Adaptive Optimization and Control in Online Advertising

Proceedings of the 58th IEEE Conf. on Decision and Control, Nice, France, December 11 - 13, 2019

Niklas Karlsson[1]

***Abstract*— Optimization in online advertising typically involves feedback control as a critical component. Here we propose a control system that maximizes the *return on investment* (ROI) for an advertiser and paces the budget delivery. It consists of an integral controller with periodic feedforward compensation of the set-point and persistent excitation. We derive stability conditions for the controller.**

Keywords: Periodic control, Floque theory, Online Advertising, ROI maximization

I. Introduction

Programmatic advertising is at the heart of the business model for companies such as Google, Facebook, and Verizon Media. A *Demand Side Platform* (DSP) is a particular business model for programmatic advertising. It is the middleman between an advertiser and one or more open exchange trading so called ad impressions. An *impression* is an opportunity of showing an ad creative, e.g., banner ad, text ad, or pre-roll video commercial to Internet users. The goal of a DSP is to manage an advertiser's advertisement budget optimally.

The optimization is challenging due to the underlying high-dimensional, nonlinear, time-varying, dynamic, and stochastic plant. An early pre-DSP publication on feedback control applied to online advertising [1] outlined several important challenges, but omitted detailed solutions. A more comprehensive overview of the control problem was presented in [2] in which bid randomization [3] was proposed to overcome complexities due to a discontinuous plant. Different approaches for control of advertising processes are proposed in [4], [5], [6], [7], [8], but none of these papers adequately consider both the dynamic and periodic nature of the plant. Statistical inference of the plant gain based on an uncertain discontinuous plant response curve combined with bid randomization is presented in [9].

Our contribution is a proposed controller structure and conditions under which it ensures global asymptotic stability.

The paper is organized as follows. In Section II we introduce the ad optimization-turned-control problem. Thereafter, in Section III, we show how the otherwise discontinuous plant is turned smooth. Section IV introduces a plant model and Section V suggests a plant identification algorithm. The core of the paper is in Section VI, where the controller structure is proposed and the stability conditions are derived. Section VII contains a simulated example of the closed loop system and Section VIII some concluding remarks.

II. Optimization turned Control Problem

The (optimization) objective is to spend an advertiser's online marketing budget in such a way that her total advertisement value, or *return on investment* (ROI), is maximized. Provided by the advertiser is a monetary budget, which is typically a fixed daily budget for the duration of a campaign flight. Given is also an optimization metric, such as, click or conversion count, which are examples of performance metrics, or brand value. The advertiser may associate a different value to each click or conversion, and almost always associates a different brand value to each impression based on user characteristics and site properties. The total value of an impression i is denoted v_i. This metric encodes the combined brand and performance value. For example, $v_i = v_{B,i} + p_{CTR,i} v_{C,i}$, where $v_{B,i}$ is the brand value of serving our ad to this user, $p_{CTR,i}$ is the probability the user will click on the ad, and $v_{C,i}$ is the advertiser-defined value of a click generated by the user.

Impressions are purchased on open exchanges where they are sold by publishers or *Supply Side Platforms* (SSPs), under a sealed second price auction. The bid submitted by the DSP on behalf of the advertiser for impression opportunity i is denoted b_i, while the highest competing bid price is denoted $b_i^\star$. The impression is awarded to the highest bidder, hence our advertiser will have an ad served to the user if $b_i \geq b_i^\star$ in which case the advertiser is charged $b_i^\star$ for the impression.

Define *total* (advertisement) *value* v of an ad campaign as the cumulative value of all awarded impressions, and *total cost* c as the cumulative cost for these impressions. The second price cost model implies that $v = \sum_i v_i \mathbb{I}_{\{b_i \geq b_i^\star\}}$ and $c = \sum_i b_i^\star \mathbb{I}_{\{b_i \geq b_i^\star\}}$, where $\mathbb{I}_X$ is the indicator function satisfying $\mathbb{I}_X = 1$, if $X =$ true, and $\mathbb{I}_X = 0$, if $X =$ false. It is known that the maximization of v subject to a budget constraint $c \leq u_{cost}^{ref}$ is achieved using a bidding strategy $b_i = uv_i$, where u is selected as the largest single value for which the budget constraint is not violated [2]. This result is illustrated graphically in Figure 1.

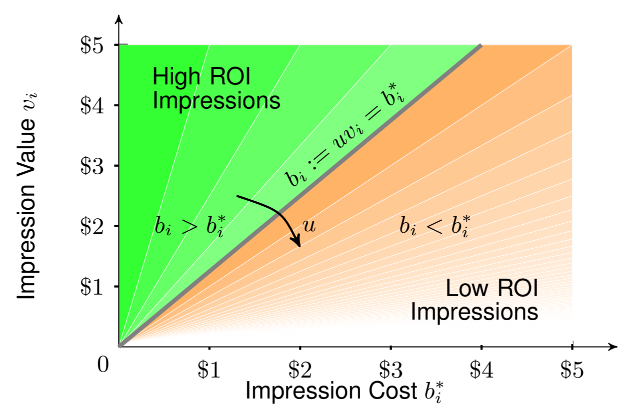

Fig. 1. The plot shows the relative ROI of different impression opportunities, where each impression is associated with an impression value v_i and a cost $b_i^\star$. High ROI impressions are located in the upper left corner, and low ROI impressions in the bottom right.

Each impression we bid for is associated with a value v_i and a cost $b_i^\star$, hence can be mapped to a coordinate in the value versus cost plot. Impressions in the upper left corner

[1]N. Karlsson is Chief Scientist and Vice President of Demand Platforms R&D at Verizon Media, 741 1st Ave., Sunnyvale, CA 94089, USA nkarlsson8 at gmail.com

with high value and low cost correspond to the highest ROI impressions since they translate to the largest value per ad dollar spend. Impressions along a straight line going through the origin all have the same ROI, and impressions in the lower right corner with low value and high cost correspond to the lowest ROI impressions. Finally, impressions in the green shaded region all have higher ROI than impressions in the orange shaded region.

