F$, and sets the reserve price for a context vector $x \in \mathcal{X}$ as*

$$r^\star(x) \\ = \arg\max_{y \geq 0} \int_0^y F^-(z - \langle \beta, x \rangle) dz - y \left( F^+(y - \langle \beta, x \rangle) \right) . \tag{3.4}$$

*Therefore, the benchmark reserve price in period $t$, denoted by $r_t^\star$, is $r^\star(x_t)$, and the* *corresponding optimal revenue, denoted by $REV_t^\star$, is equal to*

$$\int_{-\infty}^{\infty} z dF^-(z) + \langle \beta, x_t \rangle + \int_0^{r^\star(x_t)} F^-(z - \langle \beta, x_t \rangle) dz$$
$$- r^\star(x_t) \left( F^+(r^\star(x_t) - \langle \beta, x_t \rangle) \right) .$$

**Remark 3.3.1.** *When distribution $F$ satisfies the MHR assumption, the objective function of the optimization problem in Equation (3.4) is unimodal in the decision variable $y$, and according to [57], $r^\star(x)$ can be simplified as follows: $r^\star(x) = \arg\max_{y \geq 0} y(1 - F(y - \langle \beta, x \rangle))$. In words, the MHR assumption decouples the reserve price optimization problem for multiple agents to the much simpler monopolistic pricing for each individual agent.*

We observe this benchmark provides an optimal mapping from the feature vector $x_t$ to reserve price $r^\star(x_t)$, which remains unchanged over time as the mean vector $\beta$ and noise distribution $F$ are time-invariant. This echoes our earlier point that pricing is challenging in our contextual setting since we would need to approximate or learn the optimal mapping $r^\star(\cdot)$, whereas in non-contextual environments it is sufficient to learn a single optimal reserve price value.

We now proceed to define the regret of a policy $\pi$ (possibly random) when the regret is measured against the benchmark policy. Suppose that in any period $t$, policy $\pi$ selects reserve price $r_t^\pi$. Then, the regret of policy $\pi$ in period $t$ and its cumulative $T$-period regret are defined as:

$$\mathsf{Regret}^\pi(T) = \sum_{t \in [T]} \mathbb{E}\left[ \mathrm{REV}_t^\star - \mathsf{rev}_t(r_t^\pi) \right] , \tag{3.5}$$

where the optimal revenue $\mathrm{REV}_t^\star$ is given in Definition 3.3.1, and the expectation is taken with respect to the context distribution $\mathcal{D}$ as well as the possible randomness in the actual reserve price $r_t^\pi$. Our goal is to design a policy that obtains a low regret for any $\beta$, $F$, and context distribution $\mathcal{D}$.

## 3.4 The NPAC-S Policy

In this section, we first propose a policy called *Non-Parametric Contextual Policy against Strategic Buyers* (NPAC-S) to maximize seller's expected revenue in our strategic setting. Then, we provide insights into how our design in NPAC-S makes the policy robust to buyer strategic behavior, and in turn allows the policy to learn the mean vector $\beta$ and noise distribution $F$ efficiently. Finally, we present theoretical regret guarantees for NPAC-S against the clairvoyant benchmark described in Definition 3.3.1 that sets the optimal contextual reserve price defined in Equation (3.4).

**The NPAC-S policy.** The detailed NPAC-S policy is shown in Algorithm 3, and consists of three main components. *(i) Phased Structure:* NPAC-S partitions $T$ into consecutive phases, where each phase $\ell \geq 1$, denoted as $E_\ell$, has length $T^{1-2^{-\ell}}$. This implies $|E_1| = \sqrt{T}$ and $|E_\ell|/\sqrt{|E_{\ell-1}|} = \sqrt{T}$. Here, we can establish that the total number of phases can be upper bounded by $\lceil \log_2(\log_2(T)) \rceil + 1$. *(ii) Estimation for $\beta$, $F^-$ and $F^+$:* At the end of each phase, NPAC-S uses the submitted bids from the pervious phase and employs Ordinary Least Squares (OLS) and empirical distributions to estimate the mean vector $\beta$ as well as $F$, respectively. *(iii) Random isolation:* NPAC-S incorporates random isolation periods in which a single buyer is chosen at random, and the item is auctioned to this isolated buyer (i.e. the seller only considers the bid of the isolated buyer and ignores bid from other buyers).[9] Note that when a buyer $i$ is isolated, the buyer wins the item if and only if his bid is greater than the reserve price, and pays the reserve price if he wins. Here, the seller's pricing policy is announced to all buyers (at $t = 0$) so that buyers examine the policy and have the freedom to adopt any bidding strategy to maximize their long term discounted utility.

**Remark 3.4.1.** *Here, we comment on how one can solve the reserve price optimization problem in Equation (3.7). The key observation is that for any period $t$, $\widehat{F}_\ell(\cdot)$ is a step function with jumps at points in the finite set $\mathcal{C}_\ell := \{b_{i,\tau} - \langle \widehat{\beta}_\ell, x_\tau \rangle\}_{i\in[N],\tau\in E_{\ell-1}}$.*

[9]The seller discloses her commitment to the random isolation protocol to all buyers at $t = 0$, and it is not necessary for the seller to reveal, during an isolation period, which buyer is being isolated.

[10]For a matrix $A$, $A^\dagger$ represents its pseudo inverse, so if $A$ is invertible, we have $A^\dagger = A^{-1}$. In Lemma B.2.1 of Appendix B.2.1, we show that with high probability $\sum_\tau x_\tau x_\tau^\top$ is positive definite, and hence invertible.

**Algorithm 3** Non-Parametric Contextual Policy against Strategic Buyers (NPAC-S)

1: Initialize $\widehat{\beta}_1 = 0$, and $\widehat{F}_1^-(z) = \widehat{F}_1^+(z) = 0$ for $\forall z \in \mathbb{R}$.
2: **for** phase $\ell \geq 1$ **do**
3:     **for** $t \in E_\ell$ **do**
4:         **Isolation**: With probability $1/|E_\ell|$, choose one buyer uniformly at random and offer price

$$r_t^u \sim \text{Uniform}(0, v_{\max})\,. \tag{3.6}$$

5:         **No Isolation**: With probability $1 - 1/|E_\ell|$, set reserve price for all buyers as

$$\widehat{r}_t = \arg\max_{y \in [0, v_{\max}]} \int_0^y \widehat{F}_\ell^-(z - \langle \widehat{\beta}_\ell, x_t \rangle) dz - y \cdot \widehat{F}_\ell^+(y - \langle \widehat{\beta}_\ell, x_t \rangle)\,. \tag{3.7}$$

6:         **Observe all bids** $\{b_{i,t}\}_{i \in [N]}$
7:     **end for**
8:     **Update estimate of the mean vector** $\beta$**:** [10]

$$\widehat{\beta}_{\ell+1} = (\sum_{\tau \in E_\ell} x_\tau x_\tau^\top)^\dagger \cdot (\sum_{\tau \in E_\ell} x_\tau \bar{b}_\tau)\,, \tag{3.8}$$

where $\bar{b}_\tau = \frac{1}{N} \sum_{i \in [N]} b_{i,\tau}$.
9:     **Update the estimate of** $F^+$ **and** $F^-$**:**

$$\begin{aligned} \widehat{F}_{\ell+1}^-(z) &= N \widehat{F}_{\ell+1}^{N-1}(z) - (N-1) \widehat{F}_{\ell+1}^N(z) \\ \widehat{F}_{\ell+1}^+(z) &= \widehat{F}_{\ell+1}^N(z)\,. \end{aligned} \tag{3.9}$$

where $\widehat{F}_{\ell+1}(z)$ is defined as

$$\widehat{F}_{\ell+1}(z) = \frac{1}{N|E_\ell|} \sum_{\tau \in E_\ell} \sum_{i \in [N]} \mathbb{I}(b_{i,\tau} - \langle \widehat{\beta}_{\ell+1}, x_\tau \rangle \leq z), \tag{3.10}$$

10: **end for**

*This implies that in order to solve for $r_t$ in Equation (3.7), it suffices to conduct a grid search for $\forall y \in \mathcal{C}_\ell$. More specifically, we let $\{z^{(0)}, z^{(1)}, \ldots z^{(M)}\}$ be the ordered list (in increasing order) of all elements in $\mathcal{C}_\ell \cup \{0\}$, where $z^{(0)} := 0$ and $M := |\mathcal{C}_\ell|$ (here, we assumed that $0 \notin \mathcal{C}_\ell$ without loss of generality). Hence, $r_t$ is equal to*

$$\begin{aligned} \arg\max_{m \in [M]} \sum_{j=1}^m & \widehat{F}_\ell^-(z^{(j)} - \langle \widehat{\beta}_\ell, x_t \rangle) \cdot (z^{(j)} - z^{(j-1)}) \\ & - z^{(m)} \widehat{F}_\ell^+(z^{(m)} - \langle \widehat{\beta}_\ell, x_t \rangle)\,. \end{aligned}$$

*This shows that the complexity to solve Equation (3.7) is $\mathcal{O}(M^2)$. More detailed discussions and efficient algorithms regarding related problems can be found in [90].*

**Motivation for design of NPAC-S.** Here we provide some insights into the design of the NPAC-S policy, particularly the phased structure and the incorporation of random isolation periods.

Due to the phased structure of the algorithm, our estimates for $\beta$, $F^-$, and $F^+$ only depend on the bids and contextual features in the previous phase. Thus, corrupted bids submitted by buyers in past periods will have no impact on future estimates as well as pricing decisions. One can think of this as erasing all memory prior to the previous phase and restarting the algorithm, which can potentially reduce buyers' manipulating power on our estimates and reserve prices.

We now discuss the impact of having isolation periods. As all buyers are aware of the randomized isolation protocol, the presence of isolation periods restricts buyers from significantly corrupting their bids too often as by doing so they may suffer a substantial utility loss when they are isolated. To illustrate this point with an example, compare the following scenarios: (i) if there are no isolation periods, a buyer having the lowest valuation among all buyers may submit a bid by adding large corruption, but still ending up not being the second highest or highest bidder. Assuming that other buyers bid truthfully, such a scenario will not lead to any changes in utility of any buyer, but introduces a large outlier to the set of data points used in our estimations. In words, when no isolation occurs, buyers may be able to distort the seller's learning process without facing unfavorable consequences; (ii) during an isolation period when a buyer is isolated, corrupting her bid may perhaps result in significant utility loss, e.g., losing the item by underbidding when her true valuation is greater than the reserve price, or winning the item by overbidding when her true valuation is less than the reserve price. Therefore, randomized isolation incentivizes utility-maximizing buyers to reduce the frequency of corrupting their bids. Mathematically, we characterize this statement in the following Lemma 3.4.1.

**Lemma 3.4.1** (Bounding number of significantly corrupted bids)**.** *For $i \in [N]$ and* *phase $\ell \geq 1$ define*

$$\begin{aligned} \mathcal{S}_{i,\ell} &:= \left\{ t \in E_\ell : |a_{i,t}| \geq \frac{1}{|E_\ell|} \right\} \\ L_\ell &:= \log\left(v_{\max}^2 N |E_\ell|^4 - 1\right) / \log(1/\eta)\,, \end{aligned} \tag{3.11}$$

*where $\mathcal{S}_{i,\ell}$ is the set of all periods in phase $E_\ell$ during which buyer $i$ significantly corrupts his bids. Then, we have $\mathbb{P}\left(|\mathcal{S}_{i,\ell}| > L_\ell\right) \leq 1/|E_\ell|$.*

The proof of this lemma is shown in Appendix B.2.2.

**Bounding the regret of NPAC-S.** Here, we first present the regret of NPAC-S. Then we introduce several key results that are crucial to proving the regret bound of NPAC-S and also comment on how they resolve challenges that arise due to buyers' strategic behavior.

**Theorem 3.4.2** (Regret of NPAC-S Policy)**.** *Suppose that the length of the horizon $T \geq \max\{\left(\frac{8x_{\max}^2}{\lambda_0^2}\right)^4, 9\}$ where $\lambda_0^2$ is the minimum eigenvalue of covariance matrix $\Sigma$. Then, in the strategic setting, the $T$-period regret of the NPAC-S policy is in the order of $\mathcal{O}\Big(c_f\sqrt{dN^3\log(T)} \cdot \log\left(\log(T)\right)\left(\sqrt{T} + \frac{\sqrt{N^3\log(T)}T^{\frac{1}{4}}}{\log(1/\eta)}\right)\Big)$, where regret is computed against the benchmark policy in Definition 3.3.1 that knows the mean vector $\beta$ and noise distribution $F$. Here, recall $c_f = \sup_{z\in[-\epsilon_{\max},\epsilon_{\max}]} f(z) > 0$ where $f$ is the the pdf of $F$.*

**Remark 3.4.2.** *The proof of this theorem is presented in Appendix B.2.1. In the regret of NPAC-S, the factor $1/\log(1/\eta)$ serves as a worse case guarantee for the amount of corruption that buyers' can apply to their bids throughout the entire horizon $T$. As buyers get less patient, i.e., as $\eta$ decreases, buyers are less willing to forgo current utility in the current period. Thus, in the presence of randomized isolation periods, impatient buyers are less likely to significantly corrupt bids, which translates into lower regret. The $\log(\log(T))$ factor corresponds to the information loss due to the policy's phased structure, which "restarts" the algorithm at the beginning of each of $\mathcal{O}\left(\log(\log(T))\right)$ phases and relies only on the information of the previous phase.*

The regret of NPAC-S can be decomposed into two parts: (i) the estimation errors in $\beta$, $F^-$ and $F^+$, which result in the posted reserve price $r_t$ deviating from the optimal reserve price $r_t^\star$, and hence incur a revenue loss compared to the clairvoyant benchmark; and (ii) the revenue loss due to *allocation mismatch* in the auction outcome because of buyers' strategic bidding behaviour. Here, allocation mismatch refers to the phenomenon where a bidder would have won (lost) the auctioned item had she bid truthfully, but instead lost (won) the item as she submitted a corrupted bid in reality.

We first comment on several challenges with respect to bounding the estimation errors in $\beta$, $F^-$ and $F^+$. First, the OLS estimator and empirical distributions to estimate the mean vector and distributions $F^-$ and $F^+$, respectively are extremely vulnerable to corrupted data (outliers), and hence standard high probability bounds are invalid for our setting. Additionally, there exists a complication in terms of bounding the estimation errors in $F^-$ and $F^+$ because estimation errors for $\beta$ will further propagate into the estimation errors in $F$ and consequently impacting the estimates for $F^-$ and $F^+$. To illustrate this point, consider the ideal scenario where all bids are truthful (i.e. $v_{i,t} = b_{i,t}$ for all $i \in [N]$ and $t \in [T]$). Even in this scenario, the terms $v_{i,\tau} - \langle \widehat{\beta}_\ell, x_\tau \rangle$ in the expressions for $\widehat{F}_\ell(\cdot)$ are not realizations of $\epsilon_{i,\tau}$ due to estimation errors in the mean vector $\widehat{\beta}_\ell$. Hence, the estimate $\widehat{F}_\ell(\cdot)$ evaluated at any point $z \in \mathbb{R}$ is biased, i.e. $\mathbb{E}[\widehat{F}_\ell(z - \langle \widehat{\beta}_\ell, x_t \rangle)] \neq F(z - \langle \widehat{\beta}_{\ell+1}, x_t \rangle)$. Furthermore, the estimates $\widehat{F}_\ell^+(\cdot)$ and $\widehat{F}_\ell^-(\cdot)$ are evaluated at points which may be random variables since $\widehat{\beta}_\ell$ is a random variable that depends on the history of the previous phase.

In light of such challenges in bounding estimation errors, as one of our main contributions, the following Lemma 3.4.3 provides good estimation error guarantees for $\beta$, $F^-$ and $F^+$ in the presence of corrupted bids and the aforementioned error propagation phenomena.

**Lemma 3.4.3** (Bounding estimation errors in $\beta$, $F^-$ and $F^+$)**.** *For any phase $E_\ell$, with probability at least $1 - \Theta(1/|E_\ell|)$, the following events hold: (i) $\|\widehat{\beta}_{\ell+1} - \beta\|_1 = \mathcal{O}(\frac{1}{\sqrt{|E_\ell|}} + \frac{\log(|E_\ell|)}{\log(1/\eta)|E_\ell|})$; (ii) for any $z \in \mathbb{R}$, $|\widehat{F}_{\ell+1}^-(z) - F^-(z)| = \mathcal{O}(\frac{N^2}{\sqrt{|E_\ell|}} + \frac{N^2 \log(|E_\ell|)}{\log(1/\eta)|E_\ell|})$ and $|\widehat{F}_{\ell+1}^+(z) - F^+(z)| = \mathcal{O}\big(\frac{N}{\sqrt{|E_\ell|}} + \frac{N \log(|E_\ell|)}{\log(1/\eta)|E_\ell|}\big)$. Here, recall the discount factor $\eta \in (0,1)$.*

We refer readers to Lemma B.2.1 and Lemma B.2.2 in Appendix B.2.4 for more detailed statements on our high probability bounds regarding estimation errors in $\beta$, $F^-$ and $F^+$.

In addition to inaccurate estimates for $\beta$, $F^-$ and $F^+$, the allocation mismatch phenomenon due to strategic bidding also contributes to the regret of NPAC-S. For example, suppose that the highest valuation is greater than the reserve price. In that case, if buyers were truthful, the item would be allocated and the seller would gain positive revenue. Now, if buyers shade their bids, the auctioned item may not get allocated, resulting in zero revenue for the seller. In the following Lemma 3.4.4, we show that the number of allocation mismatch periods for each buyer is bounded with high probability.

**Lemma 3.4.4** (Bounding allocation mismatch periods)**.** *Define the following two sets of time periods:*

$$\begin{aligned} &\mathcal{B}^s_{i,\ell} = \{t \in E_\ell : v_{i,t} \geq D_t \ , \ b_{i,t} \leq D_t\} \quad \textit{and} \\ &\mathcal{B}^o_{i,\ell} = \{t \in E_\ell : v_{i,t} \leq D_t \ , \ b_{i,t} \geq D_t\} \\ &\textit{where } D_t = \max\{b^+_{-i,t}, \widehat{r}_t\}\,. \end{aligned} \tag{3.12}$$

*Here, $b^+_{-i,t}$ is the highest among all bids excluding that submitted by buyer $i$, and $\widehat{r}_t$ is the reserve price offered to all buyers if no isolation occurs (defined in Equation (3.7)). Then, for $\mathcal{B}_{i,\ell} := \mathcal{B}^s_{i,\ell} \cup \mathcal{B}^o_{i,\ell}$, we have $\mathbb{P}\left(|\mathcal{B}_{i,\ell}| \leq 2L_\ell + 4c_f + 8\log(|E_\ell|)\right) \ \geq \ 1 - 4/|E_\ell|$, and $L_\ell$ is defined in Equation (3.11). Here, the probability is taken with respect to the randomness in $\{(x_\tau, \epsilon_{i,\tau}, a_{i,\tau})\}_{\tau \in E_\ell, i \in [N]}$.*

Note that $\mathcal{B}^s_{i,\ell}$ represents the set of all periods in phase $\ell$ during which buyer $i$ should have won the item if she bid truthfully, but in reality lost due to shading her bid (i.e. allocation mismatch due to shading), while similarly $\mathcal{B}^o_{i,\ell}$ is the periods of allocation mismatch due to overbidding. Therefore, $\mathcal{B}_{i,\ell} := \mathcal{B}^s_{i,\ell} \cup \mathcal{B}^o_{i,\ell}$ can be interpreted as the set of all periods in phase $\ell$ when an allocation mismatch occurs for buyer $i$. The detailed proof is provided in Appendix B.2.3.

**NPAC-S against Truthful Buyers.** Here, we make a remark that in a hypothetical world where buyers are truthful (i.e. $v_{i,t} = b_{i,t}$ or equivalently the degree of corruption $a_{i,t} = 0$ for all $i \in [N]$, $t \in [T]$), our proposed NPAC-S policy achieves a regret of $\mathcal{O}(c_f \sqrt{dN^3 T \log(T)} \cdot \log\log(T))$ compared to the clairvoyant benchmark policy in Definition 3.3.1. Intuitively, this is easy to see because the set of all periods in phase $E_\ell$ during which a buyer $i$ significantly corrupts his bids, namely $\mathcal{S}_{i,\ell}$ defined in Lemma 3.4.1, will be empty. As a result, there will be no allocation mismatch periods, and the $1/\log(1/\eta)$ terms in the estimation errors in $\beta$, $F^-$, $F^+$ will vanish (see Lemma 3.4.3). The proof for the regret bounds of NPAC-S against truthful buyers would thus be a simplification to the proof of Theorem 3.4.2, and hence will be omitted.

## 3.5 Numerical studies

Here, we present numerical simulations to compare the performance of NPAC-S with several baseline seller policies. In particular, consider the following baseline policies: (i) NAIVE which always sets a 0 reserve price; (ii) CONTHEDGE which runs an independent version of the HEDGE algorithm for every distinct context vector (see an introduction of HEDGE for the adversarial multi-arm bandit problem in [10]). The "arms" of HEDGE correspond to potential reserve price options. Note that HEDGE is a special case of the well-known EXP3 algorithm which is a simple off-the-shelf algorithm that not only has good theoretical guarantees, but has also been applied (or its variations/generalizations have been adopted) in many areas in online advertising (see e.g. [114, 17, 65]). (iii) HO-SERP, which sets personalized reserve prices for each buyer using "rolling window" estimates of $\beta$ and $F$ w.r.t other buyers' submitted bids (see [75]). Here we consider HO-SERP as a baseline because among all seller algorithms in related works that study pricing in a contextual, stochastic, and strategic buyer setting similar to ours (see Table 3.1), HO-SERP achieves nearly the best theoretical performance. Note HO-SERP requires the noise distribution to be MHR.

To model buyers' strategic behavior, instead of restricting buyers to bid according to a specific strategy to maximizes long-term discounted utility, we instead mimic the outcome of some general class of such strategies (parameterized by $\eta$) via randomly selecting periods over the entire horizon and have buyers significantly corrupt bids in these periods. We will refer to these randomly selected periods as *corruption periods*. When this randomization procedure is repeated over many trials, we believe the average bidding outcome would serve as a relatively accurate approximation to the outcomes of a general class of strategies for utility-discounting buyers. Furthermore, inspired by Lemma 3.4.1 which suggests that the number of periods when a buyer significantly corrupts her bid is bounded, we let the selected number of corruption periods be $L_\ell$ defined in Equation (3.11) . Note that $L_\ell$ is increasing in $\eta$ and represents the fact that more patient buyers (i.e. larger $\eta$) value long term utility more and hence would be willing to corrupt bids more frequently with the aim of achieving higher future utility.

Out detailed experimental setup is as followed. We consider a horizon of length $T = 5,000$, $N = 2$ buyers, context vectors of dimension $d = 4$, $v_{\max} = 10$ and $v_{\min} = 0$. For each $\eta \in \{0.2, 0.4, 0.6, 0.8\}$, repeat the following procedure for $n = 50$ trials, each including $T$ periods:

For each phase $E_\ell$ ($\ell \geq 1$), [11] sample $L_\ell$ corruption periods uniformly at random. Then, regarding buyer's valuations, we generate $\beta \in [0,1]^d$, where each entry is sampled independently according to a uniform distribution on [0,1], i.e., $U(0,1)$. We further normalize $\beta$ with the sum of all entries so that $\|\beta\|_2 = 1$. We then generate 10 distinct contexts vectors $\mathcal{X} = \{X^j\}_{j\in[10]}$, where each entry for any distinct context vector is sampled independently from $U\left(\frac{v_{\max}}{3}, \frac{2v_{\max}}{3}\right)$. Then, for every period $t \in [T]$, sample $x_t$ uniformly at random from $\mathcal{X}$, and sample $\epsilon_{i,t}$ for all $i \in [N]$ independently from $U\left(-\frac{v_{\max}}{3}, \frac{v_{\max}}{3}\right)$. Note that our construction guarantees $v_{i,t} = \langle \beta, x_t \rangle + \epsilon_{i,t} \in [v_{\min}, v_{\max}]$, and the noise distribution is uniform which satisfies the MHR assumption (so the application of the HO-SERP is valid). If $t$ is a corruption period, we let buyers submit a bid of value 0 to model the behavior of significant bid-shading; otherwise, we let buyers bid their true valuations.[12]

[11]For fixed $T$, since length of phase $\ell \geq 1$ is $T^{1-2^{-\ell}}$, in our case when $T = 5,000$ we have 4 phases whose phase lengths are $70, 594, 1724, 2612$, respectively, where the last phase is truncated.

[12]We remark that our numerical experiments focus on buyers' bid-shading behavior. This is mainly because empirical studies found that shading is prevalent in repeated auctions on modern online advertising platforms and theoretical works have demonstrated various versions of bid-shading

For comprehensiveness, we also consider the truthful setting by repeating the above valuation generation procedure for another $n = 50$ trials and have buyers always submit their true valuations. Finally, for each of the aforementioned trials, we run the NPAC-S as well as other baseline algorithms independently and simply record the realized revenue of each algorithm across all repeated auctions.

We report the average per-period revenue loss compared to the benchmark policy (Definition 3.3.1) for each algorithm in Figure 3-1.

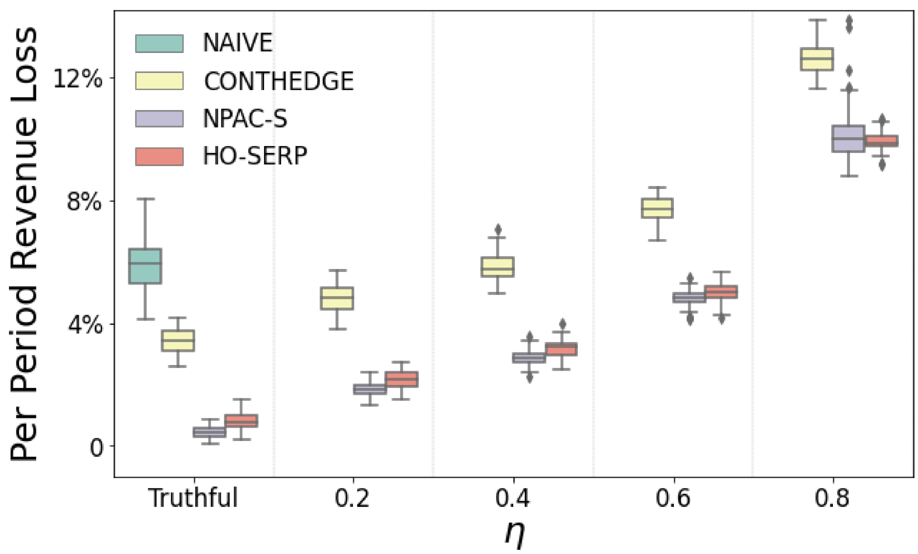


Figure 3-1: **Performance comparison with baselines.** This figure displays the average per-period revenue loss compared to the benchmark policy (Definition 3.3.1). Each box plot corresponds to $n = 50$ trials. NAIVE is only run for the truthful setting because buyers will have no incentive to bid untruthfully when there is no reserve price. CONTHEDGE is run with"arms" $\{0, 0.5, 1, \ldots 10\}$, where each arm corresponds to a reserve price option.

We observe that our proposed NPAC-S algorithm not only outperforms CONTHEDGE in all settings consistently by a $3\% \sim 4\%$ and NAIVE in the truthful setting by $6\% \sim 7\%$, NPAC-S also generally yields more stable outcomes as measured by the standard deviation of per-period revenue loss across $n$ trials. Compared to HO-SERP, NPAC-S slightly outperforms HO-SERP in the truthful setting and for $\eta = 0.2$, $0.4$, $0.6$. Nevertheless, we point out that our experimental setting inherently

strategies can help buyers achieve near-optimal performances in a variety of practical settings, such as buyers being constrained by a limited budget or target return on investment (see e.g. [113, 61, 17]).

favors HO-SERP since performance guarantees of this algorithm relies on the noise distribution being MHR, which is the case for our uniform noise. Moreover, the comparison with HO-SERP also demonstrates the advantages of NPAC-S from a practical viewpoint, since NPAC-S, unlike HO-SERP, sets a single reserve price for all buyers and still matches or improves upon the performance of HO-SERP.

## 3.6 Additional discussions and future research

### Heterogeneous discount rates

We remark that our NPAC-S algorithm is agnostic to buyer discount rates as well as adopted bidding strategies w.r.t. such rates. It is not difficult to see that in the case where buyers have heterogeneous discount rates, Lemma 3.4.1 which bounds the number of periods during which buyers significantly corrupted bids still holds since our proof only concerns each individual buyer. This implies that our NPAC-S policy can also handle heterogeneous buyer discount rates and result in a regret similar to that in the current homogeneous setup.

### Future research

Our current model in Section 2.2 studies homogeneous buyers who share an identical preference vector $\beta$ and noise distribution $F$. One immediate future research direction is to handle a heterogeneous buyer setup where these model parameters differ across buyers. This setting is challenging because the optimal contextual reserve price no longer admits a closed form as illustrated in Eq. (3.4). Another research question of significant practical and theoretical importance is contextual dynamic reserve price optimization under repeated first price auctions (FPA) facing strategic buyers. In an FPA setting, bidding truthfully is no longer a weakly dominant strategy for buyers, and therefore making it difficult for the seller to recover buyer preferences from submitted bids. It would be interesting to see if techniques such as enforcing a phased structure and including randomized isolation periods as in NPAC-S can still limit buyers' strategic behavior.

# Chapter 4

# Learning to price against financially constrained buyers

*This chapter is based on [55], which is joint work with Negin Golrezaei, Patrick Jaillet, and Vahab Mirrokni.*

## 4.1 Introduction

In the previous Chapter 3, we considered an ad platform's revenue maximization problem through designing reserve-price based ad auctions, while perceiving the advertiser-platform interaction and automated procurement algorithm components in the autobidding ecosystem together as buyers who dynamically make bidding decisions according to certain algorithms. In this chapter, we investigate a similar problem of repeatedly selling ads to algorithmic buyers, but instead of strategic buyers as in Chapter 3, here we consider buyers being subject to real-world financial constraints.

In practice, to efficiently utilize limited monetary resources that are allocated to a certain campaign, advertisers' strategies in the procurement process are typically subject to financial constraints, which generally include budget and *return-on-investment* (ROI) constraints. Budget constraints primarily reflect advertisers' monetary limits due to organizational planning, whereas ROI constraints enforces the desired performance/return on the amount of capital spent [76, 61, 13]. The presence of such financial constraints, along with the increasing availability of real time data, motivates buyers' deployment of complex algorithms to procure impressions. Such financial constraint and algorithm driven buyer behavior introduces significant challenges to sellers' design of selling mechanisms, primarily due to the fact that buyer algorithms adapt quickly and constantly to data generated by buyer-seller interactions, and also sellers' lack of information on buyers' model primitives such as target ROI, budget, buyer algorithm, etc. In this chapter, we address the following question:

*From the perspective of a seller (e.g. ad platform), what is an optimal selling strategy against a buyer who adopts value-maximizing algorithms under both budget and ROI constraints?*

We study the setting where a seller repeatedly sells items to a single budget and ROI constrained buyer through a posted price mechanism. This single-buyer setup is primarily motivated by ad platforms' targeting practices that enable advertisers to target users who may be more interested in their ads, as such practices along with advertisers' heterogeneous targeting criteria lead to a very small number of advertisers/buyers per ad impression, justifying our single-buyer setup. Throughout the repeated mechanism, the seller posts a price for an impression during each period, and the buyer decides on whether to accept and pay the posted price for the sold impression. Our key focus lies in the practical two-sided learning setup where buyers adopt learning algorithms under both budget and ROI constraints, and the seller sets prices algorithmically based on past transactions. The key challenge for the seller's problem of interest is two-fold: the seller does not know the buyer parameters such as target ROI, budget or algorithm, and buyer actions constantly adapt to the past buyer-seller algorithmic interactions. The goal of this chapter is to design a revenue-maximizing seller pricing strategy against algorthmic and financially constrained buyers in such a limited information setting.

The main contribution of this chapter is that we propose a simple seller algorithm that does not require explicitly learning buyer's parameters nor reverse engineering the buyer's learning algorithms. We show that our algorithm is feasible in achieving high revenue under limited information by exploiting a salient property of the seller revenue function against financially-constrained buyers. In particular, we summarize our contributions as followed:

**Main contributions.** We first characterize the seller revenue function against a clairvoyant budget and ROI constrained buyer who always best responds to posted prices. To begin with, we show that the buyer's best response to a posted price is a "threshold strategy", i.e. the buyer accepts the sold item if her valuation exceeds a certain threshold that depends on the posted price. With this characterization of buyer best response, we show that the seller revenue function against a best-responding clairvoyant buyer admits a salient "bell-shaped" structure: as the seller increases prices, the corresponding per-period seller revenue first monotonically increases and decreases. We argue that such a structure is exploitable by the seller to extract revenue even without knowing buyer model primitives such as value distribution, budget rate, and target ROI.

We exploit this bell-shaped structure and design an episodic binary search seller pricing algorithm. In each episode, the algorithm sets a single price, and then moves on to the next episode with an updated price based on a binary search procedure w.r.t. the realized revenue of previous prices. We also characterize general buyer-algorithm adaptiveness properties that allow buyers to adapt quickly to prices in seller episodes, and present regret analyses against buyer algorithms that are adaptive to seller prices in the sense of our defined adaptiveness properties. Moreover, we argue that seller regret of our proposed algorithm is driven by the agent (i.e. seller or buyer) who incurs a larger loss in terms of learning error.

Finally, we analyze example buyer algorithms which satisfy the aforementioned adaptiveness properties and aim to maximize total value under both budget and ROI constraints. In particular, we consider clairvoyant buyers who best respond in each period, as well as buyers who make decisions based on machine-learned advice that take the form of value distribution estimates. For each of these buyers, we show that both buyer and seller regret are sublinear.

## Related works

**Mechanism design for budget and ROI constrained buyers.** One relevant line of research addresses the mechanism design problem for budget or ROI constrained buyers. As one of the pioneering works regarding mechanism for financially constrained buyers, [79] derives the optimal mechanism for symmetric buyers and public budget information. On the contrary, a more recent paper [95] studies the general multidimensional mechanism design setting against buyers with private budgets. Regarding ROI constrained buyers, [61] shows that the optimal mechanism for symmetric ROI-constrained buyers is either second-price auctions with reduced reserve prices or subsidized second-price auctions. The work also derives an optimal mechanism for asymmetric ROI buyers. There is also a wide range of work that study dynamic mechanism design for budget constrained buyers, and we refer the reader to the survey [20] and references therein. There have also been recent developments for designing auctions in a setup called *autobidding*, where advertisers simultaneously participate in parallel auction to maximize total value while subject to a coupled

ROI constraint across all auctions (see e.g. [4, 37, 11, 33]). All aforementioned works focus on the static mechanism design problem, whereas in this chapter we address the topic of designing repeated posted price mechanisms to sell to both budget and ROI constrained buyers.

**Selling to strategic or learning buyers.** [77] studies the scenario where the seller sells items through a repeated posted price mechanism to a single truthful buyer who accepts the price if her valuation is greater than the offered price. The work presents optimal algorithms in the settings where the buyer's valuations are fixed, stochastic and adversarial, respectively. [6] also concerns selling through a posted price mechanism, but to a strategic buyer who may choose not to accept a price bellow her valuation (or accept a price above her valuation). The work presents learning algorithms in both the fixed valuation and stochastic valuation settings under the assumption that discount their utilities over time. Other related works include [59] which studies the dynamic pricing problem for repeated contextual second price auctions facing multiple strategic buyers. The work proposes learning algorithms that are robust to buyers' strategic behavior under various seller information structures and provides corresponding performance guarantees. [54] relaxes several assumptions for one of the settings in [59], and presents an algorithm with improved performance guarantees. Finally, [15] considers the dynamic mechanism design problem against strategic buyers, and further identifies a class of problems in which the optimal mechanism is to simply repeat some static mechanism over time. The closest previous work to this chapter is [25], where it studies the pricing problem against a single unconstrained quasi-linear buyer who adopts a certain class of learning algorithms, which they refer to as "mean-based" algorithms (e.g. Follow the Perturbed Leader algorithm and EXP3), the seller can extract the buyer's entire surplus; see [38] for an extensions. We remark that all works discussed here do not consider constrained buyers, and therefore this chapter distinguishes itself by studying the pricing problem against buyers with both budget and ROI constraints, which further allows us to characterize special structures of seller revenue (see Section 4.3).

## 4.2 Preliminaries

**Notation.** Let $\mathbb{R}_+$ be all non-negative real numbers, and $\mathbb{R}_{++}$ be all strictly positive real numbers. For integer $N \in \mathbb{N}$, denote $[N] = \{1, 2, \dots, N\}$ and $\Delta_N = \left\{\boldsymbol{p} \in [0,1]^N : \sum_{n\in[N]} p_n = 1\right\}$ be the $N$-dimensional probability simplex. Finally, denote $\|\cdot\|$ as the Euclidean norm.

**Model setup:** Consider a seller repeatedly selling items to a buyer over $T$ periods through a posted price mechanism: in each period $t$, the seller posts a price $d_t$ for the item to be sold, and the buyer makes a take it or leave it decision $z_t \in \{0,1\}$ based on her value $v_t$ of the item, where $z_t = 1$ when the buyer takes the item at price $d_t$, and 0 otherwise.

We assume the seller commits to a finite price set $\mathcal{D} = \{D_m\}_{m\in[M]}$ where $1 \geq D_1 > \cdots > D_M > 0$ from which she chooses the posted prices $\{d_t\}_{t\in[T]}$, and we assume the the buyer's valuations are drawn independently each period from a distribution over $\boldsymbol{g} = (g_1 \dots g_N) \in \Delta_N$ ($g_n \in \mathbb{R}_{++}$ for all $n \in [N]$) over a finite support $\mathcal{V} = \{V_n\}_{n\in[N]}$ where $1 \geq V_1 > \cdots > V_N > 0$ such that $\mathbb{P}(v_t = V_n) = g_n$ for any period $t \in [T]$.

**ROI and budget constrained buyers:** The buyer aims to maximize total acquired value over $T$ periods, while subject to an ROI constraint with the target ROI of $\gamma \geq 1$ and a budget constraint with budget rate $\rho \in (0,1)$.[1] Mathematically, using the shorthand notation $\boldsymbol{d}_{1:T}$ for the sequence of prices $\{d_t\}_{t\in[T]}$, the buyer's hindsight optimization problem can be written as followed

$$
\begin{aligned}
\mathrm{OPT}(\boldsymbol{d}_{1:T}) = \max\nolimits_{\boldsymbol{z}\in[0,1]^T} \quad & \mathbb{E}\left[\sum\nolimits_{t\in[T]} v_t z_t\right] \\
\text{s.t.} \quad & \mathbb{E}\left[\sum\nolimits_{t\in[T]} \left(v_t - \gamma d_t\right) z_t\right] \geq 0 \\
& \mathbb{E}\left[\sum\nolimits_{t\in[T]} d_t z_t\right] \leq \rho T\,.
\end{aligned}
\tag{4.1}
$$

We remark that both budget and ROI constraints are studied in expectation. Such "soft" constraints are useful in practice due to the fact that real-world advertisers

[1] Note that in the literature another common buyer objective is to optimize linear utility that takes the form $\sum_{t\in[T]}(v_t - \alpha d_t)z_t$ for some parameter $\alpha \geq 0$. We point out that all results in this paper can be extended easily to such linear objectives.

typically engage in many different online advertising campaigns, so it is reasonable to maintain these financial constraints in an average sense. We note that such soft financial constraints are also studied in mechanism design and online learning literature such as [107, 61].

We denote the optimal hindsight buyer decision sequence to Equation (4.1) as $\{z_t^*(\boldsymbol{d}_{1:T})\}_{t\in[T]}$. When all prices are equal, i.e. $d_t = d$ for all $t$, we use the shorthand notation $\mathrm{OPT}(d)$ and $\{z_t^*(d)\}_{t\in[T]}$. Note that optimal hindsight decisions $\{z_t^*(\boldsymbol{d}_{1:T})\}_{t\in[T]}$ may possibly be fractional, which can be implemented by randomization.

The buyer's target ROI $\gamma$ and budget rate $\rho$ are private to the buyer and unknown to the seller. Also, both the seller and the buyer do not know the valuation distribution $\boldsymbol{g}$.

**Seller's benchmark revenue and regret.**

The seller does not know the buyer's model primitives, namely the buyer's valuation distribution $\boldsymbol{g}$, target ROI $\gamma$ and budget rate $\rho$. Furthermore, the seller only observes the buyer's decision $z_t \in \{0, 1\}$, and does not observe buyer values. Under such information structure, we focus on non-anticipative seller pricing strategies that post prices based on historical data, i.e. in each period $t$, the decision $z_t$ can only depend on $\{(d_\tau, z_\tau)\}_{\tau\in[t-1]}$. We evaluate the performance of any sequence of pricing decision $\{d_t\}_{t\in[T]} \in \mathcal{D}^T$ by benchmarking its realized revenue, namely $\sum_{t\in[T]} d_t z_t$, to the maximum revenue that could have been obtained if (i) the seller had set a fixed price over all $T$ periods and (ii) the buyer makes optimal hindsight decisions given her ROI and budget constraints. Mathematically, assume the seller fixes price $d \in \mathcal{D}$ over all $T$ periods, and the buyer's optimal decisions are $\{z_t^*(d)\}_{t\in[T]}$. Then, the seller's benchmark revenue is $\max_{d\in\mathcal{D}} \mathbb{E}[d \sum_{t\in[T]} z_t^*(d)]$ and her regret can be defined as follows

$$\mathrm{Reg}_{\mathrm{sell}} = \max_{d\in\mathcal{D}} \mathbb{E}\left[d \sum_{t\in[T]} z_t^*(d)\right] - \sum_{t\in[T]} \mathbb{E}\left[d_t z_t\right], \tag{4.2}$$

where the expectation is taken w.r.t. $\{v_t\}_{t\in[T]}$ and randomness in the buyer's strategy (and thus randomness in $\{z_t^*(d)\}_{t\in[T]}$).

**Remark 4.2.1.** *The seller's regret resembles that of an $M$-arm multi-arm bandit (MAB) problem (see [82] for a detailed introduction), where we can view each price $d \in \mathcal{D}$ as an arm and $d \cdot z_t$ as the reward by pulling arm $m$. Nevertheless, we point out that our setting is more complex than the vanilla MAB setting as the seller's reward $d \cdot z_t$ for setting price $d$ during period $t$ not only depends on the seller algorithm which determines prices based on historical observations , but also the buyer's algorithm to optimize Equation (4.1).*

We point out that the benchmark revenue in the seller's regret of Equation (4.2) is strong, as it represents the maximum seller revenue when both the buyer and seller have complete information and act optimally, i.e. if the seller knows everything about the buyer, in each period she myopically posts a revenue-maximizing price under best buyer response.

Our goal is to develop a seller pricing algorithm to minimize regret when facing a buyer who optimizes Equation (4.1) via running some online learning algorithm (to be discussed in later sections).

## 4.3 Seller's revenue and regret

In this section, we present a reformulation for the seller's benchmark revenue in the seller's regret (Equation (4.2)), and then further characterize special structures of this reformulation which will later motivate the design of our pricing algorithm.

### 4.3.1 Reformulating the seller's benchmark revenue

Recall the seller's benchmark revenue in Equation (4.2) which depends on the buyer's best response decision sequence over the entire horizon $T$ under a fixed price. To present our reformulation of this benchmark, we first show that for any price $d$, although the buyer's hindsight optimal decisions $\{z_t^*(d)\}_{t\in[T]}$ may seemingly be interdependent across periods due to the coupling of budget and ROI constraints over the entire horizon, the optimal buyer decision in each period $t$ simply requires the buyer to myopically make a decision $z_t$ that maximizes single-period expected value under "single-period budget and ROI constraints", namely $\mathbb{E}\left[(v_t - \gamma d)\, z_t\right] \geq 0$ and $\mathbb{E}\left[d z_t\right] \leq \rho$.

Formally, consider the following myopic buyer optimization problem: for a given posted price $d$, let $\boldsymbol{x} \in [0,1]^N$ be some vector whose $n$th entry $x_n$ denotes the probability of accepting the price when the buyer's realized value is $V_n$. Then, the myopic buyer optimization problem can be written as Equation (4.3) whose optimal solution is shown in the following Lemma 4.3.1 (see proof in Appendix C.2.1).

$$
\begin{aligned}
U(d) = \max_{\boldsymbol{x} \in [0,1]^N} & \sum_{n \in [N]} g_n V_n x_n \\
\text{s.t.} & \sum_{n \in [N]} g_n \left(V_n - \gamma d\right) x_n \geq 0 \\
& d \sum_{n \in [N]} g_n x_n \leq \rho \,.
\end{aligned}
\tag{4.3}
$$

**Lemma 4.3.1.** *For any price $d$, the optimal solution to Equation (4.3) is unique, and takes the form $\boldsymbol{x}_d = (1, 1, \ldots q, 0, 0 \ldots 0) \in [0,1]^N$ for some $q \in (0,1]$.*

The special form of the optimal solution of Equation (4.3) suggests a buyer strategy that accepts all items when buyer value is beyond a certain threshold. We formalize such a strategy in the following definition.

**Definition 4.3.1** (Threshold strategy)**.** *For a given vector $\boldsymbol{x}$ that takes the form $\boldsymbol{x} = (1, 1, \ldots q, 0, 0 \ldots 0) \in [0,1]^N$ where $q \in (0,1]$ is the $n$th entry, we say a buyer adopts a threshold strategy w.r.t. $\boldsymbol{x}$ if, regardless of the posted price, she accepts the item when her value is $V_1 \ldots V_{n-1}$; accepts w.p. $q$ when her value is $V_n$; and rejects the item otherwise.*

As an example, for $N = 4$ and some vector $\boldsymbol{x} = (1, 1, 0.3, 0)$, the buyer adopts a threshold strategy w.r.t. $\boldsymbol{x}$ if she accepts the item when her value is $V_1$ or $V_2$; accepts w.p. 0.3 when her value is $V_3$, and rejects when her value is $V_4$.

With Lemma 4.3.1 and the notion of threshold strategies in Definition 4.3.1, we can formally define the buyer's best response to a given price $d$:

**Definition 4.3.2** (Buyer best response)**.** *We say a buyer best responds to a posted price $d$ if she adopts a threshold strategy w.r.t. $\boldsymbol{x}_d \in [0,1]^N$ which is the optimal solution to $U(d)$ (see Lemma 4.3.1).*

Note that in order for to best respond to a posted price, the buyer would need to know the value distribution $\boldsymbol{g}$.

Our main result for this subsection is illustrated in the following theorem, which states that buyer's hindsight optimal decision sequence $\{z_t^*(d)\}_{t\in[T]}$ for $\mathrm{OPT}(d)$ in Equation (4.1) simply requires the buyer to independently best respond to the posted price in each period.

**Proposition 4.3.2.** *Given a single price $d$ posted across all periods, the optimal buyer decision in each period $t$ is to best respond according to a threshold strategy w.r.t. $\boldsymbol{x}_d$ (Definition 4.3.2), where $\boldsymbol{x}_d \in [0,1]^N$ is the unique optimal threshold solution to $U(d)$ (Equation (4.3)). Further, the best response buyer decision induces a per-period expected revenue*

$$\mathsf{rev}(d) := d \sum_{n\in[N]} g_n x_{d,n}\,. \tag{4.4}$$

*Then,* $\max_{d\in\mathcal{D}} \mathbb{E}\left[d\sum_{t\in[T]} z_t^*(d)\right] = T\max_{d\in\mathcal{D}} \mathsf{rev}(d)$ *and thus* $Reg_{sell} = T\max_{d\in\mathcal{D}} \mathsf{rev}(d) - \sum_{t\in[T]} \mathbb{E}\left[d_t z_t\right]$.

We refer readers to the proof in Appendix C.2.2.

### 4.3.2 Structure of Benchmark Seller Revenue

Here, we present a special underlying structure of the seller revenue $\mathsf{rev}(d)$ defined in Equation (4.4) which will motivate our pricing algorithm in the next Section 4.4. The goal of this section is to develop efficient ways to identify $\arg\max_{d\in\mathcal{D}} \mathsf{rev}(d)$ by avoiding exploring each possible price in $\mathcal{D}$. In the rest of the paper, we make the following assumption to rule out trivial problem instances (e.g. cases when the optimal solution $\boldsymbol{x}_d$ corresponding to some $d \in \mathcal{D}$ has all 0 entries or when one of the constraints are redundant):

**Assumption 4.3.1.** *For any $d \in \mathcal{D}$, assume $V_N - \gamma d < 0 < V_1 - \gamma d$ and $\sum_{n\in[N]}(V_n - \gamma d)g_n \neq 0$. Furthermore, assume $D_M < \rho < D_1$.*

To begin with, we categorize all prices $d \in \mathcal{D}$ based on whether constraints are binding under the corresponding optimal solution $\boldsymbol{x}_d$.

**Definition 4.3.3.** *For price $d$ let $\boldsymbol{x}_d$ be the optimal threshold-based solution to $U(d)$ in Equation (4.3). Then we call $d$*

- ***Non-binding**, if under $\boldsymbol{x}_d$, both constraints are non binding, i.e., $d\sum_{n\in[N]} g_n x_{d,n} < \rho$ and $\sum_{n\in[N]} (V_n - \gamma d)\, g_n x_{d,n} > 0$;*
- ***Budget binding** if under $\boldsymbol{x}_d$, the budget constraints is binding, i.e. $d\sum_{n\in[N]} g_n x_{d,n} = \rho$ and $\sum_{n\in[N]}(V_n - \gamma d) g_n x_{d,n} > 0$;*
- ***ROI binding** if under $\boldsymbol{x}_d$, the ROI constraint is binding, i.e. $\sum_{n\in[N]}(V_k - \gamma d) g_n x_{d,n} = 0$ and $d\sum_{n\in[N]} g_n x_{d,n} \leq \rho$.*

It is apparent that any price $d \in \mathcal{D}$ must belong to at least one of these categories. Also, if a price is non-binding, it cannot be budget binding or ROI binding.

Our main result of this subsection is the following Theorem 4.3.3, which states that as we traverse $\mathcal{D}$ in increasing price order, prices are first non-binding and the revenue $\mathsf{rev}(d)$ increases in $d$; then prices become budget binding, where revenue remains constant at $\mathsf{rev}(d) = \rho$; finally prices become ROI binding, where $\mathsf{rev}(d)$ decreases in $d$. The proof can be found in Appendix C.2.3.

**Theorem 4.3.3** (Bell-shaped Structure of the Revenue Function)**.** *Suppose that Assumption 4.3.1 holds. Then, the following hold*

1. *For any non-binding prices $d, \widetilde{d}$, if $d < \widetilde{d}$ then $\mathsf{rev}(d) < \mathsf{rev}(\widetilde{d})$.*
2. *If $d$ is budget binding, any price $\widetilde{d} > d$ cannot be non-binding, which means $\widetilde{d}$ is budget binding or ROI binding.*
3. *If $d$ is ROI binding, then any $\widetilde{d} > d$ must also be ROI binding. Furthermore, $\mathsf{rev}(d) > \mathsf{rev}(\widetilde{d})$.*

We provide an illustration of Theorem 4.3.3 in Figure 4-1 that depicts the "non-binding $\rightarrow$ budget binding $\rightarrow$ ROI binding" transition phenomenon, as well as a corresponding revenue "increase $\rightarrow$ plateau $\rightarrow$ decrease", as we traverse prices in increasing order. We note that for specific model primitives $\boldsymbol{g}, \gamma, \rho$, there may exist no budget binding prices (as shown in right subfigure in Figure 4-1), meaning that there are scenarios in which it is impossible for the buyer to extract the entire buyer budget. Nevertheless, this transition phenomena suggests that we can efficiently identify the maximizing revenue $\arg\max_{d \in \mathcal{D}} \mathsf{rev}(d)$ by utilizing a simple binary search approach. Hence, we utilize this structure of $\mathsf{rev}(d)$ to motivate our pricing algorithm.

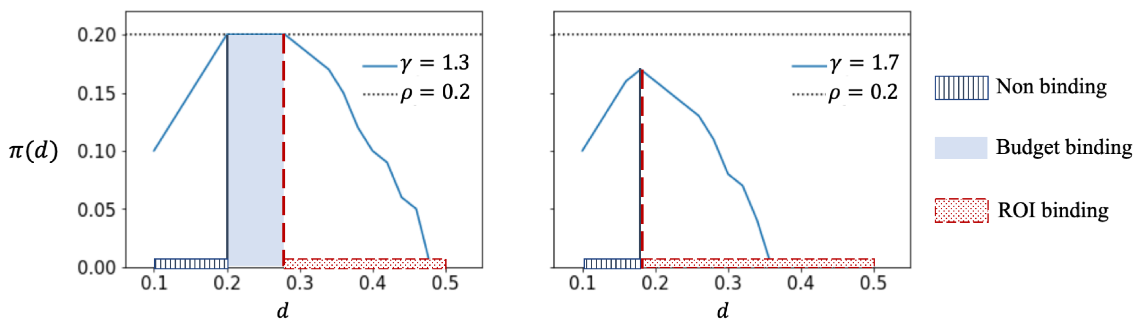


Figure 4-1: **Seller revenue function bell-shape structure.** Model primitives: number of unique buyer valuations $N = 6$, valuation set $\mathcal{V} = (0.6, 0.5, 0.4, 0.3, 0.2, 0.1)$, valuation distribution $\boldsymbol{g} = (0.1, 0.1, 0.2, 0.1, 0.2, 0.3)$, seller price set $\mathcal{D} = (0.5, 0.48 \ldots 0.1)$, buyer budget rate $\rho = 0.2$. The left and right subfigures correspond to target ROI $\gamma = 1.3$ and 1.7 respectively. In both cases, prices transition from non-binding to budget binding, and finally to ROI binnding. Revenue $\mathsf{rev}(d)$ increases as in $d$ when prices are non-binding, decreases in $d$ when prices are ROI binding, and remains at $\rho$ when prices are budget binding. Note that when $\gamma = 1.7$, there are no budget binding prices.

## 4.4 Pricing algorithm against an ROI and budget constrained Buyer

The main challenge the seller faces is her lack of knowledge on the buyer's model primitives, namely the buyer's valuation distribution $\boldsymbol{g}$, target ROI $\gamma$ and budget rate $\rho$. Furthermore, the seller has limited information feedback as she only observes whether the buyer accepted the price or not, i.e., the seller only observes the outcome $z_t \in \{0, 1\}$. This lack of information makes it very difficult for the seller to estimate the buyer's model primitives. Nevertheless, we propose a simple pricing algorithm that bypasses this lack of knowledge via exploiting the price transition phenomenon as characterized in Theorem 4.3.3 and Figure 4-1. We demonstrate later in subsection 4.4.1 that this algorithm achieves good performance when facing a general class of algorithms that is adaptive to nonstationary environments.

Our proposed pricing algorithm consists of an exploration phase and an exploitation phase. During the exploration phase, the algorithm searches for a revenue maximizing price $\arg\max_{d\in\mathcal{D}} \mathsf{rev}(d)$ through an episodic structure: the seller initiates the first episode $\mathcal{E}_1$, and fixes the price chosen in this episode $D_1$ for $E$ consecutive periods. At the end of the episode (i.e. after $E$ periods since the beginning of the episode), the seller records the average per-period revenue $\hat{\mathsf{rev}}(D_1) = \frac{D_1}{E}\sum_{t\in\mathcal{E}_1} z_t$, where $z_t \in \{0, 1\}$ indicates whether the buyer takes the price at time $t \in \mathcal{E}_1$. The process then repeats as the seller moves on to episodes $\mathcal{E}_2, \ldots$ This exploration phase eventually terminates when the seller has explored enough prices. The seller's pricing decision in each episode is governed by a binary search procedure over the price set $\mathcal{D}$, such that every price is chosen at most once across all episodes, and the exploration phase will have $\mathcal{O}(\log(M))$ episodes. Our pricing algorithm is detailed in Algorithm 4.

We note that our proposed algorithm does not try to learn the buyer's model primitives. We further point out that such a binary-search approach is a natural choice to identify revenue-optimal prices in the simplest monopolistic pricing setting under a typical unimodal assumption, [2] and one may wonder whether this approach can have good performances against a much more complex setting where the buyer is ROI and budget constrained and aims to learn her optimal bidding strategy. Surprisingly, in the next section we are in fact able to show this simple approach achieves good performances against buyers who are adaptive to price changes.

[2] In monopolistic pricing, the revenue-optimal price $p^*$ is charachterized by $d^* = \arg\max_d dF(d)$, where $F$ is the cdf of buyer valuations. A typical assumption is such that the function $dF(d)$ is unimodal.

**Algorithm 4** Episodic Binary Search

**Input:** Exploration episode length $E$.

Initialize iteration index iter $= 1$.

**Exploration episodes**:

Set $D_1$ for $E$ consecutive periods, and record per-period revenue $\hat{\mathsf{rev}}(D_1)$. Then set $D_M$ for $E$ consecutive periods, and record average per-period revenue $\hat{\mathsf{rev}}(D_M)$.

Set $m^* \leftarrow \arg\max_{m \in \{1, M\}} \hat{\mathsf{rev}}(D_m)$ $\mathrm{L} = 1$, $\mathrm{R} = M$, med $= \lfloor \frac{\mathrm{L}+\mathrm{R}}{2} \rfloor$.

**while** L $<$ R **do**

  iter $\leftarrow$ iter $+ 1$.

  **if** per-period revenue $\hat{\mathsf{rev}}(D_k)$ is not recorded for $k = \mathrm{med}, \mathrm{med}+1$ **then**

    Set price $D_k$ for $E$ consecutive periods and record per-period revenue $\hat{\mathsf{rev}}(D_k)$ for $k = \mathrm{med}, \mathrm{med}+1$

  **end if**

  **if** $\hat{\mathsf{rev}}(D_{\mathrm{med}}) < \hat{\mathsf{rev}}(D_{\mathrm{med}+1})$ **then**

    Set $m^* \leftarrow \arg\max_{m \in \{m^*, \mathrm{med}+1\}} \hat{\mathsf{rev}}(D_m)$, $\mathrm{L} \leftarrow \mathrm{med} + 1$, med $\leftarrow \lfloor \frac{\mathrm{L}+\mathrm{R}}{2} \rfloor$

  **else**

    Set $m^* \leftarrow \arg\max_{m \in \{m^*, \mathrm{med}\}} \hat{\mathsf{rev}}(D_m)$, $\mathrm{R} \leftarrow \mathrm{med} - 1$, med $\leftarrow \lfloor \frac{\mathrm{L}+\mathrm{R}}{2} \rfloor$

  **end if**

**end while**

**Exploitation episode**:

Set price $D_{m^*}$ for the remaining periods.

For notation convenience, we denote $\mathcal{E}_h$ as the collection of periods in episode $h$. Finally, we remark that the exploration episode length $E$ is deterministic and depends on the total number of periods $T$.

### 4.4.1 Regret Analysis of Pricing Algorithm

In this section, we provide theoretical guarantees for our proposed pricing algorithm against buyer algorithms whose induced decisions approximate single-round best responses (see Definition 4.3.2) in the average sense. We formally define algorithms with such properties as follows:

**Definition 4.4.1** ($\xi$-Adaptive Buyer Algorithms)**.** *We say a buyer algorithm is $\xi$-adaptive to seller algorithm 4 for some $\xi \in (0, 1)$ if the induced decisions $\{z_t\}_{t \in [T]}$ in* *any exploration or exploitation episode $\mathcal{E}_h$ satisfies*

$$\left| \frac{D_h}{|\mathcal{E}_h|} \sum_{t \in \mathcal{E}_h} z_t - \mathsf{rev}(D_h) \right| \leq \frac{\phi(|\mathcal{E}_h|)}{|\mathcal{E}_h|} \tag{4.5}$$

*with probability (w.p.) at least $1 - 1/T$ for some increasing error function $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ and $\phi(x) = \mathcal{O}(x^{1-\xi})$. Here $D_h$ is the price set in episode $h$, and $\mathsf{rev}(\cdot)$ is the per-period revenue function under buyer best response defined in Equation (4.4).*

The term $\left| \frac{D_h}{|\mathcal{E}_h|} \sum_{t \in \mathcal{E}_h} z_t - \mathsf{rev}(D_h) \right|$ is the seller's average revenue loss, relative to the revenue from optimal buyers, over a certain period with a fixed price $D_h$. Alternatively, the term can be viewed as the buyer's deviation from best responding since $\frac{\mathsf{rev}(D_h)}{D_h} = \sum_{n \in [N]} g_n x_{D_h, n}$ is the optimal probability with which the buyer should take price $D_h$.

The main result of this subsection is presented in Theorem 4.4.1, which characterizes the performance of our pricing algorithm against any $\xi$-adaptive buyer algorithm. The proof of Theorem 4.4.1 can be found in Appendix C.3.1.

**Theorem 4.4.1** (Pricing against $\xi$-adaptive buyers)**.** *Consider the seller runs Algorithm 4 against an $\xi$-adaptive buyer algorithm (Definition 4.4.1). Fix $\epsilon \in (0, \xi)$ independent of $T$. Then by setting exploration episode length $E = T^{1-\xi+\epsilon}$ in seller algorithm 4, for large enough $T$ under Assumption 4.3.1 the seller's regret is bounded as*

$$\begin{aligned} Reg_{sell} \leq & 2 \left( \lfloor \log_2(M) \rfloor + 1 \right) \cdot T^{1-\xi+\epsilon} + \phi(T) \\ & + \left( \lfloor \log_2(M) \rfloor + 1 \right)^2 / 2 \, , \end{aligned} \tag{4.6}$$

*where $\phi$ is the error function defined Equation (4.5).*

The first term $T^{1-\xi+\epsilon}$ in the seller's regret (see Equation (4.2)) characterizes the number of periods required for the buyer's algorithm to approximate the best-responding decisions in each episode facing a fixed price; the second term $\phi(T)$ represents the buyer's deviation from the best response. Finally, we point out that although in Theorem 4.6 we set the exploration episode length to be $E = T^{1-\xi+\epsilon}$, the seller does not need to know the exact value of $\xi$ as a lower bound would be sufficient: if the seller knows some lower bound for $\xi$, say $\xi' < \xi$, she can set $E = T^{1-\xi'}$, and the final seller regret would become $\text{Reg}_{\text{sell}} \leq 2\left(\lfloor \log_2(M) \rfloor + 1\right) \cdot T^{1-\xi'} + \phi\left(T\right) + \left(\lfloor \log_2(M) \rfloor + 1\right)^2/2$ for large enough $T$.

Another interesting observation for the seller regret is that its dependence on the price set dimension $M$ is logarithmic, meaning that our Algorithm 4 is robust w.r.t. the size of the seller's decision set. In fact, later in Section 4.6, we discuss that this nice logarithmic dependence on $M$ allows us to easily handle continuous price sets without causing decay in seller performance by using a simple discretization approach.

## 4.5 Example of adaptive and buyer-regret minimizing algorithms

In this section, we present simple examples of buyer algorithms that are adaptive in the sense of Definition 4.4.1, and also aim to satisfy budget and ROI constraints (Equation (4.1)) while attaining low buyer regret, where the regret of the buyer is defined as

$$\text{Reg}_{\text{buy}} = \text{OPT}(\boldsymbol{d}_{1:T}) - \sum_{t\in[T]} \mathbb{E}\left[v_t z_t\right] . \tag{4.7}$$

Here $\{z_t\}_{t\in[T]}$ is the sequence of buyer binary decisions produced by the buyer algorithm. Also recall OPT is the buyer's optimal hindsight total value described in Equation (4.1). In the following subsections, we consider a clairvoyant buyer who best responds in each period as well as a buyer who possess machine-learned (ML) advice with which she uses to make decisions. We then further characterize seller regret of our proposed Algorithm 4 against such buyers.

### 4.5.1 Best-responding buyer

As a warm-up buyer example, we first consider a clairvoyant buyer who knows her value distribution $\boldsymbol{g}$, which means the buyer has nothing to learn from the data and thus can best respond in the sense of Definition 4.3.2 during each period to maximize value under both budget and ROI constraints (Equation (4.1). We show in the following lemma that best responding is adaptive (see proof in Appendix C.4.1).

**Lemma 4.5.1** (Best-responding is 1/2-adaptive)**.** *There exists some $T_0 \in \mathbb{N}$ such that for all $T > T_0$, best responding is $\frac{1}{2}$-adaptive (Definition 4.4.1).*

Combining Lemma 4.5.1 and Theorem 4.4.1, we present the regret of Algorithm 4 against a best responding buyer in the following theorem whose proof can be found in Appendix C.4.2

**Theorem 4.5.2** (Seller's regret against best responding buyer)**.** *Assume the buyer always best responds, then for a fixed $\epsilon \in (0, \frac{1}{2})$ independent of $T$, if the seller sets prices with episode length $E = T^{\frac{1}{2}+\epsilon}$ using Algorithm 4, then for large enough $T$, the seller's regret is bounded as $Reg_{sell} \leq \mathcal{O}(T^{\frac{1}{2}+\epsilon})$. On the other hand, the buyer also incurs $\mathcal{O}(T^{\frac{1}{2}+\epsilon})$ regret, and both budget and ROI constraints are satisfied.*

In this clairvoyant buyer setting, since the buyer is not learning and always best responds, the $T^{\frac{1}{2}}$ constituent in the seller regret is due to learning error from the seller. In the next section, we introduce a buyer who is non-clairvoyant and also constantly learns how to respond, and further discuss how buyer and seller learning errors simultaneously impact seller regret.

