Abstract—We consider the problem of designing optimal online-ad investment strategies for a single advertiser, who invests at multiple sponsored search sites simultaneously, with the objective of maximizing his average revenue subject to the advertising budget constraint. A greedy online investment scheme is developed to achieve an average revenue that can be pushed to within $O(\epsilon)$ of the optimal, for any $\epsilon > 0$, with a tradeoff that the temporal budget violation is $O(1/\epsilon)$. Different from many existing algorithms, our scheme allows the advertiser to asynchronously update his investments on each search engine site, hence applies to systems where the timescales of action update intervals are heterogeneous for different sites. We also quantify the impact of inaccurate estimation of the system dynamics and show that the algorithm is robust against imperfect system knowledge.

I. INTRODUCTION

Over the past years, online advertisement has grown significantly and has become one of the most important advertising approaches. According to Emarketer analysts [1], online advertising spending will reach 46.5 billion dollars in 2013 and is expected to surpass 60 billion dollars in 2016. Among the many online advertising methods, sponsored search has been a very successful mechanism and has served as a major income source for many search engines such as Google and Yahoo!

Due to the popularity of sponsored search, designing efficient algorithms for it has received much attention, and there has been a large body of previous work in this area. These works can roughly be divided into two categories. The first category of work aims at designing optimal algorithms for maximizing the search engines’ revenue. [2] proposes an online ad allocation algorithm based on online matching [3], and proves that it achieves an $1 - 1/e$ competitive ratio. [4] uses a primal-dual approach and designs a scheme that achieves the same competitive ratio. [5] looks at the problem of finding revenue-optimal rules for ranking ads. [6] formulates the problem as a stochastic optimization problem and proposes an algorithm to achieve near-optimal performance. [7] designs algorithms for associating queries with budget constrained advertisers and conducts system implementations. The second category focuses on designing optimal bidding strategies for the advertisers. [8] studies properties of greedy bidding schemes and develops a balanced bidding algorithm. [9] considers a single advertiser trying to maximize the number of clicks given a certain budget and tackles the problem with Markov decision process (MDP). [10] formulates the optimal bidding problem as an MDP with large number of bidders. [11] considers a single advertiser and proposes a greedy bidding scheme based on stochastic approximation. However, we note that the aforementioned works only consider optimizing the performance of the advertiser or the search engine in the ad auction process. Hence, they focus mainly on designing algorithms for systems with a single search engine site. In practice, however, an advertiser can typically utilize multiple sites simultaneously for advertising. In order to maximize his revenue, one must jointly optimize the investments at all sites.

In this paper, we consider a single advertiser who is trying to advertise his product by utilizing a set of sponsored search engines (called sites below), e.g., Google or Yahoo!. At every time, the advertiser decides the advertising configuration at each site, e.g., total budget and maximum pay-per-click, associating ads to different keywords, and the bidding strategies, and decides how much money to allocate to advertising at each search engine. According to the advertiser’s customized setting and the sites’ ad allocation and billing mechanisms, the amount of money at a search engine site will be consumed at a certain rate, and an amount of revenue will be generated to the advertiser. After the money is completely depleted at a site, the advertiser updates his investment at the site and resets his advertising configurations, based on his investment strategy and his observation of the site’s performance. The objective of the advertiser is to find an investment policy to maximize his time average revenue subject to his advertising budget.

This problem captures many features of the online advertising procedure at popular sponsored search engine sites such as Google’s AdWords [12] and Bing ads [13]. However, solving this problem is challenging. First of all, the advertiser has to learn over time about how to allocate his investment across different search engine sites subject to his budget constraint. With the many different features and ad mechanisms of the different sites and the many options of the advertiser, this is a nontrivial task. Second, due to the intrinsic stochastic nature of the online advertising process, the advertising performance as well as the time to re-update the advertising configurations at each site are highly dynamic and can be very different from site to site. This imposes the constraint that the investment strategy must be able to handle asynchronous system operations. However, most existing optimal control algorithms, e.g., [14] [15], require simultaneous update of the control actions and hence do not apply in this case. Finally, optimizing time average metrics defined over time-varying update intervals
typically involves optimizing sum of ratios of functions. This is in general non-convex and problems of this kind are generally very hard to solve.

In order to tackle this problem and to resolve the aforementioned difficulties, we adopt the recently developed Lyapunov technique for renewal systems and asynchronous control [16][17]. This approach has three important components: (1) a virtual deficit queue with carefully defined arrival and service rates as ratios of the corresponding system metrics, (2) a Lyapunov drift defined over variable size time intervals, and (3) a drift-plus-penalty-ratio principle in decision making. These three components enable the development of a simple yet near-optimal asynchronous control scheme called AdInvestment (AI), which allows asynchronous updates of the investments and the advertising configurations. We show that the AI algorithm achieves an average revenue that can be pushed to within $O(\epsilon)$ of the maximum, for any $\epsilon > 0$, with a tradeoff of an $O(1/\epsilon)$ temporal budget violation.

We further investigate the performance of AI with imperfect system knowledge and quantify its impact on the algorithm performance.

This paper is organized as follows: In Section II we state our model and problem formulation. In Section III we derive a characterization of the maximum revenue. Then, we present the dynamic AdInvestment strategy (AI) for the advertiser in Section IV. We investigate the performance loss of AI due to imperfect knowledge of the system dynamics in Section V. Simulation results are presented in Section VI. We conclude the paper in Section VII.

II. SYSTEM MODEL

In this section, we present our system model. We consider a single advertiser who is trying to advertise his products at $N$ sponsored search websites (called sites below), e.g., Google or Bing, denoted by $S = \{S_1, \ldots, S_N\}$ (see Fig. 1). We assume that the system operates in continuous time, i.e., $t \in \mathbb{R}_+$.

Fig. 1. A single advertiser allocates his advertising budget across multiple search engine sites (“G,” “Y” and “B” in this example). The objective of the advertiser is to achieve maximum revenue subject to his advertising budget.

A. Advertiser Actions

The advertiser first decides how much money to invest at each site and specifies the operation configuration, e.g., the maximum pay-per-click, or which ads to play for the keywords. Then, when the money deposited at a site is depleted, the advertiser updates his investment amount and the operation configuration for that search site, and the advertising process continues.

Note that this operating mode is quite common in today’s sponsored search engines, e.g., Google Adwords [12]. Under the advertiser’s investment policy, each site goes through a sequence of asynchronous advertising (and freeze) intervals, defined as the intervals during which the advertiser has nonzero (zero) budget left at that site (Fig. 2 shows an example). In the following, we will conveniently call an advertising interval plus the subsequent freeze interval a frame.

