

# A Stochastic Approximation Method to Compute Bid Prices in Network Revenue Management Problems

Huseyin Topaloglu

School of Operations Research and Information Engineering,  
Cornell University, Ithaca, New York 14853, USA  
topaloglu@orie.cornell.edu

December 17, 2007

## **Abstract**

We present a stochastic approximation method to compute bid prices in network revenue management problems. The key idea is to visualize the total expected revenue as a function of the bid prices and to use sample path-based derivatives to search for a good set of bid prices. We deal with the discrete nature of the network revenue management setting by formulating a smoothed version of the problem, which assumes that it is possible to accept a fraction of an itinerary request. We show that the iterates of our method converge to a stationary point of the total expected revenue function of the smoothed version. Computational experiments demonstrate that the bid prices obtained by our method outperform the ones obtained by standard benchmark methods and our method is especially advantageous when the bid prices are not recomputed frequently.

The notion of bid prices forms a powerful tool to construct good policies for network revenue management problems. The idea is to associate a bid price with each flight leg that captures the opportunity cost of a unit of capacity. An itinerary request is accepted if and only if there is enough capacity and the revenue from the itinerary request exceeds the sum of the bid prices associated with the flight legs in the requested itinerary; see Williamson (1992) and Talluri and van Ryzin (2004).

Bid prices are traditionally computed by solving a deterministic linear program. However, this linear program only uses the expected numbers of the itinerary requests that are to arrive until the departure time and does not attempt to capture the probability distributions or temporal dynamics of the arrivals of the itinerary requests. In this paper, we propose a new stochastic approximation method to compute bid prices. The key idea is to visualize the total expected revenue as a function of the bid prices and to search for a good set of bid prices by using sample path-based derivatives. Since the sample path-based derivatives of the total revenue function depend on what itinerary requests arrive in what order, our stochastic approximation method explicitly captures the temporal dynamics of the arrivals of the itinerary requests.

Focusing on a class of policies characterized by a number of parameters and utilizing stochastic approximation methods to search for a good set of values for the parameters is a common approach in stochastic optimization. Nevertheless, this approach, combined with the discrete nature of the decisions in the network revenue management setting, brings unique challenges. First, if we change the bid price of a flight leg by an infinitesimal amount, then the number of itinerary requests that we accept can only change by an integer amount, which implies that the sample path-based derivative of the total revenue function either does not exist or is equal to zero. We deal with this difficulty by formulating a smoothed version of the problem. The smoothed version assumes that the leg capacities are continuous and we can accept a fraction of an itinerary request depending on how much the revenue from the itinerary request exceeds the sum of the bid prices associated with the flight legs in the requested itinerary. It is, of course, not realistic to assume that we can accept a fraction of an itinerary request and we drop this assumption when implementing the seat allocation policy in practice. Second, even in the smoothed version of the problem, the sample path-based derivative of the total revenue function may not exist when there are multiple flight legs whose capacities are simultaneously binding. We deal with this difficulty by perturbing the leg capacities by small random amounts. These two modifications ensure that the sample path-based derivative of the total revenue function exists with probability one (w.p.1) and we can develop a convergent stochastic approximation method to search for a good set of bid prices.

Our work draws on two papers in particular. Mahajan and van Ryzin (2001) propose a stochastic approximation method to choose stocking levels for substitutable products. Since the total profit is not a concave function of the stocking levels, they formulate a smoothed version of their problem by assuming that the stocking levels can take fractional values and this allows them to develop a convergent stochastic approximation method. Our smoothed version is inspired by their work. On the other hand, van Ryzin and Vulcano (2006) propose a stochastic approximation method to choose protection levels in network revenue management problems. They show that the sample path-based derivative of the total revenue function does not exist when the protection level of a virtual class coincides with the capacity

on the flight leg. They resolve this difficulty by perturbing the leg capacities by small random amounts whenever they make a decision. Our perturbations of the leg capacities are inspired by their work.

