Model Reference Adaptive Control of Advertising Systems

Proceedings of the 2017 American Control Conference, Seattle, WA, May 24 - 26, 2017

Jiaxing Guo[1] and Niklas Karlsson[2]

***Abstract*— Internet advertising is a relatively new area where feedback control has become critically important for scalable optimization. But using feedback control in this space is challenging due to nonlinear, time-varying, and uncertain plants. In this paper we propose a high-fidelity model reference adaptive controller for reference tracking of budget-constrained advertisement campaigns.**

I. INTRODUCTION

Advertising, which is a US$600 billion industry [**?**], has for online applications in recent years come to rely heavily on feedback control. Each advertiser wish to spend an advertisement budget in such a way that their specific branding and/or performance objective is optimized. Cooperation is not permitted and the advertisers compete over ad impressions (opportunities to show advertisement to Internet users). In short, each advertiser wish to serve ads to those Internet users that generate the highest return on investment.

The allocation of ad impressions is handled in impression exchanges where any advertiser may submit bids for any opportunity to show an ad, but where only the highest bidder is awarded the impression. The optimization problem turns into a problem of estimating the return on investment of each impression opportunity. Given the extremely large number of Internet users browsing Internet every day and the large number of advertisers, it is an extraordinarily high-dimensional problem. In addition to the scale, also time-varying and stochastic traffic patterns and user-behavior add complexity to the optimization problem.

Feedback control has played a critical role in solving the above type of optimization problems for more than ten years. See e.g. [**?**] for an early but high-level overview. The first comprehensive description of control problems in online advertising, presented from a feedback control perspective, was published in [**?**]. Because of time-variabilities and nonlinearities it is not surprising that adaptive control is considered, and [**?**] presents an adaptive controller involving bid randomization. Adaptive estimation in online advertisement is also promising, which was demonstrated in [**?**], and in the forthcoming paper we propose a *model reference adaptive control* (MRAC) scheme (e.g. [**?**], [**?**], [**?**]) for reference tracking of budget-constrained advertisement campaigns. The effectiveness of the proposed adaptive controller is validated by a high-fidelity advertisement model developed in [**?**] and [**?**]. The experiment study on AOL advertising optimization platform–$\mathrm{AdLearn}^{\mathrm{TM}}$ is in progress.

The paper is organized as follows. We formally define the problem in Section II. In Section III we normalize and linearize the plant model. This allows us in Section IV to design an adaptive controller. We demonstrate the control performance on simulated ad campaigns in Section V. Finally, in Section VI we wrap up the paper with some concluding remarks and ideas of future work.

II. PROBLEM STATEMENT

For ad campaigns competing with other bidders on the Internet, we implement an optimal bidding strategy [**?**] manipulated by a campaign-level signal u_p to optimize return on investment for the ad campaign with budget constraint. To ensure ad campaign cost tracks a given budget, a feedback control design is needed for the campaign-level signal u_p. Based on the advertising system analysis in [**?**], we may model the relationship between the signal u_p and the ad campaign cost c as a discrete-time model:

$$c(tT) = f(u_p((t-1)T))h_{seas}(tT)e^{\varepsilon(tT)}, \tag{1}$$

for $t = 0, 1, 2, \ldots$ and T being a fixed sampling period, where $f(u_p(tT))$ is nonlinear and unknown, $h_{seas}(tT) > 0$ is periodic, and $\varepsilon(tT)$ denotes the noise. Note, the cost $c(tT)$ represents the spend of the advertiser within the time interval $[(t-1)T, tT)$ and the signal $u_p((t-1)T)$ remains constant in the time interval $[(t-1)T, tT)$. Due to system latency in reporting the cost, we are unable to obtain the cost $c(tT)$ at time tT, while we may only obtain a cost measurement defined as uploaded cost $c^{uploaded}(tT)$ at time tT:

$$c^{uploaded}(tT) = \sum_{m=0}^{d} \alpha_m c((t-m)T), \tag{2}$$

where d represents maximum delay in reporting cost, and $\alpha_m \geq 0$, $m = 0, \ldots, d$, are unknown coefficients satisfying

$$\sum_{m=0}^{d} \alpha_m = 1. \tag{3}$$

In this paper, we develop an adaptive control scheme for the campaign-level signal u_p to handle the nonlinearities and uncertainties of the campaign cost plant to make the ad campaign cost track a desired budget reference.

III. LINEARIZED SYSTEM MODEL

Before proceeding the control design for u_p, in this section we investigate the characteristics of the uploaded cost plant (2) to establish a model for control.

