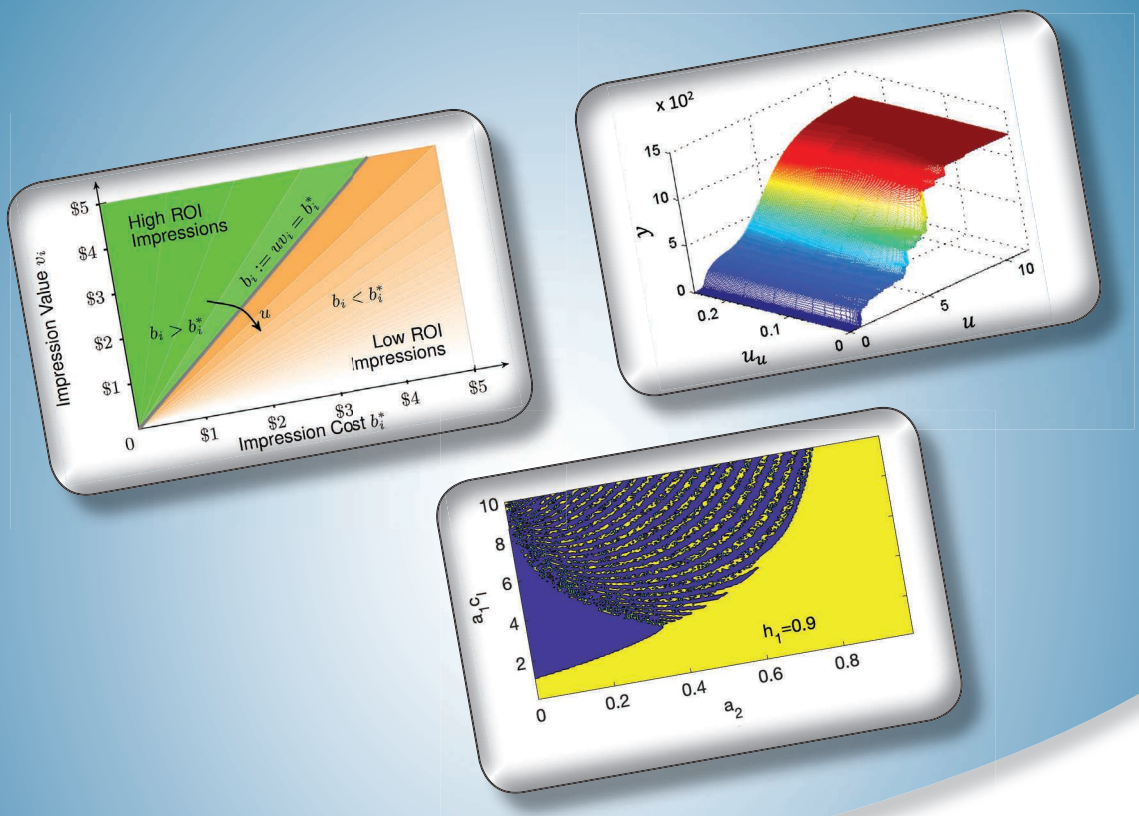
Feedback Control in Programmatic Advertising



THE FRONTIER OF OPTIMIZATION IN REAL-TIME BIDDING

NIKLAS KARLSSON

Feedback control is critical in the scalable optimization of Internet advertising, and it is, therefore, an enabling technology. However, it is challenging to model the plant and design the controller because the plant is nonlinear, time varying, stochastic, and poorly known. A closed-loop system model easily becomes unrealistic or extremely complicated and intractable to analyze.

Some of the technical challenges are discussed at length in this article. However, first consider the business significance of online advertising as a whole. Worldwide digital advertisement (*ad* for short) spending is growing at a double-digit rate and is projected to reach US$385 billion in 2020 [1]. The same projection for the United States alone is US$81 billion, out of which US$70 billion is programmatic advertising [2].

As discussed in "Summary," the goal of this article is to uncover why and how feedback is used in online advertising. Feedback control is used in most areas of the industry. However, a particularly interesting application is found in the subdiscipline of *programmatic advertising*, which is the automated optimization of advertising. An advertiser's goal in programmatic

Digital Object Identifier 10.1109/MCS.2020.3005013
Date of current version: 16 September 2020

1066-033X/20©2020IEEE

advertising is to exploit an automated system and bid in real time on advertising inventory sold on one or more open exchanges, such that an ad is shown to the right person at the right time and place. The inventory consists of impressions, which provide an opportunity to show an ad creative (for example, a banner ad, text ad, or preroll video commercial) to an Internet user. Programmatic advertising is at the heart of the business model for companies such as Google, Facebook, and Verizon Media.

Before diving deep into feedback control, it is helpful to understand the impression allocation process of programmatic advertising. The basic process is illustrated in Figure 1. An Internet user loads a webpage to consume content. It may be a social media webpage, news site, or any other Internet destination. The publisher of the webpage presents the requested content to the user (for example, a news article) and often makes money via the webpage traffic by also selling space on the page for advertising. This so-called ad inventory is frequently sold on an open exchange using an auction. A publisher may then leverage a *supply-side platform* (SSP) to sell its ad inventory. An SSP is a technology platform that enables web publishers to manage their advertising space inventory, fill it with ads, and receive revenue.

Many advertisers have an interest in serving ads to each user. A *demand-side platform* (DSP) on behalf of an advertiser (sometimes via an agency) computes and submits a bid in real time for the opportunity to show its ad. A DSP is a system that allows buyers of digital advertising inventory to manage multiple ad- and data-exchange accounts through one interface and with built-in optimization capabilities. The highest-bidding advertiser wins the impression opportunity and serves the ad. A DSP is a business model specific for programmatic advertising.

Programmatic advertising in general (and DSPs specifically) addresses many problems and technologies. It is not practical to cover everything in a single article. This article considers one popular DSP optimization problem; demonstrates how to turn it into a control problem; and illustrates, in a tutorial fashion, how to model the plant, design the controller, and establish stability conditions of the closed-loop system. The optimization problem is a value-maximization problem subject to a single constraint. Bidding across campaigns is noncooperative, and we assume, without loss of generality, that the DSP represents a single

Summary

The use of feedback control in online advertising is a hidden technology, and the goal of this article is to uncover why and how feedback is used. First, programmatic advertising is introduced as a system for the automated optimization of online advertising based on real-time bidding. The decision variables are given by the bids for a specific ad campaign, which amount to millions or billions every day. A popular optimization problem is defined and decomposed into separate prediction and control problems.

Thereafter, physical reasoning is used to justify a plant model that forms the basis for control design. The model is dynamic, discontinuous, periodic, and stochastic. To overcome the challenges associated with discontinuities, a bid-randomization technique is introduced. The randomization of bids makes plant linearization possible, and a control system is then proposed to take advantage of this. It consists of pure integral (I)-error feedback control, dynamic periodic feedforward control, and persistent excitation. Conditions for the stability of the expected trajectory and the variance of the trajectory are derived.

The stability conditions involve the loop gain, which is the product of the controller and plant gains. To ensure stability, a system-identification algorithm is proposed to estimate the plant gain. It is a recursive least-squares algorithm and completes an adaptive control system. The control system is simulated, and the impact of adding bid randomization and persistent excitation is demonstrated. The examples demonstrate the complementary benefit of bid randomization, persistent excitation, dynamic periodic feedforward control, and pure I-error feedback control.

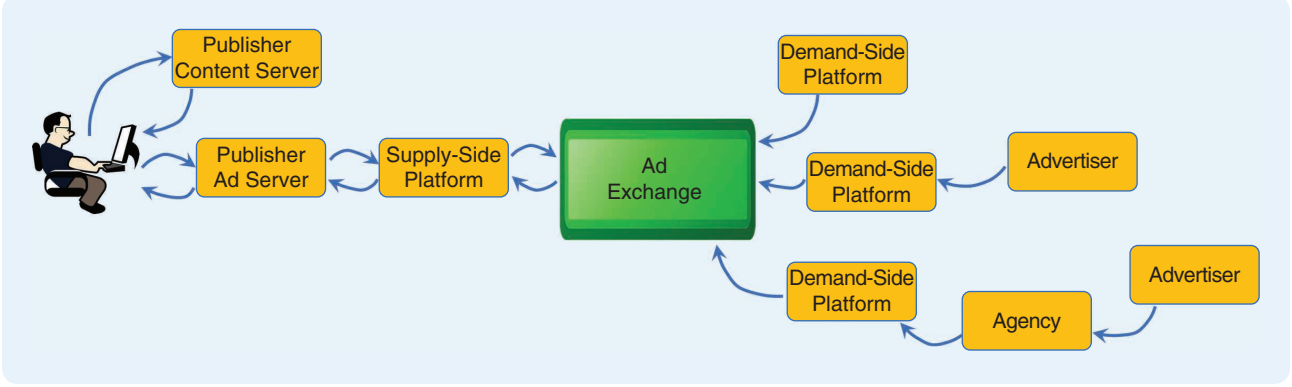


FIGURE 1 The basic advertisement allocation process.

ad campaign. Millions or billions of impression opportunities are available every day, and each bid for an impression is a decision variable. Scalability is, therefore, an important consideration in solving the problem. In this article, the problem is first decomposed into separate impression valuation and control subproblems. Thereafter, the control problem is discussed in depth, and it is shown how to overcome some of the obstacles related to plant modeling and control design.

Value maximization depends on how the advertiser defines advertising value. Each ad campaign is associated with its own definition and involves some combination of branding and performance. *Branding* is about reaching new Internet users, developing a brand recognition, and planting the seeds for future customers. *Performance*, on the other hand, is about generating near-term sales and other user engagement. The most basic pure branding objective is a value function defined as the total number of served impressions. Similarly, the most trivial pure performance objective is a value function defined as the total number of conversions (where a *conversion* is defined to be a user action, such as a product purchase).

The most common constraint of the optimization problem (and the only constraint considered in this article) relates to the total spend. An advertiser specifies a total campaign budget and wishes to spend the budget evenly over time. The advertiser does not want to spend everything on the first day, or the last day, of a campaign flight. Other constraints may involve how much to spend across different inventory sources, how much (at most) to pay for individual impressions, or how much (at most) to pay, on average, per click or per action. Although outside the scope of this article, the principal methodology used for the single-constraint case can be generalized to problems involving multiple constraints.

Feedback control has been used in online advertising for at least 15 years. However, most published work is not based on solid control engineering principles. The lack of a solid foundation is unfortunate and a missed opportunity. One explanation is that few algorithm developers in the online advertising industry have a background in control theory. Online advertising engineers frequently reinvent techniques that were developed in the control community a long time ago, or they struggle with problems that have published and well-documented solutions. Because of the algorithm developers' inadequate background in feedback systems, it is also common that control problems are not recognized as such or the solutions end up being nonrobust or unstable.

As an example, system behavior that is triggered by feedback is often mischaracterized simply as process noise. The true cause may be instability, limit cycles, bifurcations, chaos, or other behaviors that are introduced by human-made feedback algorithms. The limited appreciation of what feedback can result in (for better or worse) and the mischaracterization of the system behavior have a dramatic impact on the evolution of algorithms. Algorithms are made increasingly complex to address what is misperceived as process noise. The added complexity not only makes the systems harder to maintain, but it also introduces additional nonlinear behavior. Moreover, the incorrect understanding of the dynamics compromises both performance and robustness. Recognizing the true cause of system behavior may guide the engineer to a simple solution with better performance and robustness.

An early pre-DSP publication on feedback control applied to online advertising [3] outlines several important challenges and provides a glimpse into a possible plant model and control design but omits details. A more comprehensive overview of the control problem is presented in [4], in which bid randomization [5] is proposed to overcome complexities due to a discontinuous plant. Bid randomization is used in [6] as a secondary feedback control signal in a model reference adaptive controller to regulate the effective loop gain. In [7], a model reference adaptive controller is proposed for a single control signal to implement a high-fidelity reference tracking system for a budget-constrained campaign.

An extensive and insightful overview of optimal real-time bidding for display advertising is provided in [8]. Many aspects of the problem are covered in some detail. However, the treatment related to feedback control is thin and somewhat ad hoc. An updated treatment of the control problem by the same author in [9] adds detail. However, it is largely an empirical study of the observed performance of an ad hoc actuator combined with a standard proportional-integral-derivative controller. The ability to model and simulate the plant is a key aspect of developing a controller. A systematic methodology for offline modeling of advertising plants is proposed in [10]. The methodology is used to simulate realistic plants in a testbed for control algorithms.

A primitive control system that considers the periodic aspect of the plant is proposed in [11]. It is a proportional-integral (PI)-error feedback controller with a periodic proportional gain. An improved solution (and the basis of the control design in this article) is described in [12]. That control system consists of a periodic feedforward controller, pure integral (I)-error feedback controller, persistent-excitation controller, and plant gain estimator. It is shown how the relationship between the loop gain (the product of the controller and plant gains), plant seasonality, and latency dictates the closed-loop stability.

A robust method to estimate the seasonality of an online advertising process is proposed in [13], and a Bayesian plant gain estimator involving a prior estimate of the response curve and bid randomization is presented in [14]. In [15], a framework is provided to analyze the budget-pacing problem by means of variational calculus techniques. A continuous-time approach is taken, and a simple

characterization of the optimality conditions is described, which leads to an implementation in the form of a feedback control system. However, it is strictly a theoretical method, and its practical value is unclear.

Feedback control solves many problems in programmatic advertising that are not covered in this article (for example, control of the click-through rate, conversion rate, or any other event rate). A solution for event-rate control of a discontinuous plant is developed in [16] and [17]. The solution includes a novel approach of making the discontinuous plant effectively continuous. The proposed smoothing process is inspired by the method described in this article, but with fundamental differences since it does not perturb the bids. The most important parameters, variables, and functions used in this article are listed in Table 1 with their names, symbols, units, and a brief description.

OPTIMIZATION PROBLEM FORMULATION

Consider an advertising campaign managed by a DSP on behalf of an advertiser. The objective is to spend an advertising budget in a way that maximizes the expected total advertising value. Set Ω represents all impression opportunities i made available for sale in one or more real-time marketplaces. The decision variables are given by the bid prices $b_i \in \mathbb{R}$ for all $i \in \Omega$, which are used to compete for the impression inventory. *Budget constraint* ξ and *impression value* v_i are both defined by the advertiser. Typically, v_i depends on a user's likelihood to engage with the ad (performance) or a user's exposure to the ad (branding).

Any DSP or advertiser is permitted to participate in the bidding, and the marketplaces are, therefore, referred to as open. Furthermore, each impression $i \in \Omega$ is allocated in a sealed auction; that is, participating bidders independently submit bids with no knowledge of other bids. The highest bidder wins and pays an amount equal to the second-highest bid, which corresponds to a so-called *second-price cost model* [18].

The *highest competing* bid of impression i is unknown and represented by the variable $B_i^\star \in \mathbb{R}$, where the uppercase letter indicates it is random. A lowercase letter $b_i^\star$ denotes its expected value (if it is a future event) or a realization (if it is a historical event).

The cumulative distribution function of $B_i^\star$ is denoted $F_{B_i^\star}(b)$, while the probability density function is $f_{B_i^\star}(b) = dF_{B_i^\star}(b)/db$. Both are unknown but assumed continuous, with support for nonnegative bid values. The impression is awarded if $b_i \geq b_i^\star$, where $b_i^\star$ (according to the described convention) is the realized value of $B_i^\star$. In this case, the advertiser is charged an *ad cost* $b_i^\star$ (consistent with the second-price cost model).

Typically, the impression value $v_i \in \mathbb{R}_{\geq 0}$ of an awarded impression i is a priori unknown and must be estimated. However, to keep things focused on feedback control, it is assumed in this article to be a known quantity. Let V denote the *total value* and C denote the *cumulative ad cost*, where

$$V = \sum_{i \in \Omega} v_i \mathbb{I}_{\{b_i \geq B_i^\star\}}, \tag{1}$$

$$C = \sum_{i \in \Omega} B_i^\star \mathbb{I}_{\{b_i \geq B_i^\star\}}, \tag{2}$$

and where $\mathbb{I}_X$ is the indicator function satisfying $\mathbb{I}_X = 1$ if $X =$ true (and $\mathbb{I}_X = 0$ if $X =$ false).

Mathematically, the optimization problem is defined by

$$\underset{\{b_i \in \mathbb{R} \mid \forall i \in \Omega\}}{\text{maximize}} \; \mathrm{E}V \tag{3}$$

subject to

$$\mathrm{E}C \leq \xi, \tag{4}$$

where $\xi \in \mathbb{R}_{\geq 0}$ is the specified advertising budget. In real applications, advertisers often care about additional metrics. For example, an advertiser may set a maximum acceptable average cost per conversion or click or an upper bound of how much to pay for an individual impression.

AN OPTIMAL BIDDING MECHANISM

The derivation of an optimal bidding mechanism makes use of the following relationships.

Lemma 1

$$\mathrm{E}(\mathbb{I}_{\{b_i \geq B_i^\star\}}) = F_{B_i^\star}(b_i), \tag{5}$$

$$\mathrm{E}(B_i^\star \mathbb{I}_{\{b_i \geq B_i^\star\}}) = b_i F_{B_i^\star}(b_i) - \int_0^{b_i} F_{B_i^\star}(z)\, dz. \tag{6}$$

Proof

1) Identity (5) is obtained from $\mathrm{E}(\mathbb{I}_{\{b_i \geq B_i^\star\}}) = \int_{-\infty}^{\infty} \mathbb{I}_{\{b_i \geq z\}} f_{B_i^\star}(z)\, dz = \int_{-\infty}^{b_i} f_{B_i^\star}(z)\, dz = F_{B_i^\star}(b_i)$.
2) Identity (6) is obtained via integration by parts from $E(B_i^\star \mathbb{I}_{\{b_i \geq B_i^\star\}}) = \int_0^{\infty} z \mathbb{I}_{\{b_i \geq z\}} f_{B_i^\star}(z)\, dz = \int_0^{b_i} z f_{B_i^\star}(z)\, dz = b_i F_{B_i^\star}(b_i) - \int_0^{b_i} F_{B_i^\star}(z)\, dz$, which completes the proof. ■

Since the number of impression opportunities is extraordinarily large, so is the dimensionality of the optimization problem. This, combined with the unknown distribution of the highest competing bid $B_i^\star$, makes a centralized and plan-based solution impractical. Instead, the Lagrangian method is used to rewrite the problem as two largely independent subproblems that can be solved with feedback-based and scalable methods. The result is given by the following theorem.

Theorem 1

The optimal bid price b_i^{opt} solving (3) and (4) takes the form

$$b_i^{\text{opt}} = u v_i,$$

TABLE 1 **The important parameters, variables, and functions used in the article.**

	Name	Symbol	Unit	Description
Bid optimization	Impression	i	—	Index representing an impression opportunity
	Bid price	B_i, b_i	US$	Bid used by our advertiser for impression i
	Bid uncertainty	$b_{u,i}, u_u$	—	Relative standard deviation used to randomly perturb bid b_i
	Highest competing bid price	B_i^*, b_i^*	US$	Highest bid used by any other advertiser for impression i
	Impression value	v_i	US$	Performance + branding value of impression i
	Total campaign value	V, v	US$	Total value of awarded impressions
	Total campaign cost	C, c	US$	Total cost of awarded impressions
	Campaign budget	ξ	US$	Budget available for acquisition of impressions
	Lagrange multiplier	λ	—	Multiplier associated with constrained optimization problem
Plant smoothing	Control signal	U, u	—	Random final and nominal control, $\mathrm{E}(U) = u$
	Uncertainty signal	u_u	—	Relative standard deviation of final control, $\mathrm{Std}(U) = u_u u$
	Response	Y	Varies	Bayesian response, $Y = \mathrm{E}\left(\sum_j Y_j^a \mathbb{I}_{\{U \geq U_j\}} \mid Y_j^a, U_j\right)$
	Response step location	U_j, u_j	—	Bayesian expected step location, $\mathrm{E}(U_j) = u_j$
	Response step height	Y_j^a, y_j^a	Varies	Bayesian expected step height, $\mathrm{E}(Y_j^a) = y_j^a$
	Step location standard deviation	σ_j	—	Bayesian relative standard deviation of U_j, $\mathrm{Std}(U_j) = \sigma_j u_j$
	Win rate	W_j	—	Bayesian impression win rate
	Win rate sensitivity	dW_j/du	—	Bayesian impression win rate sensitivity
	Response sensitivity	dY/du	—	Bayesian response sensitivity, also known as plant gain
Feedback control	Time point	t	—	Time point index, $t = 1, 2, \ldots$
	Sampling time	Δ	h	Time difference between two consecutive time points
	Time steps per day	T	—	Integer satisfying $T\Delta = 24$ h
	Setpoint value	$\bar{u}_c(t)$	US$	Budget per sample period
	Adjusted setpoint value	$u_c(t)$	US$	Feedforward adjusted budget per sample period
	Servo control signal	$u_0(t)$	—	Pre-excitation control signal
	Excitation signal	$w_u(t)$	—	Artificially generated white noise perturbation
	Excitation standard deviation	σ_u	—	Standard deviation of excitation signal
	Uncertainty signal	$u_u(t)$	—	Relative standard deviation used to perturb each bid
	Control signal	$u(t)$	—	Multiplicative bid adjustment
	Realized ad spend	$\acute{y}(t)$	US$	Ad spend attributed to $u(t)$
	Observed ad spend	$y(t)$	US$	Observed ad spend in time interval t
	Tracking error	$e(t)$	US$	Error signal, $e(t) = u_c(t) - y(t)$
	Response curve	$g(u)$	US$	Preseasonality adjusted expected response, $\acute{y}(t) \propto g(u)$
	Seasonality function	$h(t)$	—	Time-of-day periodic function, $\acute{y}(t) \propto 1 + h(t)$
	Process noise	$w_m(t)$	—	Multiplicative stochastic noise, $\acute{y}(t) \propto 1 + w_m(t)$
	Noise standard deviation	σ_m	—	Standard deviation of process noise

(*Continued*)

TABLE 1 **The important parameters, variables, and functions used in the article (*Continued*).**

	Name	Symbol	Unit	Description
Feedback control	Plant offset	a_0	US$	Intercept of linearized response curve, $g(u) \approx a_0 + a_1 u$
	Plant gain	a_1	US$	Slope of linearized response curve, $g(u) \approx a_0 + a_1 u$
	Plant latency	a_2	—	Latency $a_2 = e^{-\Delta/T_P}$ with plant time constant T_P
	Excitation latency	a_3	—	Latency $a_3 = e^{-\Delta/T_{PE}}$ with excitation time constant T_{PE}
	Integral gain	c_I	US^{-1}	Gain of integral-error controller
	State variables	x_1, x_2, x_3, x_4	US$, —, —, —	Plant state x_1 and controller states x_2, x_3, x_4

for all $i \in \Omega$ and where $u \geq 0$ is the largest value for which $\mathrm{EC} \leq \xi$. If there is no finite value of u such that $\mathrm{EC} = \xi$, then $b_i^{\mathrm{opt}} \to \infty$.

Proof

The Lagrangian of (3) and (4) is

$$\mathcal{L} = \mathrm{E}V - \lambda(\mathrm{EC} - \xi), \tag{7}$$

and the Lagrangian sufficiency theorem [19] states that, if there exists b_i, for all $i \in \Omega$ and $\lambda \geq 0$, such that the b_is maximize $\mathcal{L}$, $\mathrm{EC} \leq \xi$ and $\lambda(\mathrm{EC} - \xi) = 0$, then these values of b_i solve the original optimization problem.

Assume λ is known and compute the optimal values of b_i. Thereafter, use λ as a knob adjusted by a feedback control system until the constraint is satisfied.

Combine (1) and (5) to obtain the expected total value as

$$\begin{aligned}\mathrm{E}V &= \mathrm{E}\Big(\sum_{i\in\Omega} v_i \mathbb{I}_{\{b_i \geq B_i^*\}}\Big) \\ &= \sum_{i\in\Omega} v_i F_{B_i^*}(b_i).\end{aligned}$$

Similarly, express the expected cost using (2) and (6) as

$$\begin{aligned}\mathrm{EC} &= \mathrm{E}\Big(\sum_{i\in\Omega} B_i^* \mathbb{I}_{\{b_i \geq B_i^*\}}\Big) \\ &= \sum_{i\in\Omega} \Big(b_i F_{B_i^*}(b_i) - \int_0^{b_i} F_{B_i^*}(z)\,dz\Big).\end{aligned}$$

Substitute the derived expressions for $\mathrm{E}V$ and EC into (7) and rearrange the terms as

$$\begin{aligned}\mathcal{L} &= \mathrm{E}V - \lambda\mathrm{EC} + \lambda\xi \\ &= \sum_{i\in\Omega} \Big(v_i F_{B_i^*}(b_i) - \lambda\Big(b_i F_{B_i^*}(b_i) - \int_0^{b_i} F_{B_i^*}(z)\,dz\Big)\Big) + \lambda\xi.\end{aligned} \tag{8}$$

The solution is given by the b_is that maximize $\mathcal{L}$. Consider the following two cases:

- If $\lambda = 0$, then $\mathcal{L} = \Sigma_{i\in\Omega} v_i F_{B_i^*}(b_i)$. However, $v_i \geq 0$ while $F_{B_i^*}(b_i)$ is nonnegative and nondecreasing. Hence, $b_i^{\mathrm{opt}} = \mathrm{argmax}_{b_i}\mathcal{L} \to \infty$, which also means the budget constraint is not violated for any bid price.
- If $\lambda > 0$, then a simple rearrangement of (8) yields

$$\mathcal{L} = \lambda \sum_{i\in\Omega} \Big(\Big(\frac{v_i}{\lambda} - b_i\Big)F_{B_i^*}(b_i) + \int_0^{b_i} F_{B_i^*}(z)\,dz\Big) + \lambda\xi. \tag{9}$$

For a given λ, the bid price optimization for different impression opportunities is independent, and the optimal bid for impression i must satisfy

$$b_i^{\mathrm{opt}} = \operatorname*{argmax}_{b_i} \mathcal{L}_i,$$

where

$$\mathcal{L}_i = \Big(\frac{v_i}{\lambda} - b_i\Big)F_{B_i^*}(b_i) + \int_0^{b_i} F_{B_i^*}(z)\,dz.$$

It is straightforward to show that

$$\frac{\partial \mathcal{L}_i}{\partial b_i} = \Big(\frac{v_i}{\lambda} - b_i\Big) f_{B_i^*}(b_i).$$

Since $f_{B_i^*}(b_i) \geq 0$, it follows that

$$\frac{\partial \mathcal{L}_i}{\partial b_i} \begin{cases} \geq 0 & \text{if } b_i \leq v_i/\lambda, \\ \leq 0 & \text{if } b_i > v_i/\lambda. \end{cases}$$

As a result, $\mathcal{L}_i$ is maximized at the optimal bid price

$$b_i^{\mathrm{opt}} = \frac{v_i}{\lambda} \quad \text{if } \lambda > 0.$$

Optimality requires that $\mathrm{EC} \leq \xi$ and $\lambda(\mathrm{EC} - \xi) = 0$, where $\lambda \geq 0$. It is easily shown that $\mathrm{E}V$ and EC are both nonincreasing functions of λ. Therefore, optimal bidding corresponds to the smallest nonnegative value of λ for which $\mathrm{EC} \leq \xi$ and $\lambda(\mathrm{EC} - \xi) = 0$. If $\lambda > 0$, then the optimal bid price can be expressed as $b_i^{\mathrm{opt}} = uv_i$, where the control

signal $u \geq 0$ is the largest value for which $\mathrm{EC} \leq \xi$. If there is no finite value of u such that $\mathrm{EC} = \xi$, then $b_i^{\mathrm{opt}} \to \infty$. This completes the proof. ■

Theorem 1 shows that only v_i and u are needed to determine the optimal bid. Although they are unknown a priori and must be estimated, they can be computed separately. In an implementation, v_i is replaced by an estimate $\hat{v}_i$ produced by an impression valuation system, which performs statistical inference based on historical data and large-scale machine learning algorithms. Control signal u is computed by a campaign control system as the largest value for which (4) is not violated. The bid is then computed as the product between u and $\hat{v}_i$ according to

$$b_i = u\hat{v}_i, \tag{10}$$

as shown in Figure 2. Impression valuation consumes granular impression engagement data but does not depend on the campaign constraints. Campaign control consumes only campaign-level information, such as the budget constraint and observed campaign-level spending. The controller increases u if the campaign is trending toward $\mathrm{EC} < \xi$, decreases u if it is trending toward $\mathrm{EC} > \xi$, and keeps u unchanged if it is trending toward $\mathrm{EC} = \xi$.

