

# Adaptive Spend Control for Programmatic Advertising via Input Reparameterization

Niklas Karlsson, Tina Fan, Dan Kirlin, and Abhirup Mukherjee

**Abstract**—Budget delivery in programmatic advertising requires controlling spend across advertising campaigns whose plant gains vary significantly across campaigns and evolve over time, rendering fixed-gain controllers inadequate. This paper proposes an adaptive spend control architecture comprising four components: an input reparameterization that transforms the nonlinear spend-to-control relationship into a near-linear one, an online plant gain estimator employing convex warm-start initialization and exponential smoothing, an adaptive integral controller that maintains consistent loop gain by adjusting inversely to the estimated plant gain, and a seasonality compensator that decouples the feedback loop from predictable periodic disturbances. Employing Lyapunov theory, we establish exponential convergence guarantees for the control loop acting on an idealized deterministic plant. Closed-loop simulations demonstrate that robust performance is maintained under realistic conditions including stochastic noise, nonlinear response curves, imperfect seasonality models, and mid-campaign plant dynamics changes, achieving full budget delivery across campaigns with 5x to 6x plant gain variation.

**Index Terms** — Adaptive control, system identification, programmatic advertising, budget pacing, Lyapunov theory

## I. INTRODUCTION

Programmatic advertising involves the use of automated software and data-driven algorithms to buy and sell digital advertisements in real-time auctions to serve an *impression*, which is an ad creative shown to an internet user. The global market for programmatic advertising was valued at around \$650 billion in 2025 [1], underscoring the scale and economic significance of the underlying optimization infrastructure.

*Demand-Side Platforms* (DSPs) are tools advertisers use to manage and purchase ad space. Advertiser’s intention in using a DSP is to spend an ad campaign budget optimally in some sense — for example to maximize the total number of clicks or sales. From a DSP’s perspective, this requires solving a complex, dynamic constrained optimization problem in which it must estimate the value of each impression, predict the competition, and determine an optimal bid — all under strict latency constraints and in the face of incomplete information about user behavior and market dynamics. The central challenge of programmatic advertising lies in optimizing bidding decisions across an enormous volume of auctions — often billions per day — each of which must be evaluated and acted upon in a matter of milliseconds.

The extremely high-dimensional optimization problem can be made tractable by modularizing the problem into three sub-problems, one of which is a control problem. The core

control problem involves dynamically adjusting bidding aggressiveness to regulate spend rate. However, this seemingly straightforward task is complicated by several factors: (1) the relationship between the control signal and the resulting spend is nonlinear, (2) campaign-specific plant gains vary by factors of 5x or more across different campaigns and evolve dynamically within campaigns over time, (3) external disturbances such as seasonal traffic patterns and competitive dynamics continuously perturb the system, and (4) the system operates under hard constraints with limited observability of underlying dynamics.

Feedback control in programmatic advertising is critically important, and an active research area. Early work includes [2] and [3], where the authors present a high level introduction to the control problem in online advertising and outline several critical challenges. Meanwhile [4] proposes a solution to address specific problems, but does not sufficiently consider both the dynamic and periodic nature of the plant. [5] presents a statistical inference of the plant gain based on an uncertain discontinuous plant response curve combined with bid randomization while [6] proposes an adaptive estimation and control framework that leverages *Recursive Least Squares* (RLS) estimation alongside persistent excitation to learn the true plant of the system in real-time [7], [8]. The biggest challenge in the latter’s approach is the two parameter plant model which requires persistent excitation.

The contribution of this paper sidesteps the complexities highlighted above by re-parameterizing the control signal [9], yielding a plant with advantageous properties that enable a simpler one-parameter plant model and eliminate the need for persistent excitation [7], [8]. This re-parameterization permits a streamlined online plant gain estimator that recursively identifies the plant gain, coupled with an adaptive feedback controller that adjusts its gain based on the plant gain estimate, ensuring appropriate responsiveness across diverse campaigns. We establish via Lyapunov theory exponential convergence guarantees for the feedback control loop acting on an idealized deterministic plant and we demonstrate through closed-loop simulations that the adaptive controller maintains robust performance when the ideal conditions are relaxed, including stochastic noise, nonlinear dynamics, and imperfect seasonality models.

## II. PROBLEM FORMULATION

### A. Optimization Problem

Let  $\Omega$  denote the set of all available impression opportunities sold individually and sequentially to the highest bidder in an open impression exchange. Each impression  $i \in \Omega$  is

N. Karlsson, T. Fan, D. Kirlin, and A. Mukherjee are with DemandTech at Amazon Ads {nikkarl, fmingyan, dankirl, abhiru} at amazon.com). This work was supported by Amazon Inc.

sold based on a first or second price cost model [10]. For a first price impression ( $i \in \Omega_1$ ) the winner pays an amount equal to its own bid, whereas for a second price impression ( $i \in \Omega_2$ ) the winner pays an amount equal to the second highest bid. Let  $\Omega_1 \cap \Omega_2 = \emptyset$  and  $\Omega_1 \cup \Omega_2 = \Omega$ .

The objective is to maximize the cost-discounted profit  $\mathcal{J} := \mathbb{E}V - \alpha \mathbb{E}C$ , where  $\mathbb{E}V$  is the total expected daily advertiser value,  $\mathbb{E}C$  is the total expected daily cost of awarded impressions, and  $\alpha \in (0, 1]$  is a configured cost-discount parameter. The profit maximization is subject to a spend constraint  $\mathbb{E}C \leq \xi$ , where  $\xi \in [0, \infty)$  is an advertiser specified daily budget.