### 4.5.2 Buyer with machine-learned (ML) advice

In a real world scenario, buyers typically do not know their value distribution; e.g. buyers may be unaware of the likelihood of conversion of their ad impressions. However, the emergence of data-driven tools for online advertising platforms have provided buyers with additional analytics, or so-called ML advice, to help buyers estimate ad conversion. In this subsection, we consider a buyer who possesses ML advice in the form of distribution estimates of $\boldsymbol{g}$ with which she uses to approximate best responses against posted prices. Formally, we characterize such ML-advice-driven buyer responses as followed:

**Definition 4.5.1** (Approximate best response with ML advice.)**.** *Assume in each period $t$, the buyer obtains ML advice $\hat{\boldsymbol{g}}_t \in \Delta_N$ that only depends on historical data $\{v_\tau\}_{\tau\in[t]}$ s.t. $\|\hat{\boldsymbol{g}}_t - \boldsymbol{g}\| < \ell_t$ where $\ell_t$ is some estimation error. Then, the buyer solves for the optimal solution $\hat{\boldsymbol{x}}_t$ in Equation (4.3) via replacing the true distribution $\boldsymbol{g}$ with $\hat{\boldsymbol{g}}_t$, and then adopts a threshold strategy w.r.t. $\hat{\boldsymbol{x}}_t$ (see Definition 4.3.1).*

We remark that ML advice in the form of distributional estimates is very common. For model-based approaches, ML algorithms assume distributions take a certain parametric form and then uses data to estimate unknown distribution parameters; see e.g. [47] for an intro on maximum likelihood estimation. For more general non-parametric approaches, ML advice concerns using empirical estimates (or so-called histogram estimates), which we will later discuss in Theorem 4.5.4.

The following lemma relates ML advice driven approximate responses to our notion of buyer adaptivity in Definition 4.4.1, with which we are able to quantify seller regret in light of Theorem 4.4.1. The detailed proof can be found in Appendix C.4.3

**Theorem 4.5.3** (Seller regret against approximate best responding buyer with ML advice)**.** *Assume the buyer approximate best responds with ML advice (Definition 4.5.1) and there exists some $L \in (0,1)$ s.t. in each exploration or exploitation episode $h$ of Algorithm 4 the estimation errors, denoted by $\ell_t$'s, satisfy $\lim_{t\to\infty} \ell_t = 0$ and $\sum_{t\in\mathcal{E}_h} \ell_t \le \widetilde{\phi}(|\mathcal{E}_h|)$ for some increasing function $\widetilde{\phi} : \mathbb{R}_+ \to \mathbb{R}^+$ and $\widetilde{\phi}(x) \le \mathcal{O}(x^{1-L})$. Then this buyer algorithm is $\xi$-adaptive for $\xi = \min\{\frac{1}{2}, L\}$. Further, by setting exploration episode length $E = T^{1-\xi+\epsilon}$ for some $\epsilon \in (0,\xi)$ independent of $T$, the seller regret is exactly that in Equation (4.6) of Theorem 4.4.1 for large enough $T$. On the buyer side, we have $Reg_{buy} \le \mathcal{O}(T^{1-\xi})$ and the induced buyer decisions $\{z_t\}_{t\in[T]}$ satisfy*

$$\frac{1}{T}\mathbb{E}\left[\sum_{t\in[T]} (v_t - \gamma d_t)\, z_t\right] \ge -\Theta(T^{-L})$$

*and*

$$\frac{1}{T}\mathbb{E}\left[\sum_{t\in[T]} d_t z_t\right] \le \rho + \Theta(T^{-L})\,.$$

We remark that best responding buyers considered in Section 4.5.1 can be viewed as a special case of buyers with ML advice where the advice is perfect, i.e. $\ell_t = 0$ for all $t$ so $\widetilde{\phi}(x) \equiv 0$ and consequently $L = 1$. This recovers our results in Theorem 4.5.2.

Here, we also quickly discuss the aggregate impact of buyer and seller learning error on the seller regret of our proposed Algorithm 4. In particular, the constituent $T^{1-\xi} = T^{1-\min\{\frac{1}{2},L\}}$ in the seller regret arises from learning errors of both the buyer and the seller. We can view the seller's learning rate to be in the order of $t^{-\frac{1}{2}}$, and the buyer learning rate to be of order $t^{-L}$, and thus we see that the seller regret is governed by the agent that learns at a slower rate: if the buyer is learning more slowly, i.e. $L < \frac{1}{2}$, then the seller regret is driven by the buyer learning loss; a similar argument applies for the case when the buyer learns more quickly.

To conclude this section, we present a concrete example for buyers with ML advice: consider the simple ML advice that is an empirical estimate of the buyer's value distribution:

$$\hat{\boldsymbol{g}}_t = \frac{1}{t}\cdot\Big(\sum_{\tau\in[t]}\mathbb{I}\{v_\tau = V^1\},\dots,\sum_{\tau\in[t]}\mathbb{I}\{v_\tau = V^N\}\Big)\,. \tag{4.8}$$

Then, both the buyer and seller regret are characterized in the following theorem (see proof in Appendix C.4.4).

**Theorem 4.5.4** (Seller regret against approximate best responding buyer with empirical distribution estimates)**.** *When the buyer approximate best responds with ML advice in the form of empirical estimates as defined in Equation (4.8), Theorem 4.5.3 holds for $L = \xi = \frac{1}{2}$ w.p. at least $1 - 1/T$.*

## 4.6 Additional discussions and future research

### Continuous price set

We remark that our main results in this paper, specifically the analyses of Algorithm 4 and the corresponding seller regret, can be easily extended to handle continuous seller price sets, as the seller regret in Theorem 4.4.1 only depends logarithmically on $M$ which we recall to be the size of a discrete price set. Assuming the price decision set is $[0,1]$, the approach that the seller can take is to discretize the decision set into $\mathcal{D} = \{\frac{1}{T}, \frac{2}{T} \dots 1\}$ with size $|\mathcal{D}| = T$. Recall $\pi(d)$ defined in Equation (4.4) is the expected per-period seller revenue under buyer best response, and define $d^* = \arg\max_{d\in[0,1]} \mathsf{rev}(d)$ to be the optimal price w.r.t. the continuous set, such that the seller regret is now $\text{Reg}_{\text{sell}} = T \cdot \pi(d^*) - \sum_{t\in[T]} \mathbb{E}[d_t z_t]$ (see Proposition 4.3.2). Then, for a price $\widetilde{d} \in \mathcal{D}$ in the discretized set $\mathcal{D}$ that is close to $d^*$ such that $|\widetilde{d} - d^*| < \frac{1}{T}$, similar to our proof in Theorem 4.5.3 we can show that the optimal solutions $\boldsymbol{x}_d$ and $\boldsymbol{x}_{\widetilde{d}}$ to the per-period buyer optimization problem $U(d)$ and $U(\widetilde{d})$ (see Equation (4.1)), respectively, are also close to one another. Further, we can show that $\mathsf{rev}(d^*) - \mathsf{rev}(\widetilde{d}) \leq O(\frac{1}{T})$. Therefore, via running Algorithm 4 w.r.t. the discretized price set $\mathcal{D}$, our seller regret when facing a $\xi$-adaptive buyer (Definition 4.4.1) can be bounded as

$$
\begin{aligned}
\text{Reg}_{\text{sell}} &= T \max_{d\in[0,1]} \mathsf{rev}(d) - \sum_{t\in[T]} \mathbb{E}[d_t z_t] \\
&= \underbrace{T\left(\pi(d^*) - \max_{d\in\mathcal{D}} \mathsf{rev}(d)\right)}_{\text{discretization error}} + T \max_{d\in\mathcal{D}} \mathsf{rev}(d) - \sum_{t\in[T]} \mathbb{E}[d_t z_t] \\
&\leq T(\pi(d^*) - \mathsf{rev}(\widetilde{d})) + T \max_{d\in\mathcal{D}} \mathsf{rev}(d) - \sum_{t\in[T]} \mathbb{E}[d_t z_t] \\
&\leq \mathcal{O}(1) + T \max_{d\in\mathcal{D}} \mathsf{rev}(d) - \sum_{t\in[T]} \mathbb{E}[d_t z_t] \\
&\leq \mathcal{O}(1) + \mathcal{O}(\log(T) T^{1-\xi+\epsilon} + \phi(T)),
\end{aligned}
$$

where the final inequality follows from the seller regret (Equation (4.6)) in Theorem 4.4.1 by setting the price set size $M = T$. That being said, the discretization error introduced to the seller regret is only in the order of $\mathcal{O}(1)$, and this is due to the the fact that the bell-shape structure of seller's revenue (Theorem 4.3.3) along with our seller algorithm yields a seller regret that is logarthmic in the discrete price set size.

## Future directions

One natural future research direction that is of both theoretical and practical interest involves designing pricing algorithms when facing multiple financially constrained buyers. The multi-buyer analogue to our single-buyer posted price setup in this work is to set a single reserve price in each period over time where constrained buyers compete in a second-price auction (see e.g. setup in [54] for non-constrained buyers). The key challenge lies in the fact that in this multi-buyer setup we no longer have the salient bell-shape structure in the seller revenue function, and more importantly buyer algorithmic interactions introduce significant difficulties to the analyses of seller regret. Similar challenges that arise from selling to multiple learning buyers have also been discussed (but not resolved) in related works such as [25, 38].

# Chapter 5

# Improving individual advertiser welfare with ML-advice

*This chapter is based on [33], which is joint work with Yuan Deng, Negin Golrezaei, Patrick Jaillet, and Vahab Mirrokni.*

## 5.1 Introduction

Online advertisers have access to a vast array of digital advertising channels, such as social media, web display, keyword search, etc., from which they can procure ad impressions and drive user traffic. One possible way for these channels to improve overall attractiveness and retention is to design appropriate ad auction mechanisms that enhance advertisers' total welfare, which reflects the aggregate advertiser-perceived value for procured ad impressions on the channel.

For instance, consider advertisers whose ad campaign objective is to maximize ad clicks that direct users to landing pages of their services or products, as described in [53]. These advertisers' perceived value of procured ad impressions is their click conversion rate, and thereby ad channels' welfare maximization goal translates into improving the aggregate realized click conversion among all participating advertisers.

Academic literature has developed various approaches to improve total welfare, one of which involves predicting advertiser values by applying machine learning tools to data on users' interactions with ads. In the instance where welfare corresponds to click conversion, channels use ML algorithms to produce predictions (i.e., ML advice) on click conversion rates for impressions. See [99, 103, 88] or [112] for a comprehensive survey on click predictions.

Having obtained ML advice on advertiser values, recent works such as [37, 11, 36] motivate the approach to augment existing ad auctions by directly setting personalized reserve prices for advertisers using such ML advice, and show theoretical guarantees on total welfare improvement.

Nevertheless, these results have two important issues. First, the pursuit of improving total welfare may not necessarily guarantee all individual advertisers benefit equally and may come at the expense of certain individual advertisers' welfare. For instance, larger advertisers acquiring more impressions while smaller advertisers receive fewer impressions could potentially harm the businesses of smaller advertisers and compromise the overall health of the channel in the long term. Second, welfare improvement guarantees are presented in a price of anarchy (POA) fashion, which measures the worst-case outcome total welfare compared to the maximum achievable (or efficient) welfare. However, these POA bounds are independent of advertiser bidding strategies and thus do not shed light on how particular advertiser bidding strategies in ad auctions impact individual or total welfare.

In light of these insufficiencies of existing results, in this work, we address the following questions:

> *Given an advertisers' bidding strategy to procure impressions in ad auctions, how can platforms characterize the potential welfare loss of this individual advertiser? How should ad channels utilize machine-learned advice that predict advertiser values to improve individual welfare?*

We study a prototypical *autobidding setting* where advertisers compete simultaneously in numerous multi-slot position auctions that are run in parallel, and aim to maximize total advertiser value under *return-on-ad-spent (ROAS)* constraints that restrict total spend of a bidder to be less than her total acquired value across all auctions in an average sense; see similar setups in [4, 37, 11, 89]. On the other hand, ad platforms possess ML advice that predicts advertisers' real values with a certain degree of accuracy/quality. Under this setup, our main contributions and organization of the paper is described as followed:

**Strategy-dependent individual welfare guarantee metric for individual advertisers.** In Section 5.2, we present a novel individual welfare metric that measures the difference between two specific welfare outcomes of an individual advertiser: fixing a certain bidding strategy, the worst case welfare over all auction outcomes under which all bidders' ROAS constraints are satisfied; and the welfare this individual bidder would have obtained in the global welfare maximizing outcome. Our metric achieves two key goals: 1. it characterizes individual welfare loss, and 2. allows platforms to uncover the relationship between advertiser strategies and individual welfare guaranties. Our proposed metric is the first of its kind to achieve these two goals.

**Individual welfare guarantees in VCG auctions by setting personalized**

**reserves with ML advice.** In Section 5.3, we illustrate through examples that setting ML advice as personalized reserves as in [37, 11, 36] surprisingly improves individual welfare guarantees under our aforementioned individual welfare metric. Then, in Section 5.4 where we consider parallel VCG auctions, and focus on an individual autobidder who adopts uniform bidding (Proposition 5.2.1), we formally show that augmenting such auctions with ML-advice-based reserves allows us to present an individual welfare lower bound guarantee for this advertiser that increases in the advertiser's uniform bid multiplier, the quality of ML advice, and the relative market share of this advertiser compared to competitors (Theorem 5.4.1). Together with results in [37] stating ML-advice-based reserves can improve total welfare, we conclude that incorporating ML advice as personalized reserves achieves "best of both worlds" by simultaneously benefiting total and individual welfare.

**Impossibility result: VCG yields the best individual welfare guarantees among a broad class of auctions.** In Section 5.5, we show an impossibility result that says no allocation-anonymous, truthful, and possibly randomized auction format with ML advice of given quality can achieve a strictly better individual welfare guarantee than the VCG auction coupled with ML advice of the same quality; see Theorem 5.5.1. In particular, for any allocation-anonymous, truthful, and possibly randomized auction, we construct a problem instance with personalized reserves based on ML-advice of given quality, and show that there must be at least 1 bidder whose welfare is at most the welfare lower bound guarantee we presented under VCG (i.e. Theorem 5.4.1).

**Extending individual welfare guarantee to GSP and GFP.** We extend individual welfare guarantee results to GSP and GFP auctions, and show that a similar individual welfare lower bound guarantee for VCG continues to hold (Theorem 5.6.2). We compare these lower bound guarantees in GSP and GFP with that of VCG and identify conditions under which VCG outperforms (or underperforms) GSP/GFP in terms of our individual welfare metric with the same ML advice quality.

**Numerical results.** We present numerical studies using semi-synthetic data derived from auction logs of a search ad platform to showcase individual welfare improvement via setting personalized reserve prices with ML-advice. We demonstrate that as ML-advice quality improves, more advertisers' welfare would approach what they would have obtained under the efficient outcome.

## Related Works

**Autobidding and total welfare maximization.** The most relevant works to this paper are [37, 11, 89], where they consider the same autobidding setting (i.e. value-maximizers with ROAS constraints) as ours. [37, 11, 89, 36] all present techniques to improve price-of-anarchy bounds for the total welfare of any feasible outcome in which all bidders' ROAS constraints are satisfied: where [37] relies on additive boosts on bid values, [11, 36] utilizes approximate reserve prices derived from ML-advice, and [89] develops randomized allocation and payment rules. Our work distinguishes itself from these works as we focus on welfare and individual welfare guarantees on the individual bidder level, and also sheds light on how autobidders' uniform bidding strategies affect individual welfare loss. We point out that our proof techniques also differ from those in [37, 11, 89, 36] as our individual individual welfare guarantees require novel analyses on the value-expenditure tradeoffs individual bidders' would face when they are tempted to outbid others to acquire more value; see discussion in Section 5.4 for more details.

**Exploiting machine-learned advice.** ML advice has been utilized in various applications to improve learning and decision making. For example, [108] exploits ML advice to develop algorithms for the multi-shop ski-rental problem, [87] adopts ML advice for the caching problem, and [69] studies online page migration with ML advice. However, although many works in online advertising studied predictive models for advertiser values, click through rates, etc (see e.g. [99, 83, 103]), the literature on applying such predictions (or more generally, ML advice) to the mechanism design problem has been scarce. See also [30, 56] for works that exploit sample information (unstructured ML advice) in online decision-making. One related work along this direction is [91], which develops a theoretical framework to optimize reserve prices in a posted price mechanism by utilizing prediction inputs on bid values. In this work, we do not optimize for reserves, and motivate the simple approach of setting reserves using ML advice to improve individual advertiser welfare. Finally, we note that our work contributes to the area of exploiting ML advice to designing mechanisms for improving welfare guarantees for individual bidders.

## 5.2 Preliminaries

We describe our model in the context of sponsored search as in Section 5.1, but remark that all results and insights apply to general online advertising setups such as web display, e-commerce, social newsfeed, etc. Consider $N$ bidders (i.e. advertisers) participating in $M$ parallel position auctions $(\mathcal{A}_j)_{j\in[M]}$, where each auction $\mathcal{A}_j$ is instantiated by a user keyword search query. An auction $\mathcal{A}_j$ sells to bidders $L_j \geq 1$ ad slots that are ordered by visual prominence, or equivalently the likelihood of the user viewing the slot, on the webpage, represented by click-through-rates (CTR) $1 \geq \mu_j(1) \geq \mu_j(2) \geq \ldots \geq \mu_j(L_j) \geq 0$, where $\mu_j(\ell)$ is the likelihood of the user of auction $j$ viewing slot $\ell \in [L_j]$ (see an intro to position auctions in e.g. [80, 106, 46]). A bidder $i \in [N]$ possess a private value-per-click equal to $v_{i,j} > 0$ for auction $\mathcal{A}_j$ that represents her perceived value conditioned on the user viewing her ad, so her attained utility for winning slot $\ell \in [L]$ is $\mu_j(\ell) \cdot v_{i,j}$.

In the following subsection 5.2.1 we present an overview for position auction mechanisms; in subsection 5.2.2 we describe bidder objectives and actions; and finally in subsection 5.2.3 we introduce definitions regarding bidder individual welfare guarantees.

### 5.2.1 Preliminaries for a single position auction

A (possibly randomized) position auction $\mathcal{A}$ with $L \geq 1$ slots is characterized by a tuple $(\mathcal{X}, \mathcal{P}, \boldsymbol{\mu})$, where $\mathcal{X}$ is an allocation rule, $\mathcal{P}$ is a payment rule, and CTRs $\boldsymbol{\mu} = (\mu(\ell))_{\ell\in[L]} \in [0,1]^L$ that satisfies $1 \geq \mu(1) \geq \mu(2) > \ldots \geq \mu(L) \geq 0$. Let $N$ bidders with private value per-clicks $\boldsymbol{v} = (v_i)_{i\in[N]}$ participate in auction $\mathcal{A}$ by submitting a bid profile $\boldsymbol{b} = (b_i)_{i\in[N]} \in \mathbb{R}_+^N$, and we describe the payment and allocation rules as followed:

The allocation rule $\mathcal{X} : \mathbb{R}_+^N \to \{0,1\}^{N\times L}$ maps bid profile $\boldsymbol{b} \in \mathbb{R}_+^N$ to an outcome $\boldsymbol{x} = \mathcal{X}(\boldsymbol{b}) \in \{0,1\}^{N\times L}$ which may possibly be random. The entry $x_{i,\ell} = 1$ if bidder $i$ is allocated slot $\ell \in [L]$, and 0 otherwise. Here, each slot $\ell$ is at most allocated to one bidder so $\sum_{i\in[N]} x_{i,\ell} \leq 1$ for any $\ell$. Further, under outcome $\boldsymbol{x} \in \{0,1\}^{N\times L}$, bidder $i$ who has value $v_i$ attains a total welfare of $W_i(\boldsymbol{x}) = v_i \sum_{\ell\in[L]} \mu(\ell) x_{i,\ell}$. That is, if bidder $i$ is allocated slot $\ell \in [L]$ (i.e., $x_{i,\ell} = 1$), her welfare is $W_i(\boldsymbol{x}) = \mu(\ell) v_i$. The payment rule $\mathcal{P} : \mathbb{R}_+^N \to \mathbb{R}_+^N$ maps bids $\boldsymbol{b}$ to payments $\mathcal{P}(\boldsymbol{b}) \in \mathbb{R}_+^N$ where $\mathcal{P}_i(\boldsymbol{b})$ is the payment of bidder $i$. In this work, we focus on the class of auctions that are *ex-post individual rational* (IR), i.e. the payment for any bidder is less than her submitted bid, or mathematically $\mathcal{P}_i(b_i, \boldsymbol{b}_{-i}) \leq b_i$ for any $\boldsymbol{b}_{-i} \in \mathbb{R}_+^{N-1}$. We note that the classic VCG, GSP and GFP auctions are ex-post IR. Further, we assume bidders who submit a 0 bid value will not be allocated any slots and incur no payment.

Having introduced general position auction allocation and payment rules, in the following we define three special auction classes, namely truthful auction (Definition 5.2.1), allocation anonymous auctions (Definition 5.2.2), and personalized reserve augmented allocation anonymous auctions (Definition 5.2.3).

**Definition 5.2.1** (Truthful auction)**.** *Consider position auction $\mathcal{A} = (\mathcal{X}, \mathcal{P}, \boldsymbol{\mu})$ where we recall $\mathcal{X}, \mathcal{P}$ are possibly random allocation and payment rules, and $\boldsymbol{\mu} \in [0,1]^L$ are CTRs. Then we say the auction is truthful if for any bidder $i \in [N]$, her value $v_i \in \arg\max_{b\geq 0} \mathbb{E}\left[W_i(\mathcal{X}(b, \boldsymbol{b}_{-i})) - \mathcal{P}_i(b, \boldsymbol{b}_{-i})\right]$ for any competing bid profile $\boldsymbol{b}_{-i}$, where the expectation is taken w.r.t. possible randomness in $(\mathcal{X}, \mathcal{P})$, and recall welfare $W_i(\boldsymbol{x}) = v_i \sum_{\ell\in[L]} \mu(\ell) x_{i,\ell}$ with $\boldsymbol{x} = \mathcal{X}(b, \boldsymbol{b}_{-i})$.*

Note that the well-known VCG auctions is truthful. In truthful auctions it is a weakly dominant strategy for a bidder to bid her true value when her objective is to maximize quasi-linear utility. In the next Subsection 5.2.2 we study bidders whose objectives are not necessarily quasi-linear, so that truthful bidding is no longer weakly optimal in truthful auctions.

We next define allocation-anonymous auctions, in which if two bidders swap their bids, the probability of each bidder winning any slot will also be swapped, or in other words, the outcome of the position auction only depends on solicited bid values, and independent of bidders' identity.

**Definition 5.2.2** (Allocation anonymous auctions)**.** *Consider position auction* $\mathcal{A} = (\mathcal{X}, \mathcal{P}, \boldsymbol{\mu})$ *and any permutation* $\sigma : [N] \to [N]$ *of* $\{1 \dots N\}$*, as well as the permuted bid profile* $\boldsymbol{b}' = (b_{\sigma(i)})_{i \in [N]}$*. Let* $\boldsymbol{x} = \mathcal{X}(\boldsymbol{b})$*,* $\boldsymbol{x}' = \mathcal{X}(\boldsymbol{b}')$ *be the (possibly random) outcomes under* $\boldsymbol{b}, \boldsymbol{b}'$*, respectively. Then, we say* $\mathcal{X}$ *is allocation anonymous if for any bidder* $i \in [N]$ *and slot* $j \in [L]$*, we have* $\mathbb{P}(x_{\sigma(i),j} = 1) = \mathbb{P}(x'_{i,j} = 1)$*.*

The classic VCG, GSP, and GFP are all allocation anonymous. The following presents an illustrative example of allocation-anonymity for GSP.

**Example 5.2.1** (Example for allocation anonymous auctions)**.** *Consider a single GSP auction with 2 slots and 3 bidders who submitted a bid profile* $\boldsymbol{b} = (0.1, 0.2, 0.3)$*. As GSP allocates slots by ranking bidders' submitted bids, the outcome under bid profile* $\boldsymbol{b}$ *is* $\boldsymbol{x} = \begin{pmatrix} 0,0 \\ 0,1 \\ 1,0 \end{pmatrix}$*. Next, consider some permutation* $\sigma$ *that maps* $\{1, 2, 3\}$ *to* $\{3, 1, 2\}$*. That is,* $\sigma(1) = 3$*,* $\sigma(2) = 1$ *and* $\sigma(3) = 2$*. Under this permutation, the corresponding permuted bid profile* $\boldsymbol{b}' = (0.3, 0.1, 0.2)$*, which results in the outcome* $\boldsymbol{x}' = \begin{pmatrix} 1,0 \\ 0,0 \\ 0,1 \end{pmatrix}$*. Then, it is easy to check that* $\mathbb{P}(x_{\sigma(i),j} = 1) = \mathbb{P}(x'_{i,j} = 1) = \begin{cases} 1 & \text{if } (i,j) = (1,1) \text{ or } (3,2) \\ 0 & \text{otherwise} \end{cases}$*. In particular, because* $\sigma(1) = 3$ *we have* $\mathbb{P}(x_{3,1} = 1) = \mathbb{P}(x'_{1,1} = 1) = 1$*, and because* $\sigma(3) = 2$ *we have* $\mathbb{P}(x_{2,2} = 1) = \mathbb{P}(x'_{3,2} = 1) = 1$*.*

Finally, we describe augmenting allocation anonymous auctions with personalized reserves.

**Definition 5.2.3** (Personalized-reserve augmented allocation anonymous auctions)**.** *Fix position auction* $\mathcal{A} = (\mathcal{X}, \mathcal{P}, \boldsymbol{\mu})$*, and some vector of personalized reserve prices*

*$\boldsymbol{r} \in \mathbb{R}_+^N$ where $r_i$ is the reserve price for bidder $i \in [N]$. Then, the augmented auction is $\mathcal{A}' = (\mathcal{X}', \mathcal{P}', \boldsymbol{\mu})$ whose payment $\mathcal{X}'$ and allocation $\mathcal{P}'$ are characterized via the following procedure for any bid profile $\boldsymbol{b} \in \mathbb{R}_+^N$:*

- *$\mathcal{X}'$: Define bid profile $\boldsymbol{b}' = (b_i \cdot \mathbb{I}\{b_i \geq r_i\})_{i\in[N]}$. Then $\mathcal{X}'(\boldsymbol{b}) = \mathcal{X}(\boldsymbol{b}')$.* [1]

- *$\mathcal{P}'$: If $i \in [N]$ is not allocated a slot under outcome $\mathcal{X}'(\boldsymbol{b})$, $\mathcal{P}'_i(\boldsymbol{b}) = 0$. Otherwise, let $\ell_i \in [L]$ be the slot allocated to bidder $i$ under $\mathcal{X}'(\boldsymbol{b})$. Then, $\mathcal{P}'_i(\boldsymbol{b}) = \max\{\mathcal{P}_i(\boldsymbol{b}'), \mu(\ell_i) \cdot r_i\}$.*

Recall that a 0 bid will always result in no allocation and 0 payment, so $\mathcal{X}'$ can be effectively viewed as excluding all bidders who do not clear their reserves, and implement the allocation rule $\mathcal{X}$ with respect to the remaining bidders. We remark in later sections, the personalized reserve prices relevant in this work (based on ML-advice) guarantee all bidders clear their reserves so that no bidders will be excluded from ranking. Finally, we refer readers to Example D.1.1 for an illustration of augmenting anonymous VCG, GSP, and GFP auctions with personalized reserves.

### 5.2.2 Autobidders' objectives and bidding strategies

In this subsection we describe the scope for bidders' objectives as well as bidding strategies of interest. We recall the setup with $N$ bidders participating in $M$ parallel position auctions $(\mathcal{A}_j)_{j\in[M]}$, where $\boldsymbol{v}_j \in \mathbb{R}_+^N$ are bidders' values in auction $\mathcal{A}_j = (\mathcal{X}_j, \mathcal{P}_j, \boldsymbol{\mu}_j)$ (see definitions in Subsection 5.2.1). We use the following notations for convenience:

[1] This allocation is known as an *eager implementation* of personalized reserve prices, where any high-ranked slots are always allocated before a lower-rank slot gets allocated. There also exists a *lazy implementations*, where we first rank all bids, and then allocate slots to each bidder following this ranking if the bidder clears her reserve. Note that the lazy implementation may leave "holes" in allocation, e.g. the first and third slots are allocated while leaving the second slot un-allocated. It will become clear later that all results in this work hold for both eager and lazy implementation of personalized reserve prices.

| | |
|---|---|
| $\boldsymbol{b}_i \in \mathbb{R}_+^M$: bids submitted by $i$ | $\boldsymbol{b}_j^\top \in \mathbb{R}_+^N$: bids submitted to $\mathcal{A}_j$ |
| $\mathcal{X}_j(\boldsymbol{b}_j^\top) \in \{0,1\}^{N \times L_j}$: outcome of $\boldsymbol{b}_j^\top$ in $\mathcal{A}_j$ | $\mathcal{P}_j(\boldsymbol{b}_j^\top) \in \mathbb{R}_+^N$: payment vector of $\mathcal{A}_j$ |
| $\mathcal{X}_{i,\ell,j}(\boldsymbol{b}_j^\top) \in \{0,1\}$: indicator of $i$ winning slot $\ell$ in $\mathcal{A}_j$ | $\mathcal{P}_{i,j}(\boldsymbol{b}_j^\top) \in \mathbb{R}_+$: payment of $i$ in $\mathcal{A}_j$ |

Let $\mathcal{X}(\boldsymbol{b}) := (\mathcal{X}_j(\boldsymbol{b}_j^\top))_{j\in[M]}$. Then, bidder $i$'s welfare in auction $\mathcal{A}_j$ namely $W_{i,j}(\mathcal{X}_j(\boldsymbol{b}_j^\top))$, and her total welfare over all auctions namely $W_i(\mathcal{X}(\boldsymbol{b}))$, are defined as

$$W_i(\mathcal{X}(\boldsymbol{b})) := \sum_{j\in[M]} W_{i,j}(\mathcal{X}_j(\boldsymbol{b}_j^\top)) \quad \text{and} \quad W_{i,j}(\mathcal{X}_j(\boldsymbol{b}_j^\top)) = \sum_{\ell=1}^{L_j} \mu_j(\ell)\cdot v_{i,j}\cdot \mathcal{X}_{i,\ell,j}(\boldsymbol{b}_j^\top)\,. \tag{5.1}$$

We study the setting where each bidder is subject to a *return-on-ad-spent (ROAS)* constraint, which requires her total expenditure across all auctions to be less than her total acquired value.[2] Mathematically, fix some competing bid profile $\boldsymbol{b}_{-i} \in \mathbb{R}_+^{(N-1)\times M}$, the ROAS constraint of bidder $i$ is

$$\mathbb{E}\left[W_i(\mathcal{X}(\boldsymbol{b}_i, \boldsymbol{b}_{-i}))\right] \geq \mathbb{E}\left[\mathcal{P}_i(\boldsymbol{b}_i, \boldsymbol{b}_{-i})\right] \quad \text{where } \mathcal{P}_i(\boldsymbol{b}_i, \boldsymbol{b}_{-i}) := \sum_{j\in[M]} \mathcal{P}_{i,j}(\boldsymbol{b}_j^\top)\,. \tag{5.2}$$

Here, the expectation is taken w.r.t. possible randomness in the allocation and payment rules of auctions $(\mathcal{A}_j)_{j\in[M]}$. When allocation and payment for $(\mathcal{A}_j)_{j\in[M]}$ are deterministic (e.g. for VCG, GSP and GFP), we omit the expectation for simplicity.

We call a bidder an *autobidder* when she aims to maximize welfare $\mathbb{E}\left[W_i(\mathcal{X}(\boldsymbol{b}_i, \boldsymbol{b}_{-i}))\right]$ subject to the ROAS constraint in Eq. (5.2). The following proposition states that an autobidder's optimal bidding strategy in truthful auctions, facing any competing bid profile, is *uniform bidding*:

[2]A more general concept related to ROAS is *return-on-investment (ROI)*, where each bidder $i$ has a target ROI ratio $\mathrm{T}_i$ such that her constraint in Eq. (5.2) is instead written as $W_i(\mathcal{X}(\boldsymbol{b}_i, \boldsymbol{b}_{-i})) \geq \mathrm{T}_i \cdot \sum_{j\in[M]} p_{i,j}$; see e.g. [62, 55]. In this paper, since we study worst-case instances, we can scale all bidder $i$'s values by $\mathrm{T}_i$ so it is without loss of generality to consider ROAS constraints.

**Proposition 5.2.1** (Uniform bidding is optimal for autobidders in truthful auctions)**.** *Let all auctions* $(\mathcal{A}_j)_{j\in[M]}$ *be identical truthful auctions (see Definition 5.2.1), and bidder* $i \in [N]$ *is an autobidder who aims to maximize welfare* $\mathbb{E}\left[W_i(\mathcal{X}(\boldsymbol{b}_i, \boldsymbol{b}_{-i}))\right]$ *subject to the ROAS constraint in Eq. (5.2) for any fixed competing bids* $\boldsymbol{b}_{-i} \in \mathbb{R}_+^{N-1}$*. Then, there exists some constant uniform multiplier* $\alpha_i^* \geq 1$ *s.t. the uniform bidding profile* $\alpha_i^* \boldsymbol{v}_i$ *is i's optimal strategy:*

$$\alpha_i^* \cdot \boldsymbol{v}_i \;\in\; \arg\max_{\boldsymbol{b}_i \in \mathbb{R}_+^M} \;\mathbb{E}\left[W_i(\mathcal{X}(\boldsymbol{b}_i, \boldsymbol{b}_{-i}))\right] \;\; s.t. \quad \mathbb{E}\left[W_i(\mathcal{X}(\boldsymbol{b}_i, \boldsymbol{b}_{-i}))\right] \geq \mathbb{E}\left[\mathcal{P}_i(\boldsymbol{b}_i, \boldsymbol{b}_{-i})\right], . \tag{5.3}$$

*Further, adopting any uniform bid multiplier* $\alpha_i < 1$ *is weakly dominated by truthful bidding, i.e.*
$\mathbb{E}\left[W_i(\mathcal{X}(\alpha_i \boldsymbol{v}_i, \boldsymbol{b}_{-i}))\right] \leq \mathbb{E}\left[W_i(\mathcal{X}(\boldsymbol{v}_i, \boldsymbol{b}_{-i}))\right]$ *for any* $\boldsymbol{b}_{-i} \in \mathbb{R}_+^{N-1}$*.*

This is a well-known result that has been proved and adopted in many related works such as [4, 37, 11, 89] and we will omit the proof here. We remark that autobidding represents advertisers' conversion maximizing behavior while respecting constraints on spend. We note that our methodologies and insights in the paper can be extended to autobidders with a more general private costs objective $\mathbb{E}\left[W_i(\mathcal{X}(\boldsymbol{b}_i, \boldsymbol{b}_{-i}))\right] - \rho\mathbb{E}\left[\mathcal{P}_i(\boldsymbol{b}_i, \boldsymbol{b}_{-i})\right]$ where $\rho \geq 0$ is some private cost, but for simplicity we assume in the rest of the paper $\rho = 0$. In light of this proposition, we will assume that all autobidders will adopt bid multiplier at least 1 in truthful auctions.

Finally, we conclude by introducing the notion of feasible bid profiles:

**Definition 5.2.4** (Feasible bid profiles)**.** *For a given set of parallel auctions* $(\mathcal{A}_j)_{j\in[M]}$*, we say that a bid profile* $\boldsymbol{b} \in \mathbb{R}_+^{N\times M}$ *is feasible if Eq. (5.2) holds for all bidders, and denote all feasible bid profiles as* $\mathcal{F}$*. Further, fix a bidder i and her bids* $\boldsymbol{b}_i \in \mathbb{R}_+^M$*. Then let* $\mathcal{F}_{-i}(\boldsymbol{b}_i) = \{\boldsymbol{b}_{-i} \in \mathbb{R}_+^{(N-1)\times M} : (\boldsymbol{b}_i, \boldsymbol{b}_{-i}) \in \mathcal{F}\}$*.*[3]

In words, $\mathcal{F}_{-i}(\boldsymbol{b}_i)$ is all competing bid profiles for bidder $i$ that guarantee all bidders' ROAS constraints are satisfied. It is easy to see that in ex-post IR auctions,

[3]Here, we exclude the trivial zero bid profiles in the sets $\mathcal{F}$ and $\mathcal{F}_{-i}(\boldsymbol{b}_i)$.

any bidder can satisfy her ROAS constraint by simply bidding truthfully, since a bidder's payment is always no greater than her submitted bid. This implies that the feasible set of bid profiles $\mathcal{F}$ is never empty and contains the truthful bid profile.

### 5.2.3 Efficient auction outcomes and individual welfare guarantees

Let $\ell^*_{i,j}$ be the ranking of bidder $i$ in auction $\mathcal{A}_j$ when ranked according to true values $\boldsymbol{v}_j \in \mathbb{R}^N_+$. Then we call the outcome $\boldsymbol{x}^* = (\boldsymbol{x}^*_j)_{j\in[M]}$ with $x^*_{i,\ell,j} = \mathbb{I}\{\ell = \ell^*_{i,j}\}$, the *efficient outcome*. Note that $\boldsymbol{x}^*$ yields the largest total welfare because allocation of slots in each auction follows ranking of bidder true values for that auction. Similar to our definition for welfare for any outcome in Eq. (5.1), let

$$\mathrm{OPT}_{i,j} = \mu_j(\ell^*_{i,j}) \cdot v_{i,j}, \quad \mathrm{OPT}_i = \sum_{j\in[M]} \mathrm{OPT}_{i,j}, \quad \text{and} \quad \mathrm{OPT} = \sum_{i\in[N]} \mathrm{OPT}_i \quad (5.4)$$

be the welfare of bidder $i$ in auction $j$, total welfare contribution of bidder $i$, and total welfare, respectively, under the efficient outcome. Here we let $\mu_j(\ell) = 0$ for all $\ell > L_j$.

Despite truthful bidding is always feasible (see Definition 5.2.4 for feasible bid profiles and discussions on truthful bidding thereof), autobidders may not necessarily bid truthfully (even in truthful auctions) due to the presence of ROAS constraints, and instead adopt arbitrary strategies to optimize personal welfare which may deviate real auction outcomes from the efficient outcome. For certain individual bidders. Such deviations can potentially lead to significant welfare losses compared to the welfare she would have attained under the efficient one, whereas some other bidders may be significantly better off. These biases are not only largely unfavorable to bidders, but also to auction platforms as they may incentivize bidders to leave the platforms.

It is thus important for auction platforms to characterize to what extent the auctions ensure individual advertiser welfare, and better understand how individual welfare relates to advertiser strategies. In the following Definition 5.2.5, we present a individual welfare metric that measures how much a bidder's realized welfare can fall short from the welfare she would have obtained in the efficient outcome if she adopts a certain strategy.

**Definition 5.2.5** ($\delta$-approximate). *Fix a bidder $i \in [N]$ and her bids $\boldsymbol{b}_i \in \mathbb{R}_+^M$. Then we say auctions $(\mathcal{A}_j)_{j\in[M]}$ are $\delta$-approximate for some $\delta \in [0,1]$ and any bidder $i \in [N]$ if*

$$\min_{\boldsymbol{b}_{-i}\in\mathcal{F}_{-i}(\boldsymbol{b}_i)} \frac{\mathbb{E}\left[W_i(\mathcal{X}(\boldsymbol{b})\right]}{\mathrm{OPT}_i} \;\geq\; \delta\,, \tag{5.5}$$

*where $\mathcal{F}_{-i}(\cdot)$ is defined in Definition 5.2.4, the total welfare $W_i$ under outcome $\mathcal{X}(\boldsymbol{b})$ is defined in Eq. (5.1), and the expectation is taken w.r.t. possible randomness in the allocation and payment rules of auctions $(\mathcal{A}_j)_{j\in[M]}$.*

The above individual welfare metric provides a quantitative answer to the following question: fixing a bid profile $\boldsymbol{b}_i$ for bidder $i$, among all outcomes induced by competing bid profiles $\boldsymbol{b}_{-i} \in \mathcal{F}_{-i}(\boldsymbol{b}_i)$ that ensure every bidders' ROAS constraint is satisfied (see Definition 5.2.4), what proportion of the welfare under the efficient outcome can be retained under the worst case outcome? We remark that under this interpretation, our individual welfare metric is reminiscent of the notion of price of anarchy (POA) which measures the worst-case total welfare achieved amongst all equilibrium compared to the optimal total welfare (see e.g. [100] for a detailed introduction on POA).

## 5.3 Incorporating ML advice for bidder values as personalized reserve prices

With modern machine learning (ML) models and frameworks, online ad platforms can utilize available historical data (e.g. bid logs, keyword characteristics, user profiles, etc.) to produce predictions on autobidders' values (which we refer to as ML advice) ; see e.g. [99, 103]. In this work, we specifically focus on ML advice that take the form of a *lower-confidence bound* of true advertiser values. Our key approach to incorporate this type of ML advice in our autobidding setting, is via simply setting personalized reserve prices to be the lower confidence bound for each bidder's value. To motivate this approach, we start with an example.

### 5.3.1 Motivating Example

Consider 2 bidders competing in two (single-slot) second-price auctions (i.e. $L_1 = L_2 = 1$) with corresponding CTRs $\mu_1(1) = \mu_2(1) = 1$. Bidders' values are indicated in the following table with some $v > 0$.

| | Auction 1 | Auction 2 |
|---|---|---|
| bidder 1 | $v_{1,1} = v$ | $v_{1,2} = 0$ |
| bidder 2 | $v_{2,1} = \frac{v}{2}$ | $v_{2,2} = v$ |

Suppose that both bidders are autobidders who adopt uniform bidding strategies (see Proposition 5.2.1), and in particular, suppose that bidder 1 sets her bid multiplier to be $\alpha_1 = 1$. Then when her competitor bidder 2 sets a multiplier $\alpha_2 > 2$, bidder 2 will win both auctions and acquire a total value/welfare of $v_{2,1} + v_{2,2} = \frac{3}{2}v$ while submitting a payment of $\alpha_1(v_{1,1} + v_{1,2}) = v$. In this case, bidder 2 satisfies her ROAS constraint and extracts all bidder 1's welfare, leaving her with no value. We also highlight that this bid mulitplier profile constitutes an equilibrium, [4] because bidder 1 cannot raise her bid multiplier to outbid bidder 2 for auction 1, since with $\alpha_2 > 2$ bidder 1 would violate her ROAS constraint if she bids more than $\alpha_2 v_{2,1} > v$.

Now suppose that for each value $v_{i,j}$ $(i, j \in [2])$, the platform possesses a lower-confidence type of ML advice, namely $(\underline{v}_{i,j})_{i,j\in[N]}$ such that $\beta v_{i,j} \leq \underline{v}_{i,j} < v_{i,j}$ for all $v_{i,j} > 0$ for some $\beta > \frac{1}{2}$, and sets personalized reserve price $r_{i,j} = \underline{v}_{i,j}$. If bidder 2 attempts to win both auctions by setting $\alpha_2 > 2$, her payment will be at least $\max\{\beta v_{2,1}, \alpha_1 v_{1,1}\} + \max\{\beta v_{2,2}, \alpha_1 v_{1,2}\} = v + \beta v > \frac{3}{2}v$, violating her ROAS constraint. Therefore, via setting personalized reserves with ML advice, bidder 1's competitor

[4]At an equilibrium bid multiplier profile, every bidder best responds to other bidders' bid multipliers while maintaining ROAS constraint satisfaction.

is prohibited from outbidding her in auction 1, and hence safeguarding bidder 1's welfare.

**Key takeaway from Example 5.3.1.** The main observation from the above example is that without reserve prices, bidder 2 acquires a large margin for her ROAS constraint by winning auction 2 where payment required to win the auction is low. Therefore, she can raise her bid to outbid bidder 1 in auction 1 without violating her overall ROAS constraint. In other words, bidder 2 can compensate the high expenditure in auction 1 with her acquired value margin in auction 2. By setting personalized reserve prices properly, the platform can increase bidder 2's payment in auction 2, which in turn decreases the manipulative power of bidder 2. More generally, without reserve prices, bidders with large total values across auctions can overbid and consequently manipulate the outcome of certain auctions by compensating the incurring costs with acquired values from other auctions. Therefore, setting personalized reserve prices makes such overbidding behavior more costly, and thus reduces the overall manipulative power of bidders.

### 5.3.2 Setting personalized reserve prices using ML advice

Here, we focus on the following notion of *approximate reserve prices* with which we can reduce bidders' manipulative power as examplified in Example 5.3.1.

**Definition 5.3.1** ($\beta$-accurate ML advice and approximate reserve prices)**.** *Suppose there exists ML advice* $(\underline{v}_{i,j})_{i,j\in[N]}$ *in the form a lower-confidence bound. If* $\underline{v}_{i,j} \in [\beta v_{i,j}, v_{i,j})$ *with some* $\beta \in (0,1)$ *for any bidder* $i \in [N]$ *and auction* $j \in [M]$*, we say the ML advice is* $\beta$*-accurate. Further, if the platform sets* $r_{i,j} = \underline{v}_{i,j}$*, we say reserve prices* $\boldsymbol{r}$ *are* $\beta$*-approximate.*[5]

The gap between the lower bound $\beta v_{i,j}$ and the true value $v_{i,j}$ in Definition (5.3.1) represents the inaccuracies of the platform's ML advice. In other words, $\beta$ can be perceived as a quality measure of the platform's ML advice for advertiser value, such that larger $\beta$ represents better advice quality.

[5] Note that any $\beta$-approximate reserve prices are also $\beta'$-approximate if $\beta' < \beta$.

Further, ML-advice in online advertising settings generally concerns predicting advertiser values with historical conversion data and produces confidence intervals of advertiser values (see e.g. [101, 24, 72, 31]). We remark that these confidence intervals can be viewed as a special case of the lower-confidence type of ML advice in Definition (5.3.1): suppose the auctioneers utilize some ML model to predict the true value $v_{i,j}$ of bidder $i$ in auction $j$, and produce a confidence interval $(\underline{v}_{i,j}, \bar{v}_{i,j}) \ni v_{i,j}$ with $\underline{v}_{i,j}, \bar{v}_{i,j} > 0$. The auctioneer can then choose personalized reserve $r_{i,j} = \underline{v}_{i,j}$, which is $\beta$-approximate for $\beta = \underline{v}_{i,j}/\bar{v}_{i,j} \in (0, 1)$ because $\beta v_{i,j} < \beta \bar{v}_{i,j} = \underline{v}_{i,j} = r_{i,j} < v_{i,j}$.

Furthermore, in Definition 5.3.1, it is assumed that the ML advice $\underline{v}_{i,j}$ is a true lower bound on the bidder $i$'s value in auction $j$. This assumption can be relaxed by considering the ML advice that are accurate with high probability. Suppose the we possess some prediction $\hat{v}_{i,j}$ for $v_{i,j}$ that satisfies $|\hat{v}_{i,j} - v_{i,j}| < \eta$ with high probability (w.h.p) for some known $\eta$, then the confidence interval $(\hat{v}_{i,j} - \eta, \hat{v}_{i,j} + \eta)$ contains $v_{i,j}$ w.h.p. Them, the platform can set personalized reserve $r_{i,j} = \hat{v}_{i,j} - \eta$. Note that with such personalized reserve prices derived from probabilistic ML-advice, all results in this paper remain valid w.h.p.

We conclude with a final remark regarding allocation-anonymous auctions (Definition 5.2.2).

**Remark 5.3.1.** *We remark that implementing $\beta$-approximate personalized reserve prices in allocation-anonymous auctions does not impact anonymity, because $\beta < 1$ and thus all bidders clear there reserves. Therefore, the outcome with personalized reserves will be exactly the same as that without reserves; recall augmenting allocation-anonymous auctions with personalized reserves in Definition 5.2.3.*

## 5.4 Individual welfare guarantees for VCG with ML advice

In the motivating Example 5.3.1, we observe that ML advice and corresponding $\beta$-approximate reserves allow the parallel auctions to safeguard welfare for individual bidders by increasing payments and consequently limit the manipulative behavior of bidders who face significantly small competition in certain auctions. In this section, through the following Theorem 5.4.1, we formalize this intuition for the classic VCG auction, and present a quantitative measure for the relationship between overall individual welfare and ML advice when incorporated in the form of approximate reserves.

**Theorem 5.4.1** (Individual welfare lower bound for VCGs with $\beta$-approximate reserves)**.** *Consider the setting where $(\mathcal{A}_j)_{j\in[M]}$ are VCG auctions, and personalized reserve prices $\boldsymbol{r}$ are $\beta$-approximate as in Definition 5.3.1. Fix an autobidder $i \in [K]$ who adopts bid multiplier $\alpha_i > 1$ (see Proposition 5.2.1) so $\boldsymbol{b}_i = \alpha_i \boldsymbol{v}_i$. Then, the individual welfare guarantee in Definition 5.2.5) is bounded as:*

$$\min_{\boldsymbol{b}_{-i}\in\mathcal{F}_{-i}(\alpha_i\boldsymbol{v}_i)} \frac{W_i(\mathcal{X}(\boldsymbol{b}))}{\mathrm{OPT}_i} \;\geq\; 1 - \frac{1-\beta}{\alpha_i - 1}\cdot\frac{\mathrm{OPT}_{-i}}{\mathrm{OPT}_i}\,,$$

*where $\mathcal{F}_{-i}(\cdot)$ is defined in Definition 5.2.4, and $\mathrm{OPT}_{-i} = \sum_{j\neq i}\mathrm{OPT}_j$.*[6]

Details on implementation of VCG with personalized reserve prices can be found in Definition 5.2.3 and Example D.1.1. We defer our proof for Theorem 5.4.1 to Section 5.4.1, and here we provide some intuition for the individual welfare bound in the theorem.

In light of our interpretation for the individual welfare metric in Definition 5.2.5, Theorem 5.4.1 states that at some bid multiplier $\alpha_i$, among all outcomes induced by competing bid profiles (possibly arbitrary non-uniform bidding profiles) that ensure every bidders' ROAS constraint is satisfied, in the worst case outcome, bidder $i$ can retain at least a $1 - \frac{1-\beta}{\alpha_i-1}\cdot\frac{\mathrm{OPT}_{-i}}{\mathrm{OPT}_i}$ portion of the welfare she would have obtained under the efficient outcome. We also remark that this individual welfare bound is very general since it does not impose any assumptions on competing bidders' bidding strategies and concerns any feasible general bid profiles under which bidders' ROAS constraint is satisfied. In other words, even if bidder $i$'s competing bidders optimize

[6]We remark that the individual welfare lower bound in Theorem 5.4.1 applies only to bidders whose welfare under the efficient outcome is nonnegative, i.e. $\mathrm{OPT}_i > 0$.

arbitrary objectives through complex bidding strategies, Theorem 5.4.1 holds valid as long as the resulting bid profile is feasible.

The key message from Theorem 5.4.1 is that with more accurate ML advice (i.e. larger $\beta$), auctions can set larger approximate reserves, and hence improve individual welfare guarantees for each individual bidder. We also provide some intuition for the term $\frac{1-\beta}{\alpha_i-1}\frac{\mathrm{OPT}_{-i}}{\mathrm{OPT}_i}$ in the bound. Increasing $\beta$ (i.e. increasing reserve prices via improving ML quality) or increasing the bid multiplier $\alpha_i$, raises the cost for competitors to outbid bidder $i$ in certain auctions, and hence makes it more difficult to cover her expenditures that arise from significant overbidding. This reduces competitors' manipulative power, and in turn improves the welfare guarantees for bidder $i$. Note that this aligns with the intuition we obtained in Example 5.3.1. On the other hand, $\frac{\mathrm{OPT}_i}{\mathrm{OPT}_{-i}}$ can be perceived as the relative market share of bidder $i$ w.r.t. competing bidders. Our result shows that with a small market share, the bidders become vulnerable to manipulative behavior of others, resulting in low individual welfare guarantees.

The following Corollary 5.4.2 presents a sufficient condition for the ML advice accuracy to achieve some predesignated level of individual welfare, whereas the following Theorem 5.4.3 states the individual welfare bound in Theorem 5.4.1 is tight; see Appendix D.2.1 for corresponding proofs.

**Corollary 5.4.2** (ML advice accuracy level to achieve $\delta$-approximation of individual welfare)**.** *Let $(\mathcal{A}_j)_{j\in[M]}$ be VCG auctions, and assume all bidders adopt uniform bidding strategies with bid multipliers $(\alpha_i)_{i\in[N]} \in (1,\infty)^N$, where we denote $\underline{\alpha} = \min_{i\in[N]} \alpha_i > 1$. Then, with $\beta$ accurate ML advice such that $\beta \geq 1-(1-\delta)\cdot(\underline{\alpha}-1)\cdot\min_{i\in[K]}\frac{\mathrm{OPT}_i}{\mathrm{OPT}_{-i}}$, the auctions $(\mathcal{A}_j)_{j\in[M]}$ are $\delta$-approximate for all bidders.*

**Theorem 5.4.3** (Matching individual welfare lower bound)**.** *For any $\beta\in(0,1)$, $\alpha>1$, and $R \geq \frac{1-\beta}{\alpha-1}$, there exists values $\boldsymbol{v}\in\mathbb{R}_+^{N\times M}$ and $\beta$-approximate reserves $\boldsymbol{r}\in\mathbb{R}_+^{N\times M}$, such that there is a bidder $i$ with multiplier $\alpha_i=\alpha$ and relative market share $\frac{\mathrm{OPT}_i}{\mathrm{OPT}_{-i}}=R$, who has an individual welfare guarantee $\min_{\boldsymbol{b}_{-i}\in\mathcal{F}_{-i}(\alpha_i\boldsymbol{v}_i)}\frac{W_i(\mathcal{X}(\boldsymbol{b}))}{\mathrm{OPT}_i} = 1-\frac{1-\beta}{\alpha_i-1}\cdot\frac{\mathrm{OPT}_{-i}}{\mathrm{OPT}_i}$.*

We conclude by comparing our individual welfare result in Theorem 5.4.1 with related results in the literature: we point out that although our autobidding setup described in Section 5.2 and the notion of approximate reserves (Definition 5.3.1) are the same as those in [37, 11], our analyses and proof techniques are different, primarily because we focus on the welfare guarantees for individual bidders, where as [37, 11] investigates total welfare for all bidders. In particular, in our proof we fix a bidder $i$ and carefully analyze the amount of expenditure that could be covered by each competitor who outbids bidder $i$ in auctions where $i$ has the highest value, whereas the aforementioned related works takes an aggregate view to lower bound total welfare of all bidders; see more details in the next subsection 5.4.1. Nevertheless, [11] shows that approximate reserves improve the total welfare of all bidders, and therefore along with Theorem 5.4.1, we can see that incorporating $\beta$-accurate ML advice as approximate reserves not only benefits total welfare, but also enhances individual welfare.

### 5.4.1 Proof for Theorem 5.4.1 and a tighter individual welfare guarantee

In this subsection, we first present the proof for the individual welfare lower bound in Theorem 5.4.1. From the proof, we further motivate a stronger individual welfare lower bound that depends on the total welfare of at most $\min\{M, N-1\}$ of bidder $i$'s competitors, instead of $\mathrm{OPT}_{-i}$ which sums up the welfare of all $N-1$ of bidder $i$'s competitors and can potentially be enormous due to large $N$. We will later remark that this strengthened bound, despite its improvement to the bound in Theorem 5.4.1, may be difficult to compute in practice.

#### Proof for Theorem 5.4.1

First, to prove Theorem 5.4.1, we rely on the definition of an advertisers' loss in welfare compared to her welfare contribution under the efficient outcome, formally defined as followed:

**Definition 5.4.1** (Welfare loss w.r.t. efficient outcome). *For any bidder $i \in [N]$ and outcome $\boldsymbol{x} = (\boldsymbol{x}_j \in \{0,1\}^{N\times L_j})_{j\in[M]}$, let $\mathcal{L}_i(\boldsymbol{x}) = \{j \in [M] : W_{i,j}(\boldsymbol{x}) < \mathrm{OPT}_{i,j}\}$ be the set of auctions in which bidder $i$'s acquired welfare is less than that of her welfare* *under the efficient outcome. Then, we define the welfare loss of bidder $i$ under outcome $\boldsymbol{x}$ w.r.t. the efficient outcome $\boldsymbol{x}^*$ as:*

$$\text{LOSS}_i(\boldsymbol{x}) = \sum_{j \in \mathcal{L}_i(\boldsymbol{x})} \left(\text{OPT}_{i,j} - W_{i,j}(\boldsymbol{x})\right) . \tag{5.6}$$

**Remark 5.4.1.** *For any outcome $\boldsymbol{x}$, let $\ell_{i,j}$ be the position (i.e. ranking) of bidder $i$ in auction $j$, and recall that $\ell^*_{i,j}$ is the position of bidder $i$ in auction $j$ under the efficient outcome $\boldsymbol{x}^*$. Then, the set $\mathcal{L}_i(\boldsymbol{x}) = \{j \in [M] : W_{i,j}(\boldsymbol{x}_j) < \text{OPT}_{i,j}\}$ (where $W_{i,j}(\boldsymbol{x}_j)$ is bidder $i$'s welfare in $\mathcal{A}_j$ as defined in Eq.(5.1)) can also be interpreted as the set of auctions where bidder $i$'s ranking under $\boldsymbol{x}$ is lower than her ranking under $\boldsymbol{x}^*$, or in other words the set of auctions that incur a welfare loss w.r.t. $\boldsymbol{x}^*$. Hence we can also rewrite $\mathcal{L}_i(\boldsymbol{x}) = \{j \in [M] : \ell_{i,j} > \ell^*_{i,j}\}$.*

The following proposition connects the notion of welfare loss (as in Definition 5.4.1) and individual welfare (as in Definition 5.2.5) by showing an upper bound on welfare loss can be directly translated into a welfare lower bound that corresponds to our individual welfare guarantee.

**Proposition 5.4.4** (Translating loss to individual welfare guarantee)**.** *Assume for bidder $i \in [N]$ and outcome $\boldsymbol{x} = (\boldsymbol{x}_j \in \{0,1\}^{N \times L_j})_{j \in [M]}$ we have $\text{LOSS}_i(\boldsymbol{x}) \leq B$ for some $B > 0$. Then, $\frac{W_i(\boldsymbol{x})}{\text{OPT}_i} \geq 1 - \frac{B}{\text{OPT}_i}$.*

The proof of this proposition is presented in Section D.2.2. Now, in light of this proposition, we proceed to prove Theorem 5.4.1 by bounding bidder $i$'s welfare loss for auctions where she obtains a slot that is lower in position than what she would have obtained under the efficient outcome.

*Proof of Theorem 5.4.1.* Fix any feasible competing bid profile $\boldsymbol{b}_{-i} \in \mathcal{F}_{-i}(\alpha_i \boldsymbol{v}_i)$ under which every bidders' ROAS constraint is satisfied; see Definition 5.2.4. Denote the corresponding outcome as $\boldsymbol{x} = \mathcal{X}(\boldsymbol{b})$, and $\ell_{k,j}$, $\ell^*_{k,j}$ to be the position of any bidder $k \in [N]$ in auction $j \in [M]$ under outcome $\boldsymbol{x}$ and the efficient outcome, respectively.

Consider any auction $j \in \mathcal{L}_i(\boldsymbol{x}) = \{j \in [M] : \ell_{i,j} > \ell^*_{i,j}\}$ (see Remark 5.4.1), i.e. in auction $\mathcal{A}_j$, bidder $i$ acquires a position (under $\boldsymbol{x}$) bellow her position in the efficient outcome $\boldsymbol{x}^*$. This implies there must exist competing bidders in auction $\mathcal{A}_j$ whose values are smaller than that of bidder $i$'s, but obtains a higher position, making bidder $i$ lose welfare. Motivated by this, we let $\mathcal{B}_i(k;\boldsymbol{x})$ denote the set of all auctions in which bidder $k$'s value is lower than $i$'s but acquires a higher position than $i$:

$$\mathcal{B}_i(k;\boldsymbol{x}) = \left\{j \in [M] : \mathrm{OPT}_{i,j} > 0,\ v_{k,j} < v_{i,j} \text{ and } \ell_{k,j} \leq \ell^*_{i,j} < \ell_{i,j}\right\} \tag{5.7}$$

where we recall $\mathrm{OPT}_{i,j}$ is the welfare of bidder $i$ in auction $j$ under the efficient outcome. Further, we can find a collection of $i$'s competitors whose $\mathcal{B}_i(\ \cdot\ ;\boldsymbol{x})$ "covers" all auctions $\mathcal{L}_i(\boldsymbol{x})$ in which $i$ loses welfare. We call this collection of competitors a covering, and formally define the collection of all coverings, called $\mathcal{C}_i(\boldsymbol{x})$, as followed:

$$\mathcal{C}_i(\boldsymbol{x}) = \{\mathcal{C} \subseteq [N]/\{i\} : (\mathcal{B}_i(k;\boldsymbol{x}))_{k\in\mathcal{C}} \text{ is a maximal set cover of } \mathcal{L}_i(\boldsymbol{x})\}\ . \tag{5.8}$$

Here, for any set $\mathcal{S}$, we say $\mathcal{S}_1 \dots \mathcal{S}_n$ a maximal set cover of $\mathcal{S}$ if $\mathcal{S} \subseteq \bigcup_{n'\in[n]} \mathcal{S}_{n'}$ but $\mathcal{S} \subsetneq \bigcup_{n'\in[n]} \mathcal{S}_{n'}/\mathcal{S}_{n''}$ for any $n'' \in [n]$. In words, $\mathcal{B}_i(k;\boldsymbol{x})$ is the set of auctions in which bidder $k$ has a smaller value than bidder $i$ but acquires a higher position, and any $\mathcal{C} \in \mathcal{C}_i(\boldsymbol{x})$ is a subset of $i$'s competitors who are responsible for all welfare losses of bidder $i$ in auctions of $\mathcal{L}_i(\boldsymbol{x})$.

Fix any covering $\mathcal{C} \in \mathcal{C}_i(\boldsymbol{x})$, and some bidder $k \in \mathcal{C}$. We first state the following inequality that bounds the welfare loss of bidder $i$ caused by competitor $k \in \mathcal{C}$ in the covering (we will prove this inequality later).

$$\sum_{j\in\mathcal{B}_i(k;\boldsymbol{b})} \left(\mu(\ell^*_{i,j}) - \mu(\ell_{i,j})\right) v_{i,j} \leq \frac{1-\beta}{\alpha_i - \beta} \sum_{j\in[M]} \mu(\ell_{k,j}) v_{k,j} \tag{5.9}$$

Summing the above over all competitors $k \in \mathcal{C}$, we have

$$\begin{aligned}
\textsc{loss}_i(\boldsymbol{x}) &= \sum_{j\in\mathcal{L}_i(\boldsymbol{x})} \left(\mu(\ell^*_{i,j}) - \mu(\ell_{i,j})\right) v_{i,j} \overset{(a)}{\leq} \sum_{k\in\mathcal{C}} \sum_{j\in\mathcal{B}_i(k;\boldsymbol{b})} \left(\mu(\ell^*_{i,j}) - \mu(\ell_{i,j})\right) v_{i,j} \\
&\overset{(b)}{\leq} \frac{1-\beta}{\alpha_i-\beta} \sum_{k\in\mathcal{C}} \sum_{j\in[M]} \mu(\ell_{k,j}) v_{k,j} = \frac{1-\beta}{\alpha_i-\beta} \sum_{k\in\mathcal{C}} W_k(\boldsymbol{x}) \\
&\leq \frac{1-\beta}{\alpha_i-\beta} W_{-i}(\boldsymbol{x}) \\
&\overset{(c)}{\leq} \frac{1-\beta}{\alpha_i-\beta} \left(\mathrm{OPT}_{-i} + \textsc{loss}_i(\boldsymbol{x})\right) \\
\Longrightarrow \textsc{loss}_i(\boldsymbol{x}) &\leq \frac{1-\beta}{\alpha_i-1} \mathrm{OPT}_{-i}\,.
\end{aligned} \tag{5.10}$$

Here, in (a) we used the fact that $\mathcal{L}_i(\boldsymbol{x}) \subseteq \bigcup_{k\in\mathcal{C}} \mathcal{B}_i(k;\boldsymbol{x})$ (see Eq. (5.8)); in (b) we applied Eq. (5.9); (c) follows from $\mathrm{OPT} \geq \sum_{i\in[N]} W_i(\boldsymbol{x})$ where OPT is the total efficient welfare and $\sum_{i\in[N]} W_i(\boldsymbol{x})$ is the total welfare under outcome $\boldsymbol{x}$, so further

$$\begin{aligned}
\mathrm{OPT}_{-i} &\geq W_{-i}(\boldsymbol{x}) + W_i(\boldsymbol{x}) - \mathrm{OPT}_i \\
&= W_{-i}(\boldsymbol{x}) + \sum_{j\in\mathcal{L}_i(\boldsymbol{x})} \left(W_{i,j}(\boldsymbol{x}) - \mathrm{OPT}_{i,j}\right) + \sum_{j\in[M]/\mathcal{L}_i(\boldsymbol{x})} \left(W_{i,j}(\boldsymbol{x}) - \mathrm{OPT}_{i,j}\right) \\
&\overset{(e)}{\geq} W_{-i}(\boldsymbol{x}) + \sum_{j\in\mathcal{L}_i(\boldsymbol{x})} \left(W_{i,j}(\boldsymbol{x}) - \mathrm{OPT}_{i,j}\right) \\
&= W_{-i}(\boldsymbol{x}) - \textsc{loss}_i(\boldsymbol{x})\,.
\end{aligned} \tag{5.11}$$

where in (e) we used the fact that $W_{i,j}(\boldsymbol{x}) \geq \mathrm{OPT}_{i,j}$ in any auction $j \in [M]/\mathcal{L}_i(\boldsymbol{x})$. Finally, applying Proposition 5.4.4 w.r.t. upper bound of $\textsc{loss}_i(\boldsymbol{x})$, and noting that the feasible competing bid profile is arbitrary, we obtain the desired welfare guarantee lower bound.