By bidding along the straight line, we win the impressions in the green region and the optimization problem is effectively decoupled into an impression value computation problem and a control problem. The value computation problem is really a prediction problem and involves large scale machine learning based on user features and historical engagement data to compute the best possible estimate of v_i for which we are about to bid. Independently, a feedback controller adjusts u dynamically based on delivered budget. In this paper we assume the estimate of v_i is available and focus on the control problem. An important property using bidding mechanism $b_i = uv_i$ is that $v(u)$, $c(u)$, and $v(u)/c(u)$ are monotonic functions of u. It ensures we deal with a convex problem. Convexity is in general lost for other bidding mechanisms.

Because of the natural time-of-day pattern in Internet users' presence online there is a dramatic seasonality in the available number of impression opportunities. At the daily high there are about ten times more people online compared to near the daily low. A naive control approach does not take the supply seasonality into account and strive to spend the same amount of the budget every hour of the day. This leads to overbidding for impressions nighttime when there are few impressions available to buy. Furthermore, Internet traffic is random with an approximately scale-invariant (Poisson-like) standard deviation.

Finally, due to the auction-based allocation of impressions, the relationship between control signal and ad spend (or cost) is described by a discontinuous, staircase-like relationship as illustrated in Figure 2. The awarded number of impressions

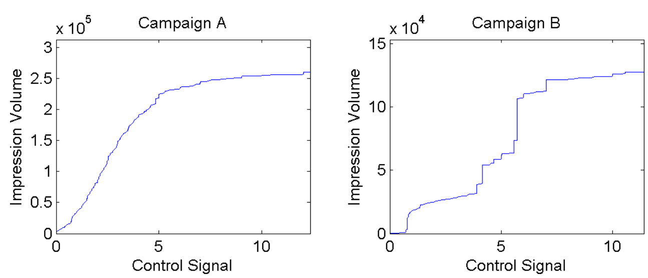

Fig. 2. Two representative examples of awarded number of impressions versus control signal u. Campaign A is a well-behaved campaign with an approximately smooth relationship while Campaign B is a challenging campaign with pronounced large steps in the relationship.

as a function of u is shown for two real ad campaigns. Campaign A is well-behaved while Campaign B is not.

In the remainder of the paper, we assume the controller is implemented in discrete time. Control signal $u(t)$ is updated at discrete time points $t = 1, 2, \ldots$ (without loss of generality) based on a set-point signal $\bar{u}_c(t) > 0$, defined by the advertiser, and an observed marginal ad spend $y(t)$. However, the advertiser cares primarily about the daily pacing and value maximization hence the optimization system may distribute the spend throughout the day in the most economical way.

III. Plant Smoothing

A plant that is not approximately continuous leads to an extremely complex closed-loop dynamics. Fortunately, we may always use Heisenberg bidding [2], [3] to turn a plant effectively continuous. It is a bid randomization technique by which each computed *nominal bid* $b_{0,i}$ $(= uv_i)$ is perturbed randomly before submitted to the auction exchange. It can be implemented with other probability distributions, but the gamma distribution is a particularly convenient choice since it has support for all positive values of $b_{0,i}$ and brings along many useful properties as a member of the family of exponential distributions.

In this case, the distribution is parameterized by $b_{0,i}$ and *uncertainty signal* u_u, to generate a *final bid price* b_i used in the market clearing. In particular, b_i is a realization of a random variable B_i defined by $B_i \sim \text{Gamma}\left(1/u_u^2, 1/(b_{0,i}u_u^2)\right)$ if $b_{0,i}, u_u > 0$, and $B_i = b_{0,i}$, otherwise. In terms of the shape parameter α and the inverse scale parameter β, Heisenberg bidding is defined by $\alpha = 1/u_u^2$ and $\beta = 1/(b_{0,i}u_u^2)$. Hence, $\mathrm{E}B_i = b_{0,i}$ and $\mathrm{Var}(B_i) = b_{0,i}^2 u_u^2$. In other words, $\mathrm{Std}(B_i)/\mathrm{E}B_i = u_u$, where $\mathrm{Std}(B_i)$ is the standard deviation of B_i.

The input-output relationship of the plant can be made arbitrarily smooth by adjusting the uncertainty signal. This is illustrated in Figure 3, which shows the result of adding

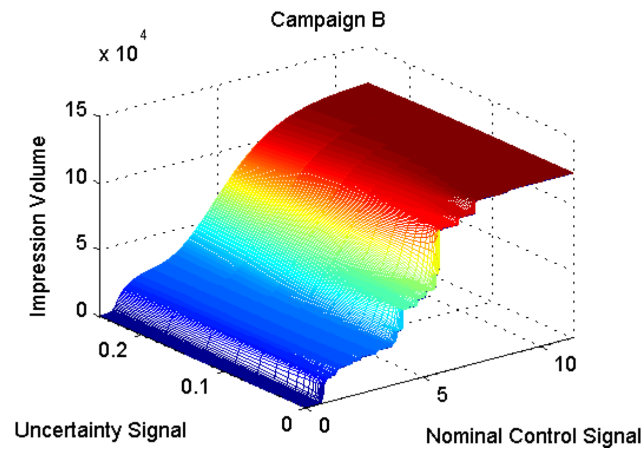

Fig. 3. The impact of using bid uncertainty on Campaign B in Figure 2. Note that the nominal control signal versus impression volume relationship is discontinuous only when the uncertainty signal equals zero.

the dimension of uncertainty signal to Campaign B in the previous figure.

