Multivariable feedback control for multi-constraint optimization in online advertising

Niklas Karlsson[1]

Abstract— **Online advertising is typically implemented via real-time bidding, and advertising campaigns are then defined as extremely high-dimensional optimization problems. To solve these problems in light of large scale and significant uncertainties, the optimization problems are modularized in a manner that makes feedback control a critical component of the solution. The control problem, however, is challenging due to plant uncertainties, nonlinearities, time-variance, and noise. Multi-constraint optimization problems are especially difficult to solve via feedback control because of the dynamic interaction across feedback loops. This paper demonstrates how one particular multi-constraint problem can be solved using a cascade feedback controller. The inner loop is managed by a linear time-periodic feedforward controller combined with a linear time-invariant feedback controller. Meanwhile, the outer loop is managed by a linear time-invariant feedforward-feedback controller. This paper is concerned with the outer loop controller and derives sufficient conditions for stability of the nonlinear closed loop system by expressing it as a Lure' system and by engaging the circle criterion. The solution is evaluated in a simulated environment based on artificial data.**

Index Terms — **Nonlinear control, Lure' system, circle criterion, stability, programmatic advertising**

I. Introduction

Online advertising is an important industry segment and at the core of the business model for companies such as Amazon, Google, and Meta. A *Demand Side Platform* (DSP) is an example of such a business model, and provides the service to efficiently spend online advertisement budgets on behalf of an advertiser. It implements advanced algorithms to compute and submit bids in real time for ad *impressions*, which are opportunities to show an ad creative to Internet users. The bidding strategy is designed to solve an ad campaign optimization problem.

There is an astronomically large number of impressions to bid on, which makes the optimization problem extremely high-dimensional, but the dimension can typically be reduced by reformulating the problem as a three player non-cooperative game. The three players that collectively produce bids on behalf of an advertiser are represented by impression valuation, campaign control, and bid shading optimization.

Impression valuation computes the expected value of an impression conditioned on it being awarded to the campaign [1]–[4]. Campaign control makes bid adjustments to satisfy campaign delivery constraints (more on that later) [5]–[9]. And bid shading optimization computes the final bid by taking into account how much other campaigns are expected to bid for the same impression [10]–[14].

[1]Niklas Karlsson is a fellow of IEEE and is a Senior Principal Research Scientist with Demand Tech, Amazon Ads, Mountain View, CA 94301 USA (e-mail: nkarlsson8@gmail.com).

All areas of bidding in online advertising are subject to intense research in both academia and industry [15]. This paper contributes to this body of work and is focused on the campaign control sub-problem. Campaign control is a critical feature of the bid computation, but is challenging due to plant uncertainties, nonlinearities, delays, time-variance, and noise. Multi-constraint optimization problems [16]–[18] are especially difficult to solve via feedback control due to the dynamic interaction across feedback loops.

Our contribution is a cascade feedback controller that solves the most commonly encountered multi-constraint problem in online advertising and a set of sufficient conditions for stability of the resulting nonlinear closed loop system. To the best of our knowledge, this is the first stability result for any of the multi-constraint problems described in [16]–[18] under a dynamic and nonlinear plant. Note, [16]–[18] only presents necessary conditions for optimality and does not examine under what conditions the closed loop system is stable. The stability results in this paper are derived by transforming the system into a Lure' system [19] and by engaging the circle criterion. The solution is successfully evaluated in a simulated environment based on artificial data produced to capture the dominant properties of a real online advertising campaign.

A cascade controller approach to the problem examined in this paper is explored also in [20]. The authors refer to their approach as "sequential pacing" and adopts a different parameterization. In contrast to our work, [20] evaluates the stability of their proposed algorithm neglecting the plant dynamics and based on a continuous-time approximation.

II. Optimization Problem

Consider managing an ad campaign via bidding on impressions sold in an open impression exchange. The objective is to solve a constrained optimization problem. Let Ω denote the set of all available impression opportunities $i \in \Omega$, where each impression is sold sequentially based on a first or second price cost model [21]. The cost model for each impression opportunity is decided by the seller of the impression and is known before a bid is computed. 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. Assume Let $\Omega_1 \cap \Omega_2 = \emptyset$ and $\Omega_1 \cup \Omega_2 = \Omega$. The decision variables are given by bid prices $b_i \in [0, \infty)$ produced by the optimization engine. Let $v_i \in [0, \infty)$ denote the expected value of impression i and $c_i \in [0, \infty)$ the cost of the impression, if awarded. The total expected cost for awarded impressions is denoted $\mathrm{E}C$, and the total expected value is denoted $\mathrm{E}V$.