[1]J. Guo is Senior Research Scientist of R&D at AOL Platforms, 395 Page Mill Road, Palo Alto, CA 94306, USA jiaxing.guo@teamaol.com

[2]N. Karlsson is Vice President of R&D at AOL Platforms, 395 Page Mill Road, Palo Alto, CA 94306, USA niklas.karlsson@teamaol.com

Normalized uploaded cost model. An off-line study using experimental advertising data has been conducted in [**?**] to generate a periodic function $\hat{h}_{seas}(tT) > 0$ to estimate the function $h_{seas}(tT)$ in the cost model (1). Dividing the estimated periodic function $\hat{h}_{seas}(tT)$ calibrated in [**?**] on both sides of the uploaded cost model (2) and defining a normalized uploaded cost $\bar{c}^{uploaded}(tT)$:

$$\bar{c}^{uploaded}(tT) = \frac{c^{uploaded}(tT)}{\hat{h}_{seas}(tT)}, \quad (4)$$

we have the normalized uploaded cost model as

$$\bar{c}^{uploaded}(tT) = \sum_{m=0}^{d} \bar{\alpha}_m(tT)\frac{c((t-m)T)}{h_{seas}((t-m)T)}, \quad (5)$$

where

$$\bar{\alpha}_m(tT) = \frac{\alpha_m h_{seas}((t-m)T)}{\hat{h}_{seas}(tT)}, \quad (6)$$

for $m = 0, \ldots, d$. In view of the cost model (1), we can further express the normalized uploaded cost model (5) as

$$\bar{c}^{uploaded}(tT) = \sum_{m=0}^{d} \bar{\alpha}_m(tT) f(u_p((t-m-1)T)). \quad (7)$$

Denoting $\bar{c}^{uploaded}(tT)$ as $y(tT)$:

$$y(tT) \triangleq \bar{c}^{uploaded}(tT), \quad (8)$$

we can obtain a state space representation for the normalized uploaded cost model (7):

$$\begin{cases} x_1((t+1)T) = x_2(tT), \\ x_2((t+1)T) = x_3(tT), \\ \vdots \\ x_{d+1}((t+1)T) = f(u_p(tT))e^{\varepsilon((t+1)T)}, \end{cases}$$
$$y(tT) = \bar{\alpha}_d(tT)x_1(tT) + \bar{\alpha}_{d-1}(tT)x_2(tT) + \cdots + \cdots + \bar{\alpha}_0(tT)x_{d+1}(tT). \quad (9)$$

Linear time-invariant normalized uploaded cost model. For AOL advertising optimization platform $\mathrm{AdLearn}^{\mathrm{TM}}$, the period of $h_{seas}(tT)$ is 24-hours, the sampling period is $T = 0.25$ hour, and the maximum delay in reporting cost is $d = 1$. So, if the estimate $\hat{h}_{seas}(tT)$ is accurate, the parameter $\bar{\alpha}_m(tT)$ can be approximated as $\bar{\alpha}_m(tT) \approx \alpha_m$, which is a constant, for $m = 0, 1$. Without loss of generality, in this control design, we assume the parameter $\bar{\alpha}_m(tT)$ in (9) is a constant; i.e., $\bar{\alpha}_m(tT) = \bar{\alpha}_m$, with $\bar{\alpha}_m$ being some unknown constant, for $m = 0, \ldots, d$. Moreover, around an operating point u_{p0}, the nonlinear function $f(u_p)$ is approximated as

$$f(u_p) \approx b_p u_p + f_{p0}, \quad (10)$$

where

$$b_p = \left.\frac{\partial f(u_p)}{\partial u_p}\right|_{u_{p0}}, \quad (11)$$

$$f_{p0} = f(u_{p0}) - \left.\frac{\partial f(u_p)}{\partial u_p}\right|_{u_{p0}} u_{p0}. \quad (12)$$

Note, the parameter b_p and the dynamics offset f_{p0} in (10) are unknown due to the uncertainties of $f(u_p(t))$. From (10), the assumption that $\bar{\alpha}_m$ is constant for $m = 0, \ldots, d$, and ignoring the noise term, we use the following linear time-invariant model to approximate the normalized uploaded cost model (9) around the operating point u_{p0}:

$$\begin{aligned} x(t+1) &= Ax(t) + Bu_p(t) + f_0, \\ y(t) &= Cx(t), \end{aligned} \quad (13)$$

where $x(t) = [x_1(t), x_2(t), \ldots, x_{d+1}(t)]^T$ is the state vector signal, and the system parameters A, B, C, and the dynamics offset f_0, are

$$A = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \end{bmatrix} \in R^{(d+1)\times(d+1)},$$

$$B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b_p \end{bmatrix} \in R^{d+1}, f_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ f_{p0} \end{bmatrix} \in R^{d+1},$$

$$C = [\bar{\alpha}_d, \bar{\alpha}_{d-1}, \ldots, \bar{\alpha}_1, \bar{\alpha}_0] \in R^{1\times(d+1)}. \quad (14)$$

Then, the input-output representation of the linear model (13) can be expressed as

$$y(t) = G_0(z)[u_p](t) + G_d(z)[u_s](t), \quad (15)$$

where $u_s(t)$ is a unit step signal; i.e., $u_s(t) = 1$, for $t = 0, 1, \ldots$, and the transfer functions are

$$G_0(z) = \frac{\bar{\alpha}_0 b_p z^d + \bar{\alpha}_1 b_p z^{d-1} + \cdots + \bar{\alpha}_d b_p}{z^{d+1}}, \quad (16)$$

$$G_d(z) = \frac{\bar{\alpha}_0 f_{p0} z^d + \bar{\alpha}_1 f_{p0} z^{d-1} + \cdots + \bar{\alpha}_d f_{p0}}{z^{d+1}}. \quad (17)$$

Control objective. Since the state signal $x(t)$ is unavailable for measurement, and only the output signal $y(t)$; i.e., the normalized uploaded cost $\bar{c}^{uploaded}(t)$ defined in (4), is measurable, we employ an output feedback MRAC design to compensate for the uncertainties of B and C and the effect of the unknown dynamics offset f_0 on the output signal $y(t)$, and make all closed-loop system signals bounded and the output signal $y(t)$ track a reference signal

$$y_m(t) = W_m(z)[r](t), \quad (18)$$

where $r(t)$ is a bounded reference input and

$$W_m(z) = \frac{1}{z}, \quad (19)$$

for the linear time-invariant model (13). Then we implement the adaptive control scheme to the original ad campaign cost plant (2) to evaluate the effectiveness of the linearization-based control design.