IV. Plant Model

The relationship between control signal $u(t)$ and *delay-free ad spend* $\acute{y}(t)$, in general, is nonlinear, time-varying, and stochastic according to $\acute{y}(t) = f(t, u(t), w_m(t))$ for some non-negative function f that is increasing in u and stochastic in w_m. However, for each time t, $\mathrm{E}(\acute{y})$ is linearizable in some small neighborhood of each u (potentially after invoking Heisenberg bidding). We assume the time-variability caused by a time-varying Internet traffic is T-periodic and scale-invariant. The expected ad spend near control signal $u(t)$ can then be written $\mathrm{E}(\acute{y}(t)) = (1 + h(t))(a_0 + a_1 u(t))$ for some values of *plant offset* a_0 and *plant gain* $a_1 > 0$, and for some T-periodic function $h(t)$ satisfying $h(t+T) = h(t)$, $\sum_{i=1}^{T} h(i) = 0$, and $h(t) > -1, \forall t$.

The ad spend is subject to approximately scale-invariant stochastic noise due to traffic variations, which is represented by multiplicative white noise. Finally, the observed ad spend $y(t)$ is modeled by $\acute{y}(t)$ subject to a first order measurement latency defined by a *plant latency time constant* T_{P}, which is mapped to the discrete-time model parameter $a_3 = e^{-\Delta/T_{\mathrm{P}}}$, where Δ is sampling time. In conclusion, the plant model is

$$\begin{aligned} y(t) - a_3 y(t-1) &= (1-a_3)(1+h(t-1)) \\ &\cdot(a_0 + a_1 u(t-1))(1+w_m(t-1)), \end{aligned} \quad (1)$$

where $w_m(t) \sim \mathrm{WN}(0, \sigma_m^2)$[1]. The system delay parameter a_3 and the seasonality function $h(t)$ can be estimated offline [10], while a_0 and a_1 are unknown a priori.

V. Plant Identification

Stability of the closed loop system is tightly coupled to the relationship between plant and controller properties. It is therefore important to estimate the unknown a_0 and a_1. Plant model (1) may be rewritten as

$$z(t) := \frac{y(t) - a_3 y(t-1)}{(1-a_3)(1+h(t-1))} = a_0 + a_1 u(t-1) + \tilde{w}_m(t-1), \quad (2)$$

where $\tilde{w}_m(t) \sim \mathrm{WN}\big(0, (a_0 + a_1 u(t))^2 \sigma_m^2\big)$. Note, $z(t)$ is a sequence of uncorrelated random variables with $\mathrm{E}\big(z(t)|u(t-1)\big) = a_0 + a_1 u(t-1)$ and $\mathrm{Var}\big(z(t)|u(t-1)\big) = (a_0 + a_1 u(t-1))^2 \sigma_m^2$. A prediction of $z(t)$ is given by $\hat{z}(t) = \hat{a}_0 + \hat{a}_1 u(t-1)$ based on estimates $\hat{a}_0$ and $\hat{a}_1$.

Assume $u(t)$ operates in a reasonably small neighborhood of some operating point. Then $\tilde{w}_m(t)$ has approximately constant variance and $\theta := [a_0, a_1]^T$ can be estimated in closed form by minimizing the sum of squares of residuals, $z(t) - \hat{z}(t)$, based on measurement $u(i), y(i)$, $i = 1, \ldots, t$. This solution may be implemented as a standard *Recursive Least Squares* (RLS) estimator with exponential memory loss to allow for time-varying plant parameters.

VI. Control

A state space representation of plant model (1) is

$$\begin{aligned} x_1(t+1) &= a_3 x_1(t) + (1-a_3)(1+h(t)) \\ &\cdot(a_0 + a_1 u(t))(1+w_m(t)) \quad (3) \\ y(t) &= x_1(t) \quad (4) \end{aligned}$$

To solve the control (aka, optimization) problem, we propose using pure *integral* (I) error feedback servo control with feedforward adjustment of the reference signal, and open loop persistent excitation of the servo control signal. The feedforward controller modifies the deterministic reference input signal $\bar{u}_c(t)$ based on known values of a_3 and $h(t)$. The idea is to distribute a daily ad budget throughout the day according to the seasonality of impression supply. This avoids a situation where the control system raises the control signal and hikes the bid prices during times of the day when there is a limited impression supply. The feedforward controller dynamics is described by

$$\begin{aligned} x_3(t+1) &= a_3 x_3(t) + (1-a_3)(1+h(t))\bar{u}_c(t) \quad (5) \\ u_c(t) &= x_3(t) \quad (6) \end{aligned}$$

The I-controller computes a servo control signal $u_0(t)$ based on a tracking error signal $e(t)$ with the goal of minimizing or reducing the marginal and cumulative tracking error and leveraging an approximately constant $u_0(t)$; and thereby maximizing the ROI of the campaign. The feedback controller dynamics is defined by

$$\begin{aligned} e(t) &= u_c(t) - y(t) \quad (7) \\ x_2(t+1) &= x_2(t) + c_{\mathrm{I}} e(t) \quad (8) \\ u_0(t) &= x_2(t) + c_{\mathrm{I}} e(t) \quad (9) \end{aligned}$$

To guarantee stability, the *integral gain* c_{I} must be chosen as a function of a_1, or in practice as a function of an estimate $\hat{a}_1$ thereof. However, throughout the analysis in this section we assume c_{I} is a given constant.

Plant identification requires persistent excitation and to ensure this, we complement the I-controller with a perturbation system that computes a final control signal $u(t)$ from

$$\begin{aligned} x_4(t+1) &= a_4 x_4(t) + (1-a_4) w_u(t) \quad (10) \\ u(t) &= u_0(t)(1+x_4(t)), \quad (11) \end{aligned}$$

where $w_u(t)$ is an artificially generated $\mathrm{WN}(0, \sigma_u^2)$ perturbation signal. In practice, we parameterize the excitation via a *persistent excitation time constant* T_{PE} and compute $a_4 = e^{-\Delta/T_{\mathrm{PE}}}$. Using a dynamic in the persistent excitation, as opposed to only white noise, reduces the impact of an imperfect knowledge of the plant latency (modeled by parameter a_3).

