Optimising Budget Management via Primal-Dual Approximation with Constrained Polynomial Weights Update

Dmitrii Moor
dmitriim@spotify.com
Spotify, London, United Kingdom

Per Berglund*
pberglund@spotify.com
Spotify, Stockholm, Sweden

Hannes Karlbom*
hannesk@spotify.com
Spotify, Stockholm, Sweden

Zhenwen Dai*
zhenwend@spotify.com
Spotify, London, United Kingdom

Kyle Kretschman
kylek@spotify.com
Spotify, New York, USA

Mounia Lalmas
mounia@acm.org
Spotify, London, United Kingdom

ABSTRACT

Budget management is an essential capability in many online applications, including running advertising campaigns in sponsored search and allocating promotional content in recommender systems (RS). Most existing approaches to optimising budget spendings rely on improving *worst-case* approximation guarantees in the respective online knapsack packing problem or leveraging online learning techniques to improve an *average* system performance. However, worst-case approaches often underperform in practice, as extreme scenarios are uncommon, while online learning methods may lack robustness in highly non-stationary environments. In our work, we bridge these two approaches by developing an online budget pacing algorithm that preserves worst-case guarantees while improving the average allocative efficiency of budget management systems.

Specifically, we frame the optimal budget allocation problem within an online learning perspective, where strategies chosen by the online learner are constrained by a primal-dual approximation algorithm defining the environment. We identify a feedback loop between the online learner's strategy and environment's response and propose a method to break this loop, optimising average allocative efficiency without compromising worst-case guarantees. To evaluate our approach, we conduct extensive offline experiments using both synthetic and real-world data, covering 320M promotional impressions across 20 different markets of a major industrial RS. Our results demonstrate that our method significantly enhances total allocated value and ROI across all markets.

ACM Reference Format:
Dmitrii Moor, Per Berglund, Hannes Karlbom, Zhenwen Dai, Kyle Kretschman, and Mounia Lalmas. 2025. Optimising Budget Management via Primal-Dual Approximation with Constrained Polynomial Weights Update. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

*The authors contributed equally to the paper.

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

To build a sustainable business, many online platforms rely on sponsored content, such as advertisements or promotional material, to augment users' search or recommendation feeds. These sponsored campaigns typically come with fixed budgets, which platforms allocate over a set period to maximize exposure and engagement among targeted user groups. Efficient budget utilization is crucial for online platforms to maximize value per dollar spent, attract new campaigns, and grow their sponsored revenue streams.

While numerous studies focus on optimizing budget spending, most existing approaches fall into two categories. The first relies on online approximation algorithms that offer theoretical *worst-case* guarantees [1–3], while the second leverages online learning techniques to improve budget allocation performance over time *on average* [4–7]. The former approach, often based on the primal-dual approximation of the online knapsack packing problem [2], requires minimal prior domain knowledge. This makes it robust to distributional shifts, which frequently occur in practice due to seasonality and the constant influx of new campaigns. However, prioritizing worst-case guarantees may not lead to strong average performance, as worst-case scenarios are not always typical. In contrast, online learning-based approaches improve the average efficiency of budget management [4, 8]. However, they tend to be less robust in highly dynamic environments where budgets, targeting parameters, and campaign durations fluctuate significantly. In such cases, maintaining worst-case guarantees can enhance system resilience to these distributional shifts.

In our work, we bridge the gap between these two research streams by designing an algorithm that retains the robustness of primal-dual approximations while improving average campaign performance through online learning. Specifically, we approach the budget management problem from an online learning perspective, where the platform (as the online learner) dynamically allocates shares of overall impressions to different campaigns at each time step informed by past campaign performance.

We introduce a refined primal-dual algorithm that leverages these provisional shares to enhance allocative efficiency while preserving worst-case guarantees. Intuitively, our algorithm acts as a *non-stationary environment* that interacts with the online learner. This environment observes the learner's strategy but adapts its own allocation strategy to maintain worst-case guarantees. If misaligned, this interaction can create a feedback loop that negatively impacts system performance. To address this, we propose a method to break the feedback loop by aligning the strategies of both the learner and the environment, thereby improving allocative efficiency while ensuring robustness against distributional changes.

To evaluate our approach, we conduct experiments on a large-scale dataset comprising 320 million sponsored requests from an industrial RS operating across 20 markets. Our results demonstrate that our method significantly enhances allocative efficiency and return-on-investment (ROI) for campaigns in all 20 markets.

The key contributions of this work include:

- A novel integration of the primal-dual approximation technique with online learning, improving allocative efficiency while maintaining robustness.
- A systematic resolution of the feedback loop between the online learner's strategy and the primal-dual environment, enhancing performance stability.
- Extensive evaluations on synthetic and industrial datasets show up to a 4.7% increase in total allocated value and an 8% boost in ROI, demonstrating the effectiveness of our approach across diverse markets.

The rest of the paper is organized as follows: Section 2 reviews related work. Section 3 formalizes our model and outlines primal-dual approximation technique. Section 4 presents our algorithm, while Section 5 discusses experimental results. Finally, Section 6 concludes with future research directions.

2 RELATED WORK

Constrained Optimisation for Allocation Problems. Using constrained optimisation techniques to solve allocation problems is a well-established approach. Brantley et. al. [7] employed a control-theoretic method to rank items in RS under long-term constraints. Unlike our work, their approach does not incorporate robustness guarantees, instead optimising performance on average. Similarly, ranking with fairness or diversity constraints has been explored in [9–11]. These methods, however, focus on constructing a single ranking rather than addressing the temporal allocation problem. Balseiro at. al. [12–14] examined budget management in ad auctions. In contrast, our work is not confined to a specific auction setting and does not emphasize incentives or equilibrium analysis.

Approximation Algorithms. Agrawal et. al. [15] model the ad allocation problem as bipartite graph matching and designed an algorithm that achieves near-optimal solutions. However, their approach is limited to the offline setting, whereas our focus is on online optimisation. In [2], the authors consider the temporal allocation problem from an online knapsack packing perspective, deriving a primal-dual approximation algorithm to solve it. They later enhance their original method by incorporating stochastic forecast into the primal-dual algorithm [1]. While these algorithms offer strong theoretical guarantees, we argue that their average performance is highly dependent on the quality of the forecast, which can lead to arbitrarily poor outcomes. Our work builds on these approaches to ensure consistently strong average performance.

Online Learning Approaches. Recent research has shifted focus from worst-case performance to improving average performance. Spaeh at. al. [16] introduced an online ad allocation algorithm that remains robust against inaccurate predictions. Our work generalizes this by developing a more comprehensive constrained online learning framework that incorporates budget constraints while maintaining worst-case guarantees. Primal-dual approaches with predictions have been explored in [3]. In contrast to our work, the authors focused on leveraging uncertainty in the impression-level predictions rather than enhancing the performance of primal-dual technique with campaign-level decisioning. In [8], the authors proposed an online learning heuristic for budget throttling. This is different from the bid pacing approach adopted in our work, and lacks the theoretical guarantees. Ding et. al. [4] investigate multi-armed bandits for budget pacing. Contrary to our work, they adopt the perspective of a single bidder selecting the bid (i.e., arm) instead of trying to allocate multiple campaigns in the most efficient way.

3 FORMAL MODEL & PRELIMINARIES

In this section, we define the key components of our budget allocation model for sponsored campaigns on an online platform.

Let $N \in \mathbb{N}$ be the number of users arriving on the platform over a fixed time horizon $T > 0$, and let $M \in \mathbb{N}$ denote the number of sponsored campaigns. We assume that each user is shown a single sponsored slot, which can be allocated to one of the of M campaigns. Each campaign j (where $j = 1, ..., M$) has a fixed *budget* $B_j > 0$. The *value* of showing campaign j to user i is represented by $v_{ij} \in [0, 1]$, a normalized score that can be either self-reported through a bidding interface or estimated by the platform.

We define $\pi_{ij} \in [0, 1]$ as the probability that user i clicks on the content of campaign j when it is displayed.[1] Since users arrive at different time steps $t = 1, ..., T$, we introduce π^t_{ij}, which represents the click probability at time t if the user is present at that time:

$$\pi^t_{ij} = \begin{cases} \pi_{ij}, & \text{if user } i \text{ arrives at time } t \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

We introduce binary decision variables $a^t_{ij} \in \{0, 1\}$ to indicate whether campaign j is allocated to user i at time t. The total number of impressions allocated to campaign j up to time t is given by

$$Q^t_j = \sum_{\tau=1}^{t} \sum_{i=1}^{N} a^t_{ij}.$$

The *budget utilisation* of campaign j at time t is then defined as:

$$g^t_j = \frac{Q^t_j}{B_j}.$$

Finally, if campaign j is allocated to user i, the campaign incurs a fixed cost $p \in \mathbb{R}_{\geq 0}$, which is paid to the platform.[2]

3.1 Offline Setting: Optimal Allocation Problem

We now formally define the optimal allocation problem. Our goal is to design an allocation algorithm that maximizes the total expected

[1] Practically, we can build a parametric model (e.g., DNN) to estimate these probabilities.
[2] In the auction setting, we can let the price being equal to the reported value v_{ij}.