Assumptions. To proceed the MRAC design, we assume (A1) all zeros of $G_0(z)$ are stable; (A2) an upper bound

$\bar{d}$ of the maximum delay d is known; (A3) high-frequency gain $\bar{\alpha}_0 b_p \neq 0$ and the sign of $\bar{\alpha}_0 b_p$ is known. The assumption (A1) is needed for stable plant-model matching, the assumptions (A2) and (A3) are used for designing controller structure and adaptive laws. Note, we verify that the assumptions can hold true for the ad campaign cost plant in the simulation study.

IV. ADAPTIVE CONTROL DESIGN

In this section, we develop an adaptive control algorithm for the linear time-invariant model (13) and discuss possible robustness modifications when implementing to the nonlinear uploaded cost model (2).

A. *Model Reference Adaptive Control Scheme*

The feedback control signal $u_p(t)$ is chosen as

$$u_p(t) = \theta_1^T(t)\omega_1(t) + \theta_2(t)r(t) + \theta_3(t), \quad (20)$$

where $\omega_1(t) = \frac{a(z)}{\Lambda(z)}[u_p](t)$, with $\Lambda(z) = z^{\bar{d}}$ and $a(z) = [z^{\bar{d}-1}, \ldots, z, 1]^T \in R^{\bar{d}}$, and $\theta_1(t)$, $\theta_2(t)$, and $\theta_3(t)$ are estimates of nominal parameters θ_1^*, θ_2^*, and θ_3^*. In particular, $\theta_3(t)$ is for compensation of the effect of the unknown dynamics offset f_0 of (13). The existence of the nominal parameters θ_1^*, θ_2^*, and θ_3^* are ensured by the following plant-model matching property.

Proposition 4.1: *There exist θ_1^*, θ_2^*, and θ_3^*, such that, when $\theta_1(t) = \theta_1^*$, $\theta_2(t) = \theta_2^*$, and $\theta_3(t) = \theta_3^*$, the control signal $u_p(t)$ in (20) ensures signal boundedness and output tracking $\lim_{t\to\infty}(y(t) - y_m(t)) = 0$ for the system (13).*

Proof: It can be shown that there exist $\theta_1^* = [\theta_{11}^*, \theta_{12}^*, \ldots, \theta_{1\bar{d}}^*]^T$ and $\theta_2^* = \frac{1}{\bar{\alpha}_0 b_p}$, where

$$\theta_{11}^* = -\frac{\bar{\alpha}_1}{\bar{\alpha}_0}, \;\; \theta_{12}^* = -\frac{\bar{\alpha}_2}{\bar{\alpha}_0}, \;\; \ldots, \;\; \theta_{1d}^* = -\frac{\bar{\alpha}_d}{\bar{\alpha}_0}, \quad (21)$$

$$\theta_{1d+1}^* = 0, \;\; \ldots, \;\; \theta_{1\bar{d}}^* = 0, \quad (22)$$

such that the matching equation

$$1 - \theta_1^{*T}\frac{a(z)}{\Lambda(z)} = \theta_2^* W_m^{-1}(z) G_0(z) \quad (23)$$

holds true. Operating both sides of (23) on $u_p(t)$, we have

$$u_p(t) - \theta_1^*\omega_1(t) = \theta_2^* W_m^{-1}(z)\left[y - G_d(z)[u_s]\right](t). \quad (24)$$

Substituting the nominal control signal $u_p(t) = \theta_1^*\omega_1(t) + \theta_2^* r(t) + \theta_3^*$ in (24) and from (23), we obtain the tracking error $e(t) = y(t) - y_m(t)$ as

$$e(t) = \bar{\alpha}_0 b_p W_m(z)\left[\left(1 - \theta_1^{*T}\frac{a(z)}{\Lambda(z)}\right) G_0^{-1} G_d[u_s] + \theta_3^*\right](t). \quad (25)$$

Based on final value theorem, (21) and (22), we have

$$\lim_{t\to\infty}\left(1 - \theta_1^{*T}\frac{a(z)}{\Lambda(z)}\right) G_0^{-1} G_d(z)[u_s](t) = \frac{\sum_{m=0}^{d}\bar{\alpha}_m f_{p0}}{\bar{\alpha}_0 b_p}.$$

In view of (25), and choosing $\theta_3^* = -\frac{\sum_{m=0}^{d}\bar{\alpha}_m f_{p0}}{\bar{\alpha}_0 b_p}$, it follows that $\lim_{t\to\infty} e(t) = 0$ exponentially. ∇

Since the nominal controller parameters θ_1^*, θ_2^*, and θ_3^* are unknown, we need to employ the adaptive controller (20), where the controller parameters $\theta_1(t)$, $\theta_2(t)$, and $\theta_3(t)$ are updated by adaptive laws developed as follows.