In practice, an advertiser imposes an upper bound on the bid price. A natural choice is to never bid more than the impression value, which, in an implementation, corresponds to a bound $b_i \leq \hat{v}_i$ or $u \leq 1$. However, $\hat{v}_i$ is an unreliable estimate of v_i, and an advertiser often prioritizes full delivery of the budget and specifies a significantly higher maximum permitted bid $b_i^{\max}$ that must not be exceeded.

Computing $\hat{v}_i$ is outside the scope of this article, and the remainder explores the control problem. See "Optimal Bidding Intuitively" for an intuitive explanation of the optimal bidding mechanism or the appendix of [4] for a different derivation of the optimal mechanism.

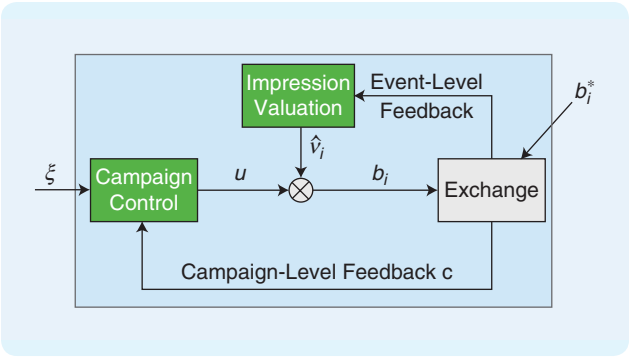


FIGURE 2 The optimal bidding mechanism. The block diagram illustrates the modularized nature of the bid calculation. Impression valuation consumes granular impression engagement data and produces an impression value estimate. Campaign control consumes the budget constraint and observed campaign-level spend to produce a control signal. The bid price is calculated as the product of the impression value estimate and control signal.

FEEDBACK CONTROL PROBLEM FORMULATION

The control objective is to compute the optimal u in (10) based on feedback as a time-varying signal $u(t) \in \mathbb{R}_{\geq 0}$. Consider a time-sampled implementation of the control system with equidistant time points indexed $t = 1, 2, \ldots$, where the sampling time Δ is, for example, 5 min.

The optimal bidding for the original problem (3) and (4) is associated with one specific value of the Lagrangian multiplier λ (independent of i). Hence, optimality corresponds to a value of u that is constant for all impression opportunities and all time points. However, although not captured by problem statements (3) and (4), an advertiser also desires an even delivery of the budget, which suggests the optimal $u(t)$ may vary over time. The even delivery requirement is handled by mapping the total campaign budget ξ into a setpoint value $\bar{u}_c(t) \in \mathbb{R}_{\geq 0}$, which is the budget per sample period. For example, $\bar{u}_c(t)$ may equal ξ divided by the total number of sample periods in the campaign flight, which, most often, is one month long. The setpoint value, together with an observed ad spend per sample period $y(t)$, is used to adjust $u(t)$. Together with an impression value estimate $\hat{v}_i$ for impression opportunity i, the final bid is computed as $b_i(t) = u(t)\hat{v}_i$, in accordance with (10).

The advertiser demands a certain daily delivery and value maximization but does not have a preference about the intraday budget delivery. In light of this, the argument that $u(t)$ should be constant holds without caveats within each day. It falls upon the control system to distribute each daily budget across the day by adjusting $\bar{u}_c(t)$ so that the error feedback system produces a control signal $u(t)$ that converges to a constant value while delivering the daily budget.

An initially too-high value of control signal $u(t)$ results in a premature delivery of the daily budget. The high value implies that some unnecessary expensive impressions are acquired. At some point during the day, it becomes necessary to reduce the control signal to avoid a spend in excess of the budget. When this reduction takes place, an opportunity cost occurs from high-yield impressions that are not purchased. Similarly, if $u(t)$ at the beginning of the day is too small, then there comes a time before the end of the day when $u(t)$ must increase significantly to catch up and deliver the daily budget. In this phase, unnecessarily expensive impressions are purchased.

If different days have the same distributions of $\{v_i, B_i^{\cdot} \mid i \in \Omega\}$, the same number of impression opportunities, and the same daily budget, then the optimization problems for the different days are identical and translate to the same optimal and constant

control signal $u(t)$. In reality, days are similar but not identical. It is, therefore, reasonable to anticipate that the optimal value of $u(t)$ is approximately (but not precisely) the same for different days. The best computed control signal for a previous day can normally be used as a good initial control signal for the next day. The main goal is to design a servo controller for daily spend with a control signal having a short transient, little overshoot, and small steady-state oscillations.

Optimal Bidding Intuitively

The (optimization) objective is to spend an advertiser's online marketing budget in such a way that the total advertisement value, or *return on investment* (*ROI*), is maximized. In the special case of a second-price cost model and a single campaign-level budget constraint, the optimal bidding mechanism can be constructed convincingly using a simple geometric construction. This provides intuition and a better appreciation of the problem at hand.

A monetary budget is provided by the advertiser, which is typically a fixed daily budget for the duration of a campaign flight. An optimization metric related to performance or branding is also given. Click or conversion counts are examples of performance metrics, while impression count is an example of a branding value metric. The advertiser may associate a different value to each click or conversion and a different branding value to each impression based on user characteristics and site properties. The total value of an impression i is denoted v_i. The impression value encodes both branding and performance value. For example,

$$v_i = v_{B,i} + p_{CTR,i} v_{C,i},$$

where $v_{B,i}$ is the branding value of serving an ad to a user, $p_{CTR,i}$ is the probability the user clicks on the ad, and $v_{C,i}$ is the advertiser-defined value of a click generated by the user. The total advertisement value v of an ad campaign is the cumulative value of all awarded impressions, and the total cost c is the cumulative cost for these impressions. The second-price cost model implies that

$$v = \sum_i v_i \mathbb{I}_{\{b_i \geq b_i^*\}},$$
$$c = \sum_i b_i^* \mathbb{I}_{\{b_i \geq b_i^*\}}.$$

Each impression on which to bid is associated with a value v_i and a cost b_i^* and, hence, can be mapped to a coordinate in a value-versus-cost plot, as shown in Figure S1. Impressions in the upper-left corner with high value and low cost correspond to the highest-ROI impressions since they translate to the largest value per ad dollar spent. 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 ROIs than impressions in the orange-shaded region.

By bidding along the straight line, all impressions in the green region are awarded, and the optimization problem is effectively decoupled into an impression value computation problem and a control problem. The optimal bidding strategy takes the form

$$b_i = u v_i,$$

and the optimal u is given by the largest fixed value for which the budget constraint is not violated. This bidding strategy is consistent with the solution derived in the "Optimal Bidding Mechanism" section. Note that to compute the optimal bid, it is necessary to know the impression value v_i (or an estimate thereof) but not the highest competing bid price b_i^*. By observing the aggregate cost of awarded impressions, the control signal may be adjusted toward its optimal value, with no need to observe or predict the cost of individual impressions.

The impression 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 impressions to bid on. A feedback controller independently adjusts u dynamically based on the delivered budget. The focus in this article is on the control problem, and v_i is assumed known. An important property using the bidding mechanism $b_i = u v_i$ is that $v(u)$, $c(u)$, and $v(u)/c(u)$ are monotonic functions of u. This guarantees the optimization problem is convex with u as an optimization variable.

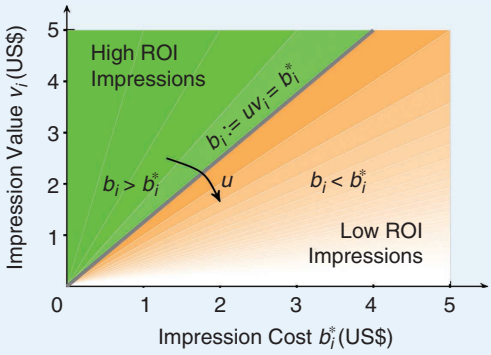


FIGURE S1 The impression value versus cost plot and optimal bidding strategy. The plot shows the relative *return on investment* (ROI) of different impression opportunities, where each impression is associated with an impression value v_i and a cost b_i^*. High ROI impressions are located in the upper-left corner, and low ROI impressions are located in the bottom-right corner. Impressions in the green-shaded region are awarded, and impressions in the orange-shaded region are not. (From [12].)

PLANT MODELING AND SMOOTHING

Modeling

The plant is defined by the map from control signal $u(t)$ and process noise $w_m(t)$ to observed spend $y(t)$. It can be represented by a concatenation of an actuator and a sensor process.

In the actuator, $u(t)$ is combined with impression value estimates to produce bid prices that participate in the bidding for impressions. Influenced by $w_m(t)$, the process output is a realized (but not yet observed) ad spend $\acute{y}(t) \in \mathbb{R}_{\geq 0}$. Actuator dynamics are insignificant, and the map from $u(t)$ to $\acute{y}(t)$ is, therefore, modeled as static. However, the map exhibits three key challenges.

First, the auction-based allocation mechanism of impressions implies that the map is discontinuous. Denote the relationship $\acute{y}(t) \propto g(u(t))$, where $\propto$ represents *proportional to*, and $g(u)$ is a nonnegative and nondecreasing function. It is described by graphs similar to what is shown in Figure 3, which depicts the relationship for two real campaigns. In Figure 3(a), the steps in the staircase are so small that the curve is effectively smooth, while in (b), the steps are large and cannot be ignored during the control design. As a consequence of the discontinuity, there exists no fixed-point solution for most setpoint values $\bar{u}_c(t)$.

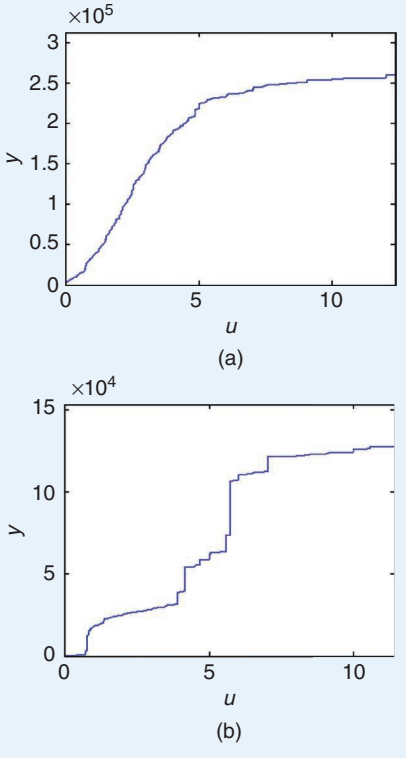


FIGURE 3 Two representative response curves. The graphs illustrate the characteristic staircase shape of spend y versus control signal u. In (a), the staircase steps are so small that the curve is effectively smooth; in (b), the steps are significant and cannot be ignored during the control design. (From [12].)

The next challenge is related to the natural time-of-day pattern in Internet users' online presence and (to a lesser extent) to (imperfect) periodic and otherwise time-varying bidding strategies among competing bidders. This leads to a dramatic daily seasonality in the available number of impression opportunities and the expected number of impressions that are awarded at each value of $u(t)$. Denote the relationship $\acute{y}(t) \propto 1 + h(t)$, where the seasonality function $h(t)$ is T periodic and satisfies $h(t) > -1$. The impression rate per hour for a fixed control signal is approximately 10 times higher in the middle of the day than it is late at night, which implies that $\max_t h(t)$ is approximately 10 times larger than $\min_t h(t)$.

The third challenge concerns the random behavior of Internet traffic, which is approximately independent and scale-invariant (Poisson-like) stochastic noise. The relationship is described by $\acute{y}(t) \propto 1 + w_m(t)$, where process noise $w_m(t)$ is modeled by a zero-mean *white noise* (WN) process $w_m(t) \sim \mathrm{WN}(0, \sigma_m^2)$, and σ_m is the standard deviation of the noise. As such, the distribution of the number of impression opportunities is skewed with a variance that is larger near the daily peak in traffic than the daily low.

Combined, the actuator model is

$$\acute{y}(t) = (1 + h(t))g(u(t))(1 + w_m(t)),$$

where it is assumed that $\Sigma_{i=1}^{T} h(i) = 0$ without a loss of generality. The model factorizes the seasonality and the control signal-dependent function by considering the dominant effects of the plant. A more accurate model would allow for the seasonality $h(\cdot)$ and the response curve $g(\cdot)$ to depend explicitly on both t and u and w_m to be colored noise with time-varying standard deviation. However, this leads to a more challenging system-identification problem, with few benefits in practice.

In the sensor, the spend rate $y(t)$ is generated from the realized ad spend $\acute{y}(t)$ and fed to the control system. More precisely, awarded impressions are identified across a fleet of geographically noncolocated ad servers. The aggregate cost of the impressions is computed, and designated data feeds connected to the control system are populated. The process represents the lag introduced by the engineering platform. A well-designed and properly dimensioned engineering infrastructure can reduce the lag but not remove it completely. The dynamics are approximately linear time invariant and are modeled using a first-order linear filter operating on $\acute{y}(t)$ according to

$$y(t+1) = a_2 y(t) + (1 - a_2)\acute{y}(t),$$

where plant latency parameter a_2 satisfies $0 \leq a_2 < 1$. Note, the static gain of the sensor equals one, which means

that all realized ad spend is eventually observed. There is no leakage and no falsely reported ad spend in the reporting system.

It is convenient to characterize the latency in terms of a plant time constant T_P rather than a_2, which are related according to $a_2 = e^{-\Delta/T_P}$. Parameterizing the model using T_P rather than a_2 makes the plant model approximately sampling time independent.

Smoothing

A plant without an almost continuous response curve relationship leads to extremely complex closed-loop dynamics. Fortunately, Heisenberg bidding [4], [5] can be used to turn a discontinuous plant in a bidding system effectively continuous. It is a bid-randomization technique by which each computed *nominal bid* $b_i = u\hat{v}_i$ is perturbed randomly before it is submitted to the auction exchange. It can be implemented with other probability distributions. However, the gamma distribution (see "Useful Probability Distributions") is a particularly good choice since it has support for all positive values of b_i and possesses other useful properties as a member of the family of exponential distributions. See "Useful Probability Distributions" for basic facts about the gamma distribution and "Heisenberg Bidding" for other use cases of Heisenberg bidding.

The perturbed bid for impression opportunity i used in the actual bidding is a realization of a random variable B_i defined by

$$B_i \sim \text{Gamma}\left(\frac{1}{b_{u,i}^2}, \frac{1}{b_i b_{u,i}^2}\right) \quad \text{if} \quad b_i, b_{u,i} > 0, \tag{11}$$

and $B_i = b_i$ otherwise. The gamma distribution is conveniently parameterized by nominal bid b_i and *bid uncertainty* $b_{u,i}$. In terms of shape parameter α and inverse scale parameter β, the distribution is defined by $\alpha = 1/b_{u,i}^2$ and $\beta = 1/(b_i b_{u,i}^2)$. The expected value and variance of B_i are given by $\text{E}(B_i) = b_i$ and $\text{Var}(B_i) = b_i^2 b_{u,i}^2$. Furthermore, the relative standard deviation $\text{Std}(B_i)/\text{E}(B_i) = b_{u,i}$, implies that $b_{u,i}$ is a scale parameter of the bid distribution.

In a general application of bid randomization, a different value of $b_{u,i}$ can be used for each impression opportunity. The value may, for example, depend on the confidence in $\hat{v}_i$ as an estimate of v_i. This provides a lever that supports exploration, which is the process of exposing a wide range of different users to an ad to learn how effective the ad is. Note, unless some impressions are awarded, it is impossible to improve the accuracy of the impression value estimate of similar impressions. If $b_{u,i}$ is large, then it is likely that at least some realized bids are large enough to win impressions. Thompson sampling [20] is based on this idea, and it is also the idea behind Heisenberg bidding-based exploration and exploitation [4].

A less granular application of Heisenberg bidding is to use the same bid uncertainty $b_{u,i} = u_u$ for all impression opportunities. Bid uncertainty is, then, a campaign-level signal, and the randomization can be interpreted as a perturbation of u (for each individual bid) during analysis. Because bid uncertainty is a scale parameter of the gamma bid distribution, this reinterpretation is particularly straightforward. Indeed, if X is a gamma random variable defined by $X \sim \text{Gamma}(\alpha, \beta)$, then, for any $k > 0$, it holds true that $kX \sim \text{Gamma}(\alpha, \beta/k)$ [21]. Hence, whenever

$$U \sim \text{Gamma}\left(\frac{1}{u_u^2}, \frac{1}{u u_u^2}\right) \quad \text{if} \quad u, u_u > 0, \tag{12}$$

and $U = u$ otherwise, then

$$B_i := U\hat{v}_i \sim \text{Gamma}\left(\frac{1}{u_u^2}, \frac{1}{u\hat{v}_i u_u^2}\right) \quad \text{if} \quad u, \hat{v}_i, u_u > 0,$$

and $B_i = u\hat{v}_i \equiv b_i$. The distribution of the final perturbed bid is the same whether the nominal control signal u is perturbed first and then multiplied by v_i or the nominal bid price $b_i = uv_i$ is perturbed after the nominal control signal u has already been multiplied by v_i.

By interpreting u and u_u as two campaign-level control signals, it is insightful to study the impact on the relationship between u and y for different values of u_u. The mapping from u to y is referred to as the (primary) input–output relationship of the plant, which can be made arbitrarily smooth by adjusting u_u. Figure 4 illustrates this capability by adding the dimension of uncertainty signal u_u to the response curve in Figure 3(b). The campaign-level bid uncertainty can be used as a secondary campaign-level control signal that is updated based on feedback to achieve a desired relationship between u and y. This was attempted in [6], where the control objective was to regulate the effective loop gain.

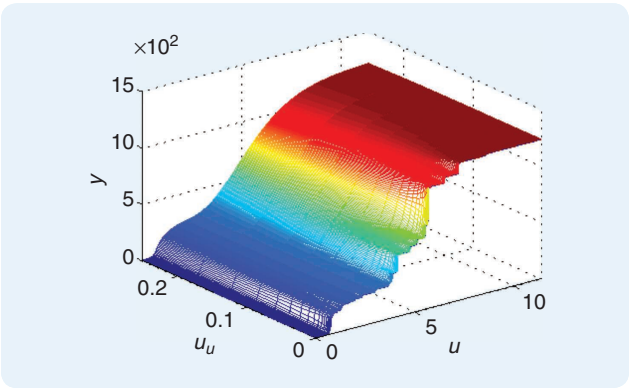


FIGURE 4 A plant smoothing surface. The plot shows how the spend rate y of the campaign in Figure 3(b) depends on the nominal control signal u and uncertainty signal u_u. Note how the relationship between y and u is discontinuous only when u_u equals zero. (From [12].)

Alternatively, a constant value of u_u can be used, which is treated as a configuration parameter. This *passive* plant smoothing makes Heisenberg bidding particularly safe since it avoids adding hard-to-analyze dynamics to the closed-loop system. It is also the choice made in this article. The uncertainty signal is simply assumed to have a fixed nonzero value to guarantee a smooth response curve. The challenges and benefits of adjusting u_u over time are outside the scope of this article.

Once the relationship between $u(t)$ and $\dot{y}(t)$ is made smooth in the expected sense, it is linearizable in a small

Useful Probability Distributions

BINOMIAL DISTRIBUTION

The binomial distribution with parameters n and p is a discrete probability distribution. If the random variable X follows the binomial distribution, we write $X \sim \text{Binomial}(n, p)$. The probability mass function of x is

$$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x} \tag{S1}$$

for $x = 0, 1, \ldots, n$. Parameters $n \in \{1, 2, \ldots\}$ and $p \in [0, 1]$ are referred to as the *number of trials* and *success probability in each trial*, respectively. The expected value of X is $\mathrm{E}(X) = np$, while the variance is $\mathrm{Var}(X) = np(1-p)$.

BETA DISTRIBUTION

The beta distribution with parameters α and β is a continuous probability distribution. If the random variable X follows the beta distribution, we write $X \sim \text{Beta}(\alpha, \beta)$. The probability density function of x is

$$f_X(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)} \tag{S2}$$

for $0 < x < 1$, where $B(\alpha, \beta)$ is the beta function (also called the Euler integral) defined by $B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx$. Parameters $\alpha > 0$ and $\beta > 0$ are referred to as shape parameters. The expected value and variance of X are $\mathrm{E}(X) = \alpha/(\alpha+\beta)$ and $\mathrm{Var}(X) = \alpha\beta/[(\alpha+\beta)^2(\alpha+\beta+1)]$.

GAMMA DISTRIBUTION

In consideration of its role in Heisenberg bidding, the gamma distribution deserves a longer introduction. It is a two-parameter family of continuous probability distributions with support on $[0, \infty)$. It is a versatile probability distribution with many applications.

It has been used in queuing models, climatology, and actuarial sciences. Examples of events that have been modeled by gamma distributions include rainfall, the size of insurance claims, load and data latencies on web servers, and signal power fading in wireless communication. The exponential distribution, Erlang distribution, and chi-squared distribution are all special cases of the gamma distribution.

Different parameterizations are in common use. However, in Bayesian statistics, it appears most common to define a gamma distribution in terms of *shape* parameter $\alpha > 0$ and *inverse scale* parameter $\beta > 0$ (also called the rate parameter). If the random variable X follows the gamma distribution, we write $X \sim \text{Gamma}(\alpha, \beta)$. The probability density function of x is

$$f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \tag{S3}$$

for $x > 0$, where $\Gamma(\alpha)$ is the gamma function defined by $\Gamma(\alpha) = \int_0^\infty e^{-t} t^{\alpha-1} dt$. The expected value and variance of X are

$$\mathrm{E}(X) = \frac{\alpha}{\beta},$$
$$\mathrm{Var}(X) = \frac{\alpha}{\beta^2}.$$

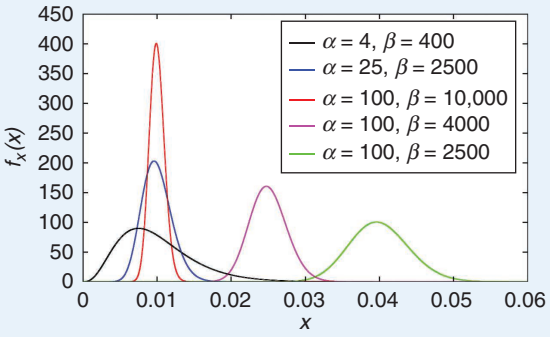


FIGURE S2 Gamma probability density functions. The figure shows five illustrative gamma probability density functions, parameterized by the shape α and inverse scale β parameters.

In Bayesian inference, the gamma distribution is the conjugate prior to many likelihood distributions, for example, the Poisson, exponential, normal (with known mean), Pareto, gamma with known shape α, inverse gamma with known shape parameter, and Gompertz with known scale parameter.

Figure S2 shows five examples of gamma probability density functions parameterized by the shape and inverse scale parameters. Note that the distributions with $(\alpha, \beta) = (4, 400), (\alpha, \beta) = (25, 2500)$, and $(\alpha, \beta) = (100, 10{,}000)$ have the same expected value but different variance.

As a member of the exponential family of distributions, the gamma distribution is closely related to familiar distributions, such as Bernoulli, Poisson, normal, and Dirichlet. As a consequence, multiple conversions are possible among these distributions.

neighborhood of each u. The static relationship locally can, then, be expressed as

$$\mathrm{E}(\dot{y}(t)) = (1 + h(t))(a_0 + a_1 u(t)) \tag{13}$$

for some values of *plant offset* a_0 and *plant gain* $a_1 > 0$ and for some T-periodic function $h(t)$ satisfying $h(t+T) = h(t)$, $\Sigma_{i=1}^{T} h(i) = 0$, and $h(t) > -1$ for all t. Some useful mathematical results to help in the analysis are provided in "The Mathematics of Plant Smoothing."

An Effective Plant Model

Passive plant smoothing can be implemented inside the actuator out of sight of the feedback control system that computes the nominal control signal. The feedback controller only needs to know that the plant model is linearizable and (13) is approximately true in a local neighborhood of the operating point. From the vantage point of the control engineer, the linearized plant model is

$$y(t+1) = a_2 y(t) + (1 - a_2)(1 + h(t))(a_0 + a_1 u(t))(1 + w_m(t)), \tag{14}$$

where $w_m(t) \sim \mathrm{WN}(0, \sigma_m^2)$. Plant parameters a_0 and a_1 are control signal dependent and, in practice, also time varying. However, the described plant smoothing ensures they depend on the control signal in a smooth manner. As long as the control signal is not excessively volatile and the linearization does not change rapidly over time for each value of u, (14) can be used as the basis for an adaptive control system, where a_0 and a_1 are estimated online.