The plant is represented by state x_1 and the controller by states x_2, x_3, and x_4. We judge the performance of the closed loop system based on performance signals $e(t)$ and $u(t)$. The desired behavior is for $e(t)$ to stay close to zero and $u(t)$ close to a constant. This corresponds to a situation where the maximum value for the advertiser is generated, and where the daily delivery pacing is even. The controller defined by (5)-(11) is summarized in Algorithm 1.

Algorithm 1: Control

Parameters: $c_{\mathrm{I}}, \sigma_u, h(t), a_3$; (in practice, c_{I} is a function of a_1)
Input: $\bar{u}_c(t), y(t)$; (in practice, $\hat{a}_1(t)$ is also an input used to compute $c_{\mathrm{I}} = c_{\mathrm{I}}(\hat{a}_1)$)
State: x_2, x_3, x_4
Initialization ($t = 0$):
  $x_2(0) = x_{2,0}$
  $x_3(0) = x_{3,0}$
  $x_4(0) = 0$
For each instant of time, $t = 1, 2, \ldots$, compute:
  $u_c(t) = x_3(t)$
  $e(t) = u_c(t) - y(t)$
  $u_0(t) = x_2(t) + c_{\mathrm{I}} e(t)$
  $u(t) = u_0(t)\big(1 + x_4(t)\big)$
  $w_u(t) \sim \mathrm{WN}(0, \sigma_u^2)$
  $x_2(t+1) = x_2(t) + c_{\mathrm{I}} e(t)$
  $x_3(t+1) = a_3 x_3(t) + (1-a_3)\big(1+h(t)\big)\bar{u}_c(t)$
  $x_4(t+1) = a_4 x_4(t) + (1-a_4) w_u(t)$

Now, analyze the closed-loop dynamics and derive fixed point solutions and limit cycles with stability conditions. Plug (4) and (6) into (7) and combine the result with (9)

[1] A *white noise* (WN) process is a random process of random variables that are uncorrelated, have mean zero, and a finite variance.

and (11), to obtain $e(t)$ and $u(t)$ as functions of the states

$$\begin{aligned} e(t) &= -x_1(t) + x_3(t) && (12) \\ u(t) &= -c_\text{I}(1 + x_4(t))x_1(t) + (1 + x_4(t))x_2(t) \\ &\quad + c_\text{I}(1 + x_4(t))x_3(t). && (13) \end{aligned}$$

Next, plug (13) into (3) and (12) into with (8); and define $\varphi_0(t) := (1 - a_3)(1 + h(t))$ and $\varphi_1(x_4, w_m) := x_4(t) + w_m(t) + x_4(t)w_m(t)$. Note that $\varphi_0(t)$ is deterministic and periodic; while $\varphi_1(x_4, w_m)$ is time-invariant, stochastic, and nonlinear in the uncorrelated random variables x_4 and w_m. This leads to the closed loop state update equations

$$\begin{aligned} x_1(t+1) &= \big(a_3 - a_1 c_\text{I}\varphi_0(t)\big(1 + \varphi_1(x_4, w_m)\big)\big)x_1(t) \\ &\quad + a_1\varphi_0(t)\big(1 + \varphi_1(x_4, w_m)\big)x_2(t) \\ &\quad + a_1 c_\text{I}\varphi_0(t)\big(1 + \varphi_1(x_4, w_m)\big)x_3(t) \\ &\quad + a_0\varphi_0(t)\big(1 + w_m(t)\big) && (14) \\ x_2(t+1) &= x_2(t) - c_\text{I}x_1(t) + c_\text{I}x_3(t) && (15) \\ x_3(t+1) &= a_3x_3(t) + \varphi_0(t)\bar{u}_c(t) && (16) \\ x_4(t+1) &= a_4x_4(t) + (1 - a_4)w_u(t) && (17) \end{aligned}$$

Note, x_3 and x_4 evolve independently of x_1, x_2, and w_m; and have closed form solutions

$$x_3(t) = a_3^t x_3(0) + \sum_{i=1}^{t} a_3^{i-1}\varphi_0(t-i)\bar{u}_c(t-i) \quad (18)$$

$$x_4(t) = a_4^t x_4(0) + (1 - a_4)\sum_{i=1}^{t} a_4^{i-1}w_u(t-i) \quad (19)$$

It suffices to analyze x_1, x_2 while treating x_3, x_4 as exogenous input signals. Define

$$\begin{aligned} \tilde{x}_1(t) &= c_\text{I}\big(x_1(t) - x_3(t)\big) && (20) \\ \tilde{x}_2(t) &= x_2(t) - \frac{\bar{u}_c - a_0}{a_1} && (21) \end{aligned}$$

Combining (20)-(21) with (14)-(15) yields the state update equations for $\tilde{x}_1(t)$ and $\tilde{x}_2(t)$

$$\begin{aligned} \tilde{x}_1(t+1) &= \big(a_3 - a_1 c_\text{I}\varphi_0(t)\big(1 + \varphi_1(x_4, w_m)\big)\big)\tilde{x}_1(t) \\ &\quad + a_1 c_\text{I}\varphi_0(t)\big(1 + \varphi_1(x_4, w_m)\big)\tilde{x}_2(t) \\ &\quad + c_\text{I}\varphi_0(t)\varphi_1(x_4, w_m)\bar{u}_c - c_\text{I}\varphi_0(t)\varphi_1(x_4, w_m)a_0 \\ &\quad + c_\text{I}\varphi_0(t)\varphi_1(x_4, w_m)w_m(t) \\ \tilde{x}_2(t+1) &= -\tilde{x}_1(t) + \tilde{x}_2(t), \end{aligned}$$

and combined with (12) and (13) we obtain the following output equations for $u(t)$ and $e(t)$

$$\begin{aligned} u(t) &= (1 + x_4(t))\left(-\tilde{x}_1(t) + \tilde{x}_2(t) + \frac{\bar{u}_c - a_0}{a_1}\right) && (22) \\ e(t) & \quad -\frac{1}{c_\text{I}}\tilde{x}_1(t). && (23) \end{aligned}$$