Tracking error equation. Operating both sides of the matching equation (23) on $u_p(t)$ and substituting the adaptive controller (20), we obtain the tracking error as

$$e(t) = \bar{\alpha}_0 b_p W_m(z)[\tilde{\theta}^T\omega](t) + f_p(t), \quad (26)$$

where $\tilde{\theta}(t) = \theta(t) - \theta^*$, $\theta(t) = [\theta_1^T(t), \theta_2(t), \theta_3(t)]^T$, $\theta^* = [\theta_1^{*T}, \theta_2^*, \theta_3^*]^T$, $\omega = [\omega_1^T, r, u_s]^T$, and

$$f_p(t) \triangleq \bar{\alpha}_0 b_p W_m(z)\left[\left(1 - \theta_1^{T*}\frac{a(z)}{\Lambda(z)}\right) G_0^{-1} G_d[u_s] + \theta_3^*\right](t).$$

Note, with the nominal parameters θ_1^*, θ_2^*, and θ_3^*, $f_p(t)$ decays to 0 exponentially as shown in Proposition 4.1.

Estimation error. We introduce an estimation error $\epsilon(t)$:

$$\epsilon(t) = e(t) + \rho(t)\xi(t), \quad (27)$$

where $\rho(t)$ is the estimate of $\rho^* \triangleq \bar{\alpha}_0 b_p$ and

$$\xi(t) = \theta^T(t)\zeta(t) - W_m(z)[\theta^T\omega](t), \quad (28)$$

$$\zeta(t) = W_m(z)[\omega](t). \quad (29)$$

Substituting the tracking error $e(t)$ (26) in (27), and ignoring the decaying term $f_p(t)$, we obtain

$$\epsilon(t) = \rho^*\tilde{\theta}^T(t)\zeta(t) + \tilde{\rho}(t)\xi(t), \quad (30)$$

where $\tilde{\rho}(t) = \rho(t) - \rho^*$.

Adaptive laws. Based on the estimation error model (30), we choose adaptive laws for the parameters $\theta(t)$ and $\rho(t)$ as

$$\theta(t+1) = \theta(t) - \frac{\text{sign}(\bar{\alpha}_0 b_p)\Gamma\zeta(t)\epsilon(t)}{m^2(t)}, \quad (31)$$

$$\rho(t+1) = \rho(t) - \frac{\gamma\xi(t)\epsilon(t)}{m^2(t)}, \quad (32)$$

where the error signal $\epsilon(t)$ is computed from (27), the adaptation gains satisfy $0 < \Gamma = \Gamma^T < \frac{2}{|\bar{\alpha}_0 b_p|} I_{\bar{d}+2}$ with $I_{\bar{d}+2}$ being a $(\bar{d}+2) \times (\bar{d}+2)$ identity matrix and $0 < \gamma < 2$, and the normalization signal $m(t)$ is

$$m(t) = \sqrt{1 + \zeta^T(t)\zeta(t) + \xi^2(t)}. \quad (33)$$

Stability property. Introduce a positive definite function

$$V(\tilde{\theta}, \tilde{\rho}) = |\rho^*|\tilde{\theta}^T\Gamma^{-1}\tilde{\theta} + \gamma^{-1}\tilde{\rho}^2. \quad (34)$$

The time increment of $V(\tilde{\theta}, \tilde{\rho})$, along (31) and (32), is

$$V(\tilde{\theta}(t+1), \tilde{\rho}(t+1)) - V(\tilde{\theta}(t), \tilde{\rho}(t)) = -\left(2 - \frac{\bar{\alpha}_0 b_p \zeta^T(t)\Gamma\zeta(t) + \gamma\xi^2(t)}{m^2(t)}\right)\frac{\epsilon^2(t)}{m^2(t)}. \quad (35)$$

From the conditions that $0 < \Gamma = \Gamma^T < \frac{2}{|\bar{\alpha}_0 b_p|} I_{\bar{d}+2}$ and $0 < \gamma < 2$, we obtain

$$V(\tilde{\theta}(t+1), \tilde{\rho}(t+1)) - V(\tilde{\theta}(t), \tilde{\rho}(t)) \leq -\beta_1\frac{\epsilon^2(t)}{m^2(t)}, \quad (36)$$

for some constant $\beta_1 > 0$, which implies that $\theta(t) \in L^\infty$, $\rho(t) \in L^\infty$, $\frac{\epsilon(t)}{m(t)} \in L^2 \cap L^\infty$, $\theta(t+1) - \theta(t) \in L^2 \cap L^\infty$, and $\rho(t+1) - \rho(t) \in L^2 \cap L^\infty$. Based on the above properties, we can show that the controller (20) with the parameters updated by the adaptive laws (31) and (32) can guarantee the closed-loop system signal boundedness and asymptotic output signal tracking; i.e., $\lim_{t\to\infty}(y(t) - y_m(t)) = 0$ asymptotically, for the model (13). The proof can be carried out by using a similar way as described in [**?**].