The paper by Bertsimas and de Boer (2005) was the first to use stochastic approximation ideas to find good protection levels in network revenue management problems. Since the protection levels are restricted to be integers and the total expected revenue is a nonsmooth and nonconcave function of the protection levels, the authors in this paper use finite differences instead of sample path-based derivatives to guide their search. This approach yields good protection levels, but it does not have a convergence guarantee. Furthermore, it turns out that computing the finite differences can be computationally intensive. As mentioned above, van Ryzin and Vulcano (2006) resolve these difficulties by perturbing the leg capacities and by assuming that it is possible to accept a fraction of an itinerary request. Their approach has a convergence guarantee and is quite efficient since it results in a smooth total expected revenue function whose sample path-based derivatives are relatively easy to compute.

Our work differs from the two papers above in two important aspects. First, we focus on bid price policies, whereas the earlier papers focus on protection level policies. Protection levels are considered as primal controls where one makes the seat allocation decisions by putting limits on how many seats can be sold to different virtual classes. On the other hand, bid prices are considered as dual controls where one makes the seat allocation decisions by assigning opportunity costs to the seats that act as price barriers. There does not seem to be a consensus in the literature as to whether one class of policies perform better than the other. At any rate, the relative performance of the two classes of policies is generally not too crucial from a practical standpoint since the information technology infrastructures that can support the protection level and bid price policies are quite different. The choice between the two classes of policies is usually dictated by the legacy information technology infrastructure of a particular airline. Once one focuses attention to one of the two classes of policies, the goal is to find the best policy within a particular class. Consequently, it is important to develop methods that can find good bid prices as well as methods that can find good protection levels. Section 3.1.2 in Talluri and van Ryzin (2004) provide a detailed discussion of both protection level and bid price policies.

Second, it turns out that we have to deal with two sources of nondifferentiability in the total revenue function when we visualize it as a function of the bid prices. As mentioned above, the sample path-based derivative of the total revenue function may not exist when there are multiple flight legs whose capacities are simultaneously binding. We overcome this difficulty by perturbing the leg capacities. This is the same source of nondifferentiability that arises when we visualize the total revenue as a function of the protection levels, and van Ryzin and Vulcano (2006) show that perturbing the leg capacities is enough to recover the differentiability of the total revenue function with respect to the protection levels. On the other hand, the total revenue function may still be nondifferentiable with respect to the bid prices even after perturbing the leg capacities. Specifically, if we change the bid price of a flight leg by an infinitesimal amount, then the number of itinerary requests that we accept can only change by an integer amount and the sample path-based derivative of the total revenue function with respect to the bid prices still either does not exist or is equal to zero. We overcome this difficulty by assuming that we can accept a fraction of an itinerary request and this fraction is given by a smooth acceptance

function provided by the user. As a result, one has to smooth the network revenue management problem twice to compute bid prices by using stochastic approximation methods. This is in contrast with the work of van Ryzin and Vulcano (2006), where the authors show that only perturbing the leg capacities is enough to recover the differentiability of the total revenue function with respect to the protection levels and they do not need a smooth acceptance function. There are also a few algorithmic differences between our work and the paper by van Ryzin and Vulcano (2006). The most interesting one of these is that van Ryzin and Vulcano (2006) use projections to ensure the nestedness of the protection levels, but this is not required when we work with bid price policies. Furthermore, van Ryzin and Vulcano (2006) establish that the iterates of their algorithm converge in probability, whereas we establish that the iterates of our algorithm converge w.p.1.

There are many other papers that use stochastic approximation methods to find good policies for complex stochastic optimization problems. For example, L'Ecuyer and Glynn (1994) focus on queueing, Fu (1994), Glasserman and Tayur (1995), Bashyam and Fu (1998) and Kunnumkal and Topaloglu (2007) focus on inventory control and van Ryzin and McGill (2000), Karaesmen and van Ryzin (2004) and van Ryzin and Vulcano (2004) focus on revenue management settings. Kushner and Clark (1978) and Bertsekas and Tsitsiklis (1996) cover the theory of stochastic approximation methods.