Now, it remains to prove Eq. (5.9) that bounds the welfare loss of bidder $i$ caused by competitor $k \in \mathcal{C}$ in the covering. Denote $p_{k,j}$ as the payment of bidder $k$, and $\hat{b}_{\ell,j}$ as the $\ell$th largest bid in any auction $j \in [M]$. Then in some auction $j \in \mathcal{B}_i(k;\boldsymbol{b})$ recall from Eqs. (5.7) and (5.8) that $v_{k,j} < v_{i,j}$ but $\ell_{k,j} \leq \ell^*_{i,j} < \ell_{i,j}$. Thus bidder $k$'s payment is lower bounded as followed: for $j \in \mathcal{B}_i(k; \boldsymbol{b})$

$$
\begin{aligned}
& p_{k,j} \\
\geq\ & \sum_{\ell=\ell_{k,j}}^{L_j} \left(\mu(\ell) - \mu(\ell+1)\right) \hat{b}_{\ell+1,j} \\
\overset{(a)}{=}\ & \sum_{\ell=\ell_{k,j}}^{\ell^*_{i,j}-1} \left(\mu(\ell) - \mu(\ell+1)\right) \hat{b}_{\ell+1,j} + \sum_{\ell=\ell^*_{i,j}}^{\ell_{i,j}-1} \left(\mu(\ell) - \mu(\ell+1)\right) \hat{b}_{\ell+1,j} + p_{i,j} \\
\overset{(b)}{\geq}\ & \left(\mu(\ell_{k,j}) - \mu(\ell^*_{i,j})\right) v_{i,j} + \alpha_i \left(\mu(\ell^*_{i,j}) - \mu(\ell_{i,j})\right) v_{i,j} + \beta \cdot \mu(\ell_{i,j}) v_{i,j} \\
=\ & \mu(\ell_{k,j}) v_{i,j} + (\alpha_i - 1) \left(\mu(\ell^*_{i,j}) - \mu(\ell_{i,j})\right) v_{i,j} - (1-\beta) \cdot \mu(\ell_{i,j}) v_{i,j}\,.
\end{aligned}
\tag{5.12}
$$

Here , (a) follows from the VCG payment rule (see Example D.1.1); (b) follows from the fact that bidder $i$'s ranking is $\ell_{i,j}$, so any bidder who is ranked before position $\ell_{i,j}$ submits a bid greater than bidder $i$'s bid $b_{i,j} = \alpha_i v_{i,j}$, i.e. $\hat{b}_{\ell,j} \geq b_{i,j} = \alpha_i v_{i,j} > v_{i,j}$ for any $\ell \leq \ell_{i,j}$.

On the other hand, we have

$$
\sum_{j \in \mathcal{B}_i(k;\boldsymbol{b})} p_{k,j} + \sum_{j \notin \mathcal{B}_i(k;\boldsymbol{b})} p_{k,j} \leq \sum_{j \in \mathcal{B}_i(k;\boldsymbol{b})} \mu(\ell_{k,j}) v_{k,j} + \sum_{j \notin \mathcal{B}_i(k;\boldsymbol{b})} \mu(\ell_{k,j}) v_{k,j}
$$

$$
p_{k,j} \geq \beta \cdot \mu(\ell_{k,j}) v_{k,j} \quad \forall j \in [M]\,,
$$

where the first inequality follows from bidder $k$'s ROAS constraint; the second inequality follows from the fact that any winning bidder's payment must be greater than her $\beta$-approximate reserves. Combining the above inequalities and rearranging we get

$$
\sum_{j \in \mathcal{B}_i(k;\boldsymbol{b})} p_{k,j} \leq \sum_{j \in \mathcal{B}_i(k;\boldsymbol{b})} \mu(\ell_{k,j}) v_{k,j} + (1-\beta) \cdot \sum_{j \notin \mathcal{B}_i(k;\boldsymbol{b})} \mu(\ell_{k,j}) v_{k,j}\,, \tag{5.13}
$$

Summing Eq.(5.12) over all $j \in \mathcal{B}_i(k;\boldsymbol{b})$ and combining with Eq. (5.13), we get

$$\begin{aligned}
&(\alpha_i - 1) \cdot \sum_{j \in \mathcal{B}_i(k;\boldsymbol{b})} \left(\mu(\ell^*_{i,j}) - \mu(\ell_{i,j})\right) v_{i,j} \\
\leq\ & (1-\beta) \cdot \left( \sum_{j \in \mathcal{B}_i(k;\boldsymbol{b})} \mu(\ell_{i,j}) v_{i,j} + \sum_{j \notin \mathcal{B}_i(k;\boldsymbol{b})} \mu(\ell_{k,j}) v_{k,j} \right) + \sum_{j \in \mathcal{B}_i(k;\boldsymbol{b})} \mu(\ell_{k,j}) \left(v_{k,j} - v_{i,j}\right) \\
\overset{(a)}{\leq}\ & (1-\beta) \cdot \left( \sum_{j \in \mathcal{B}_i(k;\boldsymbol{b})} \mu(\ell_{i,j}) v_{i,j} + \sum_{j \notin \mathcal{B}_i(k;\boldsymbol{b})} \mu(\ell_{k,j}) v_{k,j} + \sum_{j \in \mathcal{B}_i(k;\boldsymbol{b})} \mu(\ell_{k,j}) \left(v_{k,j} - v_{i,j}\right) \right) \\
\overset{(b)}{\leq}\ & (1-\beta) \cdot \left( \sum_{j \in \mathcal{B}_i(k;\boldsymbol{b})} \mu(\ell_{i,j}) v_{i,j} + \sum_{j \in [M]} \mu(\ell_{k,j}) v_{k,j} - \sum_{j \in \mathcal{B}_i(k;\boldsymbol{b})} \mu(\ell^*_{i,j}) v_{i,j} \right) .
\end{aligned}$$

In (a), we used the fact that $\beta \in (0,1]$ and $v_{k,j} - v_{i,j} < 0$ for any $k \in \mathcal{C} \subseteq \mathcal{C}_i(\boldsymbol{x})$; see definition of $\mathcal{C}_i(\boldsymbol{x})$ in Eq. (5.8); and (b) follows from $\ell_{k,j} \leq \ell^*_{i,j}$ for any $k \in \mathcal{C} \subseteq \mathcal{C}_i(\boldsymbol{x})$. Rearranging terms we obtain the desired Eq. (5.9). $\square$

**A tighter fariness lower bound guarantee**

We recognize that the individual welfare lower bound guarantee in Theorem 5.4.1, namely $1 - \frac{1-\beta}{\alpha_i - 1} \cdot \frac{\mathrm{OPT}_{-i}}{\mathrm{OPT}_i}$, may be negative and hence meaningless for a small advertiser, i.e. advertiser $i$ whose market share $\frac{\mathrm{OPT}_i}{\mathrm{OPT}_{-i}}$ is very small which may be a result of a very large number of bidders $N$. Nevertheless, in light of the proof above for Theorem 5.4.1, we can in fact present a tighter individual welfare guarantee that replaces $\mathrm{OPT}_{-i}$ in the numerator, i.e. total welfare summed over $N-1$ competitors of bidder $i$, by the total welfare of a potentially much smaller subset of bidder $i$'s competitors.

Analogous to the definitions $\mathcal{L}_i(\boldsymbol{x}), \mathcal{B}_i(k;\boldsymbol{x})$ and $\mathcal{C}_i(\boldsymbol{x})$ for any outcome $\boldsymbol{x}$ defined in Definition 5.4.1, Eqs. (5.7) and (5.8), respectively, we slightly abuse notation and define their outcome-independent counterparts.

$$\begin{aligned}
&\mathcal{L}_i = \{j \in [M] : \mathrm{OPT}_{i,j} > 0\} \\
&\mathcal{B}_i(k) = \{j \in [M] : \mathrm{OPT}_{i,j} > 0,\ v_{k,j} < v_{i,j}\} \\
&\mathcal{C}_i = \{\mathcal{C} \subseteq [N]/\{i\} : (\mathcal{B}_i(k))_{k \in \mathcal{C}} \text{ is a maximal set cover of } \mathcal{L}_i\} \ ,
\end{aligned} \tag{5.14}$$

where we recall $\mathrm{OPT}_{i,j}$ is the welfare of bidder $i$ in auction $\mathcal{A}_j$ under the efficient outcome as defined in Eq. (5.4). In words, $\mathcal{L}_i$ is the set of auctions in which bidder $i$ can potentially lose welfare, $\mathcal{B}_i(k)$ is the set of auctions in which competitor $k$ can potentially cause $i$ to lose welfare, and any covering of competitors $\mathcal{C} \in \mathcal{C}_i$ can potentially cause $i$ to lose welfare in all auctions of $\mathcal{L}_i$.

We remark that $\mathcal{C}_i$ only depends on bidder values $(v_{i,j})_{i\in[N],j\in[M]}$, and it is easy to see that any covering $\widetilde{\mathcal{C}} \in \mathcal{C}_i$ has cardinality at most $\min\{M, N-1\}$. To exemplify the covering set $\mathcal{C}_i$, consider an instance consisting of 2 single slot VCG auctions (i.e. SPA) and 3 bidders with the following advertiser values

| | SPA 1 | SPA 2 | SPA 3 |
|---|---|---|---|
| bidder 1 | $v_{1,1} = 2$ | $v_{1,2} = 5$ | $v_{1,3} = 0$ |
| bidder 2 | $v_{2,1} = 1$ | $v_{2,2} = 1$ | $v_{2,3} = 10$ |
| bidder 3 | $v_{3,1} = 0$ | $v_{3,2} = 4$ | $v_{3,3} = 10$ |

Then under the efficient outcome, bidder 1 wins auctions 1 and 2. However, it may be possible that bidder 2 solely outbids the other bidders to win both auctions 1 and 2, inducing a covering $\{2\}$ for bidder 1; or bidder 2 wins auction 1 while bidder 3 wins auction 2, inducing a covering $\{2,3\}$ for bidder 1. Therefore, the covering set $\mathcal{C}_i = \{\{2\},\{2,3\}\}$.

The following proposition states that any covering $\mathcal{C} \in \mathcal{C}_i(\boldsymbol{x})$ that contributes to all bidder $i$'s welfare loss under $\boldsymbol{x}$, must be a subset of an element of the set $\mathcal{C}_i$.

**Proposition 5.4.5.** *Let $\boldsymbol{x}$ be any outcome and denote $\mathcal{C}_i(\boldsymbol{x})$ be the corresponding set of coverings defined in Eq. (5.8). Then, for any covering $\mathcal{C} \in \mathcal{C}_i(\boldsymbol{x})$, there must exist an $\widetilde{\mathcal{C}} \in \mathcal{C}_i$ such that $\mathcal{C} \subseteq \widetilde{\mathcal{C}}$, where $\mathcal{C}_i$ is defined in Eq. (5.14).*

In the following theorem, we present a tighter individual welfare guarantee than that of Theorem 5.4.1 by replacing $\mathrm{OPT}_{-i}$ with the welfare of some covering for bidder $i$, which may potentially be a very small subset of all bidder $i$'s competitors that includes at most $\min\{M, N-1\}$ bidders.

**Theorem 5.4.6.** *Consider $(\mathcal{A}_j)_{j\in[M]}$ are VCG auctions, and personalized reserve prices $\boldsymbol{r}$ are $\beta$-approximate as in Definition 5.3.1. Fix an autobidder $i \in [K]$ who adopts bid multiplier $\alpha_i > 1$ (see Proposition 5.2.1) so $\boldsymbol{b}_i = \alpha_i \boldsymbol{v}_i$. Then, the individual welfare guarantee in Definition 5.2.5) is bounded as:*

$$\min_{\boldsymbol{b}_{-i}\in\mathcal{F}_{-i}(\alpha_i\boldsymbol{v}_i)} \frac{W_i(\mathcal{X}(\boldsymbol{b}))}{\mathrm{OPT}_i} \;\geq\; 1 - \frac{1-\beta}{\alpha_i - \beta}\cdot\frac{\max_{\widetilde{\mathcal{C}}\in\mathcal{C}_i}\sum_{k\in\widetilde{\mathcal{C}}} W_k(\mathcal{X}(\boldsymbol{v}_{\widetilde{\mathcal{C}}}, \mathbf{0}))}{\mathrm{OPT}_i}\,,$$

*where $\mathcal{F}_{-i}(\cdot)$ is defined in Definition 5.2.4, $\boldsymbol{v}_{\widetilde{\mathcal{C}}} = (\boldsymbol{v}_k)_{k\in\widetilde{\mathcal{C}}}$, and $\mathcal{X}(\boldsymbol{v}_{\widetilde{\mathcal{C}}}, \mathbf{0})$ is the outcome when bidders in covering $\widetilde{\mathcal{C}}$ bid truthfully while others bid 0; equivalently, this is the total efficient welfare when participation in all auctions are restricted to bidders in $\widetilde{\mathcal{C}}$ only.*

*Proof.* Let $\boldsymbol{b} \in \mathcal{F}$ be any feasible bid profile, and let $\boldsymbol{x} = \mathcal{X}(\boldsymbol{b})$ be the corresponding outcome. Also, let $\mathcal{C}_i(\boldsymbol{b})$ be the set of coverings defined in Eq. (5.8), and consider any $\mathcal{C} \in \mathcal{C}_i(\boldsymbol{x})$. In Eq. (5.10) within the proof of Theorem 5.4.1, we showed $\textsc{loss}_i(\boldsymbol{x}) \leq \frac{1-\beta}{\alpha_i - 1}\sum_{k\in\mathcal{C}} W_k(\boldsymbol{x})$, so

$$\begin{aligned}\textsc{loss}_i(\boldsymbol{x}) \;&\leq\; \frac{1-\beta}{\alpha_i-1}\sum_{k\in\mathcal{C}} W_k(\boldsymbol{x}) \;\overset{(a)}{\leq}\; \frac{1-\beta}{\alpha_i-1}\sum_{k\in\widetilde{\mathcal{C}}} W_k(\boldsymbol{x})\\ &\leq\; \frac{1-\beta}{\alpha_i-1}\max_{\widetilde{\mathcal{C}}\in\mathcal{C}_i}\sum_{k\in\widetilde{\mathcal{C}}} W_k(\mathcal{X}(\boldsymbol{v}_{\widetilde{\mathcal{C}}}, \mathbf{0}))\,,\end{aligned}$$

where in (a) we let $\widetilde{\mathcal{C}} \in \mathcal{C}_i$ defined in Eq. (5.14) such that $\mathcal{C} \subseteq \widetilde{\mathcal{C}}$ according Proposition 5.4.5. Rearranging and applying Proposition 5.4.4 yields the desired lower bound. □

We conclude by making the remark that although the individual welfare lower bound in Theorem 5.4.6 may potentially be stronger that in Theorem 5.4.1, it comes at the cost of significantly increased computational complexity due to the maximum over all coverings. Nevertheless, this improved bound may still be practical for instances with relatively small number of auctions, since the cardinality of any covering is upper bounder by the number of auctions.

### 5.4.2 Applicability of the individual welfare guarantee when all bidders bid uniformly

We recognize that as the individual welfare lower bound in Theorem 5.4.1 monotonically increases in the bid multiplier $\alpha_i$, it is tempting for bidder $i$ to apply a very large multiplier $\alpha_i$. Nevertheless, in this section we describe a potential tradeoff between large multipliers (i.e. better individual welfare guarantees in light of Theorem 5.4.1) and ROAS feasibility in the practical scenario where all bidders are autobidders and adopt uniform bidding.

To illustrate, we see that for large multiplier $\alpha_i$, the set of competing bids $\mathcal{F}_{-i}(\alpha_i \boldsymbol{v}_i)$ may only include very small bid values (e.g. the bid profile where each competing bidder (under)bids some small $\epsilon > 0$ close to 0 in each auction), at which bidder $i$ faces nearly no competition so that the ROAS constraint can be trivially satisfied for every bidder. In light of this discussion, we consider a more practical scenario where all competing bidders are also autobidders and adopt uniform bidding, or equivalently, a refinement of $\mathcal{F}_{-i}(\alpha_i \boldsymbol{v}_i)$ in which each competing bidder $j \neq i$, similar to bidder $i$, also adopts uniform bidding with bid multiplier $\alpha_j \geq 1$. We define $\mathcal{F}^u_{-i}(\boldsymbol{b}_i) = \mathcal{F}_{-i}(\boldsymbol{b}_i) \cap \{(\alpha_j \boldsymbol{v}_j)_{j \neq i} : \alpha_j \geq 1\}$ that represents the set of uniform competing bids for bidder $i$ that ensure ROAS constraint satisfaction for every bidder. From Theorem 5.4.1, it is easy to see

$$\min_{\boldsymbol{b}_{-i} \in \mathcal{F}^u_{-i}(\alpha_i \boldsymbol{v}_i)} \frac{W_i(\mathcal{X}(\boldsymbol{b}))}{\mathrm{OPT}_i} \overset{(i)}{\geq} 1 - \frac{1-\beta}{\alpha_i - 1} \cdot \frac{\mathrm{OPT}_{-i}}{\mathrm{OPT}_i}\,, \tag{5.15}$$

where (i) follows from $\min_{\boldsymbol{b}_{-i} \in \mathcal{F}^u_{-i}(\alpha_i \boldsymbol{v}_i)} \frac{W_i(\mathcal{X}(\boldsymbol{b}))}{\mathrm{OPT}_i} \geq \min_{\boldsymbol{b}_{-i} \in \mathcal{F}_{-i}(\alpha_i \boldsymbol{v}_i)} \frac{W_i(\mathcal{X}(\boldsymbol{b}))}{\mathrm{OPT}_i}$ because $\mathcal{F}^u_{-i}(\alpha_i \boldsymbol{v}_i) \subseteq \mathcal{F}_{-i}(\alpha_i \boldsymbol{v}_i)$. Nevertheless, in light of Eq. (5.15), when all bidders bid uniformly, an excessively large $\alpha_i$ may let bidder $i$ incur large payments that significantly exceed her values, resulting in non-existence of competing uniform bids $\boldsymbol{b}_{-i}$ that can ensure satisfaction of every bidders' ROAS constraints, i.e. $\mathcal{F}^u_{-i}(\alpha_i \boldsymbol{v}_i)$ being empty. In other words, there exists a tradeoff between large multipliers (i.e. better individual welfare guarantees) and ROAS feasibility when all bidders bid uniformly. The following

Lemma 5.4.7, along with a technical definition of "well-separated" values per Definition 5.4.2, addresses this tradeoff by characterizing how large the multiplier $\alpha_i$ can be set that still ensures the existence of uniform competing bids within $\mathcal{F}^u_{-i}(\alpha_i \boldsymbol{v}_i)$.

**Definition 5.4.2** ($\Delta$-separated values)**.** *We say values $\boldsymbol{v} \in \mathbb{R}^{N\times M}_{\geq 0}$ are $\Delta$-separated for some $\Delta > 1$ if any value $v_{i,j}$ is at least $\Delta$ times as much as any value that is less than $v_{i,j}$ in the same auction $j$, i.e. $v_{i,j} \geq \Delta \cdot \max\{v_{k,j} : k \in [N], v_{k,j} < v_{i,j}\}$ for any bidder $i$ and auction $j$.*[7]

**Lemma 5.4.7** (Valid regions for uniform bid multiplier)**.** *Let $(\mathcal{A}_j)_{j\in[M]}$ be VCG auctions and assume bidders values are $\Delta$-separated (Definition 5.4.2) in every auction for some $\Delta > 1$, then $\mathcal{F}^u_{-i}(\alpha_i \boldsymbol{v}_i) \neq \varnothing$ for any $\alpha_i \in [1, \Delta)$.*

The proof of this lemma is presented in Appendix D.2.3. We also remark that the upper bound $\Delta$ in Lemma 5.4.7 is sufficient, meaning that there may exist larger values of $\alpha_i$ that can ensure the set $\mathcal{F}^u_{-i}(\alpha_i \boldsymbol{v}_i) \neq \varnothing$ nonempty. To better visualize the structure of $\mathcal{F}^u_{-i}(\alpha_i \boldsymbol{v}_i)$, as well as our individual welfare guarantee in Theorem 5.4.1 and Eq. (5.15), we present the following example.

**Example 5.4.1.** *Consider 2 bidders bidding in 3 single-slot VCG auctions in which each slot is associated with CTR equal to 1. Bidder values are $\boldsymbol{v}_1 = (4, 3, 1)$ and $\boldsymbol{v}_2 = (1, 4, 3)$, while personalized reserves are set to be $\boldsymbol{r}_i = \beta \boldsymbol{v}_i$ for $\beta = 0.7$ and $i = 1, 2$. It is easy to check that with the presence of personalized reserves, no bidder can significantly overbid and win all auctions (otherwise she will incur large payments and thus violate their ROAS constraints), and therefore each bidder will obtain non-zero value. This aligns with our intuition presented in Sections 5.3 that states personalized reserves benefit individual welfare.*

*In the left subgraph of Figure 5-1, we color the region of all pairs of uniform bid multipliers $(\alpha_1, \alpha_2) \in [1, \infty)^2$ that induce feasible bid profiles $(\boldsymbol{b}_1, \boldsymbol{b}_2) \in \mathcal{F}$, where*

[7] Definition 5.4.2 also captures values which are "additively separated". In particular, take some $d > 0$ such that $d < \min\{v_{i,j} : v_{i,j} \neq 0\}$ and also $v_{i,j} - d \geq \max\{v_{k,j} : k \in [N], v_{k,j} < v_{i,j}\}$ for any bidder $i$ and auction $j$. Then, by taking $\Delta \in \min_{v_{i,j}:v_{i,j}\neq 0} \left\{\frac{v_{i,j}}{v_{i,j}-d}\right\}$, the values are $\Delta$-separated according to Definition 5.4.2 because $\frac{1}{\Delta} v_{i,j} \geq v_{i,j} - d \geq \max\{v_{k,j} : k \in [N], v_{k,j} < v_{i,j}\}$ for all $v_{i,j}$. This suggests Definition 5.4.2 is quite general to capture value separation scenarios.

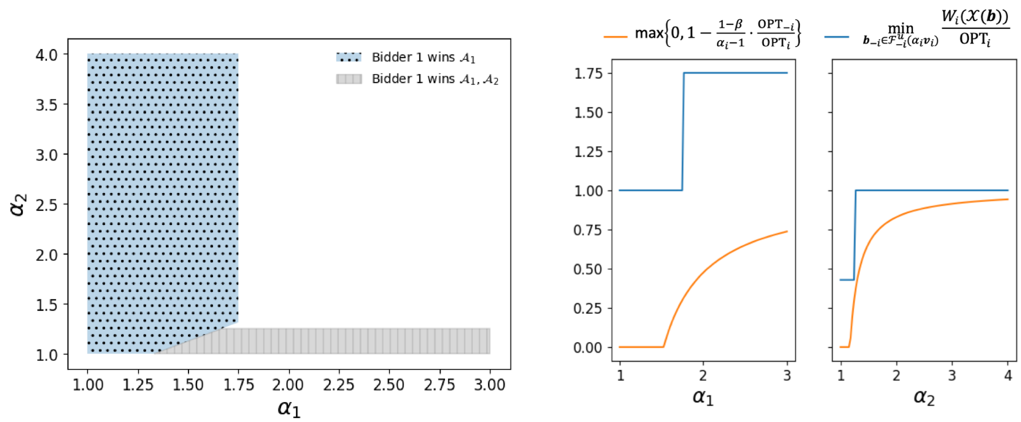


Figure 5-1: Left: Two colored regions represent uniform bid multipliers $(\alpha_1, \alpha_2) \in [1, \infty)^2$ that lead to feasible bid profiles $(\boldsymbol{b}_1, \boldsymbol{b}_2) \in \mathcal{F}$. Right: Comparison between the individual welfare guarantee of Theorem 5.4.1, namely $1 - \frac{1-\beta}{\alpha_i - 1} \cdot \frac{\mathrm{OPT}_{-i}}{\mathrm{OPT}_i}$, and the worst case welfare for each bidder $i$ (normalized by $\mathrm{OPT}_i$) among all feasible bid profiles when both bidders adopt uniform bidding, namely $\min_{\boldsymbol{b}_{-i} \in \mathcal{F}^u_{-i}(\alpha_i \boldsymbol{v}_i)} \frac{W_i(\mathcal{X}(\boldsymbol{b}))}{\mathrm{OPT}_i}$.

*the blue dotted region corresponds to bid profiles under which bidder 1 wins only $\mathcal{A}_1$, and the grey vertically-dashed region corresponds to bid profiles under which bidder 1 wins $\mathcal{A}_1$ and $\mathcal{A}_2$. From this subgraph, we can see that $\mathcal{F}^u_{-1}(\alpha_1 \boldsymbol{v}_1) = \{\alpha_2 \boldsymbol{v}_2 : \alpha_2 \in$ any colored vertical line segments at $\alpha_1\}$ and similarly $\mathcal{F}^u_{-2}(\alpha_2 \boldsymbol{v}_2) = \{\alpha_1 \boldsymbol{v}_1 : \alpha_1 \in$ any colored horizontal line segments at $\alpha_2\}$. On the right subgraph of Figure 5-1, for each bidder $i = 1, 2$, we plot the individual welfare guarantee $1 - \frac{1-\beta}{\alpha_i} \cdot \frac{\mathrm{OPT}_{-i}}{\mathrm{OPT}_i}$ as well as $\min_{\boldsymbol{b}_{-i} \in \mathcal{F}^u_{-i}(\alpha_i \boldsymbol{v}_i)} \frac{W_i(\mathcal{X}(\boldsymbol{b}))}{\mathrm{OPT}_i}$ which is the worst case welfare among all outcomes induced by uniform bid profiles that satisfy both bidders' ROAS constraints.*

On the left subgraph of Figure 5-1, we observe that it is easier for bidder 2 to ensure a non-empty feasibility set $\mathcal{F}^u_{-2}(\alpha_2 \boldsymbol{v}_2)$ at large $\alpha_2$ values than bidder 1 to ensure non-empty $\mathcal{F}^u_{-1}(\alpha_1 \boldsymbol{v}_1)$ at large $\alpha_1$; e.g. for large $\alpha_2$ such as $\alpha_2 = 3$, bidder 1 can take any $\alpha_1 \in [1, 1.75]$, but for large $\alpha_1 = 3$, bidder 2 can only take $\alpha_2 \in [1, 1.25]$. Nevertheless on the right subgraph, we see that bidder 2's realized welfare is much closer to her theoretical lower bound guarantee than that of bidder 1. Therefore this highlights a tradeoff between uniform multiplier feasibility and welfare guarantee.

## 5.5 VCG yields best individual welfare guarantee among broad class of auctions

Having presented an individual welfare guarantee in the previous Section 5.4 that improves according to the platform's ML advice accuracy, a natural question is that for a given level of accuracy $\beta$, can one achieve a universally better individual welfare guarantee than that of Theorem 5.4.1 via considering auction formats other than VCG? In this section, we demonstrate that the answer is negative when we restrict the auction to a broad class of truthful mechanisms (possibly randomized) with anonymous allocations (see Definition 5.2.2). Here, we again emphasize that truthfulness is w.r.t. quasi-linear utility maximizers (see Definition 5.2.1).

In the following theorem, we show that no allocation-anonymous, truthful auction $\mathcal{A}$ when augmented by $\beta$-approximate reserves (see Definition 5.2.3), can universally outperform VCG, i.e. for any $\mathcal{A}$ there exists a problem instance in which a bidder has a welfare guarantee at most the individual welfare lower bound for VCG of Theorem 5.4.1.

**Theorem 5.5.1.** *Let $\mathcal{A}$ be any single-slot auction format (with position bias $\mu = 1$) that is allocation-anonymous, truthful, and possibly randomized. Then, there exists an instance of $M$ parallel auctions $(\mathcal{A}_j)_{j\in[M]}$ of format $\mathcal{A}$, $N$ bidders with values $\boldsymbol{v} \in \mathbb{R}_+^{N\times M}$, $\beta$-approximate reserves $\boldsymbol{r} \in \mathbb{R}_+^{N\times M}$, and an autobidder $i$ with multiplier $\alpha_i > 1$ (see Proposition 5.2.1), such that there is a feasible bid profile $\boldsymbol{b} \in \mathcal{F}$ in which $\boldsymbol{b}_i = \alpha_i \boldsymbol{v}_i \in \mathbb{R}_+^M$ that results in the following welfare upper bound for autobidder $i$: $\frac{\mathbb{E}[W_i(\mathcal{X}(\boldsymbol{b}))]}{\mathbb{E}[\mathrm{OPT}_i]} \leq 1 - \frac{1-\beta}{\alpha_i - 1} \cdot \frac{\mathbb{E}[\mathrm{OPT}_{-i}]}{\mathbb{E}[\mathrm{OPT}_i]}$ . Here the expectation is taken w.r.t. possible randomness in $\mathcal{A}$.*

Details on implementation allocation-anonymous auctions with personalized reserve prices can be found in Definition 5.2.3 and Example D.1.1. Our proof strategy for Theorem 5.5.1 is to construct a "bad" autobidding instance for any auction $\mathcal{A}$ of interest that yields low individual welfare for one specific bidder: we show that in this autobidding instance, there is some bidder $i$ who has a welfare upper bound as stated in the theorem. The construction of this bad autobidding instance is motivated by Example 5.3.1, in which the key source of low welfare for an individual bidder $i$ comes from the fact that competing bidders outbid $i$ in auctions where $i$'s value is high, and cover their expenditures with value acquired from other auctions where they have no competition. Following this idea, since the bad instance in Theorem 5.5.1 requires us to maximize individual welfare for a specific bidder $i$, we can achieve this by having auctions where each of $i$'s competitors is the only bidder submitting a nonzero bid, and with these "no-competition" auctions competitors can cover their expenditures for outbidding bidder $i$ in auctions where $i$'s value is largest.

**Proof sketch for Theorem 5.5.1.** For any auction $\mathcal{A}$ that is allocation-anonymous, truthful and possibly randomized, we consider a "bad" autobidding instance $(N, M, \mathcal{A}, \boldsymbol{r}, \boldsymbol{v})$ where $N = K+1$ bidders labeled $B_1 ... B_K, B_0$ compete in $M = 2K+1$ auctions with single-slots for some $K \in \mathbb{N}$, and bidders' values are shown in the following table. Reserves are set to be $r_{i,j} = \beta v_{i,j}$ for some $\beta \in (0,1)$ and are $\beta$-approximate (see Definition 5.3.1). Bidder $B_0$'s multiplier is fixed at $\alpha_0 > 1$.

| | $A_1$ | $A_2$ | $\dots$ | $A_K$ | $A_{K+1}$ | $A_{K+2}$ | $\dots$ | $A_{2K}$ | $A_{2K+1}$ |
|---|---|---|---|---|---|---|---|---|---|
| $B_1$ | $\frac{\alpha_0 v+\epsilon}{\rho}$ | $\frac{\alpha_0 v+2\epsilon}{\rho}$ | $\dots$ | $\frac{\alpha_0 v+K\epsilon}{\rho}$ | $\gamma$ | $0$ | $\dots$ | $0$ | $0$ |
| $B_2$ | $\frac{\alpha_0 v+2\epsilon}{\rho}$ | $\frac{\alpha_0 v+3\epsilon}{\rho}$ | $\dots$ | $\frac{\alpha_0 v+\epsilon}{\rho}$ | $0$ | $\gamma$ | $\dots$ | $0$ | $0$ |
| $\vdots$ | $\vdots$ | $\vdots$ | | $\vdots$ | $\vdots$ | $\vdots$ | | $\vdots$ | $\vdots$ |
| $B_K$ | $\frac{\alpha_0 v+K\epsilon}{\rho}$ | $\frac{\alpha_0 v+\epsilon}{\rho}$ | $\dots$ | $\frac{\alpha_0 v+(K-1)\epsilon}{\rho}$ | $0$ | $0$ | $\dots$ | $\gamma$ | $0$ |
| $B_0$ | $v$ | $v$ | $\dots$ | $v$ | $0$ | $0$ | $\dots$ | $0$ | $y$ |

In the table, we choose $\epsilon = \mathcal{O}(1/K^3)$ and suitable parameters $\rho, \gamma, v, y > 0$ to satisfy certain conditions, one of which guarantees $B_0$'s value is the highest in auctions $A_1 \dots A_K$ . With the above instance, we consider the specific outcome $\boldsymbol{x}$ where bidders $1, \dots, K$ adopt bid multiplier $\rho$, in which case bidder $B_0$ has the lowest bid in auctions

$A_1 \dots A_K$ (since $\alpha_0 > 1$. Then, our proof of Theorem 5.5.1 is to show bidder $B_0$ can acquire welfare at most the upper bound in Theorem 5.5.1. The proof consists of 3 parts:

(1) Under outcome $\boldsymbol{x}$, we upper bound bidder $B_0$'s expected acquired welfare in auctions $A_1 ... A_K$. This acquired welfare should be small, since other bidders are outbidding bidder $B_0$ in these auctions, by covering their expenditures via the value acquired in auctions $A_{K+1}, ..., A_{2K}$, respectively.

(2) We show that bidder $B_0$ satisfies her ROAS constraint, which holds valid due to the fact that she is acquiring value in auction $A_{2K+1}$ for suitable $y$ facing no competition.

(3) We show that any bidder $i \in [K]$ satisfies her ROAS constraint. In this part, when we take appropriate parameters $\epsilon \to 0$ and $\rho \to \alpha_0$, we first show that the total expected acquired value of bidder $i$ minus her total expected cost over all $2K+1$ auctions is approximately $1 - \sum_{j \in [K]} \mathbb{P}(\text{bidder } i \text{ wins auction } j) + \mathcal{O}(K^2 \epsilon)$ and since $\epsilon = \mathcal{O}(1/K^3)$, we only need to show $1 \geq \sum_{j \in [K]} \mathbb{P}(\text{bidder } i \text{ wins auction } j)$. Our proof for this claim exploits the specific structure of our problem instance construction: we recognize that when bidders $1, \dots, K$ use bid multiplier $\rho$, the bid profiles for auctions $A_1 ... A_K$ are a cyclic permutation of the set $\{b_0, \dots, b_K\} = \{\alpha_0 v, \alpha_0 v + \epsilon, \dots, \alpha_0 v + K\epsilon\}$. Therefore by allocation-anonymity of $\mathcal{A}$, the expected outcome in each of the auctions in $A_1, \dots, A_K$ are symmetric over bidders $1, \dots, K$, or in other words, $\mathbb{P}(\text{bidder } i \text{ wins auction } j)$ depends only on the bid value of bidder $i$ in auction $j$. The cyclic structure thus implies

$$
\begin{aligned}
&\sum_{j \in [K]} \mathbb{P}(\text{bidder } i \text{ wins auction } j) \\
&= \sum_{k \in [K]} \mathbb{P}(\text{bid value } b_k \text{ wins auction } \mathcal{A} \text{ given competing bids } b_{-k}) \leq 1 .
\end{aligned}
$$

Here, we also point out that any $\beta$ approximate reserves do not affect allocation in auctions $A_1 ... A_K$, simply because any bid value in $\{b_0, \dots, b_K\}$ is greater than the largest reserve price among agents, namely $\beta v$, since $\alpha_0 > 1 > \beta$. In other words, under the specific outcome $\boldsymbol{x}$, allocation anonymity of any auction in $A_1, \ldots, A_K$ is preserved with personalized reserves $r_{i,j} = \beta v_{i,j}$ due to our construction; see also Remark 5.3.1. For a technical re-statement of Theorem 5.5.1 and its proof, please refer to Appendix D.3.2.

## 5.6 Extensions: Individual welfare guarantees for GSP and GFP with ML advice

In this section, we extend our individual welfare guarantees for the VCG auction in Theorem 5.4.1 to the GSP and GFP auctions, which are both non-truthful. For technical purposes, we assume that bidder values are "well-separated" as defined in Definition 5.4.2.

Further, as discussed in Section 5.2.2, uniform bidding (i.e. setting the same bid multiplier for all auctions) is only optimal in truthful auctions. In GSP and GFP, one can construct instances where non-uniform bidding strictly outperforms uniform bidder (for more details see e.g. [38]). Thus, for GSP and GFP autobidding instances, we impose no assumptions on the bid values of bidders other than being undominated: we say a bid value $\boldsymbol{b}_i \in \mathbb{R}_+^M$ is undominated for bidder $i$ if there is no other bid value $\boldsymbol{b}_i' \in \mathbb{R}_+^M$ that strictly outperforms $\boldsymbol{b}_i$ in welfare under all competing bid profiles. Mathematically, $\nexists \boldsymbol{b}_i' \in \mathbb{R}_+^M$ such that $W_i(\mathcal{X}(\boldsymbol{b}_i, \boldsymbol{b}_{-i})) < W_i(\mathcal{X}(\boldsymbol{b}_i', \boldsymbol{b}_{-i}))$ for all $\boldsymbol{b}_{-i} \in \mathbb{R}_+^{(N-1)\times(M)}$. The following lemma lower bounds undominated bids in the presence of $\beta$-approximate reserves.

**Lemma 5.6.1** (Lemma 4.7 & 4.9 of [11])**.** *Consider the setting where $(\mathcal{A}_j)_{j\in[M]}$ are all GSP auctions or GFP auctions, and reserve prices $\boldsymbol{r}$ are $\beta$-approximate. Denote $\mathcal{U} \subseteq \mathcal{F}$ to be the set of bid profiles in which each bid is undominated and satisfies all bidders' ROAS constraints. Then for any $\boldsymbol{b} \in \mathcal{U}$, $b_{i,j}$ must satisfy $b_{i,j} \geq r_{i,j} \geq \beta v_{i,j}$ for any bidder $i \in [N]$ and auction $\mathcal{A}_j$.*

Finally, our main theorem for this section is the following:

**Theorem 5.6.2.** *Consider the setting where $(\mathcal{A}_j)_{j\in[M]}$ are all GSP auctions or GFP auctions. Suppose reserve prices $\boldsymbol{r}$ are $\beta$-approximate, and values $\boldsymbol{v}$ are $\Delta$-separated s.t. $\beta > \frac{\Delta}{2\Delta-1}$. Consider any undominated bid profile $\boldsymbol{b}\in\mathcal{U}\subseteq\mathcal{F}$ where $\mathcal{U}$ is the set of all undominated bids under which every bidder's ROAS constraint is satisfied (see Equation (5.2)). Then, the individual welfare guarantee (Definition 5.2.5) is bounded as*

$$\min_{\boldsymbol{b}\in\mathcal{U}} \frac{W_i(\mathcal{X}(\boldsymbol{b}))}{\mathrm{OPT}_i} \;\geq\; 1 - \frac{1-\beta}{\beta - \frac{\Delta}{2\Delta-1}} \cdot \frac{\mathrm{OPT}_{-i}}{\mathrm{OPT}_i}\,.$$

Details on implementation of GSP and GFP with personalized reserve prices can be found in Definition 5.2.3 and Example D.1.1. The proof for Theorem 5.6.2 is presented in Appendix D.4.1. Comparing the individual welfare guarantees in Theorem 5.4.1 for VCG and Theorem 5.6.2 for GSP/GFP, we observe when values are $\Delta$-separated and ML advice is $\beta$-accurate, when bidders adopt small enough uniform multipliers in VCG (i.e. $\alpha_i - 1 < \beta - \frac{\Delta}{2\Delta-1}$), GSP/GFP provides a better individual welfare guarantee compared to VCG, whereas for large multipliers (i.e. $\alpha_i - 1 > \beta - \frac{\Delta}{2\Delta-1}$), individual welfare in VCG dominates that in the considered non-truthful auctions.

## 5.7 Additional discussions and future research

This chapter focuses on presenting theoretical welfare guarantees on the individual advertiser level. One important question is how ad channels can measure such welfare guarantees via a data-driven metric in practice without knowing certain parameters present in our individual welfare guarantees such as $\beta, \mathrm{OPT}_i, \mathrm{OPT}_{-i}$? For instance, how should an ad platform choose between VCG or GSP/GFP auctions (based on comparing their individual welfare guarantees), when ML-advice quality $\beta$ and value separation parameter $\Delta$ are unknown (recall Theorem 5.6.2 and discussions thereof)? Therefore one important future research direction is to develop data-driven metrics to measure individual welfare guarantees in practical setups, and consequently demonstrate the accuracy of such metrics w.r.t. our presented theoretical welfare lower bounds.

# Appendix A

# Supplementary material for Chapter 2

## A.1 Additional material for Section 2.3

### A.1.1 Proof of Lemma 2.3.1

*Proof.* Fix any option $\mathcal{I} \in \{\mathcal{I}_B, \mathcal{I}_R, \mathcal{I}_G\}$ defined in Eq. (2.2), and let $(\widetilde{\boldsymbol{\gamma}}, \widetilde{\boldsymbol{\rho}}) \in \mathcal{I}$ be the optimal solution to CH-OPT($\mathcal{I}$). Note that for the per-channel ROI only option $\mathcal{I}_R$, we have $\widetilde{\rho}_j = \infty$ and for the per-channel budget only we have $\widetilde{\gamma}_j = 0$ for all $j \in [M]$. Further, for any realization of value-cost pairs over all auctions $\boldsymbol{z} = (\boldsymbol{v}_j, \boldsymbol{d}_j)_{j\in[M]}$, recall the optimal solution $\boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)$ to $V_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)$ for each channel $j \in [M]$ as defined in Eq. (2.4).

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

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

where we used the definitions $V_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j) = \boldsymbol{v}_j^\top \boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)$ and $D_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j) = \boldsymbol{d}_j^\top \boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)$ in Eq. (2.5). This implies $\left(\boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)\right)_{j\in[M]}$ satisfies the ROI constraint in GL-OPT. A similar analysis implies $\left(\boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)\right)_{j\in[M]}$ also satisfies the budget constraint in GL-OPT. Therefore, $\left(\boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)\right)_{j\in[M]}$ is feasible to GL-

OPT. So

$$\text{GL-OPT} \geq \sum_{j\in[M]} \mathbb{E}\left[\boldsymbol{v}_j^\top \boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)\right] = \sum_{j\in M} \left[V_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)\right] = \text{CH-OPT}(\mathcal{I})\,. \quad \text{(A.1)}$$

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

### A.1.2 Proof of Theorem 2.3.4

*Proof.* In light of Lemma 2.3.1, we only need to show CH-OPT($\mathcal{I}_B$) $\geq$ GL-OPT. Let $\widetilde{\boldsymbol{x}}(\boldsymbol{z}) = \{\widetilde{\boldsymbol{x}}_j(\boldsymbol{z}_j))\}_{j\in[N]}$ be the optimal solution to GL-OPT, and define $\widetilde{\gamma}_j = 0$ and $\widetilde{\rho}_j = \mathbb{E}\left[\boldsymbol{d}_j^\top \widetilde{\boldsymbol{x}}_j(\boldsymbol{z}_j))\right]$ to be the corresponding expected spend for each channel $j$ under the optimal solution $\widetilde{\boldsymbol{x}}(\boldsymbol{z})$ to GL-OPT, respectively.

We first argue that $(\widetilde{\gamma}_j, \widetilde{\rho}_j)_{j\in[M]}$ is feasible to CH-OPT($\mathcal{I}_B$). Recall the optimal solution $\boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)$ to $V_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)$ for each channel $j \in [M]$ as defined in Eq. (2.4), as well as the definitions $V_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j) = \boldsymbol{v}_j^\top \boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)$ and $D_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j) = \boldsymbol{d}_j^\top \boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)$ in Eq. (2.5). Then, we have

$$\mathbb{E}\left[D_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)\right] \;=\; \mathbb{E}\left[\boldsymbol{d}_j^\top \boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j); \boldsymbol{z}_j\right] \;\overset{(i)}{\leq}\; \widetilde{\rho}_j \;=\; \mathbb{E}\left[\boldsymbol{d}_j^\top \widetilde{\boldsymbol{x}}_j(\boldsymbol{z}_j)\right]\,, \quad \text{(A.2)}$$

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

$$\sum_{j\in[M]} \mathbb{E}\left[D_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)\right] \leq \sum_{j\in[M]} \mathbb{E}\left[\boldsymbol{d}_j^\top \widetilde{\boldsymbol{x}}_j(\boldsymbol{z}_j)\right] \overset{(i)}{\leq} \rho\,. \quad \text{(A.3)}$$

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

On the other hand, we have

$$V_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j) \;=\; \boldsymbol{v}_j^\top \boldsymbol{x}_j^*(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j) \;\overset{(i)}{\geq}\; \boldsymbol{v}_j^\top \widetilde{\boldsymbol{x}}_j(\boldsymbol{z}_j) \quad \text{(A.4)}$$

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

$$\begin{aligned}\sum_{j \in M} \mathbb{E}\left[V_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)\right] \; &\geq \; \sum_{j \in M} \mathbb{E}\left[\boldsymbol{v}_j^\top \widetilde{\boldsymbol{x}}_j(\boldsymbol{z}_j)\right] \; \overset{(i)}{\geq} \; \gamma \sum_{j \in M} \mathbb{E}\left[\boldsymbol{d}_j^\top \widetilde{\boldsymbol{x}}_j(\boldsymbol{z}_j)\right] \\ \overset{(ii)}{\geq} \; \gamma \sum_{j \in [M]} \mathbb{E}\left[D_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)\right]\end{aligned} \tag{A.5}$$

where (i) follows from feasibility of $\widetilde{\boldsymbol{x}}(\boldsymbol{z}) = \{\widetilde{\boldsymbol{x}}_j(\boldsymbol{z}_j))\}_{j \in [N]}$ to GL-OPT since it is the optimal solution; (ii) follows from Eq. (A.2). Hence combining Eq. (A.3) (A.5) we can conclude that $(\widetilde{\gamma}_j, \widetilde{\rho}_j)_{j \in [M]}$ is feasible to CH-OPT$(\mathcal{I}_B)$.

Finally, we have CH-OPT$(\mathcal{I}_B) \geq \sum_{j \in M} \mathbb{E}\left[V_j(\widetilde{\gamma}_j, \widetilde{\rho}_j; \boldsymbol{z}_j)\right] \geq \sum_{j \in M} \mathbb{E}\left[\boldsymbol{v}_j^\top \widetilde{\boldsymbol{x}}_j(\boldsymbol{z}_j)\right] =$ GL-OPT, where the last inequality follows from (A.5), and the final equality is because we assumed $\widetilde{\boldsymbol{x}}(\boldsymbol{z}) = \{\widetilde{\boldsymbol{x}}_j(\boldsymbol{z}_j))\}_{j \in [N]}$ is the optimal solution to GL-OPT. □

### A.1.3 Proof of Corollary 2.3.5

*Proof.* In light of Lemma 2.3.1, we only need to show CH-OPT$(\mathcal{I}_G) \geq$ GL-OPT. Let $(\widetilde{\boldsymbol{\gamma}}, \widetilde{\boldsymbol{\rho}}) \in \mathcal{I}_B$, and by definition of $\mathcal{I}_B$ in Eq. (2.2) we have $\widetilde{\gamma}_j = 0$ for all $j \in [M]$. Since $(\widetilde{\boldsymbol{\gamma}}, \widetilde{\boldsymbol{\rho}})$ is feasible to CH-OPT$(\mathcal{I}_B)$, it is also feasible to CH-OPT$(\mathcal{I}_G)$ since these two problems share the same ROI and budget constraints. Because they also share the same objectives, we have

$$\text{CH-OPT}(\mathcal{I}_G) \geq \text{CH-OPT}(\mathcal{I}_B) = \text{GL-OPT} \tag{A.6}$$

where the final equality follows from Theorem 2.3.4. □

## A.2 Additional material for Section 2.4

### A.2.1 Proof of Proposition 2.4.1

*Proof.* Define $\bar{\mu}_T = \frac{1}{T}\sum_{t \in [T]} \mu_t$ as well as $\bar{\lambda}_T = \frac{1}{T}\sum_{t \in [T]} \lambda_t$ . Let $(\rho_j^*)_{j \in [M]}$ be the optimal per-channel budgets to CH-OPT$(\mathcal{I}_B)$, then we have the following weak duality statement for any $\lambda, \mu \geq 0$:

$$\text{CH-OPT}(\mathcal{I}_B) = \sum_{j\in[M]} V_j(\rho_j^*) \overset{(i)}{\leq} \sum_{j\in[M]} V_j(\rho_j^*) + \lambda \sum_{j\in[M]} \left(V_j(\rho_j^*) - \gamma\rho_j^*\right) + \mu\left(\rho - \sum_{j\in[M]}\rho_j^*\right)$$
$$\overset{(ii)}{\leq} \sum_{j\in[M]} \mathcal{L}(\rho_j^*, \lambda, \mu) + \rho\mu \tag{A.7}$$

where (i) follows from feasibility of $(\rho_j^*)_{j\in[M]}$ to CH-OPT$(\mathcal{I}_B)$; (ii) follows from the definition of the Lagrangian function in Eq. (2.7). We can further bound the advertiser's regret as followed:

$$\begin{aligned}
&T \cdot \text{GL-OPT} - \mathbb{E}\left[\sum_{t\in[T]}\sum_{j\in[M]} V_j(\rho_{j,t})\right] \\
\overset{(i)}{=}\ & M\bar{V}K + (T-K)\cdot \text{CH-OPT}(\mathcal{I}_B) - \mathbb{E}\left[\sum_{t>K}\sum_{j\in[M]} V_j(\rho_{j,t})\right] \\
\overset{(ii)}{\leq}\ & M\bar{V}K + (T-K)\cdot \sum_{j\in[M]} \mathbb{E}\left[\mathcal{L}_j(\rho_j^*, \bar{\lambda}_T, \bar{\mu}_T) + \rho\bar{\mu}_T\right] - \mathbb{E}\left[\sum_{t>K}\sum_{j\in[M]} V_j(\rho_{j,t})\right] \\
\overset{(iii)}{<}\ & M\bar{V}K + \rho\sum_{t>K}\mathbb{E}\left[\mu_t\right] + \sum_{t>K}\sum_{j\in[M]}\mathbb{E}\left[\mathcal{L}_j(\rho_j^*, \lambda_t, \mu_t)\right] \\
& - \sum_{t>K}\sum_{j\in[M]}\mathbb{E}\left[\mathcal{L}_j(\rho_{j,t}, \boldsymbol{c}_t) - \lambda_t\left(V_j(\rho_{j,t}) - \gamma\rho_{j,t}\right) + \mu_t\rho_{j,t}\right] \\
\overset{(iv)}{\leq}\ & M\bar{V}K + \sum_{j\in[M]}\sum_{t>K}\mathbb{E}\left[\mathcal{L}_j(\rho_j^*(t), \boldsymbol{c}_t) - \mathcal{L}_j(\rho_{j,t}, \boldsymbol{c}_t)\right] + \sum_{t>K}\left(\lambda_t g_{1,t} + \mu_t g_{2,t}\right)\ .
\end{aligned} \tag{A.8}$$

In (i), we first used the trivial upper bound $\bar{V}$ on conversion functions in each channel for the first $K$ periods where we choose each of the arms once and pull that arm across all channels. Then we applied GL-OPT = CH-OPT$(\mathcal{I}_B)$ from Theorem 2.3.4; (ii) follows from the weak duality statement in Eq. (A.7) by taking $\lambda = \bar{\lambda}_T = \frac{1}{T}\sum_{t\in[T]}\lambda_t$ and $\mu = \bar{\mu}_T = \frac{1}{T}\sum_{t\in[T]}\mu_t$; (iii) follows from the definition of the Lagrangian function in Eq. (2.7) where we have $\mathcal{L}_j(\rho_j, \boldsymbol{c}; \boldsymbol{z}_j) = (1+\lambda)V_j(\rho_j; \boldsymbol{z}_j) - (\lambda\gamma + \mu)\rho_j$, and $\mathcal{L}_j(\rho_j, \boldsymbol{c}) =$

$\mathbb{E}\left[\mathcal{L}_j(\rho_j, \boldsymbol{c}; \boldsymbol{z}_j)\right]$; in (iv) we used the definition that $\rho_j^*(t) = \arg\max_{\rho_j \geq 0} \mathcal{L}_j(\rho_j, \boldsymbol{c}_t)$ to be the optimal budget that maximizes the Lagrangian w.r.t. the dual variables $\boldsymbol{c}_t = (\lambda_t, \mu_t)$, and the definition $g_{1,t} = \sum_{j\in[M]} V_j(\rho_{j,t}; \boldsymbol{z}_t) - \gamma\rho_{j,t}$ and $g_{2,t} = \rho - \sum_{j\in[M]} \rho_{j,t}$ as defined in Algorithm 1. □

### A.2.2 Proof for Lemma 2.4.2

*Proof.* Here, we would like to show that $\sum_{t>K} \mathbb{E}\left[\lambda_t g_{1,t} + \mu_t g_{2,t}\right] \leq \mathcal{O}\left(\eta T + \frac{1}{\eta}\right)$, where we recall that $g_{1,t} = \sum_{j\in[M]} \left(V_j(\rho_{j,t}; \boldsymbol{z}_t) - \gamma\rho_{j,t}\right)$, $g_{2,t} = \rho - \sum_{j\in[M]} \rho_{j,t}$ and $\eta > 0$ the step size defined in Algorithm 1.

Now from lemma A.2.3, we have for any $\lambda \geq 0$ and $\mu \geq 0$,

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

Telescoping the above from $t = K+1$ to $t = T$, we get

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

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

### A.2.3 Proof of Lemma 2.4.3

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

Fix any channel $j$ which consists of $m_j$ parallel auctions, and recall that we assumed the orderding $\frac{v_{j,1}}{d_{j,1}} > \frac{v_{j,2}}{d_{j,2}} > \cdots > \frac{v_{j,m_j}}{d_{j,m_j}}$ for any realization $\boldsymbol{z}_j$. Then, with the option where the per-channel ROI is set to 0 (i.e. omitted) $V_j(\rho_j; \boldsymbol{z}_j)$ is exactly the LP relaxation of a 0-1 knapsack, whose optimal solution $\boldsymbol{x}_j^*(\rho_j; \boldsymbol{z}_j)$ is well known to be unique [32], and takes the form for any auction index $n \in [m_j]$:

$$x^*_{j,n}(\rho_j; \boldsymbol{z}_j) = \begin{cases} 1 & \text{if } \sum_{n' \in [n]} d_{j,n'} \leq \rho_j \\ \frac{\rho_j - \sum_{n' \in [n-1]} d_{j,n'}}{d_{j,n}} & \text{if } \sum_{n' \in [n]} d_{j,n'} > \rho_j \\ 0 & \text{otherwise} \end{cases} \tag{A.9}$$

where we denote $d_{j,0} = 0$. With this form, we can see that when $\rho_j$ lies in between $\sum_{n' \in [n-1]} d_{j,n'}$ and $\sum_{n' \in [n]} d_{j,n'}$, the conversion functions is

$$\begin{aligned} V_j(\rho_j; \boldsymbol{z}_j) &= \boldsymbol{v}_j^\top \boldsymbol{x}^*_j(\rho_j; \boldsymbol{z}_j) = \sum_{n' \in [m_j]} v_{j,n'} \cdot x^*_{j,n'}(\rho_j; \boldsymbol{z}_j) \\ &= \sum_{n' \in [n-1]} v_{j,n'} \cdot 1 + v_{j,n} \cdot \frac{\rho_j - \sum_{n' \in [n-1]} d_{j,n'}}{d_{j,n}} = \frac{v_{j,n}}{d_{j,n}} \rho_j + b_{j,n} \end{aligned}$$

where we denote $d_{j,0} = v_{j,0} = 0$ and also $b_{j,n} = \sum_{n' \in [n-1]} v_{j,n'} - \frac{v_{j,n}}{d_{j,n}} \cdot \left( \sum_{n' \in [n-1]} d_{j,n'} \right)$. Hence, by using indicators to specify where $\rho_j$ lies, we have

$$\begin{aligned} V_j(\rho_j; \boldsymbol{z}_j) &= \boldsymbol{v}_j^\top \boldsymbol{x}^*_j(\rho_j; \boldsymbol{z}_j) \\ = \sum_{n \in [m_j]} &\left( \frac{v_{j,n}}{d_{j,n}} \rho_j + b_{j,n} \right) \mathbb{I} \left\{ d_{j,0} + \cdots + d_{j,n-1} \leq \rho_j \leq d_{j,0} + \cdots + d_{j,n} \right\} . \end{aligned} \tag{A.10}$$

Further, it is easy to check that any two line segments, say $[X_{n-1}, X_n]$ and $[X_n, X_{n+1}]$ where we write $X_n = d_{j,0} + \cdots + d_{j,n}$, intersect at $\rho_j = X_n$, because $\frac{v_{j,n}}{d_{j,n}} \rho_j + b_{j,n} = \frac{v_{j,n+1}}{d_{j,n+1}} \rho_j + b_{j,n+1}$ at $\rho_j = X_n$. Hence, from Eq. (A.10) we can conclude $V_j(\rho_j; \boldsymbol{z}_j)$ is continuous, which further implies it is piece-wise linear and strictly increasing. Further, the ordering $\frac{v_{j,1}}{d_{j,1}} > \frac{v_{j,2}}{d_{j,2}} > \cdots > \frac{v_{j,m_j}}{d_{j,m_j}}$ implies that the slopes on each segment $[X_n, X_{n+1}]$ decreases as $n$ increases, which implies $V_j(\rho_j; \boldsymbol{z}_j)$ is concave.

Since $V_j(\rho_j) = \mathbb{E}\left[ V_j(\rho_j; \boldsymbol{z}_j) \right]$, where the expectation is taken w.r.t. randomness in $\boldsymbol{z}_j$, and since the $\boldsymbol{z}_j$ is sampled from some discrete distribution $\boldsymbol{p}_j$ on finite support $F_j$, $V_j(\rho_j)$ is simply a weighted average over all $\left( V_j(\rho_j; \boldsymbol{z}_j) \right)_{\boldsymbol{z}_j \in F_j}$ with weights in $\boldsymbol{p}_j$, so $V_j(\rho_j)$ is also continuous, piece-wise linear, strictly increasing, and concave, and thus can be written as in Lemma 2.4.3:

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

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

Finally, according to the definition of $\mathcal{L}_j(\rho_j, \boldsymbol{c}) = \mathbb{E}\left[\mathcal{L}_j(\rho_j, \boldsymbol{c}; \boldsymbol{z}_j)\right]$ and $\mathcal{L}_j(\rho_j, \boldsymbol{c}; \boldsymbol{z}_j) = (1+\lambda)V_j(\rho_j; \boldsymbol{z}_j) - (\lambda\gamma + \mu)\rho_j$ as defined in Eq. (2.7), we have

$$\mathcal{L}_j(\rho_j, \boldsymbol{c}) = (1+\lambda)V_j(\rho_j) - (\lambda\gamma + \mu)\rho_j\,, \tag{A.11}$$

which implies $\mathcal{L}_j(\rho_j, \boldsymbol{c})$ is continuous, piece-wise linear, and concave in $\rho_j$ because $V_j(\rho_j)$ is continuous, piece-wise linear, and concave as shown above. Combining Eq. (A.11) and the representation of $V_j(\rho_j)$ in Lemma (2.4.3), we have

$$\mathcal{L}_j(\rho_j, \boldsymbol{c}) = \sum_{n\in[S_j]} (\sigma_{j,n}(\boldsymbol{c})\rho_j + (1+\lambda)b_{j,n})\,\mathbb{I}\{r_{j,n-1} \le \rho_j \le r_{j,n}\}\,. \tag{A.12}$$

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

### A.2.4 Proof for Lemma 2.4.4

*Proof.* We first present some definitions for convenience: denote $\boldsymbol{c} = (\lambda, \mu)$, $\boldsymbol{c}_t = (\lambda_t, \mu_t)$, and $\boldsymbol{z}_t = (\boldsymbol{v}_t, \boldsymbol{d}_t)$. For any realization $\boldsymbol{z} = (\boldsymbol{v}, \boldsymbol{d}) = (\boldsymbol{v}_j, \boldsymbol{d}_j)_{j\in[M]} \in F_1 \times \ldots F_M$

of values and costs across channels, define the dual function

$$
\begin{aligned}
\mathcal{L}_d(\boldsymbol{c};\boldsymbol{z}) &= \max_{\boldsymbol{\rho}\in[0,\rho]^M} \sum_{j\in[M]} V_j(\rho_j;\boldsymbol{z}) + \lambda \sum_{j\in[M]} \left(V_j(\rho_j;\boldsymbol{z}) - \gamma\rho_j\right) + \mu\left(\rho - \sum_{j\in[M]}\rho_j\right) \\
&= \mu\rho + \max_{(\rho_j)_{j\in[M]}\in[0,\rho]^M} \sum_{j\in[M]} \mathcal{L}_j(\rho_j,\boldsymbol{c};\boldsymbol{z}_j)\,.
\end{aligned}
\tag{A.13}
$$

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

$$
(\boldsymbol{c}'-\boldsymbol{c})^\top \nabla\mathcal{L}_d(\boldsymbol{c};\boldsymbol{z}) \leq \mathcal{L}_d(\boldsymbol{c}';\boldsymbol{z}) - \mathcal{L}_d(\boldsymbol{c};\boldsymbol{z})\,. \tag{A.14}
$$

We prove the lemma by induction that

$$
\begin{aligned}
\|(\lambda_t,\mu_t)\| &= \|\boldsymbol{c}_t\| \\
\leq\; C_F &:= 1 + \frac{\max_{\boldsymbol{z}\in F_1\times\ldots F_M} \boldsymbol{e}^\top\boldsymbol{v} + 1}{\min_{\boldsymbol{z}\in F_1\times\ldots F_M}\left\{\sum_{j\in[M]}\left(V_j(\rho_{(\boldsymbol{z}),j};\boldsymbol{z}) - \gamma\rho_j\right), \rho - \sum_{j\in[M]}\rho_{(\boldsymbol{z}),j}\right\}}\,.
\end{aligned}
\tag{A.15}
$$

Here, we utilized the existence of some $(\rho_{(\boldsymbol{z}),j})_{j\in[M]} \in [0,\rho]^M$ for any $\boldsymbol{z}$ according to Proposition A.2.1 which states at per-channel budget profile $(\rho_{(\boldsymbol{z}),j})_{j\in[M]}$ Slater's condition holds, i.e. $\sum_{j\in[M]}\left(V_j(\rho_{(\boldsymbol{z}),j};\boldsymbol{z}) - \gamma\rho_j\right) > 0$ and $\rho - \sum_{j\in[M]}\rho_{(\boldsymbol{z}),j} > 0$. Slater's condition also implies $C_F > 1$.

The base case for $t=1$ is satisfied trivially since we take $\lambda_1 = \gamma_1 = 0$ and thus $\|\boldsymbol{c}_1\| = 0 < 1 < C_F$. Now assume $\|\boldsymbol{c}_t\| \leq C_F$ for some $t \geq 1$, and we will show $\|\boldsymbol{c}_{t+1}\| \leq C_F$ by considering two different case:

**Case (A):** $\mathcal{L}_d(\boldsymbol{c}_t;\boldsymbol{z}_t) > \mathcal{L}_d(\boldsymbol{0};\boldsymbol{z}_t) + \eta^2M^2\left((\bar{V}+\gamma\rho)^2+\rho^2\right)$. Define $\boldsymbol{g}_t = (g_{1,t}, g_{2,t})$ where we recall $g_{1,t} = \sum_{j\in[M]}(V_{j,t}(\rho_{j,t}) - \gamma\rho_{j,t})$ and $g_{2,t} = \rho - \sum_{j\in[M]}\rho_{j,t}$. From Eq.

(A.53) in the statement of Lemma A.2.3, we have for any $\lambda \geq 0$ and $\mu \geq 0$

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

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

$$\begin{aligned}
\|\boldsymbol{c}_{t+1}\|^2 \;&\leq\; \|\boldsymbol{c}_t\|^2 + 2\eta\,(\boldsymbol{c}_t)^\top \boldsymbol{g}_t + \eta^2 M^2 \left((\bar{V} + \gamma\rho)^2 + \rho^2\right) \\
&\overset{(i)}{\leq}\; \|\boldsymbol{c}_t\|^2 + \eta\,(\mathcal{L}_d(\mathbf{0}; \boldsymbol{z}_t) - \mathcal{L}_d(\boldsymbol{c}_t; \boldsymbol{z}_t)) + \eta^2 M^2 \left((\bar{V} + \gamma\rho)^2 + \rho^2\right) \,,
\end{aligned} \tag{A.16}$$

where in (i) we applied Eq. (A.14) with $\boldsymbol{z} = \boldsymbol{z}_t$ and used the fact that $\boldsymbol{g}_t = (g_{1,t}, g_{2,t}) = \nabla \mathcal{L}_d(\boldsymbol{c}_t; \boldsymbol{z}_t)$.

Then from Eq. (A.16) and the assumption $\mathcal{L}_d(\boldsymbol{c}_t; \boldsymbol{z}_t) > \mathcal{L}_d(\mathbf{0}; \boldsymbol{z}_t) + \eta^2 M^2 \left((\bar{V} + \gamma\rho)^2 + \rho^2\right)$ we have

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

where the final inequality follows from the induction hypothesis.

**Case (B):** $\mathcal{L}_d(\boldsymbol{c}_t; \boldsymbol{z}_t) \leq \mathcal{L}_d(\mathbf{0}; \boldsymbol{z}_t) + \eta^2 M^2 \left((\bar{V} + \gamma\rho)^2 + \rho^2\right)$. Under this case, we have

$$
\begin{aligned}
& \mathcal{L}_d(\mathbf{0};\boldsymbol{z}_t) + \eta M^2 \bar{V}^2 \\
\geq\ & \mathcal{L}_d(\boldsymbol{c}_t;\boldsymbol{z}_t) \\
\overset{(i)}{=}\ & \mu_t\rho + \max_{\boldsymbol{\rho}\in[0,\rho]^M} \sum_{j\in[M]} \mathcal{L}_j(\rho, \boldsymbol{c}_t;\boldsymbol{z}_t) \\
\overset{(ii)}{\geq}\ & \mu_t\rho + \mathcal{L}(\boldsymbol{\rho}_{(\boldsymbol{z}_t)}, \boldsymbol{c}_t;\boldsymbol{z}_t) \\
=\ & \sum_{j\in[M]} V_j(\rho_{(\boldsymbol{z}_t),j};\boldsymbol{z}_t) + \lambda_t \sum_{j\in[M]} \left(V_j(\rho_{(\boldsymbol{z}_t),j};\boldsymbol{z}_t) - \gamma\rho_j\right) + \mu_t\left(\rho - \sum_{j\in[M]} \rho_{(\boldsymbol{z}_t),j}\right) \\
\geq\ & \lambda_t \sum_{j\in[M]} \left(V_j(\rho_{(\boldsymbol{z}_t),j};\boldsymbol{z}_t) - \gamma\rho_j\right) + \mu_t\left(\rho - \sum_{j\in[M]} \rho_{(\boldsymbol{z}_t),j}\right) \\
\overset{(iii)}{\geq}\ & (\lambda_t + \mu_t)\cdot \min_{\boldsymbol{z}\in F_1\times\ldots F_M} \left\{\sum_{j\in[M]} \left(V_j(\rho_{(\boldsymbol{z}),j};\boldsymbol{z}) - \gamma\rho_j\right), \rho - \sum_{j\in[M]} \rho_{(\boldsymbol{z}),j}\right\}.
\end{aligned} \tag{A.17}
$$

Here, (i) follows from the definition of the dual function in Eq. (A.13); in (ii) we recall the definition of $\boldsymbol{\rho}_{(\boldsymbol{z})} \in [0,\rho]^M$ that satisfies the Slater's condition in Proposition A.2.1; in (iii), we used the fact that $\lambda_t, \mu_t \geq 0$ and also under Slater's condition we have for any $\boldsymbol{z}$, $\sum_{j\in[M]}\left(V_j(\rho_{(\boldsymbol{z}),j};\boldsymbol{z}) - \gamma\rho_j\right) > 0$ and $\rho - \sum_{j\in[M]} \rho_{(\boldsymbol{z}),j} > 0$.