The decision variables of the optimization problem are given by bid prices  $b_i \in [0, \infty)$ , for all  $i \in \Omega$ . The expected value of impression  $i$  (if awarded) is denoted  $v_i \in [0, \infty)$ . It represents the expected performance and/or branding value, and may encode the probability of a click or conversion (a sale), and the branding value of reaching prospective future customers with the advertisement message.

Given a win rate function  $F_i(b)$ , which is the cumulative density of the highest competing bid price for impression  $i$ , it is well-known (see e.g. [11], [12]) that the optimal bidding strategy under mild assumptions satisfies

$$b_i^{opt} = \begin{cases} \operatorname{argmax}_{b \geq 0} (b_i^u - b)F_i(b), & i \in \Omega_1, \\ b_i^u, & i \in \Omega_2, \end{cases}$$

where the *private value*,  $b_i^u$  is defined by

$$b_i^u = \frac{v_i}{\alpha + \lambda}; \quad (1)$$

and  $\mathbb{E}C \leq \xi$ ,  $\lambda(\mathbb{E}C - \xi) = 0$ , and  $\lambda \geq 0$ . Optimal bidding requires knowledge of  $v_i$ ,  $F_i(b)$ , and  $\lambda$ . Typically,  $v_i$  is computed in real time based on an offline-trained performance prediction model as a function of user/contextual information. Similarly,  $F_i(b)$  is normally trained offline based on historical clearing price data, whereas  $(b_i^u - b)F_i(b)$  is maximized online using non-convex optimization. There exists a rich literature on the modeling of  $v_i$  and  $F_i(b)$  [13], [14], and we shall assume they are both known (produced by separate systems outside the scope of this paper). It remains to find the value of  $\lambda$ , which is a unitless shadow price or Lagrange multiplier. A scalable and popular approach towards identifying  $\lambda$  is to use feedback control, which is the topic of this paper.

### B. Control Problem

It is easily shown that  $\mathbb{E}V$  and  $\mathbb{E}C$  are non-negative monotonic decreasing functions of  $\lambda$ , and that the optimal  $\lambda$  equals the smallest non-negative value, for which  $\mathbb{E}C \leq \xi$  [6]. The goal is to devise a feedback control algorithm that computes a Lagrange multiplier  $\lambda(t)$  converging towards a constant at which  $\lambda(\mathbb{E}C - \xi) = 0$  and  $\mathbb{E}C \leq \xi$ . The feedback algorithm must ensure closed loop stability and sufficiently fast convergence towards the optimal value.

Consider a time-sampled implementation of the control system with equidistant time points indexed  $t = 0, 1, \dots$ , where the sampling time  $\Delta$  is, for example, 1/30 hours. The number of sampling intervals per day equals  $T_{24} = 24/\Delta$ .

Distribute the daily budget uniformly throughout the day as  $\bar{u}_c(t) = \xi/T_{24}$ , and let this represent the *command signal*. Furthermore, let the observed spend in time interval  $t$  be denoted  $y(t)$ . It follows that  $\xi = \sum_{t=0}^{T_{24}-1} \bar{u}_c(t)$  and  $c = \sum_{t=0}^{T_{24}-1} y(t)$ , where  $c$  is an observation of the daily spend  $C$  (note,  $c$  is a realization of  $C$ , which is random).

A challenge is that the relationship between  $\lambda$  and the incremental spend contributing to  $\mathbb{E}C$  is a priori unknown, nonlinear, time-varying, and stochastic. The relationship moreover differs a lot from one campaigns to the next. A fixed-gain controller has to be designed conservatively to ensure stability across campaigns with varying plant gains and will exhibit substantially slower convergence for campaigns with lower plant gains.

### III. PLANT MODEL

The environment is modeled as a nonlinear, time-varying, stochastic plant with a one-step delay. Specifically, the control signal  $\lambda(t)$  applied at time  $t$  exerts its full effect on the output  $y(t+1)$  at the subsequent time step. The one-step delay assumption is consistent with the operational latency inherent in programmatic advertising auctions, where bid price adjustments propagate through the auction clearing mechanism before influencing realized spend.

The nominal plant model is given by

$$y(t+1) = (1 + h(t))g(\lambda(t))(1 + w(t)), \quad (2)$$

where  $h(t)$  is a  $T_{24}$ -periodic seasonality function satisfying  $h(t + T_{24}) = h(t)$ ,  $\sum_{\tau=0}^{T_{24}-1} h(\tau) = 0$ , and  $h(t) > -1$  for all  $t$ , and  $w(t)$  is zero-mean white noise representing stochastic fluctuations in Internet traffic and the competitive landscape. Steady-state spend-to-control relationship  $g(\cdot)$  is non-negative and monotonically decreasing, and continuously differentiable; but is otherwise unknown, motivating the online estimation approach developed in Section IV. The seasonality component  $h(t)$  may be estimated offline from historical traffic data and refined online via system identification techniques [15].

This model captures the essential dynamics relevant to budget pacing control while remaining tractable for analysis.

### IV. DESIGN

We propose a controller architecture that comprises four components: a seasonality compensator, a feedback linearizer, an online plant gain estimator, and an adaptive feedback controller. Each component addresses distinct challenges in achieving robust budget delivery across heterogeneous campaign dynamics. Figure 1 illustrates the proposed control system architecture as a block diagram.