Data: Pacing coefficients λ_j, $j = 1, \ldots, M$
1 $a^t_{ij} \leftarrow 0$ for all i, j, t
2 **for** $t = 1, \ldots T$ **do**
3   **for** $i = 1, \ldots, N$ **do**
4         $b^t_{ij} \leftarrow \pi^t_{ij} v_{ij} - \lambda_j, \quad \forall j = 1, \ldots, M$
5         $j^* \leftarrow \arg\max_j \{b^t_{ij} \ : \sum_{t=1}^T \sum_{i=1}^N a^t_{ij} < B_j\}$
6   **if** $b^t_{ij^*} < 0$ **then**
7   Do not allocate user i
8   **else**
9             $a^t_{ij^*} \leftarrow 1$ //Allocate the highest bidder
10 **return** $a^t_{ij}, \forall t, i, j$

Algorithm 1: Optimal Offline Allocation Algorithm

sponsored value while satisfying budget and feasibility constraints:[3]

$$\max_{a^t_{ij} \in \{0,1\}} \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{j=1}^{M} \pi^t_{ij} v_{ij} a^t_{ij} \tag{2}$$

$$\text{s.t.,} \quad \sum_{t=1}^{T} \sum_{i=1}^{N} p a^t_{ij} \leq B_j \quad \forall j = 1, \ldots, M, \tag{3}$$

$$\sum_{j=1}^{M} a^t_{ij} \leq 1 \quad \forall i = 1, \ldots, N, \forall t = 1, \ldots, T. \tag{4}$$

The objective defined in Equation (2) seeks to maximise the total expected sponsored value allocated over the time horizon T. Constraint (3) ensures that campaign budgets are not exceeded. Without loss of generality, we set $p = 1$ by normalizing the budgets accordingly. Constraint (4) ensures that each user i receives at most one campaign allocation at any given time t. We will now illustrate the optimal solution to Problem (2)-(4) in an *offline* setting, where all constraints (3)-(4) are known in advance.

Algorithm 1 implements a simple bidding algorithm that takes *pacing coefficients* λ_j as input and determines the allocation a^t_{ij} for all users, campaigns, and time steps. Initially, all allocation variables a^t_{ij} are zero (line 1). As the users arrive (lines 2-3), the algorithm 1 assigns each user to the campaign j^* with the highest *paced bid* (lines 4-5):

$$b^t_{ij} = \pi^t_{ij} v_{ij} - \lambda_j.$$

A campaign is allocated as long as its bid is non-negative and its budget has not been exceeded (lines 6-9).

We now demonstrate that if the pacing coefficients λ_j are equal to the Lagrangian multipliers of the budget Constraint (3), then Algorithm 1 provides an optimal solution to the allocation problem.

Statement 1. *Let λ_j, $j = 1, \ldots, M$ be the Lagrangian multipliers of Constraints (3). Then, Algorithm 1 solves Problem (2)-(4) optimally.*

Proof. We provide the proof in the Appendix. □

In practice, both users and campaigns arrive dynamically over time. As a result, the full set of constraints (3)-(4) is not known in advance, making it impossible to solve Problem (2)-(4) optimally. In the following section, we explore how the existing primal-dual approximation technique [1] can be leveraged to derive a close-to-optimal solution in this *online setting*.

[3] This formulation is common and similar to the one presented in [2].

3.2 Online Setting: Primal-Dual Approximation

Data: Step size Δ_a, $\Pi \leftarrow \min_{i,j,t}\{\pi^t_{ij} v_{ij}\}$
1 Set $a^t_{ij}, \lambda_j, \gamma_{it} \leftarrow 0$ for all i, j, t; $A \leftarrow \log|1 + \frac{2}{\Pi} \max_{ijt} \pi^t_{ij} v_{ij}|$
2 **for** $t = 1, \ldots T$ **do**
3   **for** $i = 1, \ldots, N$ **do**
4         $s_j \leftarrow -\pi^t_{ij} v_{ij}$ for all $j = 1, \ldots, M$
5   **while** $\exists j : s_j < 0$ **do**
6             $a^t_{ij} \leftarrow a^t_{ij} + \Delta_a, \ \forall j = 1, \ldots, M$
7             $\lambda_j \leftarrow \frac{\Pi}{2} \exp\{\frac{A}{B_j} \sum_{t=1}^T \sum_{i=1}^N a^t_{ij}\} - \frac{\Pi}{2}$
8             $\gamma_{it} \leftarrow \frac{\Pi}{2} \exp\{A \sum_{j=1}^M a^t_{ij}\} - \frac{\Pi}{2}$
9             $s_j \leftarrow \lambda_j + \gamma_{it} - \pi^t_{ij} v_{ij}$
10         $j^* \leftarrow \arg\min_j \{s_j : \sum_{t=1}^T \sum_{i=1}^N \pi^t_{ij} v_{ij} a^t_{ij} < B_j, a^t_{ij} \leq 1\}$
11         $a^t_{ij^*} \leftarrow 1; \quad a^t_{ij} \leftarrow 0 \ \ \forall j \neq j^*$
12 **return** $a^t_{ij}, \forall t, i, j$

Algorithm 2: Online Primal-Dual Allocation Algorithm

Algorithm 2 presents an approximate solution for Problem (2)-(4) in the online setting. It takes as input the step size Δ_a and the minimal acceptable score, defined as $\Pi = \min_{i,j,t}\{\pi^t_{ij} v_{ij}\}$, and outputs the allocation a^t_{ij}. Initially, all primal and dual variables are set to zero (line 1). As users arrive (lines 2-3), Algorithm 2 incrementally increases the allocation variables a^t_{ij} for all campaigns $j = 1, \ldots, M$ (line 6), while simultaneously updating the dual variables λ_j and γ_i (lines 7-8). Consequently, the respective *slack variables* s_j increase (line 9) until they all become non-negative (line 5). Finally, the algorithm allocates the campaign with the smallest slack $s_j = 0$, provided that the budget constraint is not violated (lines 10-11).

From line 7 of Algorithm 2, we observe that as the budget utilisation g^t_j of campaign j increases, the corresponding pacing coefficient λ_j also increases:

$$\lambda^t_j \leftarrow \frac{\Pi}{2} \exp\left\{A \cdot \underbrace{\frac{1}{B_j} \sum_{t=1}^{T} \sum_{i=1}^{N} a^t_{ij}}_{\text{budget utilisation } g^t_j}\right\} - \frac{\Pi}{2}, \quad \forall j = 1, \ldots, M. \tag{5}$$

As a result, the slack variable s_j for campaign j increases in future time steps, reducing the likelihood of the campaign "winning" impressions. This behaviour is intuitive: when a campaign's budget is nearly exhausted, it becomes more selective, prioritizing only high-value opportunities. Figure 1 visualizes this dependency.

We now show that Algorithm 2 is $O\big(\log|1+\frac{2}{\Pi}\max_{i,j,t}\{\pi^t_{ij} v_{ij}\}|\big)$-competitive.

Statement 2. *Algorithm 2 outputs an approximately optimal solution to Problem (2)-(4) that is $O\big(\log|1 + \frac{2}{\Pi}\max_{i,j,t}\{\pi^t_{ij} v_{ij}\}|\big)$-competitive.*

Proof. We provide the proof in the Appendix. □

4 ENHANCING PRIMAL-DUAL SOLUTION WITH ONLINE LEARNING

While Algorithm 2, discussed in Section 3.2 provides an approximately optimal allocation, its guarantees (Statement 2) are based on *worst-case* guarantees. These guarantees do not depend on the

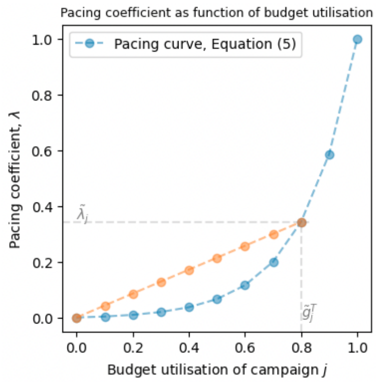


Figure 1: Blue: Pacing curve from Equation (5). As the budget utilisation increases, the pacing coefficient gets larger, filtering out the low-valued opportunities. Orange: Linear segment of the pacing curve from our refined Algorithm 3. If the budget utilisation is smaller than its anticipated value $\tilde{g}_j^T$, the pacing coefficient increases faster.

specific distribution of the setting, making them robust but potentially conservative. In practice, however, the Budget Management System (BMS) can continuously observe the performance of ongoing campaigns. This real-time feedback allows the system to refine its allocation strategy using online learning techniques, ultimately improving overall performance *on average*.

In this section, we introduce an online learning algorithm that preserves the worst-case guarantees established in Section 3.2 while enhancing the system's average performance. This is achieved through a modification of the polynomial weights update (PWU) algorithm [17–19].

To achieve this, we first present a refined version of Algorithm 2, which incorporates an additional input representing the anticipated (or provisional) performance of each campaign j. This refinement ensures that the computed allocation is at least as effective as the one found by Algorithm 2. Next, we reformulate Problem (2)-(4) from an online learning perspective [18] and demonstrate how to derive an online learning strategy that aligns with the allocation decisions of our refined primal-dual algorithm.

4.1 Refined Primal-Dual Approximation

We extend Algorithm 2 by incorporating an additional input representing the anticipated performance of each campaign. This enables the algorithm to produce an allocation that is at least as effective as the one derived from the standard primal-dual approach discussed in Section 3.2. Algorithm 3 outlines our modified approach.

We define $\tilde{g}_j^T \in [0,1]$ as the *anticipated budget utilisation* of campaign j by time T (see input data of Algorithm 3). The algorithm begins by initializing all primal and dual variables to zero (line 1). In line 2, we compute the *anticipated pacing coefficient* $\tilde{\lambda}_j$ for each campaign j based on the anticipated budget utilisation $\tilde{g}_j^T$ and Equation (5). As the users arrive (lines 3-4) we update the actual utilisation g_j^t for each campaign j (line 5), and initialise the slack variables s_j (line 6). We then iteratively increase the allocation variables a_{ij}^t (lines 7-8) while recalculating the dual variables λ_j^t and γ_{it}. The key difference from Algorithm 2 lies in the update rule for these dual variables. Specifically, if the utilization g_j^t of campaign j remains below its anticipated value $\tilde{g}_j^T$, we increase the corresponding variable λ_j^t linearly (line 10), see the orange segment in Figure 1. Otherwise, we revert to the update rule of Algorithm 2 (lines 13-14). [4] Finally, we allocate campaign j^* with the smallest slack value $s_{j^*} = 0$ (lines 16-17).