The literature on computing bid prices in network revenue management problems is also related to our work. Simpson (1989) and Williamson (1992) were the first to use the aforementioned deterministic linear program to compute bid prices. Talluri and van Ryzin (1998) give a careful study of the policies characterized by bid prices and show that these policies are asymptotically optimal as the leg capacities and the expected numbers of the itinerary requests increase linearly with the same rate. Talluri and van Ryzin (1999) propose a randomized version of the deterministic linear program that uses actual samples of the numbers of the itinerary requests as opposed to expected values. Bertsimas and Popescu (2003) use a method that captures the total opportunity cost of the leg capacities consumed by an itinerary request more accurately. Adelman (2007) works with the linear programming representation of the dynamic programming formulation of the network revenue management problem to compute bid prices. Topaloglu (2006) uses a relaxation strategy to decompose the network revenue management problem by the flight legs and computes bid prices by concentrating on one flight leg at a time.

We make the following research contributions in this paper. 1) We propose a new method to compute bid prices that uses sample path-based derivatives of the total revenue function. Our method does not require computing expectations explicitly and allows having arbitrary nonstationarities and correlations in the arrivals of the itinerary requests. 2) We show that the iterates of our method converge to a stationary point of the total expected revenue function. 3) Computational experiments demonstrate that the bid prices obtained by our method may outperform the ones obtained by standard benchmark methods. The performance gap becomes particularly noticeable when there are multiple fare classes with large differences in the fares. In addition, our method is especially advantageous when the bid prices are not recomputed frequently.

The paper is organized as follows. Section 1 formulates a basic optimization problem that maximizes the total expected revenue by choosing the bid prices. Section 2 describes a smoothed version of this

problem that can be solved by using sample path-based derivatives and shows how to compute sample path-based derivatives in a tractable manner. Section 3 gives an algorithm to solve the smoothed version and shows that the iterates of this algorithm converge to a stationary point of the total expected revenue function. Section 4 presents computational experiments.

## 1 PROBLEM FORMULATION

We have a set of flight legs that can be used to satisfy the itinerary requests that arrive randomly over time. Whenever an itinerary request arrives, we have to decide whether to accept or reject it. An accepted itinerary request generates a revenue and consumes the capacities on the relevant flight legs. A rejected itinerary request simply leaves the system. We are interested in maximizing the total expected revenue from the accepted itinerary requests.

We use  $\mathcal{L}$  to denote the set of flight legs in the airline network and  $\mathcal{J}$  to denote the set of possible itineraries. If we accept a request for itinerary  $j$ , then we generate a revenue of  $\tilde{r}_j$  and consume  $\tilde{a}_{ij}$  units of capacity on flight leg  $i$ . If flight leg  $i$  is not in itinerary  $j$ , then we naturally have  $\tilde{a}_{ij} = 0$ . The itinerary requests arrive sequentially and we index them by  $t = 1, 2, \dots$ . We use  $x_{it}$  to denote the remaining capacity on flight leg  $i$  just before making the decision for itinerary request  $t$ . Therefore, the initial capacity on flight leg  $i$  is  $x_{i1}$  and  $x_1 = \{x_{i1} : i \in \mathcal{L}\}$  is a part of the problem data.

We assume that the total number of itinerary requests is bounded by a finite integer  $\tau$ . In this case, we can characterize the arrivals of the itinerary requests by the stochastic process  $\tilde{\omega} = \{J_t : t = 1, \dots, \tau\}$  with the interpretation that itinerary request  $t$  is for itinerary  $J_t$  and the value of the random variable  $J_t$  becomes known just before making the decision for itinerary request  $t$ . If we let  $r_t = \tilde{r}_{J_t}$  and  $a_{it} = \tilde{a}_{iJ_t}$  for notational brevity, then the random variables  $r_t$  and  $a_t = \{a_{it} : i \in \mathcal{L}\}$  capture all of the information related to itinerary request  $t$  and we can alternatively characterize the arrivals of the itinerary requests by the stochastic process  $\omega = \{(r_t, a_t) : t = 1, \dots, \tau\}$ . Throughout the paper, we work with the stochastic process  $\omega$  rather than  $\tilde{\omega}$ , although these two stochastic processes are equivalent for our purposes. As far as  $\omega$  is concerned, we only assume that  $|r_t| \leq B_r$  and  $a_{it} \in \{0, \dots, B_a\}$  w.p.1 for a finite scalar  $B_r$  and a finite integer  $B_a$ . Other than these assumptions,  $\omega$  can be a general stochastic process involving arbitrary nonstationarities and correlations among the itinerary requests. Since  $a_{it}$  can take values larger than one, our problem formulation allows having group requests. We also note that since we do not necessarily have exactly  $\tau$  itinerary requests in all sample paths of  $\omega$ , we allow having  $(r_t, a_t) = (0, 0)$  for all  $t = \tau^0 + 1, \dots, \tau$  for some random variable  $\tau^0$  taking values in  $\{1, \dots, \tau\}$ . In this case, accepting the last  $\tau - \tau^0$  itinerary requests would neither generate revenue nor consume the leg capacities.