The seasonality compensator is  $T_{24}$ -periodic and computes an adjusted command signal  $u_c(t)$  based on  $\bar{u}_c(t)$  and  $h(t)$ . The daily budget  $\xi$  is constant or changes only rarely. As a result, the command signal  $\bar{u}_c(t)$  is a mostly constant setpoint. Furthermore, the seasonality profile  $h(t)$  can be estimated at high accuracy offline and is here assumed known [15]. The feedback linearizer is a static nonlinear map from a new control signal  $u$  to the bid modifier  $\lambda$  designed to achieve a linear-like relationship between  $u$  and

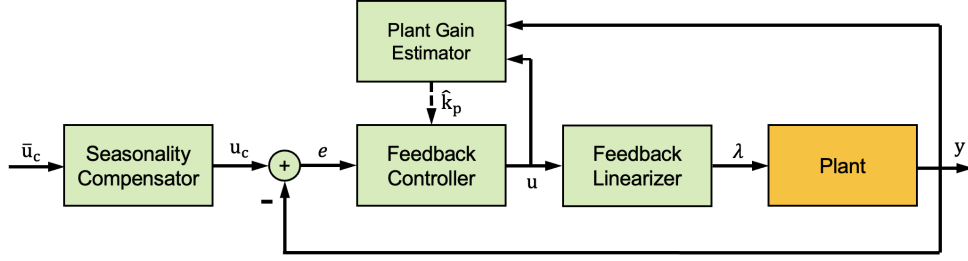


Fig. 1. Block diagram of the proposed closed-loop adaptive control system.

$y$ . The plant gain estimator computes a plant gain estimate  $\hat{k}_p(t)$  recursively based on  $u$  and  $y$ . This estimate is used to adjust the aggressiveness of the feedback controller that dynamically updates a control signal  $u(t)$  based on an error signal  $e(t) = u_c(t) - y(t)$ .

Finally, the private value  $b_i^u$  and bid  $b_i$  for each individual impression opportunity is computed based on  $u$  (via  $\lambda$ ) and used for bidding. The interconnected control system is designed so that  $y(t)$  converges to the  $T_{24}$ -periodic adjusted command signal  $u_c(t)$  and  $u(t)$  converges to a constant, which are necessary conditions for optimality [11], [12].

#### A. Seasonality Compensation

Periodic variations in Internet traffic, competitive landscape, and impression quality introduce predictable disturbances that, if unaddressed, degrade delivery accuracy and performance. We employ a feedforward compensator to cancel this effect by scaling the reference signal according to the estimated seasonality:

$$u_c(t) = (1 + h(t-1))\bar{u}_c(t), \quad (3)$$

where  $\bar{u}_c(t)$  denotes the nominal setpoint reference and the index shift by  $-1$  accounts for the plant delay. This feedforward action decouples the feedback loop from periodic disturbances, enabling convergence to the delivery target. Previous work [6] demonstrated that this seasonality compensation enables closed-loop stability.

#### B. Input Reparameterization (“Feedback Linearizer”)

The nonlinear relationship between Lagrange multiplier  $\lambda$  and realized spend  $y$  presents fundamental challenges for linear control design. To address this, we introduce a change of variables via the transformation

$$u = (\alpha + \lambda)^{-1}, \quad (4)$$

which maps  $\lambda \in [0, \infty)$  to  $u \in (0, 1/\alpha]$ , with inverse  $\lambda = u^{-1} - \alpha$ ,  $u > 0$ . Under this parameterization, the private bid value for impression  $i$  becomes  $b_i^u = uv_i$ .

**Definition 1 (Plant Gain).** Under the parameterization (4), the steady-state plant gain at  $u = u^*$  is defined as  $k_p(u^*) := \partial y / \partial u|_{u=u^*}$  representing the local sensitivity of spend rate to control signal changes.

The transformation yields a spend-to-control relationship that can be locally approximated by a linear (proportional) function.

#### C. Plant Gain Estimator

Campaign-specific plant gains  $k_p(u^*)$  vary by orders of magnitude across and within campaigns. Consequently, a fixed-gain controller cannot simultaneously achieve adequate responsiveness for low plant gain scenarios and stability for high gain scenarios. We therefore employ an online recursive estimator to identify  $k_p$  in real time.

**Initialization.** At launch, the plant gain estimate is set as

$$\hat{k}_p(0) = \frac{\xi}{T_{24}u_0}, \quad (5)$$

where  $\xi$  is the daily budget and  $u_0$  is the initial control signal.

**First 24 hours.** During the initial operating period, the estimate is formed as a convex combination of the prior and a cumulative ratio estimator:

$$\hat{k}_p(t) = w_0(t)\hat{k}_p(0) + (1 - w_0(t)) \frac{\sum_{\tau=0}^t y(\tau)}{\beta + \sum_{\tau=0}^t u(\tau)}, \quad (6)$$

where  $w_0(t) = (24 - t\Delta)/24$  is a linearly decaying weight that transitions reliance from the prior to observed data as evidence accumulates, and where  $\beta > 0$  is a regularization parameter to ensure numerical stability when control action is small (i.e.,  $\sum_{\tau=0}^t u(\tau) \approx 0$ ).

**Day 2 onwards.** Following the warm-up period, the estimate is updated daily using exponential smoothing over the preceding day’s observations:

$$\hat{k}_p(\text{day}) = w_1\hat{k}_p(\text{day} - 1) + (1 - w_1) \frac{\sum_{\text{day}} y(\tau)}{\beta + \sum_{\text{day}} u(\tau)}, \quad (7)$$

where  $w_1 \in [0, 1)$  is a fixed plant gain memory coefficient parameter. The summation is computed over the full day of active hours.

The plant gain estimate is updated daily to reduce the impact on the closed-loop system from an imperfect seasonality model, and to avoid significant fluctuations in the integral gain computed by the feedback controller (Section IV-D), which may cause biased delivery of the campaign budget.