On the other hand, we have $\eta^2 M^2 \left((\bar{V}+\gamma\rho)^2 + \rho^2\right) < 1$ since in the statement of the lemma we assumed $\eta < \frac{1}{M\sqrt{(\bar{V}+\gamma\rho)^2+\rho^2}}$. Also, $\mathcal{L}_d(\mathbf{0};\boldsymbol{z}_t) = \max_{\boldsymbol{\rho}\in[0,\rho]^M} \sum_{j\in[M]} V_j(\rho_j;\boldsymbol{z}_t) = \max_{\boldsymbol{z}\in F_1\times\ldots F_M} \boldsymbol{e}^\top \boldsymbol{v}$. Hence combining this with Eq. (A.17) we get

$$
\lambda_t + \mu_t < \frac{\max_{\boldsymbol{z}\in F_1\times\ldots F_M} \boldsymbol{e}^\top\boldsymbol{v} + 1}{\min_{\boldsymbol{z}\in F_1\times\ldots F_M}\left\{\sum_{j\in[M]}\left(V_j(\rho_{(\boldsymbol{z}),j};\boldsymbol{z}) - \gamma\rho_j\right), \rho - \sum_{j\in[M]}\rho_{(\boldsymbol{z}),j}\right\}} := C_F - 1\,, \tag{A.18}
$$

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

$$
\lambda_{t+1} = \left(\lambda_t - \eta g_{1,t}\right)_+ \leq \lambda_t + \eta|g_{1,t}| \leq C_F - 1 + \eta M(\bar{V}+\gamma\rho) < C_F \tag{A.19}
$$

where in the final inequality we used the fact that $\eta M(\bar{V} + \gamma\rho) < 1$ since in the statement of the lemma we assumed $\eta < \frac{1}{M\sqrt{(\bar{V}+\gamma\rho)^2+\rho^2}}$. A similar bound holds for $\mu_{t+1}$. $\square$

### A.2.5 Proof for Lemma 2.4.5

*Proof.* 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. (2.9). Then we define the following

$$
\begin{aligned}
&\Delta_k(\boldsymbol{c}) = \max_{\rho_j \in [0,\rho]} \mathcal{L}_j(\rho_j, \boldsymbol{c}) - \mathcal{L}_j(a_k, \boldsymbol{c}) \\
&\mathcal{C}_n = \left\{\boldsymbol{c} \in \{\boldsymbol{c}_t\}_{t\in[T]} : r_{j,n} = \arg\max_{\rho_j \geq 0} \mathcal{L}_j(\rho_j, \boldsymbol{c})\right\} \text{ for } n = 0 \ldots S_j \\
&\mathcal{C}(\underline{\sigma}) = \left\{\boldsymbol{c} \in \{\boldsymbol{c}_t\}_{t\in[T]} : \sigma_j^-(\boldsymbol{c}) > \underline{\sigma},\ |\sigma_j^+(\boldsymbol{c})| > \underline{\sigma}\right\} \text{ for } n = 0, \ldots, S_j \\
&m_k(\boldsymbol{c}) = \frac{8\log(T)}{\Delta_k^2(\boldsymbol{c})} \text{ for } \forall (k, \boldsymbol{c}) \text{ s.t. } \Delta_k(\boldsymbol{c}) > 0\,.
\end{aligned}
\tag{A.20}
$$

Here, the "adjacent slopes" $\sigma_j^-(\boldsymbol{c})$ and $\sigma_j^+(\boldsymbol{c})$, which are defined in Eq.(2.18), represent the slopes that are adjacent to the optimal budget $\arg\max_{\rho_j \in [0,\rho]} \mathcal{L}_j(\rho_j, \boldsymbol{c})$ for any context $\boldsymbol{c} = (\lambda, \mu)$. Further, $S_j$ and $\{r_{j,n}\}_{j\in[S_j]}$ are defined in Lemma 2.4.3. Here we state in words the meanings of $\Delta_k(\boldsymbol{c})$, $\mathcal{C}(\underline{\sigma})$ and $\mathcal{C}_n$, respectively.

- $\Delta_k(\boldsymbol{c})$ denotes the loss in contextual bandit rewards when pulling arm $a_k$ under context $\boldsymbol{c}$.
- $\mathcal{C}_n$ is the set including all context $\boldsymbol{c}_t$ under which the optimal per-channel budget $\arg\max_{\rho_j \geq 0} \mathcal{L}_j(\rho_j, \boldsymbol{c}_t)$ is taken at the $n$th "turning point" $r_{j,n}$ (see Lemma 2.4.3).
- $\mathcal{C}(\underline{\sigma})$ is the set of all contexts, in which the adjacent slopes to the optimal point w.r.t. the context $\boldsymbol{c}$, namely $\arg\max_{\rho_j \geq 0} \mathcal{L}_j(\rho_j, \boldsymbol{c})$, have magnitude greater than $\underline{\sigma}$, or in other words, the adjacent slopes are steep.

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

$$\mathcal{U}_j^-(\boldsymbol{c}) = [r_{j,n-1}, r_{j,n}] \text{ and } \mathcal{U}_j^+(\boldsymbol{c}) = [r_{j,n}, r_{j,n+1}] \quad \text{if} \quad \boldsymbol{c} \in \mathcal{C}_n\,. \tag{A.21}$$

In other words, if $\boldsymbol{c} \in \mathcal{C}_n$, per the definition of $\mathcal{C}_n$ above, $\arg\max_{\rho_j \in [0,\rho]} \mathcal{L}_j(\rho_j, \boldsymbol{c})$ is located at the $n$th "turning point" $r_{j,n}$, then $\mathcal{U}_j^-(\boldsymbol{c})$ and $\mathcal{U}_j^-(\boldsymbol{c})$ 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 followed

$$\begin{aligned}
&\sum_{t>K} \mathcal{L}_j(\rho_j^*(t), \boldsymbol{c}_t) - \mathcal{L}_j(\rho_{j,t}, \boldsymbol{c}_t) = X_1 + X_2 + X_3 \quad \text{where} \\
&X_1 = \sum_{t>K: \boldsymbol{c}_t \notin \mathcal{C}(\underline{\sigma})} \sum_{k \in [K]} \Delta_k(\boldsymbol{c}_t) \mathbb{I}\{\rho_{j,t} = a_k, N_{k,t} \leq m_k(\boldsymbol{c}_t)\} \\
&X_2 = \sum_{t>K: \boldsymbol{c}_t \in \mathcal{C}(\underline{\sigma})} \sum_{k \in [K]} \Delta_k(\boldsymbol{c}_t) \mathbb{I}\{\rho_{j,t} = a_k, N_{k,t} \leq m_k(\boldsymbol{c}_t)\} \\
&X_3 = \sum_{k \in [K]} \sum_{t>K} \Delta_k(\boldsymbol{c}_t) \mathbb{I}\{\rho_{j,t} = a_k, N_{k,t} > m_k(\boldsymbol{c}_t)\}\,.
\end{aligned} \tag{A.22}$$

In Section A.2.5, we show that $X_1 \leq \widetilde{\mathcal{O}}(\delta T + \underline{\sigma} T + \frac{1}{\delta})$; in Section A.2.5 we show that $X_2 \leq \widetilde{\mathcal{O}}(\delta T + \frac{1}{\delta \underline{\sigma}})$; in Section A.2.5 we show that $X_3 \leq \widetilde{\mathcal{O}}(\frac{1}{\delta T})$.

**Remark A.2.1.** *In the following sections A.2.5, A.2.5 and A.2.5 where we bound $X_1$, $X_2$, and $X_3$, respectively, we assume the optimal per-channel $\rho_j^*(t) = \arg\max_{\rho_j \in [0,\rho]} \mathcal{L}_j(\rho_j, \boldsymbol{c}_t)$ 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 followed:*

$$\begin{aligned}
&\mathcal{L}_j(\rho_j^*(t), \boldsymbol{c}_t) - \mathcal{L}_j(\rho_{j,t}, \boldsymbol{c}_t) = \mathcal{L}_j(\rho_j^*(t), \boldsymbol{c}_t) - \mathcal{L}_j(a_t^*, \boldsymbol{c}_t) + \mathcal{L}_j(a_t^*, \boldsymbol{c}_t) - \mathcal{L}_j(\rho_{j,t}, \boldsymbol{c}_t) \\
&\quad \textit{where } a_t^* = \arg\max_{a_k \in \mathcal{A}(\delta)} \mathcal{L}_j(a_k, \boldsymbol{c}_t)
\end{aligned}$$

*The first term will yield an error in the order of $\mathcal{O}(\delta)$ due to the Lagrangian func-*

*tion being unimodal, piecewise linear liner, which implies* $|a_t^* - \rho_j^*(t)| \leq \delta$ *so that* $\mathcal{L}_j(\rho_j^*(t), \boldsymbol{c}_t) - \mathcal{L}_j(a_t^*, \boldsymbol{c}_t) = \mathcal{O}(\delta)$. *Hence, this "discretization error" will accumulate to a magnitude of* $\mathcal{O}(\delta T)$ *over* $T$ *periods, which leads to an additional error that is already accounted for in the statement of the lemma.*

**Bounding $X_1$.**

Our strategy to bound $X_1$ consists of 4 steps, namely bounding the loss of arm $a_k$ at each context $\boldsymbol{c} \notin \mathcal{C}(\underline{\sigma})$ when $a_k \in \mathcal{U}_j^-(\boldsymbol{c})$ lies on the left adjacent region of the optimal budget; $a_k < \min\ \mathcal{U}_j^-(\boldsymbol{c})$ lies to the left of the left adjacent region; $a_k \in \mathcal{U}_j^+(\boldsymbol{c})$ lies on the right adjacent region of the optimal budget; and $a_k > \max\ \mathcal{U}_j^+(\boldsymbol{c})$ lies to the right of the right adjacent region. Here we recall the adjacent regions are defined in Eq.(A.21).

**Step 1: $a_k \in \mathcal{U}_j^-(\boldsymbol{c}_t)$.** For arm $k$ such that $a_k \in \mathcal{U}_j^-(\boldsymbol{c}_t)$, recall Lemma 2.4.3 that $\mathcal{L}_j(a, \boldsymbol{c}_t)$ is linear in $a$ for $a \in \mathcal{U}_j^-(\boldsymbol{c}_t)$, so $\Delta_k(\boldsymbol{c}_t) = \sigma_j^-(\boldsymbol{c}_t) \cdot (\rho_j^*(t) - a_k) \leq \underline{\sigma}\rho$ where we used the condition that $\boldsymbol{c}_t \notin \mathcal{C}(\underline{\sigma})$ so the adjacent slopes have magnitude at most $\underline{\sigma}$, and $\rho_j^*(t) \leq \rho$. Thus, summing over all such $k$ we get

$$\begin{aligned}
&\sum_{t>K:\boldsymbol{c}_t\notin\mathcal{C}(\underline{\sigma})} \sum_{k\in[K]:a_k\in\mathcal{U}_j^-(\boldsymbol{c}_t)} \Delta_k(\boldsymbol{c}_t)\mathbb{I}\{\rho_{j,t} = a_k, N_{k,t} \leq m_k(\boldsymbol{c}_t)\} \\
\leq\ &\sum_{t>K:\boldsymbol{c}_t\notin\mathcal{C}(\underline{\sigma})} \sum_{k\in[K]:a_k\in\mathcal{U}_j^-(\boldsymbol{c}_t)} \underline{\sigma}\rho \cdot \mathbb{I}\{\rho_{j,t} = a_k, N_{k,t} \leq m_k(\boldsymbol{c}_t)\} \ \leq\ \underline{\sigma}\rho T = \mathcal{O}(\underline{\sigma}T)\,.
\end{aligned} \tag{A.23}$$

**Step 2: $a_k < \min\ \mathcal{U}_j^-(\boldsymbol{c}_t)$.** For arm $k$ such that $a_k < \min\ \mathcal{U}_j^-(\boldsymbol{c}_t)$, we further split contexts into groups $\mathcal{C}_n$ for $n = 0 \ldots S_j$ (defined in Eq. (A.20)) 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 2-2 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

$$
\begin{aligned}
&\sum_{t>K:\boldsymbol{c}_t\in\mathcal{C}_n/\mathcal{C}(\underline{\sigma})}\quad\sum_{k\in[K]:a_k<\min\mathcal{U}_j^-(\boldsymbol{c}_t)}\Delta_k(\boldsymbol{c}_t)\mathbb{I}\{\rho_{j,t}=a_k,N_{k,t}\le m_k(\boldsymbol{c}_t)\}\\
\overset{(i)}{=}\;&\sum_{t>K:\boldsymbol{c}_t\in\mathcal{C}_n/\mathcal{C}(\underline{\sigma})}\quad\sum_{k\in[K]:a_k<r_{j,n-1}}\Delta_k(\boldsymbol{c}_t)\mathbb{I}\{\rho_{j,t}=a_k,N_{k,t}\le m_k(\boldsymbol{c}_t)\}\\
=\;&\sum_{t>K}\sum_{\boldsymbol{c}\in\mathcal{C}_n/\mathcal{C}(\underline{\sigma})}\quad\sum_{k\in[K]:a_k<r_{j,n-1}}\Delta_k(\boldsymbol{c})\mathbb{I}\{\boldsymbol{c}_t=\boldsymbol{c},\rho_{j,t}=a_k,N_{k,t}\le m_k(\boldsymbol{c})\}\\
\overset{(ii)}{\le}\;&\sum_{t>K}\sum_{\boldsymbol{c}\in\mathcal{C}_n/\mathcal{C}(\underline{\sigma})}\Big(\Delta_{k'}(\boldsymbol{c})\mathbb{I}\{\boldsymbol{c}_t=\boldsymbol{c}\}\\
&\qquad+\sum_{k\in[K]:a_k<r_{j,n-1}-\delta}\Delta_k(\boldsymbol{c})\mathbb{I}\{\boldsymbol{c}_t=\boldsymbol{c},\rho_{j,t}=a_k,N_{k,t}\le m_k(\boldsymbol{c})\}\Big)\\
\overset{(iii)}{\le}\;&\left((1+C_F)s_{j,n-1}\delta+\rho\underline{\sigma}\right)T+\sum_{k\in[K]:a_k<r_{j,n-1}-\delta}\quad\sum_{\boldsymbol{c}\in\mathcal{C}_n/\mathcal{C}(\underline{\sigma})}\Delta_k(\boldsymbol{c})Y_k(\boldsymbol{c})
\end{aligned}
\tag{A.24}
$$

where in the final equality we defined $Y_k(\boldsymbol{c})=\sum_{t>K}\mathbb{I}\{\boldsymbol{c}_t=\boldsymbol{c},\rho_{j,t}=a_k,N_{k,t}\le m_k(\boldsymbol{c})\}$. In (i) we used the fact that the left end of the left adjacent region, i.e. $\min\mathcal{U}_j^-(\boldsymbol{c}_t)$ is exactly $r_{j,n-1}$ because for context $\boldsymbol{c}_t\in\mathcal{C}_n$ the optimal budget $\arg\max_{\rho_j\in[0,\rho]}\mathcal{L}_j(\rho_j,\boldsymbol{c}_t)$ is at the $n$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 $a_1<a_2<\cdots<a_K$. Note that in (ii) we separate out the arm $a_{k'}$ because its distance to the optimal per-channel may be less than $\delta$ 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 $\delta$ away from the optimal per-channel budget; (iii) follows from the fact that under a context $\boldsymbol{c}\in\mathcal{C}_n/\mathcal{C}(\underline{\sigma})$, we have $\arg\max_{\rho_j\in[0,\rho]}\mathcal{L}_j(\rho_j,\boldsymbol{c})=r_{j,n}$ so

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

where in (iv) we used $\boldsymbol{c}\in\mathcal{C}_n/\mathcal{C}(\underline{\sigma})$ implies $\sigma_j^-(\boldsymbol{c})\le\underline{\sigma}$, as well as all $r_{j,n}\le\rho$ 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 $\sigma_{j,n-1}(\boldsymbol{c}) = (1+\lambda)s_{j,n-1} - (\mu + \gamma\lambda) \leq (1+C_F)s_{j,n-1}$ where $C_F$ is defined in Lemma 2.4.4.

We now bound $\sum_{\boldsymbol{c} \in \mathcal{C}_n/\mathcal{C}(\underline{\sigma})} \Delta_k(\boldsymbol{c}) Y_k(\boldsymbol{c})$ in Eq. (A.24). It is easy to see the following inequality for any sequence of context $\boldsymbol{c}_{(1)}, \ldots, \boldsymbol{c}_{(\ell)} \in \{\boldsymbol{c}_t\}_{t \in [T]}$ (This is a slight generalization of an inequality result shown in [12]):

$$Y_k(\boldsymbol{c}_{(1)}) + \cdots + Y_k(\boldsymbol{c}_{(\ell)}) \leq \max_{\ell'=1\ldots\ell} m_k(\boldsymbol{c}_{(\ell')}) \,. \tag{A.25}$$

This is because

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

For simplicity denote $L = |\mathcal{C}_n/\mathcal{C}(\underline{\sigma})|$, and order contexts in $\boldsymbol{c} \in \mathcal{C}_n/\mathcal{C}(\underline{\sigma})$ as $\{\boldsymbol{c}_{(\ell)}\}_{\ell \in [L]}$ s.t. $\Delta_k(\boldsymbol{c}_{(1)}) > \Delta_k(\boldsymbol{c}_{(2)}) > \cdots > \Delta_k(\boldsymbol{c}_{(L)})$, or equivalently $m_k(\boldsymbol{c}_{(1)}) < m_k(\boldsymbol{c}_{(2)}) < \cdots < m_k(\boldsymbol{c}_{(L)})$ according to Eq.(A.20). Then multiplying Eq. (A.25) by by $\Delta_k(\boldsymbol{c}_{(\ell)}) - \Delta_k(\boldsymbol{c}_{(\ell+1)})$ (which is strictly positive due to the ordering of contexts), and summing $\ell = 1 \ldots L$ we get

$$\begin{aligned}
\sum_{\boldsymbol{c} \in \mathcal{C}_n/\mathcal{C}(\underline{\sigma})} \Delta_k(\boldsymbol{c}) Y_k(\boldsymbol{c}) = \sum_{\ell \in [L]} \Delta_k(\boldsymbol{c}_{(\ell)}) Y_k(\boldsymbol{c}_{(\ell)}) \;&\leq\; \sum_{\ell \in [L]} m_k(\boldsymbol{c}_{(\ell)}) \left(\Delta_k(\boldsymbol{c}_{(\ell)}) - \Delta_k(\boldsymbol{c}_{(\ell+1)})\right) \\
&\overset{(i)}{=} \; 8\log(T) \sum_{\ell \in [L-1]} \frac{\Delta_k(\boldsymbol{c}_{(\ell)}) - \Delta_k(\boldsymbol{c}_{(\ell+1)})}{\Delta_k^2(\boldsymbol{c}_{(\ell)})} \;\overset{(ii)}{\leq}\; 8\log(T) \int_{\Delta_k(\boldsymbol{c}_{(L)})}^{\infty} \frac{dz}{z^2} \\
&= \; \frac{8\log(T)}{\Delta_k(\boldsymbol{c}_{(L)})} \;\overset{(iii)}{=}\; \frac{8\log(T)}{\min_{\boldsymbol{c} \in \mathcal{C}_n/\mathcal{C}(\underline{\sigma})} \Delta_k(\boldsymbol{c})} \,.
\end{aligned} \tag{A.26}$$

Here (i) follows from the definition of $m_k(\boldsymbol{c})$ in Eq. (A.20) where $m_k(\boldsymbol{c}) = \frac{8\log(T)}{\Delta_k^2(\boldsymbol{c})}$; both (ii) and (iii) follow from the ordering of contexts so that $\Delta_k(\boldsymbol{c}_{(1)}) > \Delta_k(\boldsymbol{c}_{(2)}) > \cdots > \Delta_k(\boldsymbol{c}_{(L)})$. Note that for any $\boldsymbol{c} \in \mathcal{C}_n/\mathcal{C}(\underline{\sigma})$ and arm $k$ such that $a_k < r_{j,n-1}$, we have

$$
\begin{aligned}
\Delta_k(\boldsymbol{c}) &= \mathcal{L}_j(r_{j,n}, \boldsymbol{c}) - \mathcal{L}_j(r_{j,n-1}, \boldsymbol{c}) + \mathcal{L}_j(r_{j,n-1}, \boldsymbol{c}) - \mathcal{L}_j(a_k, \boldsymbol{c}) \\
&> \mathcal{L}_j(r_{j,n-1}, \boldsymbol{c}) - \mathcal{L}_j(a_k, \boldsymbol{c}) \\
&\overset{(i)}{\geq} \sigma_{j,n-1}(\boldsymbol{c})(r_{j,n-1} - a_k) \\
&\overset{(ii)}{\geq} \left(\sigma_{j,n-1}(\boldsymbol{c}) - \sigma_{j,n}(\boldsymbol{c})\right)(r_{j,n-1} - a_k) \\
&\overset{(iii)}{=} (1+\lambda)\left(s_{j,n-1} - s_{j,n}\right)(r_{j,n-1} - a_k) \\
&> \left(s_{j,n-1} - s_{j,n}\right)(r_{j,n-1} - a_k)\,,
\end{aligned}
\tag{A.27}
$$

where in (i) we recall the slope $\sigma_{j,n-1}(\boldsymbol{c})$ is defined in Lemma 2.4.3 and further (i) follows from concavity of $\mathcal{L}_j(\rho j, \boldsymbol{c})$ in the first argument $\rho_j$; in (ii) we used the fact that $\sigma_{j,n}(\boldsymbol{c}) \geq 0$ since the optimal budget $\arg\max_{\rho_j \in [0,\rho]} \mathcal{L}_j(\rho_j, \boldsymbol{c})$ is taken at the $n$th turning point, and is the largest turning point whose left slope is non-negative from Lemma 2.4.3; (iii) follows from the definition $\sigma_{j,n'}(\boldsymbol{c}) = (1+\lambda)s_{j,n'} - (\mu + \gamma\lambda)$ for any $n'$.

Finally combining Eqs. (A.24), (A.26) and (A.27), and summing over $n = 1 \dots S_j$

we get

$$
\begin{aligned}
&\sum_{t>K:\boldsymbol{c}_t\notin\mathcal{C}(\underline{\sigma})}\ \sum_{k\in[K]:a_k<\min\mathcal{U}_j^-(\boldsymbol{c}_t)} \Delta_k(\boldsymbol{c}_t)\mathbb{I}\{\rho_{j,t}=a_k, N_{k,t}\le m_k(\boldsymbol{c}_t)\}\\
&= \sum_{n\in[S_j]}\ \sum_{t>K:\boldsymbol{c}_t\in\mathcal{C}_n/\mathcal{C}(\underline{\sigma})}\ \sum_{k\in[K]:a_k<\min\mathcal{U}_j^-(\boldsymbol{c}_t)} \Delta_k(\boldsymbol{c}_t)\mathbb{I}\{\rho_{j,t}=a_k, N_{k,t}\le m_k(\boldsymbol{c}_t)\}\\
&\le \sum_{n\in[S_j]}\left((1+C_F)s_{j,n-1}\delta+\rho\underline{\sigma}\right)T+\sum_{n\in[S_j]}\ \sum_{k\in[K]:a_k<r_{j,n-1}-\delta}\frac{8\log(T)}{(s_{j,n-1}-s_{j,n})\,(r_{j,n-1}-a_k)}\\
&\overset{(i)}{\le} \sum_{n\in[S_j]}\left((1+C_F)s_{j,n-1}\delta+\rho\underline{\sigma}\right)T+\sum_{n\in[S_j]}\ \sum_{\ell=1}^{K}\frac{8\log(T)}{(s_{j,n-1}-s_{j,n})\,\ell\delta}\\
&\le \sum_{n\in[S_j]}\left((1+C_F)s_{j,n-1}\delta+\rho\underline{\sigma}\right)T+\frac{8\log(T)\log(K)}{\delta}\sum_{n\in[S_j]}\frac{1}{(s_{j,n-1}-s_{j,n})}\\
&= \widetilde{\mathcal{O}}(\delta T+\underline{\sigma}T+\frac{1}{\delta})\,.
\end{aligned}
\tag{A.28}
$$

Note that (i) follows because for all $a_k<r_{j,n-1}-\delta$, the $a_k$'s distances from $r_{j,n-1}$ are at least $\delta, 3\delta, 2\delta\ldots$. In the last equation, we hide all logarithmic factors using the notation $\widetilde{\mathcal{O}}$, 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 $\boldsymbol{p}_j$ for value-cost pairs; see definitions of these absolute constants in Lemmas 2.4.3 and 2.4.4.

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

Therefore, combining Eqs. (A.23) and (A.28) we can conclude

$$
X_1\le\widetilde{\mathcal{O}}(\delta T+\underline{\sigma}T+\frac{1}{\delta})\,. \tag{A.29}
$$

**Bounding $X_2$.**

We first rewrite $X_2$ as followed

$$
\begin{aligned}
X_2 &= \sum_{t>K:\boldsymbol{c}_t\in\mathcal{C}(\underline{\sigma})}\sum_{k\in[K]}\Delta_k(\boldsymbol{c}_t)\mathbb{I}\{\rho_{j,t}=a_k, N_{k,t}\leq m_k(\boldsymbol{c}_t)\}\\
&= \sum_{t>K}\sum_{n\in[S_j]}\sum_{k\in[K]}\sum_{\boldsymbol{c}\in\mathcal{C}_n\cap\mathcal{C}(\underline{\sigma})}\Delta_k(\boldsymbol{c})\mathbb{I}\{\boldsymbol{c}_t=\boldsymbol{c},\rho_{j,t}=a_k, N_{k,t}\leq m_k(\boldsymbol{c})\}\\
&\overset{(i)}{=} \sum_{n\in[S_j]}\sum_{k\in[K]}\sum_{\boldsymbol{c}\in\mathcal{C}_n\cap\mathcal{C}(\underline{\sigma})}\Delta_k(\boldsymbol{c})Y_k(\boldsymbol{c})\\
&\overset{(ii)}{=} \sum_{n\in[S_j]}\sum_{\boldsymbol{c}\in\mathcal{C}_n\cap\mathcal{C}(\underline{\sigma})}\sum_{k\in\{k_n^-,k_n^+\}}\Delta_k(\boldsymbol{c})Y_k(\boldsymbol{c})+\sum_{n\in[S_j]}\sum_{\boldsymbol{c}\in\mathcal{C}_n\cap\mathcal{C}(\underline{\sigma})}\sum_{k\in[K]/\{k_n^-,k_n^+\}}\Delta_k(\boldsymbol{c})Y_k(\boldsymbol{c})\\
&\overset{(iii)}{\leq} T\delta(1+C_F)\sum_{n\in[S_j]}(s_{j,n}+s_{j,n+1})+\sum_{n\in[S_j]}\sum_{\boldsymbol{c}\in\mathcal{C}_n\cap\mathcal{C}(\underline{\sigma})}\sum_{k\in[K]/\{k_n^-,k_n^+\}}\Delta_k(\boldsymbol{c})Y_k(\boldsymbol{c})\,.
\end{aligned}
\tag{A.30}
$$

where in (i) we define $Y_k(\boldsymbol{c})=\sum_{t>K}\mathbb{I}\{\boldsymbol{c}_t=\boldsymbol{c},\rho_{j,t}=a_k,N_{k,t}\leq m_k(\boldsymbol{c})\}$; in (ii) we separate out two arms $k_n^-$ and $k_n^+$ defined as followed: recall for context $\boldsymbol{c}\in\mathcal{C}_n\cap\mathcal{C}(\underline{\sigma})$, the optimal budget $\arg\max_{\rho_j\in[0,\rho]}\mathcal{L}_j(\rho_j,\boldsymbol{c})=r_{j,n}$ is taken at the $n$th turning point per the definition of $\mathcal{C}_n$ in Eq. (A.20), 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 $\Delta_{k_n^-}(\boldsymbol{c})=\sigma_j^-(\boldsymbol{c})(r_{j,n}-a_{k_n^-})\leq\sigma_j^-(\boldsymbol{c})\delta\leq(1+C_F)s_{j,n-1}\delta$, where in the final inequality follows from the definition of $\sigma_j^-(\boldsymbol{c})=\sigma_{j,n}(\boldsymbol{c})=(1+\lambda)s_{j,n}-(\mu+\gamma\lambda)\leq(1+\lambda)s_{j,n}\leq(1+C_F)s_{j,n}$ where $C_F$ is defined in Eq. (2.4.4). A similar bound holds for $\Delta_{k_n^+}(\boldsymbol{c})$.

Then, following the same logic as Eqs. (A.25), (A.26), (A.27) in Section A.2.5 where we bound $X_1$, we can bound $\sum_{\boldsymbol{c}\in\mathcal{C}_n\cap\mathcal{C}(\underline{\sigma})}\Delta_k(\boldsymbol{c})Y_k(\boldsymbol{c})$ as followed for any arm $k\in[K]/\{k_n^-,k_n^+\}$, i.e. arms who are at least $\delta$ away from the optimal per-channel budget w.r.t. $\boldsymbol{c}$:

$$\sum_{\boldsymbol{c}\in\mathcal{C}_n\cap\mathcal{C}(\underline{\sigma})} \Delta_k(\boldsymbol{c})Y_k(\boldsymbol{c}) \leq \frac{8\log(T)}{\min_{\boldsymbol{c}\in\mathcal{C}_n\cap\mathcal{C}(\underline{\sigma})}\Delta_k(\boldsymbol{c})}\,. \tag{A.31}$$

Now, the set $k\in[K]/\{k_n^-,k_n^+\}$ in Eq. (A.30) 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 $\boldsymbol{c}\in\mathcal{C}_n\cap\mathcal{C}(\underline{\sigma})$,

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

where the final inequality follows from the definition of $\mathcal{C}(\underline{\sigma})$ in Eq. (A.20) such that $\sigma_j^-(\boldsymbol{c}) \geq \underline{(}\sigma)$ for $\boldsymbol{c}\in\mathcal{C}(\underline{\sigma})$. Hence combining this with Eq. (A.31) we have

$$\begin{aligned}\sum_{k\in[K]:a_k<r_{j,n}-\delta}\ \sum_{\boldsymbol{c}\in\mathcal{C}_n\cap\mathcal{C}(\underline{\sigma})} \Delta_k(\boldsymbol{c})Y_k(\boldsymbol{c}) &\leq \sum_{k\in[K]:a_k<r_{j,n}-\delta} \frac{8\log(T)}{\underline{\sigma}(r_{j,n}-a_k)} \\ \overset{(i)}{\leq} \sum_{\ell=1}^{K}\frac{8\log(T)}{\underline{\sigma}\ell\delta} \leq \frac{8\log(T)\log(K)}{\underline{\sigma}\delta}\,,\end{aligned} \tag{A.32}$$

where (i) follows because for all $a_k < r_{j,n}-\delta$, the $a_k$'s distances from $r_{j,n-1}$ are at least $\delta, 3\delta, 2\delta\ldots$. Symmetrically, we can show an identical bound for the set $\{k\in[K]: a_k > r_{j,n}+\delta\}$. Hence, combining Eqs. (A.30) and (A.32) we can conclude

$$X_2 \leq \widetilde{\mathcal{O}}\left(\delta T + \frac{1}{\delta\underline{\sigma}}\right). \tag{A.33}$$

Here, similar to our bound in Eq. (A.28) for bounding $X_1$, we hide all logarithmic factors using the notation $\widetilde{\mathcal{O}}$, 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 $\boldsymbol{p}_j$ for value-cost pairs; see definitions of these absolute constants in Lemma 2.4.3 and 2.4.4.

**Bounding $X_3$.**

We first define

$$\bar{\mathcal{L}} = (1+\gamma)\,\rho C_F + (1+C_F)\bar{V} \tag{A.34}$$

where $C_F$ is specified in Lemma 2.4.4. Recalling the definition $\Delta_k(\boldsymbol{c}) = \max_{\rho_j\in[0,\rho]} \mathcal{L}_j(\rho_j,\boldsymbol{c}) - \mathcal{L}_j(a_k,\boldsymbol{c})$ in Eq. (A.20), and $-(1+\gamma)\rho C_F \le \mathcal{L}_j(\rho_j,\boldsymbol{c}) \le (1+C_F)\bar{V}$ for any $\rho_j \in [0,\rho]$ and context $\boldsymbol{c}$ (see Lemma 2.4.4), it is easy to see

$$\Delta_k(\boldsymbol{c}) \le \bar{\mathcal{L}} \quad \forall k\in[K], \forall \boldsymbol{c}\,. \tag{A.35}$$

Then we bound $X_3$ as followed

$$\begin{aligned} X_3 &= \sum_{k\in[K]}\sum_{t>K}\mathbb{E}\left[\Delta_k(\boldsymbol{c})\mathbb{I}\{\rho_{j,t}=a_k, N_{k,t}>m_k(\boldsymbol{c})\}\right] \\ &\overset{(i)}{\le} \bar{\mathcal{L}}\cdot\sum_{k\in[K]}\sum_{t>K}\mathbb{P}\left(\rho_{j,t}=a_k, N_{k,t}>m_k(\boldsymbol{c}_t)\right) \\ &\overset{(ii)}{\le} \bar{\mathcal{L}}\cdot\sum_{k\in[K]}\sum_{t>K}\mathbb{P}\Big(\hat{V}_{j,t}(a_k) - \frac{\lambda_t\gamma+\mu_t}{1+\lambda_t}a_k + \texttt{UCB}_{j,t}(a_k) \ge \hat{V}_{j,t}(\rho_j^*(t)) \\ &\qquad - \frac{\lambda_t\gamma+\mu_t}{1+\lambda_t}\rho_j^*(t) + \texttt{UCB}_{j,t}(\rho_j^*(t)), N_{k,t}>m_k(\boldsymbol{c}_t)\Big)\,, \end{aligned} \tag{A.36}$$

where (i) follows from Eq. (A.35); in (ii), recall that we choose arm $\rho_{j,t} = a_k$ because the estimated UCB rewards of arm $a_k$ is greater than that of any other arm including $\rho_j^*(t)$ according to Eq. (2.10) in UCB-SGD (Algorithm 2), or mathematically, $\hat{V}_{j,t}(a_k) - \frac{\lambda_t\gamma+\mu_t}{1+\lambda_t}a_k + \texttt{UCB}_{j,t}(a_k) \ge \hat{V}_{j,t}(\rho_j^*(t)) - \frac{\lambda_t\gamma+\mu_t}{1+\lambda_t}\rho_j^*(t) + \texttt{UCB}_{j,t}(\rho_j^*(t))$. Here we also used the fact that $\rho_j^*(t)$ lies in the arm set $\mathcal{A}(\delta)$ for all $t$ (see Remark A.2.1).

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

$$\hat{R}_n(a_k) = \hat{V}_{j,\tau}(a_k) \text{ for } \tau = \min\{t\in[T] : N_{k,t} = n\} \tag{A.37}$$

where we recall $\hat{V}_{j,\tau}(a_k)$ is the estimated conversion for arm $a_k$ in channel $j$ during period $\tau$ as defined in Algorithm 1. In other words, $\tau$ 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. (A.36) as followed:

$$\begin{aligned}
&\mathbb{P}\Big(\hat{V}_{j,t}(a_k) - \frac{\lambda_t\gamma + \mu_t}{1+\lambda_t}a_k + \mathtt{UCB}_{j,t}(a_k) \\
&\qquad \geq \hat{V}_{j,t}(\rho_j^*(t)) - \frac{\lambda_t\gamma + \mu_t}{1+\lambda_t}\rho_j^*(t) + \mathtt{UCB}_{j,t}(\rho_j^*(t)),\ N_{k,t} > m_k(\boldsymbol{c}_t)\Big) \\
\leq\ &\mathbb{P}\Big(\max_{n:m_k(\boldsymbol{c}_t)<n\leq t}\Big\{\hat{R}_n(a_k) + \mathtt{UCB}_n(a_k) - \frac{\lambda_t\gamma + \mu_t}{1+\lambda_t}a_k\Big\} \\
&\qquad \geq \min_{n':1\leq n'\leq t}\Big\{\hat{R}_{n'}(\rho_j^*(t)) + \mathtt{UCB}_{n'}(\rho_j^*(t)) - \frac{\lambda_t\gamma + \mu_t}{1+\lambda_t}\rho_j^*(t)\Big\}\Big) \qquad \text{(A.38)} \\
\leq\ &\sum_{n=\lceil m_k(\boldsymbol{c}_t)\rceil+1}^{t}\sum_{n'=1}^{t}\mathbb{P}\Big(\hat{R}_n(a_k) + \mathtt{UCB}_n(a_k) - \frac{\lambda_t\gamma + \mu_t}{1+\lambda_t}a_k \\
&\qquad > \hat{R}_{n'}(\rho_j^*(t)) + \mathtt{UCB}_{n'}(\rho_j^*(t)) - \frac{\lambda_t\gamma + \mu_t}{1+\lambda_t}\rho_j^*(t)\Big)
\end{aligned}$$

Now, when the event

$$\Big\{\hat{R}_n(a_k) + \mathtt{UCB}_n(a_k) - \frac{\lambda_t\gamma + \mu_t}{1+\lambda_t}a_k > \hat{R}_{n'}(\rho_j^*(t)) + \mathtt{UCB}_{n'}(\rho_j^*(t)) - \frac{\lambda_t\gamma + \mu_t}{1+\lambda_t}\rho_j^*(t)\Big\}$$

occurs, it is easy to see that one of the following events must also occur:

$$\begin{aligned}
&\mathcal{G}_{1,n} = \big\{\bar{R}_n(a_k) \geq V(a_k) + \mathtt{UCB}_n(a_k)\big\} \qquad \text{for } n \text{ s.t. } m_k(\boldsymbol{c}_t) < n \leq t \\
&\mathcal{G}_{2,n'} = \big\{\bar{R}_{n'}(\rho_j^*(t)) \leq V(\rho_j^*(t)) - \mathtt{UCB}_n(\rho_j^*(t))\big\} \qquad \text{for } n' \text{ s.t. } 1 \leq n' \leq t \qquad \text{(A.39)} \\
&\mathcal{G}_3 = \Big\{V_j(\rho_j^*(t)) - \frac{\lambda_t\gamma + \mu_t}{1+\lambda_t}\rho_j^*(t)\ <\ V_j(a_k) - \frac{\lambda_t\gamma + \mu_t}{1+\lambda_t}a_k + 2\cdot\mathtt{UCB}_n(a_k)\Big\}
\end{aligned}$$

Note that for $n > m_k(\boldsymbol{c}_t)$, we have $\mathtt{UCB}_n(a_k) = \sqrt{\frac{2\log(T)}{n}} < \sqrt{\frac{2\log(T)}{m_k(\boldsymbol{c}_t)}} = \frac{\Delta_k(\boldsymbol{c}_t)}{2}$ since we defined $m_k(\boldsymbol{c}) = \frac{8\log(T)}{\Delta_k^2(\boldsymbol{c})}$ in Eq. (A.20). Therefore

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

where (i) follows from the definition of $\Delta_k(\boldsymbol{c}) = \max_{a \in \mathcal{A}(\delta)} \mathcal{L}(a, \boldsymbol{c}) - \mathcal{L}(a_k, \boldsymbol{c})$ in Eq. (A.20) for any context $\boldsymbol{c}$. This implies that event $\mathcal{G}_3$ in Eq. (A.39) cannot hold for $n > m_k(\boldsymbol{c}_t)$. Therefore

$$\begin{aligned} & \mathbb{P}\left(\hat{R}_n(a_k) + \texttt{UCB}_n(a_k) - \frac{\lambda_t\gamma + \mu_t}{1+\lambda_t} a_k > \hat{R}_{n'}(\rho_j^*(t)) + \texttt{UCB}_{n'}(\rho_j^*(t)) - \frac{\lambda_t\gamma + \mu_t}{1+\lambda_t} \rho_j^*(t)\right) \\ \leq \ & \mathbb{P}\left(\mathcal{G}_{1,n} \cup \mathcal{G}_{2,n'}\right) . \end{aligned} \tag{A.40}$$

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. (A.36) (A.38), (A.40) we can conclude

$$\begin{aligned} X_3 \ & \leq \ \sum_{k \in [K]} \sum_{t > K} \sum_{n = \lceil m_k(\boldsymbol{c}_t) \rceil + 1}^{t} \sum_{n'=1}^{t} \left(\mathbb{P}\left(\mathcal{G}_{1,n}\right) + \mathbb{P}\left(\mathcal{G}_{2,n'}\right)\right) \\ & \leq \ \sum_{k \in [K]} \sum_{t > K} \sum_{n = \lceil m_k(\boldsymbol{c}_t) \rceil + 1}^{t} \sum_{n'=1}^{t} \frac{2\bar{V}}{T^4} \\ & \leq \ \frac{2K\bar{V}}{T} = \mathcal{O}\left(\frac{1}{\delta T}\right) . \end{aligned} \tag{A.41}$$

□

### A.2.6 Proof for Theorem 2.4.6

*Proof.* Starting from Proposition 2.4.1, we get

$$\begin{aligned}
& T \cdot \text{GL-OPT} - \mathbb{E}\left[\sum_{t\in[T]}\sum_{j\in[M]} V_j(\rho_{j,t})\right] \\
\leq\ & M\bar{V}K + \sum_{j\in[M]}\sum_{t>K}\mathbb{E}\left[\mathcal{L}_j(\rho_j^*(t),\boldsymbol{c}_t) - \mathcal{L}_j(\rho_{j,t},\boldsymbol{c}_t)\right] + \sum_{t>K}(\lambda_t g_{1,t} + \mu_t g_{2,t}) \qquad \text{(A.42)} \\
\stackrel{(i)}{\leq}\ & M\bar{V}K + \mathcal{O}\left(\underline{\sigma}T + \delta T + \frac{1}{\underline{\sigma}\delta}\right) + \mathcal{O}\left(\eta T + \frac{1}{\eta}\right),
\end{aligned}$$

where in (i) we applied Lemmas 2.4.5 and 2.4.2. 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 2.4.3, $V_j(\rho_j)$ is concave for all $j \in [M]$, so

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

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

On the other hand, Lemma A.2.2 shows that

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

where in (i) we again applied concavity of $V_j(\rho_j)$.

For the budget constraint, again Lemma A.2.2 shows

$$-\mathcal{O}(1/\sqrt{T}) \;\leq\; \frac{1}{T}\sum_{t\in[T]}\mathbb{E}\left[g_{2,t}\right] \;=\; \frac{1}{T}\sum_{t\in[T]}\left(\rho - \sum_{j\in[M]}\mathbb{E}\left[\rho_{j,t}\right]\right) \;=\; \rho - \sum_{j\in[M]}\mathbb{E}[\bar{\rho}_{T,j}]\,.$$

$\square$

### A.2.7 Proof of Lemma 2.4.7

*Proof.* Recall in SGD-UCB-II for all periods $t \in T_1$ in Phase 1, we set the per-channel budget for each channel as some constant $\beta$ which we input to the algorithm, so that $g_{1,t} = \sum_{j\in M}\left(V_j(\beta;\boldsymbol{z}_t) - \gamma\beta\right)$ for all $t \in T_1$. Consider the hypothetical version of SGD-UCB-II where we ignore the budget balance condition $B_t + M\rho > \rho T$ in step 2 of the algorithm, and terminate phase 1 only when the condition $\sum_{t'\in[t-1]} g_{1,t'} > \sqrt{T}\log(T)$ in Step 1 holds. Denote this hypothetical stopping time as $\widetilde{T}_1$, defined to be

$$\widetilde{T}_1 = \min\left\{t \in [T] : \sum_{t'\in[t]}\underbrace{\sum_{j\in M}\left(V_j(\beta;\boldsymbol{z}_{t'}) - \gamma\beta\right)}_{:=h_{t'}} > \sqrt{T}\log(T)\right\} \tag{A.43}$$

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

$$\mathbb{P}\left(T_1 \geq R\right) \leq \mathbb{P}\left(\widetilde{T}_1 \geq R\right) \tag{A.44}$$

Now, it is not difficult to see $\{h_t\}_t$ where $h_t$ defined in Eq. (A.43) are i.i.d. random variables, since the only randomness comes from the realization of value-cost pairs $\{\boldsymbol{z}_t\}_{t\in[T]}$ which are i.i.d.. Further, for any $t \in [T]$ we have

$$h_t := \sum_{j\in M}\left(V_j(\beta;\boldsymbol{z}_t) - \gamma\beta\right) \overset{(i)}{\geq} (\xi-\gamma)\beta \implies \bar{h} := \mathbb{E}[h_t] \geq (\xi-\gamma)\beta > 0 \tag{A.45}$$

where (i) follows from Claim A.2.1 where we also defined $\xi := \min_{j\in[M]} \min_{\boldsymbol{z}_j \in F_j} \frac{v_{j,1}}{d_{j,1}} > \gamma$ under Assumption 2.4.3; in the final inequality we let $\bar{h}$ be the mean of the i.i.d. random variables $\{h_t\}_{t\in[T]}$.

We note that $\widetilde{T}_1 > R$ implies that the sum of the first $R$ $h_t$'s do not exceed $\sqrt{T}\log(T)$ (see definition of $\widetilde{T}_1$ in Eq. (A.43)). Hence, we have

$$
\begin{aligned}
\mathbb{P}\left(\widetilde{T}_1 > R\right) &\leq \mathbb{P}\left(\sum_{t=1}^{\lceil R\rceil} h_t \leq \sqrt{T}\log(T)\right) \\
&= \mathbb{P}\left(\sum_{t=1}^{\lceil R\rceil} (h_t - \bar{h}) \leq \sqrt{T}\log(T) - \lceil R\rceil \cdot \bar{h}\right) \\
&\overset{(i)}{\leq} \mathbb{P}\left(\sum_{t=1}^{\lceil R\rceil} (h_t - \bar{h}) \leq \sqrt{T}\log(T) - R(\xi-\gamma)\beta\right) \\
&\overset{(ii)}{=} \mathbb{P}\left(\sum_{t=1}^{\lceil R\rceil} (h_t - \bar{h}) \leq -\sqrt{T}\log(T)\right) \\
&\overset{(iii)}{\leq} \exp\left(-\frac{T\log^2(T)}{2\lceil R\rceil M^2 \bar{V}^2}\right) \\
&\overset{(iv)}{\leq} \frac{1}{T}.
\end{aligned} \tag{A.46}
$$

Here (i) follows from $\bar{h} \geq (\xi-\gamma)\beta$ in Eq. (A.45); (ii) follows from the definition that $R = 2\sqrt{T}\log^3(T)$ and $\beta = \frac{1}{\log(T)}$ so $R(\xi-\gamma)\beta = 2(\xi-\gamma)\sqrt{T}\log^2(T) \geq 2\sqrt{T}\log(T)$ for large enough $T$ such that $\log(T) > \frac{1}{\xi-\gamma}$; (iii) follows from Azuma Hoefding's inequality given that $h_t \in [0, M\bar{V}]$ for any $t \in [T]$; and finally (iv) follows from $T \geq \lceil R\rceil$, and $\log(T) > M^2\bar{V}^2$ for large enough $T$. Hence, combining Eqs. (A.46) and (A.44) yields the desired statement of the lemma. □

### A.2.8 Proof of Theorem 2.4.8

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

**Part I. Bounding global budget constraint violation.** The design of the

SGD-UCB-II algorithm ensures that the per-channel budget decisions never sum up to exceed $\rho T - M\rho$, so

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

**Part II. Bounding global ROI constraint violation.** Recall the event $\mathcal{E} = \{T_1 \geq 2\sqrt{T}\log^3(T)\}$ defined in Lemma 2.4.7 where $T_1$ is the end period of Phase 1 in the SGD-UCB-II algorithm (see step 10). We consider two scenarios, namely when event $\mathcal{E}$ holds and doesn't hold. Recall $g_{1,t} = \sum_{j\in[M]}(V_j(\rho_{j,t};\boldsymbol{z}_t) - \gamma\rho_{j,t}) \in [-\gamma M\rho, MV]$.

When event $\mathcal{E}$ holds, we consider the following naive bound

$$\mathbb{E}\left[\sum_{t\in[T_2]} g_{1,t} \;\middle|\; \mathcal{E}\right] \geq -T\gamma M\rho\,. \tag{A.47}$$

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

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

where we recall in Phase 1, i.e. the first $T_1$ periods we set each per-channel budget to be $\beta$, and the final inequality holds for large enough $T$. This implies that the budget balance termination condition in step 2 of Algorithm 2 is not met, so Phase 1 terminates because the ROI buffer condition in step 1 of Algorithm 2 is met, i.e.

$$\sum_{t\in[T_1]} g_{1,t} > \sqrt{T}\log(T)\,. \tag{A.48}$$

Now in periods $t = T_1 + 1 \dots T_2$, following the proof of Lemma A.2.2, we have

$$\sum_{t=T_1+1}^{T_2} g_{1,t} \geq \frac{1}{\eta}\left(\lambda_{T_1} - \lambda_{T_2}\right) \overset{(i)}{\geq} -\frac{C_F}{\eta} = -C_F\sqrt{T}\,. \tag{A.49}$$

Here, (i) follows from Lemma 2.4.4 such that $0 \leq \lambda_t \leq C_F$ for any $t \in [T]$ for some absolute constant $C_F > 0$ that only depends on the support of value-cost pairs $F = F_1 \times \ldots \times F_M$, and recall the SGD step size $\eta = 1/\sqrt{T}$. Hence, combining Eqs. (A.47), (A.48), (A.49), we have

$$\begin{aligned}
\mathbb{E}\left[\sum_{t\in[T_2]} g_{1,t}\right] &= \mathbb{E}\left[\sum_{t\in[T_2]} g_{1,t} \,\middle|\, \mathcal{E}\right]\mathbb{P}\left(\mathcal{E}\right) + \mathbb{E}\left[\sum_{t\in[T_2]} g_{1,t} \,\middle|\, \mathcal{E}^c\right]\mathbb{P}\left(\mathcal{E}^c\right) \\
&\geq -T\gamma M\rho\cdot\mathbb{P}\left(\mathcal{E}\right) + \mathbb{E}\left[\sum_{t\in[T_2]} g_{1,t} \,\middle|\, \mathcal{E}^c\right]\mathbb{P}\left(\mathcal{E}^c\right) \\
&\overset{(i)}{\geq} -\gamma M\rho + \left(\sqrt{T}\log(T) - C_F\sqrt{T}\right)\mathbb{P}\left(\mathcal{E}^c\right) \\
&\overset{(ii)}{\geq} 0\,,
\end{aligned} \tag{A.50}$$

where (i) follows from Lemma 2.4.7 which states $\mathbb{P}\left(\mathcal{E}\right) \leq \frac{1}{T}$; in (ii) we used the fact that for large enough $T$, we have $\log(T) > -C_F$, and $\mathbb{P}\left(\mathcal{E}^c\right) \geq 1 - 1/T \geq 1/2$ so $-\gamma M\rho + \frac{1}{2}\left(\sqrt{T}\log(T) - C_F\sqrt{T}\right) > 0$ since $\log(T) > C_F + \frac{2\gamma M\rho}{\sqrt{T}}$ for large $T$. Therefore Eq. (A.50) implies

$$\mathbb{E}\left[\sum_{t\in[T_2]}\sum_{j\in[M]}\left(V_j(\rho_{j,t}) - \gamma\rho_{j,t}\right)\right] \geq 0\,. \tag{A.51}$$

Finally, we have

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

where (i) follows from concavity of $V_j(\rho_j)$ according to Lemma 2.4.3.

**Part III. Bounding conversion error.** We first show that $T - T_2 \leq M + \frac{C_F}{\rho}\sqrt{T}$

where $C_F$ is an absolute constant independent of $T$ defined in Lemma 2.4.4 and $T_2$ is the last period of the algorithm.

If $T_2 = T$, the inequality $T - T_2 \le M + \frac{C_F}{\rho}\sqrt{T}$ holds trivially. Assume $T_2 < T$, the by the algorithm's termination criteria we have $\rho T \le B_{T_2} + M\rho = M\rho + \sum_{t\in[T_2]}\sum_{j\in[M]} \rho_{j,t}$, where we recall the definition of the spend balance $B_t = \sum_{t'\in[t]}\sum_{j\in[M]} \rho_{j,t'}$. Now, recalling $g_{2,t} = \rho - \sum_{j\in[M]} \rho_{j,t}$, we have

$$\begin{aligned} \rho T \;&\le\; T_2\rho - \sum_{t\in[T_2]} g_{2,t} + M\rho \;\overset{(i)}{\le}\; T_2\rho + \frac{C_F}{\eta} + M\rho \\ \Longrightarrow & T - T_2 \le M + \frac{C_F}{\rho}\sqrt{T}\,. \end{aligned} \tag{A.52}$$

where (i) follows from the proof of Lemma A.2.2 such that $\sum_{t=1}^{T_2} g_{2,t} \ge \frac{1}{\eta}\left(\lambda_1 - \lambda_{T_2}\right) \overset{(ii)}{\ge} -\frac{C_F}{\eta} = -C_F\sqrt{T}$, and (ii) follows from Lemma 2.4.4 such that $0 \le \lambda_t \le C_F$ for any $t \in [T]$ for some absolute constant $C_F > 0$ that only depends on the support of value-cost pairs $F = F_1 \times \ldots \times F_M$, and recall the SGD step size $\eta = 1/\sqrt{T}$.

So far, we have shown that $T - T_2 \le M + \frac{C_F}{\rho}\sqrt{T}$. Next, we bound $T \cdot \text{GL-OPT} - \mathbb{E}\left[\sum_{t\in[T_2]}\sum_{j\in[M]} V_j(\rho_{j,t})\right]$ using this result. Recalling the event $\mathcal{E} = \{T_1 > 2\sqrt{T}\log^3(T)\}$ defined in Lemma 2.4.7, we have

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

Here, (i) follows from Lemma 2.4.7 s.t. $\mathbb{P}(\mathcal{E}) \leq \frac{1}{T}$; in (ii) we used the fact that under event $\mathcal{E}^c$ we have $T_1 \leq 2\sqrt{T}\log^3(T)$ and also Eq. (A.52) that bounds $T - T_2$; in (iii), the term $(T_2 - T_1) \cdot \text{GL-OPT} - \sum_{t=T_1+1}^{T_2} \sum_{j\in[M]} V_j(\rho_{j,t})$ represents the convergence loss for UCB-SGD in Algorithm 1, which is in the order of $\mathcal{O}(T^{2/3})$ according to Theorem 2.4.6.

Finally, using concavity of $V_j(\rho_j)$ as illustrated in Lemma 2.4.3, we have GL-OPT$-$ $\mathbb{E}\left[\sum_{j\in[M]} V_j(\bar{\rho}_j)\right] \leq \text{GL-OPT} - \frac{1}{T}\mathbb{E}\left[\sum_{t\in[T_2]} \sum_{j\in[M]} V_j(\rho_{j,t})\right] \leq \mathcal{O}(-T^{1/3})$ □

### A.2.9 Additional Results for Section 2.4

Assumption 2.4.2 ensures that for any realization $\boldsymbol{z} = (\boldsymbol{v}, \boldsymbol{d})$ there must be some per-channel budget allocation that allows the advertiser to satisfy her ROI constraints as illustrated in the following proposition

**Proposition A.2.1** (Slater's condition)**.** *Assume Assumption 2.4.2 holds. Let* $\boldsymbol{z} = (\boldsymbol{v}, \boldsymbol{d}) = (\boldsymbol{v}_j, \boldsymbol{d}_j)_{j\in[M]} \in F_1 \times \dots F_M$ *be any realization of values and costs across all channels, then there exists some per-channel budget allocation* $(\rho_{(\boldsymbol{z}),j})_{j\in[M]} \in [0, \rho]^M$ *s.t.* $\sum_{j\in M} V_j(\rho_{(\boldsymbol{z}),j}; \boldsymbol{z}) > \gamma \sum_{j\in M} \rho_{(\boldsymbol{z}),j}$ *and* $\sum_{j\in[M]} \rho_{(\boldsymbol{z}),j} < \rho$.

*Proof.* Under Assumption 2.4.2, it is easy to see for any realization of value-cost pairs $\boldsymbol{z} = (\boldsymbol{v}_j, \boldsymbol{d}_j)_{j\in[M]}$ there always exists a channel $j \in [M]$ in which there is an auction $n \in [m_j]$ whose value-to-cost ration is at least $\gamma$, i.e. $v_{j,n} > \gamma d_{j,n}$. Since we assumed the ordering $\frac{v_{j,1}}{d_{j,1}} > \frac{v_{j,2}}{d_{j,2}} > \dots > \frac{v_{j,m_j}}{d_{j,m_j}}$, we know that $\frac{v_{j,1}}{d_{j,1}} \geq \frac{v_{j,n}}{d_{j,n}} > \gamma$. Now, in Eq. (A.10) within the proof of Lemma 2.4.3, we showed

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

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

**Lemma A.2.2** (Approximate constraint satisfaction)**.** *Assume Assumption 2.4.2 holds. If Algorithm 1 is run with stepsize $\eta = 1/\sqrt{T}$ such that $\eta < \frac{1}{M\sqrt{(\bar{V}+\gamma\rho)^2+\rho^2}}$, then we have*

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

*where we recall $g_{1,t} = \sum_{j\in[M]}\left(V_j(\rho_{j,t};\boldsymbol{z}_t) - \gamma\rho_{j,t}\right)$, $g_{2,t} = \rho - \sum_{j\in[M]}\rho_{j,t}$, and $C_F > 0$ is an absolute constant defined in Lemma 2.4.4 that depends only on the support $F = F_1 \times \ldots F_M$.*

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

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

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

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

From Lemma 2.4.4, we have $\lambda_T \leq C_F$ for some absolute constant $C_F > 0$, therefore we have

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

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

**Lemma A.2.3.** *Let $(\lambda_t, \mu_t)_{t\in[T]}$ be the dual variables generated by Algorithm 1. Then for any $\lambda, \mu \geq 0$ we have*

$$\begin{aligned}
(\lambda - \lambda_{t+1})^2 &\leq (\lambda - \lambda_t)^2 + 2\eta\left(\lambda - \lambda_t\right) g_{1,t} + \eta^2 M^2 (\bar{V} + \gamma\rho)^2 \\
(\mu - \mu_{t+1})^2 &\leq (\mu - \mu_t)^2 + 2\eta\left(\mu - \mu_t\right) g_{1,t} + \eta^2 M^2 \rho^2 \bar{V}^2 ,
\end{aligned} \tag{A.53}$$

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

*Proof.* We start with the first inequality w.r.t. $\lambda_t$'s. We have

$$(\lambda_t - \lambda)\, g_{1,t} = (\lambda_{t+1} - \lambda)\, g_{1,t} + (\lambda_t - \lambda_{t+1})\, g_{1,t}\,. \tag{A.54}$$

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

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

by rearranging terms in $(\lambda_{t+1} - (\lambda_t - \eta g_{1,t})) \cdot (\lambda - \lambda_{t+1}) \geq 0$, we get

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

where the equality can be checked easily by expending terms apart. Plugging the above back into Eq. (A.54) we get

$$\begin{aligned}(\lambda_t - \lambda)\, g_{1,t} &\leq (\lambda_t - \lambda_{t+1})\, g_{1,t} + \frac{1}{2\eta}\left((\lambda - \lambda_t)^2 - (\lambda - \lambda_{t+1})^2 - (\lambda_{t+1} - \lambda_t)^2\right) \\ &\overset{(i)}{\leq} \frac{\eta}{2} g_{1,t}^2 + \frac{1}{2\eta}\left((\lambda - \lambda_t)^2 - (\lambda - \lambda_{t+1})^2\right) \\ &\overset{(ii)}{\leq} \frac{\eta M^2 (\bar{V} + \gamma\rho)^2}{2} + \frac{1}{2\eta}\left((\lambda - \lambda_t)^2 - (\lambda - \lambda_{t+1})^2\right)\,,\end{aligned} \tag{A.55}$$

where (i) follows from $(\lambda_t - \lambda_{t+1})\, g_{1,t} - \frac{1}{2\eta}(\lambda_{t+1} - \lambda_t)^2 \leq \frac{\eta}{2} g_{1,t}^2$ using the basic inequality $ab \leq \frac{1}{2}a^2 + \frac{1}{2}b^2$ with $a^2 = \frac{1}{\eta}(\lambda_{t+1} - \lambda_t)^2$ and $b^2 = \eta g_{1,t}^2$; (ii) follows from the fact that $V_{j,t}(\rho_{j,t}) \leq \bar{V}$ for any $j \in [M]$ and $t \in [T]$ so $-M\gamma\rho \leq g_{1,t} \leq M\bar{V}$ so $g_{1,t}^2 \leq M^2(\bar{V} + \gamma\rho)^2$. Rearranging terms yields the first inequality in Eq. (A.53).

Following the same arguments above we can show for any $\mu \geq 0$ we have

$$(\mu_t - \mu)\, g_{1,t} \leq \frac{\eta M^2 \rho^2}{2} + \frac{1}{2\eta}\left((\mu - \mu_t)^2 - (\mu - \mu_{t+1})^2\right)\,.$$

□

**Proposition A.2.4.** *Under Assumption 2.4.2, the advertiser's per-channel only budget optimization problem, namely* CH-OPT($\mathcal{I}_B$) *is a convex problem.*

*Proof.* Recalling the CH-OPT($\mathcal{I}_B$) in Eq. (2.3) and the definition of $\mathcal{I}_B$ in Eq. (2.2), we can write CH-OPT($\mathcal{I}_B$) as

$$\begin{aligned} \text{CH-OPT}(\mathcal{I}_B) = \max_{(\rho_j)_{j\in[M]}\in\mathcal{I}} \quad & \sum_{j\in M} V_j(\rho_j) \\ \text{s.t.} \quad & \sum_{j\in M} V_j(\rho_j) \geq \gamma \sum_{j\in M} \rho_j \\ & \sum_{j\in[M]} \rho_j \leq \rho\,. \end{aligned} \tag{A.56}$$

Here, we used the definition $V_j(\rho_j) = \mathbb{E}\left[V_j(\rho_j; \boldsymbol{z}_j)\right]$ in Eq. (2.5), and $D_j(\rho_j; \boldsymbol{z}_j) = \rho_j$ for any $\boldsymbol{z}_j$ under the budget depletion Assumption 2.4.1. According to Lemma 2.4.3, $V_j(\rho_j)$ is concave in $\rho_j$ for any $j$, so the objective of CH-OPT($\mathcal{I}_B$) maximizes a concave function. For the feasibility region, assume $\rho_j$ and $\rho'_j$ are feasible, then defining $\rho''_j = \theta\rho_j + (1-\theta)\rho'_j$ for any $\theta \in [0,1]$, we know that

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

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

**Claim A.2.1.** *Assume Assumption 2.4.3 holds, then for any channel $j \in [M]$ and value-cost realization $\boldsymbol{z}_j = (\boldsymbol{v}_j, \boldsymbol{d}_j) \in F_j$, we have $\frac{v_{j,1}}{d_{j,1}} > \gamma$. This further implies that $\xi := \min_{j\in[M]}\min_{\boldsymbol{z}_j\in F_j}\frac{v_{j,1}}{d_{j,1}} > \gamma$. Further, let $\beta = \frac{1}{\log(T)}$. Then, for large enough $T$ we* *have* $V_j(\beta; \boldsymbol{z}_j) = \frac{v_{j,1}}{d_{j,1}}\beta \geq \xi\beta$ *for any realization* $\boldsymbol{z}_j \in F_j$*.*

*Proof.* Under Assumption 2.4.3, we know that for any realization of value-cost pairs $\boldsymbol{z}_j = (\boldsymbol{v}_j, \boldsymbol{d}_j)$ there always exists an auction $n \in [m_j]$ in this channel whose value-to-cost ratio is at least $\gamma$, i.e. $v_{j,n} > \gamma d_{j,n}$. Under the ordering $\frac{v_{j,1}}{d_{j,1}} > \frac{v_{j,2}}{d_{j,2}} > \cdots > \frac{v_{j,m_j}}{d_{j,m_j}}$, we have $\frac{v_{j,1}}{d_{j,1}} > \gamma$. In Eq. (A.10) within the proof of Lemma 2.4.3, we showed that for any realization $\boldsymbol{z}_j \in F_j$, $V_j(\rho_j; \boldsymbol{z}_j) = \frac{v_{j,1}}{d_{j,1}}\rho_j$ for all $\rho_j \leq d_{j,1}$. Hence we know that when $T$ is large enough such that $\beta = \frac{1}{\log(T)} < \min_{j\in[M]} \min_{\boldsymbol{z}_j \in F_j} d_{j,1}$, we always have $V_j(\beta; \boldsymbol{z}_j) = \frac{v_{j,1}}{d_{j,1}}\beta \geq \xi\beta$. □

## A.3 Additional material for Section 2.5

### A.3.1 Proof of Lemma 2.5.2

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

**Claim A.3.1.** *Recall* $v_{j,n}(1) > \ldots > v_{j,n}(L_{j,n}) > 0$ *and* $d_{j,n}(1) > \ldots > d_{j,n}(L_{j,n}) > 0$ *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* $\boldsymbol{z}_j = (\boldsymbol{v}_j, \boldsymbol{d}_j)$*, for any* $n \in [m_j]$ *we have* $\frac{v_{j,n}(\ell-1)-v_{j,n}(\ell)}{d_{j,n}(\ell-1)-d_{j,n}(\ell)}$ *decreases in* $\ell$ *then* $\frac{v_{j,n}(\ell)}{d_{j,n}(\ell)}$ *also decreases in* $\ell$*.*

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

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

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

$$\begin{aligned}\frac{v_{j,n}(\ell)}{d_{j,n}(\ell)} > \frac{v_{j,n}(\ell+1)}{d_{j,n}(\ell+1)} &\Longrightarrow \frac{d_{j,n}(\ell+1) - d_{j,n}(\ell)}{d_{j,n}(\ell)} > \frac{v_{j,n}(\ell+1) - v_{j,n}(\ell)}{v_{j,n}(\ell)} \\ &\Longrightarrow \frac{d_{j,n}(\ell) - d_{j,n}(\ell+1)}{d_{j,n}(\ell)} < \frac{v_{j,n}(\ell) - v_{j,n}(\ell+1)}{v_{j,n}(\ell)} \\ &\Longrightarrow \frac{v_{j,n}(\ell)}{d_{j,n}(\ell)} < \frac{v_{j,n}(\ell) - v_{j,n}(\ell+1)}{d_{j,n}(\ell) - d_{j,n}(\ell+1)}.\end{aligned} \tag{A.57}$$

Since $\frac{v_{j,n}(\ell-1)-v_{j,n}(\ell)}{d_{j,n}(\ell-1)-d_{j,n}(\ell)}$ decreases in $\ell$ we have

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

where (i) follows from Eq. (A.57), 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. □

Now we prove Lemma 2.5.2. Similar to the proof of Lemma 2.4.3, we only need to show for any realization $\boldsymbol{z}_j = (\boldsymbol{v}_j, \boldsymbol{d}_j)_{j\in[M]}$, the conversion function $V_j^+(\rho_j; \boldsymbol{z}_j) = \boldsymbol{v}_j^\top \boldsymbol{x}_j^{*,+}(\rho_j; \boldsymbol{z}_j)$ where $\boldsymbol{x}_j^{*,+}(\rho_j; \boldsymbol{z}_j)$ is defined as Eq. (2.21) is piecewise linear, continuous, strictly increasing and concave.