Latency parameter a_2 is approximately constant, and seasonality function $h(t)$ is approximately time invariant. They can be estimated offline in advance or refined online [13]. In this article, they are assumed known.

CONTROL DESIGN AND CLOSED-LOOP DYNAMICS

In preparation for a closed-loop analysis of the system dynamics, plant model (14) can be expressed in state-space form as the following first-order stochastic difference equations

$$x_1(t+1) = a_2 x_1(t) + (1 - a_2)(1 + h(t))(a_0 + a_1 u(t))(1 + w_m(t)), \tag{15}$$

$$y(t) = x_1(t), \tag{16}$$

where $w_m(t) \sim \mathrm{WN}(0, \sigma_m^2)$. The model is time periodic and subject to multiplicative process noise. Moreover, it is locally linear in the state and control signal, thanks to the previously described plant smoothing. The plant is defined by $h(t)$, σ_m, a_0, a_1, and a_2. Function $h(t)$ and parameter a_2 are assumed known, while parameters σ_m, a_0, and a_1 are not.

To solve the control (optimization) problem, consider pure 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 is used to distribute a daily ad budget throughout the day according to the seasonality of the 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. It is achieved by modifying the deterministic, and often constant, setpoint input signal $\bar{u}_c(t)$ based on known values of a_2 and $h(t)$ and computing an adjusted setpoint value $u_c(t)$. The feedforward controller dynamics are described by

$$x_3(t+1) = a_2 x_3(t) + (1 - a_2)(1 + h(t))\bar{u}_c(t), \tag{17}$$

$$u_c(t) = x_3(t). \tag{18}$$

The pure I controller computes a servo control signal $u_0(t)$ based on a tracking error signal $e(t)$ to minimize or reduce the marginal and cumulative tracking error via an approximately constant $u_0(t)$, thereby maximizing the produced advertising value of the campaign. The feedback controller dynamics are

$$e(t) = u_c(t) - y(t), \tag{19}$$

$$x_2(t+1) = x_2(t) + c_I e(t), \tag{20}$$

$$u_0(t) = x_2(t) + c_I e(t), \tag{21}$$

where c_I is the *I gain* of the feedback controller. This feedback mechanism has several benefits. First, it is well known that I-error feedback control guarantees zero steady-state error under certain conditions on the setpoint signal. Furthermore, a *pure* I controller is a low-pass filter, which makes the closed-loop system less sensitive to high-frequency process noise and more robust to a challenging response curve than a system using, for example, a *PI* feedback controller. In many other applications where the plant latency is long, a pure I control may not permit satisfactorily fast closed-loop dynamics. However, the plant dynamics for the problem considered here are reasonably fast and permit a high-gain I controller. The absence of proportional and derivative control action is, therefore, not a concern. Finally, using pure I control implies only one feedback controller parameter needs to be tuned. This significantly simplifies the commissioning of the control system.

To guarantee stability, it is later shown that c_I must be chosen as a function of a_1 or, in practice, as a function of an estimate $\hat{a}_1$ thereof. For now, assume c_I is a given constant. The estimation of a_1 is discussed later. However, to make the identification of a_1 possible in light of an approximately constant $u_0(t)$, persistent excitation must be ensured. Complementing the I controller with an open-loop, persistent-excitation controller achieves this. The excitation controller computes a final control signal $u(t)$ from

Heisenberg Bidding

Imagine occasionally winning an impression that is not supposed to be won based on the computed bid price. This is like quantum tunneling in programmatic advertising. *Heisenberg bidding* was proposed in [5] as a robust and scalable solution to this figurative quantum tunneling. It is an engineered mechanism that makes an otherwise losing (winning) bid for an ad impression a winning (losing) bid with a small probability. It is achieved without coordination among the bidders, which makes it a scalable mechanism suitable in programmatic advertising with a large number of noncooperating bidders. Heisenberg bidding is a bid-randomization scheme to support enhanced control, estimation, and optimization in online advertising. It has some resemblance to Thompson sampling [20] but with a different origin and motivation.

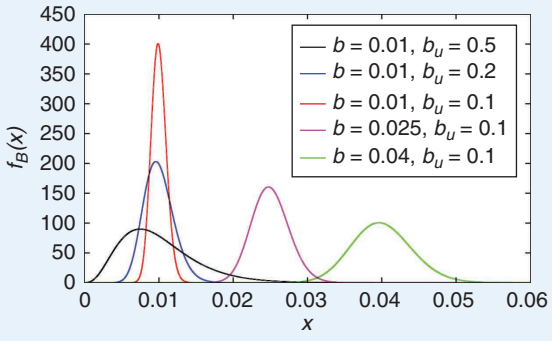


FIGURE S3 Gamma probability density functions. The figure shows five illustrative gamma probability density functions, parameterized by the nominal bid b and bid uncertainty b_u parameters.

Heisenberg bidding randomly perturbs a *nominal bid price* b according to a distribution defined by b and a *bid uncertainty* b_u. It generates a final bid price used in the market clearing. The randomization can be implemented with other probability distributions. However, the gamma distribution has a distinct advantage, as discussed in this article. See "Useful Probability Distributions" for details on the gamma distribution and "The Mathematics of Plant Smoothing" for some of the inference possible when bid randomization is based on a gamma distribution.

With an implementation based on the gamma random number generator, the perturbed bid is a realization of B, where

$$B \sim \text{Gamma}\left(\frac{1}{b_u^2}, \frac{1}{bb_u^2}\right), \quad \text{if} \quad b, b_u > 0, \tag{S4}$$

and $B = b$, otherwise. Heisenberg bidding based on the gamma distribution (and in terms of the shape parameter α and the inverse scale parameter β) is defined by $\alpha := 1/b_u^2$ and $\beta := 1/(bb_u^2)$. Consequently,

$$\text{E}(B) = \frac{1}{b_u^2} bb_u^2 = b,$$
$$\text{Var}(B) = \frac{1}{b_u^2} (bb_u^2)^2 = b^2 b_u^2.$$

Note that b_u is the unitless relative standard deviation of B, that is, $b_u \equiv \text{Std}(B)/\text{E}(B)$, where $\text{Std}(B) = \sqrt{\text{Var}(B)}$.

Figure S3 shows the same five gamma distributions that were shown in Figure S2 but parameterized by b and b_u. The new parameters separately have a meaningful physical interpretation. The nominal bid b is a measure of the center of the bid distribution, while the bid uncertainty b_u captures the spread. A large value of b_u translates to a large relative spread, and $b_u = 0$ translates to no spread at all, in which case the final bid is identical to the nominal bid.

$$x_4(t+1) = a_3 x_4(t) + \sqrt{1 - a_3^2}\, w_u(t), \tag{22}$$
$$u(t) = u_0(t)(1 + x_4(t)), \tag{23}$$

where excitation signal $w_u(t)$ is an artificially generated $\text{WN}(0, \sigma_u^2)$ perturbation, a_3 is an excitation latency design parameter, and σ_u is the excitation standard deviation. The specific structure of the perturbation system is chosen deliberately to reduce the need for customization, as the control system is used by different campaigns. Note, large DSPs manage tens of thousands (or more) of ad campaigns simultaneously.

The design of the excitation controller implies that $\text{Var}(x_4) = \text{Var}(w_u) \equiv \sigma_u^2$ at the steady state, where σ_u is a design parameter. Since x_4 perturbs the servo control signal u_0 multiplicatively, the relative standard deviation of the control signal induced by the persistent excitation equals σ_u, regardless what the values of a_3 and u_0 are. Hence, while the operating point of $u(t)$ for one campaign may be hundreds or thousands times larger or smaller than that of another campaign, the relative standard deviation of the control signals and the final bid prices remain the same for all campaigns. A random process (as opposed to a deterministic signal) is used to eliminate the risk of unintended synchronized excitation of competing campaigns. In practice, a_3 is not set directly but, rather, computed from a manually selected *persistent-excitation time constant* T_{PE} via the relationship $a_3 = e^{-\Delta/T_{\text{PE}}}$.

The choice of a first-order, linear time-invariant, persistent-excitation controller makes it easy and intuitive to tune. Using a dynamic in persistent excitation (as opposed to only WN) improves the system identification robustness to an imperfect plant latency model. If the plant latency is accurately described by the proposed first-order, linear time-invariant model and defined by a specific value of a_2 or T_P, then it is reasonable to use WN for more excitation ($a_3 = 0$). But if the latency model is poor

A typical implementation of Heisenberg bidding perturbation is illustrated as a block diagram in Figure S4. Inputs b and b_u are usually produced by an ad optimization system involving feedback control and machine learning algorithms. These algorithms may require large amounts of input data and many computations. They are, therefore, typically implemented as a discrete-time system with updates every few minutes or 1 h. Producing a gamma random number based on b and b_u, on the other hand, is very efficient and can be done (in runtime) for each bid calculation with, at most, a small impact on ad serving latency. If necessary, normalized perturbations can be precomputed to make the latency impact insignificant.

An important property of Heisenberg bidding is how it can turn an otherwise discontinuous system continuous without adding dynamics: a bidder may participate in millions of auctions every few minutes. Sometimes, the highest competing bid b_i^* for a large number of impression opportunities i is approximately proportional to the campaign's impression value estimate $\hat{v}_i$. Given how the control signal, according to (10), is a multiplicative factor (and because of how the market clearing mechanism awards an impression), this leads to a situation where the number of awarded impressions (or the spend) as a function of a nonrandomized control signal u has a staircase shape, possibly with large steps. As a result, for most values of a desired response, there exists no fixed-point solution, as illustrated in Figure 3(b).

Heisenberg bidding removes steps in the plant response curve in the expected sense by randomizing the control signal for each individual impression opportunity. It creates guarantees for the existence, in the expected sense, of a fixed-point solution of a dynamic system without necessarily introducing new dynamics. Furthermore, it makes it possible to gracefully detect if a cluster of highest competing bid prices is moving up or down, since this would be reflected in an increasing or decreasing win rate.

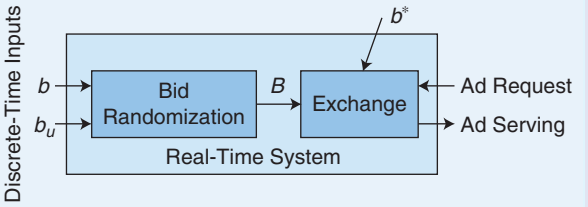


FIGURE S4 A block diagram of a typical implementation of Heisenberg bidding in an online advertising system.

Heisenberg bidding can also be used to regulate the effective plant gain and support improved controllability and robustness of a feedback system [6], [14]. Moreover, it can be used as a powerful means of adaptive and low-cost exploration, which is the process of exposing a wide range of users to an advertisement to learn how effective it is [4]. This makes it possible to simultaneously and safely test many ads without cumbersome and limiting segmentation and without relying on conventional restrictive A/B testing solutions.

Last, but not least, Heisenberg bidding enables a continuous-system analysis of the dynamics. Indeed, analyzing the dynamics, stability, and robustness of an uncertain, time-varying, discontinuous system is extremely challenging. If results are derived at all for a discontinuous system, they often lack practical value. Turning the system continuous first not only makes it possible to leverage the large number of tools developed for continuous systems, but the analytical results are also typically more useful in practice.

and $u(t)$ contains a significant high-frequency component (for example, due to excitation with a small value of a_3), then plant model (14) offers little predictive power, which makes the regression of a_0 and a_1 particularly difficult.

An alternative persistent-excitation method is to implement a *pseudorandom binary sequence* (*PRBS*) to excite $u_0(t)$ [22]. This approach is a good choice if the plant is truly linear in the range of the PRBS excitation and the latency model described by a_2 is accurate. In reality, the response curve may deviate from linearity, whereby the linearization computed using the PRBS potentially differs significantly from the linearization at $u_0(t)$. Moreover, the real latency may differ substantially from the model. Binary excitation is, therefore, likely to require simultaneous identification of the plant gain and latency model, which makes an already challenging system-identification problem more challenging.

The plant is represented by state x_1 and the controller by states x_2, x_3, and x_4. Ultimately, the performance of the closed-loop system is judged 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 translates to an even daily budget delivery and optimal value creation for the advertiser. The controller defined by (17)–(23) is summarized in Algorithm 1.

In a real implementation, it is important to add integrator windup protection to the I controller [23] and restrict the control signal between zero and some upper bound $u_{\max}$ set by the advertiser or estimated programmatically. However, to reduce the clutter, Algorithm 1 does not incorporate these aspects of the control system. Estimating $u_{\max}$ is a research project in itself. It may be a threshold value of u indicating where $g(u)$ has approximately reached its maxima. Since function $g(\cdot)$ is observed and approximately known only near the current operating point, it is not trivial to estimate the threshold value.

The closed-loop dynamics of the interconnected plant and controller system remain to be analyzed. In the rest of

The Mathematics of Plant Smoothing

MOTIVATION AND SETUP

It is intuitively obvious that Heisenberg bidding smoothens the response curve of a plant and improves the conditions for feedback control of an otherwise discontinuous process. However, it is useful to develop a mathematical foundation of Heisenberg bidding-based smoothing. This creates additional avenues of enhanced inference and control. The content discussed here is best appreciated if basic facts about relevant statistical distributions are well understood. Key facts are provided in "Useful Probability Distributions," and a review of that is recommended before proceeding. Some of the results here, plus a few additional results, are available in [14].

For each impression opportunity, the actual used control signal is generated via Heisenberg bidding-based randomization; that is, it is an independent sample from a probability density function defined by

$$U \sim \text{Gamma}\left(\frac{1}{u_u^2}, \frac{1}{u u_u^2}\right), \quad \text{if } u, u_u > 0,$$

and $U = u$ otherwise. As indicated, the distribution is parameterized by nominal control signal u and uncertainty signal u_u. Recall that u_u is the relative standard deviation of the used control signal for each individual bid.

The response curve describes how a campaign metric, such as spend or awarded number of impressions, depends on a nominal control signal u with or without smoothing. Without smoothing ($u_u = 0$), it is, for a second-price cost model, defined by a staircase function. Each step $j = 1, 2, \ldots$ of the staircase is defined by a location U_j and a height Y_j^a, which are known only probabilistically and, therefore, represented as random variables. Note, U_j and Y_j^a are random only in the Bayesian sense, and, strictly speaking, they are unknown (possibly time-varying) parameters. Given the auction-based allocation of impressions, the response staircase is, for the case without smoothing, defined by

$$Y = \sum_j Y_j^a \mathbb{I}_{\{u \geq U_j\}}, \quad \text{if } u_u = 0,$$

in which case Y is random only in the Bayesian sense. In general, the response is defined by

$$Y = \text{E}\left(\sum_j Y_j^a \mathbb{I}_{\{U \geq U_j\}} \,\middle|\, Y_j^a, U_j\right), \quad \text{if } u_u \geq 0, \tag{S5}$$

where, again, Y is random only in the Bayesian sense. Since U, by design, is an independent random variable, it follows that

$$Y = \sum_j Y_j^a \text{E}(\mathbb{I}_{\{U \geq U_j\}} \,|\, U_j), \quad \text{if } u_u \geq 0. \tag{S6}$$

However, $\text{E}(\mathbb{I}_{\{U \geq U_j\}} \,|\, U_j) = \Pr(U \geq U_j \,|\, U_j)$, which is the probability that an auction for a staircase case step j impression is won by outbidding all other bidders. In fact, it is the success rate of a Bernoulli experiment. The success rate is referred to as the *win rate* and is denoted

$$W_j := \Pr(U \geq U_j \,|\, U_j), \tag{S7}$$

and, consequently,

$$Y = \sum_j Y_j^a W_j \quad \text{if } u_u \geq 0. \tag{S8}$$

Assume the belief of U_j is given by

$$U_j \sim \text{Gamma}\left(\frac{1}{\sigma_j^2}, \frac{1}{u_j \sigma_j^2}\right),$$

where u_j and σ_j are the mean and relative standard deviations of our belief about the location of the step. Furthermore, let the Bayesian expected value of Y_j^a (the mean of its belief function) be

$$\text{E}(Y_j^a) = y_j^a.$$

The distribution of Y_j^a and its variance are not needed to derive the results here. However, they are certainly useful to extend beyond what is covered here. See [14] for some results where, for example, the variance of Y_j^a matters.

BAYESIAN ESTIMATION

In the remainder, closed-form expressions of $\text{E}(W_j), \text{E}(dW_j/du)$, $\text{E}(Y)$, and $\text{E}(dY/du)$ are derived. The expected response $\text{E}(Y)$ is a sensible point estimate of the response and can be computed for different values of u and u_u. It can be leveraged for controller initialization. Similarly, the expected response sensitivity $\text{E}(dY/du)$ can be utilized by an error feedback controller as a point estimate of the plant gain to reduce the risk of closed-loop instability.

The scaling property of the gamma distribution states that whenever $X \sim \text{Gamma}(\alpha, \beta)$, then $\beta X \sim \text{Gamma}(\alpha, 1)$ [21]. It follows that (S7) can rewritten as

$$\begin{aligned} W_j &= \Pr\left(\frac{U}{u_u^2 u} \geq \frac{U_j}{u_u^2 u} \,\middle|\, U_j\right) \\ &= \Pr\left(\acute{U} \geq \frac{U_j}{u_u^2 u} \,\middle|\, U_j\right), \end{aligned}$$

where $\acute{U} \sim \text{Gamma}(1/u_u^2, 1)$. Since the cumulative density function $F_{\acute{U}}(\acute{u}_j) = \Pr(\acute{U} \leq \acute{u}_j)$ for any $\acute{u}_j$, the win rate is

$$W_j = 1 - F_{\acute{U}}\left(\frac{U_j}{u_u^2 u} \,\middle|\, U_j\right). \tag{S9}$$

Define the *win rate sensitivity* as the derivative of W_j with respect to u. It is straightforward to show that

$$\frac{dW_j}{du} = f_{\acute{U}}\left(\frac{U_j}{u_u^2 u} \,\middle|\, U_j\right) \frac{U_j}{u_u^2 u^2}. \tag{S10}$$

To turn win rate sensitivity into a more useful format, consider the following theorem.

Theorem S1

If $X \sim \text{Gamma}(\alpha, \beta)$, then $f_X(x)x = \alpha f_{\dot{X}}(X)/\beta$, where $\dot{X} \sim \text{Gamma}(\alpha+1, \beta)$.

Proof

Replace $f_X(x)$ with the expression defining the gamma probability density function

$$f_X(x)x = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} x$$
$$= \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)\beta}\left[\frac{\beta^{\alpha+1}}{\Gamma(\alpha+1)} x^\alpha e^{-\beta x}\right].$$

The expression in $[\cdot]$ is the probability density function of a $\text{Gamma}(\alpha+1, \beta)$ random variable, and $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$ [27]. Hence,

$$f_X(x)x = \frac{\alpha}{\beta} f_{\dot{X}}(x),$$

where $\dot{X} \sim \text{Gamma}(\alpha+1, \beta)$, which completes the proof. ■

A direct application of Theorem S1 to (S10) yields

$$\frac{dW_j}{du} = \frac{1}{u_u^2 u} f_{\dot{U}}\left(\frac{U_j}{u_u^2 u}\middle| U_j\right), \tag{S11}$$

where $\dot{U} \sim \text{Gamma}(1/u_u^2 + 1, 1)$.

For a particular realization of U_j, the win rate and win rate sensitivity are denoted w_j and dw_j/du, respectively, and obtained by replacing U_j by u_j in (S9) and (S11).

The *response sensitivity* is defined as the derivative of Y with respect to u. Since Y_j^a is independent of u, it is

$$\frac{dY}{du} = \sum_j Y_j^a \frac{dW_j}{du}. \tag{S12}$$

In the context of feedback control, response sensitivity represents the plant gain at different values of the control signal value u. Knowledge of dY/du and its statistical properties can be used to select a controller gain that yields a stable and robust closed-loop system.

The proofs of the following theorems regarding the Bayesian expected values of $W_j, Y, dW_j/du$, and dY/du provide a good illustration of how it is sometimes possible to compute intimidating integrals by identifying kernels of standard random variables. Sometimes, this requires a variable transformation.

Theorem S2: Expected Win Rate

Assume a randomized bidding strategy $U \sim \text{Gamma}(1/u_u^2, 1/(u_u^2 u))$ with $u, u_u > 0$. If the highest competing bid price, u_j, is only known probabilistically as a realization of the random variable $U_j \sim \text{Gamma}(1/\sigma_j^2, 1/(\sigma_j^2 u_j))$ for known values of $\sigma_j, u_j > 0$, then the *expected win rate* is

$$\text{E}(W_j) = F_{Z_0}\left(\frac{u_u^2 u}{\sigma_j^2 u_j + u_u^2 u}\right), \tag{S13}$$

where $Z_0 \sim \text{Beta}(1/\sigma_j^2, 1/u_u^2)$.

Proof

The expected win rate is

$$\text{E}(W_j) = \int_0^\infty w_j(u_j) f_{U_j}(u_j)\, du_j,$$

where w_j is a realization of (S9), and $f_{U_j}(u_j)$ is the probability density function of U_j, which is a $\text{Gamma}(1/\sigma_j^2, 1/(\sigma_j^2 u_j))$ random variable. Substituting the expressions for w_j and $f_{U_j}(u_j)$ yields

$$\text{E}(W_j) = \int_0^\infty \left(1 - F_{\dot{U}}\left(\frac{u_j}{u_u^2 u}\right)\right) \frac{\left(\frac{1}{\sigma_j^2 u_j}\right)^{1/\sigma_j^2}}{\Gamma\left(\frac{1}{\sigma_j^2}\right)} u_j^{1/\sigma_j^2 - 1} e^{-u_j/(\sigma_j^2 u_j)}\, du_j, \tag{S14}$$

where $F_{\dot{U}}(\acute{u})$ is the cumulative density function of $\dot{U}$, which is a $\text{Gamma}(1/u_u^2, 1)$ random variable. Hence,

$$F_{\dot{U}}\left(\frac{u_j}{u_u^2 u}\right) = \int_0^{u_j/u_u^2 u} \frac{1}{\Gamma\left(\frac{1}{u_u^2}\right)} \acute{u}^{1/u_u^2 - 1} e^{-\acute{u}}\, d\acute{u}. \tag{S15}$$

Substituting (S15) into (S14) and bringing all factors independent of $\acute{u}$ and u_j outside the integration operators,

$$\text{E}(W_j) = 1 - \frac{\left(\frac{1}{\sigma_j^2 u_j}\right)^{1/\sigma_j^2}}{\Gamma\left(\frac{1}{u_u^2}\right)\Gamma\left(\frac{1}{\sigma_j^2}\right)} \int_0^\infty \int_0^{u_j/u_u^2 u} \acute{u}^{1/u_u^2 - 1} e^{-\acute{u}} u_j^{1/\sigma_j^2 - 1} e^{-u_j/(\sigma_j^2 u_j)}\, d\acute{u} du_j.$$

Next, convert the Cartesian coordinates $(u_i, \acute{u})$ to polar coordinates (r, θ) according to

$$u_j = r\cos\theta,$$
$$\acute{u} = r\sin\theta.$$

The infinitesimal area element is $d\acute{u}du_j = rdrd\theta$, and the new integration bounds are $0 \le r \le \infty$ and $0 \le \theta \le \arctan(1/u_u^2 u)$. The expected win rate can, then, be expressed as

$$\text{E}(W_j) = 1 - \frac{\left(\frac{1}{\sigma_j^2 u_j}\right)^{1/\sigma_j^2}}{\Gamma\left(\frac{1}{u_u^2}\right)\Gamma\left(\frac{1}{\sigma_j^2}\right)} \int_0^{\arctan(1/u_u^2 u)} \times \int_0^\infty (r\sin\theta)^{1/u_u^2 - 1} e^{-r\sin\theta} (r\cos\theta)^{1/\sigma_j^2 - 1} e^{-r\cos\theta/(\sigma_j^2 u_j)} r dr d\theta.$$

Move factors independent of r in front of the inner integral to obtain

$$\mathrm{E}(W_j) = 1 - \frac{\left(\frac{1}{\sigma_j^2 u_j}\right)^{1/\sigma_j^2}}{\Gamma\left(\frac{1}{u_u^2}\right)\Gamma\left(\frac{1}{\sigma_j^2}\right)} \int_0^{\arctan(1/u_u^2 u)} (\sin\theta)^{1/u_u^2 - 1} (\cos\theta)^{1/\sigma_j^2 - 1} \int_0^{\infty} r^{1/u_u^2 + 1/\sigma_j^2 - 1} e^{-r\left(\sin\theta + \cos\theta/(\sigma_j^2 u_j)\right)} dr d\theta. \quad \text{(S16)}$$

Recognize the integrand of the inner integral as the kernel of a Gamma$(1/u_u^2 + 1/\sigma_j^2, \sin\theta + \cos\theta/(\sigma_j^2 u_j))$ distribution. Hence, the integral with respect to r over the semi-infinite interval $[0, \infty)$ must satisfy

$$\int_0^{\infty} r^{1/u_u^2 + 1/\sigma_j^2 - 1} e^{-r(\sin\theta + \cos\theta/(\sigma_j^2 u_j))} dr = \frac{\Gamma\left(\frac{1}{u_u^2} + \frac{1}{\sigma_j^2}\right)}{\left(\sin\theta + \frac{\cos\theta}{\sigma_j^2 u_j}\right)^{1/u_u^2 + 1/\sigma_j^2}}. \quad \text{(S17)}$$

Combine (S16) and (S17). The result is

$$\mathrm{E}(W_j) = 1 - \frac{\Gamma\left(\frac{1}{u_u^2} + \frac{1}{\sigma_j^2}\right)\left(\frac{1}{\sigma_j^2 u_j}\right)^{1/\sigma_j^2}}{\Gamma\left(\frac{1}{u_u^2}\right)\Gamma\left(\frac{1}{\sigma_j^2}\right)} \times \int_0^{\arctan(1/u_u^2 u)} \frac{(\sin\theta)^{1/u_u^2 - 1} (\cos\theta)^{1/\sigma_j^2 - 1}}{\left(\sin\theta + \frac{\cos\theta}{\sigma_j^2 u_j}\right)^{1/u_u^2 + 1/\sigma_j^2}} d\theta.$$

Use $B(x, y) = \Gamma(x)\Gamma(y)/\Gamma(x + y)$ [27] and divide both the numerator and denominator of the integrand by $(\cos\theta/(\sigma_j^2 u_j))^{1/u_u^2 + 1/\sigma_j^2}$ to obtain

$$\mathrm{E}(W_j) = 1 - \frac{u_j \sigma_j^2}{B\left(\frac{1}{u_u^2}, \frac{1}{\sigma_j^2}\right)} \int_0^{\arctan(1/u_u^2 u)} \times \frac{(u_j \sigma_j^2 \tan\theta)^{1/u_u^2 - 1}}{(u_j \sigma_j^2 \tan\theta + 1)^{1/u_u^2 + 1/\sigma_j^2}} \frac{1}{\cos^2\theta} d\theta. \quad \text{(S18)}$$

Next, make the variable transformation

$$z_0 = \frac{1}{u_j \sigma_j^2 \tan\theta + 1}. \quad \text{(S19)}$$

The derivative of the new variable with respect to the old variable is

$$\frac{dz_0}{d\theta} = -\frac{u_j \sigma_j^2}{(u_j \sigma_j^2 \tan\theta + 1)^2 \cos^2\theta},$$

which leads to the relationship

$$\frac{d\theta}{\cos^2\theta} = -\frac{dz_0}{u_j \sigma_j^2 z_0^2}.$$

The integration bounds convert as

$$\theta = 0 \Rightarrow z_0 = 1$$

$$\theta = \arctan\left(\frac{1}{uu_u^2}\right) \Rightarrow z_0 = \frac{uu_u^2}{u_j \sigma_j^2 + uu_u^2} =: z_0^{\star},$$

where it is noted that $z_0^{\star} \leq 1$, which means the previous lower bound is now an upper bound, and vice versa. Furthermore, one minus (S19) yields

$$1 - z_0 = \frac{u_j \sigma_j^2 \tan\theta}{u_j \sigma_j^2 \tan\theta + 1}.$$

Expression (S18) can now be written in terms of variable z_0. The result is

$$\mathrm{E}(W_j) = 1 + \frac{1}{B\left(\frac{1}{u_u^2}, \frac{1}{\sigma_j^2}\right)} \int_1^{z_0^{\star}} z_0^{1/\sigma_j^2 - 1} (1 - z_0)^{1/u_u^2 - 1} dz.$$

The integrand is recognized as the kernel of a Beta $(1/\sigma_j^2, 1/u_u^2)$ distribution. Swap the integration limits and replace the second term using the cumulative density function of the recognized beta distribution. Finally, the expected win rate is simplified to be

$$\mathrm{E}(W_j) = F_{Z_0}\left(\frac{u_u^2 u}{\sigma_j^2 u_j + u_u^2 u}\right),$$

where $Z_0 \sim \mathrm{Beta}(1/\sigma_j^2, 1/u_u^2)$, which completes the proof. ■

Theorem S3: Expected Response

Assume a randomized bidding strategy $U \sim \mathrm{Gamma}(1/u_u^2, 1/(u_u^2 u))$ with $u, u_u > 0$ and consider bidding in n segments, denoted $j = 1, \ldots, n$. Suppose the highest competing bid price u_j and the available incremental response $y_j^a, j = 1, \ldots, n$ are known only probabilistically as realizations of the random variables U_j and Y_j^a, where $U_j \sim \mathrm{Gamma}(1/\sigma_j^2, 1/(\sigma_j^2 u_j))$ and $\mathrm{E}Y_j^a = y_j^a$ for known values of σ_j, u_j, and y_j^a. If U_j and Y_j^a are independent for all j, then the expected response is

$$\mathrm{E}(Y) = \sum_{j=1}^{n} y_j^a F_{Z_0}\left(\frac{u_u^2 u}{\sigma_j^2 u_j + u_u^2 u}\right), \quad \text{(S20)}$$

where $Z_0 \sim \mathrm{Beta}(1/\sigma_j^2, 1/u_u^2)$.