Note, the proposed estimate  $\hat{k}_p$  is biased since it assumes  $y = k_p u$  whereas the true relationship satisfies  $y = m + k_p u$ . However, producing an unbiased estimate of  $k_p := \partial y / \partial u$  requires persistent excitation and is an ill-conditioned problem for small variations of the control signal.

Computing a plant gain estimate using the linear approximation with a properly chosen gain margin dramatically enhances the robustness of the interconnected adaptive control system. Figure 2 illustrates this comparison.

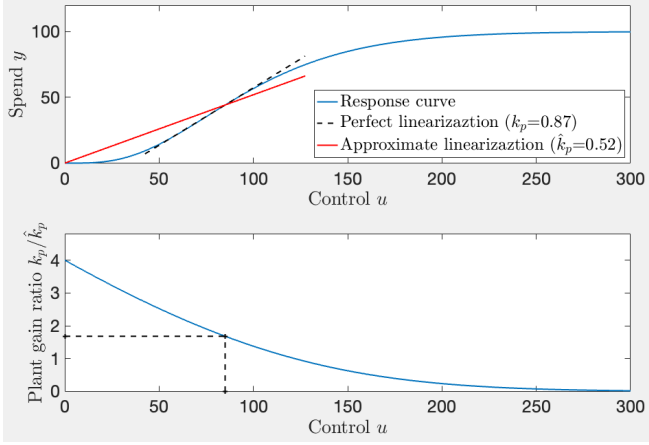


Fig. 2. Top: Illustration of a plausible spend-to-control relationship  $y = y(u)$  with perfect and approximate linearization around  $u = 85$  super-imposed. Bottom: The ratio of perfect vs approximate plant gain  $k_p/\hat{k}_p$ .

The top panel shows the spend-to-control relationship  $y = y(u)$  for a plausible campaign with the perfect and approximate linearized plant at an operating point  $u = 85$  super-imposed. At this operating point, the simplified linear model yields a plant gain of  $\hat{k}_p = 0.52$ , while the true plant gain  $k_p = 0.87$ , which is 67% larger. The bottom panel depicts the ratio  $k_p/\hat{k}_p$  at different operating points of the control signal.

#### D. Feedback Controller

The feedback controller implements pure integral error feedback updates of  $u$  with a gain inversely proportional to the estimated plant gain:

$$u(t) = \text{SAT}_{[0,1/\alpha]} [u(t-1) + k_c(t)e(t)], \quad (8)$$

where  $e(t) = u_c(t) - y(t)$  is the tracking error and  $\text{SAT}_{[a,b]}(x)$  is the projection of  $x$  onto the interval  $[a, b]$ . The adaptive gain is defined as  $k_c(t) = k_{lg}/\hat{k}_p(t)$ , where the loop gain  $k_{lg}$  is selected based on the desired closed-loop time constant  $T_{cl}$ , which is a design parameter:

$$k_{lg} = 1 - \exp(-\Delta/T_{cl}). \quad (9)$$

This adaptive structure ensures that the effective loop gain  $k_p k_c(t)$  remains approximately constant regardless of campaign-specific plant dynamics, achieving consistent closed-loop bandwidth across the advertiser population. Configuring the controller in terms of a time constant rather than the loop gain directly ensures that the control design provides sampling-time invariance.

#### V. STABILITY ANALYSIS

In this stability analysis of the feedback controller we consider an idealized plant with successful (biased) plant gain estimation subject to the following assumptions. A stability analysis of the interconnected adaptive system under stochastic noise is left as future work.

##### A. Assumptions

**Assumption 1** (Noise-Free Operation). *Deterministic system: no measurement noise or stochastic disturbances except known seasonality  $h(t)$ .*

**Assumption 2** (Known Seasonality). *The seasonality function satisfies  $|h(t)| \leq h_{\max} < 1$ , and the reference signal  $u_c(t) = (1 + h(t-1))\bar{u}_c$ , where  $\bar{u}_c$  is constant.*

**Assumption 3** (Bounded Relative Plant Gain Estimation Error).  *$\left| \frac{\hat{k}_p(t)}{k_p(u(t))} - 1 \right| \leq \epsilon < 1$  where  $\epsilon$  is constant.*

**Assumption 4** (Feedback Linearization).  *$y(t) \approx k_p(u(t-1))u(t-1)$  where plant gain  $k_p > 0$ .*

**Assumption 5** (No Saturation). *Control  $u(t) \in (0, 1/\alpha)$ .*

##### B. Error Dynamics

The plant and controller dynamics obey

$$\begin{aligned} y(t) &= [1 + h(t-1)]k_p(u(t-1))u(t-1) \\ u(t) &= u(t-1) + k_c(t)e(t). \end{aligned}$$

where the controller gain  $k_c(t) = k_{lg}/\hat{k}_p(t)$  and the error signal  $e(t) = u_c(t) - y(t)$ . Define  $\bar{e}(t) := e(t)/[1 + h(t-1)]$ . It follows that  $\bar{e}(t) = \bar{u}_c - k_p(u(t-1))u(t-1)$ .