The obtained state update and output equations can be expressed in matrix form

$$\begin{aligned} \tilde{x}(t+1) &= A(t, x_4, w_m)\tilde{x}(t) + B\epsilon_1(t, x_4, w_m) && (24) \\ z(t) &= C(x_4)\tilde{x}(t) + D\epsilon_2(x_4) && (25) \end{aligned}$$

where $\tilde{x}(t) = [\tilde{x}_1(t), \tilde{x}_2(t)]^T$, $z(t) = [u(t), e(t)]^T$, and

$$\begin{aligned} A(t, x_4, w_m) &= \begin{bmatrix} a_3 - a_1 c_\text{I}\varphi_0\big(1 + \varphi_1\big), & a_1 c_\text{I}\varphi_0\big(1 + \varphi_1\big) \\ -1, & 1 \end{bmatrix} && (26) \\ B &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} && (27) \\ C(x_4) &= \begin{bmatrix} -1 - x_4(t), & 1 + x_4(t). \\ -1/c_\text{I}, & 0 \end{bmatrix} && (28) \\ D &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} && (29) \\ \epsilon_1(t, x_4, w_m) &= c_\text{I}\varphi_0(t)\big(\varphi_1(x_4, w_m)(\bar{u}_c - a_0) + a_0 w_m\big) && (30) \\ \epsilon_2(x_4) &= \frac{\bar{u}_c - a_0}{a_1}(1 + x_4) && (31) \end{aligned}$$

State-transition matrix A is a T-periodic random matrix, which depends on the exogenous random input signals x_4 and v_m via φ_1, but not on the state $\tilde{x}$. Matrix C is a random matrix, which depends only on the exogenous random input signal x_4. B and D are constant vectors, while ϵ_1 and ϵ_2 are scalar input signals constructed from x_4 and v_m.

Equations (24) and (25) encode a wealth of information dictating the performance, stability, and robustness of the closed loop system. The following result is useful towards extracting some of this information.

Lemma 6.1: $\mathrm{E}(\varphi_1) = \mathrm{E}(\epsilon_1) = 0$, $\mathrm{E}(\epsilon_2) = (\bar{u}_c - a_0)/a_1$.

Proof: First, $\mathrm{E}(\varphi_1(x_4, w_m)|x_4) = \mathrm{E}(x_4 + w_m + x_4 w_m|x_4) = x_4 + \mathrm{E}(w_m) + x_4\mathrm{E}(w_m) = x_4$. It is trivial to show that $\mathrm{E}x_4 = 0$, and by invoking the Law of Total Expectation[2] we obtain $\mathrm{E}(\varphi_1(x_4, w_m)) = \mathrm{E}(\mathrm{E}(\varphi_1(x_4, w_m)|x_4)) = 0$.

Thereafter, by using the linearity property of the expectation operator and invoking $\mathrm{E}(\varphi_1) = 0$, we obtain $\mathrm{E}(\epsilon_1) = c_\text{I}\varphi_0\big((\bar{u}_c - a_0)\mathrm{E}(\varphi_1) + a_0\mathrm{E}(w_m)\big) = 0$.

Finally, the expected value of ϵ_2 is trivially obtained $\mathrm{E}(\epsilon_2) = \mathrm{E}\left((\bar{u}_c - a_0)(1 + x_4)/a_1\right) = (\bar{u}_c - a_0)/a_1$, which completes the proof. ■

We are now equipped to prove the following key result.

Theorem 6.1: If $\bar{u}(t)$ is constant, then $\mathrm{E}\big([x_1, x_2, x_3, x_4]\big)$ has a limit cycle $[x_1^\star, x_2^\star, x_3^\star, x_4^\star]$ given by

$$\begin{bmatrix} x_1^\star(t) \\ x_2^\star(t) \\ x_3^\star(t) \\ x_4^\star(t) \end{bmatrix} = \begin{bmatrix} \bar{u}_c \sum_{i=1}^{\infty} a_3^{i-1}\varphi_0(t-i) \\ (\bar{u}_c - a_0)/a_1 \\ \bar{u}_c \sum_{i=1}^{\infty} a_3^{i-1}\varphi_0(t-i) \\ 0 \end{bmatrix},$$

while the performance signal $\mathrm{E}(z^T) := \mathrm{E}([u, e])$ has a fixed point solution $[u^\star, e^\star]$ given by

$$z^\star := \begin{bmatrix} u^\star \\ e^\star \end{bmatrix} = \begin{bmatrix} \frac{\bar{u}_c - a_0}{a_1} \\ 0 \end{bmatrix}.$$

The limit cycle and fixed point solution are globally asymptotically stable if and only if the eigenvalues of $\bar{A}(1, 0, 0) := A(T, 0, 0)A(T-1, 0, 0)\cdots A(1, 0, 0)$ are strictly inside the unit circle.