The policy characterized by bid prices  $\lambda = \{\lambda_i : i \in \mathcal{L}\}$  accepts an itinerary request if and only if there is enough capacity and the revenue from the itinerary request exceeds the sum of the bid prices associated with the flight legs in the requested itinerary. Therefore, as a function of the remaining leg capacities and itinerary requests, the decision function of this policy can be written as

$$\tilde{u}_t(x_t, \omega, \lambda) = \mathbf{1}(x_t \geq a_t, r_t \geq \sum_{i \in \mathcal{L}} a_{it} \lambda_i), \quad (1)$$

where  $\mathbf{1}(\cdot)$  is the indicator function and  $x_t = \{x_{it} : i \in \mathcal{L}\}$  are the remaining leg capacities just before making the decision for itinerary request  $t$ . If the policy accepts itinerary request  $t$ , then we have  $\tilde{u}_t(x_t, \omega, \lambda) = 1$ . Otherwise, we have  $\tilde{u}_t(x_t, \omega, \lambda) = 0$ .

As a function of the remaining leg capacities and itinerary requests, the cumulative revenue function of the policy characterized by bid prices  $\lambda$  can be written recursively as

$$\tilde{R}_t(x_t, \omega, \lambda) = r_t \tilde{u}_t(x_t, \omega, \lambda) + \tilde{R}_{t+1}(x_t - \tilde{u}_t(x_t, \omega, \lambda) a_t, \omega, \lambda), \quad (2)$$

with  $\tilde{R}_{\tau+1}(\cdot, \cdot, \lambda) = 0$ . We note that the definition of the cumulative revenue function above is similar to the one in van Ryzin and Vulcano (2006), but the structure of our decision function is different. If we use the policy characterized by bid prices  $\lambda$ , then the total revenue that we generate from all itinerary requests is given by the random variable  $\tilde{R}_1(x_1, \omega, \lambda)$ . Therefore, we can find a good set of bid prices by solving the problem

$$\max_{\lambda} \mathbb{E}\{\tilde{R}_1(x_1, \omega, \lambda)\}. \quad (3)$$

One approach to solve this problem is to use the sample path-based derivatives of  $\tilde{R}_1(x_1, \omega, \lambda)$ . However, the difficulty with using the sample path-based derivatives is that if  $\lambda_i$  is perturbed by an infinitesimal amount, then the result of the decision function in (1) either does not change or changes by one. This implies that the derivative of  $\tilde{R}_1(x_1, \omega, \lambda)$  with respect to  $\lambda_i$  is either equal to zero or does not exist, in which case it is impossible to obtain useful sample path-based derivatives. In the next section, we resolve this difficulty by formulating a smoothed version of problem (3).

## 2 SMOOTHED PROBLEM AND SAMPLE PATH-BASED DERIVATIVES

In this section, we first formulate a smoothed version of problem (3) that can be solved by using sample path-based derivatives. The idea behind the smoothed version is to assume that the leg capacities are continuous and we can accept a fraction of an itinerary request. After this, we show how to compute the sample path-based derivatives of the cumulative revenue function for the smoothed version.