Data: Step size Δ_a, anticipated budget utilisation $\tilde{g}_j^T \quad \forall j = 1, \dots, M$

1 Initialise $a_{ij}^t \leftarrow 0, \lambda_j \leftarrow 0, \gamma_{it} \leftarrow 0$ for all i, j, t

2 $\tilde{\lambda}_j \leftarrow \frac{\Pi}{2} \exp\{A\tilde{g}_j^T\} - \frac{\Pi}{2}, \ \forall j = 1, \dots, M$ # Anticipated pacing coeff

3 **for** $t = 1, \dots T$ **do**

4   **for** $i = 1, \dots, N$ **do**

5         $g_j^t \leftarrow \frac{1}{B_j} \sum_{\tau=1}^{t} \sum_{i=1}^{N} a_{ij}^{\tau}, \ \forall j = 1, \dots, M$

6         $s_j \leftarrow -\pi_{ij}^t v_{ij}$ for all $j = 1, \dots, M$

7   **while** $\exists j : s_j < 0$ **do**

8             $a_{ij}^t \leftarrow a_{ij}^t + \Delta_a, \ \forall j = 1, \dots, M$

9   **if** $g_j^t < \tilde{g}_j^T$ **then**

10                 $\lambda_j^t \leftarrow \lambda_j + \frac{\tilde{\lambda}_j}{\tilde{g}_j^T} \frac{a_{ij}^t}{B_j}$

11                 $\gamma_{it} \leftarrow (\pi_{ij}^t v_{ij} - \tilde{\lambda}_j) a_{ij}^t$

12   **else**

13                 $\lambda_j^t \leftarrow \frac{\Pi}{2} \exp\{\frac{A}{B_j} \sum_{t=1}^{T} \sum_{i=1}^{N} a_{ij}^t\} - \frac{\Pi}{2}$

14                 $\gamma_{it} \leftarrow \frac{\Pi}{2} \exp\{A \sum_{j=1}^{M} a_{ij}^t\} - \frac{\Pi}{2}$

15             $s_j \leftarrow \lambda_j^t + \gamma_{it} - \pi_{ij}^t v_{ij}$

16         $j^* \leftarrow \arg\min_j \{s_j : \sum_{t=1}^{T} \sum_{i=1}^{N} \pi_{ij}^t v_{ij} a_{ij}^t < B_j, a_{ij}^t \leq 1\}$

17         $a_{ij^*}^t \leftarrow 1; \quad a_{ij}^t \leftarrow 0 \ \forall j \neq j^*; \quad \lambda_{j^*} \leftarrow \lambda_j^t$

18 **return** $a_{ij}^t, \forall t, i, j$

Algorithm 3: Refined Primal-Dual Allocation Algorithm

Definition 1. *We say that the anticipated budget utilisation $\tilde{g}_j^T$ is **compatible** with Algorithm 3 if it does not exceed the realised budget utilisation g_j^T produced by the algorithm, i.e., $\tilde{g}_j^T \leq g_j^T$.*

The following statement demonstrates that Algorithm 3 is at least as effective as Algorithm 2:

Statement 3. *For any $\Pi \geq 0.064$ and any compatible anticipated budget utilisation $\tilde{g}_j^T$, $j = 1, \dots, M$ Algorithm 3 computes an approximately optimal solution to Problem (2)-(4) that is at least as good as the one produced by Algorithm 2.*

Proof. We provide the proof in the Appendix. □

We note that the threshold $\Pi \geq 0.064$ is not overly restrictive, as the low-value opportunities are typically filtered out during the candidate generation step to ensure a positive user experience.

4.2 Online Learning Perspective

We now analyze Problem (2)-(4) from an online learning perspective [18, 20]. In this framework, we assume that at the beginning of each time step t the *learner* (i.e., the BMS) selects the shares of impressions to allocate across different campaigns. This decision is informed by the observed performance of campaigns from previous time steps $t-1, t-2, \dots,$ and so on.

The *environment* [20], as defined by Algorithm 3, observes the BMS's allocation choices and generates the actual distribution of

[4] Observe, that this approach is different from the one discussed in [1] in the shape of the pacing curve.

impressions along with the corresponding rewards. The objective of the BMS is to maximize the expected total reward over time T.

We define the system's set of *actions* as the set of campaigns, $j = 1, ..., M$, and assume that at each step t, the system selects a *strategy* $p^t = (p_1^t, ..., p_M^t) \in \Delta(M)$, representing a probability distribution over these actions. Here, $p_j^t \in [0, 1]$ denotes the share of impressions the system intends to allocate to campaign j at time t. Notably, the share p_j^t assigned to campaign j uniquely determines the anticipated budget utilisation $\tilde{g}_j^\tau(p_j^t)$ for that campaign at future time steps $\tau = t, ..., T$:

$$\tilde{g}_j^\tau(p_j^t) = \min\left\{1, \frac{Q_j^{t-1} + p_j^t \cdot N \cdot (\tau - t + 1)}{B_j}\right\}, \quad \forall j = 1, ..., M. \quad (6)$$

Here, Q_j^{t-1} is the number of impressions already allocated to campaign j by time t (Section 3), and $p_j^t N(\tau - t + 1)$ denotes the expected number of impressions that will be allocated to j by time τ.

We now extend Definition 1 to define *compatible strategies*:

Definition 2. *We say that the strategy p^t of the learner is* ***compatible*** *with Algorithm 3 if the respective anticipated budget utilisation $\tilde{g}_j^T(p^t)$ is compatible with Algorithm 3, i.e.,*

$$\tilde{g}_j^T(p_j^t) \leq \frac{1}{B_j}\sum_{\tau=1}^{T}\sum_{i=1}^{N} a_{ij}^\tau, \quad \forall j = 1, ..., M. \quad (7)$$

The *environment* can now be formally defined by Algorithm 3, parametrized by $\tilde{g}_j^T(p_j^t)$. This environment observes the strategy p^t chosen by the BMS, along with the corresponding anticipated utilisation $\tilde{g}_j^T(p_j^t)$, and then allocates impressions according to the rules of Algorithm 3. This process generates an *allocative distribution* $q^t(p^t) = (q_1^t(p^t), ..., q_M^t(p^t)) \in \Delta(M)$ over the M campaigns, as well as the *rewards* $r^t(p^t, q^t) = (r_1^t, ..., r_M^t) \in [0, 1]^M$:

$$q_j^t = \frac{\sum_{i=1}^{N} a_{ij}^t}{N} \qquad \forall j = 1, ..., M,\ \forall t = 1, ..., T, \quad (8)$$

$$r_j^t = \frac{\sum_{i=1}^{N} \pi_{ij}^t v_{ij} a_{ij}^t}{\sum_{i=1}^{N} a_{ij}^t} \quad \forall j = 1, ..., M,\ \forall t = 1, ..., T, \quad (9)$$

Here, a_{ij}^t is computed by Algorithm 3, parametrised by $\tilde{g}^T(p^t)$. Notably, similar to traditional online learning settings [18, 20], the environment in our domain is non-stationary. As a result, the rewards r_j^t may depend on the distribution p^t chosen by the BMS, and can fluctuate over time.

We can now reformulate the objective of Problem (2)-(4) from the learner's perspective:

$$\max_{p^t \in \Delta(M)} \sum_{t=1}^{T}\sum_{j=1}^{M} q_j^t(p^t) r_j^t(p^t, q^t(p^t)). \quad (10)$$

In other words, the goal of the BMS is to determine a distribution p^t of impressions across campaigns that maximise the expected total allocated value, as determined by Algorithm 3, parametrised by $\tilde{g}^T(p^t)$. Crucially, while the BMS selects p^t, the actual allocative distribution q^t may differ, as it is fully dictated by the environment.

Definition 3. *We say that strategy p^t is* ***consistent*** *if $p_j^t = q_j^t$ for all $j = 1, ..., M$.*

In simple terms, the strategy of the BMS is consistent if it aligns with the allocative distribution determined by the environment. We will now argue that to maintain the worst-case guarantees established in Statement 3, the system must focus solely on selecting consistent strategies that maximise the objective in Equation (10).

First, observe that maximising (10) is equivalent to optimising:

$$\max_{p^t \in \Delta(M)} \log\Big(\sum_{t=1}^{T}\frac{1}{T}\sum_{j=1}^{M} q_j^t(p^t) r_j^t\big(p^t, q^t(p^t)\big)\Big). \quad (11)$$

Now, instead of optimising Objective (11) directly, we derive its lower bound and optimize that bound instead:[5]

$$\max_{p^t \in \Delta(M)} \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{M} p_j^t \Big[\underbrace{\log r_j^t(p^t, q^t(p^t)) + \log q_j^t(p^t)}_{\text{Instantaneous reward } \tilde{r}_j^t.}\Big] + H(p^t), \quad (12)$$

where $H(p^t)$ represents the entropy of p^t, and $\tilde{r}_j^t$ is the *instantaneous reward* at time t (see Appendix A.1.1 for derivation of Equation (12)). Unlike Problem (10), where the distribution over actions q_j^t is determined by the environment, formulation (12) closely follows the standard online learning framework [20]. Here, the probabilities p_j^t are directly specified by the learner, and the objective is regularized by the entropy term $H(p^t)$.

Increasing p_j^t has two opposing effects on the instantaneous reward. On the one hand, increasing p_j^t leads to a weak increase in the anticipated utilisation $\tilde{g}_j^T(p_j^t)$, see Equation (6). This results in a weakly higher anticipated coefficient $\tilde{\lambda}_j$ (line 2 of Algorithm 3), leading to a weak increase in the average allocated value r_j^t for campaign j (Equation 9). Consequently, the term $\log r_j^t$ in Equation (12) must increase. On the other hand, increasing $\tilde{\lambda}_j$ reduces the probability q_j^t of campaign j being allocated (as shown in Algorithm 3). As a result, the term $\log q_j^t$ in Equation (12) must decrease.