Proof: The trajectory of $x_3(t)$ is deterministic and given by (18). Since $0 \le a_3 < 1$, $\bar{u}(t)$ is constant, and $\varphi_0(t)$ is bounded, it can be written as

$$x_3(t) = a_3^t x_3(0) + x_3^\star(t) - \bar{u}_c \sum_{i=t+1}^{\infty} a_3^{i-1}\varphi_0(t-i),$$

[2] Law of Total Expectation: If X_1 and X_2 are random variables and the expectation of X_1 is defined, then $\mathrm{E}(X_1) = \mathrm{E}(\mathrm{E}(X_1|X_2))$.

where $a_3^t x_3(0) \to 0$ and $x_3^\star(t) < \bar{u}_c \max_\tau(|\varphi_0(\tau)|)\sum_{i=1}^\infty a_3^{i-1} = \bar{u}_c \max_\tau(|\varphi_0(\tau)|)/(1-a_3) < \infty$. Finally, $\left|\sum_{i=t+1}^\infty a_3^{i-1}\varphi_0(t-i)\bar{u}_c(t-i)\right| < \bar{u}_c \max_\tau(|\varphi_0(\tau)|)\sum_{i=t+1}^\infty a_3^{i-1} \to 0$ as $t \to \infty$. Hence, $x_3(t) \to x_3^\star(t)$, which is a bounded T-periodic function. In other words, $x_3(t)$ has a globally asymptotically stable limit cycle defined by $x_3^\star$.

Next, since $\mathrm{E}\big(w_u(t)\big) = 0$ for all t, it follows from (19) that $\mathrm{E}\big(x_4(t)\big) = a_4^t x_4(0)$, which converges to zero. Hence $\mathrm{E}\big(x_4(t)\big)$ has a globally asymptotically stable fixed point solution $x_4^\star = 0$.

The Law of Total Expectation yields $\mathrm{E}(\tilde{x}(t+1)) = \mathrm{E}\big(\mathrm{E}(\tilde{x}(t+1)|\tilde{x}(t))\big)$. Replace $\tilde{x}(t+1)$ with the right hand side of (24) and recognize that $\mathrm{E}\big(A(t,x_4,w_m)\tilde{x}(t) + B\epsilon_1(t,x_4,w_m)|\tilde{x}(t)\big) = \mathrm{E}(A(t,x_4,w_m))\tilde{x}(t) + B\mathrm{E}(\epsilon_1(x_4,w_m))$. Since $\mathrm{E}(\varphi_1) = 0$, it is easily shown that $\mathrm{E}(A(t,x_4,w_m)) = A(t,0,0)$ and $\mathrm{E}(\epsilon_1(x_4,w_m)) = 0$. Hence, $\mathrm{E}\big(\tilde{x}(t+1)\big) = A(t,0,0)\mathrm{E}(\tilde{x}(t))$, which has the origin as a fixed point solution. However, the solution is not obviously unique since $A(t,0,0)$ potentially does not have full rank.

Since $A(t,0,0)$ is T-periodic, we invoke Floquet theory in discrete time, which states that the system is asymptotically stable if and only if the eigenvalues of $\bar{A}(t,0,0) := A(t+T-1,0,0)A(t+T-2,0,0)\cdots A(t,0,0)$ are strictly inside the unit circle. On the other hand, due to Floquet, $\bar{A}(t,0,0)$ and $\bar{A}(1,0,0)$ have the same eigenvalues and are mapped to each other via a pure rotation.

Now, invert the mapping in (20) and (21) to obtain

$$\begin{aligned} x_1(t) &= \frac{1}{c_{\text{I}}}\tilde{x}_1(t) + x_3(t) \\ x_2(t) &= \tilde{x}_2(t) + \frac{\bar{u}_c - a_0}{a_1}. \end{aligned}$$

Since $[0,0]^T$ is a fixed point solution of $\mathrm{E}(\tilde{x}(t))$, it follows that $x_3^\star$ is a limit cycle of $\mathrm{E}(x_1(t))$ and $(\bar{u}_c - a_0)/a_1$ a fixed point solution of $\mathrm{E}(x_2(t))$.

The limit cycle of $[\mathrm{E}(u(t)), \mathrm{E}(e(t))]^T$ is derived with help of the Law of Total Expectation, which implies $\mathrm{E}(z(t)) = \mathrm{E}\big(\mathrm{E}(z(t)|\tilde{x}(t))\big)$. Replace $z(t)$ with the right hand side of (25) and recognize that $\mathrm{E}\big(C(x_4)\tilde{x}(t) + D\epsilon_2(x_4)|\tilde{x}(t)\big) = \mathrm{E}(C(x_4))\tilde{x}(t) + D\mathrm{E}(\epsilon_2(x_4))$. Furthermore, note that $\mathrm{E}(C(x_4)) = C(0)$ and $\mathrm{E}(\epsilon_2(x_4)) = \epsilon_2(0)$. Hence,

$$\mathrm{E}(z(t)) = C(0)\mathrm{E}(\tilde{x}(t)) + D\epsilon_2(0). \qquad (32)$$

However, $D = [1,0]^T$ and $\epsilon_2(0) = (\bar{u}_c - a_0)/a_1$. Since the origin is a fixed point of $\mathrm{E}(\tilde{x}(t))$, it follows that $u^\star := (\bar{u}_c - a_0)/a_1$ and zero are fixed point solutions of $\mathrm{E}(u(t))$ and $\mathrm{E}(e(t))$, respectively. The above limit cycle for $x_1(t)$ and fixed point solutions for $x_2(t)$, $u(t)$, and $e(t)$ are globally asymptotically stable if and only if the eigenvalues of $\bar{A}(1,0,0)$ are strictly inside the unit circle. This completes the proof. ∎

It is in general not possible to derive a simple formula for the eigenvalues of $\bar{A}(1,0,0)$ as a function of parameters a_3, a_1, c_{I}, and $h(t)$. However, it is straightforward for any set of parameters, and with help of a computer, to evaluate the product $A(T,0,0)A(T-1,0,0)\cdots A(1,0,0)$, and then compute the eigenvalues of this matrix. It is an insightful exercise to do this for a few examples. Note for $A(t,0,0)$ that parameters a_1 and c_{I} always appear together as a product. This reduces the degrees of freedom as we analyze the dynamics. We only need to consider combinations of the triplet a_3, a_1c_{I}, and $h(t)$ in order to gain a complete understanding of the stability of any configuration.