We use $p_n[k]$ to denote the amount of new investment the advertiser makes to search site $S_n$ for the $k^{th}$ advertising interval. Note that this new investment will start a new advertising interval. To allow the advertiser to temporarily suspend his investment at a site, we assume that the advertiser can always adopt a “freeze” option. Under this freeze option, the advertiser will not make further new investment into the search site for some time $T^f_n[k]$ after the advertising interval. We call this period the freeze interval. We assume that for all $k$, the pair $(p_n[k], T^f_n[k])$ is chosen from the feasible (investment, freeze time) set $\mathcal{P}_n$ for all $n$. We assume that $\mathcal{P}_n$ includes the option $p_n[k] = 0$, and the constraints $0 \leq p_n[k] \leq p_{\text{max}}$ and $0 \leq T^f_n[k] \leq T^f_{\text{max}}$. One such example can be $\mathcal{P}_n = [0, p] \times [0, T]$. We also use $m_n[k]$ to denote the operation configuration, for instance, the maximum pay-per-click together with the maximum per-day payment, or the targeted user groups, or what ads to display for the corresponding keywords, or the bidding strategy to use at that site for that frame. We assume that $m_n[k] \in \mathcal{M}_n$, where $\mathcal{M}_n$ is the set of feasible operation configurations at site $S_n$. We assume that both $\mathcal{P}_n$ and $\mathcal{M}_n$ are compact sets.

Given the investment amount and the operating configuration, the corresponding search sites will use their own pricing and ad allocation mechanisms to allocate advertisement opportunities to the advertisers, e.g., the generalized second price auction [18] or the VCG mechanism [19], and to generate revenue for the advertiser, e.g., by bringing visiting traffic or by directly bringing product orders to the advertiser, until the advertiser’s new investment is fully depleted. However, due to the dynamic nature of the advertising process, the outcome at each site will inevitably be time-varying, e.g., the volume of the visiting traffic to the search site and the set of keywords that are queried. Such a dynamic condition
will affect the site’s advertising performance, resulting in different amounts of generated profit to the advertiser and time-varying speeds in depleting the remaining advertisement deposit at the site. To capture this dynamics, we denote the duration taken to deplete the \(k\)th investment at site \(S_n\) by \(T_{\text{ad}}^n[k]\), and assume that \(T_{\text{ad}}^n[k]\) is an independent random variable conditioning on the advertiser’s investment \(p_n[k]\) and the operating configuration \(m_n[k]\). We further assume that there exists a function \(F_n[k] \equiv F_n(p_n[k], m_n[k])\), which maps \((p_n[k], m_n[k])\) to the expected length of \(T_{\text{ad}}^n[k]\), i.e.,

\[
F_n : (p_n[k], m_n[k]) \mapsto \mathbb{E}\{T_{\text{ad}}^n[k]\}.
\]  

The function \(F_n(\cdot, \cdot)\) is assumed to be known to the advertiser. We assume that under any feasible actions, the frame size at each site \(S_n\) is both upper and lower bounded, i.e.,

\[
0 < T_{\text{min}}^n \leq T_{\text{ad}}^n[k] + T_{\text{fr}}^n[k] \leq T_{\text{max}}^n.
\]

We denote \(R_n[k]\) the total revenue the advertiser earns during advertising interval \(k\) from site \(S_n\). Similar to \(T_{\text{ad}}^n[k]\), \(R_n[k]\) is also a random variable which is conditionally independent of others given \(p_n[k]\) and \(m_n[k]\). We also assume that there exists a function \(G_n(\cdot, \cdot) \equiv G_n(p_n[k], m_n[k])\) such that:

\[
G_n : (p_n[k], m_n[k]) \mapsto \mathbb{E}\{R_n[k]\}.
\]  

The function \(G_n(\cdot, \cdot)\) satisfies \(G_n(0, \cdot) = 0\), \(G_n(\cdot, \cdot) \geq 0\), \(G_n(p, m) \leq \nu p\) for some constant \(\nu\), and \(G_n(p, m) < G_{\text{max}}\) for all \(n\) and all \(m, p\). We also assume that it is known to the advertiser.

In the following of the paper, for notational convenience, we define common upper and lower bounds by dropping the index “\(n\)”, i.e., \(T_{\text{max}} = \max_n T_{\text{max}}^n\), \(T_{\text{min}} = \min_n T_{\text{min}}^n\), \(R_{\text{max}} = \max_n R_n\) and \(p_{\text{max}} = \max_n p_n\).

### B. Budget Constraint and Objective

It can be seen from the above that the advertiser’s time average revenue is given by:

\[
\text{Profit}_n^\ast \triangleq \lim_{k \to \infty} \frac{1}{K} \sum_{k=0}^{K-1} R_n[k].
\]  

In practice, advertisers typically also have limited advertising budgets. We model this by requiring that the time average advertising expenditure is no more than some pre-specified budget value \(B_n\), i.e.,

\[
\sum_{n} \lim_{K \to \infty} \frac{1}{K} \sum_{k=0}^{K-1} p_n[k] \leq B_n.
\]

Here the term on the left-hand-side (LHS) is the time average advertising expenditure of the advertiser.

The advertiser’s objective is to find an investment strategy that determines how to make the ad-investments, i.e., choosing \(p_n[k]\), \(T_{\text{fr}}^n[k]\) and \(m_n[k]\), so as to maximize his time average revenue \(\text{Profit}_n^\ast\) subject to the average advertisement budget constraint \(B_n\). Below, we call a strategy that chooses \((p_n[k], T_{\text{fr}}^n[k]) \in \mathcal{P}_n\) and \(m_n[k] \in \mathcal{M}_n\), and satisfies the budget constraint \(B_n\) a feasible strategy. Then, we use \(\text{Profit}_n^\ast\) to denote the maximum time average budget-constrained profit achievable over all feasible strategies.

### C. Discussion of the Model

Our model is general and does not assume specific structures of the system. The operation configuration \(m_n[k]\) can be used to describe the specific bidding policy used during the advertising interval, i.e., automatic bidding or customized bidding scheme, e.g., [1]. In this case, the functions \(G_n(\cdot, \cdot)\) and \(F_n(\cdot, \cdot)\) represent the expected revenue and interval length under such bidding schemes. Therefore, the strategies developed in this paper can indeed be implemented with the existing bidding schemes. Under such general settings, solving this problem is challenging, since the performance metrics in \(4\) and \(5\) are equivalent to sum of ratios of functions and are generally non-convex. Moreover, different search sites may need updates asynchronously, whereas most existing optimal control algorithms [14, 15] require synchronous action updates at all components of the system.

### III. Characterizing Optimal Revenue

In this section, we first present a theorem which characterizes the advertiser’s maximum revenue. The main idea of the theorem is that the maximum budget-constrained revenue can in principle be computed by solving a complex nonlinear optimization problem (typically nonconvex). The theorem will later be used for our analysis. In the theorem, the parameter \(V > 1\) is a constant introduced for our later analysis.

**Theorem 1**: The maximum time average revenue \(\text{Profit}_n^\ast\) satisfies \(V \text{Profit}_n^\ast \leq \Phi^\ast\), where \(\Phi^\ast\) is the optimal value of the following optimization problem:

\[
\text{max} : \Phi \triangleq \sum_{n} \sum_{h=0}^{\infty} \alpha_n^h G_n(p_n^h, m_n^h) \\
\text{s.t.} \sum_{n} \sum_{h=0}^{\infty} \alpha_n^h F_n(p_n^h, m_n^h) + \sum_{n} \alpha_n^h T_{\text{fr}}^h_n \leq B_n \\
\sum_{h=0}^{\infty} \alpha_n^h = 1, \alpha_n^h \geq 0, \forall n
\]

Proof: The proof can be done by using an argument based on Caratheodory’s theorem as in [14]. Omitted for brevity. In Theorem 1, the values \(\{\alpha_n^h\}_{h=0}^{\infty}\) can be viewed as the probability of choosing the investment-freeze-configuration triple \((p_n^h, T_{\text{fr}}^h_n, m_n^h)\) in every frame for search site \(S_n\). With this interpretation, we immediately have the following corollary.