By incorporating the plant and controller dynamics in the error signal we obtain

$$\begin{aligned} \bar{e}(t+1) &= \frac{[1 + h(t)][\bar{u}_c - k_p(u(t))u(t)]}{1 + h(t)} \\ &= \bar{e}(t)[1 - k_p(u(t))k_c(t)[1 + h(t-1)]] \end{aligned} \quad (10)$$

It follows that the error dynamics satisfy  $\bar{e}(t+1) = \varphi(t)\bar{e}(t)$ , where  $\varphi(t)$  is the characteristic function defined by

$$\varphi(t) = 1 - [1 + h(t-1)]k_p(u(t))k_c(t) \quad (11)$$

##### C. Main Convergence Theorem

**Theorem 6** (Exponential Convergence). *Under Assumptions 1–5, if  $k_{lg} < \frac{2(1-\epsilon)}{1+h_{\max}}$ , then  $|\bar{e}(t)| \leq \lambda^t |\bar{e}(0)|$ , where  $\lambda = \max\{|\varphi_{\min}|, |\varphi_{\max}|\}$  with*

$$\varphi_{\min} = 1 - \frac{(1 + h_{\max})k_{lg}}{1 - \epsilon} \quad (12)$$

$$\varphi_{\max} = 1 - \frac{(1 - h_{\max})k_{lg}}{1 + \epsilon} \quad (13)$$

*Proof:* Consider the Lyapunov function  $V(\bar{e}) = \bar{e}^2 > 0$  for  $\bar{e} \neq 0$ . This function satisfies  $V(0) = 0$  and  $V(e) \rightarrow \infty$  as  $|e| \rightarrow \infty$ . The first difference is  $\Delta V = \bar{e}(t+1)^2 - \bar{e}(t)^2 = [\varphi(t)^2 - 1]\bar{e}(t)^2$ . For  $\Delta V < 0$ , require  $|\varphi(t)| < 1$  for all  $t$ .

Define gain ratio  $\rho(t) = k_p(u(t))/\hat{k}_p(t)$ . From Assumption 3, it follows that  $(1 + \epsilon)^{-1} \leq \rho(t) \leq (1 - \epsilon)^{-1}$ . Then  $k_p(u(t))k_c(t) = k_{lg}\rho(t)$ , which implies  $\varphi(t) = 1 - [1 + h(t-1)]k_{lg}\rho(t)$ . The extreme values of  $h(t-1)$  and  $\rho(t)$  yield  $\varphi_{\min} = 1 - (1 + h_{\max})k_{lg}(1 - \epsilon)^{-1}$  and  $\varphi_{\max} = 1 - (1 - h_{\max})k_{lg}(1 + \epsilon)^{-1}$ .

It follows that the stability condition for  $|\varphi_{\min}| < 1$ :

$$-1 < 1 - \frac{(1 + h_{\max})k_{lg}}{1 - \epsilon} < 1 \implies k_{lg} < \frac{2(1 - \epsilon)}{1 + h_{\max}}$$

whereas for  $|\varphi_{\max}| < 1$  it is automatically satisfied when  $k_{lg} < 2(1 - \epsilon)(1 + h_{\max})^{-1}$  since  $(1 - h_{\max})/(1 + \epsilon) < (1 + h_{\max})/(1 - \epsilon)$  for  $\epsilon, h_{\max} < 1$ .

Under stability condition,  $|\varphi(t)| \leq \lambda$  where  $\lambda = \max\{|\varphi_{\min}|, |\varphi_{\max}|\}$ . Then  $V(\bar{e}(t)) \leq \lambda^{2t} V(\bar{e}(0))$ , which implies  $|\bar{e}(t)| \leq \lambda^t |\bar{e}(0)|$ . This completes the proof.  $\square$

## VI. SIMULATION RESULTS

To evaluate the proposed adaptive control design under a changing environment and noise, we conduct a closed loop simulation study based on the plant model described in Section III. The following sections discuss two examples covering representative scenarios in advertising.

### A. Setup

Consider a 30-day ad campaign with daily delivery targets of \$10 for days 1–10, \$5 for days 11–20, and \$15 for days 21–30. The spend controller updates once every  $\Delta = 2/60$  hours (2 minutes). At each time step, assume the observed incremental spend  $y(t)$  and the control signal  $u(t)$  satisfy

$$y(t) = \gamma(t)\Delta(1 + h(t))f(t, u(t))(1 + w(t))/24 \quad (14)$$

where  $\gamma(t)$  represents the daily spend potential,  $f(t, u)$  is the (normalized) spend versus control signal relationship, which is monotonic non-decreasing in  $u$  and at most slowly varying or subject to significant rare changes as a function of  $t$ . Seasonality function  $h(t) = 0.6\sin(2\pi t/24 - 2.8) + 0.3\sin(4\pi t/24 + 2.5)$  is a two-term Fourier series that simulates time-of-day traffic patterns, assuming 24-hour periodicity. Disturbance signal  $w(t)$  is independent and identically distributed white noise with standard deviation  $\sigma$ .

In addition to stochastic noise, we assume the campaign is subject to the following deterministic, yet unpredictable, disturbances:

- Persistent regime shift. At the start of day 15 there is an abrupt change in the competitive landscape: the daily spend potential undergoes a step change from  $\gamma = 100$  to  $\gamma = 170$ , and the response function  $f(t, u(t))$  undergoes a sudden transition. Each regime persists for half the flight, the plant gain differs by approximately  $6\times$  across the two regimes, and the controller must adapt to this change.
- Large non-Gaussian disturbance episodes. Specifically, in the time interval  $150 \leq t\Delta \leq 155$ , delivery is boosted by a factor of  $10\times$  to simulate a brief traffic surge; and in the interval  $250 \leq t\Delta \leq 255$ , delivery drops to 0 to simulate a feedback data outage. Each episode spans 150 consecutive samples with a non-zero mean, and therefore departs from the zero-mean white-noise assumption.