Example 6.1: Assume $h(t) = h_1\sin(2\pi t\Delta/24)$, where $\Delta = 5/60$ hours, and consider six different values of the seasonality amplitude ($h_1 = 0, 0.5, 0.7, 0.9, 0.975$, and 1). For each value of h_1, we compute the eigenvalues of $\bar{A}(1,0,0)$ for $40,000$ combinations of a_3 and a_1c_{I}, and determine the stability of each combination. The result is shown using heat maps in Figure 4. Blue color means unstable, and yellow means stable. For values of h_1 different from 0, the heat maps have three regions. In the blue connected region, all combinations of a_3 and a_1c_{I} lead to instability while in the fully yellow region all combinations correspond to stability. Finally, in the third region (blue and yellow mixed) some configurations are stable and some are unstable. In particular, it corresponds to systems where tiny alterations of a_3 or a_1c_{I} may turn a stable system unstable, or an unstable system stable. The stable configurations in this region are structurally non-robust. The only region in which we want to operate is the fully yellow one, and then on a sufficiently large distance from any unstable configuration to ensure robustness.

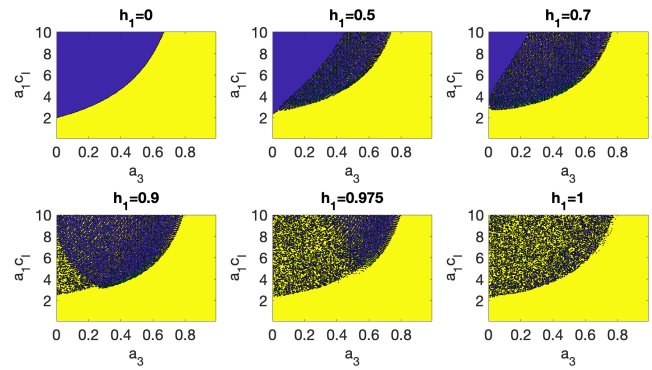

Fig. 4. Stability (yellow) and instability (blue) regions when $h(t) = h_1\sin(2\pi t\Delta/24)$ and $\Delta = 5/60$ hours.

Theorem 6.2: If $a_3 = 0$, then the limit cycle and fixed point solution in Theorem 6.1 are globally asymptotically stable if and only if $\left|\prod_{i=1}^T\left(1 - a_1c_{\text{I}}\varphi_0(t)\right)\right| < 1$.

Proof: With the same reasoning as in the proof of Theorem 6.1, we establish that $\mathrm{E}\big(\tilde{x}(t+1)\big) = A(t,0,0)\mathrm{E}(\tilde{x}(t))$. According to Floquet, $\mathrm{E}(\tilde{x}(t)) = 0$ is globally asymptotically stable if and only if the eigenvalues of $\bar{A}(1,0,0)$ are strictly inside the unit circle. Since $a_3 = 0$, we know from (26) that

$$A(t,0,0) = \begin{bmatrix} -a_1c_{\text{I}}\varphi_0(t), & a_1c_{\text{I}}\varphi_0(t) \\ -1, & 1 \end{bmatrix}.$$

It is easy to confirm by matrix multiplication and factorization that for any τ, it is always true that $A(\tau+1,0,0)A(\tau,0,0) = \big(1 - a_1c_{\text{I}}\varphi_0(\tau)\big)A(\tau+1,0,0)$. Repeating this for $\tau = 1,2,\ldots,T-1$ yields

$$\bar{A}(1,0,0) := \left(\prod_{i=1}^{T-1}\Big(1 - a_1c_{\text{I}}\varphi_0(i)\Big)\right)A(T,0,0).$$

The eigenvalues of $\bar{A}(1,0,0)$ are obtained as the solutions to $\det(\lambda I - \bar{A}(1,0,0)) = 0$. It is straight-forward to obtain

$$\det(\lambda I - \bar{A}(1,0,0)) = \lambda\left(\lambda - \prod_{i=1}^{T}\Big(1 - a_1c_{\text{I}}\varphi_0(i)\Big)\right).$$

The solutions are $\lambda_1 = 0$, which is always inside the unit circle, and $\lambda_2 = \prod_{i=1}^{T}\Big(1 - a_1 c_\text{I}\varphi_0(i)\Big)$, which is inside the unit circle if and only if $\Big|\prod_{i=1}^{T}\Big(1 - a_1 c_\text{I}\varphi_0(i)\Big)\Big| < 1$, which proves the condition for asymptotic stability of $\mathrm{E}(\tilde{x}(t)) = 0$. The same condition defines stability for $\mathrm{E}(x)$ and $\mathrm{E}(z)$ as outlined in the proof of Theorem 6.1. ■

VII. Simulation Results

Consider a plant where the delay-free ad spend $\acute{y}(t) = f(t, u(t), w_m(t)) = 500\big(1 + 0.8\sin(2\pi t\Delta/24)\big)\big(1 + w_m(t)\big)\tanh\big(u(t) - 4\big)$, $\Delta = 5/60$ hours, $w_m(t) \sim \text{Gaussian}(0, 0.2^2)$, and $t = 0, 1, \ldots$. The relationship between u and $\mathrm{E}(\acute{y})$ is smooth and can be linearized around each operating point of u yielding $\acute{y}(t) \approx \big(a_0 + a_1 u(t)\big)\big(1 + 0.8\sin(2\pi t\Delta/24)\big)\big(1 + w_m(t)\big)$ in accordance with the plant model introduced in Section IV. Suppose the observed ad spend has a latency time constant $T_p = 0.5$ hours, which corresponds to a discrete-time latency dynamics described by $y(t) = a_3 y(t-1) + (1 - a_3) f(t, u(t), w_m(t))$, where $a_3 = e^{-\Delta/T_p}$. Let the total flight time $T = 720$ hours (one month), and the set-point signal be

$$\bar{u}_c = \begin{cases} 450, & \text{if } t\Delta \leq 237 \text{ or } t\Delta > 482 \\ 100, & \text{if } 237 < t\Delta \leq 482 \end{cases}$$

Configure the controller (Algorithm 1) using $a_4 = e^{-\Delta/T_{PE}}$, where $T_{PE} = 0.05$ and the plant identification system (Section V) with an exponential memory loss $\lambda_{\text{RLS}} = e^{-\Delta/T_{RLS}}$ with $T_{RLS} = 5$. Furthermore, assume the simulated persistent excitation noise $w_u(t) \sim \text{Gaussian}(0, 0.04^2)$. Finally, let $c_\text{I} = 0.05/\hat{a}_1$.