The objective is to produce bid prices b_i, for all $i \in \Omega$, to maximize the cost-discounted profit $\mathcal{J} := \mathrm{E}V - \rho \mathrm{E}C$, subject to a spend constraint $\mathrm{E}C \leq \xi_1$ and a cost-benefit ratio (CBR) constraint $\mathrm{E}C \leq \xi_2 \mathrm{E}V$; where $\rho \in (0, 1]$ is a known cost-discount parameter, and $\xi_1, \xi_2 > 0$ are prescribed constraints on spend and CBR.

It is well-known (see e.g. [16]–[18]) that a necessary condition for optimality is that $b_i = b_i^{opt}$, for all i, where

$$b_i^{opt} = \begin{cases} \underset{b \in [0,\infty)}{\operatorname{argmax}} \ (b_i^u - b) w_i(b), & i \in \Omega_1, \\ b_i^u, & i \in \Omega_2, \end{cases}$$

and where $w_i(b)$ is the winrate (the probability impression i is awarded at bid price b), the *private value* b_i^u is defined by

$$b_i^u = \frac{1 + \xi_2 \lambda_2}{\lambda_1 + \lambda_2} v_i, \tag{1}$$

and where $\mathrm{E}C \leq \xi_1$, $(\lambda_1 - \rho)(\mathrm{E}C - \xi_1) = 0$, $\mathrm{E}C \leq \xi_2 \mathrm{E}V$, $\lambda_2(\mathrm{E}C - \xi_2 \mathrm{E}V) = 0$, $\lambda_1 \geq \rho$, and $\lambda_2 \geq 0$.

Assume $w_i(b)$ and v_i are known. In practice, they are estimated by separate systems not in scope of this paper. It remains to determine the optimal values of λ_1 and λ_2. The optimal bidding strategy for various generalizations of the above problem is derived in [16], [17] and [18].

III. Control Problem

The problem of finding the optimal λ_1 and λ_2 can in principle be achieved by adjusting their values in a feedback loop until they converge and satisfy constraints $\mathrm{E}C \leq \xi_1$, $(\lambda_1 - \rho)(\mathrm{E}C - \xi_1) = 0$, $\mathrm{E}C \leq \xi_2 \mathrm{E}V$, and $\lambda_2(\mathrm{E}C - \xi_2 \mathrm{E}V) = 0$. However, using λ_1 and λ_2 as the control levers in what is a multi-input multi-output control problem, is extraordinarily difficult due to the nonlinear, time-varying, and strong coupling across inputs and outputs. Note, for example, that the sensitivity in b_i^u due to changes in λ_1 (or λ_2) is highly dependent on λ_2 (or λ_1). In fact, $\partial b_i^u / \partial \lambda_2 = (\xi_2 \lambda_1 - 1) v_i / (\lambda_1 + \lambda_2)^2$ goes from being negative to being positive when λ_1 exceeds $1/\xi_2$. Hence, without careful consideration, λ_1 may turn a control system involving λ_2 from negative to positive feedback, and with instability as the result. To make the problem less challenging, a re-parametrization of the control levers is advised. To support this endeavour a few observations are in order.

The first observation is that the private value (1) can be expressed $b_i^u = m(\lambda_1, \lambda_2) v_i$, where bid modifier $m(\lambda_1, \lambda_2) := (1 + \xi_2 \lambda_2)/(\lambda_1 + \lambda_2)$. Hence, instead of identifying the optimal values of λ_1 and λ_2 separately and then computing $m(\lambda_1, \lambda_2)$, we may identify the optimal value of m directly.

The second observation is that $\mathrm{E}C$ and $\mathrm{E}C/\mathrm{E}V$ are both monotonic non-decreasing functions of m (see [6] for details). It can the be shown that the optimal bidding corresponds to the largest constant value of $m \leq \rho^{-1}$ for which $\mathrm{E}C \leq \xi_1$ and $\mathrm{E}C \leq \xi_2 \mathrm{E}V$ at the end of the flight [16].

The third and final observation is that instead of identifying the largest value of m maintaining the two inequalities, we may solve the problem in cascade by introducing an intermediate spend cap $\tilde{\xi}_1$ as follows: Find the largest $\tilde{\xi}_1 \leq \xi_1$ for which $\mathrm{E}C \leq \xi_2 \mathrm{E}V$, and simultaneously find the largest m for which $\mathrm{E}C \leq \tilde{\xi}_1$. Implemented as a feedback system, this delegates the maintenance of the spend constraint to an inner loop spend controller, and the maintenance of the CBR constraint to an outer loop CBR controller.

See Figure 1 for a block diagram of the cascade control system. In order to solve the problem using feedback, we

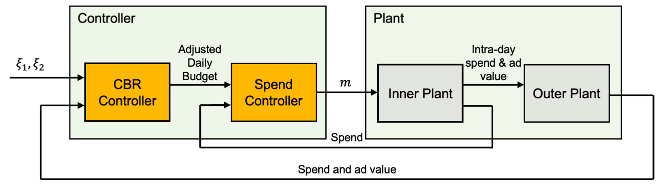

Fig. 1. Block diagram of the cascade control system architecture.

introduce time as a variable. Assume a discrete time implementation with a separate update frequencies for the two feedback systems.