Statement 4. *If there exists an optimal solution to Problem (12) that is compatible with Algorithm 3, then there also exists an optimal solution p^t that is consistent.*

Proof. We provide the proof in the Appendix. □

Statement 4 allows us to restrict the search space to only consistent strategies p^t, thereby simplifying the overall optimisation problem. As a result, we can rewrite Objective (12) as follows:

$$\max_{p^t \in \Delta(M)} \sum_{t=1}^{T}\sum_{j=1}^{M} p_j^t \log r_j^t\big(p^t, q^t(p^t)\big) - D_{KL}(p^t || q^t) \quad (13)$$

$$\text{s.t.,}\ \ p_j^t = q_j^t(p^t) \ \ \forall t = 1, ..., T,\ \forall j = 1, ..., M. \quad (14)$$

Importantly, Constraint (14) implies that the solution p^t must be a *fixed point* of the allocative distribution $q^t(p^t)$. Once this constraint is satisfied, the KL-divergence term in the objective disappears, thereby increasing the overall objective value.

[5] A similar approach is used, for example, for derivation of Expectation Maximisation algorithm [21].

4.3 Constrained Polynomial Weights Update

We now illustrate how to compute strategy p^t that is approximately consistent. Recall from Algorithm 3 that campaign j "wins" impression i if:

$$\pi^t_{ij} v_{ij} - \lambda^t_j \geq \pi^t_{i\ell} v_{i\ell} - \lambda^t_\ell$$

for any campaign $\ell \neq j$. Since at the beginning of step t, the BMS has no prior knowledge of the incoming scores $\pi^t_{ij} v_{ij}$, we assume a prior belief of BMS $\pi^t_{ij} v_{ij} = r^{t-1}_j + \epsilon^t_{ij}$, where r^{t-1}_j is the average reward from the previous time step (Equation (9)) and $\epsilon^t_{ij} \sim \mathit{Gumbel}(.)$ represents random noise. From this, the allocative distribution q^t generated by Algorithm 3 as [22]:

$$q^t_j = \frac{\exp\{r^{t-1}_j - \lambda^t_j\}}{\sum_{\ell=1}^{M} \exp\{r^{t-1}_\ell - \lambda^t_\ell\}}, \quad j = 1, ..., M. \tag{15}$$

Thus, the consistency constraint (Equation 14) becomes:

$$p^t_j = \frac{\exp\{r^{t-1}_j - \lambda^t_j(\tilde{g}^t_j(p^t_j))\}}{\sum_{\ell=1}^{M} \exp\{r^{t-1}_\ell - \lambda^t_\ell(\tilde{g}^t_\ell(p^t_\ell))\}}, \quad j = 1, ..., M, \tag{16}$$

where $\lambda^t_j(\tilde{g}^t_j(p^t_j))$ is computed using Equations (5) and (6).

This strategy mimics the standard PWU algorithm applied to Problem (12) [20].[6] However, a key distinction is that p^t_j and the rewards in Equation (16) are interlinked, creating a feedback loop between the learner's strategy of the learner and the environment's generated reward.

Breaking this feedback loop is essential to ensuring a strategy that is consistent with Algorithm 3. Doing so would provide: a) robustness guarantees for the primal-dual approximation, b) enforcement of budget and feasibility constraints, and c) improved total allocated value in Problem (13)-(14). Algorithm 4 demonstrates how to break this feedback loop using fixed point iteration.

Algorithm 4 starts by initialising the anticipated budget utilisation $\tilde{g}_j$ for each campaign j using its current (ongoing) utilisation. It then iteratively refines this estimate by simulating interactions with the environment. At each time step τ, the algorithm updates the pacing coefficients λ_j following lines 10-13 of Algorithm 3. It then recomputes the learner's strategy p^τ_j and the expected number of allocations $p^\tau_j \cdot N$. Lastly, it estimates the final budget utilisation g^T_j at the end of the time horizon T. The algorithm terminates once g^T_j converges to $\tilde{g}_j$, returning the final strategy p^τ and the anticipated utilisation $\tilde{g}^t_j$.

Linking CPWU with Primal-Dual Approach. We now present our final Primal-Dual Approximation algorithm with Constrained Polynomial Weights Update. Algorithm 5 illustrates our overall approach.

The algorithm takes the tolerance level and budgets as input, and iteratively allocates the impression opportunities. Initially, in line 1, all pacing coefficients λ_j are set to zero. Since no prior information about the quality of the campaigns is available, the initial rewards r^0_j are also set to zero. At each time step t (lines 2-7), the algorithm first computes the consistent strategy p^t and the corresponding anticipated budget utilisation. This computation relies on the known ongoing budget utilisations, observed qualities r_j, and pacing coefficients λ_j (see line 3).

Next, the algorithm initialises the primal-dual environment with the computed anticipated budget utilisation $\tilde{g}$, which is used to allocate the arriving impressions through our refined version of the primal-dual approximation. Following this, the quantities Q_j and quality scores r_j are updated, along with the pacing multipliers. The algorithm then proceeds to the next iteration, $t + 1$.

[6] Here, $(r^{t-1}_j - \lambda_j(p_j))$ is correlated with the instantaneous reward $\tilde{r}^{t-1}_j$ in Equation (12). Indeed, for small values of r^t_j we have $\log r^t_j \approx 1 + r^t_j$; Furthermore, as we increase p^t_j, the respective $\lambda_j(\tilde{g}^t_j(p^t_j))$ increases, and hence, $\log(q^t_j)$ must decrease.

Data: Tolerance Δ; $Q^{t-1}_j, r^{t-1}_j, \lambda^t_j, \forall j = 1, ..., M$
$\tilde{g}_j \leftarrow Q^{t-1}_j / B_j$ for all $j = 1, ..., M$
do
  $\tilde{\lambda}_j \leftarrow \frac{\Pi}{2} \exp\{A\tilde{g}_j\} - \frac{\Pi}{2}$ for all $j = 1, ..., M$
  $\lambda_j \leftarrow \lambda^t_j$ for all $j = 1, ..., M$
  for $\tau = t, ...T$ **do**
    $p^\tau_j \leftarrow \frac{\exp\{r^{t-1}_j - \lambda_j\}}{\sum_{\ell=1}^{M} \exp\{r^{t-1}_\ell - \lambda_j\}}$ for all $j = 1, ..., M$
    $Q^\tau_j \leftarrow Q^{\tau-1}_j + p^\tau_j \cdot N$ for all $j = 1, ..., M$
    $g^\tau_j \leftarrow \min\left\{1, Q^\tau_j / B_j\right\}$, for all $j = 1, ..., M$
    if $g^\tau_j < \tilde{g}_j$ **then**
      $\lambda_j \leftarrow g^\tau_j \cdot \frac{\tilde{\lambda}_j}{\tilde{g}_j}$ for all $j = 1, ..., M$
    else
      $\lambda_j \leftarrow \frac{\Pi}{2} \exp\{A g^\tau_j\} - \frac{\Pi}{2}$ for all $j = 1, ..., M$
  $\tilde{g}_j \leftarrow \frac{\tilde{g}_j + g^T_j}{2}$
while $\max_j\{|\tilde{g}_j - g^T_j|\} \geq \Delta$ **do**
return $p^\tau_j, \forall j = 1, ..., M, \forall \tau = t, ..., T; \tilde{g}_j, \forall j = 1, ..., M.$

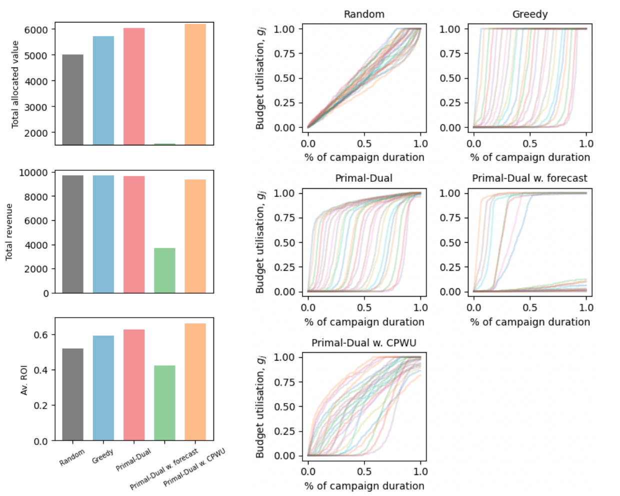
Algorithm 4: Computation of consistent strategies via fixed point iteration.

Data: Tolerance level Δ, Budgets B_j, $j = 1, ..., M$
1 Initialise $Q^0_j \leftarrow 0$, $r^0_j \leftarrow 0$, $\lambda_j \leftarrow 0$ for all $j = 1, ..., M$
2 **for** $t = 1, ...T$ **do**
3   $p^t, \tilde{g} \leftarrow$ Algorithm 4$(\Delta; Q^{t-1}_1 ... Q^{t-1}_M; r^{t-1}_1 ... r^{t-1}_M; \lambda_1 ... \lambda_M)$
4   $a^t_{ij} \leftarrow$ Algorithm 3$(\tilde{g}^t_1, ..., \tilde{g}^t_M)$, for all $j = 1, ..., M$, $i = 1, ..., N$
5   $Q^t_j \leftarrow Q^{t-1}_j + \sum_{i=1}^{N} a^t_{ij} \quad \forall j = 1, ..., M$
6   $r^t_j = \frac{\sum_{i=1}^{N} \pi^t_{ij} v_{ij} a^t_{ij}}{\sum_{i=1}^{N} a^t_{ij}} \quad \forall j = 1, ..., M$
7   **if** $Q^t_j / B_j < \tilde{g}_j$ **then**
      $\lambda_j \leftarrow Q^t_j / B_j \cdot \frac{\tilde{\lambda}_j}{\tilde{g}_j}$ for all $j = 1, ..., M$
    else
      $\lambda_j \leftarrow \frac{\Pi}{2} \exp\{A \cdot Q^t_j / B_j\} - \frac{\Pi}{2}$ for all $j = 1, ..., M$
8 **return** $a^t_{ij}, \forall t, i, j$

Algorithm 5: Primal-Dual Allocation with Constrained Polynomial Weights Updates.