Proof

The expected response is given by the expected value of (S8). Since Y_j^a and W_j are independent, the expected response satisfies

$$\mathrm{E}(Y) = \sum_{j=1}^{n} \mathrm{E}(Y_j^a)\mathrm{E}(W_j).$$

Replacing $\mathrm{E}(Y_j^a)$ with y_j^a and $\mathrm{E}(W_j)$ with (S13) yields

$$\mathrm{E}(Y) = \sum_{j=1}^{n} y_j^a F_{Z_0}\left(\frac{u_u^2 u}{\sigma_j^2 u_j + u_u^2 u}\right),$$

where $Z_0 \sim \mathrm{Beta}(1/\sigma_j^2, 1/u_u^2)$, which completes the proof. ■

Theorem S4: Expected Win Rate Sensitivity

Assume a randomized bidding strategy $U \sim \text{Gamma}(1/u_u^2, 1/(u_u^2 u))$ with $u, u_u > 0$. If the highest competing bid price, u_j, is only known probabilistically as a realization of the random variable $U_j \sim \text{Gamma}(1/\sigma_j^2, 1/(\sigma_j^2 u_j))$ for known values of $\sigma_j, u_j > 0$, then the *expected win rate sensitivity* is

$$\mathrm{E}\left(\frac{dW_j}{du}\right) = \frac{u_u^2\sigma_j^2}{u(u_u^2+\sigma_j^2+u_u^2\sigma_j^2)(u_u^2+\sigma_j^2)} f_{Z_1}\left(\frac{u_u^2 u}{\sigma_j^2 u_j + u_u^2 u}\right), \quad \text{(S21)}$$

where $Z_1 \sim \text{Beta}(1/\sigma_j^2+1, 1/u_u^2+1)$.

Proof

The expected win rate sensitivity is

$$\mathrm{E}\left(\frac{dW_j}{du}\right) = \int_0^\infty \frac{dw_j}{du}(u_j) f_{U_j}(u_j)\, du_j,$$

where dw_j/du is a realization of (S11), and $f_{U_j}(u_j)$ is the probability density function of a Gamma $(1/\sigma_j^2, 1/(\sigma_j^2 u_j))$ random variable. Substituting the expressions for dw_j/du and $f_{U_j}(u_j)$ yields

$$\mathrm{E}\left(\frac{dW_j}{du}\right) = \int_0^\infty \frac{1}{u_u^2 u} \frac{1}{\Gamma\left(\frac{1}{u_u^2}+1\right)} \left(\frac{u_j}{u_u^2 u}\right)^{1/u_u^2} \times e^{-u_j/(u_u^2 u)} \frac{\left(\frac{1}{\sigma_j^2 u_j}\right)^{1/\sigma_j^2}}{\Gamma\left(\frac{1}{\sigma_j^2}\right)} u_j^{1/\sigma_j^2 - 1} e^{-u_j/(\sigma_j^2 u_j)}\, du_j.$$

Collecting all factors independent of u_j and moving them outside the integral sign yields

$$\mathrm{E}\left(\frac{dW_j}{du}\right) = \frac{1}{\Gamma\left(\frac{1}{u_u^2}+1\right)\Gamma\left(\frac{1}{\sigma_j^2}\right)(u_u^2 u)^{1/u_u^2+1}(\sigma_j^2 u_j)^{1/\sigma_j^2}} \times \int_0^\infty u_j^{1/u_u^2+1/\sigma_j^2-1} e^{-(1/(u_u^2 u)+1/(\sigma_j^2 u_j))u_j}\, du_j.$$

The integrand is recognized as the kernel of a Gamma $(1/u_u^2+1/\sigma_j^2, 1/(u_u^2 u)+1/(\sigma_j^2 u_j))$ probability distribution. Hence, the integral over $(0,\infty)$ must equal $\Gamma(1/u_u^2+1/\sigma_j^2)/[1/(u_u^2 u)+1/(\sigma_j^2 u_j)]^{1/u_u^2+1/\sigma_j^2}$. Replacement of the integral with this expression and straightforward rearrangement of the right-hand side yields

$$\mathrm{E}\left(\frac{dW_j}{du}\right) = \frac{\Gamma\left(\frac{1}{u_u^2}+\frac{1}{\sigma_j^2}\right)}{u_u^2 u \Gamma\left(\frac{1}{u_u^2}+1\right)\Gamma\left(\frac{1}{\sigma_j^2}\right)} \left(\frac{u_u^2 u}{\sigma_j^2 u_j + u_u^2 u}\right)^{1/\sigma_j^2} \times \left(1 - \frac{u_u^2 u}{\sigma_j^2 u_j + u_u^2 u}\right)^{1/u_u^2}.$$

However, $0 < u_u^2 u/(\sigma_j^2 u_j + u_u^2 u) < 1$. Hence, the expression for $\mathrm{E}(dW_j/du)$ contains the kernel of a Beta $(1/\sigma_j^2+1, 1/u_u^2+1)$ distribution evaluated at $u_u^2 u/(\sigma_j^2 u_j + u_u^2 u)$. Replacing the kernel with an expression containing the corresponding beta probability density function yields

$$\mathrm{E}\left(\frac{dW_j}{du}\right) = \frac{\Gamma\left(\frac{1}{u_u^2}+\frac{1}{\sigma_j^2}\right)}{u_u^2 u \Gamma\left(\frac{1}{u_u^2}+1\right)\Gamma\left(\frac{1}{\sigma_j^2}\right)} B\left(\frac{1}{\sigma_j^2}+1, \frac{1}{u_u^2}+1\right) \times f_{Z_1}\left(\frac{u_u^2 u}{\sigma_j^2 u_j + u_u^2 u}\right),$$

where $Z_1 \sim \text{Beta}(1/\sigma_j^2+1, 1/u_u^2+1)$. Use the known identities $\Gamma(x+1) = x\Gamma(x)$ for $x > 0$ and $B(x,y) = \Gamma(x)\Gamma(y)/\Gamma(x+y)$ [27] together with various cancelations to obtain

$$\mathrm{E}\left(\frac{dW_j}{du}\right) = \frac{\frac{1}{u_u^2}\cdot\frac{1}{\sigma_j^2}}{u\left(\frac{1}{u_u^2}+\frac{1}{\sigma_j^2}+1\right)\left(\frac{1}{u_u^2}+\frac{1}{\sigma_j^2}\right)} f_{Z_1}\left(\frac{u_u^2 u}{\sigma_j^2 u_j + u_u^2 u}\right).$$

This can be further simplified as

$$\mathrm{E}\left(\frac{dW_j}{du}\right) = \frac{u_u^2\sigma_j^2}{u(u_u^2+\sigma_j^2+u_u^2\sigma_j^2)(u_u^2+\sigma_j^2)} f_{Z_1}\left(\frac{u_u^2 u}{\sigma_j^2 u_j + u_u^2 u}\right),$$

where $Z_1 \sim \text{Beta}(1/\sigma_j^2+1, 1/u_u^2+1)$, which completes the proof. ■

Theorem S5: Expected Response Sensitivity (Plant Gain)

Assume a randomized bidding strategy $U \sim \text{Gamma}(1/u_u^2, 1/(u_u^2 u))$ with $u, u_u > 0$ and consider bidding in n segments, denoted $j = 1, \ldots, n$. Suppose the highest competing bid price u_j and the available incremental response $y_j^a, j = 1, \ldots, n$ are known only probabilistically as realizations of the random variables U_j and Y_j^a [where $U_j \sim \text{Gamma}(1/\sigma_j^2, 1/(\sigma_j^2 u_j))$ and $\mathrm{E}Y_j^a = y_j^a$] for known values of σ_j, u_j, and y_j^a. If U_j and Y_j^a are independent for all j, then the expected response sensitivity is

$$\mathrm{E}\left(\frac{dY}{du}\right) = \sum_{j=1}^n \frac{u_u^2\sigma_j^2}{u(u_u^2+\sigma_j^2+u_u^2\sigma_j^2)(u_u^2+\sigma_j^2)} y_j^a f_{Z_1}\left(\frac{u_u^2 u}{\sigma_j^2 u_j + u_u^2 u}\right),$$

where $Z_1 \sim \text{Beta}(1/\sigma_j^2+1, 1/u_u^2+1)$.

Proof

The expected response sensitivity is given by the expected value of (S12). Since Y_j^a and dW_j/du are independent, the expected response sensitivity satisfies

$$\mathrm{E}\left(\frac{dY}{du}\right) = \sum_{j=1}^n \mathrm{E}(Y_j^a)\mathrm{E}\left(\frac{dW_j}{du}\right).$$

Replacing $\mathrm{E}(Y_j^a)$ with y_j^a and $\mathrm{E}(dW_j/du)$ with (S21) yields

$$\mathrm{E}\left(\frac{dY}{du}\right) = \sum_{j=1}^n \frac{u_u^2\sigma_j^2}{u(u_u^2+\sigma_j^2+u_u^2\sigma_j^2)(u_u^2+\sigma_j^2)} y_j^a f_{Z_1}\left(\frac{u_u^2 u}{\sigma_j^2 u_j + u_u^2 u}\right),$$

where $Z_1 \sim \text{Beta}(1/\sigma_j^2+1, 1/u_u^2+1)$, which completes the proof. ■

A deeper study of the statistical properties of the response curve is available in [14]. Specifically, closed-form expressions of the Bayesian variance of the win rate sensitivity and response sensitivity are derived. The variance provides an important measure of precision of the expected value. It is naturally used along with the expected value. However, it is excluded here in the interest of space.

Most bidders in online advertising do not have access to the landscape of competing bid prices. At best, a bidder knows whether an impression was awarded or not and the highest competing bid for each awarded impression. A bidder that is awarded only a tiny fraction of all available impression opportunities can, at best, estimate the response curve in the neighborhood of the current operating point. The number of data points is too small, and important information about the response curve is censored. In this scenario, Heisenberg bidding still smoothens the response curve, which is all that is relied upon in the body of this article. A recursive least-squares estimator is proposed to estimate the plant gain based on feedback only from our own campaign delivery.

In some cases, however, additional information is available. For example, a demand-side platform (DSP) may manage thousands of campaigns simultaneously. These may be submitting bids at different price points for a diverse set of impression opportunities across a large number of impression exchanges. Although these campaigns together may still win only a small fraction of all available impressions in competition with campaigns represented by other DSPs, the prospect of a useful response curve estimation is significantly better. The diverse bidding at different price points and for different audience segments opens up the possibility of producing response curve forecasts. Such forecasts are inherently uncertain and not necessarily accurate enough for campaign planning on their own. However, they provide useful guidance for controller initialization and control signal sensitivity computation.

For example, the closed-form expression of the expected response in Theorem S3 can help an operator of the campaign control system to initialize the control system. Depending on how risk averse the advertiser is, an initial control signal is selected above or below the control signal corresponding to an expected response that is equal to the desired response.

Furthermore, the closed-form expression of the expected response sensitivity in Theorem S5 can be fed to an adaptive control system to help maintain a desired loop gain. If a recursive plant gain estimator based on campaign feedback data is available, then an enhanced plant gain estimate can be computed by combining the recursive feedback-based estimate with the response curve-based estimate via Theorem S5.

Moreover, the closed-form expression of the expected response sensitivity can also assist in the selection of an adequate control signal uncertainty u_u. While bid randomization compromises true bid optimality, the absence—or insufficient use of—of bid randomization may lead to instability or the nonexistence of a fixed-point solution (and suboptimality). As a general principle, the control signal uncertainty should be small but large enough to make the effective plant gain, as a function of the nominal control u signal, well behaved and not wildly volatile as the nominal control is adjusted up or down.

EXAMPLE: EXPECTED RESPONSE

Assume a response curve estimate is available and given in expected sense by

$$u_{1:6} = [0.2,\ 0.3,\ 0.9,\ 2.4,\ 2.5,\ 3],$$
$$y^a_{1:6} = [10,\ 2,\ 18,\ 4,\ 4,\ 7] \cdot 10^3.$$

The jth elements of $u_{1:6}$ and $y^a_{1:6}$ define u_j and y^a_j, respectively. Consider two cases of confidence in u_j as the location of the jth step. First, $\sigma_j = 0$ corresponds to the scenario of perfect knowledge of all u_j, that is, $U_j = u_j$ for $j = 1,2,\ldots,6$. In the second case, assume our belief is described by a gamma distribution with mean u_j and relative standard deviation $\sigma_j = 0.1$ for all j; that is, the unknown U_j is in the neighborhood of u_j defined by a confidence interval having a 10% relative standard deviation.

Consider five different values of control signal uncertainty: $u_u = 0, 0.05, 0.1, 0.2,$ and 0.5. The expected response curve can be computed in closed form using Theorem S3. The simplicity of the result does not come across in the theorem. However, since the beta cumulative density function is readily available in mathematics libraries that are accessible in most programming languages, the implementation to compute the expected response $\mathrm{E}(Y)$ is both simple and computationally efficient, as shown in the following pseudocode:

$y_{\text{temp}} = 0;$
for $j = 1:6$
  $z_j = u_u^2 u / (\sigma_j^2 u_j + u_u^2 u);$
  $y_{\text{temp}} = y_{\text{temp}} + y^a_j \cdot \text{betacdf}\left(z_j, 1/\sigma_j^2, 1/u_u^2\right);$
end
$\mathrm{E}(Y(u)) = y_{\text{temp}}$

The result is drawn in Figure S5, where (a) shows the result for $\sigma_j = 0$ and (b) for $\sigma_j = 0.1$. The black curves correspond to $u_u = 0$, that is, the scenario with no plant smoothing. This special case is obtained by substituting an infinitesimal value of u_u in the closed-form expression (S20). The five curves show the impact of different control signal uncertainties $u_u = 0, 0.05, 0.1, 0.2,$ and 0.5. The larger u_u, the smoother the curve, as expected. The curves in Figure S5(b) suggest the response curves become smoother for larger σ_j. However, that conclusion is misleading since u_j and y^a_j are not actual random variables but, rather, unknown parameters that are described probabilistically using belief functions. The increased smoothness for larger σ_j only reflects the greater uncertainty of where the response curve is flat or steep. △

EXAMPLE: EXPECTED RESPONSE SENSITIVITY

Assume the same response curve estimate as in the "Example: Expected Response" section. Consider the four different

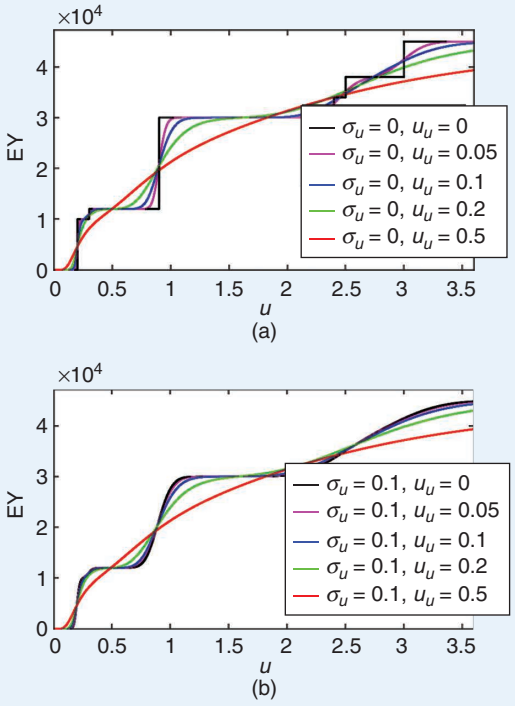


FIGURE S5 The expected response EY as a function of control signal u for different values of σ_j and u_u. The results for (a) $\sigma_j = 0$ and (b) $\sigma_j = 0.1$. The curves show the impact on the response curve using control signal uncertainties $u_u = 0$, 0.05, 0.1, 0.2, and 0.5. The larger u_u, the smoother the curve.

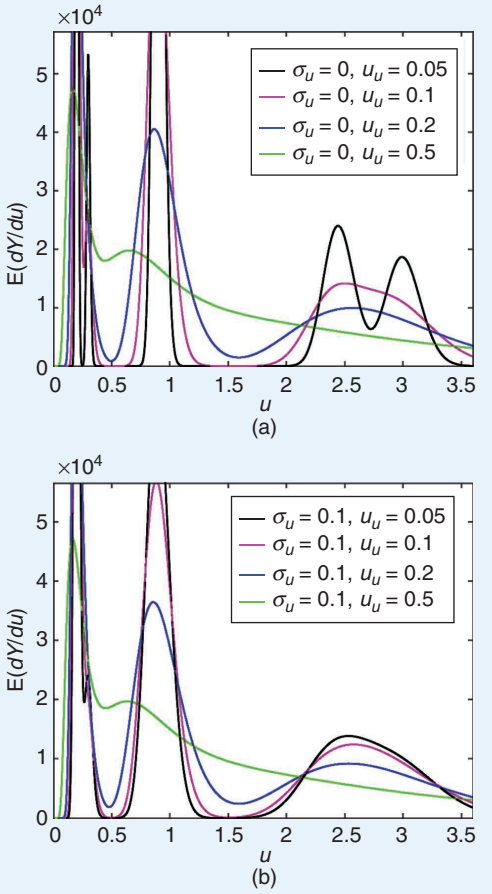


FIGURE S6 The expected response sensitivity E(dY/du) as a function of control signal u for different values of σ_j and u_u. The results for (a) $\sigma_j = 0$ and (b) $\sigma_j = 0.1$. The curves show the impact on the response sensitivity using control signal uncertainties $u_u = 0.05, 0.1, 0.2$, and 0.5. The larger u_u, the smoother the curve. For small values of u_u, the sensitivity (plant gain) spikes near the underlying step in the response curve, which is an expected behavior.

control signal uncertainty values given by $u_u = 0.05, 0.1, 0.2$, and 0.5. The expected response curve sensitivity can be computed in closed form using Theorem S5. The expression in the theorem is intimidating but simple to implement, again thanks to the availability of mathematics libraries with the relevant beta probability density functions (as shown in the following pseudocode):

$y'_{\text{temp}} = 0;$

for $j = 1:6$

$z_j = u_u^2 u / (\sigma_j^2 u_j + u_u^2 u);$

$$y'_{\text{temp}} = y'_{\text{temp}} + y_j^a \frac{u_u^2 \sigma_j^2}{u(u_u^2 + \sigma_j^2 + u_u^2 \sigma_j^2)(u_u^2 + \sigma_j^2)} \cdot \text{betapdf}\left(z_j, 1/\sigma_j^2 + 1, 1/u_u^2 + 1\right);$$

end

$$\text{E}\left(\frac{dY(u)}{du}\right) = y'_{\text{temp}}$$

The result is drawn in Figure S6, where (a) shows the result for $\sigma_j = 0$ and (b) for $\sigma_j = 0.1$. The curves show the impact on the response curve sensitivity from using control signal uncertainties $u_u = 0.05, 0.1, 0.2$, and 0.5. The larger u_u, the smoother the curve. For small values of u_u, the sensitivity (plant gain) spikes near the underlying step in the response curve, which is anticipated. The curves in Figure S6(b) suggest the response curve sensitivities are more contained for larger σ_j. However, that conclusion is misleading since u_j and y_j^a are not actual random variables but, rather, unknown parameters described probabilistically using belief functions. The increased smoothness for larger σ_j only reflects the greater uncertainty of where the response sensitivity is small or large. △

this section, the stochastic difference equations governing the system behavior are rewritten in a more manageable form. In the following section, fixed-point solutions and limit cycles of the expected trajectory are derived, together with the stability conditions of the expected trajectory. Thereafter, conditions for the stability of the covariance of the trajectory are computed.

Substituting (16) and (18) into (19) and combining the result with (21) and (23) obtains $e(t)$ and $u(t)$ as functions of the states

$$e(t) = -x_1(t) + x_3(t), \tag{24}$$

$$u(t) = -c_l(1 + x_4(t))x_1(t) + (1 + x_4(t))x_2(t) + c_l(1 + x_4(t))x_3(t). \tag{25}$$

Next, substitute (25) into (15) and (24) into (20) and define

ALGORITHM 1 Control.

Parameters: $c_l, \sigma_u, h(t), a_2$; (in practice, parameter c_l 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 $c_l = c_l(\hat{a}_1)$)

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:

$u_c(t) = x_3(t)$

$e(t) = u_c(t) - y(t)$

$u_0(t) = x_2(t) + c_l e(t)$

$u(t) = u_0(t)(1 + x_4(t))$

$w_u(t) \sim \mathrm{WN}(0, \sigma_u^2)$

$x_2(t+1) = x_2(t) + c_l e(t)$

$x_3(t+1) = a_2 x_3(t) + (1 - a_2)(1 + h(t))\bar{u}_c(t)$

$x_4(t+1) = a_3 x_4(t) + \sqrt{1 - a_3^2}\, w_u(t)$

$$\varphi_0(t) := (1 - a_2)(1 + h(t)),$$
$$\varphi_1(x_4, w_m) := x_4(t) + w_m(t) + x_4(t) w_m(t). \tag{26}$$

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

$$\begin{aligned} x_1(t+1) = {} & (a_2 - a_1 c_l \varphi_0(t)(1 + \varphi_1(x_4, w_m)))x_1(t) \\ & + a_1 \varphi_0(t)(1 + \varphi_1(x_4, w_m))x_2(t) \\ & + a_1 c_l \varphi_0(t)(1 + \varphi_1(x_4, w_m))x_3(t) \\ & + a_0 \varphi_0(t)(1 + w_m(t)), \end{aligned} \tag{27}$$

$$x_2(t+1) = x_2(t) - c_l x_1(t) + c_l x_3(t), \tag{28}$$

$$x_3(t+1) = a_2 x_3(t) + \varphi_0(t)\bar{u}_c(t), \tag{29}$$

$$x_4(t+1) = a_3 x_4(t) + \sqrt{1 - a_3^2}\, w_u(t). \tag{30}$$

The dynamical system is shown as a block diagram in Figure 5.