The CBR controller first converts the flight budget ξ_1 into daily budgets $u_r(k)$ that sum up to ξ_1. These (unadjusted) daily budgets together with a CBR constraint ξ_2 and feedback on the daily spend $y_1(k)$ and observed ad value $y_2(k)$ are used to update an adjusted daily budget $u(k) \in [0, u_r(k)]$ once per day. The goal is for $u(k)$ to converge to the largest possible value for which $y_1(k) \leq \xi_2 y_2(k)$ is satisfied on average.

The spend controller consumes $u(k)$ obtained from the CBR controller and first distributes this budget throughout the day in a performance-optimal manner using a time-periodic feedforward controller. It then uses feedback on the spend in an error feedback controller to update $m(t) \in [0, \rho^{-1}]$ at time stamps t every few minutes. The objective is for $m(t)$ to converge to the largest value for which the daily spend $y_1(k)$ does not exceed $u(k)$. Designs for this control system is readily available in the literature [9], and outside the scope of this paper. The advantage of using a daily update cadence for the CBR controller is that it turns the effective plant perceived by the CBR controller approximately time-invariant. This time-invariance makes the design of the CBR controller less challenging.

Assume the inner loop spend controller is implemented and stable. The goal is to design the outer loop CBR controller such that $u(k) \in [0, u_r]$ converges to the largest value for which $y_1(k) \leq \xi_2 y_2(k)$ in expected sense.

IV. Plant Model

The plant as perceived by the CBR controller describes the mapping $u \mapsto [y_1, y_2]$. Assume y_1 in response to adjustments of u (via intra-day adjustments of m) is observed almost immediately. In particular, at the start of day k, the controller knows how much was spent the previous day. We make the simplifying assumption that the spend controller precisely delivers any adjusted daily budget u in full; i.e.,

$$y_1(k) = q^{-1} u(k), \tag{2}$$

where q is the forward-shift operator ($q^{-1}u(k) = u(k-1)$). While not guaranteed, since the inner loop controller operates at a much higher frequency than the outer loop, the assumption is typically a good approximation in practice. The assumption is used to simplify the derivations, and modest violations have not hurt the results in closed loop simulations.

Ad value y_2 encodes user activities such as clicks and sold products triggered by served impressions. It is related to u in a nonlinear and dynamic fashion, but the dynamics is assumed to be time-invariant (partially achieved by the daily update cadence of u). Assume $u > 0$ and $\mathrm{E}y_2 > 0$, and that the spend controller has converged to a stable equilibrium. The expected CBR $\eta(u) := \mathrm{E}y_1/\mathrm{E}y_2$ can then be shown to be positive, strictly monotonic increasing, and with a limit $\eta(u) \to 0$ as $u \to 0$ [6]. In other words, the expected eventual ad value satisfies $\mathrm{E}y_2 = \mathrm{E}y_1/\eta(u) = u/\eta(u)$; however, this value generation is typically subject to a multiple days delay. The delay is a priori unknown, but is assumed to be linear and time-invariant. In particular, the observed ad value $y_2(k)$ on day k satisfies

$$y_2(k) = q^{-1}P(q)\frac{1}{\eta(u(k))}u(k) + \epsilon(k), \qquad (3)$$

where P is a proper, linear, time-invariant transfer function with steady state gain one; and where $\epsilon(k)$ is independent mean zero noise.

While not needed in the forthcoming control design, it is useful for analysis and simulation tasks to leverage additional properties of the plant, especially as it relates to how ad value and CBR depend on m.

Theorem 4.1: If the expected ad spend $c(m) := \mathrm{E}C$ is differentiable and its derivative $c'(m)$ converges to zero, as $m \to 0$, at a rate that is at least linear, then the expected ad value $v(m) := \mathrm{E}V$ and CBR $\eta(m) := c(m)/v(m)$ are given by

$$v(m) = \int_0^m \frac{1}{x}c'(x)dx, \quad v(0) = 0, \qquad (4)$$

$$\eta(m) = \frac{c(m)}{\int_0^m x^{-1}c'(x)dx}, \quad \text{if} \quad v(m) > 0. \qquad (5)$$

Proof: By virtue of the cost model and the optimal bidding strategy (see Section II), the CBR for additional impressions awarded by making an infinitesimal adjustment of the bid modifier from m to $m + dm$ equals m. The additional spend for these incremental impressions is $dc = c'(m)dm$, which means the additional expected ad value from the impressions equals $(1/m)c'(m)dm$. Hence, the ad value satisfies $v(m + dm) = v(m) + (1/m)c'(m)dm$. Since $c'(m)$ converges to zero at least linearly, $(1/m)c'(m)$ is bounded for small m. The differential equation for $v(m)$ may therefore be integrated from 0 to m, resulting in $v(m) = \int_0^m x^{-1}c'(x)dx$, which confirms (4), and (5) is a trivial consequence of (4) and $\eta(m) = c(m)/v(m)$, completing the proof. ∎

In some cases, the ad spend $c(m)$ can be approximated by a parametric function with particularly convenient properties. A versatile family of two-parametric functions that are defined for non-negative values, and that describe functions that are monotonic increasing, and bounded is given by the family of so called gamma cumulative density functions (CDFs). The versatility comes from the fact that a wide range of functional shapes can be produced by selecting different values of the parameters defining the CDF.