5 EVALUATION

To gain a better intuition about our algorithm's performance, we first generate a synthetic dataset and we compare our approach against several baselines using synthetic data. We then evaluate our method on a real-world dataset comprising 320 million impression



(a) Surplus Distribution **(b) Pacing quality**

Figure 2: Left: Total value, revenue and av. ROI of the campaigns. Right: Budget utilisation of the campaigns over time.

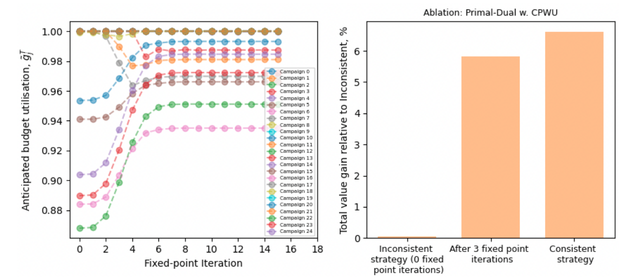


Figure 3: Left: Convergence of Algorithm 4 at $t = 0$. Right (Ablation study): Total allocated value gains when using consistent strategies relative to the inconsistent ones.

opportunities and 1.5 thousand promotional campaigns, managed by a major digital content provider in 20 existing markets.

5.1 Synthetic Data

Simulation Setup. We set the number of users to $N = 1000$, the number of campaigns to $M = 25$, and the time horizon to $T = 21$. At each step $t = 1, ..., T$, we randomly sample 50% of the N users, who then arrive sequentially on the platform at time t. We generate scores from a Beta distribution, $\pi_{ij} \sim Beta(M-j+1, j+1)$, ensuring that campaigns with higher indices j are less "clickable" (see Figure 7 in the Appendix). For simplicity, we set $v_{ij} = 1$ for all $i = 1, ..., N$, $\forall j = 1, ..., M$. Finally, campaign budgets B_j are sampled from a uniform distribution $B_j \sim U\{300, 500\}$.

Baselines & Metrics. We compare our approach against the following baselines:

- *Random*. A weak baseline that assigns sponsored items to users uniformly at random.
- *Greedy*. For each user i, this method selects the campaign j with the highest score $\pi_{ij}^t v_{ij}$ for allocation.
- *Primal-Dual* [23]. A standard Primal-Dual approach for online knapsack packing problem, as described in Algorithm 2.
- *Primal-Dual w. forecast* [1]. A variation of the Primal-Dual approach that incorporates stochastic forecasts as proposed by Buchbinder et. al. [1].
- *Primal-Dual w. CPWU*. Our proposed primal-dual approach with constrained polynomial weights update (see Algorithm 5).

To evaluate these algorithms, we primarily measure the total allocated value achieved, corresponding to our optimization objective (2)). Additionally, we assess the total revenue collected by the BMS, computed as the total budgets spent, i.e., $\sum_{j=1}^{M} Q_j^T$. Finally, we report the average ROI, defined as the mean ratio of total value to the budget spent per campaign, i.e., $\frac{1}{M}\sum_{j=1}^{M}(\sum_{t=1}^{T}\sum_{i=1}^{N}\pi_{ij}^t v_{ij} a_{ij}^t)/Q_j^T$.

Results. Figure 2 presents the results on our synthetic data. First, we observe that *Primal-Dual w. CPWU* (orange) achieves the highest allocated value among all baselines (Figure 2a, top). Specifically, our approach improves the total allocated value by +8.4% over the *Greedy* baseline (blue), compared to only a +5.6% improvement of the standard *Primal-Dual* (red) over *Greedy*.

Interestingly, while *Primal-Dual w. forecast* (green) offers better worst-case guarantees than *Primal-Dual*, its average performance is lower than even the weakest *Random* baseline (gray). This poor performance stems from the inconsistency of the forecast (Section 4.2). In our setting, the total supply is approximately equal to the total demand, meaning the anticipated budget utilisation should be $\tilde{g}_j^T \approx 1$ for all campaigns. However, setting $\tilde{g}_j^T = 1$ leads to very large pacing coefficients λ_j, causing many scores $\pi_{ij}^t v_{ij}$ to fall below λ_j [1]. As a result, fewer campaigns get allocated, reducing the overall budget spent (Figure 2a, centre), and ultimately lowering the total allocated value.[7]

To further highlight the importance of using consistent strategies, we examine the convergence of Algorithm 4 at time $t = 0$ in Figure 3 (left). The results show that the budget utilisation under a consistent strategy (i.e., after convergence) differs significantly from its initial, inconsistent estimate. As a result, *Primal-Dual w. CPWU*, when employing a consistent strategy, achieves a 6.5% higher total value compared to using inconsistent strategies (Figure 3 (right).

From Figure 2a (centre), we also observe a slight negative impact (within -2.5%) of all pacing mechanisms on revenue compared to the *Greedy* baseline. This is expected, as the objective of Problem (2)-(4) is to maximise the total value, not revenue. Finally, we find that the average ROI of *Primal-Dual w. CPWU* outperforms all other baselines, achieving +12.1% over *Greedy*–nearly doubling the +6.2% gains achieved by the standard *Primal-Dual* over *Greedy*.

We now analyze the pacing quality of all algorithms. Figure 2b shows how budget utilisation evolves over time (axes are normalised). Ideally, the budget should follow a linear diagonal trend: spending too quickly may cause missed opportunities later, while slow spending can lead underutilised budget and fewer allocations.

We observe that *Random* distributes spending almost linearly. In contrast, *Greedy* prioritizes allocating the highest-scoring campaigns first, leading to nearly full utilization of these campaigns in less than half their duration. This is suboptimal, as a slower

[7] We further illustrate this point by comparing the values of the pacing coefficients λ_j of *Primal-Dual w. forecast* vs. *Primal-Dual w. CPWU* in Figure 8 in the Appendix.

allocation could create better opportunities. *Primal-Dual* partially mitigates this issue, reducing the variance among the curves (see the third plot in Figure 2b).

Interestingly, with *Primal-Dual w. forecast*, lower-clickability campaigns experience very low utilisation. As discussed earlier, this occurs due to forecast inconsistencies–large λ_j values for low-valued campaigns reduces their chance of being allocated. Finally, *Primal-Dual w. CPWU* further improves budget spending, resulting in a more linear spending pattern compared to *Primal-Dual*.

5.2 Real-world Data

We now evaluate our approach using a real-world dataset containing past allocations of approximately 1500 sponsored campaigns run by a major digital platform. This platform delivers a personalised feed recommending music, videos and podcasts to users, with one fixed slot reserved for sponsored content. The *Greedy* algorithm serves as the platform's production allocation policy.

Our dataset consists of one full day of logged data, representing approximately 320 million impression opportunities across 20 markets (countries). For each opportunity, the full set of candidate campaigns, along with their respective scores π_{ij} and budgets B_j, is logged. This allows us to reallocate these candidates using *Greedy*, *Primal-Dual*, *Primal-Dual w. forecast* and *Primal-Dual w. CPWU* (excluding *Random* as it is a weak baseline).

To conduct our evaluation, we split our dataset into $T = 96$ intervals of 15 minutes each, updating the pacing coefficients λ_j for each campaign j at the start of every interval. We report total allocated value, ROI and budget spent, all measured relative to *Greedy*. Our findings are in Table 1.

	Total Value	ROI	Budget Spent
Primal-Dual	+2.5%	+5.3%	-3%
Primal-Dual w. forecast	-0.4%	+6.7%	-7.9%
Primal-Dual w. CPWU	**+4.7%**	**+8.3%**	-3%

Table 1: Total value, ROI and budget spent relative to *Greedy*.

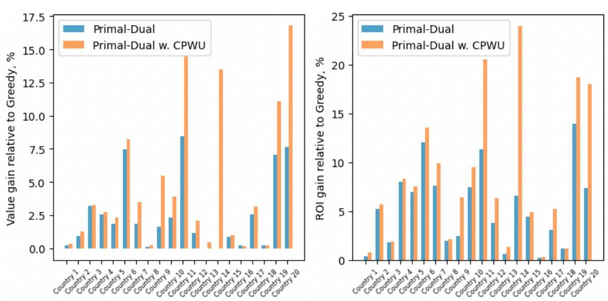


Figure 4: Total allocated value and ROI gains of Primal-Dual and Primal-Dual CPWU relative to Greedy across 20 markets.

Total Value and ROI. First, we observe that both *Primal-Dual* and *Primal-Dual w. CPWU* significantly increase total allocated value and ROI compared to *Greedy*. On average, *Primal-Dual* achieves a +2.5% improvement in total value over *Greedy* across all 20 markets, while *Primal-Dual w. CPWU* further boosts this gain to +4.7%. The ROI follows a similar trend, with *Primal-Dual* achieving +5.3% and *Primal-Dual w. CPWU* reaching +8.3%.

In contrast, *Primal-Dual w. forecast* performs poorly. As discussed in Section 5.1, its forecast may be inconsistent, causing the algorithm to become overconfident. This leads to fewer allocations (-7.9% of budget spent relative to *Greedy*, see Table 1), ultimately reducing the total allocated value (-0.4% relative to *Greedy*).