B. Robustness Modification for Adaptive Laws

When implementing the linearization-based adaptive control scheme to the nonlinear ad campaign cost plant, modeling errors caused by linearization, parameter variation, and system noise may have impacts on the closed-loop system stability and signal tracking performance. To handle the system modeling errors, we employ a parameter projection modification for the adaptive laws (31) and (32).

Parameter projection. We choose

$$\theta(t) = \theta_{temp}(t) + f_\theta(t), \tag{37}$$

$$\rho(t) = \rho_{temp}(t) + f_\rho(t), \tag{38}$$

where $\theta_{temp}(t) \triangleq [\theta_{temp11}, \ldots, \theta_{temp1\bar{d}}, \theta_{temp2}, \theta_{temp3}]^T$ and $\rho_{temp}(t)$ are updated by (31) and (32), and $f_\theta(t) \triangleq [f_{\theta 11}(t), \ldots, f_{\theta 1\bar{d}}(t), f_{\theta 2}(t), f_{\theta 3}(t)]^T$ and $f_\rho(t)$ are the projection terms. For the projection design, we assume that the nominal controller parameter θ_j^* belongs to a known interval $\theta_j^* \in [\theta_j^a, \theta_j^b]$, for $j = 11, 12, \ldots, 1\bar{d}, 2, 3$, and the nominal parameter ρ^* belongs to a known interval $\rho^* \in [\rho^a, \rho^b]$. We choose a diagonal design matrix $\Gamma = \mathrm{diag}\{\gamma_{11}, \gamma_{12}, \ldots, \gamma_{1\bar{d}}, \gamma_2, \gamma_3\}$ for (31), and select the initial estimate of $\theta_j(t)$ as $\theta_j(0) \in [\theta_j^a, \theta_j^b]$, for $j = 11, 12, \ldots, 1\bar{d}, 2, 3$, and the initial estimate of $\rho(t)$ as $\rho(0) \in [\rho^a, \rho^b]$. We set projection components $f_{\theta j}(t)$ and $f_\rho(t)$ as

$$f_{\theta j}(t) = \begin{cases} 0, & \text{if } \theta_{tempj}(t) \in [\theta_j^a, \theta_j^b], \\ \theta_j^b - \theta_{tempj}(t), & \text{if } \theta_{tempj}(t) > \theta_j^b, \\ \theta_j^a - \theta_{tempj}(t), & \text{if } \theta_{tempj}(t) < \theta_j^a, \end{cases} \tag{39}$$

for $j = 11, 12, \ldots, 1\bar{d}, 2, 3$, and

$$f_\rho(t) = \begin{cases} 0, & \text{if } \rho_{temp}(t) \in [\rho^a, \rho^b], \\ \rho^b - \rho_{temp}(t), & \text{if } \rho_{temp}(t) > \rho^b, \\ \rho^a - \rho_{temp}(t), & \text{if } \rho_{temp}(t) < \rho^a. \end{cases} \tag{40}$$

It can be shown that $\theta(t) - \theta(t-1) \in L^2$, and $\rho(t) - \rho(t-1) \in L^2$ [**?**]. Hence, the closed-loop system stability and the asymptotic output tracking can be guaranteed for the model (13). Note, if there exist certain types of modeling errors, the parameter projection scheme may ensure the system stability and the small output tracking error in the mean sense [**?**].

Since the linearization-based design may only be effective around a small neighbourhood of the operating point, we may need to limit the increment or decrement of the control signal $u_p(t)$ based on the parameter projection scheme (37)–(38). For the controller parameter $\theta_3(t)$, which is a component of $\theta(t)$ updated by (37), we choose an interval $[\theta_3^a(t), \theta_3^b(t)]$ as

$$\theta_3^a(t) = u_p(t-1) - \Delta^a(t) - \theta_1^T(t)\omega_1(t) - \theta_2(t)r(t), \tag{41}$$

$$\theta_3^b(t) = u_p(t-1) + \Delta^b(t) - \theta_1^T(t)\omega_1(t) - \theta_2(t)r(t), \tag{42}$$

where $\Delta^b(t) > 0$ and $\Delta^a(t) > 0$ are design parameters representing the maximum increment and decrement of the control signal for each step. Applying the parameter projection scheme (37) for $\theta_3(t)$ with

$$f_{\theta 3}(t) = \begin{cases} 0, & \text{if } \theta_{temp3}(t) \in [\theta_3^a(t), \theta_3^b(t)], \\ \theta_3^b(t) - \theta_{temp3}(t), & \text{if } \theta_{temp3}(t) > \theta_3^b(t), \\ \theta_3^a(t) - \theta_{temp3}(t), & \text{if } \theta_{temp3}(t) < \theta_3^a(t). \end{cases}$$

It follows that

$$u_p(t) = \theta^T(t)\omega(t) \in [u_p(t-1) - \Delta^a(t), u_p(t-1) + \Delta^b(t)], \tag{43}$$

which may enhance the effectiveness of the linearization-based design around a small neighborhood of the operating point. Note, only if the interval $[\theta_3^a(t), \theta_3^b(t)]$ defined in (41) and (42) satisfies $\theta_3^* \in [\theta_3^a(t), \theta_3^b(t)]$, the projection modification may make the closed-loop system stable. It needs further investigation to choose proper $\theta_3(0)$, $\Delta^a(t)$ and $\Delta^b(t)$ to ensure the condition holds true.