**Corollary 1**: There exists an optimal stationary and randomized investment strategy \(\Pi_S\), which works as follows: for each search site \(S_n\), after each frame, chooses an investment-freeze-configuration triple \((p_n^h, T_{\text{fr}}^h_n, m_n^h)\) with probabilities \(\{\alpha_n^h\}_{h=0}^{\infty}\) independent of other sites. In particular, \(\Pi_S\) achieves the followings:

\[
\sum_{n} V \mathbb{E}\{R_n[k]\} = \Phi^\ast.
\]
\[
\sum_n \mathbb{E}\{p_n[k]\} \leq B_{av}. \quad \Box
\]

However, as discussed above, due to the non-convex nature of problem (6), directly solving for such a randomized policy is very difficult. In the following, we instead develop an online algorithm, which allows the advertiser to asynchronously update his investment profile at different search sites and achieve near-optimal profit. We will see that our algorithm only requires greedy updates of each site and can be implemented efficiently in practice.

IV. ONLINE ALGORITHM FOR ACHIEVING NEAR OPTIMAL REVENUE

In this section, we develop an investment strategy to achieve a near optimal performance for the advertiser. Our algorithm is based on the renewal and asynchronous system optimization technique developed in [16] and [17].

A. The Virtual Deficit Queue

To start, we first define the notion of decision points. A time \( t \) is called a decision point if it marks the beginning of an advertising interval for any site. Then, we use \( \{t_d\}_{d=0}^{\infty} \) to denote the set of decision points. We see then the timeline is divided into disjoint intervals by the decision points. Fig. 2 shows five such points \( \{t_0, t_1, \ldots, t_4\} \). Since each frame will be at least \( T_{\text{min}} \) long, we see that \( \lim_{d \to \infty} t_d = \infty \) w.p. 1.

We use \( k_n(t) \) to denote the number of advertising intervals that site \( n \) has completed up to time \( t \). Note that \( k_n(t) \) also denotes the index of the current advertising interval the search site is in at time \( t \). We denote \( \delta_d \triangleq t_{d+1} - t_d \) the length of the interval between decision points \( t_d \) and \( t_{d+1} \). For our analysis, it is also useful to define the notions of effective budget consumption rate \( a_n[t_d] \), and effective revenue generating rate \( r_n[t_d] \) over the time interval \( [t_d, t_{d+1}) \) for each site \( S_n \) as follows:

\[
a_n[t_d] \triangleq \frac{p_n[k_n(t_d)]}{F_n[k_n(t_d)] + T_n^{n}[k_n(t_d)]} \quad (7)
\]

\[
r_n[t_d] \triangleq \frac{r_n[k_n(t_d)]}{F_n[k_n(t_d)] + T_n^{n}[k_n(t_d)]} \quad (8)
\]

Here \( a_n[t_d] \) and \( r_n[t_d] \) roughly denote the average budget consumption rate and the average revenue generation rate. Note that we use the expected value of the advertising interval length in both (7) and (8).

Now we define the deficit queue process \( Q(t_d) \), to keep track of the budget deficit due to temporal violation of the average budget constraint. \( Q(t_d) \) evolves according to the following dynamics:

\[
Q(t_{d+1}) = \max \left[ Q(t_d) - \mu[t_d], 0 \right] + A[t_d], \quad \text{with} \quad Q(t_0) = 0.
\]

Here \( A[t_d] \) denotes the aggregate budget consumed by all the sites during \( [t_d, t_{d+1}) \), i.e.,

\[
A[t_d] = \delta_d \sum_n a_n[t_d]. \quad (10)
\]

Similarly, \( \mu[t_d] \) denotes the amount of “allowed” average budget consumption, i.e.,

\[
\mu[t_d] = B_{av}(t_{d+1} - t_d) = \delta_d B_{av}. \quad (11)
\]

We notice that since the update of \( Q(t_d) \) uses expected lengths of the actual advertising intervals, \( Q(t_d) \) is an approximation of the true deficit. However, we will show later that guaranteeing the stability of \( Q(t_d) \) also ensures the actual budget constraint (5).

For notational convenience, let us also define \( c_{\text{max}} = T_{\text{max}} \max \left[ \sum_n \frac{\mu[t_d]}{T_n^{n}}, B_{av} \right] \). We see then \( A[t_d], \mu[t_d] \leq c_{\text{max}} \) for all \( t_d \).

B. The Dynamic Ad-Investment (AI) Algorithm

Now we present our investment algorithm. In the algorithm, \( S(t_d) \) denotes the set of search engine sites that just finish a frame and will start a new advertising interval at time \( t_d \). The algorithm works as follows:

**Ad-Investment (AI):** At each time \( t_d \), denote \( k_n = k_n(t_d) \) the frame index of \( S_n \) at time \( t_d \). For each search site \( S_n \in S(t_d) \), observe \( Q(t_d) \) and perform the following:

1) Choose \( p_n[k_n], m_n[k_n] \) and \( T_n^{n}[k_n] \) that solve the following:

\[
\max : \quad \Psi_n(t_d) \triangleq \frac{V G_n(p_n[k_n], m_n[k_n]) - Q(t_d) p_n[k_n]}{F_n(p_n[k_n], m_n[k_n]) + T_n^{n}[k_n]} \quad (12)
\]

s.t. \( (p_n[k_n], T_n^{n}[k_n]) \in \mathcal{P}_n, m_n[k_n] \in \mathcal{M}_n \).

Then, update the investment and the advertising configuration at \( S_n \) with the chosen actions.

2) Update \( Q(t_d) \) according to (9).

Denote \( \Psi_n^*(t_d) \) the optimal value of (12). We note that since \( \mathcal{P}_n \) includes the option \( p_n = 0 \) and \( G_n(0, \cdot) = 0 \), at the optimal solution we always have \( \Psi_n^*(t_d) \geq 0 \). We also note that the algorithm involves solving problem (12) for each search site separately. Hence, the advertiser can easily update his configuration at any site when the current frame completes. This is quite different from many existing algorithms that update all control actions simultaneously. Also note that (12) can often be solved efficiently in practice, especially when the actions sets \( \mathcal{P}_n \) and \( \mathcal{M}_n \) are finite. Finally, note that in every step, we use \( F_n(p, m) \) and \( G_n(p, m) \) for decision making. This requires the advertiser to have accurate prior estimation of the system dynamics. As we will see in Section V that AI can also be implemented with imperfect knowledge of the system dynamics, and still achieve good performance.

C. Performance Analysis

Define the Lyapunov function \( L(t_d) = \frac{1}{2}Q(t_d)^2 \). Then, we define a 1-slot Lyapunov drift as follows:

\[
\Delta(t_d) \triangleq L(t_{d+1}) - L(t_d). \quad (13)
\]

Using the queueing dynamic equation (9), we have the following lemma.