Definition 4.1: The gamma cumulative density function (CDF), $F(x|\alpha, \beta)$, is defined for $x \geq 0$, and given by

$$F(m|\alpha, \beta) = \int_0^m \frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}dx, \quad \text{if} \quad m > 0, \qquad (6)$$

and $F(0|\alpha, \beta) = 0$, where shape parameter $\alpha > 0$ and inverse scale parameter $\beta > 0$, and where the gamma function is defined by $\Gamma(\alpha) = \int_0^\infty e^{-x}x^{\alpha-1}dx$.

Theorem 4.2: If the expected ad spend is a scaled gamma CDF defined by $c(m) = \gamma F(m|\alpha, \beta)$, where the spend capacity $\gamma > 0$, and $\alpha, \beta > 0$; then

$$v(m) = \frac{\gamma\beta}{\alpha - 1}F(m|\alpha - 1, \beta) \quad \text{and} \qquad (7)$$

$$\eta(m) = \frac{\alpha - 1}{\beta} \cdot \frac{F(m|\alpha, \beta)}{F(m|\alpha - 1, \beta)}. \qquad (8)$$

Proof: To prove (7), use $c(m) = \gamma F(m|\alpha, \beta)$, together with the definition of $F(m|\alpha, \beta)$ from (6), in (4) to obtain

$$\begin{aligned} v(m) &= \gamma \int_0^m x^{-1}\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}dx \\ &= \frac{\gamma\beta\Gamma(\alpha - 1)}{\Gamma(\alpha)} \int_0^m \frac{\beta^{\alpha-1}}{\Gamma(\alpha - 1)}x^{\alpha-2}e^{-\beta x}dx. \end{aligned}$$

The integral expression is recognized as the gamma CDF $F(x|\alpha - 1, \beta)$, and it is well-known that $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$, hence $v(m) = \gamma\beta F(m|\alpha - 1, \beta)/(\alpha - 1)$, which completes the proof of (7). To prove (8), combine $\eta(m) = c(m)/v(m)$, $c(m) = \gamma F(m|\alpha, \beta)$, and (7); which immediately yields the result and completes the proof. ∎

V. Controller Design and Analysis

Consider a linear time-invariant controller composed of a feedforward component C_{ff} and a feedback component C_{fb} based on error signal e and feedback mechanism u

$$e(k) = C_{ff}(q)y_1(k) - \xi_2 y_2(k), \quad \text{and} \qquad (9)$$

$$u(k) = C_{fb}(q)e(k). \qquad (10)$$

Controller components $C_{ff}(q)$ and $C_{fb}(q)$ are proper, linear, and time-invariant transfer functions; and $C_{ff}(q)$ has steady state gain one. The objective is to design $C_{ff}(q)$ and $C_{fb}(q)$ so that u converges to the largest possible value for which $\mathrm{E}y_1(k) \leq \xi_2 \mathrm{E}y_2(k)$. Assume there exists a strictly positive solution u for which $\mathrm{E}y_1(k) = \xi_2 \mathrm{E}y_2(k)$. This implies $\mathrm{E}e(k) = 0$ at steady state. Recalling the definition of the expected CBR, it is noted that $\eta(u) := \mathrm{E}y_1/\mathrm{E}y_2 = \xi_2$ if the solution has been found.

However, $u = 0$ also yields $\mathrm{E}y_1(k) = \mathrm{E}y_2(k) = 0$, which corresponds to $\mathrm{E}e(k) = 0$ at steady state. This is an undesirable suboptimal equilibrium, and to prevent getting trapped in this equilibrium, we define a permissible range of values for the adjusted budget $u \in [u_{min}, u_r]$, where $u_{min} > 0$ is a configured parameter chosen so small that we confidently expect the optimal u to be larger.