C. Gain Schedule Design Based on Parameter Estimation

Recall that there is an important design condition $0 < \Gamma = \Gamma^T < \frac{2}{|\bar{\alpha}_0 b_p|} I_{\bar{d}+2}$ for the controller parameter adaptive laws (31). So the knowledge of the high frequency gain $\bar{\alpha}_0 b_p$ of the transfer function $G_0(z)$ (16) is crucial for the control design. In this subsection, we implement an adaptive parameter estimation algorithm [**?**] to estimate the system parameters, in particular the high frequency gain $\bar{\alpha}_0 b_p$.

From (16), (17), and the assumption (A2), we can parameterize the model (15) as

$$y(t) = \theta_p^{*T}\phi(t), \tag{44}$$

where

$$\theta_p^* = [\bar{\alpha}_0 b_p, \bar{\alpha}_1 b_p, \ldots, \bar{\alpha}_d b_p, 0, \ldots, 0, \bar{\alpha}_0 f_{p0}, \bar{\alpha}_1 f_{p0}, \ldots, \bar{\alpha}_d f_{p0}, 0, \ldots, 0]^T, \tag{45}$$

$$\phi(t) = \Big[z^{-1}[u_p](t), \ldots, z^{-d}[u_p](t), \ldots, z^{-\bar{d}-1}[u_p](t), z^{-1}[u_s](t), \ldots, z^{-d}[u_s](t), \ldots, z^{-\bar{d}-1}[u_s](t)\Big]^T. \tag{46}$$

To estimate the unknown parameter vector θ_p^* in (44), we introduce a parametric error $\epsilon_p(t)$:

$$\epsilon_p(t) = \theta_p^T(t)\phi(t) - y(t) = \tilde{\theta}_p^T\phi(t), \tag{47}$$

where $\theta_p(t)$ is the estimate of θ_p^* and $\tilde{\theta}_p = \theta_p - \theta_p^*$. The adaptive law for the vector $\theta_p(t)$ is chosen as

$$\theta_p(t+1) = \theta_p(t) - \frac{\Gamma_p \phi(t)\epsilon_p(t)}{m_p^2(t)}, \tag{48}$$

where $0 < \Gamma_p = \Gamma_p^T < 2I_{2(\bar{d}+1)}$ and $m_p = \sqrt{1 + \phi^T\phi}$.

Since the design parameter Γ of the adaptive law (31) needs to be chosen as $0 < \Gamma = \Gamma^T < \frac{2}{|\bar{\alpha}_0 b_p|} I_{\bar{d}+2}$, we use the estimate $\theta_{p1}(t)$ of the plant high-frequency gain $\bar{\alpha}_0 b_p$ to adjust the design parameter Γ for the controller.

With a parameter projection modification [**?**], we have the plant high-frequency gain estimate $\theta_{p1}(t) \in [\theta_{p1}^a, \theta_{p1}^b]$. Based on the assumption (A3) that $\bar{\alpha}_0 b_p \neq 0$ and the sign of $\bar{\alpha}_0 b_p$ is known, without loss of generality, we assume the lower bound $\theta_{p1}^a > 0$, so that $\theta_{p1}(t) > 0$. Then we choose a diagnosis Γ for the adaptive law (31) and update it as $\Gamma(t) = \gamma_\theta(t) I_{\bar{d}+2}$, where

$$\gamma_\theta(t) = (1-\lambda_\theta)\gamma_\theta(t-1) + \lambda_\theta \frac{2}{\beta_{p1}\theta_{p1}(t)}, \tag{49}$$

for $t = 0, 1, \ldots$, with $\gamma_\theta(0) > 0$ being an arbitrary initial gain, $\lambda_\theta \in [0,1]$ being a design forgetting factor, and $\beta_{p1} > 1$ being a constant design parameter.

V. SIMULATION STUDY

In this section, we apply the adaptive controller with the parameter projection and gain schedule modifications to a high-fidelity ad cost model developed in [**?**] and [**?**] to assess the effectiveness of the linearization-based design.

A. *High-fidelity Ad Campaign Cost Model*

For AOL advertising optimization platform $\mathrm{AdLearn}^{\mathrm{TM}}$, the controller sampling period is $T = 0.25$ hour, and the delay in reporting the cost is 0.08 hour in the mean sense, it follows that, for the uploaded cost model (2), the maximum delay is $d = 1$ and $\alpha_0 > \alpha_1 > 0$ with $\alpha_0 + \alpha_1 = 1$, so we express the uploaded cost model as

$$c^{uploaded}(t) = \alpha_0 c(t) + \alpha_1 c(t-1), \tag{50}$$

where the cost $c(t)$ is modeled as (1):

$$c(t) = f(u_p(t-1)) h_{seas}(t) e^{\varepsilon(t)}. \tag{51}$$

Cost modeling in ad exchanges. The advertiser is charged based on impressions awarded to the campaign. The campaign is usually separated into multiple segments to target desired users in ad exchanges. The impression allocation for segment i is governed by a sealed 2nd price auction [**?**]. Based on the analysis in [**?**], denoting b_i as the bid price for segment i, $b_i^\star$ as the highest competing bid price, and $n_{I,i}^{tot}$ as the available number of impressions, we may model the cost $f(u_p)$ in (51) as

$$f(u_p) = \sum_{i=1}^{n} b_i^\star \mathbb{I}_{\{u_p \hat{p}_i \geq b_i^\star\}} n_{I,i}^{tot}, \tag{52}$$

where $\hat{p}_i$ is an event rate estimate [**?**], for $i = 1, 2, \ldots, n$. In the simulation study, $b_i^\star$, $\hat{p}_i$, and $n_{I,i}^{tot}$, for $i = 1, 2, \ldots, n$, are generated based on the scheme developed in [**?**].