**Lemma 1:** At any time \( t_d \), we have:

\[
\Delta(t_d) \leq C_0 - Q(t_d) \delta_d B_{av} - A[t_d], \quad (14)
\]

where \( C_0 \triangleq \frac{1}{2} (T_{\text{max}})^2 \left[ \left( \sum_n \frac{\mu[t_d]}{T_n^{n}} \right)^2 + B_{av}^2 \right] \). \( \Box \)

**Proof:** See Appendix A.

Adding the term \(-V \delta_d \sum_n r_n[t_d] \) to both sides of (14) and rearranging terms, we obtain:

\[
\Delta(t_d) - V \delta_d \sum_n r_n[t_d]
\]
\[ C_0 - Q(t_d)\delta_d B_{av} - \delta_d \sum_n \frac{V G_n[k_n(t_d)] - Q(t_d)p_n[k_n(t_d)]}{F_n[k_n(t_d)] + T_n^a[k_n(t_d)]} \leq C_0 - Q(t_d)\delta_d B_{av} - \delta_d \sum_n \Psi_n(t_d). \]  

From (15), one sees that the Al algorithm is indeed constructed to maximize \( \Psi_n(t_d) \) on the right-hand-side (RHS) of the drift inequality for every \( S_n \in S(t_d) \). The following lemma shows that Al also approximately maximizes \( \sum_n \Psi_n(t_d) \).

**Lemma 2:** Under the Al algorithm, at every time \( t_d \), we have:

\[ \sum_n \psi_{Al}^n(t_d) \geq \max_{\Pi} \left( \sum_n \psi_{\Pi}^n(t_d) \right) - C_1. \]  

Here \( \Pi \) is any feasible investment policy in choosing \( p_n[k_n(t_d)] \), \( m_n[k_n(t_d)] \), and \( T_n^a[k_n(t_d)] \) at time \( t_d \), and \( C_1 = 2T_{max} \sum_n \Delta_{\Pi}^n \) is a \( \Theta(1) \) constant independent of \( V \).

**Proof:** See Appendix B.

Lemma 2 is important for our analysis. In practice, due to the dynamics in the advertising process, sites typically do not complete their frames at the same time. Hence, at every instance of time, only some search sites need to update their configurations. In this case, instead of a synchronous scheme is applied, some sites will have to remain idle while advertising is still going on at others. This will inevitably result in resources being wasted. Lemma 2 shows that Al roughly minimizes the RHS of the drift inequality (15) even with asynchronous updates. This guarantees the near-optimal performance of Al, as summarized in the following theorem.

**Theorem 2:** The Al algorithm achieves the following:

a) The average revenue under Al, denoted by \( \text{Profit}_{av}^{Al} \), satisfies:

\[ \text{Profit}_{av}^{Al} \geq \text{Profit}_{av}^* - \frac{C_1}{V} - \frac{C_0}{VT_{min}}. \]  

Here \( C_0 \) and \( C_1 \) are \( \Theta(1) \) constants defined in Lemmas 1 and 2.

b) At every time \( t_d \), we have:

\[ Q(t_d) \leq V\nu + 2\epsilon_{max}. \]  

Moreover, the average budget constraint (5) is satisfied with probability 1.

**Proof:** See Appendix C.

By tuning the value of \( V \), Theorem 2 implies that Al can indeed achieve a revenue that is within \( O(1/V) \) of the maximum, for any \( V \geq 1 \), with a tradeoff that \( Q(t_d) = O(V) \). Also notice that (18) does not automatically ensure that the actual deficit is deterministically bounded. This is because the actual frame length is a random variable while \( Q(t_d) \) is updated with the expected values. Despite this approximation, Part b) of Theorem 2 shows that so long as \( Q(t_d) \) is stable, (5) will be satisfied with probability 1.

\noindent \textbf{V. PERFORMANCE UNDER IMPERFECT ESTIMATION}

So far, we have assumed that the advertiser knows exactly the expected revenue that will be generated during a frame and the expected duration of the frame. In this section, we consider the more general case where the advertiser only has imperfect estimations of the functions \( G_n \) and \( F_n \), and quantify the impact of such inaccuracy on the performance of the Al algorithm.

Specifically, we assume that now the advertiser does not have the perfect knowledge of \( G_n(p,m) \) and \( F_n(p,m) \). Instead, he uses his own estimated functions \( \hat{G}_n(p,m) \) and \( \hat{F}_n(p,m) \) for algorithm design and decision making. That is, \( Q(t_d) \) and \( \Psi_n(t_d) \) will be defined with \( \hat{G}_n(p,m) \) and \( \hat{F}_n(p,m) \), and (12) will be solved with them as well. In order to quantify the impact of the estimation errors, we first introduce the quality index of the estimations, \( \rho_G \) and \( \rho_F \). That is, we say that the advertiser has an estimation quality \( (\rho_G,\rho_F) \) if for all \( S_n \),

\[ G_n(p,m) - G_n(p,m) \leq \rho_G G_n(p,m), \forall p,m, (19) \]

\[ F_n(p,m) - F_n(p,m) \leq \rho_F F_n(p,m), \forall p,m. (20) \]

Thus, \( \rho_F \) and \( \rho_G \) capture the inaccuracy levels of the estimations. We now have the following theorem regarding the performance of the Al algorithm in this case.

**Theorem 3:** Suppose the advertiser has an estimation quality \( (\rho_G,\rho_F) \) where \( \rho_G,\rho_F \in [0,1) \). Then, the Al algorithm achieves the following:

a) Revenue performance:

\[ \text{Profit}_{av}^{Al} \geq \frac{1 - \rho_F}{(1 + \rho_F)(1 + \rho_G)} \text{Profit}_{av}^* - \left( \frac{1}{V} - \frac{N\rho_G \gamma_m}{(1 + \rho_G)T_{min}} \right). \]

b) Budget control: At every time \( t_d \), we have:

\[ Q(t_d) \leq V(1 + \rho_G)\nu + 2\epsilon_{max}. \]

Here \( \epsilon_{max} \) is defined as the maximum change of \( Q(t_d) \) defined with \( \hat{G}_n(p,m) \) and \( \hat{F}_n(p,m) \), i.e.,

\[ \epsilon_{max} = (1 + \rho_F)T_{max} \max\left[ \sum_n \frac{p_n^{max}}{m} \right] B_{av}. \]

Moreover, the average budget constraint (5) is satisfied with probability 1 if we replace \( B_{av} \) with \( (1 + \rho_F)B_{av} \).

**Proof:** See Appendix D.

We note Part b) of Theorem 3 indicates that the budget constraint may be slightly violated if the estimations of \( G_n \) and \( F_n \) are inaccurate. While such violation is typical in this case, Theorem 3 shows that as long as one uses a scaled budget in the algorithm, i.e., use \( B_{av}/(1 + \rho_F) \), the original advertisement budget can easily be guaranteed. This thus provides useful guidelines for the actual algorithm implementation. Also note that Theorem 3 applies to more general cases where the system’s dynamics and the advertiser’s estimation quality both vary over time. In this case, if the advertiser’s estimation quality stays better than a certain quality \( (\rho_G,\rho_F) \) after some finite learning time, then one can invoke Theorem 3 and conclude the performance of the algorithm.

**VI. SIMULATION**

In this section, we provide simulation results for our algorithm. We consider the case when the system has two sites \( \{S_1,S_2\} \). For each site \( S_n \), we assume for simplicity that \( m_n \) only denotes the maximum pay-per-click.