In the following examples, the control signal is initialized at  $u(0) = 5$ , the estimator regularization parameter  $\beta = 10^{-6}$ , and the plant gain memory coefficient  $w_1 = 0.01$ . The controller uses a launch closed-loop time constant  $T_{cl} = 0.3$  hours during the first 24 hours and  $T_{cl} = 3$  hours thereafter.

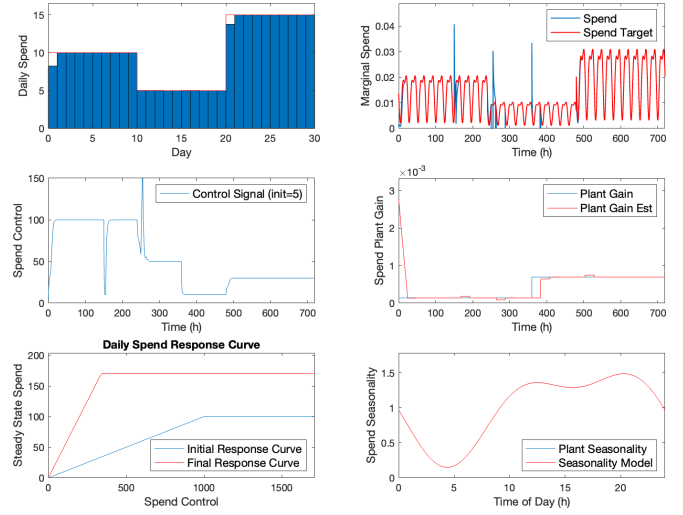


Fig. 3. Closed-loop simulation result for Example 1.

### B. Example 1: Linear Plant

The first scenario is idealized to establish a baseline and to facilitate interpretation of the closed-loop behavior. We assume no stochastic noise ( $\sigma = 0$ ), perfect seasonality compensation ( $\hat{h}(t) = h(t)$ ), and a response curve that is piecewise-linear but abruptly changing at the start of day 15. The plant gain increases by  $5\times$  and the saturation point moves from  $u = 1000$  to  $u = 340$ . This allows us to examine the closed loop effect of a persistent sudden shift in the competitive landscape.

$$f(t, u) = \begin{cases} \text{SAT}_{[0,1]}(0.1u) & \text{if } 0 < t\Delta \leq 360 \\ \text{SAT}_{[0,1]}(0.5u) & \text{if } 360 < t\Delta \leq 720 \end{cases} \quad (15)$$

The result is shown in Figure 3. The top-left panel shows the daily delivery (blue bars) versus daily delivery target (red curve). The campaign delivers in full most days. Only days subject to major changes of disturbances deviate slightly from full delivery. The top-right panel shows observed marginal spend (blue) versus intraday delivery (red) target produced by seasonality compensator. The curves overlap for most of the flight, with deviation only in brief transient periods during major campaign changes or disturbances.

The middle-left panel displays the spend control signal  $u$  which is updated based on error feedback to deliver the budget in full. It responds quickly to any external campaign change or disturbance to ensure full delivery.

The middle-right panel depicts the true plant gain (blue) and estimated plant gain (red). The true plant gain is a priori unknown to the control system. The estimated plant gain is computed based on the proposed recursive plant gain estimator. The estimate shows zero steady-state error in this scenario because the simulated plant is linear, which matches the estimator's assumed model  $y = k_p u$ , and free of stochastic noise; i.e., it converges to the true plant gain after any campaign change or disturbance.

The bottom-left panel shows daily spend versus spend control signal curve  $\gamma f(u)$  for the first 15 days (blue) and last 15 days (red). The abrupt change in the competitive landscape at the start of day 15 is apparent in the two curves.

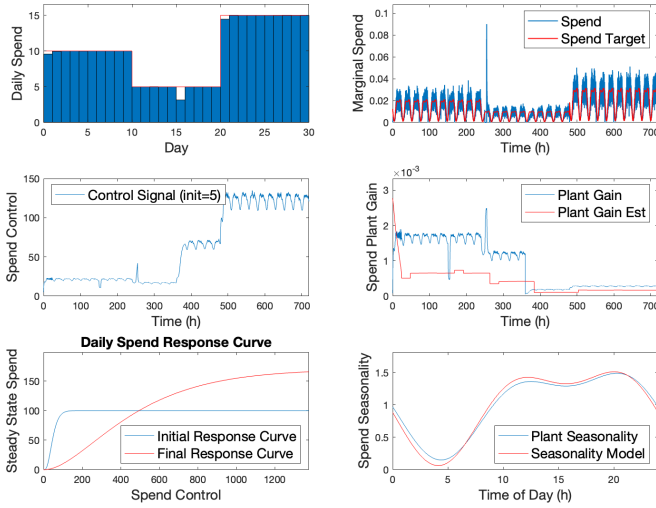


Fig. 4. Closed-loop simulation result for Example 2.

The bottom-right panel displays the true seasonality of the plant (blue) and the seasonality model (red) used by the seasonality compensator to allocate the daily budget throughout the day via feedforward control. The two curves are identical (overlapping) in this scenario because we assume perfect knowledge of the time-of-day pattern of the plant.