States x_3 and x_4 are causally independent of x_1, x_2, and w_m and have closed-form solutions:

$$x_3(t) = a_2^t x_3(0) + \sum_{i=1}^{t} a_2^{i-1} \varphi_0(t-i)\bar{u}_c(t-i), \tag{31}$$

$$x_4(t) = a_3^t x_4(0) + \sqrt{1 - a_3^2} \sum_{i=1}^{t} a_3^{i-1} w_u(t-i). \tag{32}$$

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

$$\tilde{x}_1(t) = c_l(x_1(t) - x_3(t)), \tag{33}$$

$$\tilde{x}_2(t) = x_2(t) - \frac{\bar{u}_c - a_0}{a_1}. \tag{34}$$

Combining (33) and (34) with (27) and (28) yields the state-update equations for $\tilde{x}_1(t)$ and $\tilde{x}_2(t)$:

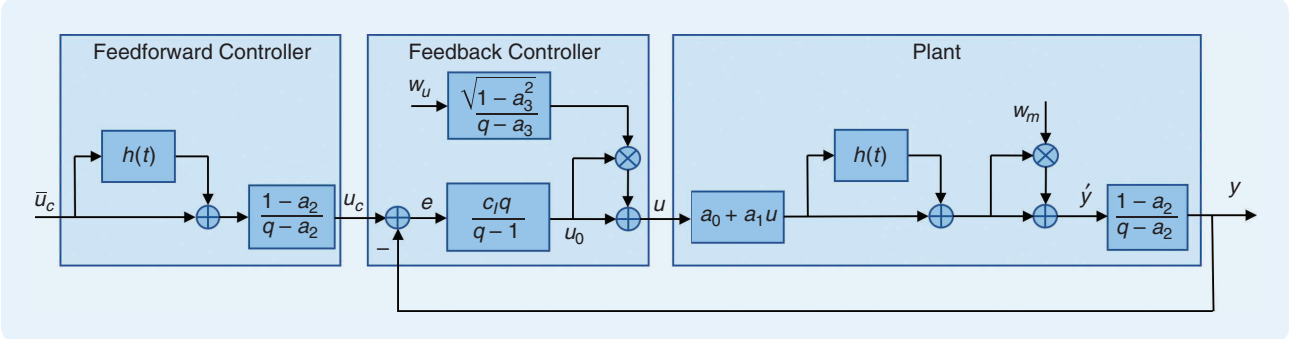


FIGURE 5 A block diagram of the nonadaptive closed-loop system. The figure depicts the interconnected system in (27)–(30), and q represents the forward-shift operator (for example, $qx(t) = x(t+1)$). It consists of a time-periodic, linear, dynamic plant with multiplicative process noise; a pure error feedback controller with artificially generated random excitation perturbations; and a time-periodic dynamic feedforward controller.

$$\begin{aligned}\tilde{x}_1(t+1) &= (a_2 - a_1 c_I \varphi_0(t)(1+\varphi_1(x_4, w_m)))\tilde{x}_1(t)\\ &\quad + a_1 c_I \varphi_0(t)(1+\varphi_1(x_4, w_m))\tilde{x}_2(t)\\ &\quad + c_I \varphi_0(t)\varphi_1(x_4, w_m)\bar{u}_c\\ &\quad - c_I \varphi_0(t)\varphi_1(x_4, w_m)a_0\\ &\quad + c_I \varphi_0(t)\varphi_1(x_4, w_m)w_m(t),\\ \tilde{x}_2(t+1) &= -\tilde{x}_1(t) + \tilde{x}_2(t).\end{aligned}$$

Combined with (24) and (25), the following output equations for $u(t)$ and $e(t)$ are obtained:

$$u(t) = (1+x_4(t))\left(-\tilde{x}_1(t) + \tilde{x}_2(t) + \frac{\bar{u}_c - a_0}{a_1}\right), \tag{35}$$

$$e(t) = -\frac{1}{c_I}\tilde{x}_1(t). \tag{36}$$

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

$$\tilde{x}(t+1) = A(t, x_4, w_m)\tilde{x}(t) + B\epsilon_1(t, x_4, w_m), \tag{37}$$

$$z(t) = C(x_4)\tilde{x}(t) + D\epsilon_2(x_4), \tag{38}$$

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

$$\begin{aligned}&A(t, x_4, w_m)\\ &= \begin{bmatrix} a_2 - a_1 c_I \varphi_0(t)(1+\varphi_1(x_4, w_m)), & a_1 c_I \varphi_0(t)(1+\varphi_1(x_4, w_m)) \\ -1, & 1 \end{bmatrix},\end{aligned} \tag{39}$$

$$B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \tag{40}$$

$$C(x_4) = \begin{bmatrix} -1 - x_4(t), & 1 + x_4(t). \\ -1/c_I, & 0 \end{bmatrix}, \tag{41}$$

$$D = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \tag{42}$$

$$\epsilon_1(t, x_4, w_m) = c_I \varphi_0(t)(\varphi_1(x_4, w_m)(\bar{u}_c - a_0) + a_0 w_m), \tag{43}$$

$$\epsilon_2(x_4) = \frac{\bar{u}_c - a_0}{a_1}(1 + x_4). \tag{44}$$

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

Difference equations (37) and (38) are more appealing than (27)–(30). However, caution is needed to avoid drawing incorrect conclusions. Not only are A and C stochastic matrices, A is also periodic, which prohibits simply applying what is known from the theory of linear time-invariant systems. To appreciate the complexity of linear periodic systems, see "Linear Periodic Systems" for a quick review of some results.

To prevent x_4 from significantly impacting the dynamics of $\tilde{x}$, the dynamics of x_4 are designed to be much faster than the dynamics of the closed-loop error feedback system. Simultaneously, the dynamics of x_4 must be sufficiently slow to allow plant identification when the assumed plant latency is only approximately correct. The time constant of the plant latency is usually a few minutes, while the desired time constant of the closed-loop error feedback system is a few hours. By selecting a persistent-excitation time constant that is similar to the plant latency time constant (and much smaller than the closed-loop system time constant), the following assumption can be made.

Assumption 1

States $x_4(t)$ and $\tilde{x}(t)$ are statistically independent.

Strictly speaking, Assumption 1 holds true only if $x_4(t)$ is WN, that is, if a_3 is zero. According to (37), $\tilde{x}(t)$ always depends on $x_4(t-1)$. Thus, if $x_4(t-1)$ and $x_4(t)$ are statistically dependent (a_3 is nonzero), then $x_4(t)$ and $\tilde{x}(t)$ are also statistically dependent. However, if a_3 is chosen sufficiently close to zero, then the statistical dependency is insignificant, allowing for the analysis as if Assumption 1 is true.

Although (37) and (38) encode everything there is to know about the closed-loop system, they do not directly reveal what is the expected behavior or the plausible variability around the expected behavior as a result of stochastic inputs. The dynamics in the expected sense are analyzed next. Thereafter, the variability is analyzed in terms of the second-order statistical moment or the variance.

EXPECTED TRAJECTORY AND STABILITY

A good starting point to describe the solution of the stochastic difference equations (37) and (38) is to compute the expected solution and determine the conditions under which this solution is stable. To derive the expected solution of (37) and (38) and stability conditions, the following results are used.

Lemma 2

At the steady state, it is always true that

$$\begin{aligned}\mathrm{E}(\varphi_1) &= 0,\\ \mathrm{E}(\epsilon_1) &= 0,\\ \mathrm{E}(\epsilon_2) &= (\bar{u}_c - a_0)/a_1.\end{aligned}$$

Proof

1) Use the law of total expectation [21] to deduce that $\mathrm{E}(\varphi_1(x_4, w_m)) = \mathrm{E}(\mathrm{E}(\varphi_1(x_4, w_m) \mid x_4))$. Furthermore, $\mathrm{E}(\varphi_1(x_4, w_m) \mid x_4) = \mathrm{E}(x_4 + w_m + x_4 w_m \mid x_4) = x_4 + \mathrm{E}(w_m) + x_4\mathrm{E}(w_m) = x_4$ since $\mathrm{E}(w_m) = 0$ by definition. Hence, $\mathrm{E}(\varphi_1(x_4, w_m)) = \mathrm{E}(x_4)$, which is easily shown to equal zero at the steady state, since x_4 is a linear combination of mean-zero random variables. See also "The Laws of Total Expectation and Variance."
2) Use the linearity property of the expectation operator to establish that $\mathrm{E}(\epsilon_1) = c_I \varphi_0((\bar{u}_c - a_0)\mathrm{E}(\varphi_1) + a_0\mathrm{E}(w_m))$. Since $\mathrm{E}(w_m) = \mathrm{E}(\varphi_1) = 0$, it follows that $\mathrm{E}(\epsilon_1) = 0$.

Linear Periodic Systems

Linear periodic systems appear innocent on the surface, and it is tempting to draw premature conclusions about their stability. However, periodic systems have remarkably complicated dynamics. It is important to appreciate the complexity and handle any periodic system carefully.

Extensive literature on linear systems is available (for example, [28]). However, with few exceptions, the coverage of the special case of linear periodic systems is brief. The reader is easily led to believe that periodic systems are not particularly important and that linear time-invariant systems are good enough. On the contrary, many important physical phenomena are cyclical by nature and must be modeled as time periodic to capture dominant dynamics. Fortunately, there exists plenty of adequate literature. For example, [29] focuses on periodic systems and provides a relatively up-to-date treatment of these intriguing systems. It is a good source for further study on this subject.

Basic facts about periodic systems are given here, and it is demonstrated how challenging linear periodic systems can be. The results are presented in the context of continuous-time systems, but the relevant takeaways carry over to discrete-time systems.

Consider the linear homogeneous system of first-order differential equations

$$\frac{dx}{dt} = A(t)x, \tag{S22}$$

where $x \in \mathbb{R}^n$, $A(t) \in \mathbb{R}^{n\times n}$ is a continuous function of $t \in \mathbb{R}$, and

$$A(t) = A(t+T) \tag{S23}$$

for some period $T > 0$.

Definition S1

A set of n linearly independent solutions of (S22), $\{\phi_1, \ldots, \phi_n\}$, is called a *fundamental set of solutions* of (S22), and the $n \times n$ matrix

$$\Phi = [\phi_1, \ldots, \phi_n] = \begin{bmatrix} \phi_{11} & \phi_{12} & \cdots & \phi_{1n} \\ \phi_{21} & \phi_{22} & \cdots & \phi_{2n} \\ \vdots & \vdots & & \vdots \\ \phi_{n1} & \phi_{n2} & \cdots & \phi_{nn} \end{bmatrix}$$

is called a fundamental matrix of (S22). ■

Theorem S6

Assume that $A(t) = A(t+T)$ and $A(t) \in \mathbb{R}^{n\times n}$ is a continuous function of $t \in \mathbb{R}$. If $\Phi(t)$ is a fundamental matrix for (S22), then so is $\Phi(t+T)$. Furthermore, for every Φ, there exists a nonsingular matrix P that is also periodic with a period T and a constant $n \times n$ matrix R, such that

$$\Phi(t) = P(t)e^{tR}. \tag{S24}$$

Proof

See, for example, [28] or any other comprehensive book on linear systems. ■

The first example from [28] demonstrates that scalar linear systems already have complicated dynamics as soon as $A(t)$ is periodic rather than constant.

EXAMPLE: SCALAR SYSTEM

Consider the scalar system

$$\dot{x} = -(\sin t + 2)x. \tag{S25}$$

Then, $A(t) = -(\sin t + 2)$, and $A(t)$ is periodic with $T = 2\pi$. As can be verified by substitution into the relation $\dot{\Phi}(t) = A(t)\Phi(t)$, a fundamental matrix for (S25) is given by

$$\Phi(t) = e^{\cos t - 1 - 2t}.$$

The Laws of Total Expectation and Variance

These useful and versatile results from mathematical statistics are sometimes known as the law of iterated expectation and the conditional variance identity.

The *law of total expectation* states that if X_1 and X_2 are any two random variables, then

$$\mathrm{E}X_1 = \mathrm{E}(\mathrm{E}(X_1 \,|\, X_2)), \tag{S28}$$

provided that the expectations exist.

The *law of total variance* states that for any two random variables X_1 and X_2,

$$\mathrm{Var}X_1 = \mathrm{E}(\mathrm{Var}(X_1 \,|\, X_2)) + \mathrm{Var}(\mathrm{E}(X_1 \,|\, X_2)), \tag{S29}$$

provided that the expectations exist. The identities are well known in the literature, and the proofs are both straightforward and readily available in textbooks and online [21].

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

Lemma 2 is all that is needed to prove the following key result.

Theorem 2

If setpoint signal $\bar{u}_c(t) \geq 0$ is constant, then $\mathrm{E}([x_1, x_2, x_3, x_4]^T)$ has a limit cycle $[x_1^\star, x_2^\star, x_3^\star, x_4^\star]^T$, where

$$\begin{bmatrix} x_1^\star(t) \\ x_2^\star(t) \\ x_3^\star(t) \\ x_4^\star(t) \end{bmatrix} = \begin{bmatrix} \bar{u}_c \sum_{i=1}^{\infty} a_2^{i-1} \varphi_0(t-i) \\ (\bar{u}_c - a_0)/a_1 \\ \bar{u}_c \sum_{i=1}^{\infty} a_2^{i-1} \varphi_0(t-i) \\ 0 \end{bmatrix},$$

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

To determine $P(t)$ and R, let $t = 0$ and $t = 2\pi$ in (S24). Note that $P(t)$ satisfies $P(2\pi) = P(0)$, and $\Phi(0) = P(0)$. We obtain $\Phi(2\pi) = e^{-4\pi} = \Phi(0)e^{2\pi R} = e^{2\pi R}$, or

$$R = -2.$$

Matrix $P(t)$ is now obtained from (S24) as $P(t) = \Phi(t)e^{-tR}$. The substitution with derived expressions of $\Phi(t)$ and R yields

$$\begin{aligned} P(t) &= e^{\cos t - 1 - 2t} e^{2t} \\ &= e^{\cos t - 1}, \end{aligned}$$

which is clearly periodic, with period $T = 2\pi$. System (S25) is next transformed by $P(t)$ via $y = P^{-1}(t)x$ into the system

$$\begin{aligned} \dot{y} &= e^{1-\cos t}[-(\sin t + 2)e^{\cos t - 1} + \sin t e^{\cos t - 1}]y \\ &= -(e^{1-\cos t})(2e^{\cos t - 1})y \\ &= -2y \\ &= Ry. \end{aligned} \tag{S26}$$

Since $R < 0$, it is concluded that the origin of (S26) [and, therefore, also of (S25)] is asymptotically stable. △

The following example from [30] illustrates that a linear periodic system may be unstable even if the eigenvalues of $A(t)$ are always in the open left half-plane for a continuous-time system (inside the unit circle for a discrete-time system).

EXAMPLE: MATHIEU EQUATION

Consider the second-order differential equation

$$\frac{dx}{dt} = \underbrace{\begin{bmatrix} 0 & 1 \\ -a_1(t) & -a_0 \end{bmatrix}}_{A(t)} x. \tag{S27}$$

When $a_0(t)$ and $a_1(t)$ are constants, a necessary and sufficient condition for the exponential stability of $x = [0, 0]^T$ is $a_0, a_1 > 0$. However, consider instead

$$\begin{aligned} a_0(t) &= a_0 > 0 \\ a_1(t) &= \alpha - \beta \cos \omega t > 0, \end{aligned}$$

which implies that (S27) represents the familiar damped Mathieu equation. The stability of (S27) depends on the values of the involved parameters. For $a_0 = 0.1, \alpha = 10.1$, $\beta = 10$, and $\omega = \pi = 3.14159$, the eigenvalues $\lambda_{1,2}$ of the state matrix $A(t)$ are

$$\lambda_{1,2} = -0.05 \pm i\frac{\sqrt{0.39 + 40(1 - \cos \pi t)}}{2},$$

which, clearly, are in the left half-plane for all t. However, the stability of (S27) is not determined by $\lambda_{1,2}$ but, rather, by the properties of transition matrix $\Phi(t, t_0)$. By recognizing that $A(t)$ is periodic with period $T = 2$, we may numerically integrate

$$\frac{d\Phi(t,0)}{dt} = A(t)\Phi(t,0),$$

where $\Phi(0,0)$ is the identity matrix. We obtain

$$\Phi(T,0) = \begin{bmatrix} 1.0988 & 0.1784 \\ 2.0683 & 1.0809 \end{bmatrix},$$

whose eigenvalues have absolute values 1.6973, 0.4824. Since $x(kT) = \Phi^k(T,0)x(0), k \in \mathbb{N}$, and $\Phi(T,0)$ has eigenvalues outside the unit circle, it follows that there always exists an initial value $x(0)$ such that $x(t)$ diverges to $-\infty$ or ∞ as $k \to \infty$. Hence, the zero equilibrium of (S27) is unstable. △

These examples illustrate that the stability properties of linear time-periodic systems are quite complicated and not always compatible with our intuition from linear time-invariant systems. Furthermore, the analysis of these properties for a given state matrix $A(t)$ is a formidable task.

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

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

Proof

First, compute the expected values of x_4, x_3, and $\tilde{x} = [\tilde{x}_1, \tilde{x}_2]^T$, in that order. Since $\mathrm{E}(w_u(t)) = 0$ for all t, and $0 \le a_3 < 1$, it follows from (32) that

$$\mathrm{E}(x_4(t)) = a_3^t x_4(0),$$

which converges to zero. Hence, $\mathrm{E}(x_4(t))$ has a globally asymptotically stable, fixed-point solution $x_4^\star = 0$.

The trajectory of $x_3(t)$ is deterministic, with a closed-form solution given by (31). Recall that $\bar{u}_c(t)$ is constant and $\varphi_0(t)$ is bounded, while $0 \le a_2 < 1$. As long as the infinite sums exist, the finite sum in the solution for $x_3(t)$ can be written as the difference between $x_3^\star(t)$ and an infinite tail sum as

$$x_3(t) = a_2^t x_3(0) + x_3^\star(t) - \bar{u}_c \sum_{i=t+1}^{\infty} a_2^{i-1} \varphi_0(t-i).$$

The first term, $a_2^t x_3(0)$, converges to zero as t goes to infinity. Furthermore, $x_3^\star(t) \ge 0$, and

$$\begin{aligned} x_3^\star(t) &\equiv \bar{u}_c \sum_{i=1}^{\infty} a_2^{i-1} \varphi_0(t-i) \\ &\le |\bar{u}_c| \max_\tau (|\varphi_0(\tau)|) \sum_{i=1}^{\infty} a_2^{i-1} \\ &= \frac{|\bar{u}_c| \max_\tau (|\varphi_0(\tau)|)}{1 - a_2} \\ &< \infty. \end{aligned}$$

Hence, $x_3^\star(t)$ is bounded and exists. It is also trivial to show that $x_3^\star(t+T) = x_3^\star(t)$. Finally, the tail sum satisfies

$$\left|\bar{u}_c \sum_{i=t+1}^{\infty} a_2^{i-1}\varphi_0(t-i)\right| < |\bar{u}_c| \max_{\tau}(|\varphi_0(\tau)|) \sum_{i=t+1}^{\infty} a_2^{i-1},$$

which converges to zero as $t \to \infty$. Therefore, $x_3(t) \to x_3^\star(t)$, which is a bounded T-periodic function. In other words, $\mathrm{E}(x_3(t))$ has a globally asymptotically stable limit cycle defined by $x_3^\star$.

Finally, consider $\tilde{x} = [\tilde{x}_1, \tilde{x}_2]^T$. The expected value of $\tilde{x}(t+1)$ can be rewritten using the law of total expectation as $\mathrm{E}(\tilde{x}(t+1)) = \mathrm{E}(\mathrm{E}(\tilde{x}(t+1) \mid \tilde{x}(t)))$. Replace $\tilde{x}(t+1)$ with the right-hand side of (37) and recognize that $\mathrm{E}(A(t, x_4, w_m)\tilde{x}(t) + B\epsilon_1(t, x_4, w_m) \mid \tilde{x}(t)) = \mathrm{E}(A(t, x_4, w_m) \mid \tilde{x}(t))\tilde{x}(t) + B\mathrm{E}(\epsilon_1(x_4, w_m) \mid \tilde{x}(t))$. Assumption 1 implies that

$$\mathrm{E}(A(t, x_4, w_m) \mid \tilde{x}(t)) = \mathrm{E}(A(t, x_4, w_m)),$$
$$\mathrm{E}(\epsilon_1(x_4, w_m) \mid \tilde{x}(t)) = \mathrm{E}(\epsilon_1(x_4, w_m)).$$

Since $\mathrm{E}(\varphi_1) = 0$, it is, furthermore, 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,

$$\mathrm{E}(\tilde{x}(t+1)) = A(t, 0, 0)\mathrm{E}(\tilde{x}(t)),$$

which has the origin as a fixed-point solution. The solution is not automatically unique since $A(t,0,0)$ potentially does not have full rank and stability is not yet proven.

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

Invert the mapping in (33) and (34) to obtain

$$x_1(t) = \frac{1}{c_l}\tilde{x}_1(t) + x_3(t), \tag{45}$$

$$x_2(t) = \tilde{x}_2(t) + \frac{\bar{u}_c - a_0}{a_1}. \tag{46}$$

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 $\mathrm{E}([u(t), e(t)]^T)$ is derived using the law of total expectation, which implies $\mathrm{E}(z(t)) = \mathrm{E}(\mathrm{E}(z(t) \mid \tilde{x}(t)))$. Replace $z(t)$ with the right-hand side of (38) and recognize that $\mathrm{E}(C(x_4)\tilde{x}(t) + D\epsilon_2(x_4) \mid \tilde{x}(t)) = \mathrm{E}(C(x_4) \mid \tilde{x})\tilde{x}(t) + D\mathrm{E}(\epsilon_2(x_4) \mid \tilde{x})$. Based on Assumption 1, it follows that $\mathrm{E}(C(x_4) \mid \tilde{x}) = \mathrm{E}(C(x_4))$ and $\mathrm{E}(\epsilon_2(x_4) \mid \tilde{x}) = \mathrm{E}(\epsilon_2(x_4))$. Furthermore, note that $\mathrm{E}(C(x_4)) = C(0)$ and $\mathrm{E}(\epsilon_2(x_4)) = \epsilon_2(0)$. Hence,

$$\mathrm{E}(z(t)) = C(0)\mathrm{E}(\tilde{x}(t)) + D\epsilon_2(0). \tag{47}$$

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 derived 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 $\bar{A}(1)$ are strictly inside the unit circle. This completes the proof. ■

It is, in general, not possible to derive a closed-form expression for the eigenvalues of $\bar{A}(1)$ as a function of

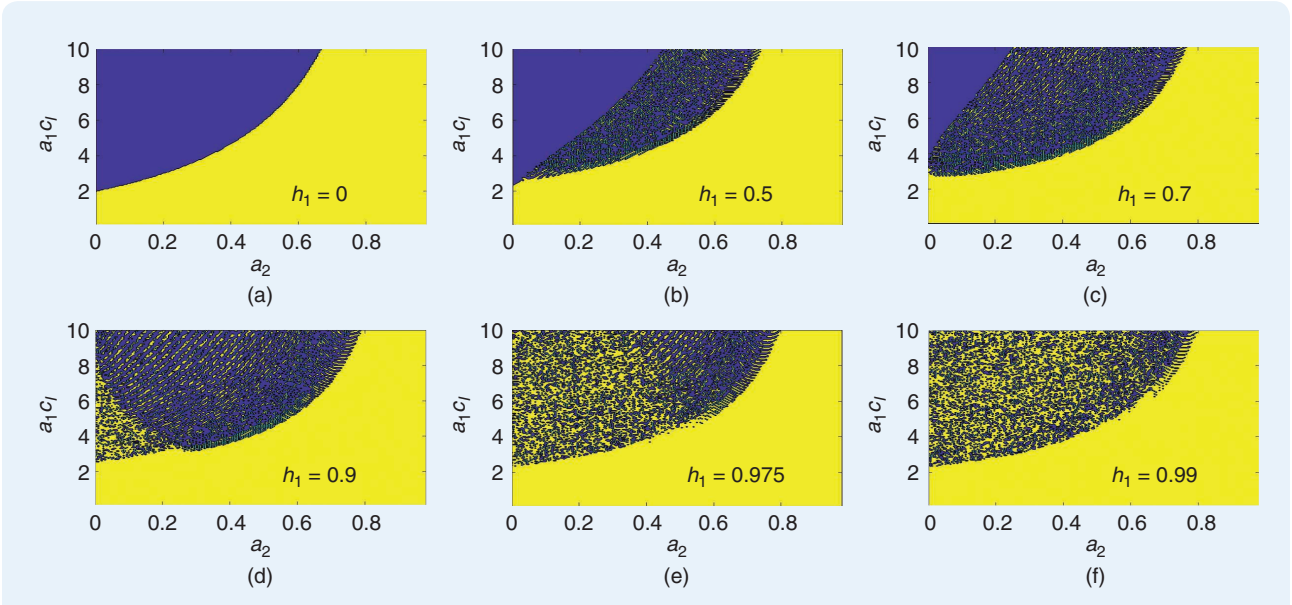


FIGURE 6 The stability regime of the expected trajectory for sinusoidal plant seasonality. The charts indicate the stability (yellow) and instability (blue) of the expected limit cycle for different combinations of a_2, a_1c_l, and h_1 when $h(t) = h_1\sin(2\pi t\Delta/24)$ and $\Delta = 5/60$ h. (From [12].)

parameters a_2, a_1, c_l, and $h(t)$. However, it is straightforward for any set of parameters (and with the 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_l always appear together as a product. This reduces the degrees of freedom as the dynamics are analyzed. It is sufficient to consider combinations of the triplet a_2, a_1c_l, and $h(t)$ to gain a complete understanding of the stability of any configuration.