Combine (2), (3), (9), and (10) to obtain

$$e = q^{-1}\left(C_{ff} - P\frac{\xi_2}{\eta(u)}\right)C_{fb}e - \xi_2\epsilon, \quad \text{and} \tag{11}$$

$$u = q^{-1}C_{fb}\left(C_{ff} - P\frac{\xi_2}{\eta(u)}\right)u - \xi_2 C_{fb}\epsilon, \tag{12}$$

where arguments k and q are omitted from signals and operators to reduce the clutter. This is a nonlinear dynamical system and it is not obvious under what conditions the system is stable. Note, for example, that P and $\eta^{-1}(u)$, as well as, C_{fb} and $\eta^{-1}(u)$ do not commute; i.e., $P\eta^{-1}(u) \neq \eta^{-1}(u)P$ and $C_{fb}\eta^{-1}(u) \neq \eta^{-1}(u)C_{fb}$. The following main result of the paper establishes sufficient conditions for stability.

Theorem 5.1: Let $H(z)$ be the discrete time transfer function (in the $\mathcal{Z}$-domain) defined by

$$H(z) = \left(z - C_{fb}(C_{ff} - P)\right)^{-1} C_{fb}P. \tag{13}$$

If there exists a $u^{opt} \in [u_{min}, u_r]$ such that $\eta(u^{opt}) = \xi_2$, no pole of $H(z)$ is outside the unit circle, and $\text{Re}[H(e^{j\omega})] > -1$ for $\pi/2 \leq \omega \leq \pi/2$; then $u = u^{opt}$ is the unique and absolutely stable solution of (11)-(12) with a finite domain of $u \in [u_{min}, u_r]$.

Proof: First rewrite (12) as

$$u = q^{-1}C_{fb}\left(C_{ff} + P\left(1 - \frac{\xi_2}{\eta(u)}\right) - P\right)u - \xi_2 C_{fb}\epsilon.$$

Define $\Psi(u) := 1 - \xi_2/\eta(u)$ and $\delta := \Psi(u)u$, which implies $u = q^{-1}C_{fb}(C_{ff} - P)u + q^{-1}C_{fb}P\delta - \xi_2 C_{fb}\epsilon$. It follows that the closed loop system dynamics is described by

$$u = \left(q - C_{fb}(C_{ff} - P)\right)^{-1}\left(C_{fb}P\delta - \xi_2 q C_{fb}\epsilon\right),$$
$$\delta = \Psi(u)u.$$

This is a feedback connection of a linear time-invariant dynamical system and a memory-less nonlinear element. The system is depicted as a block diagram in Figure 2 and is recognized as a so called Lure' system [19]. Recall that $\eta(u)$ is strictly monotonic increasing and satisfies $\lim_{u\to 0^+}\eta(u) = 0$. Consequently, $\Psi(u) := 1 - \xi_2/\eta(u)$ is defined for all $u \geq u_{min} > 0$ and is an increasing function satisfying $\lim_{u\to 0^+}\Psi(u) = -\infty$, and $\Psi(u) \leq 1$.

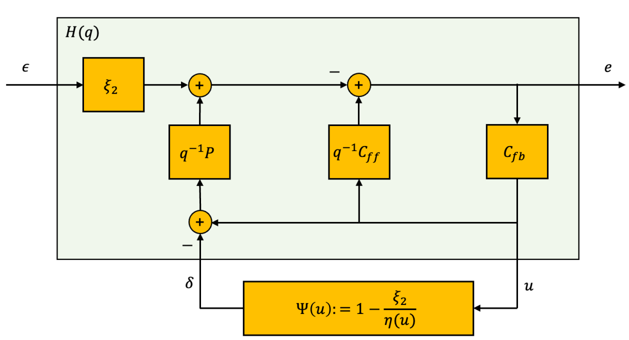

Fig. 2. Block diagram of the closed loop dynamical system depicted as a Lure system composed of a linear time-invariant operator $H(z)$ and a nonlinear memory-less function $\Psi(u)$.

For the purpose of stability analysis, we may set $\epsilon = 0$. The dynamics under examination then equals

$$u = H(z)\delta, \tag{14}$$
$$\delta = \Psi(u)u. \tag{15}$$

By assumption, there exists a $u^{opt} \in [u_{min}, u_r]$ such that $\eta(u^{opt}) = \xi_2$, which corresponds to $\Psi(u^{opt}) = 0$. Since $\Psi(u)$ is an increasing function, the solution is unique. Furthermore, since only strictly positive solutions are considered, u^{opt} uniquely also solves $\delta(u^{opt}) = \Psi(u^{opt})u^{opt} = 0$.