### 2.1 SMOOTHING THE REVENUE FUNCTION

We consider a policy that accepts a fraction of an itinerary request depending on how much the revenue from the itinerary request exceeds the sum of the bid prices associated with the flight legs in the requested itinerary. For this purpose, we let  $\theta(\cdot)$  be an increasing and differentiable function that satisfies  $\lim_{p \rightarrow \infty} \theta(p) = 1$  and  $\lim_{p \rightarrow -\infty} \theta(p) = 0$ . The policy characterized by bid prices  $\lambda$  accepts  $\theta(r_t - \sum_{i \in \mathcal{L}} a_{it} \lambda_i)$  units of itinerary request  $t$  as long as there is enough capacity. Therefore, the decision function of this policy can be written as

$$\hat{u}_t(x_t, \omega, \lambda) = \min \left\{ \min_{i \in \mathcal{L}_t^+} \{x_{it}/a_{it}\}, \theta(r_t - \sum_{i \in \mathcal{L}} a_{it} \lambda_i) \right\}, \quad (4)$$

where we let  $\mathcal{L}_t^+ = \{i \in \mathcal{L} : a_{it} > 0\}$ . It is easy to see that we have  $\hat{u}_t(x_t, \omega, \lambda) a_{it} \leq x_{it}$  for all  $i \in \mathcal{L}$  and the decision function above does not violate the leg capacities. We assume that  $\theta(\cdot)$  and its derivative

$\dot{\theta}(\cdot)$  are Lipschitz. That is, there exist finite scalars  $L_\theta$  and  $L_{\dot{\theta}}$  such that we have  $|\theta(p) - \theta(q)| \leq L_\theta |p - q|$  and  $|\dot{\theta}(p) - \dot{\theta}(q)| \leq L_{\dot{\theta}} |p - q|$  for all  $p, q \in \mathbb{R}$ . We discuss possible choices for  $\theta(\cdot)$  in Section 4.

Unfortunately, the decision function in (4) is still not differentiable with respect to  $\lambda$  whenever we have  $\min_{i \in \mathcal{L}_t^+} \{x_{it}/a_{it}\} = \theta(r_t - \sum_{i \in \mathcal{L}} a_{it} \lambda_i)$ . To overcome this difficulty, we use an approach proposed by van Ryzin and Vulcano (2006). In particular, we let  $\alpha = \{\alpha_{it} : i \in \mathcal{L}, t = 1, \dots, \tau\}$  be uniformly distributed random variables over a small interval  $[0, \epsilon]$  and perturb the leg capacities by  $\alpha_t = \{\alpha_{it} : i \in \mathcal{L}\}$  just before making the decision for itinerary request  $t$ . Therefore, we use the decision function

$$u_t(x_t, \omega, \alpha, \lambda) = \min \left\{ \min_{i \in \mathcal{L}_t^+} \{[x_{it} + \alpha_{it}]/a_{it}\}, \theta(r_t - \sum_{i \in \mathcal{L}} a_{it} \lambda_i) \right\}. \quad (5)$$

Assuming that  $\alpha$  is independent of  $\omega$  and  $\{\alpha_{it} : i \in \mathcal{L}, t = 1, \dots, \tau\}$  are independent of each other, the event that  $\min_{i \in \mathcal{L}_t^+} \{[x_{it} + \alpha_{it}]/a_{it}\} = \theta(r_t - \sum_{i \in \mathcal{L}} a_{it} \lambda_i)$  occurs with probability zero and the decision function above is differentiable with respect to  $\lambda$  w.p.1.

Similar to (2), the cumulative revenue function of the policy characterized by bid prices  $\lambda$  becomes

$$R_t(x_t, \omega, \alpha, \lambda) = r_t u_t(x_t, \omega, \alpha, \lambda) + R_{t+1}(x_t + \alpha_t - u_t(x_t, \omega, \alpha, \lambda) a_t, \omega, \alpha, \lambda), \quad (6)$$

with  $R_{\tau+1}(\cdot, \cdot, \cdot, \lambda) = 0$ . In this case, the smoothed version of the problem that we want to solve is

$$\max_{\lambda} \mathbb{E}\{R_1(x_1, \omega, \alpha, \lambda)\}. \quad (7)$$

Using the fact that the decision function in (5) is differentiable with respect to  $\lambda$  w.p.1, one can check by backward induction on (6) that  $R_1(x_1, \omega, \alpha, \lambda)$  is differentiable with respect to  $\lambda$  w.p.1 and it may be possible to solve problem (7) by using the sample path-based derivatives of  $R_1(x_1, \omega, \alpha, \lambda)$ .