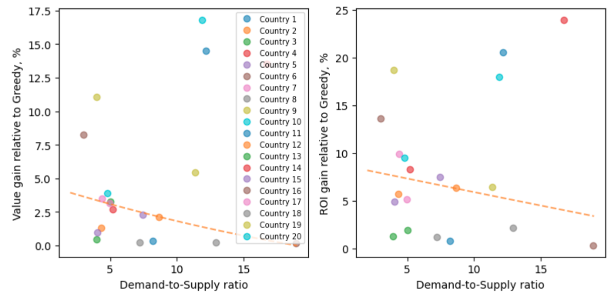


Figure 5: Total allocated value and ROI gains as a function of Demand-to-Supply ratio for 20 existing markets.

Second, we find that *Primal-Dual w. CPWU* consistently outperforms standard *Primal-Dual*, both in total value and ROI (see Table 1 and Figure 4). More importantly, these gains are nearly uniform across all 20 markets, as shown in Figure 4, where the orange bars (representing *Primal-Dual w. CPWU*) are consistently taller than the blue bars (*Primal-Dual*) in most markets.

Role of Demand/Supply Ratio. We now analyze how demand-to-supply ratio impact expected value and ROI gains. Figure 5 shows that these gains are significantly larger in markets with lower demand-to-supply ratios. This trend is intuitive: when impression opportunities outnumber the overall demand, campaigns have more flexibility to wait for better opportunities. As a result, they can allocate a larger share of high-value impressions, improving overall performance. Moreover, this finding suggests a potential revenue opportunity for the platform. In markets with higher demand-to-supply ratios, increasing prices could help optimize revenue without significantly reducing allocation efficiency.

Impact on Budget Utilisation. Finally, we observe that both *Primal-Dual* and *Primal-Dual w. CPWU* spend 3% less budget compared to *Greedy*. This is expected, as both methods allocate impressions more selectively, leading to a slower budget spend. As a result, while these approaches improve allocation efficiency, they may lead to a slightly lower total number of allocations relative to *Greedy*.

5.3 Discussion

We evaluated our approach using both synthetic and real-world data. Through synthetic experiments, we demonstrated that our algorithm significantly increases total allocated value and ROI for campaigns. These gains are further supported by notable improvements in pacing quality, ensuring that budgets are spent smoothly throughout the campaign duration. Additionally, our ablation study highlighted the importance of consistent strategy computation, leading to a +6.5% improvement over inconsistent strategies. We then validated our approach on a large-scale industrial dataset from a major digital platform. Our results show that our algorithm delivers substantial value and ROI gains in real-world settings. Notably, these gains are consistent across all 20 markets and are particularly pronounced in markets with low demand-to-supply ratios.

6 CONCLUSIONS

We proposed a novel budget management algorithm that integrates primal-dual approximation with constrained polynomial weights update (CPWU) to achieve both worst-case guarantees and improved allocative efficiency. Our approach bridges the gap between traditional worst-case optimization techniques and adaptive online learning methods by introducing a feedback-breaking mechanism, ensuring robust budget allocation in dynamic environments.

Through extensive offline evaluations, we demonstrated that our method significantly improves total allocated value and ROI compared to existing approaches. Our findings highlight that Primal-Dual with CPWU consistently outperforms standard Primal-Dual and learning-based baselines, achieving up to a +4.7% increase in allocated value and a +8.3% boost in ROI across 20 markets. Our results indicate that the gains are particularly pronounced in low demand-to-supply environments, providing insights for optimizing pricing strategies in budget-constrained systems.

Our study reaffirms the importance of balancing worst-case guarantees with learning-based adaptability in budget pacing algorithms, paving the way for more efficient and flexible budget management solutions in digital advertising and beyond. Future work could explore how to adapt the model towards incorporating additional constraints such as fairness or diversity.

7 ACKNOWLEDGEMENTS

We are thankful to Lucas Maystre, Zofia Trstanova, David Gustafsson, Tonia Danylenko, Max Lefarov, Jackie Doremus, Ben Carterette, Georges Dupret, Joel Persson, Cindy Zhang and Himan Abdollahpouri for the numerous discussions and for their suggestions that helped to significantly improve the paper.

REFERENCES

[1] N. Buchbinder, K. Jain, and J. S. Naor, "Online primal-dual algorithms for maximizing ad-auctions revenue," in *Proceedings of the 15th Annual European Conference on Algorithms*, ESA'07, (Berlin, Heidelberg), p. 253–264, Springer-Verlag, 2007.

[2] N. Buchbinder and J. S. Naor, "Online primal-dual algorithms for covering and packing," *Mathematics of Operations Research*, vol. 34, no. 2, pp. 270–286, 2009.

[3] E. Kevi and K. T. Nguyen, "Primal-dual algorithms with predictions for online bounded allocation and ad-auctions problems," in *Proceedings of The 34th International Conference on Algorithmic Learning Theory* (S. Agrawal and F. Orabona, eds.), vol. 201 of *Proceedings of Machine Learning Research*, pp. 891–908, PMLR, 20 Feb–23 Feb 2023.

[4] W. Ding, T. Qiny, X.-D. Zhang, and T.-Y. Liu, "Multi-armed bandit with budget constraint and variable costs," in *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence*, AAAI'13, p. 232–238, AAAI Press, 2013.

[5] C. Borgs, J. Chayes, N. Immorlica, K. Jain, O. Etesami, and M. Mahdian, "Dynamics of bid optimization in online advertisement auctions," in *Proceedings of the 16th International Conference on World Wide Web*, WWW '07, (New York, NY, USA), p. 531–540, Association for Computing Machinery, 2007.

[6] L. Tran-Thanh, A. C. Chapman, E. M. de Cote, A. Rogers, and N. R. Jennings, "Epsilon-first policies for budget-limited multi-armed bandits," in *AAAI Conference on Artificial Intelligence*, 2010.

[7] K. Brantley, Z. Fang, S. Dean, and T. Joachims, "Ranking with long-term constraints," in *Proceedings of the 17th ACM International Conference on Web Search and Data Mining*, WSDM '24, (New York, NY, USA), p. 47–56, Association for Computing Machinery, 2024.

[8] J. Xu, K.-c. Lee, W. Li, H. Qi, and Q. Lu, "Smart pacing for effective online ad campaign optimization," in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '15, (New York, NY, USA), p. 2217–2226, Association for Computing Machinery, 2015.

[9] R. Mehrotra, B. Carterette, Y. Li, Q. Yao, C. Gao, J. Kwok, Q. Yang, and I. Guyon, "Advances in recommender systems: From multi-stakeholder marketplaces to automated recsys," in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery Data Mining*, (New York, NY, USA), p. 3533–3534, Association for Computing Machinery, 2020.

[10] L. E. Celis, D. Straszak, and N. K. Vishnoi, "Ranking with fairness constraints," *ArXiv*, vol. abs/1704.06840, 2017.

[11] A. Singh and T. Joachims, "Fairness of exposure in rankings," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '18, (New York, NY, USA), p. 2219–2228, Association for Computing Machinery, 2018.

[12] S. Balseiro, A. Kim, M. Mahdian, and V. Mirrokni, "Budget management strategies in repeated auctions," in *Proceedings of the 26th International Conference on World Wide Web*, WWW '17, (Republic and Canton of Geneva, CHE), p. 15–23, International World Wide Web Conferences Steering Committee, 2017.

[13] S. Balseiro, Y. Deng, J. Mao, V. Mirrokni, and S. Zuo, "Robust auction design in the auto-bidding world," in *Proceedings of the 35th International Conference on Neural Information Processing Systems*, NIPS '21, (Red Hook, NY, USA), Curran Associates Inc., 2021.

[14] S. R. Balseiro, K. Bhawalkar, Z. Feng, H. Lu, V. Mirrokni, B. Sivan, and D. Wang, "A field guide for pacing budget and ROS constraints," in *Proceedings of the 41st International Conference on Machine Learning* (R. Salakhutdinov, Z. Kolter, K. Heller, A. Weller, N. Oliver, J. Scarlett, and F. Berkenkamp, eds.), vol. 235 of *Proceedings of Machine Learning Research*, pp. 2607–2638, PMLR, 21–27 Jul 2024.

[15] S. Agrawal, M. Zadimoghaddam, and V. Mirrokni, "Proportional allocation: Simple, distributed, and diverse matching with high entropy," in *Proceedings of the 35th International Conference on Machine Learning* (J. Dy and A. Krause, eds.), vol. 80 of *Proceedings of Machine Learning Research*, pp. 99–108, PMLR, 10–15 Jul 2018.

[16] F. Spaeh and A. Ene, "Online ad allocation with predictions," in *Proceedings of the 37th International Conference on Neural Information Processing Systems*, NIPS '23, (Red Hook, NY, USA), Curran Associates Inc., 2024.

[17] S. Arora, E. Hazan, and S. Kale, "The multiplicative weights update method: a meta-algorithm and applications," *Theory of Computing*, vol. 8, no. 6, pp. 121–164, 2012.

[18] N. Littlestone, "Learning quickly when irrelevant attributes abound: A new linear-threshold algorithm," in *28th Annual Symposium on Foundations of Computer Science (sfcs 1987)*, pp. 68–77, 1987.

[19] *Algorithmic Game Theory.* Cambridge University Press, 2007.

[20] A. Blum and Y. Mansour, "From external to internal regret," *Journal of Machine Learning Research*, vol. 8, no. 47, pp. 1307–1324, 2007.

[21] C. M. Bishop, *Patern Recognition and Machine Learning.* New York, NY, USA: Springer, 2007.

[22] K. E. Train, *Discrete Choice Methods with Simulation.* Cambridge University Press, 2 ed., 2009.