### C. Example 2: Nonlinear Plant

This example uses a more realistic set of conditions. First, the response curve is no longer linear but is defined by  $f(t, u) = \Gamma(u|\mu, \sigma_{rel})$ , where  $\Gamma(u|\mu, \sigma_{rel})$  denotes the cumulative distribution function of a gamma random variable with mean  $\mu$  and relative standard deviation  $\sigma_{rel}$ . Assume  $\mu$  and  $\sigma_{rel}$  vary over time as follows:

$$(\mu, \sigma_{rel}) = \begin{cases} (50, 0.5) & \text{if } 0 < t\Delta \leq 360 \\ (500, 0.7) & \text{if } 360 < t\Delta \leq 720 \end{cases} \quad (16)$$

Let  $\sigma = 0.2$  and assume the controller has an imperfect knowledge  $\hat{h}(t)$  of plant seasonality  $h(t)$ . Specifically,  $\hat{h}(t)$  overstates each harmonic amplitude by 10% and advances each harmonic phase by 0.1 radians relative to the plant, so the seasonality compensator overestimates the daily swing and places its peak slightly too early.

Figure 4 displays the closed-loop simulation result. The daily delivery (top left) is similar to Example 1, while the marginal spend (as expected) is noticeably impacted by the stochastic noise (top right). The spend control signal now exhibits a distinct time-of-day pattern driven by the imperfect seasonality model. The plant gain (middle right) is noticeably different. First, the true plant gain (blue curve) oscillates due to oscillation in the spend control signal (triggered by the imperfect seasonality profile). Furthermore, the plant gain estimate is biased because the response curve model used by the plant gain estimator ( $y \approx k_p u$ ) is now an imperfect approximation of the more accurate representation defined by  $y = m + k_p u$ . The biased plant gain estimate is not an issue under feedback control. If the loop gain  $k_{lg}$  is selected wisely to ensure stability under the uncertain plant gain estimate, convergence and zero steady-state delivery error are ensured.

### D. Robustness Analysis and Fixed-Gain Baseline

The gain margin of the closed-loop system in Example 2 provides insight into how the example's conclusions hold for plants deviating from the assumed model. An approximation of the effective gain margin  $g = 2/(k_p k_c)$  is displayed in the bottom-left panel of Figure 5, while a formal computation of exact gain margins for the time-varying periodic system being future work (based on Theorem 6).

The gain margin remains approximately constant within each regime, with medians of 67 and 107, corresponding to a closed-loop realized time constant between 1.1 and 1.8 hours (for a specified desired time constant of 3 hours). Adaptation absorbs the  $6\times$  change in plant gain into a  $1.6\times$  change in stability margin. The residual variation is due to the bias in plant gain estimation. Margins fall below this band only during the launch transient and the two disturbance episodes, reaching a worst case of 47.

In contrast, a fixed-gain controller cannot hold this invariance. Because  $g \propto 1/k_p$  for constant  $k_c$ , the margin ratio across the two regimes is the plant gain ratio, and no choice of fixed gain compress it. To demonstrate this we repeat the Example 2 simulation with the adaptive gain replaced by a constant. Two fixed values are considered:  $k_c = 15$ , corresponding to a conservative controller tuned for the highest plant gain of the flight, and  $k_c = 110$  tuned for the lowest. Figure 5 shows the result.

The conservative controller maintains a large stability margin of 82 in high-plant-gain regime, but its realized time constant degrades to 7.9 hours in low-plant-gain regime,  $4.5\times$  slower than adaptation and  $2.6\times$  slower than the 3-hour specification, so campaign under-delivers on the days following each plant dynamics change. The aggressive controller recovers delivery speed, with a realized time constant of at most 1.1 hours throughout, but its median margin in the high-plant-gain regime falls to 11, with a worse case of 7.3, substantially lower than the 60+ achieved by adaptation. The margin of both fixed designs varies by nearly  $6\times$  across the flight as predicted, and neither achieves the balance between delivery speed and stability margin that adaptation can.

## VII. CONCLUSIONS AND FUTURE WORK

This paper proposes an adaptive spend control architecture for budget delivery in programmatic advertising, combining input re-parameterization, an online plant gain estimator that recursively identifies the campaign-specific plant gain from observed data, and an adaptive feedback controller that computes its gain from the plant gain estimate.

We establish via Lyapunov theory exponential convergence guarantees for the control loop acting on an idealized deterministic plant, assuming the estimator maintains its designed conservative bias. Closed-loop simulations under realistic conditions demonstrate full budget delivery with steady-state tracking errors below 1% across campaigns exhibiting  $5\times$  to  $6\times$  plant gain variation. The adaptive controller achieves convergence times of 1.1 to 1.8 hours, up to  $4.5\times$  faster than conservatively-tuned fixed-gain control, while maintaining gain margins between 67 and 107 (a  $3.5\times$