Time-of-day pattern estimation. Based on the study conducted in [**?**], the periodic function $h_{seas}(t)$ in the model (51) possesses a 24-hour period and can be estimated by the following periodic function $\hat{h}_{seas}(t)$:

$$\hat{h}_{seas}(t) = 1 + c_1 \sin\left(\frac{2\pi T}{24} t + \phi_1\right) + c_2 \sin\left(\frac{4\pi T}{24} t + \phi_2\right), \tag{53}$$

with the sampling period $T = 0.25$ hour and the parameters c_1, ϕ_1, c_2, and ϕ_2 calibrated by the off-line empirical study in [**?**]. The noise in the model (51) may satisfy a normal distribution

$$\epsilon(t) \overset{iid}{\sim} N(0, \sigma^2), \tag{54}$$

with σ calibrated by the empirical study in [**?**].

Normalized uploaded cost model. Hence, the normalized uploaded cost model (9) used for control design is

$$\begin{cases} x_1(t+1) = x_2(t), \\ x_2(t+1) = f(u_p(t)) e^{\varepsilon(t+1)}, \end{cases}$$
$$y(t) = \bar{\alpha}_1(t) x_1(t) + \bar{\alpha}_0(t) x_2(tT), \tag{55}$$

with the output signal $y(t)$ defined in (8):

$$y(t) = \frac{c^{uploaded}(t)}{\hat{h}_{seas}(t)}, \tag{56}$$

where $f(u_p)$ is modeled as (52), the estimated periodic function $\hat{h}_{seas}(t)$ is chosen as (53), and the noise $\epsilon(t)$ satisfies the normal distribution in (54).

Control design conditions. Before applying the linearization-based adaptive control design, we need to verify that the design conditions (A1)–(A3) can hold true. Assuming that the estimate $\hat{h}_{seas}(t)$ in (53) is accurate; i.e., $\hat{h}_{seas}(t) \approx h_{seas}(t)$, and from the definition of $\bar{\alpha}_m(t)$, $m = 0, 1$ in (6), it follows that $\bar{\alpha}_0(t) \approx \alpha_0$ and $\bar{\alpha}_1(t) \approx \alpha_1$ for the normalized uploaded cost model (55). For the assumption (A1), the zero of $G_0(z)$ (16) satisfies that $|z_0| = \frac{\bar{\alpha}_1}{\bar{\alpha}_0} \approx \frac{\alpha_1}{\alpha_0} < 1$ based on the property that $\alpha_0 > \alpha_1 > 0$ with $\alpha_0 + \alpha_1 = 1$, which indicates that the assumption (A1) holds true. For the assumption (A2), we choose the upper bound of d as $\bar{d} = 1$. Based on the system analysis in [**?**], the function $f(u_p)$ is increasing with respect to u_p for certain operating region, since $b_i^\star$, $n_{I,i}^{tot}$, $\hat{p}_i \geq 0$ and $\mathbb{I}_{\{u_p \hat{p}_i \geq b_i^\star\}}$ is a step function with positive step. Hence, we have that $b_p > 0$ in view of (11), which leads to $\mathrm{sign}(\bar{\alpha}_0 b_p) = 1$ for the assumption (A3).

B. *Adaptive Controller*

Since the upper bound of the maximum delay is $\bar{d} = 1$, in view of (20), the adaptive controller $u_p(t)$ is chosen as

$$u_p(t) = \theta_1(t)\omega_1(t) + \theta_2(t) r(t) + \theta_3(t), \tag{57}$$

where $\omega_1(t) = \frac{1}{z}[u_p](t)$ and $\theta_1(t)$, $\theta_2(t)$, and $\theta_3(t)$ are updated by the adaptive laws (31) and (32) with the parameter projection and gain schedule modifications. For the parameter projection modifications (37) and (38), we choose the bounds for the parameters based on estimates of the nominal controller parameters. In particular, for $\theta_3^a(t)$ in (41) and $\theta_3^b(t)$ in (42), we may choose $\Delta^a(t) = \gamma^a |u_p(t-1)|$ and $\Delta^b(t) = \gamma^b |u_p(t-1)|$ with some $\gamma^a, \gamma^b \in (0, 1)$. For the gain schedule modification in (49), the parameter estimate $\theta_p(t)$ is updated by the adaptive law (48) with parameter projection. Note, since the upper bound of the maximum delay is $\bar{d} = 1$, the parameterized model (44) to be estimated is

$$y(t) = \theta_p^{*T} \phi(t), \tag{58}$$

with $\theta_p^{*T} = [\bar{\alpha}_0 b_p, \bar{\alpha}_1 b_p, \bar{\alpha}_0 f_{p0}, \bar{\alpha}_1 f_{p0}]^T$ and $\phi(t) = [z^{-1}[u_p](t), z^{-2}[u_p](t), z^{-1}[u_s](t), z^{-2}[u_s](t)]^T$.