Moreover, the upper bound and monotonicity of $\Psi(u)$ imply that $\theta(u - u_{opt}) \leq \delta(u) \leq u - u_{opt}$, if $u \geq u_{opt}$, and $u - u_{opt} \leq \delta(u) \leq \theta(u - u_{opt})$, if $u_{min} \leq u \leq u_{opt}$, for some $\theta > 0$. It follows that the nonlinearity $\delta(u)$ satisfies the sector condition [19] centered around $u = u_{opt}$.

$$\theta(u - u_{opt})^2 \;\leq\; (u - u_{opt})\delta(u) \;\leq\; (u - u_{opt})^2.$$

The conditions for the circle criterion (Theorem 7.2 [19]) are satisfied, which states that if $H(z)$ has no unstable poles and the Nyquist plot of $H(e^{j\omega})$ does not enter the circular disk in the imaginary plane going through points $-1/\theta$ and -1, then the feedback connection is absolutely stable with a finite domain. Since $\theta > 0$ may be arbitrarily small (determined by u_{min}), the circular disk may be very large, but still going through -1. Hence, $\text{Re}[H(e^{j\omega})] > -1$ ensures the Nyquist plot certainly does not enter the circular disk. This completes the proof. ∎

Example 5.1: To demonstrate how Theorem 5.1 can be used to examine the stability of a system, consider a plant $P = 0.5/(z - 0.5)$, and a candidate controller defined by $C_{ff} = P$ and $C_{fb} = 0.3z/(z - 1)$. Equation (13) yields the transfer function from δ to u as $H(z) = 0.15/(z^2 - 1.5z + 0.5)$. The pole-zero map and the Nyquist plot of $H(z)$, produced in Matlab, are shown in Figure 3. It is

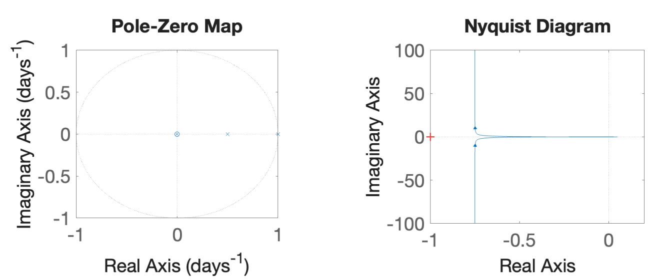

Fig. 3. The pole-zero map and the Nyquist diagram of the dynamic component of a feedback system to in Example 5.1 to demonstrate how to check for the stability of a feedback interconnection with a nonlinear memoryless function.

noted that no pole is outside the unit circle and that the Nyquist curve is on the right side of -1 for all frequencies. Hence, according to Theorem 5.1, the feedback system has a unique and absolutely stable solution with a finite domain of $u \in [u_{min}, u_r]$ for some value of $u_{min} > 0$.

Next section demonstrates the closed loop behavior of a system designed with help of Theorem 5.1.

VI. Simulation Results

To evaluate the theoretical results in Section V, we conduct a simulation study designed to capture the dominant properties of a real advertising problem. Many properties are well-understood from years of operation, but there are also edge case scenarios such as sparse discrete-valued feedback y_2 and isolated extreme disturbances. These scenarios require experimentation which is outside the scope of this paper.

A. Set-up

Consider a 30 days long ad campaign, where the inner loop spend controller updates m once every $\Delta = 2/60$ hours. Assume the observed intraday spend $y_1^{intra}(m,t)$ satisfies

$$y_1^{intra}(m,t) = \frac{\gamma\Delta}{24}\left(1 + 0.8\sin\left(\frac{2\pi t}{24}\right)\right)F(m|\alpha,\beta)(1+\epsilon_1(t)),$$

where $\gamma = 4000$, $F(m|\alpha,\beta)$ is a gamma CDF (Definition 4.1), and where $\epsilon_1(t)$ is independent and identically distributed noise with mean zero and standard deviation 0.2. Furthermore, assume $\alpha = 1/\sigma_{rel}^2$ and $\beta = 1/(\mu\sigma_{rel}^2)$, where $\mu = 6$, $\sigma_{rel} = 0.8$. The sinusoidal component of y_1^{intra} simulates the time-of-day pattern in Internet traffic with less impressions available during the night than during the day.

Theorem 4.2 can be used together with (15) to derive closed form expressions for $\mathrm{E}y_1$, η, and δ in terms of gamma CDFs. The result is depicted graphically in Figure 4. The left subplots show how $\mathrm{E}y_1$ and η are monotonic increasing functions of m, whereas the right subplots show that η is a monotonic increasing function of u and that $\delta(u)$ satisfies the sector condition $\theta(u-u_{opt})^2 \leq (u-u_{opt})\delta(u) \leq (u-u_{opt})^2$ (see also the proof of Theorem 5.1).

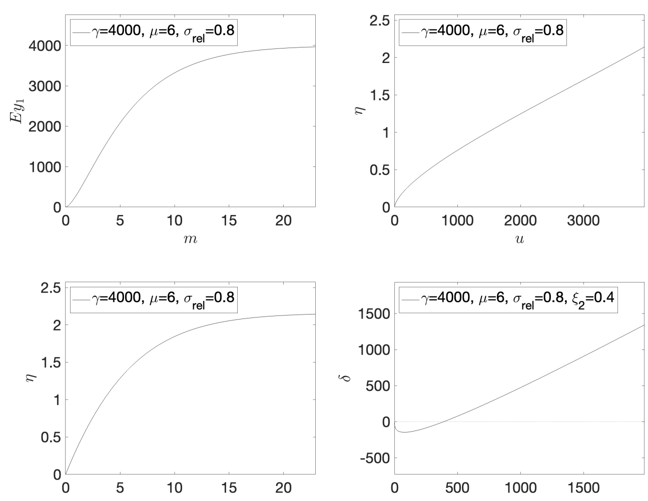

Fig. 4. The plant model for the campaign being simulated, defined by key input-output relationships.