[23] S. A. Plotkin, D. B. Shmoys, and Éva Tardos, "Fast approximation algorithms for fractional packing and covering problems," *Mathematics of Operations Research*, vol. 20, no. 2, pp. 257–301, 1995.

A APPENDIX

A.1 Proofs

First, we write down the dual for Problem (2)-(4):

$$\min_{\lambda_j, \gamma_{ij} \geq 0} \sum_{j=1}^{M} \lambda_j B_j + \sum_{t=1}^{T} \sum_{i=1}^{N} \gamma_{it} \tag{17}$$

$$\text{s.t.,} \;\; \lambda_j + \gamma_{it} \geq \pi_{ij}^t v_{ij} \quad \forall t, i, j. \tag{18}$$

Statement 1. *Let λ_j, $j = 1, ..., M$ be the Lagrangian multipliers of Constraints (3). Then, Algorithm 1 solves Problem (2)-(4) optimally.*

Proof. First, we illustrate that if all bids for user i at time t are negative, (i.e., $b_{ij}^t < 0$ for all $j = 1, ..., M$), then user i is not allocated at time t. Indeed, if $b_{ij}^t < 0$ for all $j = 1, ..., M$, then $\pi_{ij}^t v_{ij} - \lambda_j < 0$, $\forall j = 1, ..., M$. Consequently, given that $\gamma_{it} \geq 0$ we have

$$\gamma_{it} - (\pi_{ij}^t v_{ij} - \lambda_j) = \gamma_{it} - b_{ij}^t > 0, \tag{19}$$

for all $j = 1, ..., M$. Equivalently, $\lambda_j + \gamma_{it} > \pi_{ij}^t v_{ij}$, $\forall j = 1, ..., M$. Thus, from complementary slackness of Constraint (18) it follows that $a_{ij}^t = 0$ for all $j = 1, ..., M$. This implies that the user i is not allocated at time t.

We now illustrate that as long as there exists at least one promoter with a positive bid $b_{ij}^t > 0$ for user i at time t, then the user must be allocated a promotion. Assume contrary, i.e., $a_{ij}^t = 0 \; \forall j = 1, ..., M$. Then, from complementary slackness of Constraint (4) it follows that $\gamma_{it} = 0$. This implies that $\lambda_j \geq \pi_{ij}^t v_{ij}$ for all $j = 1, ..., M$, or equivalently, $\pi_{ij}^t v_{ij} - \lambda_j = b_{ij}^t \leq 0$ for all $j = 1, ..., M$, which is a contradiction.

Next, we show that if there exists at least one promoter with a positive bid, then the user i must be allocated to the highest bidder. Let $j^* \in \arg\max_j \{b_{ij}^t\}$. We now assume contrary, i.e., that user i is allocated to promoter $j \neq j^*$. In this case, $a_{ij^*}^t = 0$, and $a_{ij}^t = 1$. Consequently, by complementary slackness we have $\gamma_{it} = \pi_{ij}^t v_{ij} - \lambda_j$. This also means that

$$\gamma_{it} \geq \pi_{ij^*}^t v_{ij^*} - \lambda_{j^*} = b_{ij^*}^t > b_{ij}^t = \pi_{ij}^t v_{ij} - \lambda_j = \gamma_{it} \tag{20}$$

which is a contradiction. Therefore, user i must be allocated to j^*.

Finally, if the budget of promoter j is exhausted, we remove the bidder from the allocation problem to satisfy Constraint (3). Q.E.D. □

Statement 2. *Algorithm 2 outputs an approximately optimal solution to Problem (2)-(4) that is $O\big(\log|1 + \frac{2}{\Pi} \max_{i,j,t}\{\pi_{ij}^t v_{ij}\}|\big)$-competitive.*

Proof. We first show that the solution produced by Algorithm 2 is always feasible. Indeed, at the first iteration all $a_{ij}^t = 0$ which is a feasible solution for Problem (2)-(4). At each iteration of Algorithm 2 (i.e., the loops in lines 2-3) the budget constraint (3) is always satisfied. Also, Algorithm 2 only allocates at most a single campaign j^* to the user (line 11), which implies feasibility of Constraint (4).

We now illustrate that at each iteration the gap between the primal solution and the dual solution is bounded. Indeed, as we update the dual variables a_{ij}^t (line 6 of Algorithm 2), the respective change of the primal objective (17) is:

$$\frac{\partial \mathcal{P}}{\partial a_{ij}^t} = B_j \frac{\partial \lambda_j}{\partial a_{ij}^t} + \frac{\partial \gamma_{it}}{\partial a_{ij}^t}. \tag{21}$$

We let

$$\lambda_j = \frac{\Pi}{2} \exp\left\{\frac{A}{B_j} \sum_{t=1}^{T} \sum_{i=1}^{N} a_{ij}^t\right\} - \frac{\Pi}{2}, \tag{22}$$

$$\gamma_{it} = \frac{\Pi}{2} \exp\left\{A \sum_{j=1}^{M} a_{ij}^t\right\} - \frac{\Pi}{2}. \tag{23}$$

Consequently,

$$\frac{\partial \lambda_j}{\partial a_{ij}^t} = \frac{A}{B_j} \frac{\Pi}{2} \exp\left\{\frac{A}{B_j} \sum_{t=1}^{T} \sum_{i=1}^{N} a_{ij}^t\right\}, \tag{24}$$

$$\frac{\partial \gamma_{it}}{\partial a_{ij}^t} = A \frac{\Pi}{2} \exp\left\{A \sum_{j=1}^{M} a_{ij}^t\right\}. \tag{25}$$

Therefore, we can re-write:

$$\frac{\partial \lambda_j}{\partial a_{ij}^t} = \frac{A}{B_j}\Bigg[\underbrace{\frac{\Pi}{2} \exp\left\{\frac{A}{B_j} \sum_{t=1}^{T} \sum_{i=1}^{N} a_{ij}^t\right\} - \frac{\Pi}{2}}_{\lambda_j \;\; \text{(see Equation (22))}} + \frac{\Pi}{2}\Bigg] = \frac{A}{B_j}\Big[\lambda_j + \frac{\Pi}{2}\Big]. \tag{26}$$

Similarly, we obtain

$$\frac{\partial \gamma_{it}}{\partial a_{ij}^t} = A\Big[\gamma_{it} + \frac{\Pi}{2}\Big]. \tag{27}$$

Now, we can re-write Equation (21) as follows:

$$\frac{\partial \mathcal{P}}{\partial a_{ij}^t} = A\Big[\lambda_j + \gamma_{it} + \Pi\Big] \leq A\Big[\pi_{ij}^t v_{ij} + \Pi\Big] \leq 2A\pi_{ij}^t v_{ij} = 2A \frac{\partial \mathcal{D}}{\partial a_{ij}^t}. \tag{28}$$

Here, $\lambda_j + \gamma_{it}$ gradually increases until the respective constraint (18) becomes binding. Thus, $\lambda_j + \gamma_{it} \leq \pi_{ij}^t v_{ij}$ in lines 6-9 of Algorithm 2. Consequently, the gap between the primal and dual solutions is always bounded by $2A$.

Finally, let us derive the constant A. Assume that the budget constraint (3) is binding for some j, i.e., $\sum_{t=1}^{T} \sum_{i=1}^{N} a_{ij}^t = B_j$. In this case, from Equation (22) we obtain:

$$\lambda_j = \frac{\Pi}{2} \exp\{A\} - \frac{\Pi}{2}. \tag{29}$$

From Constraint (18) we know that

$$\lambda_j \geq \pi_{ij}^t v_{ij} - \gamma_{it} \geq \pi_{ij}^t v_{ij} \quad \forall t, i. \tag{30}$$

Consequently, λ_j must be at least $\max_{t,i}\{\pi_{ij}^t v_{ij}\}$. Therefore, from Equation (29) we obtain:

$$A \geq \log(1 + \frac{2}{\Pi} \max_{i,j,t}\{\pi_{ij}^t v_{ij}\}) \tag{31}$$

Thus, we let $A = \log\big[1 + \frac{2}{\Pi} \max_{i,j,t}\{\pi_{ij}^t v_{ij}\}\big]$. Q.E.D. □

Lemma 1. *Function $f(q_1, ..., q_M, \lambda_1, ..., \lambda_M) = (f_1, ..., f_M)$ is Lipschitz-continuous.*

Proof. Consider first a single component of the function $\sigma_j(.)$. We first show that all partial derivatives of $\sigma_j(.)$ are bounded.