The \( G_n(p,m) \) and \( F_n(p,m) \) functions are given by:

\[ F_n(p,m) = \frac{k_n p}{m}, \quad G_n(p,m) = \gamma_n \sqrt{\frac{p}{m} m^q}. \]  

(22)
The function \( F_n \) is chosen to model the fact that the expected duration of the advertising interval is proportional to \( \frac{p}{m} \), i.e., roughly the number of times the advertiser wins the ad opportunity. The choice of function \( G_n \) assumes that whenever an ad opportunity arises, the advertiser wins the opportunity with \( m^\theta \) success probability, and that it roughly takes \( p/m \) ad displays to generate \( \sqrt{p/m} \) revenue. For both \( S_n \), we assume that \( m_n = \{0.1, 0.2\} \) and that the feasible (investment-freeze) set is given by:

\[
P_n = \{(p = 0, T = 5), (p = 5, T = 0), (p = 5, T = 5), (p = 10, T = 0), (p = 10, T = 5)\}.
\]

For each action, we assume that the actual revenue is uniformly distributed in \([0.8G_n, 1.2G_n]\). Also, the advertising interval length is uniformly distributed in \([0.8F_n, 1.2F_n]\). We set \( \kappa_1 = 1 \) and \( \kappa_2 = 2; \gamma_1 = 1 \) and \( \gamma_2 = 2; B_{av} = 0.2 \) and we use \( q = 0.2 \). We simulate the algorithm for \( V \in \{5, 10, 20, 50, 100, 200\} \).

First, we see from Fig. 3 that as the \( V \) value increases, the average revenue quickly converges to the optimal when there is no estimation errors. In the imperfect estimation case, we simulate \( \text{Al} \) with a scaled budget \( B_{av}/(1 + \rho F) \). In this case, we see from the right plot that the actual budget never exceeds \( B_{av} = 0.2 \). Also, the left plot shows that the revenue performance of \( \text{Al} \) is always above \( \frac{1 - \rho F}{(1 + \rho F)(1 + \rho G)} \text{Profit}_n^p \) (indicated by the black dotted line), which implies that it is also always no less than \( \frac{1 - \rho F}{B_{av}/(1 + \rho F)} \) fraction of the optimal revenue with budget constraint \( B_{av}/(1 + \rho F) \). These results are consistent with Theorems 2 and 3.

To evaluate the effect of \( V \) on budget violation, Fig. 4 also plots the average deficit queue size. We see that it increases linearly in \( V \). The right plot of Fig. 4 also shows the time when the two sites update their configurations, for the case when \( V = 20 \) and there is no estimation error. It can be seen that under \( \text{Al} \), the site updates take place asynchronously.

VII. CONCLUSION

In this paper, we design optimal investment strategies for a single advertiser who utilizes multiple sponsored search engine sites simultaneously for advertising. An asynchronous investment strategy \( \text{Ad-Investment (Al)} \) is proposed. The algorithm allows the advertiser to asynchronously update his investment configuration at each site. We show that the algorithm achieves an average revenue that is within \( O(\varepsilon) \) of the optimal, for any \( \varepsilon > 0 \), while guaranteeing that the temporal budget violation is \( O(1/\varepsilon) \). We also analyze the performance of \( \text{Al} \) under imperfect system estimation and show that it is robust against estimation errors.

APPENDIX A – PROOF OF LEMMA 1

Here we prove Lemma 1:

**Proof:** Squaring both sides of (9) and using the fact that \((\max[x, 0])^2 \leq x^2 \) for all \( x \in \mathbb{R} \), we have:

\[
Q^2(t_{d+1}) \leq Q^2(t_d) + (\mu[t_d])^2 + (A[t_d])^2 - 2Q(t_d)(\mu[t_d] - A[t_d]).
\]

Now using the facts that \( a_n[t_d] = n_{\text{max}}/T_{\text{min}} \) and \( \delta_d \leq T_{\text{max}} \), we see that at any time \( t_d \),

\[
\frac{1}{2}(A[t_d])^2 + \frac{1}{2}(\mu[t_d])^2 \\
\leq C_0 \leq \frac{1}{2}(T_{\text{max}})^2 \left(\sum_n \frac{n_{\text{max}}}{T_{\text{min}}}^2 + B_{av}^2\right).
\]

Multiplying both sides of (23) by \( \frac{1}{2} \) and using the definition of \( C_0 \), we see that the lemma follows.

APPENDIX B – PROOF OF LEMMA 2

**Proof:** To prove the lemma, we first recall that the \( \Psi_n(t_d) \) function is defined as:

\[
\Psi_n(t_d) = \frac{\text{VG}_n(p_n[k_n], m_n[k_n]) - Q(t_d)p_n[k_n]}{F_n(p_n[k_n], m_n[k_n]) + T_n^p[k_n]}.
\]

Here \( k_n = k_n[t_d] \).

If at time \( t_d \), \( S_n \in S(t_d) \), then \( \Psi_n(t_d) \) is maximized over all possible pure strategies, i.e., strategies that choose a given \((p_n[k], m_n[k], T_n^p[k])\) with probability 1. Now notice that for any \( c_1, d_1, c_2, d_2 \) where \( d_1, d_2 \neq 0 \) and \( \theta \in [0, 1] \), we have:

\[
\frac{\theta c_1 + (1 - \theta)c_2}{\theta d_1 + (1 - \theta)d_2} = \frac{\theta c_1 + (1 - \theta)c_2}{d_1} - \frac{(1 - \theta)c_2}{d_2},
\]

for some \( \eta \in [0, 1] \) [20]. Therefore, if a policy \( \Pi \) mixes two pure strategies \( \Pi_1 \) and \( \Pi_2 \), i.e., uses two triples \((p_1, m_1, 1)\) and \((p_2, m_2, 2)\) with probabilities \( a_1 > 0 \) and \( a_2 > 0 \), where \( a_1 + a_2 = 1 \), then there exists \( \eta \in [0, 1] \) such that:

\[
\Psi_{n}^{\Pi}(t_d) = \eta\Psi_{n}^{1}(t_d) + (1 - \eta)\Psi_{n}^{2}(t_d) \leq \Psi_{n}^{\text{Al}}(t_d).
\]

By induction, we see that \( \Psi_n(t_d) \) is maximized under \( \text{Al} \) over all strategies for all \( S_n \in S(t_d) \).
It remains to show that for any other $S_n \notin \mathcal{S}(t_d)$, the function $\Psi_n(t_d)$ is also approximately maximized. Consider an $S_n \notin \mathcal{S}(t_d)$. Let $\tilde{t}_n$ be the last time the advertiser updates his configuration at $S_n$ before $t_a$, i.e., $S_n \in \mathcal{S}(\tilde{t}_n)$ but $S_n \notin \mathcal{S}(t_d)$ for all $t_n < t_d \leq t_a$. Note that $t_d - t_n \leq T_{\max}$. This is because the maximum frame length of any site $S_n$ is bounded by $T_{\max}$.