For simplicity we use the shorthand notation $\boldsymbol{x}_j^* = \boldsymbol{x}_j^{*,+}(\rho_j; \boldsymbol{z}_j) \in [0,1]^{\sum_{n\in[m_j]} L_{j,n}}$ as the optimal solution to $V_j^+(\rho_j; \boldsymbol{z}_j)$, defined in Eq. (2.21). By re-labeling the auction indices in channel $j \in [M]$ such that $\frac{v_{j,1}(1)}{d_{j,1}(1)} > \frac{v_{j,2}(1)}{d_{j,2}(1)} > \cdots > \frac{v_{j,m_j}(1)}{d_{j,m_j}(1)}$, we claim that $\boldsymbol{x}_j^*$ takes the following form:

$$x_{j,n}^*(\ell) = \begin{cases} 1 & \text{if } \ell = 1 \text{ and } \sum_{n'\in[n]} d_{j,n'}(1) \le \rho_j \\ \frac{\rho_j - \sum_{n'\in[n-1]} d_{j,n'}(1)}{d_{j,n}(1)} & \text{if } \ell = 1 \text{ and } \sum_{n'\in[n]} d_{j,n'}(1) > \rho_j \\ 0 & \text{otherwise} \end{cases} \tag{A.58}$$

which is analogous to that of Eq. (A.9) in the proof of Lemma 2.4.3. 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 $\boldsymbol{x}_j^*$, the rest of the proof follows exactly from that for Lemma 2.4.3.

It now remains to show that Eq. (A.58) holds. We first argue by contradiction that in any auction, no impression other than the first would get procured, i.e. $x_{j,n}^*(\ell) = 0$ for any $\ell \in 2 \ldots L_{j,n}$. Assume there exists some auction $n \in [m_j]$ and impression slot

$\ell' \in 2 \ldots L_{j,n}$ such that $x^*_{j,n}(\ell') > 0$, then by the constraint that at most 1 impression can be procured, i.e. $\sum_{\ell \in [L_{j,n}]} x^*_{j,n}(\ell) \leq 1$ in Eq. (2.21), we know that $x^*_{j,n}(1) < 1$. Also, note that $x^*_{j,n}(\ell')$ incurs a cost of $d_{j,n}(\ell') \cdot x^*_{j,n}(\ell')$ amongst the total per-channel budget $\rho_j$. If we instead use this cost on the first impression, then we will obtain a value increase of

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

where the final inequality follows from the assumption that $x^*_{j,n}(\ell') > 0$, and the multi-item auction has increasing marginal values (see Definition 2.5.1) so Claim A.3.1 holds. This contradicts the optimality of $\boldsymbol{x}^*_j$, and hence $x^*_{j,n}(\ell) = 0$ for any $\ell \in 2 \ldots L_{j,n}$, or in other words, a channel will only procure impressions ranked first. Hence, a channel's procurement problem in Eq. (2.21) can be restricted to the first impression in each auction, and thus similar to the proof of Lemma 2.4.3, is an LP-relaxation to the 0-1 knapsack with budget $\rho_j$, and $m_j$ items whose values are $v_{j,1}(1) \ldots v_{j,m_j}(1)$ with costs $d_{j,1}(1) \ldots d_{j,m_j}(1)$. $\square$

# Appendix B

# Supplementary material for Chapter 3

## B.1 Additional material for Section 3.3

### B.1.1 Proof for Proposition 3.3.1

Let $Q_t(\cdot)$ be the distributions of a buyer's valuation when we condition on the feature vector $x_t$. Further, let $Q_t^-(\cdot)$ be the distribution of $v_t^-$, which is the second highest valuation at time $t$. Then, we have $Q_t(z) = F(z - \langle \beta, x_t \rangle)$ and $Q_t^-(z) = F^-(z - \langle \beta, x_t \rangle)$. When $N \geq 2$ and all buyers bid truthfully, according to Equations (3.2) , the seller's expected revenue conditioned on $x_t$ by setting reserve price $r_t$ is:

$$\begin{aligned} \mathsf{rev}_t(r_t) &= \mathbb{E}\left[\max\{r_t, v_t^-\}\mathbb{I}\{v_t^+ \geq r_t\} \mid x_t, r_t\right] \\ &= \mathbb{E}\left[r_t \mathbb{I}\{v_t^+ \geq r_t \geq v_t^-\} + v_t^- \mathbb{I}\{v_t^+ \geq v_t^- \geq r_t\} \mid x_t, r_t\right] , \end{aligned} \tag{B.1}$$

where $v_t^+$ is the highest valuation at time $t$. The first term within the expectation, conditioned on $x_t$ and $r_t$, is

$$\mathbb{E}\left[r_t \mathbb{I}\{v_t^+ \geq r_t \geq v_t^-\} \mid x_t, r_t\right] = r_t N \left[Q_t(r_t)\right]^{N-1} \left[1 - Q_t(r_t)\right], \tag{B.2}$$

where we used the fact that $r_t$ is independent of $v_t^+$ and $v_t^-$ since the seller sets reserve price $r_t$ based on only the past history $\mathcal{H}_{t-1} = \{(r_1, v_1, x_1), (r_2, v_2, x_2), \ldots, (r_{t-1}, v_{t-1}, x_{t-1})\}$, and both $v_t^+$ and $v_t^-$, conditioned on $x_t$, are independent of the past. The second term within the expectation of Equation (B.1) is

$$\begin{aligned}
& \mathbb{E}\left[v_t^- \mathbb{I}\{v_t^+ \geq v_t^- \geq r_t\} \mid x_t, r_t\right] \\
= \;& \mathbb{E}\left[v_t^- \mathbb{I}\{v_t^- \geq r_t\} \mid x_t, r_t\right] \\
= \;& \mathbb{E}\left[(v_t^- - r_t)\mathbb{I}\{v_t^- \geq r_t\} \mid x_t, r_t\right] + r_t \mathbb{E}\left[\mathbb{I}\{v_t^- \geq r_t\} \mid x_t, r_t\right] \\
= \;& \int_0^\infty \mathbb{P}\left(v_t^- - r_t \geq z\right) dz + r_t\left[1 - Q_t^-(r_t)\right] \\
= \;& \int_{r_t}^\infty \left[1 - Q_t^-(z)\right] dz + r_t\left[1 - Q_t^-(r_t)\right] \\
= \;& \mathbb{E}\left[v_t^- \mid x_t, r_t\right] - \int_0^{r_t} \left[1 - Q_t^-(z)\right] dz + r_t\left[1 - Q_t^-(r_t)\right] \\
= \;& \mathbb{E}\left[v_t^- \mid x_t\right] + \int_0^{r_t} Q_t^-(z) dz - r_t Q_t^-(r_t)\,. 
\end{aligned} \tag{B.3}$$

Note that the integration starts from 0 because all valuations are considered to be positive. Since $F^-(\tilde{z}) := NF^{N-1}(\tilde{z}) - (N-1)F^N(\tilde{z})$ for any $\tilde{z} \in \mathbb{R}$, we have

$$Q_t^-(r_t) \;=\; N\left[Q_t(r_t)\right]^{N-1}\left[1 - Q_t(r_t)\right] + \left[Q_t(r_t)\right]^N\,. \tag{B.4}$$

Hence, combining Equations (B.1), (B.2), (B.3), and (B.4), we have

$$\begin{aligned}
\mathsf{rev}_t(r_t) \;&=\; \mathbb{E}\left[v_t^- \mid x_t\right] + \int_0^{r_t} Q_t^-(z) dz - r_t\left[Q_t(r_t)\right]^N \\
&=\; \mathbb{E}\left[v_t^- \mid x_t\right] + \int_0^{r_t} F^-(z - \langle\beta, x_t\rangle) dz - r_t\left[F^+(r_t - \langle\beta, x_t\rangle)\right] \\
&=\; \int_{-\infty}^\infty z dF^-(z) + \langle\beta, x_t\rangle + \int_0^{r_t} F^-(z - \langle\beta, x_t\rangle) dz - r_t\left[F^+(r_t - \langle\beta, x_t\rangle)\right]\,.
\end{aligned}$$

□

## B.2 Additional material for Section 3.4

### B.2.1 Proof of Theorem 3.4.2

We first introduce some definitions that we will extensively rely on throughout our proof of Theorem 3.4.2. We start off with the "good" events $\xi_{\ell+1}$, $\xi_{\ell+1}^-$ and $\xi_{\ell+1}^+$ for

$\ell \geq 1$ in which the estimates of $\beta$, $F^-$ and $F^+$ are accurate:

$$\xi_{\ell+1} = \left\{ \|\widehat{\beta}_{\ell+1} - \beta\|_1 \leq \frac{\delta_\ell}{x_{\max}} \right\} \tag{B.5}$$

$$\text{where} \;\; \delta_\ell := \frac{\sqrt{2d\log(|E_\ell|)}\epsilon_{\max}x_{\max}^2}{\lambda_0^2\sqrt{N|E_\ell|}} + \frac{\sqrt{d}\,(NL_\ell a_{\max}+1)\,x_{\max}^2}{|E_\ell|\lambda_0^2}\,, \tag{B.6}$$

$$\xi_{\ell+1}^- = \left\{ \left|\widehat{F}_{\ell+1}^-(z) - F^-(z)\right| \;\leq\; 2N^2\left(\gamma_\ell + c_f\delta_\ell + \frac{c_f + NL_\ell}{|E_\ell|}\right)\right\}\,, \tag{B.7}$$

$$\xi_{\ell+1}^+ = \left\{ \left|\widehat{F}_{\ell+1}^+(z) - F^+(z)\right| \;\leq\; N\left(\gamma_\ell + c_f\delta_\ell + \frac{c_f + NL_\ell}{|E_\ell|}\right)\right\}\,, \tag{B.8}$$

where $a_{\max}$ is the maximum possible corruption, $\gamma_\ell = \sqrt{\log(|E_\ell|)}/\sqrt{2N|E_\ell|}$, $\lambda_0^2$ is the minimum eigenvalue of covariance matrix $\Sigma$, and $c_f = \sup_{z\in[-\epsilon_{\max},\epsilon_{\max}]} f(z) \geq \inf_{z\in[-\epsilon_{\max},\epsilon_{\max}]} f(z) > 0$. Furthermore,

$$L_\ell = \frac{\log\left(v_{\max}^2 N|E_\ell|^4 - 1\right)}{\log(1/\eta)} = \mathcal{O}\left(\frac{\log(|E_\ell|)}{\log(1/\eta)}\right)\,,$$

where $|E_\ell| = T^{1-2^{-\ell}}$ is the length of the $\ell^{th}$ phase.

We also define the event that the number of periods in phase $E_\ell$ during which buyer $i$ submits significantly corrupted bids is bounded by $L_\ell$:

$$\mathcal{G}_{i,\ell} := \{|\mathcal{S}_{i,\ell}| \leq L_\ell\}\;. \tag{B.9}$$

Here, $\mathcal{S}_{i,\ell} = \left\{t \in E_\ell : |a_{i,t}| \geq \frac{1}{|E_\ell|}\right\}$ is the set of all periods in phase $E_\ell$ during which buyer $i$ extensively corrupts her bids.

We are now equipped to show Theorem 3.4.2 according to the following steps:

(i) Decompose the single period regret into $\mathcal{R}_t^{(1)}$ and $\mathcal{R}_t^{(2)}$, where $\mathcal{R}_t^{(1)}$ bounds the expected revenue loss due to the discrepancy between the actual reserve price $r_t$ and the optimal reserve price $r_t^\star$ and $\mathcal{R}_t^{(2)}$, which bounds the expected revenue loss due to allocation mismatches. Note that $\mathcal{R}_t^{(1)}$ is a result of the estimation inaccuracies in $\beta$, $F^-$ and $F^+$.

(ii) Bound $\mathcal{R}_t^{(1)}$ using Lemmas 3.4.1, B.2.1, B.2.2, and B.2.3.

(iii) Bound $\mathcal{R}_t^{(2)}$ using Lemmas 3.4.1 and 3.4.4.

(iv) Sum up $\mathcal{R}_t^{(1)}$ and $\mathcal{R}_t^{(2)}$ to bound the cumulative expected regret over a phase $E_\ell$ and the entire horizon $T$.

**(i) Decomposing single period regret into $\mathcal{R}_t^{(1)}$ and $\mathcal{R}_t^{(2)}$:** According to the NPAC-S policy detailed in Algorithm 3, the expected revenue in period $t$ is given by

$$\begin{aligned}\mathsf{rev}_t(r_t) \;=\; \mathbb{E}\Big[ & \max\{b_t^-, \widehat{r}_t\}\mathbb{I}\{b_t^+ > \widehat{r}_t\}\mathbb{I}\{\text{no isolation in } t\} \\ & + \sum_{i\in[N]} r_t^u \mathbb{I}\{b_{i,t} > r_t^u\}\mathbb{I}\{i \text{ is isolated}\} \;\mid\; x_t, r_t\Big],\end{aligned} \tag{B.10}$$

where the expectation is taken with respect to $\{(x_\tau, \epsilon_{i,\tau}, a_{i,\tau})\}_{\tau\in[t],i\in[N]}$ and $\widehat{r}_t, r_t^u$ are defined in Equations (3.6) and (3.7) respectively. Hence, the regret is given by

$$\begin{aligned}\mathsf{Regret}_t \;&=\; \mathbb{E}\left[\mathrm{REV}_t^\star - \mathsf{rev}_t(r_t)\right] \\ &=\; \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\} - \mathsf{rev}_t(r_t)\right] \\ &=\; \Big(\mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\}\right] - \mathbb{E}\left[\max\{v_t^-, \widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}\mathbb{I}\{\text{no isolation in } t\}\right]\Big) \\ &\quad\; + \Big(\mathbb{E}\left[\max\{v_t^-, \widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}\mathbb{I}\{\text{no isolation in } t\} - \mathsf{rev}_t(r_t)\right]\Big) \\ &:=\; \mathcal{R}_t^{(1)} + \mathcal{R}_t^{(2)},\end{aligned} \tag{B.11}$$

where the expectation is taken with respect the context $x_t \sim \mathcal{D}$ and the randomness in $r_t$; $r_t^\star$ is the optimal reserve price (defined in Equation (3.4)) if the seller has full knowledge of $F$ and $\beta$; and we defined:

$$\begin{aligned}\mathcal{R}_t^{(1)} &:= \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\}\right] - \mathbb{E}\left[\max\{v_t^-, \widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}\mathbb{I}\{\text{no isolation in } t\}\right] \\ \mathcal{R}_t^{(2)} &:= \mathbb{E}\left[\max\{v_t^-, \widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}\mathbb{I}\{\text{no isolation in } t\} - \mathsf{rev}_t(r_t)\right]\end{aligned} \tag{B.12}$$

**(ii) Bounding $\mathcal{R}_t^{(1)}$:** We start by upper bounding $\mathcal{R}_t^{(1)}$ for a period $t \in E_{\ell+1}$

where $\ell \geq 1$.

$$\begin{aligned}
\mathcal{R}_t^{(1)} &= \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\}\right] - \mathbb{E}\left[\max\{v_t^-, \widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}\mathbb{I}\{\text{no isolation in } t\}\right] \\
&= \mathbb{E}\left[\left(\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\} - \max\{v_t^-, \widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}\right)\mathbb{I}\{\text{no isolation in } t\}\right] \\
&\quad + \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\}\left(1 - \mathbb{I}\{\text{no isolation in } t\}\right)\right] \\
&= \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\} - \max\{v_t^-, \widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}\right]\left(1 - \frac{1}{|E_\ell|}\right) \\
&\quad + \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\}\right] \cdot \frac{1}{|E_\ell|} \\
&\leq \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\} - \max\{v_t^-, \widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}\right] + \frac{v_{\max}}{|E_\ell|}, \qquad \text{(B.13)}
\end{aligned}$$

where the third equality is because an isolation event is independent of any other event, and the final inequality follows from a simple observation that $\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\} \leq v_{\max}$.

For simplicity, we define

$$\widetilde{\mathcal{R}}_t^{(1)} := \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\} - \max\{v_t^-, \widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\} \;\middle|\; x_t, \widehat{r}_t\right],$$

so Equation (B.13) yields

$$\mathcal{R}_t^{(1)} \leq \mathbb{E}\left[\widetilde{\mathcal{R}}_t^{(1)}\right] + \frac{v_{\max}}{|E_\ell|}, \qquad \text{(B.14)}$$

where the expectation is taken with respect to the context $x_t$ and reserve price $\widehat{r}_t$. Notice that $\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\} - \max\{v_t^-, \widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}$ is exactly the revenue difference $\mathsf{rev}_t(r_t^\star) - \mathsf{rev}_t(r_t)$ had the seller set reserve prices $r_t^\star$ or $r_t$ when all buyers bid truthfully. Hence, by applying Proposition 3.3.1 we obtain

$$\begin{aligned}
\widetilde{\mathcal{R}}_t^{(1)} &= \int_0^{r_t^\star} F^-(z - \langle\beta, x_t\rangle)dz - r_t^\star\left[F^+(r_t^\star - \langle\beta, x_t\rangle)\right] \\
&\quad - \int_0^{\widehat{r}_t} F^-(z - \langle\beta, x_t\rangle)dz + \widehat{r}_t\left[F^+(\widehat{r}_t - \langle\beta, x_t\rangle)\right].
\end{aligned}$$

Note that we can apply Proposition 3.3.1 because $\widehat{r}_t$ is the reserve price set according to the NPAC-S policy when no isolation occurs, and only depends on the current context $x_t$ and the past $\mathcal{H}_{t-1} = \{(r_1, b_1, x_1), (r_2, b_2, x_2), \dots, (r_{t-1}, b_{t-1}, x_{t-1})\}$.

By defining $y_t := \langle \beta, x_t \rangle$, $\widehat{y}_t := \langle \widehat{\beta}_\ell, x_t \rangle$ and

$$\rho_t(r, y, F^{(1)}, F^{(2)}) := \int_0^r F^{(2)}(z-y)dz - r\left[F^{(1)}(r-y)\right] \,, \tag{B.15}$$

we can rewrite $\widetilde{\mathcal{R}}_t^{(1)}$ as the following:

$$\begin{aligned}
\widetilde{\mathcal{R}}_t^{(1)} &= \mathbb{E}\left[\max\{v_t^-, r_t^\star\}\mathbb{I}\{v_t^+ > r_t^\star\} - \max\{v_t^-, \widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\} \;\middle|\; x_t, \widehat{r}_t\right] \\
&= \rho_t(r_t^\star, y_t, F^-, F^+) - \rho_t(\widehat{r}_t, y_t, F^-, F^+) \\
&= \rho_t(r_t^\star, y_t, F^-, F^+) - \rho_t(r_t^\star, \widehat{y}_t, F^-, F^+) \\
&\quad + \rho_t(r_t^\star, \widehat{y}_t, F^-, F^+) - \rho_t(r_t^\star, \widehat{y}_t, \widehat{F}_{\ell+1}^-, \widehat{F}_{\ell+1}^+) \\
&\quad + \rho_t(r_t^\star, \widehat{y}_t, \widehat{F}_{\ell+1}^-, \widehat{F}_{\ell+1}^+) - \rho_t(\widehat{r}_t, \widehat{y}_t, \widehat{F}_{\ell+1}^-, \widehat{F}_{\ell+1}^+) \\
&\quad + \rho_t(\widehat{r}_t, \widehat{y}_t, \widehat{F}_{\ell+1}^-, \widehat{F}_{\ell+1}^+) - \rho_t(\widehat{r}_t, \widehat{y}_t, F^-, F^+) \\
&\quad + \rho_t(\widehat{r}_t, \widehat{y}_t, F^-, F^+) - \rho_t(\widehat{r}_t, y_t, F^-, F^+) \,.
\end{aligned} \tag{B.16}$$

We now invoke Lemma B.2.3, where we show that when events $\xi_{\ell+1}$, $\xi_{\ell+1}^-$ and $\xi_{\ell+1}^+$ (see definition in Equation (B.5),(B.6), (B.7) and (B.8) ) happen for some phase $\ell \geq 1$, we have for $r \in \{r_t^\star, \widehat{r}_t\}$,

(i) $|\rho_t(r, y_t, F^-, F^+) - \rho_t(r, \widehat{y}_t, F^-, F^+)| \;\leq\; 3rc_f N^2 \delta_\ell$ a.s.

(ii) $\left|\rho_t(r, \widehat{y}_t, F^-, F^+) - \rho_t(r, \widehat{y}_t, \widehat{F}_{\ell+1}^-, \widehat{F}_{\ell+1}^+)\right| \;\leq\; 3rN^2\left(\gamma_\ell + c_f\delta_\ell + \frac{c_f + L_\ell}{|E_\ell|}\right)$ a.s.

Note that the first inequality bounds the impact of errors $\beta$ and the second bounds the impact of errors in the distributions. Applying these bounds in (B.16), we get

$$\begin{aligned}\widetilde{\mathcal{R}}_t^{(1)} \cdot \mathbb{I}\left\{\xi_{\ell+1} \cap \xi_{\ell+1}^- \cap \xi_{\ell+1}^+\right\} \ \leq\ & 3(r_t^\star + \widehat{r}_t)c_f N^2 \delta_\ell \\ & + 3(r_t^\star + \widehat{r}_t) N^2 \left(\gamma_\ell + c_f\delta_\ell + \frac{c_f + L_\ell}{|E_\ell|}\right) \\ & + \rho_t(r_t^\star, \widehat{y}_t, \widehat{F}_{\ell+1}^-, \widehat{F}_{\ell+1}^+) - \rho_t(\widehat{r}_t, \widehat{y}_t, \widehat{F}_{\ell+1}^-, \widehat{F}_{\ell+1}^+)\,.\end{aligned} \tag{B.17}$$

We recall that the seller's pricing decision $\widehat{r}_t$ when no isolation occurs is defined in Equation (3.7), and realize that in fact $\widehat{r}_t = \arg\max_{r\in(0,v_{\max}]} \rho_t(r, \widehat{y}_t, \widehat{F}_{\ell+1}^-, \widehat{F}_{\ell+1}^+)$. So, by the optimality of $\widehat{r}_t$ and $r_t^\star \leq v_{\max}$, we obtain the fact that $\rho_t(r_t^\star, \widehat{y}_t, \widehat{F}_{\ell+1}^-, \widehat{F}_{\ell+1}^+) - \rho_t(\widehat{r}_t, \widehat{y}_t, \widehat{F}_{\ell+1}^-, \widehat{F}_{\ell+1}^+) \leq 0$. Using this inequality in (B.17), we get

$$\begin{aligned}& \widetilde{\mathcal{R}}_t^{(1)} \cdot \mathbb{I}\left\{\xi_{\ell+1} \cap \xi_{\ell+1}^- \cap \xi_{\ell+1}^+\right\} \\ \leq\ & 6v_{\max} c_f N^2 \delta_\ell + 6 v_{\max} N^2 \left(\gamma_\ell + c_f\delta_\ell + \frac{c_f + L_\ell}{|E_\ell|}\right) \\ =\ & 12 v_{\max} c_f N^2 \delta_\ell + 6 v_{\max} N^2 \left(\frac{\sqrt{\log(|E_\ell|)}}{\sqrt{2N|E_\ell|}} + \frac{c_f + L_\ell}{|E_\ell|}\right) \\ =\ & 12 v_{\max} c_f N^2 \delta_\ell + \frac{6v_{\max}\sqrt{N^3 \log(|E_\ell|)}}{\sqrt{2E_\ell}} + \frac{6 v_{\max} N^2 (c_f + L_\ell)}{|E_\ell|}\,,\end{aligned} \tag{B.18}$$

where we used the fact that $r_t^\star, \widehat{r}_t \leq v_{\max}$ in the inequality. Note that $L_\ell = \log\left(v_{\max}^2 N |E_\ell|^4 - 1\right)/\log(\frac{1}{\eta}) = \mathcal{O}\left(\log(T)/\log(1/\eta)\right)$, since we recall that $|E_\ell| = T^{1-2^{-\ell}}$.

To complete the bound for $\mathcal{R}_t^{(1)}$ in period $t \in E_{\ell+1}$, we continue to bound Equation (B.14):

$$\begin{aligned}\mathcal{R}_t^{(1)} \ \leq\ & \mathbb{E}\left[\widetilde{\mathcal{R}}_t^{(1)}\right] + \frac{v_{\max}}{|E_\ell|} \\ =\ & \mathbb{E}\left[\widetilde{\mathcal{R}}_t^{(1)} \cdot \mathbb{I}\left\{\xi_{\ell+1} \cap \xi_{\ell+1}^- \cap \xi_{\ell+1}^+\right\}\right] + \mathbb{E}\left[\widetilde{\mathcal{R}}_t^{(1)} \cdot \mathbb{I}\left\{\xi_{\ell+1}^c \cup \left(\xi_{\ell+1}^-\right)^c \cup \left(\xi_{\ell+1}^+\right)^c\right\}\right] + \frac{v_{\max}}{|E_\ell|} \\ \leq\ & \mathbb{E}\left[\widetilde{\mathcal{R}}_t^{(1)} \cdot \mathbb{I}\left\{\xi_{\ell+1} \cap \xi_{\ell+1}^- \cap \xi_{\ell+1}^+\right\}\right] + v_{\max}\mathbb{P}\left(\xi_{\ell+1}^c \cup \left(\xi_{\ell+1}^-\right)^c \cup \left(\xi_{\ell+1}^+\right)^c\right) + \frac{v_{\max}}{|E_\ell|} \\ \leq\ & 12 v_{\max} c_f N^2 \delta_\ell + \frac{6v_{\max}\sqrt{N^3 \log(|E_\ell|)}}{\sqrt{2E_\ell}} + \frac{v_{\max}\left(6N^2(c_f + L_\ell) + 9N + 15d + 9\right)}{|E_\ell|}\,,\end{aligned} \tag{B.19}$$

where the second inequality follows from a simple observation that $\widetilde{\mathcal{R}}_t^{(1)} \leq v_{\max}$ almost surely, and the third inequality uses Equation (B.18) and Lemma B.2.4, which shows $\mathbb{P}\left(\xi_{\ell+1}^c \cup \left(\xi_{\ell+1}^-\right)^c \cup \left(\xi_{\ell+1}^+\right)^c\right) \;\leq\; (9N + 15d + 8)/|E_\ell|$,

**(iii) Bounding $\mathcal{R}_t^{(2)}$:** So far, we have bounded $\mathcal{R}_t^{(1)}$ for $t \in E_{\ell+1}$ ($\ell \geq 1$), and will move on to bound $\mathcal{R}_t^{(2)}$ defined in Equation (B.11) for $t \in E_\ell$ for any $\ell \geq 1$ . We define

$$b_{-i,t}^+ = \max_{j \neq i} b_{j,t} \quad \text{and} \quad v_{-i,t}^+ = \max_{j \neq i} v_{j,t}\,, \tag{B.20}$$

which represent the highest bid excluding that of buyer $i$, and the highest valuation excluding that of buyer $i$, respectively. We then have

$$
\begin{aligned}
&\mathcal{R}_t^{(2)} \\
=\ & \mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}\mathbb{I}\{\text{no isolation in } t\} - \mathsf{rev}_t(r_t)\right] \\
\leq\ & \mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}\mathbb{I}\{\text{no isolation in } t\}\right] \\
&\quad - \mathbb{E}\left[\max\{b_t^-, r_t\}\mathbb{I}\{b_t^+ > \widehat{r}_t\}\mathbb{I}\{\text{no isolation in } t\}\right] \\
=\ & \left(\mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}\right] - \mathbb{E}\left[\max\{b_t^-,\widehat{r}_t\}\mathbb{I}\{b_t^+ > \widehat{r}_t\}\right]\right) \cdot \left(1 - \frac{1}{|E_\ell|}\right) \\
<\ & \mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{v_t^+ > \widehat{r}_t\}\right] - \mathbb{E}\left[\max\{b_t^-,\widehat{r}_t\}\mathbb{I}\{b_t^+ > \widehat{r}_t\}\right] \\
=\ & \sum_{i\in[N]} \mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{v_{i,t} > \max\{v_{-i,t}^+,\widehat{r}_t\}\} - \max\{b_t^-,\widehat{r}_t\}\mathbb{I}\{b_{i,t} > \max\{b_{-i,t}^+\widehat{r}_t\}\}\right] \\
=\ & \sum_{i\in[N]} \mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{\max\{v_{-i,t}^+,\widehat{r}_t\} < v_{i,t} < \max\{b_{-i,t}^+\widehat{r}_t\}\}\right] \\
&\quad - \sum_{i\in[N]} \mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{\max\{b_{-i,t}^+\widehat{r}_t\} < v_{i,t} < \max\{v_{-i,t}^+,\widehat{r}_t\}\}\right] \\
&\quad + \sum_{i\in[N]} \mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{v_{i,t} > \max\{b_{-i,t}^+\widehat{r}_t\}\} - \max\{b_t^-,\widehat{r}_t\}\mathbb{I}\{b_{i,t} > \max\{b_{-i,t}^+\widehat{r}_t\}\}\right] \\
\leq\ & \sum_{i\in[N]} \mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{\max\{v_{-i,t}^+,\widehat{r}_t\} < v_{i,t} < \max\{b_{-i,t}^+\widehat{r}_t\}\}\right] \\
&\quad + \sum_{i\in[N]} \mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{v_{i,t} > \max\{b_{-i,t}^+\widehat{r}_t\}\} - \max\{b_t^-,\widehat{r}_t\}\mathbb{I}\{b_{i,t} > \max\{b_{-i,t}^+\widehat{r}_t\}\}\right] \\
\leq\ & \sum_{i\in[N]} v_{\max}\mathbb{E}\left[\mathbb{I}\{\max\{v_{-i,t}^+,\widehat{r}_t\} < v_{i,t} < \max\{b_{-i,t}^+\widehat{r}_t\}\}\right] \\
&\quad + \sum_{i\in[N]} \mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{v_{i,t} > \max\{b_{-i,t}^+\widehat{r}_t\}\} - \max\{b_t^-,\widehat{r}_t\}\mathbb{I}\{b_{i,t} > \max\{b_{-i,t}^+\widehat{r}_t\}\}\right] ,
\end{aligned}
\tag{B.21}
$$

where the first inequality follows from Equation (B.10); the third inequality is due to the fact that $\sum_{i\in[N]} \mathbb{E}\left[\max\{v_t^-,\widehat{r}_t\}\mathbb{I}\{\max\{b_{-i,t}^+\widehat{r}_t\} < v_{i,t} < \max\{v_{-i,t}^+,\widehat{r}_t\}\}\right] \geq 0$; and the last inequality holds because $\max\{v_t^-,\widehat{r}_t\} \leq v_{\max}$. To continue the bound for

Equation (B.21), we use the definition of $\mathcal{B}_{i,\ell} := \mathcal{B}^s_{i,\ell} \cup \mathcal{B}^o_{i,\ell}$ in Lemma 3.4.4, where

$$\mathcal{B}^s_{i,\ell} = \left\{t \in E_\ell : \mathbb{I}\left\{v_{i,t} > \{b^+_{-i,t}, \widehat{r}_t\}\right\} = 1 \ , \ \mathbb{I}\left\{b_{i,t} > \{b^+_{-i,t}, \widehat{r}_t\}\right\} = 0\right\}$$
$$\mathcal{B}^o_{i,\ell} = \left\{t \in E_\ell : \mathbb{I}\left\{v_{i,t} > \{b^+_{-i,t}, \widehat{r}_t\}\right\} = 0 \ , \ \mathbb{I}\left\{b_{i,t} > \{b^+_{-i,t}, \widehat{r}_t\}\right\} = 1\right\} .$$

Here, $\mathcal{B}^s_{i,\ell}$ represents the periods during which buyer $i$ could have won the auction had she bid truthfully but in reality lost since she shaded her bid (allocation mismatch due to shading), while $\mathcal{B}^o_{i,\ell}$ represents the periods when buyer $i$ would have lost the auction had she bid truthfully, but instead won the item due to overbidding (allocation mismatch due to overbidding). Hence, for any period $t \in E_\ell/\mathcal{B}_{i,\ell} = \left\{t \in E_\ell : \mathbb{I}\left\{v_{i,t} > \{b^+_{-i,t}, \widehat{r}_t\}\right\} = \mathbb{I}\left\{b_{i,t} > \{b^+_{-i,t}, \widehat{r}_t\}\right\}\right\}$ (which means in period $t \in E_\ell/\mathcal{B}_{i,\ell}$ the outcome for buyer $i$ would not have changed even if she bid truthfully), we have $\mathbb{I}\{v_{i,t} > \max\{b^+_{-i,t}, \widehat{r}_t\}\} = \mathbb{I}\{b_{i,t} > \max\{b^+_{-i,t}, \widehat{r}_t\}\}$. Therefore, defining $\mathcal{B}_\ell := \cup_{i\in[N]}\mathcal{B}_{i,\ell}$, we have

$$\begin{aligned}
&\mathcal{R}^{(2)}_t \mathbb{I}\{t \in E_\ell/\mathcal{B}_\ell\} \\
\leq\ & \sum_{i\in[N]} v_{\max}\mathbb{E}\left[\mathbb{I}\{\max\{v^+_{-i,t}, \widehat{r}_t\} < v_{i,t} < \max\{b^+_{-i,t}\widehat{r}_t\}\}\right] \\
&\quad + \sum_{i\in[N]} \mathbb{E}\Big[\max\{v^-_t, \widehat{r}_t\}\mathbb{I}\{v_{i,t} > \max\{b^+_{-i,t}\widehat{r}_t\}\} \\
&\quad - \max\{b^-_t, \widehat{r}_t\}\mathbb{I}\{b_{i,t} > \max\{b^+_{-i,t}\widehat{r}_t\}\}\Big]\mathbb{I}\{t \in E_\ell/\mathcal{B}_\ell\} \\
=\ & \sum_{i\in[N]} v_{\max}\mathbb{E}\left[\mathbb{I}\{\max\{v^+_{-i,t}, \widehat{r}_t\} < v_{i,t} < \max\{b^+_{-i,t}\widehat{r}_t\}\}\right] \\
&\quad + \sum_{i\in[N]} \mathbb{E}\left[\left(\max\{v^-_t, \widehat{r}_t\} - \max\{b^-_t, \widehat{r}_t\}\right)\mathbb{I}\{b_{i,t} > \max\{b^+_{-i,t}, \widehat{r}_t\}\}\right] \\
\leq\ & \sum_{i\in[N]} v_{\max}\mathbb{E}\left[\mathbb{I}\{\max\{v^+_{-i,t}, \widehat{r}_t\} < v_{i,t} < \max\{b^+_{-i,t}, \widehat{r}_t\}\}\right] + \mathbb{E}\left[\max\{v^-_t, \widehat{r}_t\} - \max\{b^-_t, \widehat{r}_t\}\right] \\
\leq\ & \sum_{i\in[N]} v_{\max}\mathbb{E}\left[\mathbb{I}\{\max\{v^+_{-i,t}, \widehat{r}_t\} < v_{i,t} < \max\{b^+_{-i,t}, \widehat{r}_t\}\}\right] + \mathbb{E}\left[\left(v^-_t - b^-_t\right)^+\right] .
\end{aligned}$$

The first inequality follows from Equation (B.21); the first equality follows from the fact that $t \in E_\ell/\mathcal{B}_\ell$; the second inequality holds because

$$\sum_{i\in[N]} \mathbb{I}\{b_{i,t} > \max\{b^+_{-i,t}\widehat{r}_t\}\} \ \leq \ \sum_{i\in[N]} \mathbb{I}\{b_{i,t} > b^+_{-i,t}\}\} = 1$$

The third inequality applies the fact that $\max\{a,c\} - \max\{b,c\} \leq (a-b)^+$ for any $a,b,c \in \mathbb{R}$. Denoting $i^\star := \arg\max_{i\in[N]} v_{i,t}$, we have

$$\begin{aligned}
&\sum_{i\in[N]} v_{\max}\mathbb{E}\left[\mathbb{I}\{\max\{v^+_{-i,t},\widehat{r}_t\} < v_{i,t} < \max\{b^+_{-i,t},\widehat{r}_t\}\}\right] \\
&= \ v_{\max}\mathbb{E}\left[\mathbb{I}\{\max\{v^+_{-i^\star,t},\widehat{r}_t\} < v_{i^\star,t} < \max\{b^+_{-i^\star,t},\widehat{r}_t\}\}\right]
\end{aligned}$$

since $\mathbb{I}\{\max\{v^+_{-i,t},\widehat{r}_t\} < v_{i,t}\} = 0$ if $i \neq i^\star$. Therefore

$$\mathcal{R}_t^{(2)}\mathbb{I}\{t\in E_\ell/\mathcal{B}_\ell\} \ \leq \ v_{\max}\mathbb{E}\left[\mathbb{I}\{\max\{v^+_{-i^\star,t},\widehat{r}_t\} < v_{i^\star,t} < \max\{b^+_{-i^\star,t},\widehat{r}_t\}\}\right] + \mathbb{E}\left[\left(v_t^- - b_t^-\right)^+\right], \tag{B.22}$$

To bound the first term in Equation (B.22), we again evoke the inequality $\max\{a,c\} - \max\{b,c\} = (a-b)^+$ for any $a,b,c \in \mathbb{R}$ and get $\max\{b^+_{-i^\star,t},\widehat{r}_t\} - \max\{v^+_{-i^\star,t},\widehat{r}_t\} \leq \left(b^+_{-i^\star,t} - v^+_{-i^\star,t}\right)^+$. Hence,

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{I}\{\max\{v^+_{-i^\star,t},\widehat{r}_t\} < v_{i^\star,t} < \max\{b^+_{-i^\star,t},\widehat{r}_t\}\}\right] \\
\leq\ &\mathbb{E}\left[\mathbb{I}\{\max\{b^+_{-i^\star,t},\widehat{r}_t\} - \left(b^+_{-i^\star,t} - v^+_{-i^\star,t}\right)^+ < v_{i^\star,t} < \max\{b^+_{-i^\star,t},\widehat{r}_t\}\}\right] \\
=\ &\mathbb{E}\left[\mathbb{E}\left[\mathbb{I}\{\max\{b^+_{-i^\star,t},\widehat{r}_t\} - \left(b^+_{-i^\star,t} - v^+_{-i,t}\right)^+ < v_{i^\star,t} < \max\{b^+_{-i^\star,t},\widehat{r}_t\}\} \ \middle|\ b^+_{-i^\star,t}, v^+_{-i^\star,t}\right]\right] \\
=\ &\mathbb{E}\left[\int_{\max\{b^+_{-i^\star,t},\widehat{r}_t\} - \left(b^+_{-i^\star,t} - v^+_{-i^\star,t}\right)^+ - \langle\beta,x_t\rangle}^{\max\{b^+_{-i^\star,t},\widehat{r}_t\} - \langle\beta,x_t\rangle} f(z)dz\right] \\
\leq\ &c_f\mathbb{E}\left[\left(b^+_{-i^\star,t} - v^+_{-i^\star,t}\right)^+\right].
\end{aligned} \tag{B.23}$$

Now, set $j \in [N]$ such that $b^+_{-i^\star,t} = b_{j,t}$ ($j \neq i^\star$), i.e. $j$ is the highest bidder among all buyers excluding $i^\star$. Then $b^+_{-i^\star,t} - v^+_{-i^\star,t} = b_{j,t} - v^+_{-i^\star,t} \leq b_{j,t} - v_{j,t} = -a_{j,t}$, where the inequality follows from the fact that $v^+_{-i^\star,t}$ is the highest valuation among all buyers excluding $i^\star$ (which includes $j$ as $j \neq i^\star$). Therefore, continuing the bound in Equation

(B.23), we have

$$\mathbb{E}\left[\mathbb{I}\{\max\{v^{+}_{-i^\star,t},\widehat{r}_t\} < v_{i^\star,t} < \max\{b^{+}_{-i^\star,t},\widehat{r}_t\}\}\right] \;\le\; c_f(-a_{j,t})^{+} \;\le\; c_f\sum_{i\in[N]}(-a_{i,t})^{+}\,. \tag{B.24}$$

To bound the second term in Equation (B.22), namely $\mathbb{E}\left[\left(v^{-}_t - b^{-}_t\right)^{+}\right]$, without loss of generality assume $v_{1,t} \ge v_{2,t} \ge \cdots \ge v_{N,t}$. Hence $v^{-}_t = v_{2,t}$. If $b_{2,t} \le b^{-}_t$, we have $v^{-}_t - b^{-}_t \le v_{2,t} - b_{2,t} = a_{2,t}$. Otherwise if $b_{2,t} > b^{-}_t$, then buyer 2 submitted the highest bid, so $b_{i,t} \le b^{-}_t$ for any $i \neq 2$ and thus, $v^{-}_t - b^{-}_t \le v_{1,t} - b^{-}_t \le v_{1,t} - b_{1,t} = a_{1,t}$. Hence,

$$\mathbb{E}\left[\left(v^{-}_t - b^{-}_t\right)^{+}\right] \le \max_{j\in[N]}(a_{j,t})^{+} \le \sum_{j\in[N]}(a_{j,t})^{+}\,. \tag{B.25}$$

Finally, combining Equations (B.22), (B.24), and (B.25), we have for any $t \in E_\ell$ and $\ell \ge 1$

$$\mathcal{R}^{(2)}_t\mathbb{I}\{t \in E_\ell/\mathcal{B}_\ell\} \le v_{\max}c_f\sum_{i\in[N]}(-a_{i,t})^{+} + \sum_{i\in[N]}(a_{i,t})^{+} \le (v_{\max}c_f + 1)\sum_{i\in[N]}|a_{i,t}| \tag{B.26}$$

**iv. Bounding Cumulative Regret:** We now bound the cumulative expected regret in a phase $E_{\ell+1}$ ($\ell \geq 1$) via first bounding $\sum_{t\in E_{\ell+1}} \mathcal{R}_t^{(1)}$ and $\sum_{t\in E_{\ell+1}} \mathcal{R}_t^{(2)}$ respectively.

$$
\begin{aligned}
&\sum_{t\in E_{\ell+1}} \mathcal{R}_t^{(1)} \\
\leq\ & \sum_{t\in E_{\ell+1}} \left(12 v_{\max} c_f N^2 \delta_\ell + \frac{6 v_{\max}\sqrt{N^3 \log(|E_\ell|)}}{\sqrt{2E_\ell}} + \frac{v_{\max}\left(6N^2(c_f+L_\ell)+9N+15d+9\right)}{|E_\ell|}\right) \\
=\ & |E_{\ell+1}| \left(12 v_{\max} c_f N^2 \delta_\ell + \frac{6 v_{\max}\sqrt{N^3 \log(|E_\ell|)}}{\sqrt{2E_\ell}} + \frac{v_{\max}\left(6N^2(c_f+L_\ell)+9N+15d+9\right)}{|E_\ell|}\right) \\
=\ & |E_{\ell+1}| \cdot \frac{3 v_{\max}\sqrt{2N^3 \log(|E_\ell|)}}{\sqrt{|E_\ell|}} \left(\frac{4 c_f \epsilon_{\max} x_{\max}^2 \sqrt{d}}{\lambda_0^2} + 1\right) \\
&\quad + \frac{|E_{\ell+1}|}{|E_\ell|}\left(\frac{12 v_{\max} c_f N^2 \sqrt{d}\left(N L_\ell a_{\max}+1\right) x_{\max}^2}{\lambda_0^2} + v_{\max}\left(6N^2(c_f+L_\ell)+9N+15d+9\right)\right) \\
\leq\ & c_1^1 c_f \sqrt{dTN^3\log(|E_\ell|)} + c_2^2 c_f \sqrt{d} N^3 L_\ell T^{\frac{1}{4}} \\
\leq\ & c_1 c_f \sqrt{dN^3 \log(|E_\ell|)} \left(\sqrt{T} + \frac{\sqrt{N^3\log(|E_\ell|)} T^{\frac{1}{4}}}{\log(1/\eta)}\right),
\end{aligned}
\tag{B.27}
$$

for some absolute constants $c_1^1, c_1^2, c_1 > 0$. The first inequality follows from Equation (B.19). In the second equality, we then used the definition of $\delta_\ell = \frac{\sqrt{2d\log(|E_\ell|)}\epsilon_{\max} x_{\max}^2}{\lambda_0^2\sqrt{N|E_\ell|}} + \frac{\sqrt{d}(NL_\ell a_{\max}+1)x_{\max}^2}{|E_\ell|\lambda_0^2}$, defined in Equation (B.6). In the second inequality, we relied on the construction of the length of phases in Algorithm 3, i.e. $|E_\ell| = T^{1-2^{-\ell}}$ so that $|E_{\ell+1}|/\sqrt{|E_\ell|} = \sqrt{T}$ and $|E_{\ell+1}|/|E_\ell| = T^{2^{-(\ell+1)}} \leq T^{\frac{1}{4}}$. The last inequality follows from the fact that $L_\ell = \log\left(v_{\max}^2 N |E_\ell|^4 - 1\right)/\log(\frac{1}{\eta})$.

On the other hand, to bound $\sum_{t\in E_{\ell+1}} \mathcal{R}_t^{(2)}$, we again utilize the definition of $\mathcal{B}_{i,\ell} := \mathcal{B}_{i,\ell}^s \cup \mathcal{B}_{i,\ell}^o$ and $\mathcal{B}_\ell := \cup_{i\in[N]}\mathcal{B}_{i,\ell}$ where $\mathcal{B}_{i,\ell}^s$ and $\mathcal{B}_{i,\ell}^o$ are defined in Equation (3.12) of Lemma 3.4.4. Denote $K_{\ell+1} := 2L_{\ell+1} + 4c_f + 8\log(|E_{\ell+1}|)$. Then, we have

$$\begin{aligned}
\sum_{t\in E_{\ell+1}} \mathcal{R}_t^{(2)} &= \mathbb{E}\left[\sum_{t\in\mathcal{B}_{\ell+1}} \mathcal{R}_t^{(2)}\right] + \mathbb{E}\left[\sum_{t\in E_{\ell+1}/\mathcal{B}_{\ell+1}} \mathcal{R}_t^{(2)}\right] \\
&\leq v_{\max}\mathbb{E}\left[|\mathcal{B}_{\ell+1}|\cdot\mathbb{I}\{|\mathcal{B}_{\ell+1}| \leq NK_{\ell+1}\right] + v_{\max}\mathbb{E}\left[|\mathcal{B}_{\ell+1}|\cdot\mathbb{I}\{|\mathcal{B}_{\ell+1}| > NK_{\ell+1}\}\right] \\
&\quad + (v_{\max}c_f + 1)\mathbb{E}\left[\sum_{t\in E_{\ell+1}/\mathcal{B}_{\ell+1}}\sum_{i\in[N]} |a_{i,t}|\right] \\
&\leq v_{\max}NK_{\ell+1} + v_{\max}|E_{\ell+1}|\cdot\mathbb{P}\left(|\mathcal{B}_{\ell+1}| > NK_{\ell+1}\right) \\
&\quad + (v_{\max}c_f + 1)\mathbb{E}\left[\sum_{t\in E_{\ell+1}/\mathcal{B}_{\ell+1}}\sum_{i\in[N]} |a_{i,t}|\right] \\
&\leq v_{\max}NK_{\ell+1} + 4v_{\max}N + (v_{\max}c_f + 1)\mathbb{E}\left[\sum_{t\in E_{\ell+1}/\mathcal{B}_{\ell+1}}\sum_{i\in[N]} |a_{i,t}|\right] \\
&\leq v_{\max}N(K_{\ell+1} + 4) + (v_{\max}c_f + 1)\mathbb{E}\left[\sum_{t\in E_{\ell+1}}\sum_{i\in[N]} |a_{i,t}|\right], \qquad \text{(B.28)}
\end{aligned}$$

where the first inequality follows from Equation (B.26) and uses the fact that $\mathcal{R}_t^{(2)} \leq v_{\max}$; the second inequality is because $|\mathcal{B}_{\ell+1}| \leq |E_{\ell+1}|$; the third inequality applies Lemma 3.4.4 which shows $\mathbb{P}\left(|\mathcal{B}_{i,\ell+1}| > K_{\ell+1}\right) \leq 4/|E_{\ell+1}|$, and hence $\mathbb{P}\left(|\mathcal{B}_{\ell+1}| \leq NK_{\ell+1}\right) \geq \mathbb{P}\left(\cap_{i\in[N]}\{|\mathcal{B}_{i,\ell+1}| \leq K_{\ell+1}\}\right) \geq 1 - 4N/|E_{\ell+1}|$. To bound $\mathbb{E}\left[\sum_{t\in E_{\ell+1}}\sum_{i\in[N]}|a_{i,t}|\right]$, we recall $\mathcal{S}_{\ell+1} := \cup_{i\in[N]}\mathcal{S}_{i,\ell+1}$ where $\mathcal{S}_{i,\ell+1}$ is defined in Equation (3.11), and consider the following

$$
\begin{aligned}
& \mathbb{E}\left[\sum_{t\in E_{\ell+1}}\sum_{i\in[N]}|a_{i,t}|\right] \\
\leq\ & \mathbb{E}\left[\sum_{t\in \mathcal{S}_{\ell+1}}\sum_{i\in[N]}|a_{i,t}|\right] + \mathbb{E}\left[\sum_{t\in E_{\ell+1}/\mathcal{S}_{\ell+1}}\sum_{i\in[N]}\frac{1}{|E_{\ell+1}|}\right] \\
\leq\ & Na_{\max}\mathbb{E}\left[|\mathcal{S}_{\ell+1}|\right] + N \\
=\ & Na_{\max}\mathbb{E}\left[|\mathcal{S}_{\ell+1}|\cdot(\mathbb{I}\{|\mathcal{S}_{\ell+1}|\leq NL_{\ell+1}\} + \mathbb{I}\{|\mathcal{S}_{\ell+1}|> NL_{\ell+1}\})\right] + N \\
\leq\ & Na_{\max}\left(NL_{\ell+1} + |E_{\ell+1}|\cdot\mathbb{P}\left(|\mathcal{S}_{\ell+1}|> NL_{\ell+1}\right)\right) + N \\
\leq\ & N^2 a_{\max}\left(L_{\ell+1}+1\right) + N\,,
\end{aligned} \tag{B.29}
$$

where the first inequality holds because $|a_{i,t}|\leq 1/|E_{\ell+1}|$ for all $t\in E_{\ell+1}/\mathcal{S}_{\ell+1}$ and the fourth inequality follows from Lemma 3.4.1 that shows $\mathbb{P}\left(|\mathcal{S}_{i,\ell+1}|> L_{\ell+1}\right)\leq 1/|E_{\ell+1}|$, which implies $\mathbb{P}\left(|\mathcal{S}_{\ell+1}|\leq NL_{\ell+1}\right)\geq \mathbb{P}\left(\cap_{i\in[N]}\left\{|\mathcal{S}_{i,\ell+1}|\leq L_{\ell+1}\right\}\right)\geq 1 - N/|E_{\ell+1}|$.

Hence, Equations (B.28) and (B.29) show that $\sum_{t\in E_{\ell+1}}\mathcal{R}_t^{(2)}$ is upper bounded as

$$
\begin{aligned}
\sum_{t\in E_{\ell+1}}\mathcal{R}_t^{(2)} \ &\leq\ v_{\max}N(K_\ell+4) + (v_{\max}c_f+1)\left(N^2a_{\max}\left(L_{\ell+1}+1\right)+N\right) \\
&\leq\ c_2c_fN^2\cdot\frac{\log(|E_{\ell+1}|)}{\log\left(1/\eta\right)}\,,
\end{aligned} \tag{B.30}
$$

for some absolute constant $c_2>0$. Combining this with the upper bound

$$
c_1c_f\sqrt{dN^3\log(|E_\ell|)}\left(\sqrt{T}+\frac{\sqrt{N^3\log(|E_\ell|)}T^{\frac{1}{4}}}{\log\left(1/\eta\right)}\right)
$$

shown in Equation (B.27), the expected cumulative regret in phase $E_{\ell+1}$ is

$$
\sum_{t\in E_{\ell+1}}\mathsf{Regret}_t\ \leq\ c_3c_f\sqrt{dN^3\log(T)}\left(\sqrt{T}+\frac{\sqrt{N^3\log(T)}T^{\frac{1}{4}}}{\log\left(1/\eta\right)}\right)\,,
$$

for some absolute constant $c_3>0$. Finally, since the total number of phases is upper bounded by $\lceil\log\log(T)\rceil+1$, the cumulative expected regret over the entire horizon

$T$ is

$$\begin{aligned}\mathsf{Regret}(T) &\leq v_{\max}|E_1| + \sum_{\ell=2}^{\lceil \log\log(T)\rceil} c_3 c_f \sqrt{dN^3\log(T)}\left(\sqrt{T} + \frac{\sqrt{N^3\log(T)}T^{\frac{1}{4}}}{\log(1/\eta)}\right)\\ &= \mathcal{O}\left(c_f\sqrt{dN^3\log(T)}\cdot\log(\log(T))\left(\sqrt{T}+\frac{\sqrt{N^3\log(T)}T^{\frac{1}{4}}}{\log(1/\eta)}\right)\right).\end{aligned}$$

□

### B.2.2 Proof of Lemma 3.4.1

According to the definitions of the cumulative discounted utility defined in Equation (3.1) and the NPAC-S policy in Algorithm 3, buyer $i$'s utility for submitting a bid $b \in [0, v_{\max}]$ in period $t \in [T]$ conditioning on $v_{i,t}, b^+_{-i,t}, r_t$ is given by

$$u_{i,t}(b) = \begin{cases} \left(v_{i,t} - \max\{r_t, b^+_{-i,t}\}\right)\mathbb{I}\{b > \max\{r_t, b^+_{-i,t}\}\} & \text{no isolation} \\ (v_{i,t} - r_t)\,\mathbb{I}\{b > r_t\} & i \text{ is isolated} \\ 0 & j \neq i \text{ is isolated}\end{cases}, \tag{B.31}$$

where $b^+_{-i,t}$ is the highest bid excluding that of buyer $i$, and the reserve price $r_t = \widehat{r}_t\mathbb{I}\{\text{no isolation in } t\} + r^u_t(1 - \mathbb{I}\{\text{no isolation in } t\})$ ($\widehat{r}_t$ and $r^u_t$ are defined in Equations (3.6) and (3.7) of the NPAC-S policy respectively). Note that $u_{i,t}(b)$ is a random variable that depends on the $x_t, \{\epsilon_{i,t}\}_{i\in[N]}, b^+_{-i,t}$ and $r_t$. The undiscounted utility loss $u^-_{i,t}$ for buyer $i$ if he submits a bid $b_{i,t}$ compared to bidding truthfully is $u^-_{i,t} = u_{i,t}(v_{i,t}) - u_{i,t}(b_{i,t})$.

Now, when any buyer $j \neq i$ is isolated, the utility for buyer $i$ is always 0 regardless of what he submits, so there is no utility loss due to bidding behaviour. We now consider the scenarios when no isolation occurs and when buyer $i$ is isolated, respectively, using the definition of utility in Equation (3.1).

**No isolation occurs:** The undiscounted utility loss for bidding untruthfully is

$$\begin{aligned}
& u^-_{i,t}\mathbb{I}\{\text{no isolation in } t\} \\
= \ & (u_{i,t}(v_{i,t}) - u_{i,t}(b_{i,t}))\,\mathbb{I}\{\text{no isolation in } t\} \\
= \ & \left(v_{i,t} - \max\{r_t, b^+_{-i,t}\}\right)\mathbb{I}\{v_{i,t} > \max\{r_t, b^+_{-i,t}\}\} \\
& \quad - \left(v_{i,t} - \max\{r_t, b^+_{-i,t}\}\right)\mathbb{I}\{b_{i,t} > \max\{r_t, b^+_{-i,t}\}\} \\
= \ & \left|v_{i,t} - \max\{r_t, b^+_{-i,t}\}\right|\mathbb{I}\{v_{i,t} > \max\{r_t, b^+_{-i,t}\} > b_{i,t}\} \\
& \quad + \left|v_{i,t} - \max\{r_t, b^+_{-i,t}\}\right|\mathbb{I}\{v_{i,t} < \max\{r_t, b^+_{-i,t}\} < b_{i,t}\} \\
\geq \ & 0\,.
\end{aligned} \tag{B.32}$$

**Isolating buyer $i$:** The undiscounted utility for submitting any bid $b \in \mathbb{R}$ for any given $r_t$ is $(v_{i,t} - r_t)\,\mathbb{I}\{b > r_t\}$. Hence,

$$\begin{aligned}
& u^-_{i,t}\mathbb{I}\{i \text{ is isolated}\} \\
= \ & (u_{i,t}(v_{i,t}) - u_{i,t}(b_{i,t}))\,\mathbb{I}\{i \text{ is isolated}\} \\
= \ & (v_{i,t} - r_t)\,\mathbb{I}\{v_{i,t} > r_t\} - (v_{i,t} - r_t)\,\mathbb{I}\{b_{i,t} > r_t\} \\
= \ & (v_{i,t} - r_t)\,\mathbb{I}\{v_{i,t} > r_t > b_{i,t}\} + (-v_{i,t} + r_t)\,\mathbb{I}\{v_{i,t} < r_t < b_{i,t}\}\,.
\end{aligned} \tag{B.33}$$

The NPAC-S policy offers a price $r_t$ drawn from Uniform$(0, v_{\max})$ to the isolated buyer $i$ with probability $1/|E_\ell|$, where $i$ is chosen uniformly among all buyers. So, the expected utility loss $u^-_{i,t}$ for a buyer $i \in [N]$ conditioned on the fact that the buyer lies by an amount of $a_{i,t}$ is

$$
\begin{aligned}
&\mathbb{E}[u^-_{i,t} \mid a_{i,t}] \\
=\;& \mathbb{E}[u^-_{i,t}\mathbb{I}\{i \text{ is isolated}\} + u^-_{i,t}\mathbb{I}\{\text{no isolation in } t\} \mid a_{i,t}] \\
\geq\;& \mathbb{E}[u^-_{i,t}\mathbb{I}\{i \text{ is isolated}\} \mid a_{i,t}] \\
=\;& \frac{1}{N|E_\ell|}\mathbb{E}\left[(v_{i,t} - r_t)\,\mathbb{I}\{v_{i,t} > r_t > b_t\} + (-v_{i,t} + r_t)\,\mathbb{I}\{b_t < r_t < v_{i,t}\} \mid a_{i,t}\right] \\
=\;& \frac{1}{v_{\max}N|E_\ell|}\mathbb{E}\left[\mathbb{E}\left[\int_{v_{i,t}-a_{i,t}}^{v_{i,t}} (v_{i,t} - r)dr + \int_{v_{i,t}}^{v_{i,t}+a_{i,t}} (-v_{i,t} + r)dr \;\middle|\; a_{i,t}, v_{i,t}\right] \;\middle|\; a_{i,t}\right] \\
=\;& \frac{(a_{i,t})^2}{v_{\max}N|E_\ell|}\,. 
\end{aligned}
\tag{B.34}
$$

The first inequality follows from $u^-_{i,t}\mathbb{I}\{i \text{ is isolated}\} \geq 0$ as demonstrated in Equation (B.32). Now we lower bound the total expected utility loss in phase $E_\ell$. First, by Equations (B.32) and (B.33), we know that $u^-_{i,t} \geq 0$ for $\forall i, t$. Therefore, denoting $s_{\ell+1}$ as the first period of phase $E_{\ell+1}$, for any $\tilde{z} > 0$ we have

$$
\begin{aligned}
\mathbb{E}\left[\sum_{t\in E_\ell}\eta^t u^-_{i,t}\right] \;&\geq\; \mathbb{E}\left[\sum_{t\in\mathcal{S}_{i,\ell}}\eta^t u^-_{i,t}\right] \\
&\geq\; \mathbb{E}\left[\sum_{t\in\mathcal{S}_{i,\ell}}\eta^t u^-_{i,t}\mathbb{I}\{|\mathcal{S}_{i,\ell}| \geq \tilde{z}\}\right] \\
&=\; \mathbb{E}\left[\mathbb{E}\left[\sum_{t\in\mathcal{S}_{i,\ell}}\eta^t u^-_{i,t} \;\middle|\; \{a_{i,t}\}_{t\in E_\ell}\right]\mathbb{I}\{|\mathcal{S}_{i,\ell}| \geq \tilde{z}\}\right] \\
&\geq\; \mathbb{E}\left[\sum_{t\in\mathcal{S}_{i,\ell}}\frac{\eta^t}{v_{\max}N|E_\ell|^3}\cdot\mathbb{I}\{|\mathcal{S}_{i,\ell}| \geq \tilde{z}\}\right] \\
&\geq\; \mathbb{E}\left[\sum_{t=s_{\ell+1}-|\mathcal{S}_{i,\ell}|}^{s_{\ell+1}-1}\frac{\eta^t}{v_{\max}N|E_\ell|^3}\cdot\mathbb{I}\{|\mathcal{S}_{i,\ell}| \geq \tilde{z}\}\right] \\
&\geq\; \mathbb{E}\left[\sum_{t=s_{\ell+1}-\tilde{z}}^{s_{\ell+1}-1}\frac{\eta^t}{v_{\max}N|E_\ell|^3}\cdot\mathbb{I}\{|\mathcal{S}_{i,\ell}| \geq \tilde{z}\}\right] \\
&=\; \frac{\eta^{s_{\ell+1}}\left(1-\eta^{-\tilde{z}}\right)}{(1-\eta)v_{\max}N|E_\ell|^3}\mathbb{P}\left(|\mathcal{S}_{i,\ell}| \geq \tilde{z}\right)\,, 
\end{aligned}
\tag{B.35}
$$

where the first equality holds because $|\mathcal{S}_{i,\ell}| = \sum_{t\in E_\ell} \mathbb{I}\{a_{i,t} > 1/E_\ell\}$ is a function of $\{a_{i,t}\}_{t\in E_\ell}$; the third inequality follows from Equation (B.34) and $a_{i,t} \geq 1/|E_\ell|$ for any $t \in \mathcal{S}_{i,\ell}$; and the fourth inequality is because $\eta \in (0, 1)$.

Furthermore, corrupting a bid at time $t \in E_\ell$ will only impact the prices offered by the seller in future phases, i.e., phase $\ell + 1, \ell + 2, \ldots$, so the utility gain due to lying in phase $\ell$, denoted as $U^+_{i,\ell}$ is upper bounded by $v_{\max} \sum_{t\geq s_{\ell+1}} \eta^t = v_{\max}\eta^{s_{\ell+1}}/(1 - \eta)$. Since the buyer is utility maximizing, the net utility gain due to lying in phase $\ell$ should be greater than 0, otherwise the buyer can choose to always bid 0 in phase $\ell$ which is equivalent to not participating in the auctions. Hence,

$$\mathbb{E}\left[U^+_{i,\ell} - \sum_{t\in E_\ell} \eta^t u^-_{i,t}\right] \geq 0\,.$$

Combining this with $U^+_{i,\ell} \leq v_{\max}\eta^{s_{\ell+1}}/(1 - \eta)$ and the lower bound for $\mathbb{E}\left[\sum_{t\in E_\ell} u^-_{i,t}\right]$ shown in Equation (B.35), we have

$$\frac{v_{\max}\eta^{s_{\ell+1}}}{1-\eta} \geq \frac{\eta^{s_{\ell+1}}\left(1 - \eta^{-\tilde{z}}\right)}{(1-\eta)v_{\max}N|E_\ell|^3}\mathbb{P}\left(|\mathcal{S}_{i,\ell}| \geq \tilde{z}\right)\,,$$

which holds for any $\tilde{z} > 0$. Taking $\tilde{z} = \log\left(v^2_{\max}N|E_\ell|^4 - 1\right)/\log(1/\eta)$ and by rearranging terms, the inequality above yields

$$\mathbb{P}\left(|\mathcal{S}_{i,\ell}| \geq \frac{\log\left(v^2_{\max}N|E_\ell|^4 - 1\right)}{\log(\frac{1}{\eta})}\right) \leq \frac{1}{|E_\ell|}\,.$$

□

### B.2.3 Proof of Lemma 3.4.4

Defining $\mathcal{H}_{i,t} := \{(b^+_{-i,\tau}, \widehat{r}_\tau, x_\tau)\}_{\tau\in[t]}$, we have

$$\begin{aligned}
& \mathbb{E}\left[\mathbb{I}\{t \in (E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}^s_{i,\ell}\} \mid \mathcal{H}_{i,t}\right] \\
= \; & \mathbb{P}\left(t \in (E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}^s_{i,\ell} \mid \mathcal{H}_{i,t}\right) \\
= \; & \mathbb{P}\left(v_{i,t} \geq \max\{b^+_{-i,t}, \widehat{r}_t\} \ , \ b_{i,t} < \max\{b^+_{-i,t}, \widehat{r}_t\} \ , \ a_{i,t} \in (0, 1/|E_\ell|) \mid \mathcal{H}_{i,t}\right) \\
= \; & \mathbb{P}\Big( \max\{b^+_{-i,t}, \widehat{r}_t\} - \langle x_t, \beta\rangle \leq \epsilon_{i,t} \leq \max\{b^+_{-i,t}, \widehat{r}_t\} - \langle x_t, \beta\rangle + a_{i,t} \ , \\
& \quad a_{i,t} \in (0, 1/|E_\ell|) \mid \mathcal{H}_{i,t}\Big) \\
\leq \; & \mathbb{P}\left(\max\{b^+_{-i,t}, \widehat{r}_t\} - \langle x_t, \beta\rangle \leq \epsilon_{i,t} \leq \max\{b^+_{-i,t}, \widehat{r}_t\} - \langle x_t, \beta\rangle + 1/|E_\ell| \mid \mathcal{H}_{i,t}\right) \\
= \; & \mathbb{E}\left[\int_{\max\{b^+_{-i,t}, \widehat{r}_t\} - \langle x_t, \beta\rangle}^{\max\{b^+_{-i,t}, \widehat{r}_t\} - \langle x_t, \beta\rangle + 1/|E_\ell|} f(z)dz \ \middle| \ \mathcal{H}_{i,t}\right] \\
\leq \; & \frac{c_f}{|E_\ell|} \, .
\end{aligned} \tag{B.36}$$

The last inequality uses the fact that $c_f = \sup_{\tilde{z}\in[-\epsilon_{\max}, \epsilon_{\max}]} f(\tilde{z})$.