## 2.2 COMPUTING THE SAMPLE PATH-BASED DERIVATIVES

We can compute the sample path-based derivatives of  $R_1(x_1, \omega, \alpha, \lambda)$  by using a simple recursion. We begin by introducing some new notation. We let  $\partial_i^\Lambda R_t(x_t, \omega, \alpha, \lambda)$  be the derivative of  $R_t(\cdot, \omega, \alpha, \cdot)$  with respect to the bid price of flight leg  $i$  evaluated at remaining leg capacities  $x_t$  and bid prices  $\lambda$ . Similarly, we let  $\partial_i^X R_t(x_t, \omega, \alpha, \lambda)$  be the derivative of  $R_t(\cdot, \omega, \alpha, \cdot)$  with respect to the remaining capacity on flight leg  $i$  evaluated at remaining leg capacities  $x_t$  and bid prices  $\lambda$ . That is, we have

$$\begin{aligned} \partial_i^\Lambda R_t(x_t, \omega, \alpha, \lambda) &= \left. \frac{\partial R_t(z_t, \omega, \alpha, \gamma)}{\partial \gamma_i} \right|_{(z_t, \gamma) = (x_t, \lambda)} \\ \partial_i^X R_t(x_t, \omega, \alpha, \lambda) &= \left. \frac{\partial R_t(z_t, \omega, \alpha, \gamma)}{\partial z_{it}} \right|_{(z_t, \gamma) = (x_t, \lambda)}. \end{aligned}$$

We also use  $\partial_i^\Lambda u_t(x_t, \omega, \alpha, \lambda)$  and  $\partial_i^X u_t(x_t, \omega, \alpha, \lambda)$  with similar interpretations.

Differentiating (6) with respect to the bid price of flight leg  $i$  and using the chain rule, we have

$$\begin{aligned} \partial_i^\Lambda R_t(x_t, \omega, \alpha, \lambda) &= r_t \partial_i^\Lambda u_t(x_t, \omega, \alpha, \lambda) + \partial_i^\Lambda R_{t+1}(x_t + \alpha_t - u_t(x_t, \omega, \alpha, \lambda) a_t, \omega, \alpha, \lambda) \\ &\quad - \sum_{j \in \mathcal{L}} a_{jt} \partial_i^\Lambda u_t(x_t, \omega, \alpha, \lambda) \partial_j^X R_{t+1}(x_t + \alpha_t - u_t(x_t, \omega, \alpha, \lambda) a_t, \omega, \alpha, \lambda). \quad (8) \end{aligned}$$

To compute the terms on the right side above, we differentiate (6) with respect to the remaining capacity on flight leg  $i$  to obtain

$$\begin{aligned} \partial_i^X R_t(x_t, \omega, \alpha, \lambda) &= r_t \partial_i^X u_t(x_t, \omega, \alpha, \lambda) \\ &+ \sum_{j \in \mathcal{L}} [\mathbf{1}(j = i) - a_{jt} \partial_i^X u_t(x_t, \omega, \alpha, \lambda)] \partial_j^X R_{t+1}(x_t + \alpha_t - u_t(x_t, \omega, \alpha, \lambda) a_t, \omega, \alpha, \lambda). \end{aligned} \quad (9)$$

On the other hand, differentiating (5) with respect to the bid price of flight leg  $i$ , we have

$$\partial_i^\Lambda u_t(x_t, \omega, \alpha, \lambda) = \begin{cases} -a_{it} \dot{\theta}(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) & \text{if } \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) < \min_{j \in \mathcal{L}_t^+} \{[x_{jt} + \alpha_{jt}]/a_{jt}\} \\ \text{may not exist} & \text{if } \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) = \min_{j \in \mathcal{L}_t^+} \{[x_{jt} + \alpha_{jt}]/a_{jt}\} \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