C. Simulation Result

In this simulation study, we simulate a normal scenario observed in the optimization platform $\text{AdLearn}^{\text{TM}}$, where the campaign cost is simulated by the uploaded cost model (50), with $f(u_p)$ in (52) generated by the model developed in [**?**] (including a scheme modeling abrupt increase and decrease of available impressions), the noise $\epsilon(t) \overset{iid}{\sim} N(0, 0.1^2)$, the periodic function

$$h_{seas}(t){=}1{+}0.52\sin\left(\frac{2\pi T}{24}t{+}2.34\right){+}0.17\sin\left(\frac{4\pi T}{24}t{+}0.46\right),$$

and the latency coefficients $\alpha_0 = 0.83$ and $\alpha_1 = 0.17$.

For the adaptive controller (57), the estimated time-of-day pattern periodic function is chosen as

$$\hat{h}_{seas}(t){=}1{+}0.58\sin\left(\frac{2\pi T}{24}t{+}2.5\right) + 0.20\sin\left(\frac{4\pi T}{24}t{+}0.50\right),$$

which is a good approximation of Internet traffic in the US, the parameter projection bounds for the controller parameters are chosen as $\theta_1(t) \in [-5, 0]$, $\theta_2(t) \in [0, 100]$, and $\theta_3(t) \in [\theta_3^a(t), \theta_3^b(t)]$ defined in (41) and (42) with $\Delta^a(t) = 0.7|u_p(t-1)|$ and $\Delta^b(t) = 0.4|u_p(t-1)|$. Note, $y_m(t)$ in (18) is the reference for $y(t)$ defined in (56), which indicates that the reference for the uploaded cost $c^{uploaded}(t)$ is $y_m^{uploaded}(t) = y_m(t)\hat{h}_{seas}(t)$.

In the top plot of Figure 1, the blue line shows $y_m^{uploaded}(t)$, the green markers display $c^{uploaded}(t)$, and the red markers show $c(t)$. In the bottom plot of Figure 1, the blue line shows $u_p(t)$. Figure 2 shows the parameters $\theta_1(t)$, $\theta_2(t)$, and $\theta_3(t)$. From Figure 1 and Figure 2, we can see that the adaptive controller can accommodate the large increase of the available impressions happening between hour 5 and 29 and between hour 53 and 101, and can also compensate for the large decrease of the available impressions happening between hour 101 and 150.

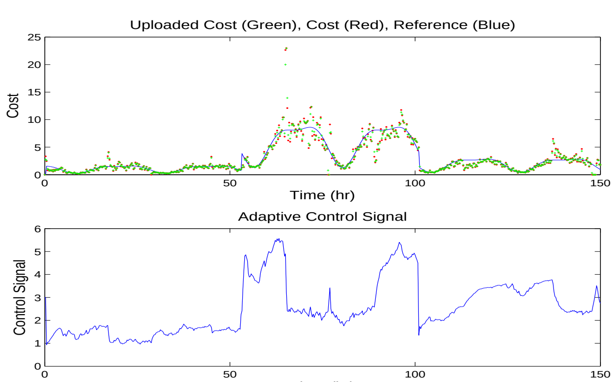

Fig. 1. Campaign cost measurements and control signal.

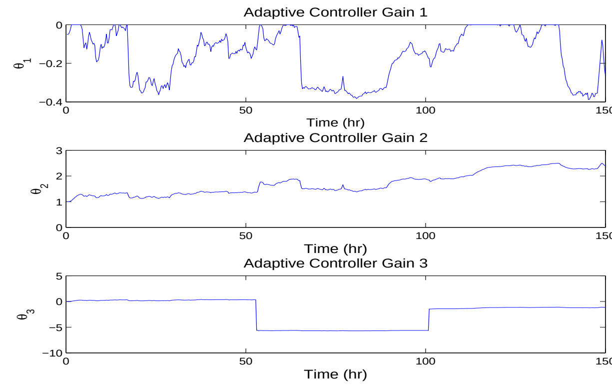

Fig. 2. Controller parameters $\theta_1(t)$, $\theta_2(t)$, and $\theta_3(t)$.

VI. CONCLUSIONS

The online ad impression allocation is handled in impression exchanges based on the 2nd price auction, so the ad campaign cost model is a highly nonlinear and time-varying plant with large uncertainties. In this paper, a linearization-based MRAC scheme has been developed to deal with the nonlinearities, uncertainties, and time-varying properties to ensure cost tracking of a desired budget. The simulation validation on a high-fidelity ad campaign cost model has verified the effectiveness of the proposed adaptive control scheme. The experiment validation on AOL advertising optimization platform $\text{AdLearn}^{\text{TM}}$ is in progress. Further research of the control of advertising systems may include designing a cost volume forecasting algorithm to estimate the plant gain and leveraging bid randomization to make the control designs more efficient and robust to real-world auction networks.