Define $\zeta_t = \mathbb{I}\{t \in (E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}^s_{i,\ell}\}$ and $\phi_t = \mathbb{E}\left[\mathbb{I}\{t \in (E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}^s_{i,\ell}\} \mid \mathcal{H}_{i,t}\right]$. Then $\mathbb{E}[\zeta_t \mid \mathcal{H}_{i,t}] = \phi_t$, which implies

$$\mathbb{E}[\zeta_t - \phi_t \mid \sum_{\tau<t} \zeta_\tau, \sum_{\tau<t} \phi_\tau] = \mathbb{E}\left[\mathbb{E}\left[\zeta_t - \phi_t \mid \mathcal{H}_{i,t}\right] \mid \sum_{\tau<t} \zeta_\tau, \sum_{\tau<t} \phi_\tau\right] = 0.$$

Hence, in the context of the multiplicative Azuma inequality described in Lemma B.3.3, by setting $z_{1,t} = \zeta_t$, $z_{2,t} = \phi_t$, $\tilde{\gamma} = 1/2$ and $A = 2\log(|E_\ell|)$ we have $|z_{1,t} - z_{2,t}| \leq 1$

$$\mathbb{P}\left(\frac{1}{2}\sum_{t\in E_\ell} \zeta_t \geq \sum_{t\in E_\ell} \phi_t + 2\log(|E_\ell|)\right) \leq \exp\left(-\log(|E_\ell|)\right) \, . \tag{B.37}$$

Now, according to Equation (B.36), we have $\phi_t \leq c_f/|E_\ell|$, so $\sum_{t\in E_\ell} \phi_t \leq c_f$. Moreover,

$|(E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}^s_{i,\ell}| = \sum_{t\in E_\ell} \zeta_t$. Hence, following Equation (B.37), we have

$$\begin{aligned}
& \mathbb{P}\left(|(E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}^s_{i,\ell}| \geq 2c_f + 4\log(|E_\ell|)\right) \\
\leq\ & \mathbb{P}\left(\frac{1}{2}\sum_{t\in E_\ell} \zeta_t \geq \sum_{t\in E_\ell} \phi_t + 2\log(|E_\ell|)\right) \\
\leq\ & \exp\left(-\log(|E_\ell|)\right) = \frac{1}{|E_\ell|}\,. \qquad \text{(B.38)}
\end{aligned}$$

When the event $\mathcal{G}_{i,t} = \{|\mathcal{S}_{i,\ell}| \leq L_\ell\}$ occurs, where $L_\ell = \log\left(v^2_{\max} N|E_\ell|^4 - 1\right)/\log(1/\eta)$, we have $|\mathcal{B}^s_{i,\ell}| \leq |\mathcal{S}_{i,\ell}| + |(E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}^s_{i,\ell}| \leq L_\ell + |(E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}^s_{i,\ell}|$. Therefore when event $\mathcal{G}_{i,t}$ occurs,

$$\begin{aligned}
& \mathbb{P}\left(|\mathcal{B}^s_{i,\ell}| \leq L_\ell + 2c_f + 4\log(|E_\ell|)\right) \\
\geq\ & \mathbb{P}\left(\left\{|\mathcal{B}^s_{i,\ell}| \leq L_\ell + 2c_f + 4\log(|E_\ell|)\right\} \bigcap \mathcal{G}_{i,t}\right) \\
\geq\ & \mathbb{P}\left(\left\{|(E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}^s_{i,\ell}| \leq 2c_f + 4\log(|E_\ell|)\right\} \bigcap \mathcal{G}_{i,t}\right) \\
\geq\ & 1 - \mathbb{P}\left(|(E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}^s_{i,\ell}| \geq 2c_f + 4\log(|E_\ell|)\right) - \mathbb{P}\left(\mathcal{G}^c_{i,t}\right) \\
\geq\ & 1 - \frac{2}{|E_\ell|}\,.
\end{aligned}$$

The second inequality follows from $|\mathcal{B}^s_{i,\ell}| \leq L_\ell + |(E_\ell/\mathcal{S}_{i,\ell}) \cap \mathcal{B}^s_{i,\ell}|$ when the event $\mathcal{G}_{i,t}$ occurs; the third inequality applies the union bound, and the final inequality follows from Equation (B.38) and Lemma 3.4.1.

Similarly, we can show the same probability upper bound for $|\mathcal{B}^o_{i,\ell}|$. Finally, using the fact that $\mathcal{B}_{i,\ell} = \mathcal{B}^s_{i,\ell} \cup \mathcal{B}^o_{i,\ell}$ and applying a union bound would yield the desired expression. $\square$

### B.2.4 Other Lemmas for proving Theorem 3.4.2

**Lemma B.2.1** (Bounding Estimation Errors in $\beta$)**.** *For any phase $E_\ell$ and $\gamma > 0$, we have*

$$\mathbb{P}\left(\|\widehat{\beta}_{\ell+1} - \beta\|_1 \leq \gamma + \frac{d\left(NL_\ell a_{\max} + 1\right)x_{\max}}{|E_\ell|\lambda_0^2}\right)$$
$$\geq\ 1 - 2d\exp\left(-\frac{N\gamma^2\lambda_0^4|E_\ell|}{2{\epsilon_{\max}}^2 x_{\max}^2 d}\right) - d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right) - \frac{N}{|E_\ell|}\,,$$

*where $\lambda_0^2$ is the minimum eigenvalue of the covariance matrix $\Sigma$, $\widehat{\beta}_{\ell+1}$ is defined in Equation (3.8), and $L_\ell = \log\left(v_{\max}^2 N|E_\ell|^4 - 1\right)/\log(1/\eta)$. Furthermore, setting $\gamma = \frac{\sqrt{2d\log(|E_\ell|)}\epsilon_{\max}x_{\max}}{\lambda_0^2\sqrt{N|E_\ell|}}$ and denoting $\delta_\ell = \gamma \cdot x_{\max} + \frac{d(NL_\ell a_{\max}+1)x_{\max}^2}{|E_\ell|\lambda_0^2}$, we have*

$$\mathbb{P}\left(\|\widehat{\beta}_{\ell+1} - \beta\|_1 \leq \frac{\delta_\ell}{x_{\max}}\right) \geq 1 - \frac{2d+N}{|E_\ell|} - d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right).$$

*Proof.* Proof of Lemma B.2.1.

The proof of Lemma B.2.1 is inspired by Lemma EC.7.2 in [19], but here we made substantial modifications to resolve the issues that arise when estimating $\beta$ in the presence of corrupted bids submitted by buyers.

First, recall that the smallest eigenvalue $\lambda_0^2$ of the covariance matrix $\Sigma$ of $x \sim \mathcal{D}$ is greater than 0. Since the second moment matrix $\mathbb{E}[x_t x_t^\top] = \Sigma + \mathbb{E}[x]\mathbb{E}[x]^\top$, we know that the smallest eigenvalue of $\mathbb{E}[x_t x_t^\top]$ is at least $\lambda_0^2 > 0$. We denote the design matrix of all the features in phase $E_\ell$ as $X \in \mathbb{R}^{|E_\ell|\times d}$, and $\bar{\epsilon}_\tau = \frac{\sum_{i\in[N]}\epsilon_{i,\tau}}{N}$ for $\forall \tau \in E_\ell$.

We first consider the case where the smallest eigenvalue of the second moment matrix $\lambda_{\min}\left(X^\top X/|E_\ell|\right) \geq \lambda_0^2/2$, which implies that $(X^\top X)^{-1}$ exists and $(X^\top X)^{-1} = (X^\top X)^\dagger$. By the definition $b_{i,t} = v_{i,t} - a_{i,t}$, and the definition of $\bar{b}_\tau$ for any $\tau \in [T]$ in

Equation (3.8) we have

$$\widehat{\beta}_{\ell+1} = \left(X^\top X\right)^{-1} X^\top \begin{pmatrix} \bar{b}_1 \\ \vdots \\ \bar{b}_t \end{pmatrix} = \left(X^\top X\right)^{-1} X^\top \begin{pmatrix} \frac{\sum_{i\in[N]} v_{i,1} - a_{i,1}}{N} \\ \vdots \\ \frac{\sum_{i\in[N]} v_{i,t} - a_{i,t}}{N} \end{pmatrix}$$

$$= \beta + \left(X^\top X\right)^{-1} X^\top \begin{pmatrix} \frac{\sum_{i\in[N]} \epsilon_{i,1} - a_{i,1}}{N} \\ \vdots \\ \frac{\sum_{i\in[N]} \epsilon_{i,t} - a_{i,t}}{N} \end{pmatrix}$$

$$= \beta + \left(X^\top X\right)^{-1} X^\top \left(\bar{\mathcal{E}} - A\right) , \tag{B.39}$$

where $\bar{\mathcal{E}}$ is the column vector consisting of all $\bar{\epsilon}_\tau := \frac{\sum_{i\in[N]} \epsilon_{i,\tau}}{N}$, and $A$ is the column vector consisting of all $\bar{a}_\tau := \frac{\sum_{i\in[N]} a_{i,\tau}}{N}$ for $\forall \tau \in [t]$. Therefore,

$$\begin{aligned} \|\widehat{\beta}_{\ell+1} - \beta\|_2 &= \|\left(X^\top X\right)^{-1} X^\top \left(\bar{\mathcal{E}} - A\right)\|_2 \\ &\leq \frac{1}{|E_\ell|\lambda_0^2} \left(\|X^\top \bar{\mathcal{E}}\|_2 + \|X^\top A\|_2\right) . \end{aligned} \tag{B.40}$$

Denote $X^j$ as the $j$th column of $X$, i.e. the $j$th row of $X^\top$ for $j = 1, 2 \ldots d$, we now bound $\|X^\top \bar{\mathcal{E}}\|_2$ and $\|X^\top A\|_2$ separately. First, since $\|X^\top \bar{\mathcal{E}}\|_2^2 = \sum_{j\in[d]} \left|\bar{\mathcal{E}}^\top X^j\right|^2$, for any $\gamma > 0$,

$$\bigcap_{j\in[d]} \left\{ \left|\bar{\mathcal{E}}^\top X^j\right| \leq \frac{|E_\ell|\lambda_0^2\gamma}{\sqrt{d}} \right\} \subseteq \left\{ \frac{1}{|E_\ell|\lambda_0^2} \cdot \|X^\top \bar{\mathcal{E}}\|_2 \leq \gamma \right\} . \tag{B.41}$$

We observe that $\bar{\mathcal{E}}^\top X^j = \frac{\sum_{\tau\in E_\ell} \sum_{i\in[N]} \epsilon_{i,\tau} X_{\tau j}}{N}$, where all $\epsilon_{i,\tau} X_{\tau j}$ are 0-mean and $\epsilon_{\max} x_{\max}$-subgaussion random variables. Therefore by Hoeffding's inequality, for any $\tilde{\gamma} > 0$

$$\mathbb{P}\left(\left|N\bar{\mathcal{E}}^\top X^j\right| \leq \tilde{\gamma}\right) \geq 1 - 2\exp\left(-\frac{\tilde{\gamma}^2}{2{\epsilon_{\max}}^2 x_{\max}^2 |E_\ell| N}\right) . \tag{B.42}$$

Replacing $\tilde{\gamma}$ with $N|E_\ell|\lambda_0^2\gamma/\sqrt{d}$ and using Equation (B.41) yields:

$$\begin{aligned}
\mathbb{P}\left(\left\{\frac{1}{|E_\ell|\lambda_0^2}\cdot\|X^\top\bar{\mathcal{E}}\|_2\le\gamma\right\}\right) \;&\ge\; \mathbb{P}\left(\bigcap_{j\in[d]}\left\{\left|\bar{\mathcal{E}}^\top X^j\right|\le\frac{|E_\ell|\lambda_0^2\gamma}{\sqrt{d}}\right\}\right)\\
&\ge\; 1-\sum_{j\in[d]}\mathbb{P}\left(\left|\bar{\mathcal{E}}^\top X^j\right|>\frac{|E_\ell|\lambda_0^2\gamma}{\sqrt{d}}\right)\\
&\ge\; 1-2d\exp\left(-\frac{N\gamma^2\lambda_0^4|E_\ell|}{2{\epsilon_{\max}}^2x_{\max}^2 d}\right)\,, \qquad \text{(B.43)}
\end{aligned}$$

where the first inequality follows from Equation (B.41), the second inequality applies the union bound, and the last inequality follows from Equation (B.42).

In the following, we show a high probability bound for $\|X^\top A\|_2^2$ by using the fact that $|a_{i,t}|\le 1/|E_\ell|$ for any $t\in E_\ell/\mathcal{S}_{i,\ell}$, where $\mathcal{S}_{i,\ell}=\{t\in E_\ell:|a_{i,t}|>1/|E_\ell|\}$, and $\mathcal{S}_{i,\ell}\le L_\ell$ with high probability.

Recall the event $\mathcal{G}_{i,\ell}=\{|\mathcal{S}_{i,\ell}|\le L_\ell\}$, and in Lemma 3.4.1 we showed that $\mathbb{P}\left(\mathcal{G}_{i,\ell}^c\right)=\mathbb{P}\left(|\mathcal{S}_{i,\ell}|>L_\ell\right)\le\frac{1}{|E_\ell|}$. We now bound $\|X^\top A\|_2$ under the occurrence of $\mathcal{G}_{i,\ell}$ for all $i$.

$$\begin{aligned}
\|X^\top A\|_2^2=\sum_{j\in[d]}\left|A^\top X^j\right|^2 \;&=\; \sum_{j\in[d]}\left(\frac{\sum_{\tau\in E_\ell}\sum_{i\in[N]}a_{i,\tau}X_{\tau j}}{N}\right)^2\\
&\le\; \sum_{j\in[d]}\left(\frac{\sum_{\tau\in E_\ell}\sum_{i\in[N]}|a_{i,\tau}|x_{\max}}{N}\right)^2\,. \qquad \text{(B.44)}
\end{aligned}$$

For periods in $S_\ell:=\cup_{i\in[N]}\mathcal{S}_{i,\ell}$, we have,

$$\frac{\sum_{\tau\in S_\ell}\sum_{i\in[N]}|a_{i,\tau}|x_{\max}}{N}\;\le\;\sum_{\tau\in S_\ell}a_{\max}x_{\max}\;\le\;NL_\ell a_{\max}x_{\max}\,, \qquad \text{(B.45)}$$

where the last inequality holds because events $\mathcal{G}_{i,\ell}$ occurs for all $i$. On the other hand, recall that $|a_{i,t}|\ge 1/|E_\ell|$ for any $i$ and $t\in\mathcal{S}_{i,\ell}$. Hence, $|a_{i,t}|\le 1/|E_\ell|$ for periods in $E_\ell/\mathcal{S}_\ell$,

$$\frac{\sum_{\tau\in E_\ell/\mathcal{S}_\ell}\sum_{i\in[N]}|a_{i,\tau}|x_{\max}}{N}\;\le\;\sum_{\tau\in E_\ell/\mathcal{S}_\ell}\frac{x_{\max}}{|E_\ell|}\;\le\;\sum_{\tau\in E_\ell}\frac{x_{\max}}{|E_\ell|}=x_{\max}\,. \qquad \text{(B.46)}$$

Combining Equations (B.44), (B.45), and (B.46), we have

$$\|X^\top A\|_2 \;\le\; \sqrt{d\left(\frac{\sum_{\tau\in[t]}\sum_{i\in[N]}|a_{i,\tau}|x_{\max}}{N}\right)^2} \;\le\; \sqrt{d}\,(NL_\ell a_{\max}+1)\,x_{\max}\,. \quad \text{(B.47)}$$

Now it only remains to show $\lambda_{\min}\left(X^\top X/|E_\ell|\right) \ge \lambda_0^2/2$ with high probability, which can be achieved by applying Lemma B.3.2. In the context of this lemma, we consider the sequence of random matrices $\{x_\tau x_\tau^\top/|E_\ell|\}_{\tau\in[E_\ell]}$, and note that $X^\top X/|E_\ell| = \sum_{\tau\in E_\ell}(x_\tau x_\tau^\top/|E_\ell|)$. We first upper bound the maximum eigenvalue of $x_\tau x_\tau^\top/|E_\ell|$, namely $\lambda_{\max}\left(x_\tau x_\tau^\top/|E_\ell|\right)$ for any $\tau\in E_\ell$ by

$$\lambda_{\max}\left(\frac{x_\tau x_\tau^\top}{|E_\ell|}\right) = \max_{\|z\|_2=1} z^\top \frac{x_\tau x_\tau^\top}{|E_\ell|} z \le \frac{1}{|E_\ell|}\max_{\|z\|_2=1}(x^\top z)^2 \le \frac{x_{\max}^2}{|E_\ell|}\,.$$

This allows us to apply the matrix Chernoff bound in Lemma B.3.2 (setting $\bar\gamma = 1/2$ in the lemma) and get

$$\begin{aligned}\mathbb{P}\left(\lambda_{\min}\left(\frac{X^\top X}{|E_\ell|}\right) \ge \frac{\lambda_0^2}{2}\right) \;&\ge\; \mathbb{P}\left(\lambda_{\min}\left(\frac{X^\top X}{|E_\ell|}\right) \ge\; \frac{1}{2}\lambda_{\min}\left(\mathbb{E}\left[\frac{X^\top X}{|E_\ell|}\right]\right)\right) \\ &\ge\; 1 - d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right)\,, \quad \text{(B.48)}\end{aligned}$$

where the first inequality follows from the fact that $\lambda_{\min}\left(\mathbb{E}[X^\top X/|E_\ell|]\right) \ge \lambda_0^2$.

Putting everything together, we get

$$
\begin{aligned}
& \mathbb{P}\left(\|\widehat{\beta}_{\ell+1}-\beta\|_1 \leq \gamma+\frac{\sqrt{d}\left(N L_\ell a_{\max}+1\right) x_{\max}}{|E_\ell|\lambda_0^2}\right) \\
\geq\ & \mathbb{P}\left(\|\widehat{\beta}_{\ell+1}-\beta\|_2 \leq \gamma+\frac{\sqrt{d}\left(N L_\ell a_{\max}+1\right) x_{\max}}{|E_\ell|\lambda_0^2}\right) \\
\geq\ & \mathbb{P}\Big(\left\{\frac{1}{|E_\ell|\lambda_0^2}\left(\|X^\top \bar{\mathcal{E}}\|_2+\|X^\top A\|_2\right) \leq \gamma+\frac{\sqrt{d}\left(N L_\ell a_{\max}+1\right) x_{\max}}{|E_\ell|\lambda_0^2}\right\} \\
& \qquad \bigcap\left\{\lambda_{\min}\left(\frac{X^\top X}{|E_\ell|}\right) \geq \frac{\lambda_0^2}{2}\right\}\Big) \\
\geq\ & \mathbb{P}\left(\left\{\frac{1}{|E_\ell|\lambda_0^2}\|X^\top \bar{\mathcal{E}}\|_2 \leq \gamma\right\} \bigcap \left(\bigcap_{i\in[N]} \mathcal{G}_{i,\ell}\right) \bigcap \left\{\lambda_{\min}\left(\frac{X^\top X}{|E_\ell|}\right) \geq \frac{\lambda_0^2}{2}\right\}\right) \\
\geq\ & 1-\mathbb{P}\left(\left\{\frac{1}{|E_\ell|\lambda_0^2}\|X^\top \bar{\mathcal{E}}\|_2 > \gamma\right\}\right)-\sum_{i\in[N]}\mathbb{P}\left(\mathcal{G}_{i,\ell}^c\right)-\mathbb{P}\left(\left\{\lambda_{\min}\left(\frac{X^\top X}{|E_\ell|}\right) \leq \frac{\lambda_0^2}{2}\right\}\right) \\
\geq\ & 1-2d\exp\left(-\frac{N\gamma^2\lambda_0^4|E_\ell|}{2{\epsilon_{\max}}^2 x_{\max}^2 d}\right)-\frac{N}{|E_\ell|}-d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right).
\end{aligned}
$$

The first inequality follows from the fact that $\|z\|_1 \leq \|z\|_2$ for any vector $z$; the second inequality follows from Equation (B.40); the third inequality follows from Equation (B.47) when the event $\cap_{i\in[N]}\mathcal{G}_{i,\ell}$ occurs; the fourth inequality applies a simple union bound; and the final inequality follows from Equations (B.43), (B.48) and Lemma 3.4.1. □

**Lemma B.2.2** (Bounding Estimation Error in $F^-$ and $F^+$)**.** *Define $\tilde{\sigma}_t$ to be the sigma algebra generated by all $\{x_\tau, a_{i,\tau}, \epsilon_{i,\tau}\}_{i\in[N],\tau\in[t]}$. Then, for any $\tilde{\sigma}_t$-measurable random variable $z$ and $\gamma > 0$, we have*

$$
\begin{aligned}
& \mathbb{P}\left(\left|\widehat{F}^-_{\ell+1}(z)-F^-(z)\right| \leq 2N^2 z_\ell\right) \\
\geq\ & 1-4\exp\left(-2N|E_\ell|\gamma^2\right)-\frac{4(d+N)}{|E_\ell|}-2d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right) \\
& \mathbb{P}\left(\left|\widehat{F}^+_{\ell+1}(z)-F^+(z)\right| \leq N z_\ell\right) \\
\geq\ & 1-4\exp\left(-2N|E_\ell|\gamma^2\right)-\frac{4(d+N)}{|E_\ell|}-2d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right),
\end{aligned}
$$

*where* $z_\ell := \gamma + c_f\delta_\ell + (c_f + L_\ell)/|E_\ell|$, $c_f = \sup_{\tilde{z}\in[-\epsilon_{\max},\epsilon_{\max}]} f(\tilde{z})$, $\delta_\ell$ *is defined in Equation (B.6), and* $L_\ell = \log(v_{\max}^2 N|E_\ell|^4 - 1)/\log(1/\eta)$.

*Proof.* Proof of Lemma B.2.2. We first bound the error in the estimate of $F$, namely $\left|\widehat{F}_{\ell+1}(z) - F(z)\right|$. Then, we use the relationship $F^-(z) = NF^{N-1}(z) - (N-1)F^N(z)$ and $F^+(z) = F^N(z)$, as well as the definition of $\widehat{F}^-_{\ell+1}(z)$ and $\widehat{F}^+_{\ell+1}(z)$ in Equation (3.9) to show the desired probability bounds.

We first upper and lower bound $\widehat{F}^-_{\ell+1}(z)$ for any $z \in \mathbb{R}$. Recall the event $\mathcal{S}_{i,\ell} = \{t \in E_\ell : |a_{i,t}| \geq 1/|E_\ell|\}$ and in Lemma 3.4.1 we showed that $\mathbb{P}(|\mathcal{S}_{i,\ell}| > L_\ell) \leq 1/|E_\ell|$. Hence, for any $i \in [N]$, we have $|a_{i,t}| \leq 1/|E_\ell|$ for all periods $\tau \in E_\ell/\mathcal{S}_{i,\ell}$, so

$$\begin{aligned}
&\sum_{\tau\in E_\ell} \mathbb{I}\left\{b_{i,\tau} - \langle\widehat{\beta}_{\ell+1}, x_\tau\rangle \leq z\right\} \\
&= \left(\sum_{\tau\in E_\ell/\mathcal{S}_{i,\ell}} \mathbb{I}\left\{b_{i,\tau} - \langle\widehat{\beta}_{\ell+1}, x_\tau\rangle \leq z\right\} + \sum_{\tau\in\mathcal{S}_{i,\ell}} \mathbb{I}\left\{v_{i,\tau} - \langle\widehat{\beta}_{\ell+1}, x_\tau\rangle \leq z\right\}\right) \\
&\quad + \left(\sum_{\tau\in\mathcal{S}_{i,\ell}} \mathbb{I}\left\{b_{i,\tau} - \langle\widehat{\beta}_{\ell+1}, x_\tau\rangle \leq z\right\} - \sum_{\tau\in\mathcal{S}_{i,\ell}} \mathbb{I}\left\{v_{i,\tau} - \langle\widehat{\beta}_{\ell+1}, x_\tau\rangle \leq z\right\}\right) . \qquad \text{(B.49)}
\end{aligned}$$

Consider the sum in first the parenthesis of Equation (B.49) and note that $b_{i,\tau} = v_{i,\tau} - a_{i,\tau} = \langle\beta, x_\tau\rangle + \epsilon_{i,\tau} - a_{i,\tau}$. Since $|a_{i,\tau}| \leq 1/|E_\ell|$ for any $i \in [N]$ and $\tau \in E_\ell/\mathcal{S}_{i,\ell}$,

$$\langle\beta, x_\tau\rangle + \epsilon_{i,\tau} - \frac{1}{|E_\ell|} \;\leq\; b_{i,\tau} \;\leq\; \langle\beta, x_\tau\rangle + \epsilon_{i,\tau} + \frac{1}{|E_\ell|}, \quad \forall\tau \in E_\ell/\mathcal{S}_{i,\ell}\,. \qquad \text{(B.50)}$$

Now, assume that the event $\xi_{\ell+1} = \left\{\|\widehat{\beta}_{\ell+1} - \beta\|_1 \leq \delta_\ell/x_{\max}\right\}$ holds. Therefore, we can upper bound the sum in first the parenthesis of Equation (B.49) as

$$
\begin{aligned}
&\sum_{\tau \in E_\ell/\mathcal{S}_{i,\ell}} \mathbb{I}\left\{b_{i,\tau} - \langle \widehat{\beta}_{\ell+1}, x_\tau\rangle \le z\right\} + \sum_{\tau \in \mathcal{S}_{i,\ell}} \mathbb{I}\left\{v_{i,\tau} - \langle \widehat{\beta}_{\ell+1}, x_\tau\rangle \le z\right\} \\
\le\ & \sum_{\tau \in E_\ell/\mathcal{S}_{i,\ell}} \mathbb{I}\left\{\epsilon_{i,\tau} \le z + \langle \widehat{\beta}_{\ell+1} - \beta, x_\tau\rangle + \frac{1}{|E_\ell|}\right\} \\
&\quad + \sum_{\tau \in \mathcal{S}_{i,\ell}} \mathbb{I}\left\{\epsilon_{i,\tau} \le z + \langle \widehat{\beta}_{\ell+1} - \beta, x_\tau\rangle + \frac{1}{|E_\ell|}\right\} \\
=\ & \sum_{\tau \in E_\ell} \mathbb{I}\left\{\epsilon_{i,\tau} \le z + \langle \widehat{\beta}_{\ell+1} - \beta, x_\tau\rangle + \frac{1}{|E_\ell|}\right\} \\
\le\ & \sum_{\tau \in E_\ell} \mathbb{I}\left\{\epsilon_{i,\tau} \le z + \delta_\ell + \frac{1}{|E_\ell|}\right\}\,,
\end{aligned}
\tag{B.51}
$$

where the first equality follows from $v_{i,\tau} = \langle \beta, x_\tau\rangle + \epsilon_{i,\tau}$ and $b_{i,\tau} = v_{i,\tau} - a_{i,\tau}$; the first inequality follows Equation (B.50); and the final inequality is due to the occurrence of the event $\xi_{\ell+1} = \left\{\|\widehat{\beta}_{\ell+1} - \beta\|_1 \le \delta_\ell / x_{\max}\right\}$. Similarly, we can also lower bound the sum in the first parenthesis of Equation (B.49):

$$
\begin{aligned}
&\sum_{\tau \in E_\ell/\mathcal{S}_{i,\ell}} \mathbb{I}\left\{b_{i,\tau} - \langle \widehat{\beta}_{\ell+1}, x_\tau\rangle \le z\right\} + \sum_{\tau \in \mathcal{S}_{i,\ell}} \mathbb{I}\left\{b_{i,\tau} - \langle \widehat{\beta}_{\ell+1}, x_\tau\rangle \le z\right\} \\
\ge\ & \sum_{\tau \in E_\ell} \mathbb{I}\left\{\epsilon_{i,\tau} \le z - \delta_\ell - \frac{1}{|E_\ell|}\right\}\,.
\end{aligned}
\tag{B.52}
$$

Furthermore, assuming events $\mathcal{G}_{i,\ell} = \{|\mathcal{S}_{i,\ell}| \le L_\ell\}$ hold for all $i \in [N]$, we can simply upper bound and lower bound the expression in the second parenthesis of Equation (B.49):

$$
-L_\ell \ \le\ \sum_{\tau \in \mathcal{S}_{i,\ell}} \mathbb{I}\left\{b_{i,\tau} - \langle \widehat{\beta}_{\ell+1}, x_\tau\rangle \le z\right\} - \sum_{\tau \in \mathcal{S}_{i,\ell}} \mathbb{I}\left\{v_{i,\tau} - \langle \widehat{\beta}_{\ell+1}, x_\tau\rangle \le z\right\} \ \le\ L_\ell\,.
\tag{B.53}
$$

Combining Equations (B.49), (B.51), (B.52), (B.53), and using the definition

$$
\widehat{F}_{\ell+1}(z) = \frac{1}{N|E_\ell|} \sum_{i \in [N]} \sum_{\tau \in E_\ell} \mathbb{I}\left\{b_{i,\tau} - \langle \widehat{\beta}_{\ell+1}, x_\tau\rangle \le z\right\},
$$

under the occurrence of events $\xi_{\ell+1}$, and $\mathcal{G}_{i,\ell}$ for all $i \in [N]$, we have

$$
\begin{aligned}
&\frac{1}{N|E_\ell|} \sum_{i\in[N]} \sum_{\tau\in E_\ell} \mathbb{I}\left\{\epsilon_{i,\tau} \le z - \delta_\ell - \frac{1}{|E_\ell|}\right\} - \frac{L_\ell}{|E_\ell|} \ \le\ \widehat{F}_{\ell+1}(z) \ \text{ and} \\
&\widehat{F}_{\ell+1}(z) \ \le\ \frac{1}{N|E_\ell|} \sum_{i\in[N]} \sum_{\tau\in E_\ell} \mathbb{I}\left\{\epsilon_{i,\tau} \ \le\ z + \delta_\ell + \frac{1}{|E_\ell|}\right\} + \frac{L_\ell}{|E_\ell|}\,. \qquad \text{(B.54)}
\end{aligned}
$$

Now, for any $\gamma > 0$,

$$
\begin{aligned}
&\mathbb{P}\left(F\left(z-\delta_\ell-\frac{1}{|E_\ell|}\right) - \widehat{F}_{\ell+1}(z) \le \gamma + \frac{L_\ell}{|E_\ell|}\right) \\
\ge\ & \mathbb{P}\left(\left\{F\left(z-\delta_\ell-\frac{1}{|E_\ell|}\right) - \widehat{F}_{\ell+1}(z) \le \gamma + \frac{L_\ell}{|E_\ell|}\right\} \bigcap \xi_{\ell+1} \bigcap \left(\bigcap_{i\in[N]} \mathcal{G}_{i,\ell}\right)\right) \\
\ge\ & \mathbb{P}\Bigg(\left\{F\left(z-\delta_\ell-\frac{1}{|E_\ell|}\right) - \frac{1}{N|E_\ell|}\sum_{i\in[N]}\sum_{\tau\in E_\ell}\mathbb{I}\left\{\epsilon_{i,\tau}\le z-\delta_\ell-\frac{1}{|E_\ell|}\right\} \le \gamma\right\} \\
&\quad \bigcap \xi_{\ell+1} \bigcap \left(\bigcap_{i\in[N]} \mathcal{G}_{i,\ell}\right)\Bigg) \\
\ge\ & \mathbb{P}\left(\left\{\sup_{\tilde z\in\mathbb{R}}\left|F(\tilde z) - \frac{1}{N|E_\ell|}\sum_{i\in[N]}\sum_{\tau\in E_\ell}\mathbb{I}\left\{\epsilon_{i,\tau}\le \tilde z\right\}\right| \le \gamma\right\} \bigcap \xi_{\ell+1} \bigcap \left(\bigcap_{i\in[N]} \mathcal{G}_{i,\ell}\right)\right) \\
\ge\ & 1 - \mathbb{P}\left(\left\{\sup_{\tilde z\in\mathbb{R}}\left|F(\tilde z) - \frac{1}{N|E_\ell|}\sum_{i\in[N]}\sum_{\tau\in E_\ell}\mathbb{I}\left\{\epsilon_{i,\tau}\le \tilde z\right\}\right| > \gamma\right\}\right) \\
&\quad - \mathbb{P}\left(\xi_{\ell+1}^c\right) - \sum_{i\in[N]}\mathbb{P}\left(\mathcal{G}_{i,\ell}^c\right) \\
\ge\ & 1 - 2\exp\left(-2N|E_\ell|\gamma^2\right) - \left(\frac{2d+N}{|E_\ell|} + d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right)\right) - \frac{N}{|E_\ell|} \\
=\ & 1 - 2\exp\left(-2N|E_\ell|\gamma^2\right) - \frac{2(d+N)}{|E_\ell|} - d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right), \qquad \text{(B.55)}
\end{aligned}
$$

where the second inequality follows from Equation (B.54), the fourth inequality uses the union bound, and the final inequality follows from the DKW inequality (Theorem B.3.1), Lemma B.2.1, and Lemma 3.4.1. We note that we can apply the DKW inequality because $\{\epsilon_{i,\tau}\}_{\tau\in E_\ell, i\in[N]}$ are $N|E_\ell|$ i.i.d. realizations of noise variables. According to the Lipschitz property of $F$ shown in Lemma B.2.5, $|F(z-\delta_\ell-1/|E_\ell|)-$

$F(z)| \leq c_f(\delta_\ell + 1/|E_\ell|)$ for $\forall z \in \mathbb{R}$. Hence, combining this with Equation (B.55), yields

$$\begin{aligned}
&\mathbb{P}\left(F(z) - \widehat{F}_{\ell+1}(z) \ \leq \ \gamma + c_f\left(\delta_\ell + \frac{1}{|E_\ell|}\right) + \frac{L_\ell}{|E_\ell|}\right) \\
\geq \ &\mathbb{P}\left(F\left(z - \delta_\ell - \frac{1}{|E_\ell|}\right) - \widehat{F}_{\ell+1}(z) \ \leq \ \gamma + \frac{L_\ell}{|E_\ell|}\right) \\
\geq \ &1 - 2\exp\left(-2N|E_\ell|\gamma^2\right) - \frac{2(d+N)}{|E_\ell|} - d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right). \qquad \text{(B.56)}
\end{aligned}$$

Similarly, $|F(z + \delta_\ell + 1/|E_\ell|) - F(z)| \leq c_f(\delta_\ell + 1/|E_\ell|)$ for $\forall z \in \mathbb{R}$, so we can show

$$\begin{aligned}
&\mathbb{P}\left(\widehat{F}_{\ell+1}(z) - F(z) \ \leq \ \gamma + c_f\left(\delta_\ell + \frac{1}{|E_\ell|}\right) + \frac{L_\ell}{|E_\ell|}\right) \\
\geq \ &\mathbb{P}\left(\widehat{F}_{\ell+1}(z) - F\left(z + \delta_\ell + \frac{1}{|E_\ell|}\right) \ \leq \ \gamma + \frac{L_\ell}{|E_\ell|}\right) \\
\geq \ &1 - 2\exp\left(-2N|E_\ell|\gamma^2\right) - \frac{2(d+N)}{|E_\ell|} - d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right). \qquad \text{(B.57)}
\end{aligned}$$

Combining Equations (B.56) and (B.57) using a union bound yields

$$\begin{aligned}
&\mathbb{P}\left(\left|\widehat{F}_{\ell+1}(z) - F(z)\right| \leq \gamma + c_f\delta_\ell + \frac{c_f + L_\ell}{|E_\ell|}\right) \\
\geq \ &1 - 4\exp\left(-2N|E_\ell|\gamma^2\right) - \frac{4(d+N)}{|E_\ell|} - 2d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right). \qquad \text{(B.58)}
\end{aligned}$$

Finally, we now bound $|\widehat{F}_t^-(z) - F^-(z)|$ and $|\widehat{F}_t^+(z) - F^+(z)|$ using the fact that

$F^-(z) = NF^{N-1}(z) - (N-1)F^N(z)$ and $F^+(z) = F^N(z)$.

$$
\begin{aligned}
|\widehat{F}^-_{\ell+1}(z) - F^-(z)| &= \left| N\widehat{F}^{N-1}_{\ell+1}(z) - (N-1)\widehat{F}^N_{\ell+1}(z) - \left(NF^{N-1}(z) - (N-1)F^N(z)\right)\right| \\
&\leq N\left|\widehat{F}^{N-1}_{\ell+1}(z) - F^{N-1}(z)\right| + (N-1)\left|\widehat{F}^N_{\ell+1}(z) - F^N(z)\right| \\
&= N\left|\left(\widehat{F}_{\ell+1}(z) - F(z)\right)\left(\sum_{n=1}^{N-1}\left(\widehat{F}_{\ell+1}(z)\right)^{n-1}(F(z))^{N-1-n}\right)\right| \\
&\quad + (N-1)\left|\left(\widehat{F}_{\ell+1}(z) - F(z)\right)\left(\sum_{n=1}^{N}\left(\widehat{F}_{\ell+1}(z)\right)^{n-1}(F(z))^{N-n}\right)\right| \\
&\leq N(N-1)\left|\widehat{F}_{\ell+1}(z) - F(z)\right| + (N-1)N\left|\widehat{F}_{\ell+1}(z) - F(z)\right| \\
&< 2N^2\left|\widehat{F}_{\ell+1}(z) - F(z)\right|. \qquad \text{(B.59)}
\end{aligned}
$$

The second equality uses $a^m - b^m = (a-b)\left(\sum_{n=1}^m a^{n-1}b^{m-n}\right)$ for any integer $m \geq 2$. The second inequality follows from $\widehat{F}_{\ell+1}(z), F(z) \in [0,1]$ for $\forall z \in \mathbb{R}$. Combining Equations (B.58) and (B.59), we get

$$
\begin{aligned}
&\mathbb{P}\left(\left|\widehat{F}^-_{\ell+1}(z) - F^-(z)\right| \leq 2N^2\left(\gamma + c_f\delta_\ell + \frac{c_f + L_\ell}{|E_\ell|}\right)\right) \\
\geq\ & 1 - 4\exp\left(-2N|E_\ell|\gamma^2\right) - \frac{4(d+N)}{|E_\ell|} - 2d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right).
\end{aligned}
$$

The probability bound for $\left|\widehat{F}^-_{\ell+1}(z) - F^-(z)\right|$ can be shown in a similar fashion by noting that similar to Equation (B.59) we can show $|\widehat{F}^+_{\ell+1}(z) - F^+(z)| < N\left|\widehat{F}_{\ell+1}(z) - F(z)\right|$.

□

**Lemma B.2.3** (Bounding the Impact of Estimation Errors on Revenue)**.** *We assume that the events*

$$
\begin{aligned}
\xi_{\ell+1} &= \left\{\|\widehat{\beta}_{\ell+1} - \beta\|_1 \leq \frac{\delta_\ell}{x_{\max}}\right\} \\
\xi^-_{\ell+1} &= \left\{\left|\widehat{F}^-_{\ell+1}(z) - F^-(z)\right| \leq 2N^2\left(\gamma_\ell + c_f\delta_\ell + \frac{c_f + L_\ell}{|E_\ell|}\right)\right\} \\
\xi^+_{\ell+1} &= \left\{\left|\widehat{F}^+_{\ell+1}(z) - F^+(z)\right| \leq N\left(\gamma_\ell + c_f\delta_\ell + \frac{c_f + L_\ell}{|E_\ell|}\right)\right\}
\end{aligned}
$$

*occur for some phase $\ell \geq 1$, where $z \in \mathbb{R}$, $\gamma_\ell = \sqrt{\log(|E_\ell|)}/\sqrt{2N|E_\ell|}$, and $\delta_\ell$ is defined in Equation (B.6). Hence for any $r \in \{r_t^\star,\ r_t\}$ where $t \in E_{\ell+1}$ we have the following:*

*(i)* $|\rho_t(r, y_t, F^-, F^+) - \rho_t(r, \widehat{y}_t, F^-, F^+)| \ \leq\ 3rc_f N^2 \delta_\ell \qquad a.s.$

*(ii)* $\left|\rho_t(r, \widehat{y}_t, F^-, F^+) - \rho_t(r, \widehat{y}_t, \widehat{F}^-_{\ell+1}, \widehat{F}^+_{\ell+1})\right| \ \leq\ 3rN^2 \left(\gamma_\ell + c_f\delta_\ell + \frac{c_f + L_\ell}{|E_\ell|}\right) \qquad a.s.$

*where $y_t = \langle \beta, x_t\rangle$, $\widehat{y}_t = \langle \widehat{\beta}_{\ell+1}, x_t\rangle$, $\widehat{\beta}_{\ell+1}, \widehat{F}^-_{\ell+1}, \widehat{F}^+_{\ell+1}$ are defined in Equations (3.8) and (3.9). The function $\rho_t$ is defined in Equation (B.15).*

*Proof.* Proof of Lemma B.2.3. **Part (i)** We consider the following:

$$
\begin{aligned}
& \left|\rho_t(r, y_t, F^-, F^+) - \rho_t(r, \widehat{y}_t, F^-, F^+)\right| \\
= \ & \left| \int_0^r \left[F^-(z - y_t) - F^-(z - \widehat{y}_t)\right] dz - r\left[F^+(r - y_t) - F^+(r - \widehat{y}_t)\right] \right| \\
\leq \ & \int_0^r \left|F^-(z - y_t) - F^-(z - \widehat{y}_t)\right| dz + r\left|F^+(r - y_t) - F^+(r - \widehat{y}_t)\right| \\
\leq \ & \int_0^r 2c_f N^2 |y_t - \widehat{y}_t| dz + rc_f N |y_t - \widehat{y}_t| \\
\leq \ & \int_0^r 2c_f N^2 \left(\|\widehat{\beta}_{\ell+1} - \beta\|_1 x_{\max}\right) dz + rc_f N \|\widehat{\beta}_{\ell+1} - \beta\|_1 x_{\max} \\
\leq \ & 3rc_f N^2 \delta_\ell \,.
\end{aligned}
$$

The first equality follows from definition of $\rho_t$ in Equation (B.15), and the second inequality applies the Lipschitz property of $F^-$ and $F^+$ using Lemma B.2.5. The third inequality follows from Cauchy's inequality: $|y_t - \widehat{y}_t| = |\langle \widehat{\beta}_{\ell+1} - \beta, x_t\rangle| \leq \|\widehat{\beta}_{\ell+1} - \beta\|_1 x_{\max}$, and the last inequality follows from the occurrence of $\xi_{\ell+1}$ and $N \geq 1$.

**Part (ii)** Similar to part (i), we have

$$
\begin{aligned}
& \left|\rho_t(r,\widehat{y}_t,F^-,F^+) - \rho_t(r,\widehat{y}_t,\widehat{F}^-_{\ell+1},\widehat{F}^+_{\ell+1})\right| \\
= & \left|\int_0^r \left[F^-(z-\widehat{y}_t) - \widehat{F}^-_{\ell+1}(z-\widehat{y}_t)\right] dz - r\left[F^+(r-\widehat{y}_t) - \widehat{F}^+_{\ell+1}(r-\widehat{y}_t)\right]\right| \\
\leq & \int_0^r \left|F^-(z-\widehat{y}_t) - \widehat{F}^-_{\ell+1}(z-\widehat{y}_t)\right| dz + r\left|F^+(r-\widehat{y}_t) - \widehat{F}^+_{\ell+1}(r-\widehat{y}_t)\right| \\
\leq & \; 3rN^2\left(\gamma_\ell + c_f\delta_\ell + \frac{c_f+L_\ell}{|E_\ell|}\right),
\end{aligned}
$$

where the last inequality follows from the occurrence of events $\xi^-_{\ell+1}$ and $\xi^+_{\ell+1}$ and $N \geq 1$. □

**Lemma B.2.4** (Bounding probabilities)**.** *The probability that not all events $\xi_{\ell+1}$, $\xi^-_{\ell+1}$ and $\xi^+_{\ell+1}$ occur for some phase $\ell \geq 1$ is bounded as*

$$
\mathbb{P}\left(\xi^c_{\ell+1} \cup \left(\xi^-_{\ell+1}\right)^c \cup \left(\xi^+_{\ell+1}\right)^c\right) \;\leq\; \frac{9N+15d+8}{|E_\ell|},
$$

*where the events $\xi_{\ell+1}$, $\xi^-_{\ell+1}$ and $\xi^+_{\ell+1}$ are defined in Equations (B.5), (B.7), and (B.8) respectively.*

*Proof.* Proof of Lemma B.2.4.

We first bound the probability of $\xi^c_{\ell+1}$, and then proceed to bound the the probability of $\left(\xi^-_{\ell+1}\right)^c$ and $\left(\xi^+_{\ell+1}\right)^c$.

Recall that $\xi_{\ell+1} = \left\{\|\widehat{\beta}_{\ell+1} - \beta\|_1 \leq \frac{\delta_\ell}{x_{\max}}\right\}$. Then,

$$
\begin{aligned}
\mathbb{P}\left(\xi^c_{\ell+1}\right) \;&\leq\; \frac{2d+N}{|E_\ell|} + d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x^2_{\max}}\right) \\
&\leq\; \frac{2d+N}{|E_\ell|} + d\exp\left(-\frac{\log(|E_\ell|)T^{\frac{1}{4}}\lambda_0^2}{8x^2_{\max}}\right) \\
&\leq\; \frac{N+3d}{|E_\ell|},
\end{aligned} \tag{B.60}
$$

where the first inequality follows from Lemma B.2.1 by taking $\gamma = \frac{\sqrt{2d\log(|E_\ell|)}\epsilon_{\max}x_{\max}}{\lambda_0^2\sqrt{N|E_\ell|}}$; the second inequality uses the fact that $|E_\ell| \geq |E_1| = \sqrt{T}$, $\quad T \geq \max\left\{\left(\frac{8x^2_{\max}}{\lambda_0^2}\right)^4, 9\right\}$, which implies $|E_\ell| \geq \log(|E_\ell|)\sqrt{|E_\ell|} \geq T^{\frac{1}{4}} \log(|E_\ell|)$. Note that here we used the fact that $\sqrt{x} \geq \log(x)$ for all $x \geq 9$.

We now bound the probability of $\left(\xi_{\ell+1}^-\right)^c$:

$$
\begin{aligned}
\mathbb{P}\left(\left(\xi_{\ell+1}^-\right)^c\right) \;&\leq\; 4\exp\left(-2N|E_\ell| \cdot \left(\frac{\sqrt{\log(|E_\ell|)}}{\sqrt{2N|E_\ell|}}\right)^2\right) + \frac{4(d+N)}{|E_\ell|} + 2d\exp\left(-\frac{|E_\ell|\lambda_0^2}{8x_{\max}^2}\right) \\
&\leq\; \frac{2(2N+3d+2)}{|E_\ell|}\,,
\end{aligned}
\tag{B.61}
$$

where the first inequality follows from Lemma B.2.2 by taking $\gamma = \gamma_\ell = \sqrt{\log(|E_\ell|)}/\sqrt{2N|E_\ell|}$, and the last inequality again uses the fact that $|E_\ell| \geq \log(|E_\ell|)\sqrt{|E_\ell|} \geq T^{\frac{1}{4}}\log(|E_\ell|)$ when $T \geq \max\left\{\left(\frac{8x_{\max}^2}{\lambda_0^2}\right)^4, 9\right\}$.

Similarly, we can bound the probability of $\left(\xi_{\ell+1}^+\right)^c$:

$$
\mathbb{P}\left(\left(\xi_{\ell+1}^+\right)^c\right) \;\leq\; \frac{2(2N+3d+2)}{|E_\ell|}\,, \tag{B.62}
$$

Finally, combining Equations (B.60), (B.61) and (B.62), we have

$$
\mathbb{P}\left(\xi_{\ell+1}^c \cup \left(\xi_{\ell+1}^-\right)^c \cup \left(\xi_{\ell+1}^+\right)^c\right) \;\leq\; \mathbb{P}\left(\xi_{\ell+1}^c\right) + \mathbb{P}\left(\left(\xi_{\ell+1}^-\right)^c\right) + \mathbb{P}\left(\left(\xi_{\ell+1}^+\right)^c\right) \;\leq\; \frac{9N+15d+8}{|E_\ell|}\,.
$$

□

**Lemma B.2.5** (Lipschitz Property for $F$, $F^-$ and $F^+$)**.** *The following hold for any $z_1, z_2 \in \mathbb{R}$:*

*(i)* $|F(z_1) - F(z_2)| \leq c_f|z_1 - z_2|$.

*(ii)* $|F^-(z_1) - F^-(z_2)| \leq 2c_f N^2|z_1 - z_2|$.

*(iii)* $|F^+(z_1) - F^+(z_2)| \leq c_f N|z_1 - z_2|$.

*Here,* $0 < c_f = \sup_{z\in[-\epsilon_{\max},\epsilon_{\max}]} f(z)$.

*Proof.* Proof of Lemma B.2.5. Without loss of generality, we assume $z_1 < z_2$. Note that $F(z) = 0$ for $\forall z \in (-\infty, -\epsilon_{\max}]$, and $F(z) = 1$ for $\forall z \in [\epsilon_{\max}, \infty)$.

**Part (i)** We consider the following cases:

1. **Case 1:** ($z_1 < z_2 \leq -\epsilon_{\max}$ or $\epsilon_{\max} \leq z_1 < z_2$): $|F(z_2) - F(z_1)| = 0 \leq c_f|z_2 - z_1|$.

2. **Case 2:** ($-\epsilon_{\max} < z_1 < z_2 < \epsilon_{\max}$): By the mean value theorem, $|F(z_2) - F(z_1)| = f(\tilde{z})|z_2 - z_1| < c_f|z_2 - z_1|$, where $\tilde{z} \in (z_1, z_2)$.

3. **Case 3:** ($z_1 \leq -\epsilon_{\max} < z_2 < \epsilon_{\max}$): We have $|z_2 - (-\epsilon_{\max})| = z_2 - (-\epsilon_{\max}) \leq z_2 - z_1$ and $F(z_1) = F(-\epsilon_{\max}) = 0$. Hence $|F(z_2) - F(z_1)| = |F(z_2) - F(-\epsilon_{\max})| = f(\tilde{z})|z_2 - (-\epsilon_{\max})| \leq c_f|z_2 - z_1|$, where $\tilde{z} \in (-\epsilon_{\max}, z_2)$ by the mean value theorem.

4. **Case 4:** ($-\epsilon_{\max} < z_1 < \epsilon_{\max} \leq z_2$): We have $|\epsilon_{\max} - z_1| = \epsilon_{\max} - z_1 \leq z_2 - z_1$ and $F(z_2) = F(\epsilon_{\max}) = 1$ . Hence $|F(z_2) - F(z_1)| = |F(\epsilon_{\max}) - F(z_1)| = f(\tilde{z})|\epsilon_{\max} - z_1| \leq c_f|z_2 - z_1|$, where $\tilde{z} \in (z_1, \epsilon_{\max})$ by the mean value theorem.

**Part (ii) & (iii)** We recall that $F^-(z) = NF^{N-1}(z) - (N-1)F^N(z)$ and $F^+(z) = F^N(z)$, so

$$
\begin{aligned}
& |F^-(z_2) - F^-(z_1)| \\
= \; & \left|NF^{N-1}(z_2) - (N-1)F^N(z_2) - \left(NF^{N-1}(z_1) - (N-1)F^N(z_1)\right)\right| \\
\leq \; & N\left|F^{N-1}(z_2) - F^{N-1}(z_1)\right| + (N-1)\left|F^N(z_2) - F^N(z_1)\right| \\
= \; & N\left|(F(z_2) - F(z_1))\left(\sum_{n=1}^{N-1} (F(z_2))^{n-1}(F(z_1))^{N-1-n}\right)\right| \\
& \quad + (N-1)\left|(F(z_2) - F(z_1))\left(\sum_{n=1}^{N} (F(z_2))^{n-1}(F(z_1))^{N-n}\right)\right| \\
\leq \; & N(N-1)\left|F(z_2) - F(z_1)\right| + (N-1)N\left|F(z_2) - F(z_1)\right| \\
< \; & 2N^2 c_f|z_2 - z_1| \,.
\end{aligned}
$$

The second equality uses $a^m - b^m = (a-b)\left(\sum_{n=1}^{m} a^{n-1}b^{m-n}\right)$ for any $a, b \in \mathbb{R}$ and integer $m \geq 2$. The second inequality follows from $F(z) \in [0,1]$ for $\forall z \in \mathbb{R}$. The final inequality follows from the Lipschitz property of $F$ shown in part (i). Following the same arguments, we can also show that $|F^+(z_2) - F^+(z_1)| \leq c_f N|z_2 - z_1|$. $\square$

## B.3 Supplementary lemmas

**Lemma B.3.1** (Dvoretzky-Kiefer-Wolfowitz Inequality ([44])). *Let $Z_1, Z_2, \dots Z_n$ be i.i.d. random variables with cumulative distribution function $F$, and denote the associated empirical distribution function as*

$$\widehat{F}(z) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}\{Z_i \leq z\} \quad , z \in \mathbb{R}\,. \tag{B.63}$$

*Then, for any $\bar{\gamma} > 0$,*

$$\mathbb{P}\left(\sup_{z\in\mathbb{R}} \left|\widehat{F}(z) - F(z)\right| \leq \bar{\gamma}\right) \geq 1 - 2\exp\left(-2n\bar{\gamma}^2\right)\,. \tag{B.64}$$

**Lemma B.3.2** (Matrix Chernoff Bound ([104])). *Consider a finite sequence of independent, random matrices $\{Z_k \in \mathbb{R}^d\}_{k\in[K]}$. Assume that $0 \leq \lambda_{\min}(Z_k)$ and $\lambda_{\max}(Z_k) \leq B$ for any $k$. Denote $Y = \sum_{k\in[K]} Z_k$, $\mu_{\min} = \lambda_{\min}(\mathbb{E}[Y])$, and $\mu_{\max} = \lambda_{\max}(\mathbb{E}[Y])$. Then for $\forall\bar{\gamma} \in (0,1)$,*

$$\mathbb{P}\left(\lambda_{\min}(Y) \leq \bar{\gamma}\mu_{\min}\right) \leq d\exp\left(-\frac{(1-\bar{\gamma})^2\mu_{\min}}{2B}\right)\,.$$

**Lemma B.3.3** (Multiplicative Azuma Inequality([78])). *Let $Z_1 = \sum_{\tau\in[\widetilde{T}]} z_{1,\tau}$ and $Z_2 = \sum_{\tau\in[\widetilde{T}]} z_{2,\tau}$ be sums of non-negative random variables, where $\widetilde{T}$ is a random stopping time with a finite expectation, and, for all $\tau \in [\widetilde{T}]$, $|z_{1,\tau} - z_{2,\tau}| \leq 1$ and $\mathbb{E}\left[(z_{1,\tau} - z_{2,\tau}) \;\middle|\; \sum_{s<\tau} z_{1,s}, \sum_{s<\tau} z_{2,s}\right] \leq 0$. Let $\tilde{\gamma} \in [0,1]$ and $A \in \mathbb{R}$. Then,*

$$\mathbb{P}\left((1-\tilde{\gamma})Z_1 \geq Z_2 + A\right) \leq \exp\left(-\tilde{\gamma}A\right)$$

# Appendix C

# Supplementary material for Chapter 4

## C.1 Additional definitions

In this section, we introduce some additional definitions that will be used throughout the appendices.

**Definition C.1.1** (Threshold vectors). *We say that an $N$-dimensional vector $\boldsymbol{x} \in \mathbb{R}^N$ is a threshold vector if it takes the form of $\boldsymbol{x} = (1 \dots 1, q, 0 \dots 0)$, where the first $J \in \{0, \dots N\}$ entries are 1's, followed by some number $q \in [0, 1)$, and trailing with $(N - J - 1)_+$ 0's.*[1] *Any threshold vector is uniquely characterized by its dimension $N$, as well as, a tuple $(J, q) \in \{0, \dots N\} \times [0, 1)$, so we denote the vector as $\psi(J, q)$. In the special case when $J = N$, take $q = 0$.*

For any two vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$, let $\min\{\boldsymbol{a}, \boldsymbol{b}\} = (\min\{a_i, b_i\})_{i \in [n]}$ be the element-wise minimum. We write $\boldsymbol{a} \preceq \boldsymbol{b}$ if and only if $a_i \leq b_i$ and $\boldsymbol{a} \succeq \boldsymbol{b}$ if and only if $a_i \geq b_i$ for all $i \in [n]$.

## C.2 Additional material for Section 4.3

### C.2.1 Proof of Lemma 4.3.1

Here, we show a more detailed version of the lemma stated as followed:

[1] For the edge case of $(1, \dots 1) \in \mathbb{R}^N$, $J = N$ and hence the number of trailing 0's is $(N - J - 1)_+ = 0$.

**Theorem C.2.1** (Detailed version of Lemma 4.3.1)**.** *For a fixed price $d$, define*

$$
\begin{aligned}
&r = \max\left\{n \in [N] : \sum_{\ell\in[n]} g_\ell\left(V_\ell - \gamma d\right) \geq 0\right\}, \quad q_{\mathtt{R}} = \frac{\sum_{k\in[r]} g_n\left(V_n - \gamma d\right)}{g_{r+1}\cdot\left|V_{r+1}-\gamma d\right|}\,, \\
&b = \max\left\{n \in [N] : d\sum_{\ell\in[n]} g_\ell \leq \rho\right\}, \quad \text{and} \quad q_{\mathtt{B}} = \frac{\rho - d\sum_{k\in[b]} g_n}{g_{b+1}\cdot d}\,,
\end{aligned}
\tag{C.1}
$$

*If we let $\boldsymbol{x}_{\mathtt{R}} = \psi(r, q_{\mathtt{R}})$ and $\boldsymbol{x}_{\mathtt{B}} = \psi(b, q_{\mathtt{B}})$ be two threshold vectors (see Definition C.1.1), then $\boldsymbol{x}_d = \min\left\{\boldsymbol{x}_{\mathtt{R}}, \boldsymbol{x}_{\mathtt{B}}\right\}$ is the unique optimal solution to $U(d)$ in Equation* (4.3)*. Furthermore, $\boldsymbol{x}_d$ is also a threshold vector characterized by tuple $(J, q)$ where*

$$
J = \min\{r, b\}, \quad q = x_{d,J+1} = \min\left\{x_{\mathtt{B},J+1}, x_{\mathtt{R},J+1}\right\}\,. \tag{C.2}
$$

*Proof.*

Our proof for Theorem C.2.1 consists of 3 steps:

- **Step 1.** We show that $\boldsymbol{x}_{\mathtt{B}}$ is the unique optimal solution to the "budget constraint only" problem:

$$
\texttt{P-Budget} = \max_{\boldsymbol{x}\in[0,1]^N} \sum_{n\in[N]} g_n V_n x_n \text{ s.t. } d\sum_{n\in[N]} g_n x_n \leq \rho\,, \tag{C.3}
$$

- **Step 2.** We show that $\boldsymbol{x}_{\mathtt{R}}$ is the unique optimal solution the "ROI constraint only" problem:

$$
\texttt{P-ROI} = \max_{\boldsymbol{x}\in[0,1]^N} \sum_{n\in[N]} g_n V_n x_n \text{ s.t. } \sum_{n\in[N]} g_n\left(V_n - \gamma d\right)x_n \geq 0\,, \tag{C.4}
$$

- **Step 3.** We show that $\boldsymbol{x}_d = \min\{\boldsymbol{x}_{\mathtt{B}}, \boldsymbol{x}_{\mathtt{R}}\}$ is feasible to $U(d)$. In other words, we show $\boldsymbol{x}_d$ is feasible to both `P-Budget` and `P-ROI`.

**Step 1.** We recognize that `P-Budget` is the linear program (LP) relaxation of a 0-1 knapsack problem, in which the items' "value-to-cost ratio", namely $\frac{g_n V_n}{d g_n} = \frac{V_n}{d}$ are ordered: $\frac{V_1}{d} > \ldots \frac{V_N}{d}$ since $V_1 > \ldots V_N > 0$. Therefore, it is a well known result that the unique optimal solution to `P-Budget` is exactly $\boldsymbol{x}_{\mathtt{B}}$ (a threshold vector) defined in the statement of Theorem C.2.1; see e.g. [32] for the optimal solution to the 0-1 knapsack LP relaxation.

**Step 2.** Let $\widetilde{\boldsymbol{x}} \in [0,1]^N$ be any optimal solution to `P-ROI`. Define $\kappa = \max\{n \in [N] : V_n \geq \gamma d\}$ so that $V_n \geq \gamma d$ for all $n \leq \kappa$. Then it is easy to see for any $n = 1 \ldots \kappa$, we have $\widetilde{x}_n = 1$. This is because if there exists some $j \leq \kappa$ such that $\widetilde{x}_j < 1$, then the solution $\boldsymbol{x} = (\widetilde{x}_1 \ldots \widetilde{x}_{j-1}, 1, \widetilde{x}_{j+1}, \ldots \widetilde{x}_N)$ is feasible and yields a strictly larger objective than $\widetilde{\boldsymbol{x}}$:

$$\sum_{n\in[N]} g_n V_n x_n - \sum_{n\in[N]} g_n V_n \widetilde{x}_n = V_j(1-\widetilde{x}_j) > 0\,. \tag{C.5}$$

Hence, the optimal solution to `P-ROI` takes the form of $\widetilde{\boldsymbol{x}} = (\underbrace{1\ldots1}_{\kappa\ 1\text{'s}}, y_{\kappa+1}, \ldots y_N) \in [0,1]^N$. Hence, we know that $\widetilde{\boldsymbol{y}} := (y_\kappa, \ldots y_N)$ must satisfy

$$\widetilde{\boldsymbol{y}} \;\in\; \arg\max_{\boldsymbol{x}\in[0,1]^{N-\kappa}} \;\sum_{k=\kappa+1}^{N} g_n V_n x_n \text{ s.t. } \sum_{k=\kappa+1}^{N} g_n\left(\gamma d - V_n\right) x_n \leq \widetilde{c}\,, \tag{C.6}$$

where we defined $\widetilde{c} = \sum_{n\in[\kappa]} g_n\left(V_n - \gamma d\right) > 0$. Note that we have $\gamma d - V_n > 0$ for all $k = \kappa+1\ldots N$, and hence the optimization problem in Equation (C.6) is again an LP relaxation of the 0-1 knapsack problem. Thus similar to Step 1, we again consider the "value-to-cost-ratios": for any $i, j \in \{\kappa+1\ldots N\}$, we have

$$V_i > V_j \iff \frac{g_i V_i}{g_i\left(\gamma d - V_i\right)} > \frac{g_j V_j}{g_j\left(\gamma d - V_j\right)}\,.$$

Hence the "value-to-cost-ratios" $\frac{g_n V_n}{g_n(\gamma d - V_n)}$ decreases in $n$ for $n \in \{\kappa+1\ldots N\}$. Therefore, the optimal solution $\widetilde{\boldsymbol{y}}$ to the 0-1 knapsack LP relaxation in Equation (C.6) is again unique, and is a threshold vector (again see [32]). Hence, the unique optimal solution to `P-ROI`is a threshold vector, and following Step 1., it is easy to see this unique optimal solution is $\boldsymbol{x}_{\mathtt{R}}$ defined in the statement of Theorem C.2.1.

**Step 3.** Since $g_n d > 0$ for all $n \in [N]$ and $\boldsymbol{x}_d = \min\{\boldsymbol{x}_{\mathtt{B}}, \boldsymbol{x}_{\mathtt{R}}\} \preceq \boldsymbol{x}_{\mathtt{B}}$, we can apply Lemma C.2.3 (i) with $a_n = g_n d$, $\boldsymbol{Z} = \boldsymbol{x}_{\mathtt{B}}$ and $\boldsymbol{Y} = \boldsymbol{x}_d$, which yields

$$d \sum_{n \in [N]} g_n x_{d,n} \leq d \sum_{n \in [N]} g_n x_{\mathtt{B},n} \leq \rho \,,$$

where the last inequality is due to the fact that $\boldsymbol{x}_{\mathtt{B}}$ is feasible to `P-Budget`. This implies $\boldsymbol{x}_d$ is also feasible to `P-Budget`.

On the other hand, again define $\kappa = \max\{n \in [N] : V_n \geq \gamma d\}$ so that $V_n \geq \gamma d$ for all $n \leq \kappa$. Then since $\boldsymbol{x}_d = \min\{\boldsymbol{x}_{\mathtt{B}}, \boldsymbol{x}_{\mathtt{R}}\} \preceq \boldsymbol{x}_{\mathtt{R}}$, and since $g_n (V_n - \gamma d) > 0$ for $n = 1 \dots \kappa$ and $g_n (V_n - \gamma d) < 0$ for $n = \kappa + 1 \dots N$, we can apply Lemma C.2.3 (ii) with $b_n = g_n (V_n - \gamma d)$, $\boldsymbol{Z} = \boldsymbol{x}_{\mathtt{R}}$ and $\boldsymbol{Y} = \boldsymbol{x}_d$, which shows

$$\sum_{n \in [N]} g_n (V_n - \gamma d) \, x_{\mathtt{R},n} \overset{(i)}{\geq} 0 \overset{(ii)}{\Longrightarrow} \sum_{n \in [N]} g_n (V_n - \gamma d) \, x_{d,k} \geq 0 \,,$$

where (i) follows from the fact that $\boldsymbol{x}_{\mathtt{R}}$ is feasible to `P-ROI` and (ii) follows from the first half of Lemma C.2.3 (ii). Hence $\boldsymbol{x}_d$ is also feasible to `P-ROI`.