As discussed above, the event that  $\theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) = \min_{j \in \mathcal{L}_t^+} \{[x_{jt} + \alpha_{jt}]/a_{jt}\}$  occurs with probability zero. For notational brevity, we arbitrarily set  $\partial_i^\Lambda u_t(x_t, \omega, \alpha, \lambda)$  to  $-a_{it} \dot{\theta}(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j)$  whenever the second case above holds, in which case we can write (10) as

$$\partial_i^\Lambda u_t(x_t, \omega, \alpha, \lambda) = \begin{cases} -a_{it} \dot{\theta}(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) & \text{if } \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) \leq \min_{j \in \mathcal{L}_t^+} \{[x_{jt} + \alpha_{jt}]/a_{jt}\} \\ 0 & \text{otherwise.} \end{cases} \quad (11)$$

The expressions in (10) and (11) agree w.p.1 and we can simply use (11) instead of (10). This makes our proofs considerably shorter. Finally, differentiating (5) with respect to the remaining capacity on flight leg  $i$ , we have

$$\begin{aligned} \partial_i^X u_t(x_t, \omega, \alpha, \lambda) &= \begin{cases} 1/a_{it} & \text{if } i \in \mathcal{L}_t^+ \text{ and} \\ & [x_{it} + \alpha_{it}]/a_{it} < \min \left\{ \min_{j \in \mathcal{L}_t^+ \setminus \{i\}} \{[x_{jt} + \alpha_{jt}]/a_{jt}\}, \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) \right\} \\ \text{does not exist} & \text{if } i \in \mathcal{L}_t^+ \text{ and} \\ & [x_{it} + \alpha_{it}]/a_{it} = \min \left\{ \min_{j \in \mathcal{L}_t^+ \setminus \{i\}} \{[x_{jt} + \alpha_{jt}]/a_{jt}\}, \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) \right\} \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \quad (12)$$

Since  $\{\alpha_{it} : i \in \mathcal{L}, t = 1, \dots, \tau\}$  are continuous random variables and are independent of each other and  $\omega$ , the event that  $i \in \mathcal{L}_t^+$  and  $[x_{it} + \alpha_{it}]/a_{it} = \min \left\{ \min_{j \in \mathcal{L}_t^+ \setminus \{i\}} \{[x_{jt} + \alpha_{jt}]/a_{jt}\}, \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) \right\}$  occurs with probability zero. Similar to (11), we arbitrarily set  $\partial_i^X u_t(x_t, \omega, \alpha, \lambda)$  to  $1/a_{it}$  whenever the second case above holds, in which case we can write (12) as

$$\partial_i^X u_t(x_t, \omega, \alpha, \lambda) = \begin{cases} 1/a_{it} & \text{if } i \in \mathcal{L}_t^+ \text{ and} \\ & [x_{it} + \alpha_{it}]/a_{it} \leq \min \left\{ \min_{j \in \mathcal{L}_t^+ \setminus \{i\}} \{[x_{jt} + \alpha_{jt}]/a_{jt}\}, \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) \right\} \\ 0 & \text{otherwise.} \end{cases} \quad (13)$$

Once again, the expressions in (12) and (13) agree w.p.1 and we can simply use (13) instead of (12).

To compute  $\partial_i^\Lambda R_1(x_1, \omega, \alpha, \lambda)$ , all we need to do is to simulate the decisions of the policy characterized by bid prices  $\lambda$  under itinerary requests  $\omega$  and perturbation random variables  $\alpha$ . In this case, the leg capacities just before making the decision for itinerary request  $t + 1$  are given recursively by

$$X_{t+1}(x_1, \omega, \alpha, \lambda) = X_t(x_1, \omega, \alpha, \lambda) + \alpha_t - u_t(X_t(x_1, \omega, \alpha, \lambda), \omega, \alpha, \lambda) a_t, \quad (14)$$