Note that at time $t_n$, the site $S_n$ is still operating under the actions chosen at time $\tilde{t}_n$, which is based on the queue size $Q(\tilde{t}_n)$. Denote the actions chosen then by $\tilde{p}_n, \tilde{m}_n$ and $\tilde{T}^*_{\tilde{m}_n}$. The actions at time $t_d$ by solving (12) with the backlog $Q(t_d)$. Note that these actions are not actually implemented because the advertiser will not update the configuration of $S_n$ at time $t_d$. Then, we have:

$$\frac{V G_n(\tilde{p}_n, \tilde{m}_n) - Q(t_d)\tilde{p}_n}{F_n(\tilde{p}_n, \tilde{m}_n) + \tilde{T}^*_{\tilde{m}_n}} = \frac{V G_n(p_n^*, m_n^*, t_d)}{F_n(p_n^*, m_n^*, t_d) + T^*_{m_n^*}} \geq \frac{V G_n(p_n^*, m_n^*, t_d)}{F_n(p_n^*, m_n^*, t_d) + \tilde{T}^*_{\tilde{m}_n}} + \frac{(Q(t_d) - Q(\tilde{t}_n))p_n}{F_n(\tilde{p}_n, \tilde{m}_n) + \tilde{T}^*_{\tilde{m}_n}}$$

$$+ \frac{(Q(t_d) - Q(\tilde{t}_n))\tilde{p}_n}{F_n(\tilde{p}_n, \tilde{m}_n) + \tilde{T}^*_{\tilde{m}_n}} \geq \frac{V G_n(p_n^*, m_n^*, t_d)}{F_n(p_n^*, m_n^*, t_d) + T^*_{m_n^*}}$$

Here the inequality (a) is because $\tilde{p}_n, \tilde{m}_n$ and $\tilde{T}^*_{\tilde{m}_n}$ maximize (12) given $\tilde{t}_n$.

Using the queueing dynamic equation (9), we see that:

$$Q(t_d) - (t_d - \tilde{t}_n)c_{\max} \leq Q(\tilde{t}_n) \leq Q(t_d) + (t_d - \tilde{t}_n)c_{\max}.$$}

Together with the fact that $t_d - t_n \leq T_{\max}$, we have:

$$|Q(t_d) - Q(\tilde{t}_n)| \leq T_{\max}c_{\max}.$$}

Using this in (28), we get:

$$\frac{V G_n(\tilde{p}_n, \tilde{m}_n) - Q(t_d)\tilde{p}_n}{F_n(\tilde{p}_n, \tilde{m}_n) + \tilde{T}^*_{\tilde{m}_n}} \geq \Psi_n(t_d) - C_{1n},$$

where $\Psi_n(t_d)$ is the value of $\Psi_n(t_d)$ with the actions $p_n^*, m_n^*$, $T^*_{m_n^*}$, and $Q(t_d)$, and $C_{1n} \triangleq \frac{2T_{\max}c_{\max}p_n^*}{F_n(p_n^*, m_n^*, t_d) + T^*_{m_n^*}}$. This shows that even for $S_n \notin \mathcal{S}(t_d)$, $\Psi_n(t_d)$ is approximately maximized.

Now by defining $C_1 = \sum_n C_{1n}$, and using a similar argument as in (27), we see that the lemma follows.

**APPENDIX C – PROOF OF THEOREM 2**

Here we prove Theorem 2.

Proof: (Theorem 2) (Revenue Performance) We first prove the revenue performance. To do so, we recall the drift inequality (15) as follows:

$$\Delta(t_d) - V\delta_d \sum_n r_n[t_d] \leq C_0 + C_1 \delta_d - Q(t_d)\delta_d B_{av} - \delta_d \sum_n \Psi_n(t_d).$$

By Lemma 2, we note that $A_1$ maximizes the term $\sum_n \Psi_n(t_d)$ over all possible investment policies for up to a finite constant $C_1$. Thus, we can plug into the RHS of (30) any control algorithm and the following holds:

$$\Delta(t_d) - V\delta_d \sum_n r_n[t_d] \leq C_0 + C_1 \delta_d - Q(t_d)\delta_d B_{av} - \delta_d \sum_n \Psi_{\text{alt}}(t_d).$$

Therefore, we plug the randomized and stationary policy $P_S$ in Corollary 1 into (30) and obtain:

$$\Delta(t_d) - V\delta_d \sum_n r_n[t_d] \leq C_0 + \delta_d C_1 - \delta_d \Phi^*.$$ (32)

Taking expectations on both sides, carrying out a telescoping sum from $d = 0$ to $D - 1$, and rearranging terms, we obtain:

$$V \sum_{d=0}^{D-1} \mathbb{E}\{\delta_d \sum_n r_n[t_d]\} \geq \sum_{d=0}^{D-1} \mathbb{E}\{\delta_d\}(\Phi^* - C_1) - DC_0 - \mathbb{E}\{L(t_0)\}.$$ (33)

Dividing both sides by $V \sum_n \mathbb{E}\{\delta_d\}$ and letting $D \to \infty$, we have:

$$\lim_{D \to \infty} \frac{\sum_{d=0}^{D-1} \mathbb{E}\{\delta_d \sum_n r_n[t_d]\}}{\sum_{d=0}^{D-1} \mathbb{E}\{\delta_d\}} \geq \Phi^*/V - C_1 - \lim_{D \to \infty} V \sum_{d=0}^{D-1} \mathbb{E}\{\delta_d\} \to C_0 - \frac{C_0 D}{V T_{\min}}.$$ (34)

It remains to show that the actual average revenue also satisfies (33). Denote $p_{z,p,m,T}$ the long term average fraction of time that the triple $(p, m, T)$ is used at $S_n$ under $A_1$. Because for any $(p, m)$, $\mathbb{E}\{R_n[k]\}$ and $\mathbb{E}\{T^{ad}[n,k]\}$ are bounded, we use the strong law of large numbers (SLLN) (21) to conclude:

$$\lim_{K \to \infty} \frac{1}{K} \sum_{k=0}^{K-1} \mathbb{E}\{R_n[k]\} = \frac{1}{K} \sum_{p,m,T} p_{z,p,m,T} G_n(p, m) T.$$ (34)

Now fix a search site $S_n$, we see then for the $d$ values that are in the middle of a frame of $S_n$, $r_n[t_d]$ remains unchanged. Let $\tilde{t}_n$ be the last time $S_n$ starts a new frame before $t_d$, and let $\mathcal{J}_n^D(p, m, T)$ be the set of frames during which the triple $(p, m, T)$ is adopted at site $S_n$ up to $t_d$. Then,

$$\sum_{d=0}^{D-1} \mathbb{E}\{\delta_d r_n[t_d]\} = \sum_{p,m,T} G_n(p, m) T \mathbb{E}\{J^{\text{ad}}[p, m] \sum_{j \in \mathcal{J}_n^D(p, m, T)} \mathbb{E}\{T^{\text{ad}}[j] + T^*_{\tilde{m}_n}\}\} + \mathbb{E}\{r_n[t_d](t_d - \tilde{t}_n)\}.$$ (35)