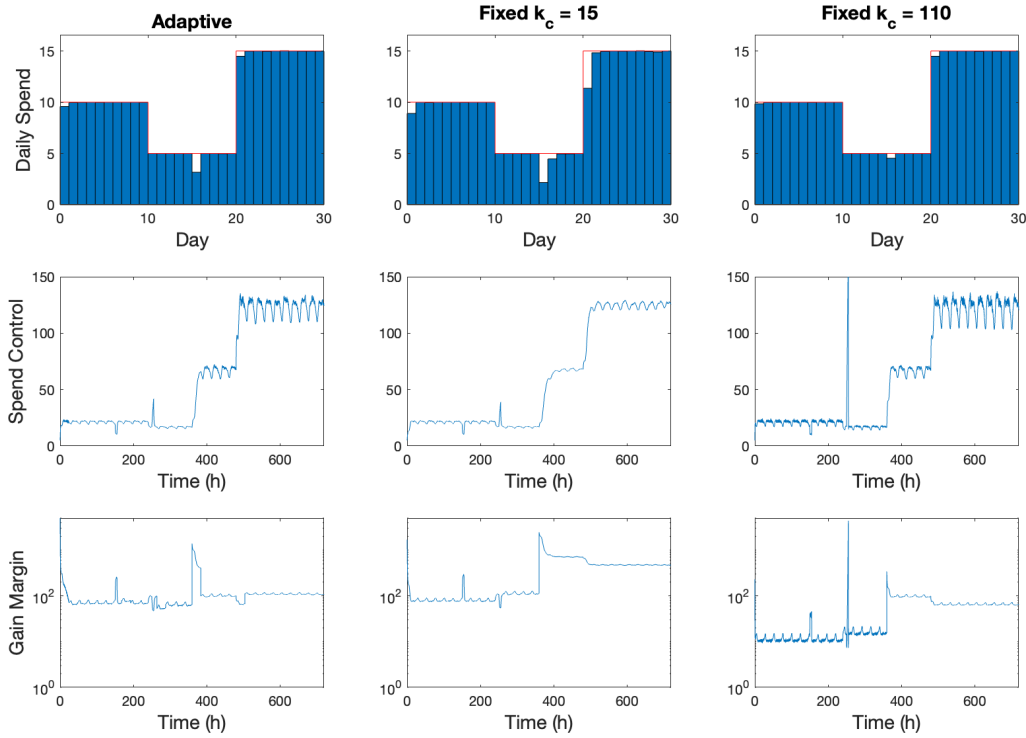


Fig. 5. Closed-loop comparison of fixed-gain and adaptive control on the Example 2 plant. Columns: adaptive, fixed  $k_c = 15$ , fixed  $k_c = 110$ . Rows: Daily spend, control signal, and gain margin on logarithmic scale.

to  $6\times$  improvement over fixed-gain baselines of comparable convergence speed).

Future work includes: (i) a stability analysis of the interconnected estimator-controller system under imperfect feedback linearization, seasonality compensation errors, and stochastic noise; (ii) large-scale A/B validation on production traffic to quantify real-world delivery improvement; and (iii) development of a joint estimator that simultaneously identifies plant gain and seasonality parameters within a single recursive framework (e.g., via an extended Kalman filter), improving robustness when the nominal seasonality model is poor.

Finally, empirical observations reveal that the plant can operate under qualitatively different dynamics for extended periods rather than experiencing isolated point disturbances. A natural research direction is to model the plant gain as a discrete-mode process, for example, via a jump-diffusion model or a multiple-model adaptive estimator, which would provide more principled handling of regime shifts.

## REFERENCES

- [1] Mordor Intelligence. Programmatic advertisement market size, share, trends & industry outlook, 2030. <https://www.mordorintelligence.com/industry-reports/global-programmatic-advertisement-market>, 2025.
- [2] Niklas Karlsson and Jianlong Zhang. Applications of feedback control in online advertising. In *Proceedings of the 2013 American Control Conference, Washington, DC, June 17 - 19*, pages 6008 – 6013, 2013.
- [3] Jun Wang, Weinan Zhang, and Shuai Yuan. Display advertising with real-time bidding (RTB) and behavioural targeting. 2016.
- [4] Weinan Zhang, Yifei Rong, Jun Wang, Tianchi Zhu, and Xiaofan Wang. Feedback control of real-time display advertising. *Proceedings of the 9th ACM International Conference on Web Search and Data Mining, New York, NY, USA*, pages 407–416, 2018.
- [5] Niklas Karlsson. Plant gain estimation in online advertising processes. In *Proceedings of the 2017 IEEE 56th Conference on Decision and Control, Melbourne, Australia*, pages 2182 – 2187, Dec 12-15 2017.
- [6] Niklas Karlsson. Feedback control in programmatic advertising: The frontier of optimization in real-time bidding. *IEEE Control Systems Magazine*, 40(5):40–77, 2020.
- [7] Karl Johan Åström and Björn Wittenmark. *Adaptive Control (2 ed.)*. Addison-Wesley, 1995.
- [8] S. S. Sastry and M. Bodson. *Adaptive control: stability, convergence, and robustness*. Prentice-Hall, Upper Saddle River, NJ, 1989.
- [9] Hassan K. Khalil. *Nonlinear Systems*. Prentice Hall, Upper Saddle River, N.J., third edition, 2002.
- [10] Vijay Krishna. *Auction Theory*. Academic Press, 2002.
- [11] Niklas Karlsson. Scalable multi-objective optimization in programmatic advertising via feedback control. In *Proceedings of the 2021 60th IEEE Conference on Decision and Control (CDC), Austin, TX, USA*, pages 1363–1370, Dec 2021.
- [12] Niklas Karlsson. Feedback control-based multiobjective optimization in programmatic advertising involving a cost per bid constraint. In *2022 IEEE 61st Conference on Decision and Control (CDC)*, pages 505–510, 2022.
- [13] Junwei Pan, Jian Xu, Alfonso Lobos Ruiz, Wenliang Zhao, Shengjun Pan, Yu Sun, and Quan Lu. Field-weighted factorization machines for click-through rate prediction in display advertising. In *Proceedings of the World Wide Web Conference, WWW '18*, pages 1349–1357, 2018.
- [14] Tian Zhou, Hao He, Shengjun Pan, Niklas Karlsson, Bharatbhushan Shetty, Brendan Kitts, Djordje Gligorijevic, San Gultekin, Tingyu Mao, Junwei Pan, Jianlong Zhang, and Aaron Flores. An efficient deep distribution network for bid shading in first-price auctions. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'21)*, 8/14-18, Singapore, 2021.
- [15] Hao He and Niklas Karlsson. Identification of seasonality in internet traffic to support control of online advertising. In *2019 American Control Conference (ACC)*, pages 3835–3840, 2019.