The rest of the proof is straightforward: `P-Budget`, `P-ROI` and $U(d)$ have the same objectives, while each of `P-Budget` and `P-ROI` has one less constraint than $U(d)$, respectively. So `P-Budget` $\geq U(d)$ and `P-ROI` $\geq U(d)$. If $\boldsymbol{x}_d = \boldsymbol{x}_{\mathtt{B}}$, because from Step 3. we know $\boldsymbol{x}_d$ is feasible to $U(d)$, then `P-Budget` $= U(d)$ and $\boldsymbol{x}_d$ is the optimal solution to both `P-Budget` and $U(d)$. Similarly, when $\boldsymbol{x}_d = \boldsymbol{x}_{\mathtt{R}}$, $\boldsymbol{x}_d$ is the optimal solution to both `P-ROI` and $U(d)$.

Finally, we argue $\boldsymbol{x}_d$ is the unique optimal solution to $U(d)$. Assume by contradiction there exists some other vector $\boldsymbol{x} \in [0,1]^N$ that is an optimal solution to $U(d)$ and $\boldsymbol{x}_d \neq \boldsymbol{x}$. Then, again if $\boldsymbol{x}_d = \boldsymbol{x}_{\mathtt{B}}$, we know that `P-Budget` $= U(d)$, and because both $\boldsymbol{x}_d, \boldsymbol{x}$ achieve total value $U(d)$, then both $\boldsymbol{x}_d, \boldsymbol{x}$ are optimal solutions to `P-Budget`, which contradicts uniqueness of the optimal solution to `P-Budget` as argued in Step 1. Similarly, we can again arrive at a contradiction for the case when $\boldsymbol{x}_d = \boldsymbol{x}_{\mathtt{R}}$. Hence, the optimal solution to $U(d)$ is unique.

□

### C.2.2 Proof of Proposition 4.3.2

The proof for this proposition consists of two steps. First, we show that the buyer's optimal hindsight problem w.r.t. a single price $d$, namely $\mathrm{OPT}(d)$ in Equation (4.3.1) is upper bounded by $T \cdot U(d)$, which is the single-period myopic optimization problem denoted in Equation (4.3). Next, we show playing the threshold strategy w.r.t. $\boldsymbol{x}_d \in [0,1]^N$ (i.e. the optimal solution to $U(d)$) every period, gives the buyer a total value exactly $T \cdot U(d)$ while simultaneously satisfying both budget and ROI constraints. Therefore playing the threshold strategy w.r.t. $\boldsymbol{x}_d$ is the optimal value maximizing strategy to the buyer under a fixed price across all periods.

**Step 1.** Recall the linear program (LP) in Equation (4.3) that denotes the buyer's single-period myopic optimization problem. It is easy to see the optimal value is bounded and the LP is feasible (consider the solution with all entries set to be 0). Then, strong duality holds, and therefore for any $d$, there exists corresponding optimal dual variables $(\lambda, \mu) \in \mathbb{R}_+^2$ s.t.

$$\begin{aligned} U(d) &= \max_{\boldsymbol{x}\in[0,1]^N} \sum_{n\in[N]} \left(g_n(1+\lambda)V_n - (\gamma\lambda+\mu)d\right) x_n + \rho\mu \\ &= \sum_{n\in[N]} \left(g_n(1+\lambda)V_n - (\gamma\lambda+\mu)d\right)_+ + \rho\mu \end{aligned} \tag{C.7}$$

On the other hand, when the sequence of posted prices stays constant at $d$, we have

$$\begin{aligned} \mathrm{OPT}(d) &\leq \max_{\boldsymbol{z}\in[0,1]^T} \sum_{t\in[T]} \mathbb{E}\left[\left((1+\lambda)v_t - (\gamma\lambda+\mu)d\right) z_t\right] + T\rho\mu \\ &\leq \sum_{t\in[T]} \mathbb{E}\left[\left((1+\lambda)v_t - (\gamma\lambda+\mu)d\right)_+\right] + T\rho\mu \\ &= T\left(\sum_{n\in[N]} g_n\left((1+\lambda)V_n - (\gamma\lambda+\mu)d\right)_+ + \rho\mu\right) \\ &= T\cdot U(d) \end{aligned} \tag{C.8}$$

**Step 2.** Let $\boldsymbol{x}_d \in [0,1]^N$ be the optimal solution to $U(d)$ in Equation (4.3). Then, the threshold strategy w.r.t $\boldsymbol{x}_d$ (see Definition 4.3.1) can be represented as

$$z_t^* = \sum_{n\in[N]} x_{d,n}\mathbb{I}\{v_t = V_n\} \tag{C.9}$$

It is easy to see $\{z_t^*\}_{t\in[T]}$ is feasible to the buyer's optimal hindsight problem $\mathrm{OPT}(d)$ because:

$$\begin{aligned}
\mathbb{E}\left[\sum_{t\in[T]}(v_t-\gamma d)\,z_t^*\right] &= \sum_{t\in[T]}\mathbb{E}\left[(v_t-\gamma d)\sum_{n\in[N]}x_{d,n}\mathbb{I}\{v_t=V_n\}\right]\\
&= \sum_{t\in[T]}\sum_{n\in[N]}g_n(V_n-\gamma d)x_{d,n} \overset{(i)}{\geq} 0
\end{aligned} \tag{C.10}$$

and

$$\begin{aligned}
\mathbb{E}\left[\sum_{t\in[T]}dz_t^*\right] &= d\sum_{t\in[T]}E\left[\sum_{n\in[N]}x_{d,n}\mathbb{I}\{v_t=V_n\}\right]\\
&= T\cdot d\sum_{t\in[T]}\sum_{n\in[N]}g_n x_{d,n}\\
&\overset{(ii)}{\leq} \rho T
\end{aligned} \tag{C.11}$$

where both (i) and (ii) hold because $\boldsymbol{x}_d$ is feasible to $U(d)$. Finally, the threshold strategy yields a total value exactly $TU(d)$ because

$$\sum_{t\in[T]}\mathbb{E}\left[v_t z_t^*\right] = \sum_{t\in[T]}\mathbb{E}\left[v_t\sum_{n\in[N]}x_{d,n}\mathbb{I}\{v_t=V_n\}\right] = T\cdot\sum_{n\in[N]}g_nV_nx_{d,n} = T\cdot U(d)\,, \tag{C.12}$$

where the final equality follows from the fact that $\boldsymbol{x}_d$ is optimal to $U(d)$.

Therefore, in light of the upper bound shown in Equation (C.8), the threshold strategy in Equation (C.9) is optimal to the buyer's hindsight problem $\mathrm{OPT}(d)$.

Finally, the seller's revenue under the buyer's optimal threshold strategy is

$d\sum_{t\in[T]} z_t^* = T \cdot \sum_{n\in[N]} g_n x_{d,n} = T \cdot \pi(d)$ where $\pi(d)$ is the per-period revenue defined in Equation (4.4). $\square$

### C.2.3 Proof of Theorem 4.3.3

Our proof relies on the following proposition

**Proposition C.2.2.** *If price $d$ is nonbinding, then the corresponding optimal solution $\boldsymbol{x}_d$ to $U(d)$ is $\boldsymbol{x}_d = (1\dots 1) \in \mathbb{R}_+^n$.*

*Proof.* We prove the claim via contradiction. Assume there is some index $k \in [N]$ such that $\boldsymbol{x}_{d,k} < 1$. Then consider the solution $\boldsymbol{x} = (x_{d,1}\dots x_{d,k-1}, y, x_{d,k+1}, \dots x_{d,n})$ where we replaced the $k$'th entry of $\boldsymbol{x}_d$ with

$$y = x_{d,k} + \epsilon, \text{ where } \epsilon := \min\left\{1 - x_{d,k}, \frac{\rho - \sum_{n\in[N]} g_n x_{d,n}}{d g_k}, \frac{\sum_{n\in[N]}(V_n - \gamma d)\, g_n x_{d,n}}{|V_k - \gamma d|\, g_k}\right\} \overset{(i)}{>} 0\,,$$

where (i) follows from the fact that $\boldsymbol{x}_d$ is nonbinding, i.e. $\rho > \sum_{n\in[N]} g_n x_{d,n}$ and $\sum_{n\in[N]}(V_n - \gamma d)\, g_n x_{d,n} > 0$. Then

$$d\sum_{n\in[N]} g_n x_n \;=\; d\sum_{n\in[N]} g_n x_{d,n} + d g_k \epsilon \;\leq\; d\sum_{n\in[N]} g_n x_{d,n} + \left(\rho - \sum_{n\in[N]} g_n x_{d,n}\right) = \rho\,.$$

On the other hand, if $V_k - \gamma d > 0$, then

$$\sum_{n\in[N]}(V_n - \gamma d)\, g_n x_{d,n} = \sum_{n\in[N]}(V_n - \gamma d)\, g_n x_{d,n} + (V_k - \gamma d)\, g_k \epsilon > \sum_{n\in[N]}(V_n - \gamma d)\, g_n x_{d,n} > 0\,.$$

If $V_k - \gamma d < 0$, then

$$\begin{aligned}\sum_{n\in[N]}(V_n - \gamma d)\, g_n x_{d,n} \;&=\; \sum_{n\in[N]}(V_n - \gamma d)\, g_n x_{d,n} + (V_k - \gamma d)\, g_k \epsilon \\ &\geq\; \sum_{n\in[N]}(V_n - \gamma d)\, g_n x_{d,n} + (V_k - \gamma d) \cdot \frac{\sum_{n\in[N]}(V_n - \gamma d)\, g_n x_{d,n}}{|V_k - \gamma d|} = 0\end{aligned}$$

where in the last equality we used $|V_n - \gamma d| = -(V_n - \gamma d)$ since $V_n - \gamma d < 0$.

The above shows $\boldsymbol{x}$ is feasible to $U(d)$. On the other hand, $\sum_{n\in[N]} V_n g_n x_{d,n} < \sum_{n\in[N]} V_n g_n x_n$, so $\boldsymbol{x}$ yields a strictly larger objective than $\boldsymbol{x}_d$, contradicting the optimality of $\boldsymbol{x}_d$. $\square$

We now return to our proof for Theorem 4.3.3.

*(1).* When both $d, \widetilde{d}$ are non-binding, Proposition C.2.2 implies $\boldsymbol{x}_d = \boldsymbol{x}_{\widetilde{d}} = (1\ldots1)$.

$$\mathsf{rev}(d) = d\sum_{n\in[N]} g_n x_{d,n} = d\sum_{n\in[N]} g_n < \widetilde{d}\sum_{n\in[N]} g_n = \widetilde{d}\sum_{n\in[N]} g_n x_{\widetilde{d},n} = \mathsf{rev}(\widetilde{d})\,.$$

*(2).* We prove this claim by contradiction. Assume $\widetilde{d}$ is non-binding and $\widetilde{d} > d$ where $d$ is budget binding. Proposition C.2.2 states that $\boldsymbol{x}_{\widetilde{d}} = (1\ldots1)$. Hence

$$\rho = \mathsf{rev}(d) = d\sum_{n\in[N]} g_n x_{d,n} \leq d\sum_{n\in[N]} g_n x_{\widetilde{d},n} < \widetilde{d}\sum_{n\in[N]} g_n x_{\widetilde{d},n} \overset{(i)}{<} \rho\,,$$

where (i) follows from the definition that $\widetilde{d}$ is non-binding. Hence we obtain a contradiction, and $\widetilde{d}$ cannot be non-binding. This means $\widetilde{d}$ must be budget or ROI binding.

*(3).* Here we show that if some price $d \in \mathcal{D}$ is ROI binding so that $\sum_{n\in[N]}(V_n - \gamma d)g_n x_{d,n} = 0$, any price $\widetilde{d} > d$ must also be ROI binding. We first claim that $\boldsymbol{x}_{\widetilde{d}} \preceq \boldsymbol{x}_d$. To show this, we use a contradiction argument by assuming $\boldsymbol{x}_{\widetilde{d}} \succeq \boldsymbol{x}_d$.

Let the threshold vector $\boldsymbol{x}_d$ be characterized by $\boldsymbol{x}_d = \psi(J, q)$ (see definition of threshold vectors in Definition C.1.1). Under Assumption 4.3.1, we note that $\boldsymbol{x}_d$ cannot have all 0 entries and hence $x_{d,1} > 0$. However, since $\sum_{n\in[N]}(V_n - \gamma d)g_n x_{d,n} = 0$, it must be the case that $V_{J+1} - \gamma d < 0$. Now, applying the ordering property for threshold vectors in the second half of Lemma C.2.3 (ii) by taking $\boldsymbol{Z} = \boldsymbol{x}_{\widetilde{d}}$, $\boldsymbol{Y} = \boldsymbol{x}_d$, and $b_i = V_i - \gamma d$ we have

$$0 = \sum_{n\in[N]} (V_n - \gamma d) g_n x_{d,n} \geq \sum_{n\in[N]} (V_n - \gamma d) g_n x_{\widetilde{d},n} > \sum_{n\in[N]} (V_n - \gamma \widetilde{d}) g_n x_{\widetilde{d},n}\,.$$

In the last inequality we used the fact that $\widetilde{d} > d$. Hence, this contradicts the feasibility of $\boldsymbol{x}_{\widetilde{d}}$, so we conclude that $\boldsymbol{x}_{\widetilde{d}} \preceq \boldsymbol{x}_d$. This further implies

$$\rho \geq \underbrace{d \sum_{n\in[N]} g_n x_{d,n}}_{=\mathsf{rev}(d)} \overset{(i)}{=} \frac{1}{\gamma}\sum_{n\in[N]} V_n g_n x_{d,n} \overset{(ii)}{>} \frac{1}{\gamma}\sum_{n\in[N]} V_n g_n x_{\widetilde{d},n} \overset{(iii)}{\geq} \underbrace{\widetilde{d}\sum_{n\in[N]} g_n x_{\widetilde{d},n}}_{=\mathsf{rev}(\widetilde{d})},$$

where (i) follows from $d$ being ROI binding, i.e. $\sum_{n\in[N]}(V_n-\gamma d)g_n x_{d,n}=0$; (ii) follows from $\boldsymbol{x}_{\widetilde{d}} \preceq \boldsymbol{x}_d$; (iii) follows from feasibility of $\widetilde{d}$ so that $\sum_{n\in[N]}(V_n-\gamma\widetilde{d})g_n x_{\widetilde{d},n} \geq 0$. Therefore, $\rho \geq \mathsf{rev}(d) > \mathsf{rev}(\widetilde{d})$.

Finally, $\rho > \mathsf{rev}(\widetilde{d})$ implies that $\widetilde{d}$ is either non-binding or ROI binding. We note that it is not possible for $\widetilde{d}$ to be non-binding, because $\widetilde{d}$ non-binding implies $\boldsymbol{x}_{\widetilde{d}} = (1\dots1)$ according Proposition C.2.2, contradicting $\boldsymbol{x}_{\widetilde{d}} \preceq \boldsymbol{x}_d$ which we showed earlier. Here we used the fact that $\boldsymbol{x}_d \neq (1\dots1)$ because $\boldsymbol{x}_d$ is ROI binding and Assumption 4.3.1 states for any $d\in\mathcal{D}$, $\sum_{n\in[N]}(V_n-\gamma d)g_n \neq 0$. □

### C.2.4 Additional lemmas for Section 4.3

**Lemma C.2.3** (Ordering property for threshold vectors)**.** *Consider $\{a_i\}_{i\in[N]} \subseteq \mathbb{R}_+^N$ and $\{b_i\}_{i\in[N]} \subseteq \mathbb{R}^N$ where there exists some $j\in[N]$ such that $b_i>0$ for all $i=1\dots j$ and $b_i<0$ for all $i=j+1,\dots m$. Let $\boldsymbol{Z},\boldsymbol{Y}\in[0,1]^N$ be two threshold vectors (see Definition C.1.1) such that $\boldsymbol{Y}=\psi(J_Y,q_Y)$, $\boldsymbol{Z}=\psi(J_Z,q_Z)$, and $\boldsymbol{Z}\succeq\boldsymbol{Y}$. Then the following hold:*

*(i)* $\sum_{i\in[N]} a_i Z_i \geq \sum_{i\in[N]} a_i Y_i$.

*(ii) If* $\sum_{i\in[N]} b_i Z_i \geq 0$ *then* $\sum_{i\in[N]} b_i Y_i \geq 0$*. Furthermore, if* $b_{J_Y+1}<0$*, then* $\sum_{i\in[N]} b_i Y_i \geq \sum_{i\in[N]} b_i Z_i \geq 0$.

*(iii) If* $\sum_{i\in[N]} b_i Y_i < 0$ *then* $\sum_{i\in[N]} b_i Z_i < 0$.

*Proof.*

**(i)** Since $a_i>0$ for all $i\in[N]$, and $\boldsymbol{Z}\succeq\boldsymbol{Y}$ (i.e. $Z_i\geq Y_i$ for all $i\in[N]$), it is easy to see $\sum_{i\in[N]} a_i Z_i \geq \sum_{i\in[N]} a_i Y_i$.

**(ii)** By the definition of threshold vectors, we have $Y_{J_Y+1} = q_Y$ while $Y_i = 0$ for all $i > J_Y + 1$. We prove the claim by contradiction by assuming $\sum_{i\in[N]} b_i Y_i < 0$.

First, it is easy to see $b_{J_Y+1} < 0$. This is because if $b_{J_Y+1} > 0$, then $b_i > 0$ for all $i = 1 \ldots J_Y+1$ by the definition of $\{b_i\}_{i\in[N]}$, and hence $\sum_{i\in[N]} b_i Y_i = \sum_{i\in[J_Y+1]} b_i Y_i \geq 0$ contradicting our assumption that $\sum_{i\in[N]} b_i Y_i < 0$ .

Next, since $\sum_{i\in[N]} b_i Y_i < 0 \leq \sum_{i\in[N]} b_i Z_i$, we have $\sum_{i\in[N]} b_i(Z_i - Y_i) \geq 0$. On the other hand,

$$\sum_{i\in[N]} b_i(Z_i - Y_i) \overset{(i)}{=} \sum_{i=J_Y+1}^{N} b_i(Z_i - Y_i) \overset{(ii)}{<} 0\,.$$

Here, (i) follows from the definition of a threshold vector so that $Y_i = 1$ for all $i = 1 \ldots J_Y$ and also $Z_i = 1$ for all $i = 1 \ldots J_Y$ due to $\boldsymbol{Z} \succeq \boldsymbol{Y}$. (ii) follows from the fact that $b_{J_Y+1} < 0$ so $b_i < 0$ for all $i \geq J_Y + 1$ due to the definition of $\{b_i\}_{i\in[N]}$. Hence, we arrive at a contradiction, which allows us to conclude the first half of the claim, i.e. $\sum_{i\in[N]} b_i Z_i \geq 0$ implies $\sum_{i\in[N]} b_i Y_i \geq 0$.

We now show the second half of the claim i.e. $b_{J_Y+1} < 0$ implies $\sum_{i\in[N]} b_i Y_i \geq \sum_{i\in[N]} b_i Z_i \geq 0$. If $b_{J_Y+1} < 0$, then $b_i < 0$ for all $i = J_Y + 1 + \ldots J_Z + 1$, and hence

$$\sum_{i\in[N]} b_i(Z_i - Y_i) = b_{J_Y+1}(Z_{J_Y+1} - Y_{J_Y+1}) + \sum_{i=J_Y+2}^{J_Z+1} b_i Z_i \overset{(i)}{<} 0\,.$$

Note that in the above inequality the summand $\sum_{i=J_Y+2}^{J_Z+1} b_i Z_i$ does not exist if $J_Y = J_Z$, and in (i) we also used the fact that $Y_i = 0$ for all $i > J_Y + 1$ using the definition of a threshold veector.

**(iii)** We again use a contradiction argument by assuming $\sum_{i\in[N]} b_i Z_i \geq 0$, and the rest of the proof is almost identical to that of (ii) so we will omit it here. □

## C.3 Additional material for Section 4.4

### C.3.1 Proof of Theorem 4.4.1

Define $G := \min_{d,\widetilde{d}\in\mathcal{D}:\mathsf{rev}(d)\neq\mathsf{rev}(\widetilde{d})} \left|\mathsf{rev}(d) - \mathsf{rev}(\widetilde{d})\right|$ to be the minimum revenue gap for all price pairs that do not yield the same revenue, where $\mathsf{rev}(d) := d\sum_{n\in[N]} g_n x_{d,n}$ for any $d \in \mathcal{D}$ is the per-period average seller revenue defined in Equation (4.4). Recall $\widehat{\mathsf{rev}}(D_h) = \frac{D_h}{|\mathcal{E}_h|}\sum_{t\in\mathcal{E}_h} z_t$ is the estimate of $\mathsf{rev}(D_h)$ for episode $h$ with fixed price $D_h$ (see Algorithm 4).

For any exploration episode $\mathcal{E}_h$ whose length is $|\mathcal{E}_h| = T^{1-\xi+\epsilon}$, we have w.p. at least $1 - 1/T$

$$\begin{aligned} &\left|\frac{\widehat{\mathsf{rev}}(D_h)}{D_h} - \frac{\mathsf{rev}(D_h)}{D_h}\right| = \left|\frac{1}{|\mathcal{E}_h|}\sum_{t\in\mathcal{E}_h} z_t - \frac{\mathsf{rev}(D_h)}{D_h}\right| \overset{(i)}{\leq} \frac{\phi(|\mathcal{E}_h|)}{|\mathcal{E}_h|} \overset{(ii)}{\leq} \frac{\phi(T)}{T^{1-\xi+\epsilon}} \\ \overset{(iii)}{\Rightarrow}\ & |\widehat{\mathsf{rev}}(D_h) - \mathsf{rev}(D_h)| \leq \frac{\phi(T)}{T^{1-\xi+\epsilon}} \end{aligned} \tag{C.13}$$

where (i) is due to the definition of $\xi$-adaptive buyers in Definition 4.4.1; (ii) is due to the fact that $\phi$ is an increasing function and the exploration episode lengths are $|\mathcal{E}_h| = T^{1-\xi+\epsilon}$; (iii) is due to the fact that all prices are less than 1.

Since $\phi(T) = \mathcal{O}(T^{1-\xi})$, there exists some $T_\epsilon \in \mathbb{N}$ such that when $T > T_\epsilon$ we have

$$\frac{\phi(T)}{T^{1-\xi+\epsilon}} < \frac{G}{2} \tag{C.14}$$

The rest of the proof relies on the following lemma:

**Lemma C.3.1.** *Assume $T > T_\epsilon$ s.t. Equation (C.14) holds. If $\widehat{\mathsf{rev}}(D_i) \geq \widehat{\mathsf{rev}}(D_j)$ for some exploration episodes $i, j$ s.t. $i \neq j$, then w.p. at least $1 - \frac{1}{T}$, $\mathsf{rev}(D_i) \geq \mathsf{rev}(D_j)$. Furthermore, the following event $\mathcal{G}$*

$$\mathcal{G} = \{\widehat{\mathsf{rev}}(D_i) \geq \widehat{\mathsf{rev}}(D_j) \Longrightarrow \mathsf{rev}(D_i) \geq \mathsf{rev}(D_j) \ \ \textit{for all exploration episodes } i \neq j\} \tag{C.15}$$

*holds with probability at least $1 - \frac{H(H-1)}{2T}$, where $H = \lfloor \log_2(M) \rfloor + 1$ is the maximum* *number of binary search iterations (i.e. number of episodes in the exploration phase).*

*Proof of Lemma C.3.1.* Because $\widehat{\mathsf{rev}}(D_i) \geq \widehat{\mathsf{rev}}(D_j)$, applying Equation (C.13) for episodes $i, j$ yields

$$\begin{aligned} & \mathsf{rev}(D_i) + \frac{\phi(T)}{T^{1-\xi+\epsilon}} \geq \widehat{\mathsf{rev}}(D_i) \geq \widehat{\mathsf{rev}}(D_j) \geq \mathsf{rev}(D_j) - \frac{\phi(T)}{T^{1-\xi+\epsilon}} \qquad \text{(C.16)} \\ \Longrightarrow \ & \frac{2\phi(T)}{T^{1-\xi+\epsilon}} \geq \mathsf{rev}(D_j) - \mathsf{rev}(D_i)\,, \end{aligned}$$

Now, contrary to our claim, suppose that $\mathsf{rev}(D_i) < \mathsf{rev}(D_j)$. We then have

$$\frac{2\phi(T)}{T^{1-\xi+\epsilon}} \geq \mathsf{rev}(D_j) - \mathsf{rev}(D_i) \geq \min_{d,\widetilde{d}\in\mathcal{D}:\mathsf{rev}(d)\neq\mathsf{rev}(\widetilde{d})} \left|\mathsf{rev}(d) - \mathsf{rev}(\widetilde{d})\right| := G\,,$$

which contradicts Equation (C.14) for $T > T_\epsilon$. As there are $H(H-1)/2$ pairs $(i,j)$ such that $i \neq j$, a simple union bound shows event $\mathcal{G}$ holds with probability at least $1 - \frac{H(H-1)}{2T}$. □

We now return to our proof of Theorem 4.4.1. We first show that under event $\mathcal{G}$ (see Equation (C.15)),the final price in the exploitation phase $D_{m^*}$ is revenue-optimal, i.e. $\max_{d\in\mathcal{D}} \mathsf{rev}(d) = \mathsf{rev}\left(D_{m^*}\right)$

We use an induction argument that shows after each iteration of the binary search procedure in the exploration phase of Algorithm 4, $\mathsf{rev}(D_m) \leq \mathsf{rev}(D_{m^*})$ for all $m \leq \mathrm{L}$ and $m \geq \mathrm{R}$. The base case is the first iteration, where we have $\mathrm{L} = 1$, $\mathrm{R} = M$. If $m^* = \mathrm{L} = 1$, then under event $\mathcal{G}$ we get

$$\widehat{\mathsf{rev}}(D_1) \geq \widehat{\mathsf{rev}}(D_M) \overset{(i)}{\Longrightarrow} \mathsf{rev}(D_1) \geq \mathsf{rev}(D_M)\,.$$

Hence after the first iteration $\mathsf{rev}(D_m) \leq \mathsf{rev}(D_{m^*})$ for any $m \leq \mathrm{L}$ and $m \geq \mathrm{R}$. The case for $m^* = \mathrm{R}$ follows from the same argument.

Now assume that the induction hypothesis holds, i.e. at the beginning of some iteration with the tuple $(\mathrm{L}, \mathrm{R}, m^*)$, we have $\mathsf{rev}(D_m) \leq \mathsf{rev}(D_{m^*})$ $m \leq \mathrm{L}$ and $m \geq \mathrm{R}$. According to Algorithm 4, we only need to show two cases in order to validate the induction procedure.

- **Case 1.** If $\hat{\mathsf{rev}}(D_{\text{med}}) < \hat{\mathsf{rev}}(D_{\text{med}+1})$, then we show $\mathsf{rev}(D_m) \le \mathsf{rev}(D_{\text{med}+1})$ for all $m = 1 \ldots \text{med} + 1$

- **Case 2.** If $\hat{\mathsf{rev}}(D_{\text{med}}) \ge \hat{\mathsf{rev}}(D_{\text{med}+1})$, then we show $\mathsf{rev}(D_m) \ge \mathsf{rev}(D_{\text{med}})$ for all $m = \text{med} + 1 \ldots M$

Note that under Case 1., $\text{med} + 1$ will be the new value of $m^*$ in the next iteration (i.e. the next induction step). So by showing $\mathsf{rev}(D_m) \le \mathsf{rev}(D_{\text{med}+1})$ for all $m = 1 \ldots \text{med} + 1$, we validate the induction hypothesis for the next induction step. A similar argument holds for Case 2.

**Case 1.** When $\hat{\mathsf{rev}}(D_{\text{med}}) < \hat{\mathsf{rev}}(D_{\text{med}+1})$, under event $\mathcal{G}$ (see Equation (C.15)) we have $\mathsf{rev}(D_{\text{med}}) \le \mathsf{rev}(D_{\text{med}+1})$. We claim that $D_{\text{med}}$ cannot be an ROI binding price. Assume the contrary that $D_{\text{med}}$ is ROI binding. Then, part (3) of Theorem 4.3.3 states $\mathsf{rev}(D_{\text{med}+1}) < \mathsf{rev}(D_{\text{med}})$, leading to a contradiction. Hence $D_{\text{med}}$ must be either a nonbinding price or a budget binding price. Applying part (1) of Theorem 4.3.3, we can then conclude that for any $m \le \text{med}$, $\mathsf{rev}(D_m) \le \mathsf{rev}(D_{\text{med}})$, so

$$\mathsf{rev}(D_m) \le \mathsf{rev}(D_{\text{med}}) \le \mathsf{rev}(D_{\text{med}+1}) \quad \forall m = 1 \ldots \text{med}\,.$$

At the end of the iteration, as we update $m^{*+} = \text{med} + 1$ (here we denote $m^{*+}$ as the updated value to distinguish from its initial value at the start of the iteration), we have $\mathsf{rev}(D_{m^{*+}}) \ge \mathsf{rev}(D_{\text{med}+1}) \ge \mathsf{rev}(D_{\text{med}}) \ldots \mathsf{rev}(D_1)$. On the other hand, since $\hat{\mathsf{rev}}(D_{m^{*+}}) = \max_{m \in \{m^*, \text{med}+1\}} \hat{\mathsf{rev}}(D_m) \ge \hat{\mathsf{rev}}(D_{m^*})$, event $\mathcal{G}$ implies

$$\mathsf{rev}(D_{m^{*+}}) \ge \mathsf{rev}(D_{m^*}) \overset{(i)}{\ge} \mathsf{rev}(D_m) \quad \forall m = \mathrm{R} \ldots M\,,$$

where (i) follows from the induction hypothesis. Therefore, we have

$$\mathsf{rev}(D_{m^{*+}}) \ge \mathsf{rev}(D_m) \quad \forall m = \mathrm{R} \ldots M \text{ and } m = 1 \ldots \text{med} + 1\,,$$

and by realizing the tuple $(\text{med} + 1, \mathrm{R}, m^{*+})$ is the initial tuple for the next iteration concludes the induction step.

**Case 2.** The case when $\hat{\mathsf{rev}}(D_{\text{med}}) \geq \hat{\mathsf{rev}}(D_{\text{med}+1})$ follows from an identical argument, and we will omit the details. This concludes the induction proof.

The above implies that when the event

$$\mathcal{G} = \{\hat{\mathsf{rev}}(D_i) \geq \hat{\mathsf{rev}}(D_j) \Longrightarrow \mathsf{rev}(D_i) \geq \mathsf{rev}(D_j) \text{ for all } i, j \in [H]\}$$

holds throughout the exploration phase, the above induction argument implies we have $\mathsf{rev}(D_{m^*}) \geq \mathsf{rev}(D_m)$for all $m \in [M]$. Hence $\mathsf{rev}(D_{m^*}) = \max_{d\in\mathcal{D}} \mathsf{rev}(d)$ w.p. at least $1 - \frac{H(H-1)}{2T}$ according to Lemma C.3.1 where $H = \lfloor \log_2(M) \rfloor + 1$.

Furthermore, we point out that in each iteration of the binary search procedure the seller explores at most two prices. Hence the entire exploration phase, which consists of all periods in exploration episodes and we denote as $\mathcal{E}$, has length at most $2E\left(\lfloor \log_2(M) \rfloor + 1\right) = 2T^{1-\xi+\epsilon}\left(\lfloor \log_2(M) \rfloor + 1\right)$ periods. Therefore, the seller's regret can be upper bounded as

$$\begin{aligned}
&\mathrm{Reg}_{\text{sell}} \\
=\ & T \max_{d\in\mathcal{D}} \mathsf{rev}(d) - \sum_{t\in[T]} \mathbb{E}\left[d_t z_t\right] \\
\leq\ & |\mathcal{E}| + \sum_{t=|\mathcal{E}|+1}^{T} \max_{d\in\mathcal{D}} \mathsf{rev}(d) - \mathbb{E}\left[d_t z_t\right] \\
\overset{(i)}{\leq}\ & |\mathcal{E}| + \sum_{t\in[T]/\mathcal{E}} \mathbb{E}\left[(\mathsf{rev}(D_{m^*}) - D_{m^*} z_t)\,\mathbb{I}\left\{\mathcal{G}\right\}\right] + (T - |\mathcal{E}|)\mathbb{P}\left(\mathcal{G}^c\right) \\
\leq\ & |\mathcal{E}| + D_{m^*}(T - |\mathcal{E}|) \cdot \mathbb{E}\left[\frac{\mathsf{rev}(D_{m^*})}{D_{m^*}} - \frac{1}{T - |\mathcal{E}|} \sum_{t\in[T]/\mathcal{E}} z_t\right] + (T - |\mathcal{E}|)\mathbb{P}\left(\mathcal{G}^c\right) \\
\overset{(ii)}{\leq}\ & |\mathcal{E}| + \phi\left(T - |\mathcal{E}|\right) + (T - |\mathcal{E}|) \cdot \mathbb{P}\left(\left|\frac{\mathsf{rev}(D_{m^*})}{D_{m^*}} - \frac{1}{T - |\mathcal{E}|} \sum_{t\in[T]/\mathcal{E}} z_t > \frac{\phi(T - |\mathcal{E}|)}{T - |\mathcal{E}|}\right|\right) \\
&\quad + T\mathbb{P}\left(\mathcal{G}^c\right) \\
\overset{(iii)}{\leq}\ & |\mathcal{E}| + \phi\left(T - |\mathcal{E}|\right) + 1 + T\mathbb{P}\left(\mathcal{G}^c\right) \\
\overset{(iv)}{\leq}\ & 2\left(\lfloor \log_2(M) \rfloor + 1\right) \cdot T^{1-\xi+\epsilon} + \phi\left(T\right) + \left(\lfloor \log_2(M) \rfloor + 1\right)^2 / 2\,.
\end{aligned}$$

In (i) we used the fact that $\max_{d\in\mathcal{D}} \mathsf{rev}(d) = \mathsf{rev}(D_{m^*})$ under event $\mathcal{G}$ and $d_t = D_{m^*}$ for all exploitation periods $t \in [T]/\mathcal{E}$; in (ii) and (iii) we used the definition of $\xi$-adaptive buyer algorithm (see Definition 4.4.1) so that for the exploitation phase $[T]/\mathcal{E}$, the event $\left|\frac{\mathsf{rev}(D_{m^*})}{D_{m^*}} - \frac{1}{T-|\mathcal{E}|}\sum_{t\in[T]/\mathcal{E}} z_t\right| \leq \frac{\phi(T-|\mathcal{E}|)}{T-|\mathcal{E}|}$ holds with probability at least $1-1/T$, and also $\phi$ is an increasing function; In (iv), we used the fact that all periods in exploration episodes $\mathcal{E}$, has length at most $2E\left(\lfloor\log_2(M)\rfloor + 1\right) = 2T^{1-\xi+\epsilon}\left(\lfloor\log_2(M)\rfloor + 1\right)$ periods, and the fact that $\mathbb{P}\left(\mathcal{G}^c\right) \leq \frac{(\lfloor\log_2(M)\rfloor+1)\cdot\lfloor\log_2(M)\rfloor}{2T}$ according to Lemma C.3.1, so $1 + T\mathbb{P}\left(\mathcal{G}^c\right) \leq \left(\lfloor\log_2(M)\rfloor + 1\right)^2/2$ given $M \geq 2$. □

## C.4 Additional material for Section 4.5

### C.4.1 Proof of Lemma 4.5.1

Recall that when the buyer best responds, she adopts the threshold strategy w.r.t $\boldsymbol{x}_d$ where $\boldsymbol{x}_d \in [0,1]^N$ is the optimal solution to $U(d)$ in Equation (4.3); see Definition 4.3.2 for best response. Further, the threshold strategy can be represented as decision

$$z_t^* = \sum_{n\in[N]} x_{d,n}\mathbb{I}\{v_t = V_n\}\,.$$

Then, for any exploration or exploitation episode $\mathcal{E}$ (whose posted price we denote as $d$), for the best response decisions $\{z_t\}_{t\in\mathcal{E}}$ defined above, we have for any $t \in \mathcal{E}$

$$\mathbb{E}[dz_t^*] = d\sum_{n\in[N]} g_n x_{d,n} = \mathsf{rev}(d)$$

where $\mathsf{rev}(d)$ is the per-period expected revenue defined in Equation (4.4). Hence, by defining

$$Y_t = dz_t^* - \mathsf{rev}(d)$$

we know that the sequence $\{Y_t\}_{t\in\mathcal{E}}$ is a martingale difference sequence such that that $|Y_t| \leq d \leq 1$ for all $t \in \mathcal{E}$. By Azuma Hoeffding's inequality (see Lemma C.5.1) we have for any $\delta \in (0,1)$

$$\mathbb{P}\left(\left|d\sum_{t\in\mathcal{E}} z_t^* - |\mathcal{E}|\cdot \mathsf{rev}(d)\right| > \sqrt{2|\mathcal{E}|\log\left(2/\delta\right)}\right) \leq \delta\,.$$

Hence, by taking $\delta = 1/T$ and considering the increasing function $\phi(x) = \sqrt{2x\log\left(2T\right)} = \mathcal{O}(x^{1/2})$, for any exploration/exploitation episode $\mathcal{E}$ (whose price we denote as $d$) we have

$$\left|\frac{d}{|\mathcal{E}|}\sum_{t\in\mathcal{E}} z_t^* - \mathsf{rev}(d)\right| \leq \frac{\phi(|\mathcal{E}|)}{|\mathcal{E}|}$$

with probability (w.p.) at least $1-1/T$. Therefore best responding is $\frac{1}{2}$-adaptive. □

### C.4.2 Proof of Theorem 4.5.2

For the seller, the regret upper bound is a direct result of Lemma 4.5.1 and Theorem 4.4.1.

On the other hand for the buyer, following the exact proof Step 2. in the proof of Proposition 4.3.2 (see Appendix C.2.2), in particular Equations (C.9), (C.10) and (C.11), we know that by best responding, the buyer's budget and ROI constraints are satisfied. Finally, we bound the buyer's regret (see Definition in (4.7)) as followed:

Let $d$ be the posted price in the final exploitation episode (see Algorithm 4). Using an argument similar to Step 1. in the proof of Proposition 4.3.2, for the linear program (LP) $U(d)$ in Equation (4.3) that denotes the buyer's single-period myopic optimization problem, it is easy to see the optimal value is bounded and the LP is feasible (consider the solution with all entries set to be 0). Then, strong duality holds, and there exists corresponding optimal dual variables $(\lambda,\mu)\in\mathbb{R}_+^2$ w.r.t. the exploitation price $d$ s.t.

$$\begin{aligned}U(d) &= \max_{\boldsymbol{x}\in[0,1]^N}\sum_{n\in[N]} g_n\left((1+\lambda)V_n - (\gamma\lambda+\mu)d\right)x_n + \rho\mu \\ &= \sum_{n\in[N]} g_n\left((1+\lambda)V_n - (\gamma\lambda+\mu)d\right)_+ + \rho\mu\end{aligned} \tag{C.17}$$

Similar to Equation (C.8), by denoting $\mathcal{E}$ to be all periods within exploration episodes, the buyer's hindsight objective can be bounded as

$$
\begin{aligned}
\mathrm{OPT}(\boldsymbol{d}_{1:T}) &\leq \max_{\boldsymbol{z}\in[0,1]^T} \sum_{t\in[T]} \mathbb{E}\left[((1+\lambda)v_t - (\gamma\lambda+\mu)d)\, z_t\right] + T\rho\mu \\
&\leq \sum_{t\in[T]} \mathbb{E}\left[((1+\lambda)v_t - (\gamma\lambda+\mu)d)_+\right] + T\rho\mu \\
&= \sum_{t\in[T]} \left( \sum_{n\in[N]} g_n\left((1+\lambda)V_n - (\gamma\lambda+\mu)d\right)_+ + \rho\mu \right) \\
&\leq (1+\lambda+\rho\mu)\cdot|\mathcal{E}| + \sum_{t\in[T]/\mathcal{E}} \left( \sum_{n\in[N]} g_n\left((1+\lambda)V_n - (\gamma\lambda+\mu)d\right)_+ + \rho\mu \right) \\
&\overset{(i)}{=} \Theta(T^{\frac{1}{2}+\epsilon}) + (T - |\mathcal{E}|)U(d)\,.
\end{aligned}
\tag{C.18}
$$

Here (i) follows from Equation (C.17) and the fact that there are at most $2\left(\lfloor \log_2(M)\rfloor + 1\right)$ exploration episodes, which implies in $\mathcal{E}$ there are at most $2T^{1-\xi+\epsilon}\left(\lfloor \log_2(M)\rfloor + 1\right) = \Theta(T^{\frac{1}{2}+\epsilon})$ periods. The buyer's regret can be thus bounded as followed

$$
\begin{aligned}
\mathrm{Reg}_{\mathrm{buy}} &= \mathrm{OPT}(\boldsymbol{d}_{1:T}) - \sum_{t\in[T]} \mathbb{E}\left[v_t z_t\right] \leq \Theta(T^{\frac{1}{2}+\epsilon}) + (T - |\mathcal{E}|)U(d) - \sum_{t\in[T]/\mathcal{E}} \mathbb{E}\left[v_t z_t\right] \\
&= \Theta(T^{\frac{1}{2}+\epsilon})
\end{aligned}
$$

where in the final equality, we used the fact that the buyer's expected utility is exactly $U(d)$ for each exploitation period when best responding as shown in Equation (C.12). $\square$

### C.4.3 Proof of Theorem 4.5.3

This proof consists of two parts, namely bounding seller's regret, and bounding buyer's regret as well the "balance" of buyer's budget and ROI constraints.

**Part 1. Bounding seller's regret.** Here, we only need to show that the buyer's strategy is $\xi$-adaptive (see Definition in 4.4.1), and the rest of the proof follows from Theorem 4.4.1.

For notation convenience, fix some exploration or exploitation episode $\mathcal{E}$, and denote the corresponding price in the episode as $d$. In light of Lemma 4.3.1, we let $\boldsymbol{x}_d \in [0,1]^N$ be the unique optimal threshold vector (see Definition C.1.1) solution to $U(d)$. According to the definition of the per-period seller expected revenue $\mathsf{rev}(d)$ under buyer best response in Equation (4.4), we can further write the seller's per-period expected revenue for episode $h$ as

$$\mathsf{rev}(d) = d \sum_{n\in[N]} x_{d,n} g_n \,. \tag{C.19}$$

Let $\mathcal{F}_t$ be the sigma algebra generated by $\{(v_\tau, d_\tau, z_\tau)\}_{\tau\in[t]}$, which characterizes all randomness in the buyer and seller's behavior up to period $t$. Recall $\hat{\boldsymbol{x}}_t$ is the optimal solution to $U(d_t)$ of Equation (4.3) via replacing the true distribution $\boldsymbol{g} \in \Delta_N$ with the estimate $\hat{\boldsymbol{g}}_t \in \Delta_N$. The buyer adopting a threshold strategy w.r.t. $\hat{\boldsymbol{x}}_t$ implies the buyer's decision to be

$$z_t = \sum_{n\in[N]} \hat{x}_{t,n} \mathbb{I}\{v_t = V_n\} \tag{C.20}$$

Since $\hat{\boldsymbol{x}}_t$ is $\mathcal{F}_{t-1}$-measurable, for $t \in \mathcal{E}$ we have

$$\mathbb{E}\left[z_t \middle| \mathcal{F}_{t-1}\right] = \sum_{n\in[N]} g_n \hat{x}_{t,n}$$

Thus, the by defining

$$Y_t = \sum_{n\in[N]} g_n \hat{x}_{t,n} - z_t \,, \tag{C.21}$$

we know that the sequence $\{Y_t\}_{t\in\mathcal{E}}$ is a martingale difference sequence such that that $|Y_t| \le 1$ for all $t$. By Azuma Hoeffding's inequality (see Lemma C.5.1) we have for any $\delta \in (0,1)$

$$\mathbb{P}\left(\widetilde{\mathcal{G}}\right) \geq 1-\delta \text{ where } \widetilde{\mathcal{G}} := \left\{ \left| \sum_{t\in\mathcal{E}} \left( \sum_{n\in[N]} g_n \hat{x}_{t,n} - z_t \right) \right| \leq \sqrt{2|\mathcal{E}|\log(2/\delta)} \right\} . \quad \text{(C.22)}$$

The remaining proof relies on the following lemma whose proof can be found in Appendix C.4.5

**Lemma C.4.1.** *Fix some price $d$ and define the following problem which is solved by the approximate best response buyer with ML advice to obtain $\hat{\boldsymbol{x}}_t$ (see Definition 4.5.1):*

$$\hat{U}_t(d) = \max_{\boldsymbol{x}\in[0,1]^N} \sum_{n\in[N]} \hat{g}_{t,n} V_n x_n \quad s.t. \sum_{n\in[N]} \hat{g}_{t,n}\left(V_n - \gamma d\right) x_n \geq 0 \quad \textit{and} \quad d \sum_{n\in[N]} \hat{g}_{t,n} x_n \leq \rho \,. \tag{C.23}$$

*Here, recall $\hat{\boldsymbol{g}}_t \in \Delta_N$ is the ML advice obtained in period $t$ which is an estimate for the true value distribution $\boldsymbol{g} \in \Delta_N$. Further, define the following values*

$$\begin{aligned} (A) &= \left(U(d) - \sum_{n\in[N]} g_n V_n \hat{x}_{t,n}\right)_+ , \quad (B) = \left(-\sum_{n\in[N]} g_n \left(V_n - \gamma d\right) \hat{x}_{t,n}\right)_+ \\ (C) &= \left(d \sum_{n\in[N]} g_n \hat{x}_{t,n} - \rho\right)_+ , \end{aligned} \tag{C.24}$$

*where we recall $U(d)$ is defined in Equation (4.3). Then, the values $(B),(C)$ are upper bounded by $\sqrt{N}\|\boldsymbol{g} - \hat{\boldsymbol{g}}_t\|$ for all $t$. Further, because the estimation error $\lim_{t\to\infty} \ell_t = 0$ there exists some $T_0 \in \mathbb{N}$ s.t. $\|\boldsymbol{g} - \hat{\boldsymbol{g}}_t\| \leq \ell_t < \frac{g_1}{2}$ for all $t > T_0$. Then, there exists an absolute constant $C$ that only depends on buyer model primitives $(\boldsymbol{g}, \boldsymbol{V}, \rho, \gamma)$ s.t. the values $(A)$ and $\|\boldsymbol{x}_d - \hat{\boldsymbol{x}}_t\|$ are upper bounded by $C\sqrt{N}\|\boldsymbol{g} - \hat{\boldsymbol{g}}_t\|$ for $t > T_0$, where $\boldsymbol{x}_d$ is the optimal solution to $U(d)$.*

We now show a high probability bound for $\frac{\mathsf{rev}(d)}{d} - \sum_{t\in\mathcal{E}} z_t$. Assume event $\widetilde{\mathcal{G}}$

(Equation (C.22)) holds, then

$$\begin{aligned}
\left|\sum_{t\in\mathcal{E}}\left(\frac{\mathsf{rev}(d)}{d}-z_t\right)\right| &= \left|\sum_{t\in\mathcal{E}}\left(\sum_{n\in[N]}x_{d,n}g_n-z_t\right)\right| \\
&\leq \left|\sum_{t\in\mathcal{E}}\left(\sum_{n\in[N]}\hat{x}_{t,n}g_n-z_t\right)\right| + \sum_{t\in\mathcal{E}}\left|\sum_{n\in[N]}\hat{x}_{t,n}g_n-\sum_{n\in[N]}x_{d,n}g_n\right| \\
&\overset{(i)}{\leq} \sqrt{2|\mathcal{E}|\log(2T)}+\sum_{t\in\mathcal{E}}\|\boldsymbol{x}_d-\hat{\boldsymbol{x}}_t\|\cdot\|\boldsymbol{g}\| \\
&\overset{(ii)}{\leq} \sqrt{2|\mathcal{E}|\log(2T)}+T_0+C\sqrt{N}\sum_{t\in\mathcal{E}:t>T_0}\ell_t \\
&\leq \sqrt{2|\mathcal{E}|\log(2T)}+T_0+C\sqrt{N}\sum_{t\in\mathcal{E}}\ell_t \\
&\overset{(iii)}{\leq} \sqrt{2|\mathcal{E}|\log(2T)}+T_0+C\sqrt{N}\widetilde{\phi}(|\mathcal{E}|) \\
&:= \phi(|\mathcal{E}|)
\end{aligned}$$

where in (i) we plugged in the Azuma-Hoeffding inequality result showed in Equation (C.22) with $\delta=\frac{1}{T}$; in (ii) we applied Lemma C.4.1 and some constant absolute constant $C$ for $t>T_0$ (defined in statement of Lemma C.4.1), and the fact that $\|\boldsymbol{g}\|\leq 1$ since $\boldsymbol{g}$ is a probability simplex; in (iii) we used the assumption that there exists some increasing function $\widetilde{\phi}$ s.t. $\sum_{t\in\mathcal{E}}\ell_t\leq\widetilde{\phi}(|\mathcal{E}|)$. Therefore w.p. at least $1-1/T$ (since $\widetilde{\mathcal{G}}$ holds w.p. at least $1-1/T$ when $\delta=1/T$), we have

$$\left|\frac{d}{|\mathcal{E}|}\sum_{t\in\mathcal{E}}z_t-\mathsf{rev}(d)\right|\leq\frac{\phi(|\mathcal{E}|)}{|\mathcal{E}|}$$

Since $\widetilde{\phi}(x)\leq\mathcal{O}(x^{1-L})$, we know that $\phi(x)=\mathcal{O}(x^{1-\xi})$ for $\xi=\min\{\frac{1}{2},L\}$. Hence, for large enough $T$ s.t. the exploration episode length $E=T^{1-\xi+\epsilon}>T_0$, the buyer's approximate best responding with ML advice is $1-\xi$-adaptive for $\xi=\min\{\frac{1}{2},L\}$.

**Part 2. Bounds for the buyer.** We first follow a similar approach as the proof of Theorem 4.5.2 to upper bound the buyer regret.

Let $d$ be the posted price in the final exploitation episode (see Algorithm 4), and denote $\mathcal{E} = \Theta(T^{1-\xi+\epsilon})$ as all periods within exploration episodes. Then using the same arguments as in Equations (C.17) and (C.18), we can show the buyer's hindsight objective can be bounded as

$$\text{OPT}(\boldsymbol{d}_{1:T}) \;\leq\; \Theta(T^{\xi+\epsilon}) + (T - |\mathcal{E}|)U(d)\,.$$

Since the buyer approximately best responds w.r.t. $\hat{\boldsymbol{x}}_t$ which is the optimal solution to the problem $\hat{U}_t(d)$ Equation (C.23), recall the buyer's decision $z_t$ can be written as in Equation (C.20):

$$z_t = \sum_{n\in[N]} \hat{x}_{t,n}\mathbb{I}\{v_t = V_n\}$$

Hence, $\mathbb{E}[v_t z_t|\mathcal{F}_{t-1}] = \sum_{n\in[N]} g_n V_n \hat{x}_{t,n}$. Let $C$ and $T_0$ be defined as in Lemma C.4.1, and thus the buyer's regret can be thus bounded as followed

$$\begin{aligned}
\text{Reg}_{\text{buy}} \;&=\; \text{OPT}(\boldsymbol{d}_{1:T}) - \sum_{t\in[T]} \mathbb{E}\left[v_t z_t\right] \\
&\leq\; \Theta(T^{\xi+\epsilon}) + \sum_{t\in[T]/\mathcal{E}} \mathbb{E}\left[(U(d) - g_n\hat{x}_{t,n})\right] \\
&\leq\; \Theta(T^{\xi+\epsilon}) + T_0 + \sum_{t\in[T]/\mathcal{E}:t>T_0} \mathbb{E}\left[(U(d) - g_n\hat{x}_{t,n})\right] \\
&\overset{(i)}{\leq}\; \Theta(T^{\xi+\epsilon}) + T_0 + C\sqrt{N} \sum_{t\in[T]/\mathcal{E}:t>T_0} \|\boldsymbol{g} - \hat{\boldsymbol{g}}_t\| \\
&\overset{(ii)}{\leq}\; \Theta(T^{\xi+\epsilon}) + T_0 + C\sqrt{N} \sum_{t\in[T]/\mathcal{E}} \ell_t \\
&\overset{(iii)}{\leq}\; \Theta(T^{\xi+\epsilon}) + T_0 + C\sqrt{N}\widetilde{\phi}(T - |\mathcal{E}|) \\
&\overset{(iv)}{=}\; \Theta(T^{\xi+\epsilon})\,.
\end{aligned}$$

In (i), we applied Lemma C.4.1 for the value (A) defined in Equation (C.24); (ii) follows from the definition of the estimation errors $\ell_t \geq \|\boldsymbol{g} - \hat{\boldsymbol{g}}_t\|$; (iii) follows from the assumption that for any exploration or exploitation episode $\mathcal{E}_h$, the total error $\sum_{t\in\mathcal{E}_h} \ell_t$ is upper bounded by $\widetilde{\phi}(T - |\mathcal{E}|$ where $\widetilde{\phi}$ is an increasing function; (iv) follows from the fact that $\widetilde{\phi}(x) \le \mathcal{O}(T^{1-L}) \le \mathcal{O}(T^{1-\xi})$.

Now we show the buyer constraint violation is small, namely

$$\frac{1}{T}\mathbb{E}\left[\sum_{t\in[T]}(v_t - \gamma d_t)\, z_t\right] \ge -\Theta(T^{-L}) \quad \text{and} \quad \frac{1}{T}\mathbb{E}\left[\sum_{t\in[T]} d_t z_t\right] \le \rho + \Theta(T^{-L})\,.$$

The proofs for both inequalities are very similar, so here we just show $\frac{1}{T}\mathbb{E}\left[\sum_{t\in[T]}(v_t - \gamma d_t)\, z_t\right] \ge -\Theta(T^{-L})$. Similar to the above where we bounded buyer's regret, we have $\mathbb{E}[(v_t - \gamma d_t)\, z_t | \mathcal{F}_{t-1}] = \sum_{n\in[N]} g_n \,(V_n - \gamma d)\, \hat{x}_{t,n}$, and thus for all exploration and exploitation episodes $\mathcal{E}_1 \dots \mathcal{E}_H$ (assuming there are $H$ episodes), we have

$$\begin{aligned}
-\mathbb{E}\left[\sum_{t\in[T]}(v_t - \gamma d_t)\, z_t\right] &= \sum_{t\in[T]} \mathbb{E}\left[-\left(\sum_{n\in[N]} g_n\,(V_n - \gamma d)\,\hat{x}_{t,n}\right)\right] \\
&\le \sum_{t\in[T]} \mathbb{E}\left[\left(-\sum_{n\in[N]} g_n\,(V_n - \gamma d)\,\hat{x}_{t,n}\right)_+\right] \\
&\overset{(i)}{\le} \sqrt{N}\sum_{t\in[T]} \ell_t \\
&= \sqrt{N}\sum_{h\in[H]}\sum_{t\in\mathcal{E}_h} \ell_t \\
&\overset{(ii)}{\le} \sqrt{N}\sum_{h\in[H]} \mathcal{O}(|\mathcal{E}_h|^{1-L}) \\
&= \Theta(T^{1-L})
\end{aligned}$$

where (i) follows from the upper bound of (B) (Equation (C.24)) in Lemma C.4.1; (ii) follows from the assumption that for any exploration and exploitation episode $\mathcal{E}_h$ the errors $\{\ell_t\}_t$ satisfy $\sum_{t\in\mathcal{E}_h} \ell_t \le \widetilde{\phi}(|\mathcal{E}_h|)$ for some increasing function $\widetilde{\phi}: \mathbb{R}_+ \to \mathbb{R}^+$ and $\widetilde{\phi}(x) \le \mathcal{O}(x^{1-L})$.

Finally, dividing both sides by $T$ yields the desired bound $\frac{1}{T}\mathbb{E}\left[\sum_{t\in[T]}(v_t - \gamma d_t)\, z_t\right] \ge -\Theta(T^{-L})$. □

### C.4.4 Proof of Theorem 4.5.4

We know that the empirical estimates $\hat{\boldsymbol{g}}_t \in \Delta_N$ for the buyer's value distribution $\boldsymbol{g} \in \Delta_N$ defined in Equation (4.8) follow a multinomial distribution, i.e. $\hat{\boldsymbol{g}}_t \sim \frac{1}{t}\text{Multinomial}(t, \boldsymbol{g})$. Therefore, applying Lemma C.5.3 by taking $\delta = 1/T^2$ , we have w.p. at least $1 - 1/T^2$ the following event holds

$$\mathcal{G}_t := \left\{ \|\hat{\boldsymbol{g}}_t - \boldsymbol{g}\| \leq \ell_t := \sqrt{\frac{2N\log(2T^2)}{t}} \right\} \tag{C.25}$$

Here we used the fact that $\|\boldsymbol{x}\| \leq \|\boldsymbol{x}\|_1$ for any vector $\boldsymbol{x}$. Hence, using a simple union bound, the event $\cup_{t\in[T]}\mathcal{G}_t$ holds w.p. at least $1 - 1/T$. Further, for any exploration or exploitation episode $\mathcal{E}$, we have

$$\sum_{t\in\mathcal{E}} \ell_t \leq \sum_{\tau\in[|\mathcal{E}|]} \sqrt{\frac{2N\log(2T^2)}{\tau}} \leq \widetilde{\phi}(|\mathcal{E}|) \tag{C.26}$$

for some increasing function $\widetilde{\phi}$ s.t. $\widetilde{\phi}(x) \leq \mathcal{O}(x^{\frac{1}{2}})$. Hence, w.p. at least $1 - 1/T$, the estimation errors $\{\ell_t\}_t$ defined above satisfy the conditions in Theorem 4.5.3 for large enough $T$, i.e. $\lim_{t\to\infty} \ell_t = 0$ and $\sum_{t\in\mathcal{E}} \ell_t \leq \widetilde{\phi}(|\mathcal{E}|)$ for any any exploration or exploitation episode $\mathcal{E}$ where increasing function $\widetilde{\phi} : \mathbb{R}_+ \to \mathbb{R}^+$ and $\widetilde{\phi}(x) \leq \mathcal{O}(x^{1-L})$. The rest of the proof directly follows from Theorem 4.5.3.

□

### C.4.5 Proof of Lemma C.4.1

Consider the region

$$\mathcal{C} = \left\{ \boldsymbol{x} \in [0,1]^N : -\sum_{n\in[N]} g_n V_n x_n \leq -U(d), -\sum_{n\in[N]} g_n (V_n - \gamma d)\, x_n \leq 0, d \sum_{n\in[N]} g_n x_n \leq \rho \right\} \tag{C.27}$$

By Lemma 4.3.1, we know that $\boldsymbol{x}_d$ is the unique optimal solution to $U(d)$, and hence $\mathcal{C}$ consists of the single point $\boldsymbol{x}_d$, namely $\mathcal{C} = \{\boldsymbol{x}_d\}$. Now consider the optimal solution

$\hat{\boldsymbol{x}}_t \in [0,1]^N$ to $\hat{U}_t(d)$ in Equation (C.23), by the Hoffman bound (Lemma C.5.2), there exists some constant $H > 0$ that only depends on $(\boldsymbol{g}, \boldsymbol{V})$ s.t.

$$\begin{aligned}
&\|\hat{\boldsymbol{x}}_t - \boldsymbol{x}_d\| \\
\leq\; & H\left(\underbrace{\left(U(d) - \sum_{n\in[N]} g_n V_n \hat{x}_{t,n}\right)_+}_{(A)} + \underbrace{\left(-\sum_{n\in[N]} g_n\left(V_n - \gamma d\right)\hat{x}_{t,n}\right)_+}_{(B)} + \underbrace{\left(d\sum_{n\in[N]} g_n \hat{x}_{t,n} - \rho\right)_+}_{C}\right)
\end{aligned} \tag{C.28}$$

where we used the inequality $\|(\boldsymbol{y})_+\| \leq \sum_{n\in[N]}(y_n)_+$ for any vector $y \in \mathbb{R}^N$. We now bound (A), (B) and (C) respectively.

**Bounding (A).** Similar to the proof of Theorem 4.5.2, strong duality holds for the LP $\hat{U}_t(d)$, and hence there exists optimal dual variables $\hat{\lambda}, \hat{\mu} \in \mathbb{R}_+$ s.t.

$$\hat{U}_t(d) = \sum_{n\in[N]} g_n V_n \hat{x}_{t,n} = \max_{\boldsymbol{x}\in[0,1]^N} \sum_{n\in[N]} \hat{g}_{t,n}\left((1+\hat{\lambda})V_n - (\gamma\hat{\lambda}+\hat{\mu})d\right)x_n + \rho\hat{\mu} \tag{C.29}$$

Since $\hat{U}_t(d) \leq 1$, it is easy to see $\hat{\mu} \in [0, 1/\rho]$, and further by considering $\boldsymbol{x} = (1, 0\ldots 0) \in \mathbb{R}^N$, we have

$$\begin{aligned}
1 \geq \hat{U}_t(d) \;&\geq\; \hat{g}_{t,1}\left((1+\hat{\lambda})V_1 - (\gamma\hat{\lambda}+\hat{\mu})d\right) \\
\overset{(i)}{\geq}\; \hat{g}_{t,1}\hat{\lambda}(V_1-\gamma d) - \hat{\mu}d \;&\overset{(ii)}{\geq}\; \frac{g_1}{2}\cdot\hat{\lambda}(V_1-\gamma d) - \frac{1}{\rho} \\
\overset{(iii)}{\Rightarrow}\; \hat{\lambda} \leq 2\left(1+\frac{1}{\rho}\right)\frac{V_1-\gamma D_1}{g_1}&,
\end{aligned} \tag{C.30}$$

where in (i) we used the fact that $\hat{g}_{t,1} \in [0,1]$; in (ii) we used the fact that $|\hat{g}_{t,1} - g_1| \leq \|\hat{\boldsymbol{g}}_t - \boldsymbol{g}\| \leq \ell_t < \frac{g_1}{2}$ for all $t > T_0$, and also $d \in [0,1]$ as well as $\hat{\mu} \in [0, 1/\rho]$; in (iii), we used Assumption 4.3.1 s.t. $V_1 - \gamma d > 0$ for all $d \in \mathcal{D}$, and $g_1 > 0$.

On the other hand, we have

$$
\begin{aligned}
U(d) \;&\leq\; \max_{\boldsymbol{x}\in[0,1]^N} \sum_{n\in[N]} g_n \left((1+\hat{\lambda})V_n - (\gamma\hat{\lambda}+\hat{\mu})d\right) x_n + \rho\hat{\mu} \\
&\overset{(i)}{\leq}\; \max_{\boldsymbol{x}\in[0,1]^N} \sum_{n\in[N]} \hat{g}_{t,n} \left((1+\hat{\lambda})V_n - (\gamma\hat{\lambda}+\hat{\mu})d\right) x_n + (1+\hat{\lambda}) \sum_{n\in[N]} |\hat{g}_{t,n} - g_n| + \rho\hat{\mu} \\
&\overset{(ii)}{\leq}\; \max_{\boldsymbol{x}\in[0,1]^N} \sum_{n\in[N]} \hat{g}_{t,n} \left((1+\hat{\lambda})V_n - (\gamma\hat{\lambda}+\hat{\mu})d\right) x_n + \rho\hat{\mu} + (1+\hat{\lambda})\sqrt{N}\|\hat{\boldsymbol{g}}_t - \boldsymbol{g}\| \\
&\overset{(iii)}{=}\; \hat{U}_t(d) + 2\left(1+\frac{1}{\rho}\right)\frac{V_1 - \gamma D_1}{g_1} \cdot \sqrt{N}\|\hat{\boldsymbol{g}}_t - \boldsymbol{g}\|
\end{aligned}
\tag{C.31}
$$

In (i), we used the fact that for all $n \in [N]$, $x_n \in [0,1]$ and $(1+\hat{\lambda})V_n - (\gamma\hat{\lambda}+\hat{\mu})d \leq (1+\hat{\lambda})V_n \leq 1+\hat{\lambda}$ since all possible values $V_n \in [0,1]$; (ii) applies Cauchy–Schwarz inequality; (iii) plugs in Equation (C.29) and (C.30).

Therefore, if $\hat{U}_t(d) = \sum_{n\in[N]} g_n V_n \hat{x}_{t,n} \geq U(d)$, then (A) $= 0$, whereas if $\hat{U}_t(d) = \sum_{n\in[N]} g_n V_n \hat{x}_{t,n} < U(d)$, Equation (C.31) implies

$$
(A) \leq 2\left(1+\frac{1}{\rho}\right)\frac{V_1 - \gamma D_1}{g_1}\sqrt{N}\|\hat{\boldsymbol{g}}_t - \boldsymbol{g}\| \tag{C.32}
$$

**Bounding (B) and (C).** The bounds for (B) and (C) are similar, and therefore we only show that for (B).

$$
\begin{aligned}
(B) \;&=\; \left(-\sum_{n\in[N]} g_n \left(V_n - \gamma d\right) \hat{x}_{t,n}\right)_+ \\
&\overset{(i)}{\leq}\; \left(-\sum_{n\in[N]} \hat{g}_{t,n} \left(V_n - \gamma d\right) \hat{x}_{t,n}\right)_+ + \left|\sum_{n\in[N]} \left(\hat{g}_{t,n} - g_n\right)\left(V_n - \gamma d\right) \hat{x}_{t,n}\right| \\
&\overset{(ii)}{\leq}\; \sum_{n\in[N]} |\hat{g}_{t,n} - g_n| \\
&\overset{(iii)}{\leq}\; \sqrt{N}\|\hat{\boldsymbol{g}}_t - \boldsymbol{g}\|
\end{aligned}
\tag{C.33}
$$

Here, (i) follows from the basic inequality sequence $(a+b)_+ \leq (a)_+ + (b)_+ \leq (a)_+ + |b|$; (ii) follows from the fact that $\hat{\boldsymbol{x}}_t$ is feasible to $\hat{U}_t(d)$ so that $\sum_{n\in[N]} \hat{g}_{t,n} (V_n - \gamma d) \hat{x}_{t,n} \geq 0$, and also $|V_n - \gamma d| \leq V_n \leq 1$ and $\hat{x}_{t,n} \in [0,1]$; (iii) follows from the Cauchy–Schwarz inequality.