In this case, we have:

$$\lim_{D \to \infty} \sum_{p,m,T} \mathbb{E}\{J^{\text{ad}}[p, m] \sum_{j \in \mathcal{J}_n^D(p, m, T)} \mathbb{E}\{T^{\text{ad}}[j] + T^*_{\tilde{m}_n}\}\} + \mathbb{E}\{r_n[t_d](t_d - \tilde{t}_n)\}.$$ (35)
Since \((t_D - \hat{t}_n) \le T^\text{max}, r_n(t_D) \le \frac{G^\text{max}}{T^\text{max}}\), and \(\lim_{D \to \infty} \sum_{d=0}^{D-1} \mathbb{E}\{\delta_d\} = \infty\), we see that the last term vanishes. Moreover, it can be shown that for all \(S_n\):
\[
\lim_{D \to \infty} \frac{\mathbb{E}\{\sum_{d=0}^{D-1} [J^D_n(p, m, T)]\}}{\sum_{d=0}^{D-1} \mathbb{E}\{\delta_d\}} = \lim_{D \to \infty} \frac{\sum_{p, m, T} z_{p, m, T} G_n(p, m) + T \cdot \mathbb{E}\{\delta_d\}}{\sum_{p, m, T} z_{p, m, T} \mathbb{E}\{\delta_d\}}.
\]

Using this in (35), we have:
\[
\lim_{D \to \infty} \frac{\sum_{n_1} \sum_{d=0}^{D-1} \mathbb{E}\{\delta_d n_1\}}{\sum_{d=0}^{D-1} \mathbb{E}\{\delta_d\}} = \frac{\sum_{p, m, T} z_{p, m, T} \mathbb{E}\{\delta_d\}}{\sum_{p, m, T} \mathbb{E}\{\delta_d\}}.
\]

This and (34) show that the average revenue satisfies:
\[
\text{Profit}^\text{av} \ge \text{Profit}^* = \frac{C_1}{V} - \frac{C_0}{V T^\text{min}}.
\]

(Budget) We now prove that the average budget constraint (5) is ensured under the Al algorithm.

We first prove (18). To see that (13) holds, note that \(c_{\text{max}}\) is the maximum change of \(Q(t_d)\) during any interval \([t_d, t_{d+1})\). Thus, if \(Q(t_d) < V \rho\) at time \(k\), then we must have \(Q(t_{d+1}) \le V \rho + c_{\text{max}}\). On the other hand, suppose now \(Q(t_d) > V \rho\). Without loss of generality, assume that \(Q(t_d)\) is the first time \(Q(t_d) > V \rho\) occurs. One see that \(Q(t_d) \le V \rho + c_{\text{max}}\). Also, we see that from now on until \(Q(t_d)\) drops below \(V \rho\) again, the advertiser will choose \(p_n = \rho\) for any \(S_n\) that requires update during this interval. This is because once \(Q(t_d) > V \rho\), choosing any \(p_n > \rho\) will result in \(\Psi_n(t_d) < 0\). This implies that \(a_n(t_d) = 0\) for all \(S_n \in S(t_d)\) during this time. Hence, the only possible increment of \(Q(t_d)\) will be due to the sites that are in the middle of their frames. \(Q(t_d)\) will also start to decrease once all the active sites complete their current frames. Now it can be seen that this increasing interval will last for at most \(T^\text{max}\) long, and for each active site, \(a_n(t_d) \le \frac{p_n}{T^\text{min}}\). Thus, the maximum possible increment after \(Q(t_d) \ge V \rho\) will be \(c_{\text{max}}\). Combining this fact with \(Q(t_d) \le V \rho + c_{\text{max}}\), we conclude that \(Q(t_d) \le V \rho + 2c_{\text{max}}\) for all \(t_d\).

This shows that the deficit queue with \(A[t_d]\) and \(\rho[t_d]\) being the arrival and service rates is bounded. Using a similar argument as in the revenue case, one can show that:
\[
\sum_{n_1} \sum_{p, m, T} z_{p, m, T} \mathbb{E}\{\delta_d n_1\} \le B_{\text{av}}.
\]

Using SLLN again, one sees that the LHS equals the average expenditure with probability 1. Hence, (5) is satisfied. 

**Appendix D – Proof of Theorem 3**

We prove Theorem 3 here.

**Proof:** (Revenue) First, we prove the revenue performance. Note that the \(Q(t_d)\) value is now updated with \(G_n(p, m)\) and \(F_n(p, m)\). For convenience, we denote this queueing process by \(\hat{Q}(t_d)\), and denote the effective revenue generation rate by \(\hat{r}_n[t_d]\). Then, using the queueing dynamic equation (9), we can similarly obtain:
\[
\Delta(t_d) - V \delta_d \sum_{n_1} \hat{r}_n[t_d] \\
\le C_2 - \hat{Q}(t_d) \delta_d B_{\text{av}} - \delta_d \sum_{n_1} V \hat{G}_n[k_n(t_d)] - \delta_d \sum_{n_1} F_n[k_n(t_d)] + T_n[k_n(t_d)] \\
\le C_2 - \hat{Q}(t_d) \delta_d B_{\text{av}} - \delta_d \sum_{n_1} \Psi_n(t_d).
\]

Here \(C_2 \ge \frac{1}{2}[(1 + \rho F) T^\text{max}]^2 \{[\sum_{n_1} \frac{p_n}{(1 - \rho F) T^\text{max}}]^2 + B_{\text{av}}\}^2\) and \(\Psi_n(t_d) = \Psi_n(t_d)\) with \(G_n(p, m)\) and \(F_n(p, m)\).

Similar to the proof of Theorem 2, we will show that the term \(\sum_{n_1} \Psi_n(t_d)\) is roughly maximized under Al with estimation errors. To do so, note that the advertiser will now choose \(p_n, m_n, T_n\) by solving (12) with \(G_n, F_n,\) and \(Q(t_d)\).

Consider any \(S_n\) and \(t_d\). Let \(t_n\) be the last time when the advertiser updates \(S_n\)’s configuration before \(t_d\). Denote the actions chosen at time \(t_n\) by Al with \(G_n\) and \(F_n\) by \(p_n, m_n,\) and \(T_n\), and denote \(p_n^*, m_n^*,\) and \(T_n^*\) any alternative feasible actions. Also, define \(\sigma_{g_n}(p, m) = G_n(p, m) - G_n(p, m)\) and \(\sigma_{f_n}(p, m) = F_n(p, m) - F_n(p, m)\) the estimation errors of the functions. Then, we have:
\[
\Psi_n(t_d) = V [G_n(p_n, m_n) + \sigma_{g_n}(p_n, m_n)] - \hat{Q}(t_d) p_n \\
\le V [G_n(p_n, m_n) + \sigma_{g_n}(p_n, m_n)] - \hat{Q}(t_d) p_n \\
\le V [G_n(p_n, m_n) + \sigma_{g_n}(p_n, m_n)] - \hat{Q}(t_d) p_n \\
\le V [G_n(p_n, m_n) + \sigma_{g_n}(p_n, m_n)] - \hat{Q}(t_d) p_n \\
\le V [G_n(p_n, m_n) + \sigma_{g_n}(p_n, m_n)] - \hat{Q}(t_d) p_n.
\]

Here inequality (a), we have used:
\[
\Psi_n^*(t_d) - V \rho G^\text{max} + 2T^\text{max} \hat{c}_{\text{max}}^\text{max}, \quad \Psi_n^*(t_d) = \frac{1}{1 + \rho F} \Psi_n^*(t_d) - V \rho G^\text{max} + 2T^\text{max} \hat{c}_{\text{max}}^\text{max}.
\]

In inequality (b), we have used the fact that \(\Psi_n^*(t_d) \ge 0\). Plug (40) and (39) back into (38), we obtain:
\[
\Delta(t_d) - V \delta_d \sum_{n_1} \hat{r}_n[t_d] \le C_2 + \delta_d C_3 \\
\hat{Q}(t_d) \delta_d B_{\text{av}} - \delta_d \frac{1}{1 + \rho F} \sum_{n_1} \Psi_n^*(t_d) + \frac{\delta_d V N \rho G^\text{max}}{(1 - \rho F) T^\text{min}}.
\]
Here $C_3 \triangleq \frac{2NT^{\max}T_{\rho G}^{\max}}{(1-\rho_F)T_{\min}} = \Theta(1)$.