The advertiser defines ad value as the number of sold products. Assume, at steady state, it is a Poisson random variable $y_2(k) = u/\eta(u) + \epsilon_2(k)$ with mean $u/\eta(u)$ and where $\epsilon_2(k)$ is mean zero random noise. The advertiser specifies daily constraints on spend and CBR, and the constraints are updated a few times throughout the flight according to

$$u_r(k) = \begin{cases} 600, & \text{if } 1 \leq k \leq 10 \\ 500, & \text{if } 11 \leq k \leq 20 \\ 700, & \text{if } 21 \leq k \leq 30 \end{cases},$$

$$\xi_2(k) = \begin{cases} 0.4, & \text{if } 1 \leq k \leq 15 \\ 0.8, & \text{if } 16 \leq k \leq 30 \end{cases}.$$

Suppose the delay between spend and ad value is given by $P(z) = 0.4z/(z-0.6)$ corresponding an average delay of 48 hours in continuous time.

The most likely violated assumptions in the real world and with most significant impact is that $\eta(u)$ is monotonic and that $P(z)$ is linear time-invariant and known. The monotonicity is based on the premise that the inner-loop spend controller is stable and non-volatile, which is sometimes hard to achieve in practice. Experiments will assess the implications of violating these assumptions, and others.

B. Control Design

Assume the inner loop spend controller is implemented using the design proposed in [9]. For the outer loop CBR controller, consider the design in Example 5.1; i.e., $C_{ff}(z) = P(z)$ and $C_{fb}(z) = 0.3/(z-1)$ which is a pure integral error feedback controller. In a real application, plant $P(z)$ is not known precisely and C_{ff} can at most be an estimate of the real plant.

A simulated outcome of the campaign defined in the previous section managed by the above control system is depicted in Figure 5.

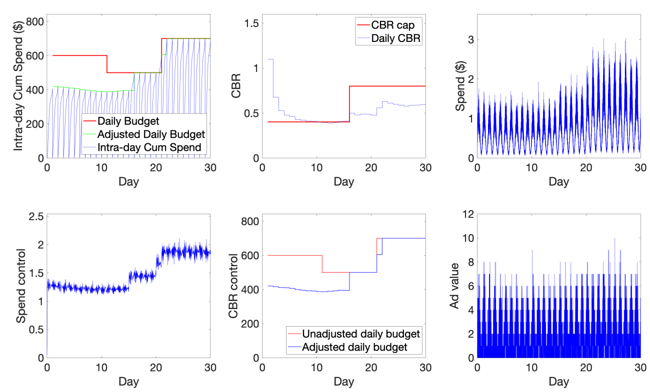

Fig. 5. Closed loop simulation results for the campaign under control by the proposed multivariable control system.

The top-left panel shows the daily budget (u_r) as a red curve, the adjusted daily budget (u) produced by the CBR controller as a green curve, and the intra-day cumulative spend as a blue curve. It is noted that the spend constraint is maintained throughout the flight.

The top-center panel shows the CBR constraint as a red curve and the daily observed CBR as a blue curve. It is noted that the CBR constraint is initially violated due to a poor initialization, but recovers throughout the flight illustrating the stability. It is noted that during the first 15 days of the flight the CBR constraint is limiting the campaign delivery, whereas in the last 15 days it is the spend constraint keeping the campaign from delivering more.

The top-right panel displays the intra-day spend, which exhibits a distinct time-of-day pattern and contains multiplicative noise. Next, the bottom-left panel presents spend control signal m, which is updated at high frequency and responds to the intra-day spend noise. Furthermore, the unadjusted daily budget (u_r) and the adjusted daily budget (u) are shown in the bottom-center panel. Finally, the intraday observed ad value is shown in the bottom-right panel. Based on the convergence of the spend and the CBR, and based on how the constraints are maintained at steady-state, we conclude for this example that the control system behaves as desired and expected.

VII. Conclusions and Future Work

We have proposed a cascade multivariable control system to manage the spend (budget utilization) and the cost-benefit ratio, CBR, (performance) of an online advertising campaign. The plant is nonlinear, time-varying, and dynamic; and is subject to noise and uncertainties. The control system is composed of an inner loop spend controller (developed separately), and an outer loop CBR controller. Besides deriving certain key plant properties for the inner loop plant, this paper is focused on the outer loop control system, which consists of a feedforward and a feedback component. The main result of the paper are sufficient conditions for stability of the CBR controller. These results are obtained by transforming the closed loop system into a Lure' system, identifying the applicable sector condition, and making use of the circle criterion. The results demonstrate how concepts from nonlinear system theory can be used to establish key stability conditions in applications such as online advertising.