We can similarly show

$$(C) \leq \sqrt{N} \|\hat{\boldsymbol{g}}_t - \boldsymbol{g}\| \tag{C.34}$$

Finally, combining Equations (C.28), (C.32), (C.33), and (C.34) yields the desired result. □

## C.5 Supplementary lemmas

**Lemma C.5.1** (Azuma–Hoeffding inequality). *beLet $Y_1 \dots Y_n$ be a martingale difference sequence with a uniform bound $|Y_j| \leq 1$ for all $j \in [n]$. Then for any $\delta \in (0, 1/e)$,*

$$\mathbb{P}\left(\left|\sum_{j\in[n]} Y_j\right| > \sqrt{2n\log(2/\delta)}\right) \leq \delta\,.$$

**Lemma C.5.2** (Hoffman bound [68]). *Consider a non-empty linear region $\mathcal{C} = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\}$ for some $\boldsymbol{b} \in \mathbb{R}^n$ and $\boldsymbol{A} \in \mathbb{R}^{m\times n}$. Then, there exists some constant $H > 0$ that only depends on $\boldsymbol{A}$ s.t. for any $\boldsymbol{y} \in \mathbb{R}^n$ we have $\inf_{\boldsymbol{z}\in\mathcal{C}} \|\boldsymbol{z} - \boldsymbol{y}\| \leq H\|(\boldsymbol{A}\boldsymbol{y} - \boldsymbol{b})_+\|$. Here $(\boldsymbol{y})_+$ is the vector that takes the positive parts for each entry in $\boldsymbol{y}$, i.e. $(\boldsymbol{y})_+ = ((y_1)_+ \dots (y_n)_+)$.*

**Lemma C.5.3** (Empirical distribution concentration inequality [110]). *Let $\boldsymbol{g} \in \Delta_N$ be a $N$-dimensional probability simplex ($N \geq 2$), and $\hat{\boldsymbol{g}}_t \sim \frac{1}{t}$Multinomial$(t, \boldsymbol{g})$. Then for any $\delta \in (0,1)$, we have*

$$\mathbb{P}\left(\|\hat{\boldsymbol{g}}_t - \boldsymbol{g}\|_1 > \sqrt{\frac{2N\log(2/\delta)}{t}}\right) \leq \delta$$

See also [98] Proposition 2. for a similar statement to that of Lemma C.5.3.

# Appendix D

# Supplementary material for Chapter 5

## D.1 Additional material for Section 5.2

**Example D.1.1** (Personalized-reserve augmented VCG, GSP, GFP auctions)**.** *Consider $M \geq 2$ parallel position auctions $(\mathcal{A}_j)_{j\in[M]}$ all of which take the form of VCG, GSP or GFP auctions. Each auction $\mathcal{A}_j$ is associated with $L_j \geq 1$ slots and CTRs $\boldsymbol{\mu}_j = (\mu_j(\ell))_{\ell\in L_j}$. Assume $N$ bidders submit bid profile $\boldsymbol{b}_j \in \mathbb{R}_+^N$ to auction $\mathcal{A}_j$, where $\widetilde{N}_j \leq N$ are cleared, i.e. greater than respective personalized reserve prices. Define $\widetilde{\boldsymbol{b}}_j \in \mathbb{R}_+^{\widetilde{N}_j}$ to be all "cleared bids", and let $\widetilde{b}_j^{(\ell)}$ be the $\ell$th highest cleared bid. Then, in $\mathcal{A}_j$ bidders who cleared their reserves are assigned slots according to the ranking of $\widetilde{\boldsymbol{b}}_j$, whereas the bidders who do not clear their reserves never get allocated any slots. The payment for a bidder $i$ who cleared her reserve and allocated slot $\ell_{i,j} \in [\min\{\widetilde{N}_j, L_j\}]$ is*

- *VCG: $p_{i,j} = \sum_{\ell=\ell_{i,j}}^{\min\{\widetilde{N}_j,L_j\}}(\mu_j(\ell) - \mu_j(\ell+1)) \cdot \max\{\widetilde{b}_j^{(\ell+1)}, r_{i,j}\}$ where $\widetilde{b}_j^{(\ell)} = 0$ when $\ell > \widetilde{N}_j$.*
- *GSP: $p_{i,j} = \mu_j(\ell_{i,j}) \cdot \max\{\widetilde{b}_j^{(\ell_{i,j}+1)}, r_{i,j}\}$.*
- *GFP: $p_{i,j} = \mu_j(\ell_{i,j}) \cdot \max\{\widetilde{b}_j^{(\ell_{i,j})}, r_{i,j}\}$.*

It is well known that for the same bid profile $\boldsymbol{b}$ and for any bidder $i$, the payment under the GFP auction is greater than equal to that under GSP auction, and the payment under the GSP auction is greater than equal to that under VCG; see e.g. [46].

## D.2 Additional material for Section 5.4

### D.2.1 Proof for Theorem 5.4.3

**Theorem D.2.1** (Restatement of Theorem 5.4.3)**.** *Consider 2 bidders competing in three SPA auctions whose values are indicated in the following table for any* $\beta \in (0,1)$ *and* $y \geq 0$.

| | *Auction 1* | *Auction 2* | *Auction 3 .* |
|---|---|---|---|
| *bidder 1* | $y$ | $v$ | *0* |
| *bidder 2* | *0* | $v-\epsilon$ | $\gamma + \frac{1}{1-\beta}\cdot\epsilon$ |

*Bidder 1's multiplier is fixed to be* $\alpha_1 > 1$*, and consider* $v = \frac{1-\beta}{\alpha_1 - 1}\cdot\gamma$ *for any* $\gamma > 0$*. The reserve prices are set to be* $r_{i,j} = \beta v_{i,j}$*. Then, we have*

$$\min_{\boldsymbol{b}\in\mathcal{F}} \frac{W_1(\mathcal{X}(\boldsymbol{b}))}{\mathrm{OPT}_1} \;=\; 1 - \frac{1-\beta}{\alpha_1 - 1}\cdot\frac{\mathrm{OPT}_{-1} - \frac{1}{1-\beta}\cdot\epsilon}{\mathrm{OPT}_1} \tag{D.1}$$

*Taking* $\epsilon \to 0$ *shows that bidder 1's welfare is equal to the individual welfare guarantee in Theorem 5.4.1.*

**Remark D.2.1.** *We remark that as* $\epsilon \to 0$, $\frac{\mathrm{OPT}_{-i}}{\mathrm{OPT}_i} = \frac{\frac{\alpha_1-1}{1-\beta}v}{y+v} \in \left[0, \frac{\alpha_1-1}{1-\beta}\right]$*, so by varying* $y \in [0,\infty)$*, the above example demonstrates our individual welfare lower bound in Theorem 5.4.1 is tight for any nontrivial market share ratio* $\frac{\mathrm{OPT}_i}{\mathrm{OPT}_{-i}} \in \left[\frac{\alpha_1-1}{1-\beta}, \infty\right)$.

*Proof.* Note that in any feasible outcome, bidder 1 must win auction 1, and bidder 2 must win auction 3. Hence for auction 2, we only need to consider the following outcome:

**Bidder 1 loses auction 2, and suffers welfare loss $v$.** This outcome can be achieved by setting $\alpha_2$ such that $\alpha_2(v-\epsilon) > \alpha_1 v$. Bidder 2 accumulates value $v+\gamma+\left(\frac{1}{1-\beta}-1\right)\epsilon$. Her payment for auction 2 is $\max\{\alpha_1 v, \beta(v-\epsilon)\}$, and for auction 3 is $\beta\left(\gamma+\frac{1}{1-\beta}\cdot\epsilon\right)$. The following shows that her ROAS constraint is satisfied:

$$
\begin{aligned}
& v+\gamma+\left(\frac{1}{1-\beta}-1\right)\epsilon-\max\{\alpha_1 v, \beta(v-\epsilon)\}-\beta\left(\gamma+\frac{1}{1-\beta}\cdot\epsilon\right) \\
\stackrel{(a)}{=}\; & v+\gamma+\left(\frac{1}{1-\beta}-1\right)\epsilon-\alpha_1 v-\beta\left(\gamma+\frac{1}{1-\beta}\cdot\epsilon\right) \\
=\; & (1-\alpha_1)\,v+(1-\beta)\gamma+\left(\frac{1}{1-\beta}-1-\beta\cdot\frac{1}{1-\beta}\right)\epsilon \\
=\; & 0\,,
\end{aligned}
$$

where in (a) we used the fact $\beta \leq 1 < \alpha_1$ and $\epsilon \to 0$. In the final equality we used the definition that $v = \frac{1-\beta}{\alpha_1-1}\cdot\gamma$. On the other hand, bidder 1's ROAS constraint is apparently satisfied.

Under this outcome, denoting bidder 1's welfare as $W_1$ we have

$$
\frac{W_1}{OPT_1} = 1-\frac{v}{OPT_1} = 1-\frac{1-\beta}{\alpha_i-1}\cdot\frac{\gamma}{\mathrm{OPT}_i} = 1-\frac{1-\beta}{\alpha_i-1}\cdot\frac{\mathrm{OPT}_{-i}-\frac{1}{\beta}\cdot\epsilon}{\mathrm{OPT}_i}
$$

□

### D.2.2 Proof of Proposition 5.4.4

*Proof.* For simplicity, denote $\delta_{i,j} = \mathrm{OPT}_{i,j} - W_{i,j}(\boldsymbol{x})$. Then,

$$
\begin{aligned}
\mathrm{OPT}_i - W_i(\boldsymbol{x}) \;&=\; \sum_{j\in[M]:\delta_{i,j}>0}\delta_{i,j}+\sum_{j\in[M]:\delta_{i,j}=0}\delta_{i,j}+\sum_{j\in[M]:\delta_{i,j}<0}\delta_{i,j} \\
&=\; \textsc{loss}_i(\boldsymbol{x})+\sum_{j\in[M]:W_{i,j}(\boldsymbol{x})>\mathrm{OPT}_{i,j}}\left(\mathrm{OPT}_{i,j}-W_{i,j}(\boldsymbol{x})\right) \leq \textsc{loss}_i(\boldsymbol{x}) \leq B\,.
\end{aligned}
$$

Rearranging and dividing both sides by $\mathrm{OPT}_i$ we get $\frac{W_i(\boldsymbol{x})}{\mathrm{OPT}_i} \geq 1-\frac{B}{\mathrm{OPT}_i}$.

Here we remark that it is possible to have $W_{i,j}(\boldsymbol{x}) > \mathrm{OPT}_{i,j}$ because bidders may overbid, and therefore win auctions/slots that they would not have won under the efficient outcome. □

### D.2.3 Proof of Lemma 5.4.7

*Proof.* Recall there is $\Delta$-separation in values. Fix a bidder $i$ and let $v_j^+$ be the smallest competitor value that is strictly greater than $v_{i,j}$ in any auction $\mathcal{A}_j$ where bidder $i$'s value is not the largest, and by definition of $\Delta$-separated values we have $v_j^+ \geq \Delta v_{i,j}$. Hence, by using any multiplier $\alpha_i \in [1, \Delta)$ and assuming competitors bid truthfully, the outcome of the auctions would be identical to that of everyone (including bidder $i$) bidding truthfully. And since truthful bidding is always feasible, we conclude that $\boldsymbol{v}_{-i} \in \mathcal{F}^u_{-i}(\alpha_i \boldsymbol{v}_i)$ for $\alpha_i \in [1, \Delta)$. □

## D.3 Additional material for Section 5.5

### D.3.1 Additional Definitions and Lemmas for Section 5.5

The following lemma shows that for anonymous and truthful auctions, the probability of the lowest bidder winning a single auction is capped by a bound that decreases as the number of bidders grow.

**Lemma D.3.1** (Lemma 3 in [89])**.** *In an anonymous and truthful auction for a single item with $N$ bidders, the bidder who submits the lowest bid wins the item with probability at most $\frac{1}{N}$.*

The following technical definition and lemma (i.e. Definition D.3.1 and Lemma D.3.2) concerns the scenario where only one bidder participates in the auction (others bid 0), and present an upper bound on the probability and cost respectively for the single bidder to win the auction.

**Definition D.3.1** (Single bidder purchase probability and bid threshold)**.** *For any allocation-anonymous and truthful auction $\mathcal{A}$, consider the setting with a single bidder* *who submits bid $b > 0$ and define*

$$\pi_{\mathcal{A}} = \lim_{b \to \infty} \mathbb{P}(\textit{bidder wins item with bid } b)\,, \tag{D.2}$$

*where the limit exists because in a truthful auction, $\mathbb{P}$(bidder wins item with bid $b$) increases in $b$ (see Definition 5.2.1 for truthful auctions). Assume this max probability is reached at some bid threshold $Q_{\mathcal{A}}$, i.e.*

$$Q_{\mathcal{A}} = \min\left\{b > 0 : \mathbb{P}(\textit{bidder wins item with bid } b) = \pi_{\mathcal{A}}\right\}\,. \tag{D.3}$$

Note that in a deterministic single-slot auction that allocates to the highest bidder, $\pi_{\mathcal{A}} = 1$, and $Q_{\mathcal{A}} \to 0$. For example, in an SPA with no reserve, the single bidder can win the auction with any arbitrarily small positive bid with probability 1.

**Lemma D.3.2** (Lemma 4 in [89])**.** *For any allocation-anonymous and truthful auction $\mathcal{A}$ with single-bidder purchase probability $\pi_{\mathcal{A}}$ and bid threshold $Q_{\mathcal{A}}$, the expected cost for a single bidder for winning the item is at most $\pi_{\mathcal{A}} \cdot Q_{\mathcal{A}}$.*

### D.3.2 Proof of Theorem 5.5.1

**Theorem D.3.3** (Restatement of Theorem 5.5.1)**.** *For any auction $\mathcal{A}$ that is allocation-anonymous, truthful, and possibly randomized,* [1] *consider an autobidding problem instance w.r.t. $\mathcal{A}$ with $M = 2K + 1$ auctions and $N = K + 1$ bidders. Fix bidder 0's bid multiplier to be $\alpha_0$ and some $\beta \in [0, 1)$. Consider the bidder values $\{v_{i,j}\}_{i\in[N], j\in[M]}$ given in the following table.*

[1] Here, we assume all auctions of interest are individually rational (IR), i.e. the payment of a bidder is always less than her submitted bid.

| | $A_1$ | $A_2$ | $\dots$ | $A_K$ | $A_{K+1}$ | $A_{K+2}$ | $\dots$ | $A_{2K}$ | $A_{2K+1}$ |
|---|---|---|---|---|---|---|---|---|---|
| $B_1$ | $\frac{\alpha_0 v+\epsilon}{\rho}$ | $\frac{\alpha_0 v+2\epsilon}{\rho}$ | $\dots$ | $\frac{\alpha_0 v+K\epsilon}{\rho}$ | $\gamma$ | 0 | $\dots$ | 0 | 0 |
| $B_2$ | $\frac{\alpha_0 v+2\epsilon}{\rho}$ | $\frac{\alpha_0 v+3\epsilon}{\rho}$ | $\dots$ | $\frac{\alpha_0 v+\epsilon}{\rho}$ | 0 | $\gamma$ | $\dots$ | 0 | 0 |
| $\vdots$ | $\vdots$ | $\vdots$ | | $\vdots$ | $\vdots$ | $\vdots$ | | $\vdots$ | $\vdots$ |
| $B_K$ | $\frac{\alpha_0 v+K\epsilon}{\rho}$ | $\frac{\alpha_0 v+\epsilon}{\rho}$ | $\dots$ | $\frac{\alpha_0 v+(K-1)\epsilon}{\rho}$ | 0 | 0 | $\dots$ | $\gamma$ | 0 |
| $B_0$ | $v$ | $v$ | $\dots$ | $v$ | 0 | 0 | $\dots$ | 0 | $y$ |

*In the table, we let* $\gamma > \frac{Q_{\mathcal{A}}}{\beta} > Q_{\mathcal{A}}$, $\epsilon = O(1/K^3)$ *and* $v = \frac{1-\beta}{\alpha_0 - 1} \cdot \pi_{\mathcal{A}} \cdot \gamma$. *Let* $\rho, y$ *and a large enough* $K$ *satisfy the following:*

$$\alpha_0 < \rho < \frac{\alpha_0}{\beta} \text{ s.t. } \frac{\alpha_0 v + K\epsilon}{\rho} < v, \qquad \text{and } y > \max\left\{\frac{Q_{\mathcal{A}}}{\alpha_0}, \frac{\alpha_0 v}{\pi_{\mathcal{A}}}\right\}, \tag{D.4}$$

*where* $Q_{\mathcal{A}}, \pi_{\mathcal{A}}$ *are defined in Definition D.3.1. Further, suppose the platform enforces personalized reference prices* $\boldsymbol{r} \in \mathbb{R}_+^{N\times M}$ *on top of auction* $\mathcal{A}$*, where* $r_{i,j} = \beta v_{i,j}$*. Then, letting the (possibly random) outcome be* $\boldsymbol{x}$ *when bidders 1, ... K all adopt the bid multiplier* $\rho$*, the ROAS constraints for all bidders are satisfied when* $K \to \infty$ *and* $\rho \to \alpha_0$*, and for bidder* 0 *we have*

$$\lim_{K\to\infty} \frac{\mathbb{E}_{\mathcal{A}}[W_0(\boldsymbol{x})]}{\mathbb{E}_{\mathcal{A}}[\mathrm{OPT}_0]} \leq 1 - \frac{1-\beta}{\alpha_0 - 1} \cdot \lim_{K\to\infty} \frac{\mathbb{E}_{\mathcal{A}}[\mathrm{OPT}_{-0}]}{\mathbb{E}_{\mathcal{A}}[\mathrm{OPT}_0]} \tag{D.5}$$

*where* $\mathbb{E}_{\mathcal{A}}$ *is taken w.r.t. the randomness in outcome* $\boldsymbol{x}$ *due to randomness in the auction* $\mathcal{A}$*.*

*Proof.* First note that bidder 0 only has competition in auctions $A_1 ... A_K$, and hence can only incur a loss (that contributes to $\textsc{loss}_0(\boldsymbol{x})$ defined in Equations (5.6)) within these auctions. Hence $\mathbb{E}_{\mathcal{A}}[\textsc{loss}_0(\boldsymbol{x})] = v\sum_{j\in[K]}\mathbb{P}(\text{bidder 0 loses auction } j)$. Then we consider the following:

$$\begin{aligned}
\mathbb{E}_{\mathcal{A}}[\textsc{loss}_0(\boldsymbol{x})] &= v\sum_{j\in[K]}\mathbb{P}(\text{bidder 0 loses auction } j)\\
&= v\sum_{j\in[K]}\left(1-\mathbb{P}(\text{bidder 0 wins auction } j)\right)\\
&\overset{(a)}{\geq} v\cdot\frac{K^2}{K+1} = \frac{1-\beta}{\alpha_0-1}\cdot\gamma\cdot\pi_{\mathcal{A}}\cdot\frac{K^2}{K+1} \overset{(b)}{=} \frac{1-\beta}{\alpha_0-1}\cdot\mathbb{E}_{\mathcal{A}}\left[\mathrm{OPT}_{-0}\right]\cdot\frac{K}{K+1}\,.
\end{aligned} \tag{D.6}$$

Here (a) holds because bidder 0 bids $\alpha_0 v$ for any auction in 1,2...K, which is strictly less than all other bidders' bids as they all adopt multipliers $\rho$ in these auctions, so from Lemma D.3.1, we have $\mathbb{P}(\text{bidder 0 wins auction } j)\leq\frac{1}{K+1}$; in (b) we used the fact that $\mathbb{E}_{\mathcal{A}}\left[\mathrm{OPT}_{-0}\right]=\sum_{j=K+1}^{2K}\mathbb{E}_{\mathcal{A}}\left[\gamma\right]=\gamma\cdot K\cdot\pi_{\mathcal{A}}$ since there is only a single non-zero bidder in auctions $A_{K+1}\ldots A_{2K}$ and each bidder submits a bid $\rho\gamma>\rho>Q_{\mathcal{A}}$ (see Definition D.3.1).

Therefore we have

$$\lim_{K\to\infty}\frac{\mathbb{E}_{\mathcal{A}}\left[W_0(\boldsymbol{x})\right]}{\mathbb{E}_{\mathcal{A}}\left[\mathrm{OPT}_0\right]} \overset{(a)}{=} 1-\lim_{K\to\infty}\frac{\mathbb{E}_{\mathcal{A}}\left[\textsc{loss}_0(\boldsymbol{x})\right]}{\mathbb{E}_{\mathcal{A}}\left[\mathrm{OPT}_0\right]} \leq 1-\frac{1-\beta}{\alpha_0-1}\lim_{K\to\infty}\frac{\mathbb{E}_{\mathcal{A}}\left[\mathrm{OPT}_{-0}\right]}{\mathbb{E}_{\mathcal{A}}\left[\mathrm{OPT}_0\right]}\,, \tag{D.7}$$

where (a) follows from the fact that in our constructed autobidding instance, bidder 0's acquired value in each auction cannot exceed that under the efficient allocation, and hence can only incur loss in welfare.

Now it only remains to show that the multiplies $(\alpha_0,\rho,\ldots\rho)\in(1,\infty)^{K+1}$ yields a feasible outcome, i.e. the ROI constraints of each bidder is satisfied in expectation. Let $V_{i,j}$ and $C_{i,j}$ be the expected value and cost of bidder $i$ in auction $A_j$, respectively.

**1. Showing bidder** $0$**'s ROI constraint is satisfied**. We show by the following: bidder 0 only incurs a non-zero expected cost in auctions $A_1\ldots A_K$ and $A_{2K+1}$, and we will show that the expected value $V_{0,2K+1}$ is lower bounded by the expected costs

$C_{0,2K+1} + \sum_{j\in[K]} C_{0,j}$.

Since $\alpha_0 y > Q_{\mathcal{A}}$, the definition of the single-bidder purchasing probability in Definition D.3.1 implies that bidder 0 acquires an expected value from auction $A_{2K+1}$ of $V_{0,2K+1} = \pi_{\mathcal{A}} y$. Further, since bidder 0 is submits the lowest bids in auctions $A_1 \dots A_K$ under bid multiplier profile $(\alpha_0, \rho \dots \rho) \in (0,\infty)^{K+1}$, from Lemma D.3.1, we have $\mathbb{P}(\text{bidder 0 wins auction } j) \le \frac{1}{K+1}$ for all $j \in [K]$. Since the payment of a bidder in an auction is at most her submitted bid (as the auction is IR), we know that $\sum_{j\in[K]} C_{0,j} \le K \cdot \frac{\alpha_0 v}{K+1} < \pi_{\mathcal{A}} y = V_{0,2K+1}$, where the inequality follows from the definition of $y$ in Equation (D.4) such that $y > \max\left\{\frac{Q_{\mathcal{A}}}{\alpha_0}, \frac{\alpha_0 v}{\pi_{\mathcal{A}}}\right\}$. This implies bidder 0's ROI constraint is satisfied.

**2. Showing bidder $i$'s ROI constraint is satisfied for any $i = 1, 2 \dots K$.** We show this by considering the following: bidder $i$ only incurs a non-zero expected cost in auctions $A_1 \dots A_K$ and $A_{K+i}$, and we will show that the expected values $V_{i,K+i} + \sum_{k\in[K]} V_{i,j}$ is lower bounded by the expected costs $C_{i,K+i} + \sum_{j\in[K]} C_{i,j}$.

- **Calculate cost $C_{i,K+i}$:** For auction $A_{K+i}$, bidder $i$'s bid is $\rho\gamma > \gamma > Q_{\mathcal{A}}$ from the definition of $\gamma$, so by Definition D.3.1, the probability of $i$ winning the item in auction $A_{K+i}$ is $\pi_{\mathcal{A}}$, and the expected cost is

$$C_{i,K+i} \le \pi_{\mathcal{A}} \cdot \max\{r_{i,K+i}, Q_{\mathcal{A}}\} \le \pi_{\mathcal{A}} \cdot \beta\gamma\,, \tag{D.8}$$

  where the final inequality follows from the definition $r_{i,K+i} = \beta\gamma$

- **Upper bound costs $\sum_{j\in[K]} C_{i,j}$:** For auctions $[K] = 1 \dots K$, bidder $i$'s total expected cost can be bounded as

$$\begin{aligned}\sum_{j\in[K]} C_{i,j} \;&\le\; \rho \sum_{j\in[K]} v_{i,j} \mathbb{P}\left(\text{bidder } i \text{ wins auction } A_j\right) \\ &=\; \alpha_0 v \sum_{j\in[K]} \mathbb{P}\left(\text{bidder } i \text{ wins auction } A_j\right) + \frac{(K+1)K}{2}\epsilon\,.\end{aligned} \tag{D.9}$$

  where the first inequality follows from a bidder's payment is at most her submitted bid since the auction is IR.

- **Calculate $V_{i,K+i}$:** Considering auction $A_{K+i}$, bidder $i$ is the only bidder, and since $\rho\gamma > \gamma > Q_{\mathcal{A}}$, the definition of the single-bidder purchasing probability in Definition D.3.1 implies that bidder $i$'s acquires an expected value from this auction of

$$V_{i,K+i} = \pi_{\mathcal{A}} \cdot \gamma \,. \tag{D.10}$$

- **Lower bound $\sum_{k\in[K]} V_{i,j}$:**

$$\sum_{k\in[K]} V_{i,j} \geq \frac{\alpha_0 v}{\rho} \sum_{j\in[K]} \mathbb{P}\left(\text{bidder } i \text{ wins auction } j\right) \,. \tag{D.11}$$

Combining Equations (D.8),(D.9),(D.10) and (D.11), we get

$$\begin{aligned}
&\sum_{j\in[K]} V_{i,j} + V_{i,K+i} - \left(\sum_{j\in[K]} C_{i,j} + C_{i,K+i}\right) \\
\geq\ & \pi_{\mathcal{A}} \cdot \gamma + \frac{\alpha_0 v}{\rho} \cdot \sum_{j\in[K]} \mathbb{P}\left(\text{bidder } i \text{ wins auction } j\right) \\
&\quad - \left(\pi_{\mathcal{A}} \cdot \beta\gamma + \alpha_0 v \cdot \sum_{j\in[K]} \mathbb{P}\left(\text{bidder } i \text{ wins auction } j\right) + \frac{(K+1)K}{2}\epsilon\right) \\
=\ & \pi_{\mathcal{A}} \cdot (1-\beta)\gamma - \left(\alpha_0 - \frac{\alpha_0}{\rho}\right) v \cdot \sum_{j\in[K]} \mathbb{P}\left(\text{bidder } i \text{ wins auction } j\right) - \frac{(K+1)K}{2}\epsilon \\
\overset{(a)}{=}\ & (\alpha_0 - 1)v - \left(\alpha_0 - \frac{\alpha_0}{\rho}\right) v \cdot \sum_{j\in[K]} \mathbb{P}\left(\text{bidder } i \text{ wins auction } j\right) - \frac{(K+1)K}{2}\epsilon \\
\overset{(b)}{\geq}\ & (\alpha_0 - 1)v - \left(\alpha_0 - \frac{\alpha_0}{\rho}\right) v - \frac{(K+1)K}{2}\epsilon \\
&,
\end{aligned} \tag{D.12}$$

where (a) follows from the definition $v = \frac{1-\beta}{\alpha_0 - 1} \cdot \pi_{\mathcal{A}} \cdot \gamma$; In (b) we used the fact that $\rho > \alpha_0 > 1$ and $\sum_{j\in[K]} \mathbb{P}\left(\text{bidder } i \text{ wins auction } A_j\right) \leq 1$ due to the following:

Consider the set of bid values $\mathcal{B} = \{\alpha_0 v, \alpha_0 v + \epsilon, \alpha_0 v + 2\epsilon \dots \alpha_0 v + K\epsilon\} \subseteq \mathbb{R}_{>0}$, and we recognize that any bid value $b_k \in \mathcal{B}$ exceeds the maximim reserve price $\beta v$ in auctions $A_1 ... A_K$. Therefore the constructed reserve prices do not affect allocation, and hence by anonymity of auction $\mathcal{A}$ there exists probabilities $\boldsymbol{q}(\mathcal{B}) = (q_0(\mathcal{B}), q_1(\mathcal{B}) \dots q_K(\mathcal{B})) \in [0,1]^{K+1}$ where

$$q_k(\mathcal{B}) = \mathbb{P}(\text{bid value } b_k \text{ wins auction } \mathcal{A} \text{ given competing bids } \boldsymbol{b}_{-k}) \text{ and } \sum_{k=0}^{K} q_k(\mathcal{B}) \leq 1.$$

We recognize that in each auction $A_1 \dots A_K$, under bid multipliers $(\alpha_0, \rho, \dots \rho) \in (1, \infty)^{K+1}$ the submitted bid profile is a cyclic permutation of $\mathcal{B}$. Therefore we know that

$$\sum_{j \in [K]} \mathbb{P}\left(\text{bidder } i \text{ wins auction } j\right) = \sum_{k=1}^{K} q_k(\mathcal{B}) \leq 1 - q_0(\mathcal{B}) \leq 1$$

Finally, by taking $\rho \to \alpha_0$ and $K \to \infty$ in Equation (D.12), and utilizing $\epsilon = O(1/K^3)$ we have

$$\lim_{\rho \to \alpha_0} \lim_{K \to \infty} \sum_{j \in [K]} V_{i,j} + V_{i,K+i} - \left( \sum_{j \in [K]} C_{i,j} + C_{i,K+i} \right) \;\geq\; 0\,.$$

This shows that bidder $i$'s ROI constraint is satisfied. □

## D.4 Additional material for Section 5.6

### D.4.1 Proof of Theorem 5.6.2

*Proof.* For convenience, define $\delta = 2 - \frac{1}{\Delta}$, so $\Delta > 1$ implies $\delta \in (1,2)$, and further $1 > \beta > \frac{\Delta}{2\Delta - 1}$ implies $\frac{1}{\delta} < \beta < 1$.

Fix a bidder $i \in [K]$ and any feasible competing bid profile $\boldsymbol{b} \in \mathcal{U}$. Denote the corresponding outcome as $\boldsymbol{x} = \mathcal{X}(\boldsymbol{b})$, where $\boldsymbol{x} = (\boldsymbol{x}_1 ... \boldsymbol{x}_M)$ where $\boldsymbol{x}_j \in \{0,1\}^{N \times L_j}$ is the outcome vector in auction $\mathcal{A}_j$. Note that by definition of $\mathcal{U}$ which is the set of undominated and feasible bids, under the outcome $\boldsymbol{x}$ all bidders' ROAS constraints are satisfied. Denote $\ell_{k,j}$, $\ell^*_{k,j}$ to be the position of bidder $k \in [N]$ in auction $j \in [M]$ under outcome $\boldsymbol{x}$ and the efficient outcome, respectively.

Recall in Eq.(5.8) the definition for the set of all "coverings" for bidder $i$, denoted as $\mathcal{C}_i(\boldsymbol{x})$:

$$\mathcal{B}_i(k;\boldsymbol{x}) = \left\{ j \in [M] : \mathrm{OPT}_{i,j} > 0,\ v_{k,j} < v_{i,j} \text{ and } \ell_{k,j} \leq \ell^*_{i,j} < \ell_{i,j} \right\}$$
$$\mathcal{C}_i(\boldsymbol{x}) = \{\mathcal{C} \subseteq [N]/\{i\} : (\mathcal{B}_i(k;\boldsymbol{x}))_{k \in \mathcal{C}} \text{ is a maximal set cover of } \mathcal{L}_i(\boldsymbol{x})\}$$

where $\mathcal{L}_i(\boldsymbol{x}) = \{j \in [M] : W_{i,j}(\boldsymbol{x}) < \mathrm{OPT}_{i,j}\}$ is the set of auctions in which bidder $i$'s acquired welfare is less than that of her welfare under the efficient outcome; see Definition 5.4.1.

Denote $p_{k,j}$ as the payment of any bidder $k$, and $\hat{b}_{\ell,j}$ as the $\ell$th largest bid in any auction $j \in [M]$. Similar to the proof of Theorem 5.4.1, fix any covering $\mathcal{C} \subseteq \mathcal{C}_i(\boldsymbol{x})$, and any bidder $k \in \mathcal{C}$, such that in some auction $j \in \mathcal{B}_i(k;\boldsymbol{x})$, we have $v_{k,j} < v_{i,j}$ but $\ell_{k,j} \leq \ell^*_{i,j} < \ell_{i,j}$. Thus following a similar deduction as Eq. (5.12) in the proof of

Theorem 5.4.1, bidder $k$'s payment is lower bounded as follows: for $j \in \mathcal{B}_i(k;\boldsymbol{x})$,

$$
\begin{aligned}
p_{k,j} &\overset{(a)}{\geq} \sum_{\ell=\ell_{k,j}}^{L_j} \left(\mu(\ell)-\mu(\ell+1)\right)\hat{b}_{\ell+1,j} = \sum_{\ell=\ell_{k,j}}^{\ell_{i,j}-1} \left(\mu(\ell)-\mu(\ell+1)\right)\hat{b}_{\ell+1,j} + p_{i,j} \\
&\overset{(b)}{\geq} \beta\left(\mu(\ell_{k,j})-\mu(\ell_{i,j})\right)v_{i,j} + \beta\cdot\mu(\ell_{i,j})v_{i,j} \\
&= \beta\mu(\ell_{k,j})\cdot v_{i,j} \\
&= \mu(\ell_{k,j})v_{i,j} + \left(\beta-\frac{1}{\delta}\right)\left(\mu(\ell^*_{i,j})-\mu(\ell_{i,j})\right)v_{i,j} - (1-\beta)\cdot\mu(\ell_{i,j})v_{i,j} \\
&\quad - (1-\beta)\mu(\ell_{k,j})v_{i,j} + \left(\frac{1}{\delta}-\beta\right)\mu(\ell^*_{i,j})v_{i,j} + \left(1-\frac{1}{\delta}\right)\mu(\ell_{i,j})v_{i,j} \qquad \text{(D.13)} \\
&\overset{(c)}{\geq} \mu(\ell_{k,j})v_{i,j} + \left(\beta-\frac{1}{\delta}\right)\left(\mu(\ell^*_{i,j})-\mu(\ell_{i,j})\right)v_{i,j} - (1-\beta)\cdot\mu(\ell_{i,j})v_{i,j} \\
&\quad - \left(1-\frac{1}{\delta}\right)\mu(\ell_{k,j})v_{i,j} + \left(1-\frac{1}{\delta}\right)\mu(\ell_{i,j})v_{i,j} \\
&= \left(\beta-\frac{1}{\delta}\right)\left(\mu(\ell^*_{i,j})-\mu(\ell_{i,j})\right)v_{i,j} - (1-\beta)\cdot\mu(\ell_{i,j})v_{i,j} \\
&\quad + \frac{1}{\delta}\mu(\ell_{k,j})v_{i,j} + \left(1-\frac{1}{\delta}\right)\mu(\ell_{i,j})v_{i,j}
\end{aligned}
$$

Here , (a) follows from the fact that for a fix bid profile, the payment of GSP or GFP for each bidder in an auction dominates that of VCG (see Example D.1.1 and discussions thereof); (b) follows from $\hat{b}_{\ell,j} \geq b_{i,j}$ for $\ell \leq \ell_{i,j}$, and since $\boldsymbol{b} \in \mathcal{U} \subseteq \mathbb{R}_+^{N\times M}$ is an undominated bid profile, Lemma 5.6.1 applies and $b_{i,j} \geq \beta v_{i,j}$. Also $p_{i,j} \geq r_{i,j} \geq \beta v_{i,j}$ be the definition of $\beta$-approximate reserves; (c) follows from the fact that $\beta > \frac{1}{\delta}$ and $\mu(\ell^*_{i,j}) \leq \mu(\ell_{k,j})$ since $\ell_{k,j} \leq \ell^*_{i,j}$ for any $k \in \mathcal{C} \subseteq \mathcal{C}_i(\boldsymbol{x})$ and $j \in \mathcal{B}_i(k;\boldsymbol{x})$; see definition in Eq. (5.8).

On the other hand, we have

$$
\begin{aligned}
&\sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} p_{k,j} + \sum_{j\notin\mathcal{B}_i(k;\boldsymbol{x})} p_{k,j} \leq \sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{k,j})v_{k,j} + \sum_{j\notin\mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{k,j})v_{k,j} \\
&p_{k,j} \geq \beta\cdot\mu(\ell_{k,j})v_{k,j} \quad \forall j \in [M],
\end{aligned}
$$

where the first inequality follows from bidder $k$'s ROAS constraint; the second inequality follows from the fact that any winning bidder's payment must be greater than her

$\beta$-approximate reserves.

Combining the above inequalities and rearranging we get

$$\sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} p_{k,j} \leq \sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{k,j})v_{k,j} + (1-\beta)\cdot \sum_{j\notin\mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{k,j})v_{k,j}\,, \tag{D.14}$$

Summing Eq.(D.13) over all $j \in \mathcal{B}_i(k;\boldsymbol{x})$ and combining with Eq. (D.14), we get

$$\begin{aligned}
&\left(\beta-\frac{1}{\delta}\right)\cdot \sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \left(\mu(\ell^*_{i,j})-\mu(\ell_{i,j})\right) v_{i,j} \\
\leq\;& (1-\beta)\cdot\left(\sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{i,j})v_{i,j} + \sum_{j\notin\mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{k,j})v_{k,j}\right) \\
&\quad + \underbrace{\sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{k,j})v_{k,j} - \frac{1}{\delta}\sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{k,j})v_{i,j} - \left(1-\frac{1}{\delta}\right)\sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{i,j})v_{i,j}}_{Y}\,.
\end{aligned} \tag{D.15}$$

We now upper bound $Y$:

$$
\begin{aligned}
&\sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{k,j})v_{k,j} - \frac{1}{\delta}\sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{k,j})v_{i,j} - \left(1-\frac{1}{\delta}\right)\sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{i,j})v_{i,j} \\
=\;& \left(1-\frac{1}{\delta}\right)\sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{k,j})v_{k,j} - \frac{1}{\delta}\sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{k,j})\left(v_{i,j}-v_{k,j}\right) \\
&\quad - \left(1-\frac{1}{\delta}\right)\sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{i,j})v_{i,j} \\
=\;& \left(1-\frac{1}{\delta}\right)\sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{k,j})\left(v_{k,j}-v_{i,j}\right) - \frac{1}{\delta}\sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{k,j})\left(v_{i,j}-v_{k,j}\right) \\
&\quad + \left(1-\frac{1}{\delta}\right)\sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \left(\mu(\ell_{k,j})-\mu(\ell_{i,j})\right)v_{i,j} \\
=\;& \left(1-\frac{1}{\delta}\right)\sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{k,j})\left(v_{k,j}-v_{i,j}\right) \\
&\quad + \sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \frac{\mu(\ell_{k,j})v_{i,j}}{\delta}\left((\delta-1)\left(1-\frac{\mu(\ell_{i,j})}{\mu(\ell_{k,j})}\right)-\left(1-\frac{v_{k,j}}{v_{i,j}}\right)\right) \\
\stackrel{(a)}{\leq}\;& \left(1-\frac{1}{\delta}\right)\sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{k,j})\left(v_{k,j}-v_{i,j}\right) + \sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \frac{\mu(\ell_{k,j})v_{i,j}}{\delta}\left((\delta-1)-\left(1-\frac{v_{k,j}}{v_{i,j}}\right)\right) \\
=\;& \left(1-\frac{1}{\delta}\right)\sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{k,j})\left(v_{k,j}-v_{i,j}\right) + \sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \frac{\mu(\ell_{k,j})v_{i,j}}{\delta}\left((\delta-2)+\frac{v_{k,j}}{v_{i,j}}\right) \\
\stackrel{(b)}{\leq}\;& \left(1-\frac{1}{\delta}\right)\sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{k,j})\left(v_{k,j}-v_{i,j}\right) \\
\stackrel{(c)}{\leq}\;& (1-\beta)\sum_{j\in\mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{k,j})\left(v_{k,j}-v_{i,j}\right)\,.
\end{aligned}
\tag{D.16}
$$

where in (a) we recall $\delta > 1$ and $\ell_{k,j} < \ell_{i,j}$ for any $k \in \mathcal{C}$ and $j \in \mathcal{B}_i(k;\boldsymbol{x})$ so that $\mu(\ell_{k,j}) > \mu(\ell_{i,j})$; (b) follows from the fact that values are $\delta$-separated, so $v_{i,j} > v_{k,j}$ for $k \in \mathcal{C}$ and $j \in \mathcal{B}_i(k;\boldsymbol{x})$ implies $\frac{v_{k,j}}{v_{i,j}} \leq \frac{1}{\Delta} = 2-\delta$; in (c) we used the fact that $\beta > \frac{1}{\delta}$ so $1-\beta < 1-\frac{1}{\delta}$, and the fact that $v_{k,j} < v_{i,j}$ for any $k \in \mathcal{C}$ and $j \in \mathcal{B}_i(k;\boldsymbol{x})$.

Combining Equations (D.15) and (D.16) we get

$$
\begin{aligned}
&\left(\beta - \frac{1}{\delta}\right) \cdot \sum_{j \in \mathcal{B}_i(k;\boldsymbol{x})} \left(\mu(\ell^*_{i,j}) - \mu(\ell_{i,j})\right) v_{i,j} \\
\leq\ & (1-\beta) \cdot \left( \sum_{j \in \mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{i,j}) v_{i,j} + \sum_{j \notin \mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{k,j}) v_{k,j} + \sum_{j \in \mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{k,j}) \left(v_{k,j} - v_{i,j}\right) \right) \\
=\ & (1-\beta) \cdot \left( \sum_{j \in \mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{i,j}) v_{i,j} + \sum_{j \in [M]} \mu(\ell_{k,j}) v_{k,j} - \sum_{j \in \mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{k,j}) v_{i,j} \right) \\
\stackrel{(a)}{\leq}\ & (1-\beta) \cdot \left( \sum_{j \in \mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell_{i,j}) v_{i,j} + \sum_{j \in [M]} \mu(\ell_{k,j}) v_{k,j} - \sum_{j \in \mathcal{B}_i(k;\boldsymbol{x})} \mu(\ell^*_{i,j}) v_{i,j} \right) . \\
\Longrightarrow & \sum_{j \in \mathcal{B}_i(k;\boldsymbol{x})} \left(\mu(\ell^*_{i,j}) - \mu(\ell_{i,j})\right) v_{i,j} \ \leq\ \frac{1-\beta}{1-\frac{1}{\delta}} \sum_{j \in [M]} \mu(\ell_{k,j}) v_{k,j}
\end{aligned}
\tag{D.17}
$$

where (a) follows from $\mu(\ell^*_{i,j}) \leq \mu(\ell_{k,j})$ due to the fact that $\ell_{k,j} < \ell_{i,j}$ for any $k \in \mathcal{C}$ and $j \in \mathcal{B}_i(k;\boldsymbol{x})$.

Summing the above over all $k \in \mathcal{C}$, and following the same arguments as in Eq.(5.10) of the proof of Theorem 5.4.1, we have

$$
\begin{aligned}
\textsc{loss}_i(\boldsymbol{x}) &= \sum_{j \in \mathcal{L}_i(\boldsymbol{x})} \left(\mu(\ell^*_{i,j}) - \mu(\ell_{i,j})\right) v_{i,j} \\
&\leq \sum_{k \in \mathcal{C}} \sum_{j \in \mathcal{B}_i(k;\boldsymbol{x})} \left(\mu(\ell^*_{i,j}) - \mu(\ell_{i,j})\right) v_{i,j} \\
&\leq \frac{1-\beta}{1-\frac{1}{\delta}} \sum_{k \in \mathcal{C}} \sum_{j \in [M]} \mu(\ell_{k,j}) v_{k,j} \\
&= \frac{1-\beta}{1-\frac{1}{\delta}} \sum_{k \in \mathcal{C}} W_k(\boldsymbol{x}) \\
&\leq \frac{1-\beta}{1-\frac{1}{\delta}} W_{-i}(\boldsymbol{x}) \\
&\leq \frac{1-\beta}{1-\frac{1}{\delta}} \left(\mathrm{OPT}_{-i} + \textsc{loss}_i(\boldsymbol{x})\right) .
\end{aligned}
\tag{D.18}
$$

Rearranging we get $\textsc{loss}_i(\boldsymbol{x}) \leq \frac{1-\beta}{\beta - \frac{1}{\delta}} \mathrm{OPT}_{-i} = \frac{1-\beta}{\beta - \frac{\Delta}{2\Delta - 1}} \mathrm{OPT}_{-i}$. Finally, applying

Proposition 5.4.4 w.r.t. upper bound of $\textsc{loss}_i(\boldsymbol{x})$ and using the fact that the competing bid profile is arbitrary, we obtain the desired welfare guarantee lower bound. $\square$

# Bibliography


[1] Programmatic advertising trends, stats, & news. https://roirevolution.com/blog/programmatic-advertising-trends-stats-news/. Accessed: 2023-03-30.

[2] Programmatic digital display advertising in 2022: Ad spend, formats, and forecast. https://www.insiderintelligence.com/insights/programmatic-digital-display-ad-spending/. Accessed: 2023-03-30.

[3] Alekh Agarwal, Daniel Hsu, Satyen Kale, John Langford, Lihong Li, and Robert Schapire. Taming the monster: A fast and simple algorithm for contextual bandits. In *International Conference on Machine Learning*, pages 1638–1646. PMLR, 2014.

[4] Gagan Aggarwal, Ashwinkumar Badanidiyuru, and Aranyak Mehta. Autobidding with constraints. In *International Conference on Web and Internet Economics*, pages 17–30. Springer, 2019.

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

[6] Kareem Amin, Afshin Rostamizadeh, and Umar Syed. Learning prices for repeated auctions with strategic buyers. In *Advances in Neural Information Processing Systems*, pages 1169–1177, 2013.

[7] Kareem Amin, Afshin Rostamizadeh, and Umar Syed. Repeated contextual auctions with strategic buyers. In *Advances in Neural Information Processing Systems*, pages 622–630, 2014.

[8] Charles Audet and John E Dennis Jr. A pattern search filter method for nonlinear programming without derivatives. *SIAM Journal on Optimization*, 14(4):980–1010, 2004.

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

[10] Peter Auer, Nicolo Cesa-Bianchi, Yoav Freund, and Robert E Schapire. Gambling in a rigged casino: The adversarial multi-armed bandit problem. In *Proceedings*

*of IEEE 36th Annual Foundations of Computer Science*, pages 322–331. IEEE, 1995.

[11] Santiago Balseiro, Yuan Deng, Jieming Mao, Vahab Mirrokni, and Song Zuo. Robust auction design in the auto-bidding world. *Advances in Neural Information Processing Systems*, 34:17777–17788, 2021.

[12] Santiago Balseiro, Negin Golrezaei, Mohammad Mahdian, Vahab Mirrokni, and Jon Schneider. Contextual bandits with cross-learning. *Advances in Neural Information Processing Systems*, 32, 2019.

[13] Santiago Balseiro, Negin Golrezaei, Vahab Mirrokni, and Sadra Yazdanbod. A black-box reduction in mechanism design with private cost of capital. *Available at SSRN 3341782*, 2019.

[14] Santiago Balseiro, Anthony Kim, Mohammad Mahdian, and Vahab Mirrokni. Budget management strategies in repeated auctions. In *Proceedings of the 26th International Conference on World Wide Web*, pages 15–23, 2017.

[15] Santiago Balseiro, Anthony Kim, and Daniel Joseph Russo. On the futility of dynamics in robust mechanism design. *Columbia Business School Research Paper Forthcoming*, 2019.

[16] Santiago R Balseiro, Yuan Deng, Jieming Mao, Vahab S Mirrokni, and Song Zuo. The landscape of auto-bidding auctions: Value versus utility maximization. In *Proceedings of the 22nd ACM Conference on Economics and Computation*, pages 132–133, 2021.

[17] Santiago R Balseiro and Yonatan Gur. Learning in repeated auctions with budgets: Regret minimization and equilibrium. *Management Science*, 65(9):3952–3968, 2019.

[18] Santiago R Balseiro, Haihao Lu, and Vahab Mirrokni. The best of many worlds: Dual mirror descent for online allocation problems. *Operations Research*, 2022.

[19] Hamsa Bastani and Mohsen Bayati. Online decision-making with high-dimensional covariates. *Available at SSRN 2661896*, 2015.

[20] Dirk Bergemann and Maher Said. Dynamic auctions. *Wiley Encyclopedia of Operations Research and Management Science*, 2010.

[21] Dimitri P Bertsekas. *Constrained optimization and Lagrange multiplier methods*. Academic press, 2014.

[22] Omar Besbes, Yonatan Gur, and Assaf Zeevi. Stochastic multi-armed-bandit problem with non-stationary rewards. *Advances in neural information processing systems*, 27, 2014.

[23] John R Birge, Yifan Feng, N Bora Keskin, and Adam Schultz. Dynamic learning and market making in spread betting markets with informed bettors. *Available at SSRN 3283392*, 2018.

[24] Petrônio L Braga, Adriano LI Oliveira, and Silvio RL Meira. Software effort estimation using machine learning techniques with robust confidence intervals. In *7th international conference on hybrid intelligent systems (HIS 2007)*, pages 352–357. IEEE, 2007.

[25] Mark Braverman, Jieming Mao, Jon Schneider, and Matt Weinberg. Selling to a no-regret buyer. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pages 523–538, 2018.

[26] Árpád Bűrmen, Janez Puhan, and Tadej Tuma. Grid restrained nelder-mead algorithm. *Computational optimization and applications*, 34(3):359–375, 2006.

[27] L Elisa Celis, Gregory Lewis, Markus Mobius, and Hamid Nazerzadeh. Buy-it-now or take-a-chance: Price discrimination through randomized auctions. *Management Science*, 60(12):2927–2948, 2014.

[28] Hongfan Chen and N Bora Keskin. Markdown policies for demand learning and strategic customer behavior. *Available at SSRN 3299819*, 2018.

[29] Wang Chi Cheung, David Simchi-Levi, and Ruihao Zhu. Learning to optimize under non-stationarity. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 1079–1087. PMLR, 2019.

[30] José Correa, Andrés Cristi, Laurent Feuilloley, Tim Oosterwijk, and Alexandros Tsigonias-Dimitriadis. The secretary problem with independent sampling. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2047–2058. SIAM, 2021.

[31] Bo Dai, Ofir Nachum, Yinlam Chow, Lihong Li, Csaba Szepesvári, and Dale Schuurmans. Coindice: Off-policy confidence interval estimation. *Advances in neural information processing systems*, 33:9398–9411, 2020.

[32] George B Dantzig. Discrete-variable extremum problems. *Operations research*, 5(2):266–288, 1957.

[33] Yuan Deng, Negin Golrezaei, Patrick Jaillet, Jason Cheuk Nam Liang, and Vahab Mirrokni. Fairness in the autobidding world with machine-learned advice. *arXiv preprint arXiv:2209.04748*, 2022.

[34] Yuan Deng, Negin Golrezaei, Patrick Jaillet, Jason Cheuk Nam Liang, and Vahab Mirrokni. Multi-channel autobidding with budget and roi constraints. *arXiv preprint arXiv:2302.01523*, 2023.

[35] Yuan Deng, Sebastien Lahaie, and Vahab Mirrokni. Robust pricing in non-clairvoyant dynamic mechanism design. *Available at SSRN*, 2019.

[36] Yuan Deng, Jieming Mao, Vahab Mirrokni, Hanrui Zhang, and Song Zuo. Efficiency of the first-price auction in the autobidding world. *arXiv preprint arXiv:2208.10650*, 2022.

[37] Yuan Deng, Jieming Mao, Vahab Mirrokni, and Song Zuo. Towards efficient auctions in an auto-bidding world. In *Proceedings of the Web Conference 2021*, pages 3965–3973, 2021.

[38] Yuan Deng, Jon Schneider, and Balasubramanian Sivan. Prior-free dynamic auctions with low regret buyers. *Advances in Neural Information Processing Systems*, 32, 2019.

[39] Mahsa Derakhshan, Negin Golrezaei, and Renato Paes Leme. Lp-based approximation for personalized reserve prices. *arXiv preprint arXiv:1905.01526*, 2019.

[40] Qin Ding, Cho-Jui Hsieh, and James Sharpnack. An efficient algorithm for generalized linear bandit: Online stochastic gradient descent and thompson sampling. In *International Conference on Artificial Intelligence and Statistics*, pages 1585–1593. PMLR, 2021.

[41] Alexey Drutsa. Reserve pricing in repeated second-price auctions with strategic bidders. *arXiv preprint arXiv:1906.09331*, 2019.

[42] Alexey Drutsa. Optimal non-parametric learning in repeated contextual auctions with strategic buyer. In *International Conference on Machine Learning*, pages 2668–2677. PMLR, 2020.

[43] Paul Dütting, Zhe Feng, Harikrishna Narasimhan, David Parkes, and Sai Srivatsa Ravindranath. Optimal auctions through deep learning. In *International Conference on Machine Learning*, pages 1706–1715. PMLR, 2019.

[44] Aryeh Dvoretzky, Jack Kiefer, Jacob Wolfowitz, et al. Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *The Annals of Mathematical Statistics*, 27(3):642–669, 1956.

[45] Kwassi Joseph Dzahini, Michael Kokkolaras, and Sébastien Le Digabel. Constrained stochastic blackbox optimization using a progressive barrier and probabilistic estimates. *Mathematical Programming*, pages 1–58, 2022.

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

[47] Scott R Eliason. *Maximum likelihood estimation: Logic and practice*. Number 96. Sage, 1993.

[48] Giovanni Fasano, Giampaolo Liuzzi, Stefano Lucidi, and Francesco Rinaldi. A linesearch-based derivative-free approach for nonsmooth constrained optimization. *SIAM journal on optimization*, 24(3):959–992, 2014.

[49] Zhe Feng, Sebastien Lahaie, Jon Schneider, and Jinchao Ye. Reserve price optimization for first price auctions in display advertising. In *International Conference on Machine Learning*, pages 3230–3239. PMLR, 2021.

[50] Zhe Feng, Swati Padmanabhan, and Di Wang. Online bidding algorithms for return-on-spend constrained advertisers. *arXiv preprint arXiv:2208.13713*, 2022.

[51] Zhe Feng, David C Parkes, and Haifeng Xu. The intrinsic robustness of stochastic bandits to strategic manipulation. *arXiv preprint arXiv:1906.01528*, 2019.

[52] Abraham D Flaxman, Adam Tauman Kalai, and H Brendan McMahan. Online convex optimization in the bandit setting: gradient descent without a gradient. *arXiv preprint cs/0408007*, 2004.

[53] Anindya Ghose and Sha Yang. An empirical analysis of search engine advertising: Sponsored search in electronic markets. *Management science*, 55(10):1605–1622, 2009.

[54] Negin Golrezaei, Patrick Jaillet, and Jason Cheuk Nam Liang. Incentive-aware contextual pricing with non-parametric market noise. *arXiv preprint arXiv:1911.03508*, 2019.

[55] Negin Golrezaei, Patrick Jaillet, Jason Cheuk Nam Liang, and Vahab Mirrokni. Bidding and pricing in budget and roi constrained markets. *arXiv preprint arXiv:2107.07725*, 2021.

[56] Negin Golrezaei, Patrick Jaillet, and Zijie Zhou. Online resource allocation with samples. *Available at SSRN*, 2022.

[57] Negin Golrezaei, Adel Javanmard, and Vahab Mirrokni. Dynamic incentive-aware learning: Robust pricing in contextual auctions. 2018.

[58] Negin Golrezaei, Adel Javanmard, and Vahab Mirrokni. Dynamic incentive-aware learning: Robust pricing in contextual auctions. *Advances in Neural Information Processing Systems*, 32, 2019.

[59] Negin Golrezaei, Adel Javanmard, and Vahab Mirrokni. Dynamic incentive-aware learning: Robust pricing in contextual auctions. *Operations Research*, 2020.

[60] Negin Golrezaei, Max Lin, Vahab Mirrokni, and Hamid Nazerzadeh. Boosted second price auctions: Revenue optimization for heterogeneous bidders. 2017.

[61] Negin Golrezaei, Ilan Lobel, and Renato Paes Leme. Auction design for roi-constrained buyers. In *Proceedings of the international conference on World Wide Web*, 2021.

[62] Negin Golrezaei, Ilan Lobel, and Renato Paes Leme. Auction design for roi-constrained buyers. In *Proceedings of the Web Conference 2021*, pages 3941–3952, 2021.

[63] Negin Golrezaei, Vahideh Manshadi, Jon Schneider, and Shreyas Sekar. Learning product rankings robust to fake users. In *Proceedings of the 22nd ACM Conference on Economics and Computation*, pages 560–561, 2021.

[64] Anupam Gupta, Tomer Koren, and Kunal Talwar. Better algorithms for stochastic bandits with adversarial corruptions. In *Conference on Learning Theory*, pages 1562–1578. PMLR, 2019.

[65] Yanjun Han, Zhengyuan Zhou, Aaron Flores, Erik Ordentlich, and Tsachy Weissman. Learning to bid optimally and efficiently in adversarial first-price auctions. *arXiv preprint arXiv:2007.04568*, 2020.

[66] Yanjun Han, Zhengyuan Zhou, and Tsachy Weissman. Optimal no-regret learning in repeated first-price auctions. *arXiv preprint arXiv:2003.09795*, 2020.

[67] Yuxuan Han, Zhipeng Liang, Yang Wang, and Jiheng Zhang. Generalized linear bandits with local differential privacy. *Advances in Neural Information Processing Systems*, 34:26511–26522, 2021.

[68] Alan J Hoffman. On approximate solutions of systems of linear inequalities. In *Selected Papers Of Alan J Hoffman: With Commentary*, pages 174–176. World Scientific, 2003.

[69] Piotr Indyk, Frederik Mallmann-Trenn, Slobodan Mitrovic, and Ronitt Rubinfeld. Online page migration with ml advice. In *International Conference on Artificial Intelligence and Statistics*, pages 1655–1670. PMLR, 2022.

[70] Adel Javanmard. Perishability of data: dynamic pricing under varying-coefficient models. *The Journal of Machine Learning Research*, 18(1):1714–1744, 2017.

[71] Adel Javanmard and Hamid Nazerzadeh. Dynamic pricing in high-dimensions. *arXiv preprint arXiv:1609.07574*, 2016.

[72] Bo Jiang, Xuegong Zhang, and Tianxi Cai. Estimating the confidence interval for prediction errors of support vector machine classifiers. *The Journal of Machine Learning Research*, 9:521–540, 2008.

[73] PK Kannan, Werner Reinartz, and Peter C Verhoef. The path to purchase and attribution modeling: Introduction to special section, 2016.

[74] Yash Kanoria and Hamid Nazerzadeh. Dynamic reserve prices for repeated auctions: Learning from bids. *Available at SSRN 2444495*, 2017.

[75] Yash Kanoria and Hamid Nazerzadeh. Incentive-compatible learning of reserve prices for repeated auctions. *Operations Research*, 69(2):509–524, 2021.

[76] Pavel Kireyev, Koen Pauwels, and Sunil Gupta. Do display ads influence search? attribution and dynamics in online advertising. *International Journal of Research in Marketing*, 33(3):475–490, 2016.

[77] Robert Kleinberg and Tom Leighton. The value of knowing a demand curve: Bounds on regret for online posted-price auctions. In *44th Annual IEEE Symposium on Foundations of Computer Science, 2003. Proceedings.*, pages 594–605. IEEE, 2003.

[78] Christos Koufogiannakis and Neal E Young. A nearly linear-time ptas for explicit fractional packing and covering linear programs. *Algorithmica*, 70(4):648–674, 2014.

[79] Jean-Jacques Laffont and Jacques Robert. Optimal auction with financially constrained buyers. *Economics Letters*, 52(2):181–186, 1996.

[80] Sébastien Lahaie, David M Pennock, Amin Saberi, and Rakesh V Vohra. Sponsored search auctions. *Algorithmic game theory*, 1:699–716, 2007.

[81] Jeffrey Larson, Matt Menickelly, and Stefan M Wild. Derivative-free optimization methods. *Acta Numerica*, 28:287–404, 2019.

[82] Tor Lattimore and Csaba Szepesvári. *Bandit algorithms.* Cambridge University Press, 2020.

[83] Kuang-chih Lee, Burkay Orten, Ali Dasdan, and Wentong Li. Estimating conversion rate in display advertising from past erformance data. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 768–776, 2012.

[84] Jinyan Liu, Zhiyi Huang, and Xiangning Wang. Learning optimal reserve price against non-myopic bidders. In *Advances in Neural Information Processing Systems*, pages 2038–2048, 2018.

[85] Haipeng Luo, Chen-Yu Wei, Alekh Agarwal, and John Langford. Efficient contextual bandits in non-stationary worlds. In *Conference On Learning Theory*, pages 1739–1776. PMLR, 2018.

[86] Thodoris Lykouris, Vahab Mirrokni, and Renato Paes Leme. Stochastic bandits robust to adversarial corruptions. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 114–122, 2018.

[87] Thodoris Lykouris and Sergei Vassilvtiskii. Competitive caching with machine learned advice. In *International Conference on Machine Learning*, pages 3296–3305. PMLR, 2018.

[88] H Brendan McMahan, Gary Holt, David Sculley, Michael Young, Dietmar Ebner, Julian Grady, Lan Nie, Todd Phillips, Eugene Davydov, Daniel Golovin, et al. Ad click prediction: a view from the trenches. In *Proceedings of the 19th ACM*

*SIGKDD international conference on Knowledge discovery and data mining*, pages 1222–1230, 2013.

[89] Aranyak Mehta. Auction design in an auto-bidding setting: Randomization improves efficiency beyond vcg. In *Proceedings of the ACM Web Conference 2022*, pages 173–181, 2022.

[90] Mehryar Mohri and Andrés Munoz Medina. Learning algorithms for second-price auctions with reserve. *The Journal of Machine Learning Research*, 17(1):2632–2656, 2016.

[91] Andres Munoz and Sergei Vassilvitskii. Revenue optimization with approximate bid predictions. *Advances in Neural Information Processing Systems*, 30, 2017.

[92] Roger B Myerson. Optimal auction design. *Mathematics of operations research*, 6(1):58–73, 1981.

[93] Anthony Nguyen and Krishnakumar Balasubramanian. Stochastic zeroth-order functional constrained optimization: Oracle complexity and applications. *INFORMS Journal on Optimization*, 2022.

[94] Renato Paes Leme, Martin Pal, and Sergei Vassilvitskii. A field guide to personalized reserve prices. In *Proceedings of the 25th international conference on world wide web*, pages 1093–1102, 2016.

[95] Mallesh M Pai and Rakesh Vohra. Optimal auctions with financially constrained buyers. *Journal of Economic Theory*, 150:383–425, 2014.

[96] David Pardoe, Peter Stone, Maytal Saar-Tsechansky, Tayfun Keskin, and Kerem Tomak. Adaptive auction mechanism design and the incorporation of prior knowledge. *INFORMS Journal on Computing*, 22(3):353–370, 2010.

[97] Tony Pourmohamad and Herbert KH Lee. The statistical filter approach to constrained optimization. *Technometrics*, 62(3):303–312, 2020.

[98] Jian Qian, Ronan Fruit, Matteo Pirotta, and Alessandro Lazaric. Concentration inequalities for multinoulli random variables. *arXiv preprint arXiv:2001.11595*, 2020.

[99] Matthew Richardson, Ewa Dominowska, and Robert Ragno. Predicting clicks: estimating the click-through rate for new ads. In *Proceedings of the 16th international conference on World Wide Web*, pages 521–530, 2007.

[100] Tim Roughgarden. Intrinsic robustness of the price of anarchy. *Journal of the ACM (JACM)*, 62(5):1–42, 2015.

[101] Durga L Shrestha and Dimitri P Solomatine. Machine learning approaches for estimation of prediction interval for the model output. *Neural networks*, 19(2):225–235, 2006.

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

[103] Eric Sodomka, Sébastien Lahaie, and Dustin Hillard. A predictive model for advertiser value-per-click in sponsored search. In *Proceedings of the 22nd international conference on World Wide Web*, pages 1179–1190, 2013.

[104] Joel A Tropp et al. An introduction to matrix concentration inequalities. *Foundations and Trends® in Machine Learning*, 8(1-2):1–230, 2015.

[105] Ilnura Usmanova, Andreas Krause, and Maryam Kamgarpour. Safe convex learning under uncertain constraints. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 2106–2114. PMLR, 2019.

[106] Hal R Varian. Position auctions. *international Journal of industrial Organization*, 25(6):1163–1178, 2007.

[107] Rahul Vaze. Online knapsack problem under expected capacity constraint. In *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*, pages 2159–2167. IEEE, 2018.

[108] Shufan Wang, Jian Li, and Shiqiang Wang. Online algorithms for multi-shop ski rental with machine learned advice. *Advances in Neural Information Processing Systems*, 33:8150–8160, 2020.

[109] Jonathan Weed, Vianney Perchet, and Philippe Rigollet. Online learning in repeated auctions. In *Conference on Learning Theory*, pages 1562–1583. PMLR, 2016.

[110] Tsachy Weissman, Erik Ordentlich, Gadiel Seroussi, Sergio Verdu, and Marcelo J Weinberger. Inequalities for the l1 deviation of the empirical distribution. *Hewlett-Packard Labs, Tech. Rep*, 2003.

[111] Yidan Xing, Zhilin Zhang, Zhenzhe Zheng, Chuan Yu, Jian Xu, Fan Wu, and Guihai Chen. Designing ad auctions with private constraints for automated bidding. *arXiv preprint arXiv:2301.13020*, 2023.

[112] Yanwu Yang and Panyu Zhai. Click-through rate prediction in online advertising: A literature review. *Information Processing & Management*, 59(2):102853, 2022.

[113] Robert Zeithammer. Research note—strategic bid-shading and sequential auctioning with learning from past prices. *Management Science*, 53(9):1510–1519, 2007.

[114] Julian Zimmert and Yevgeny Seldin. An optimal algorithm for stochastic and adversarial bandits. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 467–475. PMLR, 2019.