Now plug the stationary and randomized policy $\pi_S$ into (41), we get:

$$\Delta(t_d) - V\delta \sum_n \hat{r}_n(t_d) \leq C_2 + \delta_d C_3$$

$$-\hat{Q}(t_d)\delta_d B_{av} - \delta_d \frac{1}{1+\rho_F} \sum_n \psi_{t_n}(t_d) + \delta_d VN\rho_G G_{\max} \left(1 - \rho_F\right) T_{\min}$$

$$\leq C_2 + \delta_d C_3 + \frac{\delta_d VN\rho_G G_{\max}}{(1-\rho_F)T_{\min}}$$

$$-\hat{Q}(t_d)\delta_d B_{av} - \frac{1}{1+\rho_F} B_{av} - \delta_d \frac{1}{1+\rho_F} V\text{Profit}_{av}$$

$$\leq C_2 + \delta_d C_3 + \frac{\delta_d VN\rho_G G_{\max}}{(1-\rho_F)T_{\min}} + \frac{\delta_d VN\rho_G G_{\max} + \delta_d VN\rho_G G_{\max} + \delta_d VN\rho_G G_{\max}}{(1-\rho_F)T_{\min}}$$

Using an argument similar to the one used in the proof of Theorem 2 one can show that:

$$\lim_{D \to \infty} \frac{1}{D} \sum_{n=0}^{D-1} \mathbb{E} \left[ \delta_d r_n(t_d) \right]$$

$$= \frac{1}{1+\rho_F} \text{Profit}_{av}$$

$$\geq \frac{C_2 + T^{\min}_{C_3}}{V T^{\min}_{C_3}} - \frac{N\rho_G G_{\max}}{(1-\rho_F)T_{\min}}.$$ #

Therefore, the actual average profit satisfies:

$$\text{Profit}_{av} = \frac{1}{N} \frac{\sum_{p,m,T} \tilde{z}^{p,m,T} G_n(p,m)}{(1+\rho_G) \sum_{p,m,T} \tilde{z}^{p,m,T} G_n(p,m)},$$

$$\geq \frac{1}{1+\rho_G} \frac{\text{Profit}_{av}}{(1-\rho_G)}$$

$$\geq \frac{C_2 + T^{\min}_{C_3}}{V T^{\min}_{C_3}} - \frac{N\rho_G G_{\max}}{(1-\rho_F)T_{\min}} + \frac{N\rho_G G_{\max}}{(1-\rho_G)T_{\min}}.$$ #

Here in inequality (a) we have used the facts that $G_n(p,m) \geq \frac{G_n(p,m)}{1+\rho_G}$ and $F_n(p,m) \leq \frac{F_n(p,m)}{1-\rho_G}$ for all $n$ and all $p, m$.

(Budget) We now prove the budget performance. Using a similar argument as in the proof of Theorem 2 we see that $Q(t_d)$ is bounded by $(1+\rho_G)\nu \nu + 2\epsilon_{\max}$. However, different from Theorem 2 this does not imply that the actual expenditure does not violate the budget constraint $B_{av}$. This is so because the estimation errors make $Q(t_d)$ an approximation of the actual expected budget deficit.

To prove our result, we denote $\tilde{z}^{p,m,T}$ the fraction of time the triple $(p,m,T)$ is used under the A1 algorithm with imperfect estimations $\tilde{G}_n(\cdot, \cdot)$ and $\tilde{F}_n(\cdot, \cdot)$. Then, since $Q(t_d)$ is stable, we have:

$$\sum_n \sum_{p,m,T} \tilde{z}^{p,m,T} [F_n(p,m) + T] \leq B_{av}.$$  

Now using the fact that:

$$\sum_n \sum_{p,m,T} \tilde{z}^{p,m,T} [(1 + \rho_F) F_n(p,m) + T]$$

we conclude that:

$$\sum_n \sum_{p,m,T} \tilde{z}^{p,m,T} [F_n(p,m) + T] \leq (1 + \rho_F) B_{av}.$$  

Using a similar argument as in the proof of Theorem 2 it can be shown that the LHS of (43) corresponds to the actual average advertising expenditure. This completes the proof of the theorem.

### References

1. Emarketer.com. Us online advertising spending to surpass print in 2012. [http://www.emarketer.com/Article/US-Online-Advertising-Spending-Surpass-Print-2012/1008783], accessed on July 24 2013.

2. A. Mehta, A. Saberi, U. Vazirani, and V. Vazirani. Adwords and generalized online matching. Journal of the ACM, vol. 54, no. 5, p. 22, 2007.

3. R. Karp, U. Vazirani, and V. Vazirani. An optimal algorithm for online bipartite matching. Proceedings of the 22nd Annual ACM Symposium on Theory of Computing, pp. 352-358, 1990.

4. N. Buchbinder, K. Jain, and J. S. Naor. Online primal-dual algorithms for maximizing a-auctions revenue. Proceedings of the 15th Annual European Symposium, pp. 253-264, 2007.

5. S. Lahaye and D. Pennock. Revenue analysis of a family of ranking constraints: Optimal bidding strategies and equilibria. SSRN working papers, 2012.

6. C. Jiang, C. Beck, and R. Srikant. Bidding with limited statistical knowledge in online auctions. W-PIN+NetEcon: The joint Workshop on Pricing and Incentives in Networks and Systems, June 2013.

7. Google ads. [http://www.google.com/ads/].

8. Bing ads. [https://secure.bingads.microsoft.com/].

9. L. Georgiadis, M. J. Neely, and L. Tassiulas. Resource Allocation and Cross-Layer Control in Wireless Networks. Foundations and Trends in Networking Vol. 1, no. 1, pp. 1-144, 2006.

10. L. Huang, S. Moeller, M. J. Neely, and B. Krishnamachari. Lifto-backpressure achieves near optimal utility-delay tradeoff. IEEE/ACM Transactions on Networking, Volume 21, Issue 3, Pages 831-844, June 2013.

11. M. J. Neely. Stochastic Network Optimization with Application to Communication and Queueing Systems. Morgan & Claypool, 2010.

12. M. J. Neely. Asynchronous scheduling for energy optimality in systems with multiple servers. Proceedings of 46th Annual Conference on Information Sciences and Systems (CISS) (Invited), March 2012.

13. B. Edelman, M. Ostrovsky, and M. Schwarz. Internet advertising and sponsored search auctions: Optimal bidding strategies and equilibria. SSRN working papers, 2012.