Immediate future work includes experimental validation of the mathematical results. Real advertising campaigns are difficult to model and simulation results cannot capture all significant behaviors that are present in the real plant. Near-term future work also includes determining conditions on C_{ff} and C_{fb} to ensure stability. The control system must also be robust to model uncertainties and must be sufficiently responsive to load disturbances and measurement noise, which requires additional research. In particular, important areas of future work is to determine conditions for robustness to violations of the assumptions that $\eta(u)$ is monotonic and that $P(z)$ is linear time-invariant and known. We conjecture the circle criterion framework can be extended to establish robustness margins and uncertainty bounds.

Furthermore, the derived stability conditions are sufficient but not necessary, hence, they may be overly conservative. Future work therefore also includes a search for tighter sufficiency conditions. Finally, advertisers are increasingly interested in a multitude of objectives and constraints. These lead to additional feedback control problems that must be solved simultaneously. Consequently, additional research is needed to develop multivariable control systems that solve a more general multi-input multi-output control problem.

References

[1] Ruoxi Wang, Bin Fu, Gang Fu, and Mingliang Wang. Deep & cross network for ad click predictions. *CoRR*, abs/1708.05123, 2017.

[2] 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 2018 World Wide Web Conference*, WWW '18, pages 1349–1357.

[3] Junwei Pan, Wei Xue, Ximei Wang, Haibin Yu, Xun Liu, Shijie Quan, Xueming Qiu, Dapeng Liu, Lei Xiao, and Jie Jiang. Ads recommendation in a collapsed and entangled world. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, KDD '24, page 5566–5577. ACM, August 2024.

[4] Wei Deng, Junwei Pan, Tian Zhou, Deguang Kong, Aaron Flores, and Guang Lin. Deeplight: Deep lightweight feature interactions for accelerating ctr predictions in ad serving. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*, WSDM '21, page 922–930, New York, NY, USA, 2021. Association for Computing Machinery.

[5] 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.

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

[7] 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.

[8] Niklas Karlsson. Adaptive optimization and control in online advertising. In *Proceedings of the 2019 IEEE 58th Conference on Decision and Control, Nice, France*, Dec 11-13 2019.

[9] Niklas Karlsson. Feedback control in programmatic advertising: The frontier of optimization in real-time bidding. *IEEE Control Systems Magazine*, 40(5):40–77, 2020.

[10] Djordje Gligorijevic, Tian Zhou, Bharatbhushan Shetty, Brendan Kitts, Shengjun Pan, Junwei Pan, and Aaron Flores. Bid shading in the brave new world of first-price auctions. *Proc. of the 29th ACM International Conference on Information & Knowledge Management*, Oct 2020.

[11] Shengjun Pan, Brendan Kitts, Tian Zhou, Hao He, Bharatbhushan Shetty, Aaron Flores, Djordje Gligorijevic, Junwei Pan, Tingyu Mao, San Gultekin, and Jianlong Zhang. Bid shading by win-rate estimation and surplus maximization. 2020.

[12] Niklas Karlsson and Qian Sang. Adaptive bid shading optimization of first-price ad inventory,. In *Proceedings of the 2021 American Control Conference, New Orleans, LA, USA*, ACC '21, pages 4971–4978, May 26–28 2021.

[13] 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 *Proc. of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'21), August 14-18, 2021, Singapore*, 2021.

[14] Yanlin Qu, Ravi Kant, Yan Chen, Brendan Kitts, San Gultekin, Aaron Flores, and Jose Blanchet. Double distributionally robust bid shading for first price auctions, https://arxiv.org/abs/2410.14864, 2024.

[15] Weitong Ou, Bo Chen, Xinyi Dai, Weinan Zhang, Weiwen Liu, Ruiming Tang, and Yong Yu. A survey on bid optimization in real-time bidding display advertising. *ACM Trans. Knowl. Discov. Data*, 18(3), 2023.

[16] 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.

[17] 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.

[18] Niklas Karlsson. Feedback-control based hierarchical multi-constraint ad campaign optimization. In *2023 62nd IEEE Conference on Decision and Control (CDC)*, pages 7568–7573, 2023.

[19] Hassan K. Khalil. *Nonlinear Systems*. Prentice Hall, 3rd edition, 2002.

[20] Santiago R. Balseiro, Kshipra Bhawalkar, Zhe Feng, Haihao Lu, Vahab S. Mirrokni, Balasubramanian Sivan, and Di Wang. Joint feedback loop for spend and return-on-spend constraints. *ArXiv*, abs/2302.08530, 2023.

[21] Vijay Krishna. *Auction Theory*. Academic Press, 2002.