Example: The Stability Region of the Expected Trajectory for Sinusoidal Seasonality

Consider a sinusoidal plant seasonality

$$h(t) = h_1 \sin\left(\frac{2\pi t\Delta}{24}\right),$$

where $\Delta = 5/60$ h for six different values of the seasonality amplitude

$$h_1 = 0, 0.5, 0.7, 0.9, 0.975, \text{and } 0.99.$$

For each value of h_1, compute the eigenvalues of $\bar{A}(1)$ for 40,000 different combinations of a_2 and a_1c_l, and determine the stability of each combination. The result is shown using heat maps in Figure 6. Blue indicates instability, and yellow indicates stability. For values of h_1 that are different from zero, the heat maps have three regions. In the blue-connected region, all combinations of a_2 and a_1c_l correspond 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 this region, tiny alterations of a_2 or a_1c_l may turn a stable system unstable or an unstable system stable. The stable configurations in this region are structurally nonrobust. In conclusion, the design goal is to operate in the fully yellow region and on a sufficiently large distance from any unstable configuration to ensure robustness. △

Example: The Stability Region of the Expected Trajectory for Square Wave Seasonality

Consider a square wave plant seasonality

$$h(t) = \begin{cases} h_1, & \text{if } 0 \le \text{mod}(t,T) < T/2, \\ -h_1, & \text{otherwise} \end{cases}$$

for six different values of the amplitude

$$h_1 = 0, 0.5, 0.7, 0.9, 0.975, \text{and } 0.99.$$

For each value of h_1, compute the eigenvalues of $\bar{A}(1)$ for 40,000 different combinations of a_2 and a_1c_l, and determine the stability of each combination. The result is shown using heat maps in Figure 7. Blue indicates instability, and yellow indicates stability. The conclusions are similar to those drawn in the previous example, with some qualitative differences in terms of how the regions change as h_1 goes from zero to one. Again, the objective is to operate in the fully yellow region sufficiently far from any unstable configuration. △

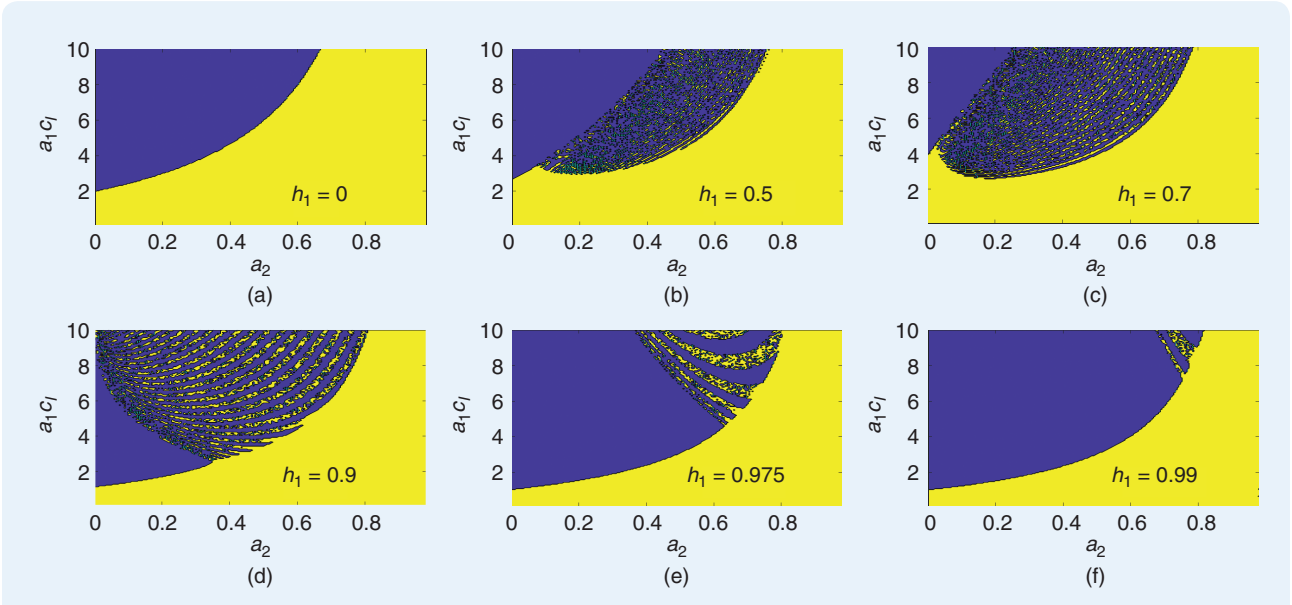


FIGURE 7 The stability regime of the expected trajectory for square wave plant seasonality. The charts indicate the stability (yellow) and instability (blue) of the expected limit cycle for different combinations of a_2, a_1c_l, and h_1 when $h(t) = h_1$ if $0 \le \text{mod}(t,T) < T/2$ and $h(t) = -h_1$ otherwise.

It is a daunting task to derive a closed-form expression of the expected trajectory or of the eigenvalues of $\bar{A}(1)$ in general, which is the motivation of establishing Theorem 2. However, in the special case $a_2 = 0$, the situation is less challenging. Since a_2, typically, is small in a real implementation, this scenario has practical value.

Theorem 3

If $a_2 = 0$, then the limit cycle and fixed-point solution in Theorem 2 are globally asymptotically stable if and only if $|\Pi_{i=1}^{T}(1 - a_1 c_I \varphi_0(i))| < 1$.

Proof

With the same reasoning as in the proof of Theorem 2, we establish that $\mathrm{E}(\tilde{x}(t+1)) = A(t,0,0)\mathrm{E}(\tilde{x}(t))$. According to Floquet, $\mathrm{E}(\tilde{x}(t)) = 0$ is globally asymptotically stable if and only if the eigenvalues of $\bar{A}(1)$ are strictly inside the unit circle. Since $a_2 = 0$, it is known from (39) that

$$A(t,0,0) = \begin{bmatrix} -a_1 c_I \varphi_0(t), & a_1 c_I \varphi_0(t) \\ -1, & 1 \end{bmatrix}.$$

It is easy to confirm by matrix multiplication and factorization that, for any integer τ, it is always true that $A(\tau+1,0,0)A(\tau,0,0) = (1 - a_1 c_I \varphi_0(\tau))A(\tau+1,0,0)$. Repeating this for $\tau = 1, 2, \ldots, T-1$ yields

$$\bar{A}(1) := \left(\prod_{i=1}^{T-1}(1 - a_1 c_I \varphi_0(i))\right) A(T,0,0).$$

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

$$\det(\lambda I - \bar{A}(1)) = \lambda\left(\lambda - \prod_{i=1}^{T}(1 - a_1 c_I \varphi_0(i))\right).$$

The solutions are $\lambda_1 = 0$ (which is always inside the unit circle) and $\lambda_2 = \Pi_{i=1}^{T}(1 - a_1 c_I \varphi_0(i))$ [which is inside the unit circle if and only if $|\Pi_{i=1}^{T}(1 - a_1 c_I \varphi_0(i))| < 1$], 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 2. ■

Corollary 1

A sufficient condition for the global asymptotic stability in Theorem 3 is

$$0 < a_1 c_I < \frac{2}{1 + \max_i h(i)}. \tag{48}$$

Proof

Assume (48) is true. Since $h(t) > -1$ for all t, the inequalities in the corollary can be written as

$$0 < a_1 c_I < \frac{2}{1 + h(t)}, \quad \text{for all } t. \tag{49}$$

Furthermore, $a_2 = 0$ implies that $\varphi_0(t) := (1 - a_2)(1 + h(t)) = 1 + h(t)$. Substitute $1 + h(t)$ with $\varphi_0(t)$ in (49) to obtain

$$0 < a_1 c_I < \frac{2}{\varphi_0(t)}, \quad \text{for all } t. \tag{50}$$

Multiply the second inequality of (50) by $\varphi_0(t)$ and subtract one to obtain

$$a_1 c_I \varphi_0(t) - 1 < 1. \tag{51}$$

Next, multiply the first inequality of (50) by $-\varphi_0(t)$ (which is less than zero) and add one to obtain

$$1 > 1 - a_1 c_I \varphi_0(t). \tag{52}$$

Combine (51) and (52) into

$$|1 - a_1 c_I \varphi_0(t)| < 1. \tag{53}$$

Since (53) is true for all t, it follows that

$$\prod_{i=1}^{T}|1 - a_1 c_I \varphi_0(i)| < 1.$$

Finally, $\Pi_{i=1}^{T}|1 - a_1 c_I \varphi_0(i)| = |\Pi_{i=1}^{T}(1 - a_1 c_I \varphi_0(i))|$. Hence, the condition for global asymptotic stability in Theorem 3 is satisfied, which completes the proof. ■

COVARIANCE TRAJECTORY AND STABILITY

It is not wise to only compute the expected solution to a stochastic system. A well-behaved expected trajectory is of little use if, for example, the variance diverges. This section derives the conditions for stability of the covariance. To support that endeavor, we use Lemma 3.

Lemma 3

At the steady state, it is always true that

$$\mathrm{Cov}(x_4) = \sigma_u^2,$$
$$\mathrm{Cov}(\varphi_1) = \sigma_u^2 + \sigma_m^2 + \sigma_m^2\sigma_u^2,$$
$$\mathrm{Cov}(\epsilon_1) = (c_I\varphi_0(t))^2((\bar{u}_c - a_0)^2(\sigma_u^2 + \sigma_m^2 + \sigma_m^2\sigma_u^2) + a_0^2\sigma_m^2),$$
$$\mathrm{Cov}(\epsilon_2) = \left(\frac{\bar{u}_c - a_0}{a_1}\right)^2\sigma_u^2.$$

Proof

1) State variable $x_4(t)$ is stochastic, but it is causally independent of other states. It evolves according to (22), and its covariance is

$$\mathrm{Cov}(x_4(t+1)) = \mathrm{Cov}(a_3 x_4(t) + \sqrt{1 - a_3^2}\, w_u(t)).$$

Moreover, $\mathrm{Cov}(w_u(t)) = \sigma_u^2$, and $\mathrm{Cov}(x_4(t), w_u(t)) = 0$. It is trivial to rewrite this as

$$\mathrm{Cov}(x_4(t+1)) = a_3^2\mathrm{Cov}(x_4(t)) + (1 - a_3^2)\sigma_u^2.$$

By design, $0 \le a_3 < 1$; hence, this is a stable difference equation. Consequently, $\mathrm{Cov}(x_4(t))$ converges to its steady state, which is defined by $\mathrm{Cov}(x_4(t+1)) = \mathrm{Cov}(x_4(t))$. Therefore, $\mathrm{Cov}(x_4(t))$ converges to σ_u^2.

2) The law of total covariance [21] dictates that

$$\mathrm{Cov}(\varphi_1(x_4, w_m)) = \mathrm{Cov}(\mathrm{E}(\varphi_1(x_4, w_m)\,|\,x_4)) + \mathrm{E}(\mathrm{Cov}(\varphi_1(x_4, w_m)\,|\,x_4)). \quad (54)$$

See also "The Laws of Total Expectation and Variance." Meanwhile, $\varphi_1(x_4, w_m) = x_4 + w_m + x_4 w_m$ according to (26), and, consequently,

$$\mathrm{E}(\varphi_1(x_4, w_m)\,|\,x_4) = x_4, \quad (55)$$

$$\mathrm{Cov}(\varphi_1(x_4, w_m)\,|\,x_4) = (1 + x_4)^2 \sigma_m^2. \quad (56)$$

Substitute (55) and (56) into (54); then, leverage $\mathrm{E}(x_4) = 0$ and the first result of this lemma $\mathrm{E}(x_4^2) = \mathrm{Cov}(x_4) = \sigma_u^2$ to obtain

$$\begin{aligned}\mathrm{Cov}(\varphi_1(x_4, w_m)) &= \mathrm{Cov}(x_4) + \mathrm{E}((1 + x_4)^2 \sigma_m^2) \\ &= \sigma_u^2 + \sigma_m^2 \mathrm{E}(1 + 2x_4 + x_4^2) \\ &= \sigma_u^2 + \sigma_m^2 + \sigma_m^2 \sigma_u^2.\end{aligned}$$

3) Based on Lemma 2 (which states that $\mathrm{E}(\epsilon_1) = 0$), it follows that $\mathrm{Cov}(\epsilon_1) = \mathrm{E}(\epsilon_1^2)$, which, combined with the definition of ϵ_1 in (43), yields

$$\begin{aligned}\mathrm{Cov}(\epsilon_1) &= (c_I \varphi_0)^2 \mathrm{E}((\bar{u}_c - a_0)\varphi_1 + a_0 w_m)^2 \\ &= (c_I \varphi_0)^2 ((\bar{u}_c - a_0)^2 \mathrm{E}(\varphi_1^2) + a_0^2 \mathrm{E}(w_m^2) \\ &\quad + a_0(\bar{u}_c - a_0)\mathrm{E}(\varphi_1 w_m)).\end{aligned}$$

However, $\mathrm{E}(w_m^2) = \sigma_m^2$, and, with help from the law of total expectation, $\mathrm{E}(\varphi_1 w_m) = \mathrm{E}(\mathrm{E}(\varphi_1 w_m\,|\,\varphi_1)) = 0$. Furthermore, it is known from Lemma 2 that $\mathrm{E}(\varphi_1) = 0$, and, from the second result of this lemma, $\mathrm{E}(\varphi_1^2) = \mathrm{Cov}(\varphi_1) = \sigma_u^2 + \sigma_m^2 + \sigma_m^2 \sigma_u^2$. Therefore,

$$\mathrm{Cov}(\epsilon_1) = (c_I \varphi_0)^2 ((\bar{u}_c - a_0)^2 (\sigma_u^2 + \sigma_m^2 + \sigma_m^2 \sigma_u^2) + a_0^2 \sigma_m^2).$$

4) Given the definition of ϵ_2 in (44), it follows that the autocovariance of ϵ_2 is

$$\mathrm{Cov}(\epsilon_2) = \mathrm{Cov}\left(\frac{\bar{u}_c - a_0}{a_1}(1 + x_4)\right).$$

However, $\mathrm{E}(x_4) = 0$, and $\mathrm{Cov}(x_4) = \sigma_u^2$; hence, trivially,

$$\mathrm{Cov}(\epsilon_2) = \left(\frac{\bar{u}_c - a_0}{a_1}\right)^2 \sigma_u^2,$$

which completes the proof of the lemma. ■

The convergence speed toward the steady-state values in Lemma 3 is driven by the convergence speed of $\mathrm{Cov}(x_4(t))$. However, the dynamics of the persistent-excitation controller are fast by design (a_3 is small). Hence, $\mathrm{Cov}(x_4(t))$ and covariances determined by $\mathrm{Cov}(x_4(t))$ approach the steady-state values rapidly. The following key result about the covariance of state and output can now be derived by utilizing Lemma 3.

Theorem 4

Globally asymptotically stable and unique limit cycles of $\mathrm{Cov}(x(t))$ and $\mathrm{Cov}(z(t))$ exist if and only if the eigenvalues of

$$\bar{\mathcal{A}}(1) := \mathcal{A}(T)\mathcal{A}(T-1)\ldots\mathcal{A}(1)$$

are strictly inside the unit circle, where

$$\mathcal{A}(t) = A_0(t) \otimes A_0(t) + A_1(t) \otimes A_1(t)\mathrm{Cov}(\varphi_1), \quad (57)$$

where $\otimes$ denotes the Kronecker product; $\mathcal{A}(t)$ is constructed from

$$A_0(t) = \begin{bmatrix} a_2 - a_1 c_I \varphi_0(t), & a_1 c_I \varphi_0(t) \\ -1, & 1 \end{bmatrix}, \quad (58)$$

$$A_1(t) = \begin{bmatrix} -1, & 1 \\ 0 & 0 \end{bmatrix} a_1 c_I \varphi_0(t), \quad (59)$$

and $\mathrm{Cov}(\varphi_1) = \sigma_u^2 + \sigma_m^2 + \sigma_m^2 \sigma_u^2$.

Proof

The autocovariance of $x_4(t)$ is stable, and its steady-state value equals σ_u^2 (Lemma 3). Meanwhile, $x_4(t)$, according to Assumption 1, is statistically independent of other states, which implies they are also uncorrelated. Moreover, $x_3(t)$ is deterministic and, therefore, also statistically uncorrelated with all other states. It follows that the stability of $\mathrm{Cov}(x(t))$ is dictated by the stability of $\mathrm{Cov}([x_1(t), x_2(t)]^T)$ or $\mathrm{Cov}(\tilde{x}(t))$ due to the relationship in (33) and (34).

Specifically, $\mathrm{Cov}(x(t))$ is globally asymptotically stable and with a unique limit cycle if and only if $\mathrm{Cov}(\tilde{x}(t))$ is globally asymptotically stable with a unique limit cycle. The dynamics of $\tilde{x}(t)$ are governed by (37). Hence, the covariance of $\tilde{x}(t)$ satisfies the difference equation

$$\mathrm{Cov}(\tilde{x}(t+1)) = \mathrm{Cov}(A(t, x_4, w_m)\tilde{x}(t) + B\epsilon_1(t, x_4, w_m)). \quad (60)$$

However, $\mathrm{E}(\tilde{x}(t)) = 0$, and it is easy to show that, when expanding the right-hand side of (60),

$$\begin{aligned}\mathrm{Cov}(\tilde{x}(t+1)) = \mathrm{E}(&A(t, x_4, w_m)\tilde{x}(t)\,\tilde{x}(t)^T A(t, x_4, w_m)^T \\ &+ B\epsilon_1(t, x_4, w_m)\,\epsilon_1(t, x_4, w_m)^T B^T),\end{aligned}$$

where Assumption 1 and the law of total expectation have been used to establish the cross term $\mathrm{E}(A(t, x_4, w_m)\tilde{x}(t)\,\epsilon_1(t, x_4, w_m)B^T) = 0$, which can be further used to rewrite the difference equation as

$$\begin{aligned}\mathrm{Cov}(\tilde{x}(t+1)) = \mathrm{E}(&A(t, x_4, w_m)\mathrm{Cov}(\tilde{x}(t))A(t, x_4, w_m)^T) \\ &+ BB^T\mathrm{Cov}(\epsilon_1),\end{aligned}$$

where $\mathrm{Cov}(\epsilon_1)$, at the steady state, is a constant and given in Lemma 3.

Next, decompose $A(t, x_4, w_m)$ into a deterministic and a stochastic component as

$$A(t, x_4, w_m) = A_0(t) + A_1(t)\varphi_1,$$

where $\varphi_1 = \varphi_1(x_4, w_m)$ is scalar and stochastic and where $A_0(t)$ and $A_1(t)$ [defined in (58) and (59)] are deterministic and T-periodic matrices. Since φ_1 is a scalar,

$$\begin{aligned}\mathrm{Cov}(\tilde{x}(t+1)) = {} & A_0(t)\mathrm{Cov}(\tilde{x}(t))A_0(t)^T + A_1(t)\mathrm{Cov}(\tilde{x}(t)) \\ & \times A_1(t)^T\mathrm{Cov}(\varphi_1) + BB^T\mathrm{Cov}(\epsilon_1),\end{aligned}$$

which is a deterministic and periodic matrix difference equation without a trivial closed-form solution.

To proceed, vectorize the matrix difference equations describing the dynamics of $\mathrm{Cov}(\tilde{x}(t))$. If X is any matrix, then $\mathrm{Vec}(X)$ denotes the vectorization of X formed by stacking the columns of X into a single column vector. It is well known from the theory of matrices [24] that, whenever the matrix dimensions of V, Y, and W match, then

$$(W^T \otimes V)\mathrm{Vec}(Y) = \mathrm{Vec}(VYW),$$

where (in the case of a 2×2 matrix V), the Kronecker product $\otimes$ is defined by

$$V \otimes W = \begin{bmatrix} V_{11}W & V_{12}W \\ V_{21}W & V_{22}W \end{bmatrix}.$$

Using Kronecker product notation, the dynamics of the covariance are

$$\begin{aligned}\mathrm{Vec}(\mathrm{Cov}(\tilde{x}(t+1))) = {} & (A_0(t) \otimes A_0(t) + A_1(t) \otimes A_1(t)\,\mathrm{Cov}(\varphi_1)) \\ & \times \mathrm{Vec}(\mathrm{Cov}(\tilde{x}(t))) + \mathrm{Vec}(BB^T)\mathrm{Cov}(\epsilon_1),\end{aligned}$$

or, simply,

$$\mathrm{Vec}(\mathrm{Cov}(\tilde{x}(t+1))) = \mathcal{A}(t)\mathrm{Vec}(\mathrm{Cov}(\tilde{x}(t))) + \mathrm{Vec}(BB^T)\mathrm{Cov}(\epsilon_1).$$

The dynamics are of the form $\mathcal{X}(t+1) = \mathcal{A}(t)\mathcal{X}(t) + \mathcal{B}(t)$, where $\mathcal{A}(t)$ and $\mathcal{B}(t)$ are bounded T-periodic matrices, which is a standard linear periodic difference equation.

Since $\mathcal{A}(t)$ and $\mathcal{B}(t)$ are T-periodic, invoke Floquet theory in discrete time, which states that the system is asymptotically stable with a unique limit cycle if and only if the eigenvalues of $\bar{\mathcal{A}}(1) := \mathcal{A}(T)\mathcal{A}(T-1)\ldots\mathcal{A}(1)$ are strictly inside the unit circle.

The covariance of the performance signal $[u(t), e(t)]^T$ is, in a similar manner, easily shown to satisfy

$$\mathrm{Cov}\left(\begin{bmatrix} u(t) \\ e(t) \end{bmatrix}\right) = \mathrm{E}(C(x_4)\mathrm{Cov}(\tilde{x}(t))C(x_4)^T) + DD^T\mathrm{Cov}(\epsilon_2(t)),$$

where $\mathrm{Cov}(\epsilon_2(t))$, at the steady state, is a constant and given by Lemma 3.

Decompose $C(x_4)$ in a deterministic and stochastic component as

$$C(x_4) = C_0 + C_1 x_4,$$

where

$$C_0 = \begin{bmatrix} -1, & 1 \\ -1/c_I, & 0 \end{bmatrix}, \tag{61}$$

$$C_1 = \begin{bmatrix} -1, & 1 \\ 0, & 0 \end{bmatrix}. \tag{62}$$

Since x_4 is a scalar,

$$\begin{aligned}\mathrm{Cov}\left(\begin{bmatrix} u(t) \\ e(t) \end{bmatrix}\right) = {} & C_0\mathrm{Cov}(\tilde{x}(t))C_0^T + C_1\mathrm{Cov}(\tilde{x}(t))C_1^T\mathrm{Cov}(x_4(t)) \\ & + DD^T\mathrm{Cov}(\epsilon_2(t)),\end{aligned}$$

where the steady-state values of $\mathrm{Cov}(x_4(t))$ and $\mathrm{Cov}(\epsilon_2(t))$ are constant. Therefore, $\mathrm{Cov}([u(t), e(t)]^T)$ is globally asymptotically stable and with a unique limit cycle if and only if $\mathrm{Cov}(\tilde{x}(t))$ is globally asymptotically stable and with a unique limit cycle. This completes the proof. ■

Similar to what holds for the expected trajectory and $\bar{A}(1)$, it is, in general, not possible to derive a simple formula for the eigenvalues of $\bar{\mathcal{A}}(1)$ as a function of parameters a_2, a_1, c_I, $h(t)$, σ_u, and σ_m. However, it is straightforward for any specific set of parameters (and with the help of a computer) to evaluate the product $\mathcal{A}(T)\mathcal{A}(T-1)\ldots\mathcal{A}(1)$, and then compute the eigenvalues of this matrix. It is an insightful exercise to complete this task in practice on the computer. Note, for $\mathcal{A}(t)$, that parameters a_1 and c_I always appear together as a product, and the impact of σ_u and σ_m is dictated by the value of $\mathrm{Cov}(\varphi_1)$. This reduces the degrees of freedom during the analysis of the dynamics. It is still necessary to consider combinations of the quadruple a_2, a_1c_I, $h(t)$, and $\mathrm{Cov}(\varphi_1)$ to gain a complete understanding of the stability of any configuration.

Example: The Stability Region of the Covariance Trajectory for Sinusoidal Seasonality

Consider a sinusoidal plant seasonality

$$h(t) = 0.7\sin\left(\frac{2\pi t\Delta}{24}\right),$$

where $\Delta = 5/60$ h, and consider four values of the covariance of φ_1

$$\mathrm{Cov}(\varphi_1) = 0,\ 0.03,\ 0.1,\ \text{and}\ 0.2.$$

For each value of $\mathrm{Cov}(\varphi_1)$, compute the eigenvalues of $\bar{\mathcal{A}}(1)$ for 40,000 different combinations of a_2 and a_1c_I and determine the stability of each combination. The result is shown using heat maps in Figure 8. Blue indicates instability, and yellow indicates stability.

The heat maps have two regions. In the blue region, all combinations of a_2 and a_1c_I lead to instability while, in the yellow region, all combinations correspond to stability. It is desirable to operate in the yellow region and

at a sufficiently large distance from any unstable configuration.

Note how the size of the blue unstable region grows as $\mathrm{Cov}(\varphi_1)$ increases. This matches our intuition that a higher level of stochastic process noise $w_m(t)$ or persistent-excitation noise $w_u(t)$ increases the chance of instability. △

A closed-form expression of the covariance of the trajectory (and an analytical expression of the stability criteria for the general case) is outside the scope of this article. However, for the special case $a_2 = 0$, Theorem 4 can be refined as follows. Since a_2 is small in a typical real implementation, this scenario has practical value.

Theorem 5

If $a_2 = 0$, then the limit cycles of $\mathrm{Cov}(x(t))$ and $\mathrm{Cov}(z(t))$ are globally asymptotically stable if and only if $\prod_{i=1}^{T}((1-a_1c_I\varphi_0(i))^2 + (a_1c_I\varphi_0(i))^2\mathrm{Cov}(\varphi_1)) < 1$.