We implement the plant identification system as a bank of seven parallel RLS estimators (Section V), for which the parameter estimate covariance matrix P is reset in a staggered fashion, one every 15 hours. That is, each estimator has its covariance matrix reset once every 105 hours. The bank of estimators generates estimates $\hat{\theta}^1, \hat{\theta}^2, \ldots, \hat{\theta}^7$, and each such estimate contains a plant gain estimate $\hat{a}_1^{(i)}$. We select $\hat{a}_1 = \underset{i\in\{1,\ldots,7\}}{\text{median}}\,\hat{a}_1^{(i)}$ for adjusting c_I to ensure desired performance and robustness trade-off.

A representative result of the closed-loop behavior is shown in Figure 5. The figure displays the control signal u, observed ad spend y, delivery error e, and plant gain estimate $\hat{a}_1$. Note how u rapidly settles down on an approximately constant value while the observed ad spend with good precision tracks the adjusted set-point signal u_c and keeping the error small. The approximately constant u and approximately zero e corresponds to near-optimal ROI and smooth pacing for the advertiser. The plant gain estimate $\hat{a}_1$ also converges towards the true value of a_1, yet slowly. Overall, the control behavior is good and according to expectation.

VIII. Conclusions

We have proposed a controller for online advertising campaigns and derived stability conditions. The time-varying plant yields a complex dynamics, but the simplicity of the controller makes tuning easy. The controller has been commissioned extensively in simulated scenarios with convincing results. Experimental commissioning is in progress with great preliminary result. In a practical implementation, the controller is supplemented with wind-up protection.

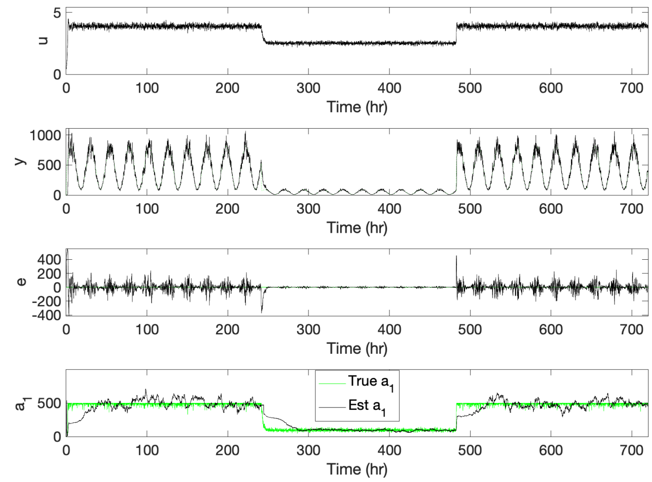

Fig. 5. Simulated closed-loop results showing control signal u, observed ad spend y, tracking error e and estimated (true) plant gain $\hat{a}_1$ (a_i). The set-point signal $\hat{u}_c$ changed from 100 to 450 at hour 250.

Future work includes robustness analysis to determine the impact on dynamics and stability from imperfect estimates of a_3 and $h(t)$, and research for enhanced online and offline algorithms to estimate plant gain, seasonality parameters, feedback latency, and plant noise level.

References

[1] 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*, ACC '13, pages 6008 – 6013, 2013.

[2] Niklas Karlsson. Control problems in online advertising and benefits of randomized bidding strategies. *European Journal of Control*, 30:31—-49, July 2016.

[3] Niklas Karlsson. *Systems and Methods for Controlling Bidding for Online Advertising Campaigns, United States Patent Application 20100262497*. USPTO, 2009.

[4] Niklas Karlsson. Adaptive control using Heisenberg bidding. In *Proceedings of the 2014 American Control Conference, Portland, USA, June 4-6, 2014*, ACC '14, pages 1304–1309, 2014.

[5] Weinan Zhang. Optimal real-time biding for display advertising. *Ph.D. Dissertation, Department of Computer Science, University College London, London, UK*, 2016.

[6] Jiaxing Guo and Niklas Karlsson. Model reference adaptive control of advertising systems. In *Proceedings of the 2017 American Control Conference, Seattle, WA, USA*, ACC '17, pages 5482 – 5487, May 24-26 2017.

[7] Weinan Zhang, Yifei Rong, Jun Wang, Tianchi Zhu, and Xiaofan Wang. Feedback control of real-time display advertising. In *Proceedings of the 9th ACM International Conference on Web Search and Data Mining, New York, NY, USA*, ACM '18, pages 407–416, 2018.

[8] Niklas Karlsson. Control of periodic systems in online advertising. In *Proceedings of the 2018 IEEE 57th Conference on Decision and Control, Miami, FL, USA*, CDC '18, Dec 17-19 2018.

[9] Niklas Karlsson. Plant gain estimation in online advertising processes. In *Proceedings of the 2017 IEEE 56th Conference on Decision and Control, Melbourne, Australia*, CDC '17, pages 2182 – 2187, Dec 12-15 2017.

[10] Hao He and Niklas Karlsson. Identification of seasonality in Internet traffic to support control of online advertising. In *Proceedings of the American Control Conference, Philadelphia, Philadelphia, USA*, ACC '19, July 10-12, 2019.