First, let us re-write:

$$\sigma_j(s_1, ..., s_M) = \min\left\{1, \frac{N}{B_j}\frac{\exp\left\{q_j - \frac{\Pi}{2}\exp\{As_j\} + \frac{\Pi}{2}\right\}}{\sum_{k=1}^{M}\exp\left\{q_k - \frac{\Pi}{2}\exp\{As_k\} + \frac{\Pi}{2}\right\}}\right\}. \tag{32}$$

Now, observe that if $\sigma_j = 1$, then $\frac{\partial \sigma_j}{\partial s_j} = 0$, and therefore, is bounded. If $\sigma_j < 1$, then we have:

$$\frac{\partial \sigma_j}{\partial s_j} = \frac{N}{B_j}\frac{\exp\left\{q_j - \frac{\Pi}{2}\exp\{As_j\} + \frac{\Pi}{2}\right\}(-\frac{\Pi}{2}\exp\{As_j\}A)\sum_{k\neq j}\exp\left\{q_k - \frac{\Pi}{2}\exp\{As_k\} + \frac{\Pi}{2}\right\}}{\left(\sum_{k=1}^{M}\exp\left\{g_k(\theta_k|W_k) - w_k\cdot\frac{\Pi}{2}\exp\{As_k\} + w_k\frac{\Pi}{2}\right\}\right)^2} = \tag{33}$$

$$\frac{N}{B_j}(-w_j\frac{\Pi}{2}\exp\{As_j\}A)\cdot\sigma_j(1-\sigma_j). \tag{34}$$

If follows then that

$$\frac{N}{B_j}(-w_j\frac{\Pi}{2}\exp\{A\}A)\cdot\sigma_j(1-\sigma_j) \leq \frac{\partial \sigma_j}{\partial s_j} \leq \frac{N}{B_j}(-w_j\frac{\Pi}{2}A)\cdot\frac{1}{4}, \tag{35}$$

which implies that the derivative is bounded. Similarly, we can show that $\frac{\partial \sigma_j}{\partial s_\ell}$ is bounded for any $\ell \neq j$:

TBD. Q.E.D. □

Statement 3. *For any $\Pi \geq 0.064$ and any compatible anticipated budget utilisation $\tilde{g}_j^T$, $j = 1, ..., M$ Algorithm 3 computes an approximately optimal solution to Problem (2)-(4) that is at least as good as the one produced by Algorithm 2.*

Proof. Observe that the only difference between Algorithm 2 and Algorithm 3 is in the update step in lines 10-11. We now show that the change in the primal objective during this update step is still bounded by the respective change in the dual objective:

$$\frac{\partial \mathcal{P}}{\partial a_{ij}^t} = \frac{\tilde{\lambda}_j}{\tilde{g}_j} + \pi_{ij}^t v_{ij} - \tilde{\lambda}_j = \tilde{\lambda}_j\left(\frac{1}{\tilde{g}_j} - 1\right) + \pi_{ij}^t v_{ij} = \tag{36}$$

$$\left(\frac{1}{\tilde{g}_j} - 1\right)\left[\frac{\Pi}{2}\exp\{A\tilde{g}_j\} - \frac{\Pi}{2}\right] + \pi_{ij}^t v_{ij} \tag{37}$$

From the Taylor expansion we know:

$$\exp\{A\tilde{g}_j\} \leq 1 + A\exp\{A\tilde{g}_j\}\tilde{g}_j. \tag{38}$$

Consequently,

$$\frac{\partial \mathcal{P}}{\partial a_{ij}^t} \leq \left(\frac{1}{\tilde{g}_j} - 1\right)\frac{\Pi}{2}A\exp\{A\tilde{g}_j\}\tilde{g}_j + \pi_{ij}^t v_{ij} \leq \tag{39}$$

$$\pi_{ij}^t v_{ij}\left[1 + \frac{1-\tilde{g}_j}{2}A\exp\{A\tilde{g}_j\}\right]. \tag{40}$$

Observe, that the term in the parentheses is maximized when $\tilde{g}_j = \frac{A-1}{A}$. Therefore,

$$\frac{\partial \mathcal{P}}{\partial a_{ij}^t} \leq \pi_{ij}^t v_{ij}\left[1 + \frac{\exp\{A-1\}}{2}\right]. \tag{41}$$

We now illustrate that for all $A \leq 3.48$ the term in the parentheses is not larger than $2A$. This would imply that our statement holds for any $\Pi \geq 2/\exp\{3.48 - 1\} \geq 0.064$. Indeed,

$$1 + \frac{\exp\{A-1\}}{2} \leq 2A \tag{42}$$

implies

$$\exp\left\{A - 1\right\} \leq 4A - 2. \tag{43}$$

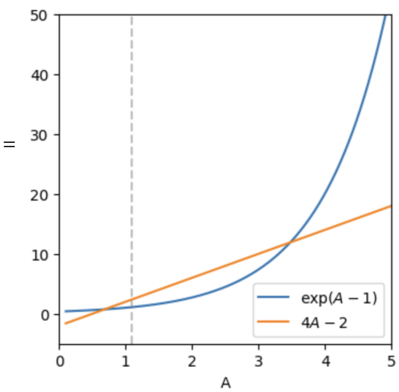


Figure 6: Numerical solution of Inequality (43).

Observe that the minimal value that A can take is $A_{min} = \log(1 + \frac{2}{1}) = \log(3)$ (see the dashed grey line in Figure 6). From Figure 6 we can also see that for all $A \leq 3.48$ Inequality (43) holds. □

Statement 4. *If there exists an optimal solution to Problem (12) that is compatible with Algorithm 3, then there also exists an optimal solution p^t that is consistent.*

Proof. Let p^t be the optimal solution to Problem (12) that is compatible with Algorithm 3. We will now assume contrary, i.e., that p^t is not consistent. W.l.o.g., we let t be the earliest time step when $p^t \neq q^t$, and we let $p_j^t > q_j^t$ for some campaign j.

We let the anticipated budget utilisation at time t be

$$\tilde{g}_j^t(p_j^t) = \min\left\{1, \frac{Q_j^{t-1} + p_j^t N}{B_j}\right\}, \tag{44}$$

and we let the actual budget utilisation be

$$g_j^t = \min\left\{1, \frac{Q_j^{t-1} + q_j^t N}{B_j}\right\}. \tag{45}$$

From monotonicity of Equations (44) and (45) we see that $p_j^t > q_j^t$ implies $\tilde{g}_j^t(p_j^t) \geq g_j^t$. Observe, that if $\tilde{g}_j^t > g_j^t$, then p_j^t cannot be compatible (see Definition 2) which is a contradiction. If instead, $\tilde{g}_j^t = g_j^t$, then it must hold that $\tilde{g}_j^t = g_j^t = 1$. In this case, one could gradually decrease p_j^t until $p_j^t = q_j^t$ without affecting the anticipated budget utilisation $\tilde{g}_j^t(p_j^t)$, and consequently, without affecting the allocations produced by Algorithm 3. This would result in a new optimal solution to Problem (12) that is compatible and also consistent. Q.E.D. □

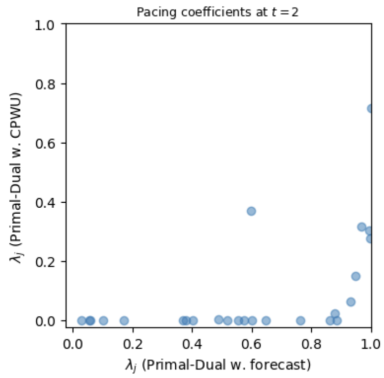


Figure 8: Pacing coefficients λ_j of the same $M = 25$ simulated campaigns at time $t = 2$. The pacing coefficients used by *Primal-Dual w. forecast* are significantly larger than the ones computed by *Primal-Dual w. CPWU*. This may lead to significantly fewer allocations achieved by *Primal-Dual w. forecast* compared to *Primal-Dual w. CPWU*.

A.1.1 Derivation of the Lower Bound.

$$\max_{p^t \in \Delta(M)} \log\Big(\sum_{t=1}^{T} \frac{1}{T} \sum_{j=1}^{M} q_j^t(p^t) r_j^t\big(p^t, q^t(p^t)\big)\Big) \geq \tag{46}$$

$$\max_{p^t \in \Delta(M)} \sum_{t=1}^{T} \frac{1}{T} \log\Big(\sum_{j=1}^{M} q_j^t(p^t) r_j^t\big(p^t, q^t(p^t)\big)\Big) \geq \tag{47}$$

$$\max_{p^t \in \Delta(M)} \frac{1}{T} \sum_{t=1}^{T} \log\Big(\sum_{j=1}^{M} p_j^t \frac{q_j^t(p^t)}{p_j^t} r_j^t\big(p^t, q^t(p^t)\big)\Big) \geq \tag{48}$$

$$\max_{p^t \in \Delta(M)} \frac{1}{T} \sum_{t=1}^{T} \sum_{j=1}^{M} p_j^t \log\Big(\frac{q_j^t(p^t)}{p_j^t} r_j^t\big(p^t, q^t(p^t)\big)\Big) \geq \tag{49}$$

$$\max_{p^t \in \Delta(M)} \frac{1}{T} \sum_{t=1}^{T} \sum_{j=1}^{M} \Big(p_j^t \log \frac{q_j^t(p^t)}{p_j^t} + p_j^t \log r_j^t\big(p^t, q^t(p^t)\big)\Big) = \tag{50}$$

$$\max_{p^t \in \Delta(M)} \frac{1}{T} \sum_{t=1}^{T} \left(\sum_{j=1}^{M} p_j^t \Big[\log r_j^t\big(p^t, q^t(p^t)\big) + \log q_j^t(p^t)\Big] + H(p^t)\right). \tag{51}$$

A.2 Simulator

Figure 7 illustrates the distributions of scores π_{ij} of three campaigns $j = 0, 12, 24$ used in our offline experiments in Section 5.1. We can see that the distributions gradually shift towards the lower levels of click probability, s.t, the campaigns with larger indices have less clickable content.

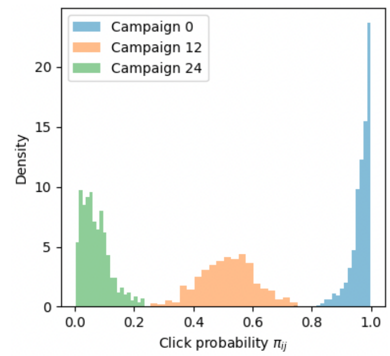


Figure 7: Simulations setup (synthetic data).

Figure 8 illustrates the pacing coefficients of all $M = 25$ simulated campaigns computed with *Primal-Dual w. CPWU* (y-axis) vs. *Primal-Dual w. forecast* (x-axis) at time $t = 2$. We can see that the pacing coefficients computed by *Primal-Dual w. forecast* are significantly larger than their counterparts computed by *Primal-Dual w. CPWU* at the same time step. This implies that the campaigns are paced much more aggressively by *Primal-Dual w. forecast* already at the very beginning of their duration. This results in a significantly fewer allocations relative to *Primal-Dual w. CPWU*, decreasing the total allocated value.