Proof

As shown in the proof of Theorem 4, $\mathrm{Cov}(\tilde{x}(t))$ satisfies

$$\mathrm{Vec}(\mathrm{Cov}(\tilde{x}(t+1))) = \mathcal{A}(t)\mathrm{Vec}(\mathrm{Cov}(\tilde{x}(t))) + \mathrm{Vec}(BB^T)\mathrm{Cov}(\epsilon_1).$$

According to Floquet, $\mathrm{Cov}(\tilde{x}(t))$ has a globally asymptotically stable limit point if and only if the eigenvalues of $\bar{\mathcal{A}}(1) \coloneqq \mathcal{A}(T)\mathcal{A}(T-1)\ldots\mathcal{A}(1)$ are strictly inside the unit circle. It follows from (57) that the product of two consecutive matrices $\mathcal{A}(t)$ is

$$\begin{aligned}\mathcal{A}(\tau+1)\mathcal{A}(\tau) = (&A_0(\tau+1)\otimes A_0(\tau+1) + A_1(\tau+1)\\ &\otimes A_1(\tau+1)\mathrm{Cov}(\varphi_1(t+1)))\cdot(A_0(\tau)\otimes A_0(\tau)\\ &+ A_1(\tau)\otimes A_1(\tau)\mathrm{Cov}(\varphi_1(t))).\end{aligned}$$

Consider next the steady state, where $\mathrm{Cov}(\varphi_1)$ is time invariant and constant. Expand the right-hand side to obtain

$$\begin{aligned}\mathcal{A}(\tau+1)\mathcal{A}(\tau) = &(A_0(\tau+1)\otimes A_0(\tau+1))(A_0(\tau)\otimes A_0(\tau))\\ &+(A_0(\tau+1)\otimes A_0(\tau+1))(A_1(\tau)\otimes A_1(\tau)\mathrm{Cov}(\varphi_1))\\ &+(A_1(\tau+1)\otimes A_1(\tau+1)\mathrm{Cov}(\varphi_1))(A_0(\tau)\otimes A_0(\tau))\\ &+(A_1(\tau+1)\otimes A_1(\tau+1))(A_1(\tau)\otimes A_1(\tau))(\mathrm{Cov}(\varphi_1))^2.\end{aligned}$$

It is well known that, for any matrices X, Y, V, W with matching dimensions, $(X\otimes Y)(V\otimes W) = (XV)\otimes(YW)$ [24]. Use this relationship to express $\mathcal{A}(\tau+1)\mathcal{A}(\tau)$ as

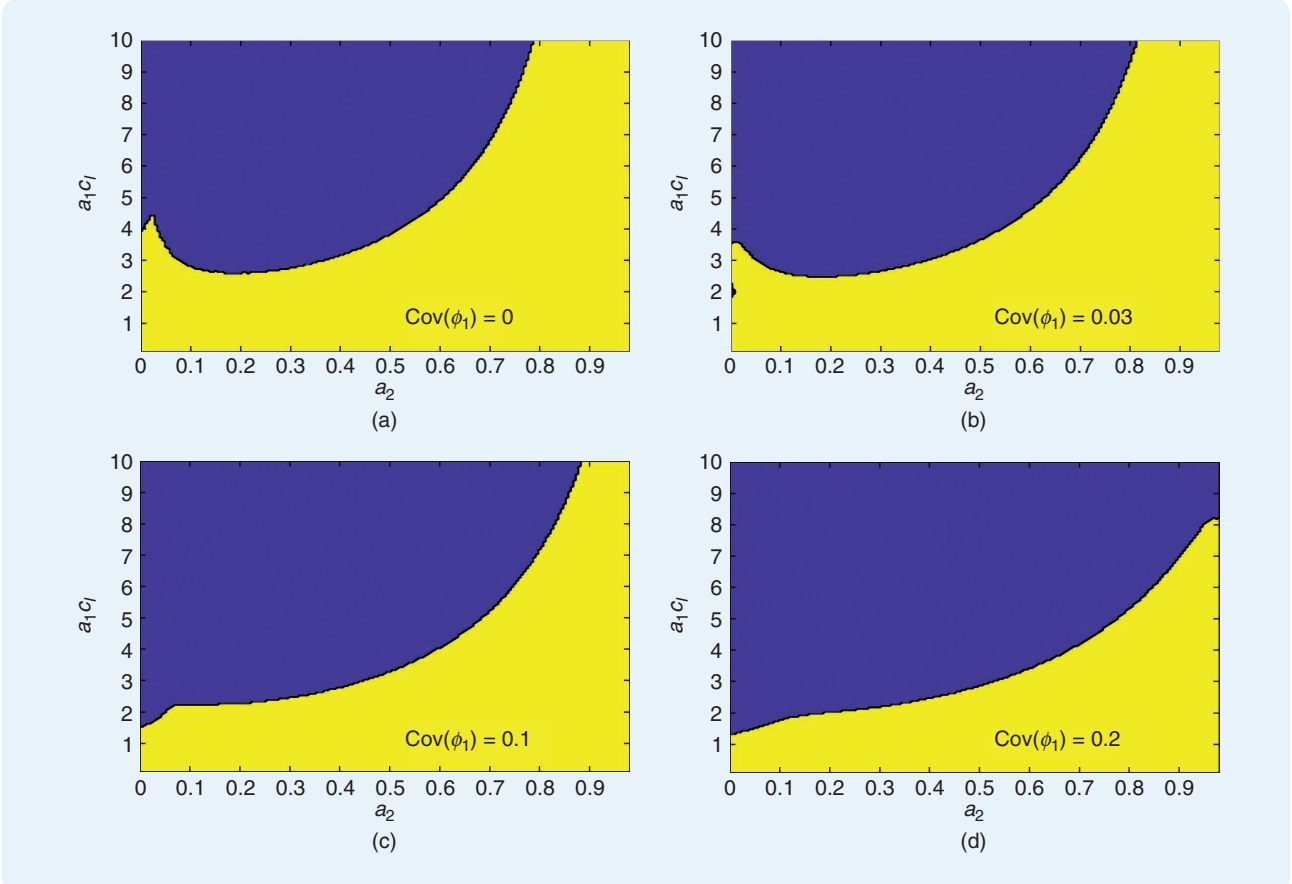


FIGURE 8 The stability region of the covariance trajectory for sinusoidal seasonality. The charts indicate the stability (yellow) and instability (blue) of the covariance limit cycle for different combinations of a_2, a_1c_I, and $\mathrm{Cov}(\varphi_1)$ when $h(t) = 0.7\sin(2\pi t\Delta/24)$ and $\Delta = 5/60$ h.

$$\begin{aligned}\mathcal{A}(\tau+1)A(\tau)=&(A_0(\tau+1)A_0(\tau))\otimes(A_0(\tau+1)A_0(\tau))\\&+(A_0(\tau+1)A_1(\tau))\otimes(A_0(\tau+1)A_1(\tau))\mathrm{Cov}(\varphi_1)\\&+(A_1(\tau+1)A_0(\tau))\otimes(A_1(\tau+1)A_0(\tau))\mathrm{Cov}(\varphi_1)\\&+(A_1(\tau+1)A_1(\tau))\otimes(A_1(\tau+1)A_1(\tau))(\mathrm{Cov}(\varphi_1))^2.\end{aligned}$$

Consider the fact that $a_2=0$. Evaluate the conventional matrix products $A_0(\tau+1)A_0(\tau)$, $A_0(\tau+1)A_1(\tau)$, $A_1(\tau+1)A_0(\tau)$ and $A_1(\tau+1)A_1(\tau)$. It is straightforward to show that

$$\begin{aligned}A_0(\tau+1)A_0(\tau)&=A_0(\tau+1)(1-a_1c_I\varphi_0(\tau)),\\A_0(\tau+1)A_1(\tau)&=-A_0(\tau+1)a_1c_I\varphi_0(\tau),\\A_1(\tau+1)A_0(\tau)&=A_1(\tau+1)(1-a_1c_I\varphi_0(\tau)),\\A_1(\tau+1)A_1(\tau)&=-A_1(\tau+1)c_I\varphi_0(\tau).\end{aligned}$$

Substitute these four matrix products in the expression for $\mathcal{A}(\tau+1)\mathcal{A}(\tau)$ to obtain

$$\begin{aligned}\mathcal{A}(\tau+1)\mathcal{A}(\tau)=&(A_0(\tau+1)\otimes A_0(\tau+1))(1-a_1c_I\varphi_0(\tau))^2\\&+(A_0(\tau+1)\otimes A_0(\tau+1))(a_1c_I\varphi_0(\tau))^2\mathrm{Cov}(\varphi_1)\\&+(A_1(\tau+1)\otimes A_1(\tau+1))(1-a_1c_I\varphi_0(\tau))^2\mathrm{Cov}(\varphi_1)\\&+(A_1(\tau+1)\otimes A_1(\tau+1))(a_1c_I\varphi_0(\tau))^2(\mathrm{Cov}(\varphi_1))^2.\end{aligned}$$

This expression can immediately be simplified as

$$\begin{aligned}\mathcal{A}(\tau+1)\mathcal{A}(\tau)=&(A_0(\tau+1)\otimes A_0(\tau+1)+A_1(\tau+1)\\&\otimes A_1(\tau+1)\mathrm{Cov}(\varphi_1))\cdot((1-a_1c_I\varphi_0(\tau))^2\\&+(a_1c_I\varphi_0(\tau))^2\mathrm{Cov}(\varphi_1)).\end{aligned}$$

By inspection, $\mathcal{A}(\tau+1)$ is recognized inside this expression; hence,

$$\begin{aligned}\mathcal{A}(\tau+1)\mathcal{A}(\tau)=&\mathcal{A}(\tau+1)((1-a_1c_I\varphi_0(\tau))^2\\&+(a_1c_I\varphi_0(\tau))^2\mathrm{Cov}(\varphi_1)).\end{aligned}$$

Repeating this for $\tau=1,2,\ldots,T-1$ yields

$$\bar{\mathcal{A}}(1)=\mathcal{A}(T)\prod_{i=1}^{T-1}((1-a_1c_I\varphi_0(i))^2+(a_1c_I\varphi_0(i))^2\mathrm{Cov}(\varphi_1)).$$

Next, compute the eigenvalues of $\mathcal{A}(T)$ by going through the painful (but straightforward) mechanical labor of solving $\det(\tilde{\lambda}I-\mathcal{A}(T))=0$. The determinant is

$$\det(\tilde{\lambda}I-\mathcal{A}(T))=\tilde{\lambda}^3(\tilde{\lambda}-(1-a_1c_I\varphi_0(T))^2-(a_1c_I\varphi_0(T))^2\mathrm{Cov}(\varphi_1)).$$

The four eigenvalues of $\mathcal{A}(T)$ are

$$\begin{aligned}\tilde{\lambda}_{1,2,3}&=0,\\\tilde{\lambda}_4&=(1-a_1c_I\varphi_0(T))^2+(a_1c_I\varphi_0(T))^2\mathrm{Cov}(\varphi_1).\end{aligned}$$

Since $\bar{\mathcal{A}}(1)$ equals $\mathcal{A}(T)$ times a scalar, eigenvalue λ_j of $\bar{\mathcal{A}}(1)$ equals eigenvalue $\tilde{\lambda}_j$ of $\mathcal{A}(T)$ times the same scalar, such that

$$\lambda_j=\tilde{\lambda}_j\prod_{i=1}^{T-1}((1-a_1c_I\varphi_0(i))^2+(a_1c_I\varphi_0(i))^2\mathrm{Cov}(\varphi_1)).$$

The four eigenvalues of $\bar{\mathcal{A}}(T)$, therefore, are

$$\begin{aligned}\lambda_{1,2,3}&=0,\\\lambda_4&=\prod_{i=1}^{T}((1-a_1c_I\varphi_0(i))^2+(a_1c_I\varphi_0(i))^2\mathrm{Cov}(\varphi_1)).\end{aligned}$$

The eigenvalues at the origin are always inside the unit circle and, therefore, stable. Hence, the system is asymptotically stable if and only if λ_4 is inside the circle. Since the system is linear and periodic, stability implies a unique limit cycle. Specifically, $\mathrm{Cov}(x(t))$ and $\mathrm{Cov}(z(t))$ have globally asymptotically stable unique limit cycles if and only if $\prod_{i=1}^{T}((1-a_1c_I\varphi_0(i))^2+(a_1c_I\varphi_0(i))^2\mathrm{Cov}(\varphi_1))<1$, which completes the proof. ■

Corollary 2

A sufficient condition for the global asymptotic stability in Theorem 5 is

$$0<a_1c_I<\frac{2}{(1+\mathrm{Std}(\varphi_1))(1+\max_i h(i))},\tag{63}$$

and $\mathrm{Std}(\varphi_1)<1$, where $\mathrm{Std}(\varphi_1)$ is the steady-state standard deviation of φ_1.

Proof

Assume (63) is true. Since $h(t)>-1$ for all t, it follows that

$$0<a_1c_I<\frac{2}{(1+\mathrm{Std}(\varphi_1))(1+h(t))},\quad\text{for all } t.\tag{64}$$

Furthermore, $a_2=0$ implies that $\varphi_0(t):=(1-a_2)(1+h(t))=1+h(t)$. Substitute $1+h(t)$ with $\varphi_0(t)$ in (64) to obtain

$$0<a_1c_I<\frac{2}{(1+\mathrm{Std}(\varphi_1))\varphi_0(t)},\quad\text{for all } t.\tag{65}$$

Multiply the right-hand inequality of (65) by $(1+\mathrm{Std}(\varphi_1))\varphi_0(t)$ and subtract one to obtain

$$a_1c_I\varphi_0(t)-1+a_1c_I\varphi_0(t)\mathrm{Std}(\varphi_1)<1.\tag{66}$$

Next, multiply the left-hand inequality of (65) by $(-1+\mathrm{Std}(\varphi_1))\varphi_0(t)$ (which is smaller than zero, since $\mathrm{Std}(\varphi_1)<1$) and add one to obtain

$$1>1-a_1c_I\varphi_0(t)+a_1c_I\varphi_0(t)\mathrm{Std}(\varphi_1).\tag{67}$$

Combine (66) and (67) into

$$|1-a_1c_I\varphi_0(t)|+a_1c_I\varphi_0(t)\mathrm{Std}(\varphi_1)<1.\tag{68}$$

It follows from the triangle inequality that

$$(1-a_1c_1\varphi_0(t))^2+(a_1c_1\varphi_0(t))^2\mathrm{Cov}(\varphi_1)<1. \tag{69}$$

Since (69) is true for all t, it follows that

$$\prod_{i=1}^{T}((1-a_1c_1\varphi_0(i))^2+(a_1c_1\varphi_0(i))^2\mathrm{Cov}(\varphi_1))<1,$$

which is the condition for global asymptotic stability in Theorem 5 and, therefore, completes the proof. ■

PLANT IDENTIFICATION

The stability of the closed-loop system depends on a_1c_1, a_2, and $h(t)$, where c_1 is a design parameter, and a_2 and $h(t)$ are approximately known a priori. Plant gain a_1, which must be estimated online, remains to be determined. Plant model (14) may be rewritten as

$$\tilde{y}(t)=a_0+a_1u(t-1)+\tilde{w}_m(t-1),$$

where

$$\tilde{y}(t)=\frac{y(t)-a_2y(t-1)}{(1-a_2)(1+h(t-1))},$$
$$\tilde{w}_m(t-1)=(a_0+a_1u(t-1))w_m(t-1),$$

and $w_m(t)\sim\mathrm{WN}(0,\sigma_m^2)$. Conditioned on the control signal sequence $u(t)$, variable $\tilde{y}(t)$ is a sequence of uncorrelated random variables with $\mathrm{E}(\tilde{y}(t)\,|\,u(t-1))=a_0+a_1u(t-1)$ and $\mathrm{Var}(\tilde{y}(t)\,|\,u(t-1))=(a_0+a_1u(t-1))^2\sigma_m^2$. If estimates of a_0 and a_1 are available and denoted $\hat{a}_0$ and $\hat{a}_1$, respectively, then a prediction of $\tilde{y}(t)$ is given by $\hat{\tilde{y}}(t)=\hat{a}_0+\hat{a}_1u(t-1)$. Furthermore, $\tilde{w}_m(t)\sim\mathrm{WN}(0,(a_0+a_1u(t))^2\sigma_m^2)$.

Assume $u(t)$ operates in a 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 by minimizing the discounted sum of the squares of the residuals, $\tilde{y}(t)-\hat{\tilde{y}}(t)$, based on measurements $u(i),y(i),i=1,\ldots,t$.

Let

$$\begin{bmatrix}\hat{a}_0(t)\\ \hat{a}_1(t)\end{bmatrix}=\underset{\hat{a}_0,\hat{a}_1}{\operatorname{argmin}}\sum_{i=0}^{t-1}\lambda^i(\tilde{y}(t+1-i)-a_0-a_1u(t-i))^2, \tag{70}$$

which can be solved using the standard *recursive least squares* estimator, where the exponential memory loss allows for time-varying plant parameters [25], [26]. The recursive implementation with $\theta:=[a_0,a_1]^T$ is outlined in Algorithm 2. It follows that $\hat{\tilde{y}}(t):=\hat{a}_0(t)+\hat{a}_1(t)u(t-i)$ is an unbiased, λ-weighted, minimum-variance prediction of $\tilde{y}(t)$.

To better appreciate the need for persistent excitation, it is insightful to consider the closed-form solution of (70), which is

$$\begin{bmatrix}\hat{a}_0(t)\\ \hat{a}_1(t)\end{bmatrix}=\begin{bmatrix}\sum_{i=0}^{t-1}\lambda^i, & \sum_{i=0}^{t-1}\lambda^iu(t-i)\\ \sum_{i=0}^{t-1}\lambda^iu(t-i), & \sum_{i=0}^{t-1}\lambda^iu(t-i)^2\end{bmatrix}^{-1}$$
$$\times\begin{bmatrix}\sum_{i=0}^{t-1}\lambda^i\tilde{y}(t-i)\\ \sum_{i=0}^{t-1}\lambda^iu(t-i)\tilde{y}(t-i)\end{bmatrix}.$$

If $u\to$ constant, then the inverted 2×2 matrix converges to a rank-deficient matrix that cannot be inverted. As a result, the identifiability of a_0 and a_1 is lost. Since estimating the slope of a response curve $g(u)$ based on samples of $g(u)$ requires data points associated with multiple u values, the result is easy to understand. Of practical importance is that the competitive landscape and impression supply, which dictate the response curve, are time varying. To detect and identify a time-varying linearization, it is necessary to excite the system on an ongoing basis.

Algorithm 2 may be enhanced by adding, for example, estimator windup protection or by implementing a bank of estimators where the covariance matrix P for each estimator is reset at different time points [25]. However, with sufficient excitation, these improvements are not critical.

The interconnected system consisting of the plant and adaptive controller is shown in Figure 9. Recall the control objective, which is to achieve a small $e(t)$ and a constant $u(t)$.

SIMULATION RESULTS

The challenges of the control problem and how they are addressed by the proposed solution are now illustrated with the help of a simulated example. In this example, the control system is implemented with discrete time updates every $\Delta=5/60$ h.

Plant dynamics and process noise are defined by

$$y(t+1)=a_2y(t)+(1-a_2)(1+h(t))g(u(t))(1+w_m(t)),$$

where

$$g(u)=20\mathbb{I}_{\{u\geq1\}}+60\mathbb{I}_{\{u\geq2.5\}}+70\mathbb{I}_{\{u\geq3\}}+1000\mathbb{I}_{\{u\geq4\}}+100\mathbb{I}_{\{u\geq5\}}+400\mathbb{I}_{\{u\geq8\}},$$
$$a_2=0.85,$$
$$h(t)=0.8\sin\left(\frac{2\pi t\Delta}{24}\right),$$
$$w_m(t)\sim\mathrm{Gaussian}(0,0.2^2).$$

ALGORITHM 2 Plant identification.

Parameters: $\lambda, a_2, h(t), \hat{\theta}_0, P_0$
Input: $u(t), y(t)$
State: $\hat{\theta}, P$
Initialization ($t=0$):
$\hat{\theta}(0)=\theta_0$
$P(0)=P_0$
For each instant of time, $t=1,2,\ldots$, compute:
$\tilde{y}(t)=(y(t)-a_2y(t-1))/((1-a_2)(1+h(t-1)))$
$\varphi(t)=[1,u(t-1)]^T$
$G(k)=P(t-1)\varphi(t)/(\lambda+\varphi^T(t)P(t-1)\varphi(t))$
$P(t)=(I-G(t)\varphi^T(t))P(t-1)/\lambda$
$\hat{\theta}(t)=\hat{\theta}(t-1)+G(t)(\tilde{y}(t)-\varphi^T(t)\hat{\theta}(t-1))$

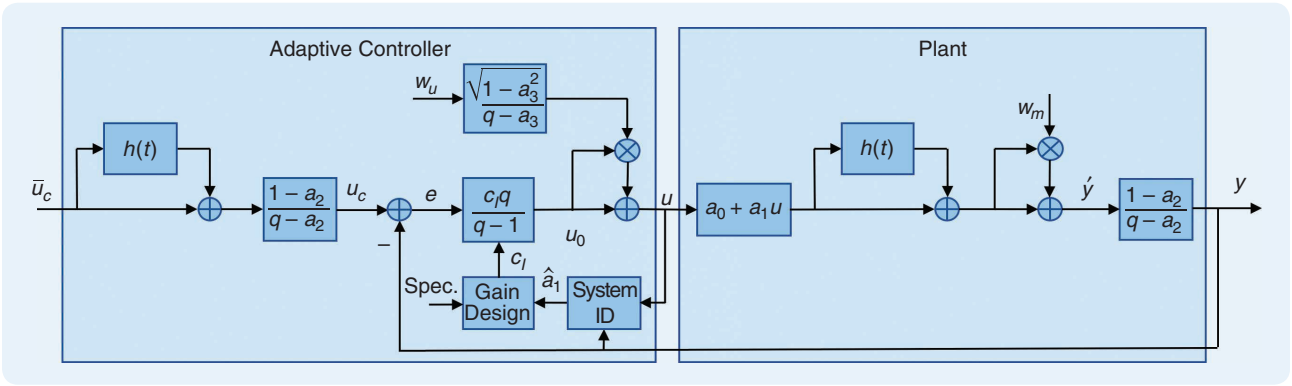


FIGURE 9 A block diagram of an adaptive closed-loop system. The figure depicts the interconnected system in (27)–(30), and q represents the forward-shift operator [for example, $qx(t) = x(t+1)$]. It consists of a time-periodic, linear dynamic plant with multiplicative process noise; a pure integral (I)-error feedback controller with artificially generated random excitation perturbations; a time-periodic dynamic feedforward controller; a system-identification module (such as a recursive least squares algorithm); and a gain design system that computes the integral gain.

The latency parameter $a_2 = 0.85$ corresponds to a plant time constant of $T_P = 0.5$ h.

Figure 10(a) shows the response curve for uncertainty signals $u_u = 0, 0.05$ and 0.1. The black staircase function represents the original response curve with no plant smoothing. Figure 10(b), on the other hand, shows the implied effective response sensitivity (plant gain) for $u_u = 0.05$ and 0.1. The slope dg/du in the case of $u_u = 0$ is not drawn; however, it equals zero everywhere except for at the discrete control signal values $u = 1, 2.5, 3, 4, 5$, and 8, where the slope is undefined (it is infinite). Pay attention to how the smoothness of both g and dg/du increase with increasing value of the bid uncertainty.

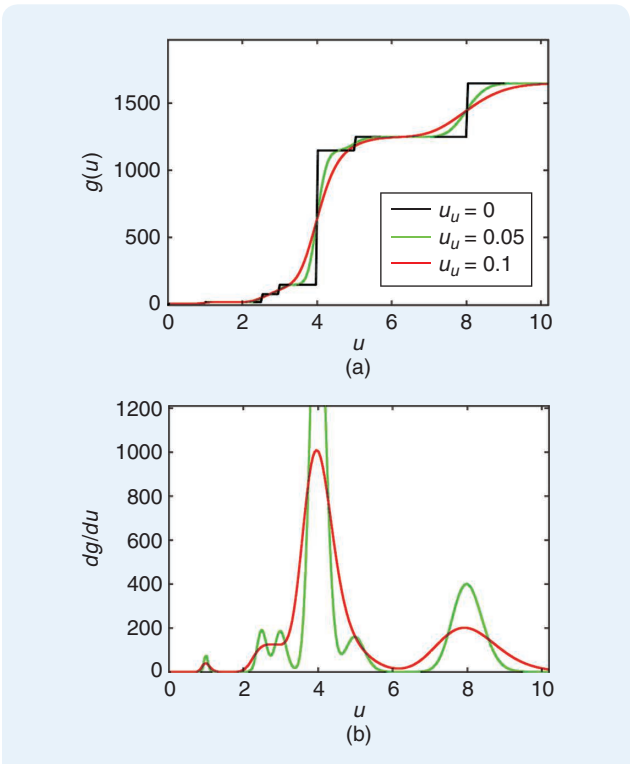


FIGURE 10 A response curve and plant smoothing. (a) The response curve for uncertainty signal $u_u = 0, 0.05$, and 0.1. The black staircase function ($u_u = 0$) represents the original response curve with no plant smoothing. (b) The implied effective response sensitivity (also known as plant gain) for $u_u = 0.05$ and 0.1.

The campaign is set to run for 300 h, and the advertiser has specified a desired delivery rate represented by a setpoint signal $\bar{u}_c$ defined by

$$\bar{u}_c = \begin{cases} 450 & \text{if } t\Delta \le 100, \\ 100 & \text{if } 100 < t\Delta \le 200, \\ 450 & \text{if } 200 < t\Delta. \end{cases}$$

Given this setpoint signal and comparing it with $g(u)$ (see Figure 10), the optimal control signal to an oracle is known to be approximately $u = 4$ whenever $t\Delta \le 100$ or $200 < t\Delta$, and $u = 3$ whenever $100 < t\Delta \le 200$. The precise value, and whether an equilibrium exists, depends on the value of the uncertainty bid. Moreover, the response curve is unknown to the control system and cannot be used to determine the control signal.

Latency parameter a_2 and seasonality function $h(t)$ are assumed known to the control system. In practice, they can often be estimated ahead of time at a reasonably high accuracy. Indeed, $h(t)$ can be estimated favorably based on the delivery of historical ad campaigns and Internet traffic patterns, while a_2 can be estimated based on the latency profiling of ad servers and data flows.

The control system relies exclusively on observations $y(t)$ to identify the plant and to power the error feedback controller. Let the desired *loop gain* be $a_1 c_l = 0.1$, which is well within the stable regime for the expected trajectory and its covariance, as indicated in the "Example: The Stability Region of the Expected Trajectory for Sinusoidal Seasonality" and "Example: The Stability Region of the Covariance Trajectory for Sinusoidal Seasonality" sections for any reasonable value of σ_u. Recall that $\mathrm{Cov}(\varphi_1) = \sigma_u^2 + \sigma_m^2 + \sigma_m^2\sigma_u^2$,

which impacts the stability regime of the state trajectory covariance. The selected loop gain is also well within the bounds prescribed by Corollaries 1 and 2 for the special case of $a_2 = 0$, which, arguably, is a more conservative case.

Moreover, a_1 is known only as an estimate computed by Algorithm 2, and the error feedback controller in Algorithm 1 is configured with c_I. Hence, the controller gain is implemented as $c_I(t) = 0.1/\hat{a}_1(t)$, where $\hat{a}_1(t)$ is the most up-to-date estimate of the plant gain.

Integral Control With and Without Plant Smoothing

Consider, first, pure I control with and without Heisenberg bidding-based plant smoothing and without excitation control. Two values of the uncertainty bid are compared: $u_u = 0$ (which is the no-plant-smoothing scenario) and $u_u = 0.05$ (which is plant smoothing with a bid randomization having 5% relative standard deviation). The absence of excitation control is configured by setting the standard deviation σ_u of the persistent excitation to zero.

One simulation of the closed-loop system for each of the two configurations is shown in Figure 11. Four quantities are drawn as time series. In Figure 11, (a) displays the control signal u, (b) the spend rate y, (c) the control error e, and (d) the plant gain (true and estimated plant gain). The blue and magenta curves correspond to $u_u = 0$ and $u_u = 0.05$, respectively. The green curve in the spend rate plot shows $u_c(t)$, which is the output of the feedforward controller

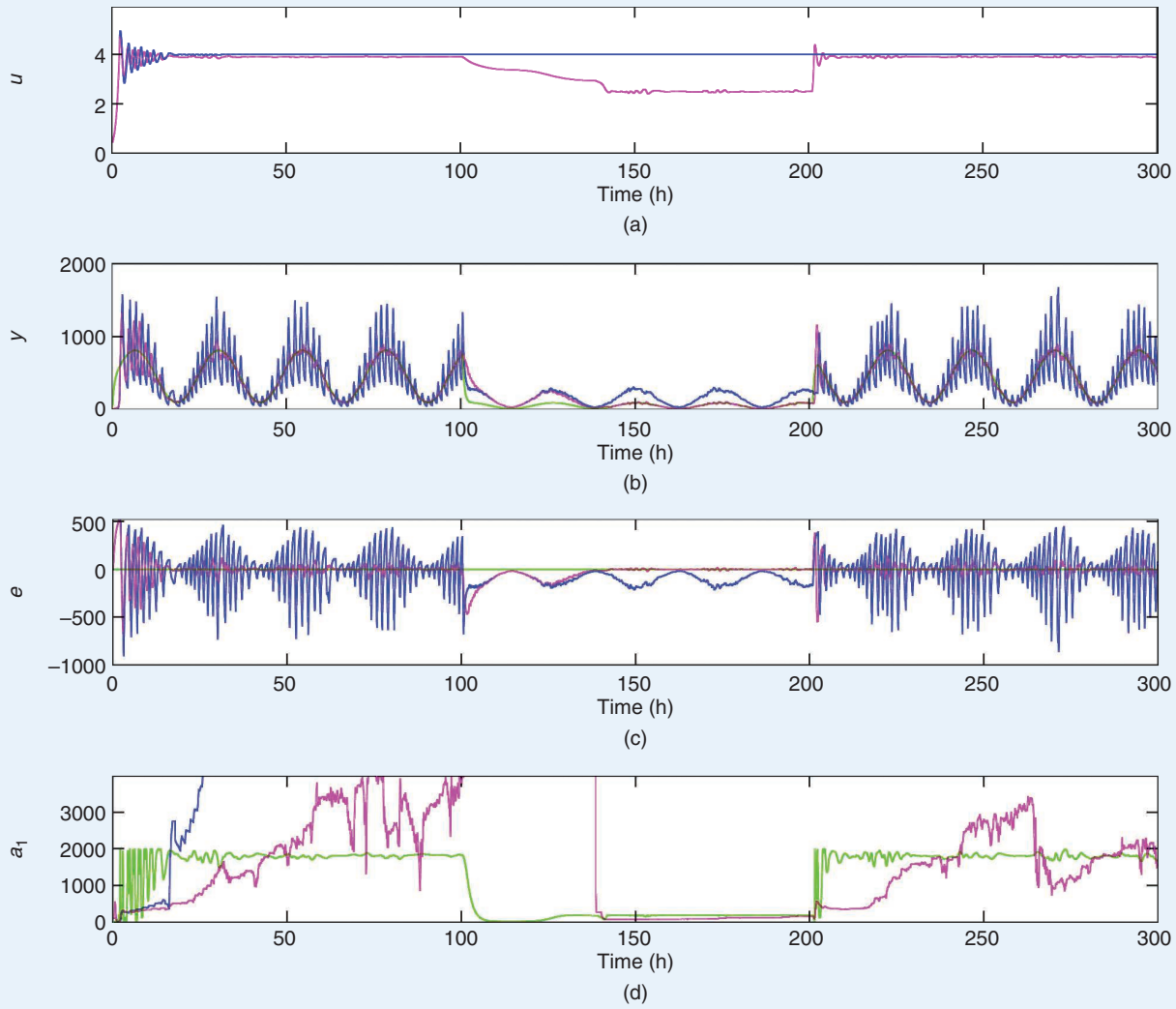


FIGURE 11 Closed-loop performance using pure integral control with and without plant smoothing (without excitation control). The blue and magenta curves correspond to bid uncertainty $u_u = 0$ and $u_u = 0.05$, respectively. (a) The control signal u. (b) The spend rate y and feedforward-adjusted setpoint signal u_c (green). (c) The control error e. (d) The plant gain estimate $\hat{a}_1$ and true plant gain a_1 (green) associated with the magenta controller. The plant gain estimation is poor for both controller configurations but impossible without plant smoothing. Alternately, pure integral control prevents a highly volatile control signal despite the challenging conditions for feedback control. Overall, none of the controller configurations are adequate.

and, also, the expected spend pattern corresponding to an optimal and constant control signal. In Figure 11(d), the green curve represents the true effective plant gain in the case $u_u = 0.05$.

Note how none of the plant gain estimates in Figure 11(d) converge. The estimate corresponding to $u_u = 0$ (blue) quickly diverges out of sight and never returns, while the estimate corresponding to $u_u = 0.05$ (magenta) is highly volatile although appears to be unbiased. This demonstrates the impact of poor excitation and a challenging response curve. The system without plant smoothing is particularly challenging, since the response curve from the controller's perspective is, then, a discontinuous staircase function. The system with plant smoothing is struggling. However, it at least manages to keep the plant gain estimate bounded. In this case, the true plant gain is always defined (as shown in the green curve). However, without persistent excitation, the plant identification is unreliable.

For the control signals in Figure 11(a), it is first noted that none of them diverge. The credit partially goes to the use of pure I control, which computes the feedback control signal $u(t)$ as a value proportional to the sum of all feedback errors $e(t)$ from the beginning of the campaign. The responsiveness, or lack thereof, is understood by remembering that the controller gain is computed as $c_I(t) = 0.1/\hat{a}_1(t)$. Hence, if $\hat{a}_1(t)$ is abnormally large, then the controller gain becomes tiny, and the control system slows.

Initially, the plant gain estimates for both $u_u = 0$ and $u_u = 0.05$ are small, which explains the volatile control signal for both controllers in the first 10 h of the campaign. After the setpoint signal is changed at times $t = 100$ and $t = 200$, it is noted that the control signal corresponding to $u_u = 0$ (blue) visibly does not respond at all. As shown in the figure, its plant gain estimate, by then, has diverged to an extremely large value, making the controller gain approximately zero. The plant gain estimate for the case $u_u = 0.05$ (magenta) is too large but not as large as for the case $u_u = 0$. Consequently, the control signal is more responsive but still sluggish.

The spend rate time series in Figure 11(b) demonstrates the dramatic effect of using plant smoothing. Without bid randomization ($u_u = 0$), the closed-loop system enters an operating mode where the spend oscillates around the desired spend pattern (see the blue and green curves). In fact, there exists no equilibrium in this case, since the desired spend is between two steps in the staircase response curve. For any fixed control signal value, the spend rate is either too large or too small, which forces the control signal to oscillate with a small (and almost invisible), high-frequency amplitude. Since no proportional or derivative action is used in the feedback controller, the control signal does not exhibit large oscillations. This demonstrates an advantage of using pure I control.

Moreover, the practical implication of the observed behavior is worse than illustrated in this basic simulation because it may trigger a ripple effect among other bidders in the network. Indeed, the market consists of many bidders. If several bidders experience a highly volatile impression allocation (or spend rate), then this triggers volatile bid prices (which are observed as volatile highest competing bid prices by other bidders). This leads to an even higher volatility in the impression allocation and spend rate, which, then, leads to a further increased volatility in bid prices and competing bid prices. Overall, in a network of noncooperating bidders, it poses a serious risk of instability.

The spend rate with bid randomization ($u_u = 0.05$) is shown by the magenta curve. It is well behaved throughout most of the scenario. Only in the time interval between times $t = 100$ and $t = 140$ is the spend rate significantly different from the desired rate. This shows that bid randomization improves the closed-loop behavior dramatically but not enough. In a real setting, both the setpoint signal and the response curve are likely to change over time. They also may change not only at distinct rare time points but also steadily over time. The response curve may change in a step-like fashion when competing bidders enter or exit the market or smoothly as a result of gradually evolving competing bids. This behavior calls for a plant gain estimation that is robust, fast, and ongoing. It also emphasizes that persistent excitation of the system is important.

Integral Control With and Without Excitation Control

Control with Heisenberg bidding-based plant smoothing is configured with an uncertainty bid $u_u = 0.05$. Two configurations of persistent excitation are evaluated: one without excitation ($\sigma_u = 0$) and one with excitation, defined by $\sigma_u = 0.02$ and $a_3 = 0.19$. The latter configuration corresponds to a control signal excitation with a standard deviation of 2% and a persistent-excitation time constant of $T_{PE} = 0.05$ h. Note, if the plant latency (modeled by a_2) is known accurately, then a smaller value of a_3 (and T_{PE}) can be used. With a perfect latency model, it is reasonable to use $a_3 = 0$ for maximum excitation. Recognize that the randomization related to Heisenberg bidding (plant smoothing) is applied per individual bid, while the randomization related to persistent excitation is applied once per sampling period.

Figure 12 shows an example of closed-loop behavior using the two configurations of the persistent-excitation controller. As in Figure 11, Figure 12 displays (a) the control signal u, (b) the spend rate y, (c) the control error e, and (d) the plant gain (true and estimated plant gain). This time, the magenta curves correspond to $\sigma_u = 0$, while black curves correspond to $\sigma_u = 0.02$ and $a_3 = 0.19$. The green curve in the spend rate plot shows $u_c(t)$, which is the output of the feedforward controller and, also, the expected spend pattern corresponding to an optimal control signal. In Figure 12(d), the green curve represents the true efficient plant gain in the case of $u_u = 0.05$, $\sigma_u = 0.02$, and $a_3 = 0.19$.

The rapidly evolving advertising industry, which includes some of the world's largest and most valuable companies, has many open control problems.

The magenta curves in this figure ($u_u = 0.05$ and $\sigma_u = 0$) are equivalent to the magenta curves in Figure 11 and used as a benchmark of the black curves involving excitation.

Note in Figure 12(d) the dramatic improvement of the plant gain estimation when control signal excitation is used. The plant gain estimate converges to a value close to the true plant gain within fewer than 24 h. After the setpoint adjustments at hours 100 and 200 when the operating point along the response curve is altered, the plant identification responds convincingly and adequately updates the plant gain estimate.

The improved plant gain estimation is achieved at the cost of a low-amplitude, approximately WN pattern in the control signal, as shown in Figure 12(a). Thanks to the much improved plant gain estimate, the spend rate in the middle subplot is also improved. Specifically, the response to the setpoint adjustments at 100 h and 200 h is dramatically enhanced. The improved control is also illustrated in Figure 12(c), where error e for the black curve is more responsive after the setpoint change at time 100 h.

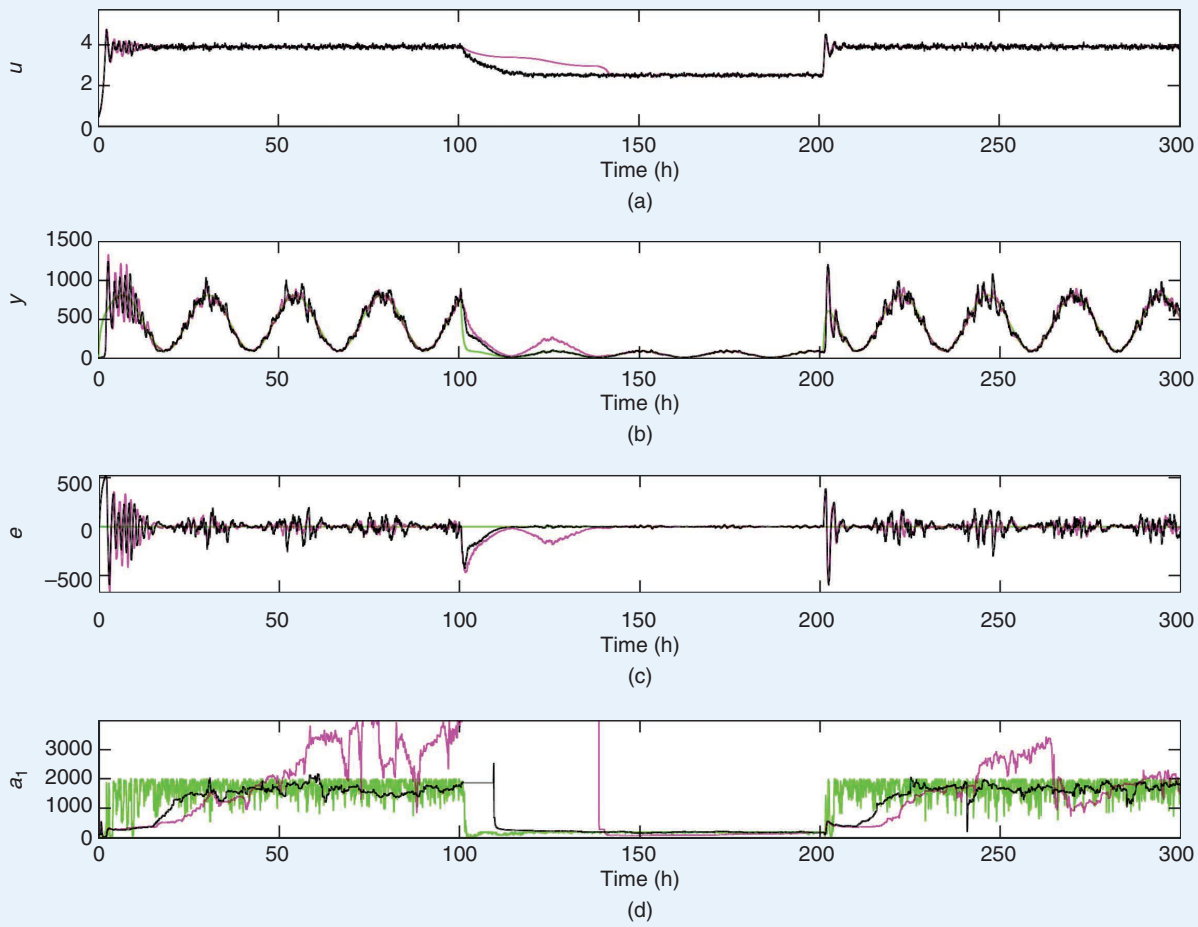


FIGURE 12 Closed-loop performance using pure integral control and plant smoothing, with and without excitation control. The magenta curves correspond to no excitation control ($\sigma_u = 0$), and the black curves correspond to excitation control defined by $\sigma_u = 0.02$ and $a_3 = 0.19$. (a) The control signal u. (b) The spend rate y and feedforward-adjusted setpoint signal u_c (green). (c) The control error e. (d) The plant gain estimate $\hat{a}_1$ and true plant gain a_1 (green) associated with the black controller. The plant gain estimation is adequately responsive and robust only with excitation control. With excitation control, the control signal stays in close proximity to its optimal value, and the spend rate operates near its optimal limit cycle. Overall, the pure integral controller with plant smoothing with persistent-excitation control is doing well.

Discontinuities and other nonlinearities, periodicities, stochastic and deterministic noise, and coupling effects are the norm for control problems related to online advertising.

This example has, hopefully, convinced the reader about the complementary benefits of Heisenberg bidding-based plant smoothing, persistent-excitation control, and pure I-error feedback control. Heisenberg bidding is the first line of defense against the challenges encountered when using feedback control in online advertising. It turns a discontinuous response curve continuous and guarantees the existence of a fixed-point solution of the closed-loop system. It is also a first step toward making plant gain estimation possible. Without bid randomization, the plant gain (strictly speaking) equals zero everywhere except for in a set of measure zero, where the plant gain is undefined.

Persistent-excitation control is the second line of defense and ensures the response curve is continuously explored in a small neighborhood of the operating point. This (in contrast to many standard techniques in adaptive estimation, such as conditional updates and covariance resets [25]) ensures the control system is never operating blindfolded but always probing the plant for changes.

Finally, pure I control is the third line of defense and reduces the sensitivity to challenging process noise and further improves the robustness to a challenging response curve. Note that an I controller is a low-pass filter and, therefore, blocks high-frequency noise. It reduces the sensitivity to steps in the response curve.

A benefit of Heisenberg bidding-based plant smoothing and persistent-excitation control not illustrated in the previous example is how they improve the response to a gradually evolving competitive landscape. Indeed, both of these mechanisms (in slightly different ways) help detect if a step in the response curve is gradually approaching from above or below. Without plant smoothing first and foremost (but, to some extent, also without persistent excitation of the control signal), an approaching step in the response curve may trigger massive spikes in the spend rate.

FINAL REMARKS ON CONTROLLER IMPLEMENTATION

As noted before, the response curve function $g(u)$ is generally unknown, and the controller has access only to the control signal $u(t)$ and noisy spend measurements $y(t)$. Since the control objective involves driving $u(t)$ toward a constant and $g(u)$ is highly nonlinear, it is not possible, based on the described measurements, to estimate the global shape of the response curve. In some cases, however, an estimate of the global response curve is available. It may describe $g(u)$ with a varying degree of certainty at different values of u, and it may be most accurate in the neighborhood of where the controller has operated.

If a response curve estimate is available and expressed as a function $g(u) = \Sigma_j Y_j^a \mathbb{I}_{\{u \geq U_j\}}$ for steps indexed $j = 1, 2, \ldots$ [where U_j, Y_j^a are known in the Bayesian sense as $U_j = \text{Gamma}(1/\sigma_j^2, 1/(u_j\sigma_j^2))$ and $\text{E}Y_j^a = y_j^a$], then the results in "The Mathematics of Plant Smoothing" can be leveraged. For example 1) Theorem S3 can be used to select an initial control signal $u(0)$ before feedback data are available, 2) Theorem S5 can be leveraged to select a suitable uncertainty bid u_u (for example, as the smallest value for which dg/du does not change dramatically for small adjustments of u), or 3) Theorem S5 can be used to produce a plant gain estimate (which may be used alone or in combination with a plant gain estimate produced by a recursive plant gain estimator, as proposed in this article).

DISCUSSION

This article introduced optimization, modeling, and control in online advertising with members of the control community as the target audience. The introduction started with a setup and the process of impression allocation. It continued with an in-depth treatment of a specific advertising optimization problem to highlight the challenges and opportunities of feedback control. The goal was to uncover why and how feedback is used and bring attention (and, hopefully, an interest) to this application of control engineering.

The rapidly evolving advertising industry, which includes some of the world's largest and most valuable companies, has many open control problems. The problems range from higher-performant and more robust estimation and control algorithms for partially solved problems to control algorithms for novel multiobjective optimization problems. As the online advertising industry advances and advertisers become more savvy, the business requirements become more elaborate and stricter. This leads to a growing list of interesting and challenging control problems.

A natural immediate extension of the results in this article is to develop online learning algorithms for the seasonality and latency models $h(t)$ and a_2, respectively. These plant properties were assumed known in this article. However, in reality, they are only approximately known. Poor

estimates hurt the closed-loop performance and robustness. Other relevant extensions are robustness margins to uncertain plant parameters or a sound algorithm to select u_u, the control signal uncertainty.

Other problems in the proximity of what was solved in this article relate to the optimal bidding and control for campaigns when the ad inventory is sold based on a first-price-cost model. Optimal bidding on first-price inventory is challenging. However, it is a timely problem to work on, since the industry trend is toward this cost model.

Discontinuities and other nonlinearities, periodicities, stochastic and deterministic noise, and coupling effects are the norm for control problems related to online advertising. Hopefully, this article has convinced the reader that these challenges can be managed systematically and analytically. The key is to use a sound blend of mathematics and physical reasoning from first principles. It also helps to think outside the box. If the plant is too challenging to control as is, consider how to modify the plant to make the control problem less difficult.

AUTHOR INFORMATION

Niklas Karlsson (nkarlsson8@gmail.com) received the M.S. degree in engineering physics from Lund University in 1995 and the M.A. degree in statistics and applied probability and the Ph.D. degree in mechanical engineering (with a specialization in control theory, dynamic systems, and robotics) from the University of California, Santa Barbara (UCSB) in 1997 and 2002, respectively. He also graduated from the Stanford Executive Program in 2017. He joined Evolution Robotics in Pasadena, California, in 2002, where he was the principal investigator of navigation and control algorithms. During his tenure at Evolution Robotics, he invented, among other things, a commercial vSLAM technology, which is used as the brain of an autonomous vacuum cleaner. In 2005, he joined Advertising.com. He joined Verizon Media by way of acquisition. He is now the chief scientist and vice president of R&D for Verizon Media's demand-side platforms, where he creates and implements the research vision around feedback control, artificial intelligence, and big data to online advertising. He is the inventor of 36 issued patents in the areas of mobile autonomous robotics and online advertising; he received the Distinguished Alumni Award from the Department of Mechanical Engineering at UCSB in 2015 and the Master Inventor Award from Verizon Media in 2017. He is a Senior Member of the IEEE.

REFERENCES

[1] J. Enberg, "Global digital ad spending 2019: Digital accounts for half of total media ad spending worldwide," eMarketer, 2019. Accessed on: Nov. 11, 2019. [Online]. Available: https://www.emarketer.com/content/global-digital-ad-spending-2019

[2] L. Fisher, "US programmatic ad spending forecast 2019: Nearly half of programmatic ad dollars now go to video," eMarketer, 2019. Accessed on: Nov. 11, 2019. [Online]. Available: https://www.emarketer.com/content/us-programmatic-ad-spending-forecast-2019

[3] N. Karlsson and J. Zhang, "Applications of feedback control in online advertising," in *Proc. American Control Conf. (ACC '13)*, Washington, D.C., June 17–19, 2013, pp. 6008–6013. doi: 10.1109/ACC.2013.6580779.

[4] N. Karlsson, "Control problems in online advertising and benefits of randomized bidding strategies," *Eur. J. Control*, vol. 30, pp. 31–49, July 2016. doi: 10.1016/j.ejcon.2016.04.007.

[5] N. Karlsson, "Systems and methods for controlling bidding for online advertising campaigns," U.S. Patent Application 2010/0262497 A1, Apr. 10, 2009.

[6] N. Karlsson, "Adaptive control using Heisenberg bidding," in *Proc. American Control Conf. (ACC '14)*, Portland, OR, June 4–6, 2014, pp. 1304–1309. doi: 10.1109/ACC.2014.6859107.

[7] J. Guo and N. Karlsson, "Model reference adaptive control of advertising systems," in *Proc. American Control Conf. (ACC '17)*, Seattle, WA, May 24–26, 2017, pp. 5482–5487. doi: 10.23919/ACC.2017.7963807.

[8] W. Zhang, "Optimal real-time bidding for display advertising," Ph.D. thesis, Dept. of Computer Science, Univ. College London, U.K., 2016. Accessed on: Jan. 2, 2020. [Online]. Available: https://discovery.ucl.ac.uk/id/eprint/1496878/1/weinan-zhang-phd-2016.pdf

[9] W. Zhang, Y. Rong, J. Wang, T. Zhu, and X. Wang, "Feedback control of real-time display advertising," in *Proc. 9th ACM Int. Conf. Web Search and Data Mining (ACM '18)*, New York, 2018, pp. 407–416. doi: 10.1145/2835776.2835843.

[10] V. Mardanlou, N. Karlsson, and J. Guo, "Statistical plant modeling and simulation in online advertising," in *Proc. American Control Conf. (ACC '17)*, Seattle, WA, May 24–26, 2017, pp. 2176–2181. doi: 10.23919/ACC.2017.7963275.

[11] N. Karlsson, "Control of periodic systems in online advertising," in *Proc. IEEE 57th Conf. Decision and Control (CDC '18)*, Miami, FL, Dec. 17–19, 2018, pp. 5928–5933. doi: 10.1109/CDC.2018.8619536.

[12] N. Karlsson, "Adaptive optimization and control in online advertising," in *Proc. IEEE 58th Conf. Decision and Control*, Nice, France, Dec. 11–13, 2019, pp. 1513–1518. doi: 10.1109/CDC40024.2019.9028857.

[13] H. He and N. Karlsson, "Identification of seasonality in Internet traffic to support control of online advertising," in *Proc. American Control Conf. (ACC '19)*, Philadelphia, July 10–12, 2019, pp. 3835–3840. doi: 10.23919/ACC.2019.8814710.

[14] N. Karlsson, "Plant gain estimation in online advertising processes," in *Proc. IEEE 56th Conf. Decision and Control (CDC '17)*, Melbourne, Australia, Dec. 12–15, 2017, pp. 2182–2187. doi: 10.1109/CDC.2017.8263968.

[15] J. Fernandez-Tapia, "An analytical solution to the budget-pacing problem in programmatic advertising," *J. Inform. Optim. Sci.*, vol. 40, no. 3, pp. 603–614, 2019. doi: 10.1080/02522667.2017.1303946.

[16] N. Karlsson and Q. Sang, "Event rate control in online advertising," in *Proc. American Control Conf. (ACC '17)*, Seattle, WA, May 24–26, 2017, pp. 4215–4220. doi: 10.23919/ACC.2017.7963603.

[17] Q. Sang, N. Karlsson, and J. Guo, "Feedback control of event rate in online advertising campaigns," *Control Eng. Pract.*, vol. 75, pp. 126–136, June 2018. doi: 10.1016/j.conengprac.2018.03.010.

[18] V. Krishna, *Auction Theory*, 2nd ed. San Diego, CA: Academic Press, 2010.

[19] D. G. Luenberger and Y. Ye, *Linear and Nonlinear Programming*. New York: Springer-Verlag, 2015.

[20] W. R. Thompson, "On the likelihood that one unknown probability exceeds another in view of the evidence of two samples," *Biometrika*, vol. 25, nos. 3–4, pp. 285–294, 1933. doi: 10.1093/biomet/25.3-4.285.

[21] G. Casella and R. L. Berger, *Statistical Inference*, 2nd ed. Belmont, CA: Duxbury, 2001.

[22] I. D. Landau and G. Zito, *Digital Control Systems: Design, Identification and Implementation (Communications and Control Engineering)*. Berlin: Springer-Verlag, 2006.

[23] K. J. Åström and T. Hägglund, *Advanced PID Control*. Research Triangle, NC: International Society of Automation (ISA), 2005.

[24] R. A. Horn, *Topics in Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1991.

[25] K. J. Åström and B. Wittenmark, *Adaptive Control*, 2nd ed. Reading, MA: Addison-Wesley, 1994.

[26] S. Haykin, *Adaptive Filter Theory*, 4th ed. Englewood Cliffs, NJ: Prentice Hall, 2002.

[27] W. W. Bell, *Special Functions for Scientists and Engineers*. New York: Dover, 2004.

[28] P. J. Antsaklis and A. N. Michel, *Linear Systems*. Boston, MA: Birkhäuser, 2006.

[29] S. Bittanti and P. Colaneri, *Periodic Systems: Filtering and Control*, 1st ed. Berlin: Springer-Verlag, 2008.

[30] K. S. Tsakalis and P. A. Ioannou, *Linear Time-varying Systems: Control and Adaptation*. Englewood Cliffs, NJ: Prentice Hall, 1993.