with  $X_1(x_1, \omega, \alpha, \lambda) = x_1$ . Using (11) and (13), we can compute  $\partial_i^\Lambda u_t(X_t(x_1, \omega, \alpha, \lambda), \omega, \alpha, \lambda)$  and  $\partial_i^X u_t(X_t(x_1, \omega, \alpha, \lambda), \omega, \alpha, \lambda)$  for all  $i \in \mathcal{L}$ ,  $t = 1, \dots, \tau$ . We can compute  $\partial_i^X R_t(X_t(x_1, \omega, \alpha, \lambda), \omega, \alpha, \lambda)$  for all  $i \in \mathcal{L}$ ,  $t = 1, \dots, \tau$  by using (9) and moving backwards through the itinerary requests. Finally, we can compute  $\partial_i^\Lambda R_t(X_t(x_1, \omega, \alpha, \lambda), \omega, \alpha, \lambda)$  for all  $i \in \mathcal{L}$ ,  $t = 1, \dots, \tau$  by using (8) and moving backwards through the itinerary requests. All of these sample path-based derivatives exist in w.p.1 sense.

To give a feel for the number of operations required to compute  $\partial_i^\Lambda R_t(X_t(x_1, \omega, \alpha, \lambda), \omega, \alpha, \lambda)$  for all  $i \in \mathcal{L}$ ,  $t = 1, \dots, \tau$ , we assume that  $|\mathcal{L}_t^+| \leq M \leq |\mathcal{L}|$  w.p.1 for a finite scalar  $M$ . Since we can compute  $\min_{i \in \mathcal{L}_t^+} \{[x_{it} + \alpha_{it}]/a_{it}\}$  in  $O(M)$  time, we can compute  $X_t(x_1, \omega, \alpha, \lambda)$  for all  $t = 1, \dots, \tau$  in  $O(\tau M + \tau |\mathcal{L}|) = O(\tau |\mathcal{L}|)$  time by using (5) and (14). This can be interpreted as the amount of time required to simulate the decisions of the policy characterized by bid prices  $\lambda$ . On the other hand, since we can compute  $\min_{i \in \mathcal{L}_t^+} \{[x_{it} + \alpha_{it}]/a_{it}\}$  in  $O(M)$  time, we can compute  $\partial_i^\Lambda u_t(X_t(x_1, \omega, \alpha, \lambda), \omega, \alpha, \lambda)$  for all  $i \in \mathcal{L}$ ,  $t = 1, \dots, \tau$  in  $O(\tau M + \tau |\mathcal{L}|) = O(\tau |\mathcal{L}|)$  time by using (11). Noting that (13) can equivalently be written as

$$\partial_i^X u_t(x_t, \omega, \alpha, \lambda) = \begin{cases} 1/a_{it} & \text{if } i \in \operatorname{argmin}_{j \in \mathcal{L}_t^+} \{[x_{jt} + \alpha_{jt}]/a_{jt}\} \\ & \text{and } [x_{it} + \alpha_{it}]/a_{it} \leq \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) \\ 0 & \text{otherwise,} \end{cases} \quad (15)$$

we can also compute  $\partial_i^X u_t(X_t(x_1, \omega, \alpha, \lambda), \omega, \alpha, \lambda)$  for all  $i \in \mathcal{L}$ ,  $t = 1, \dots, \tau$  in  $O(\tau M + \tau |\mathcal{L}|) = O(\tau |\mathcal{L}|)$  time. These imply that we can compute  $\partial_i^\Lambda R_t(X_t(x_1, \omega, \alpha, \lambda), \omega, \alpha, \lambda)$  and  $\partial_i^X R_t(X_t(x_1, \omega, \alpha, \lambda), \omega, \alpha, \lambda)$  for all  $i \in \mathcal{L}$ ,  $t = 1, \dots, \tau$  in  $O(\tau |\mathcal{L}| + \tau |\mathcal{L}| + \tau |\mathcal{L}| M) = O(\tau |\mathcal{L}| M)$  time by using (8) and (9). In contrast, Section 2.4 in van Ryzin and Vulcano (2006) shows that the sample path-based derivatives used by these authors can be computed in  $O(K + \tau)$  time, where  $K$  is a constant that is significantly smaller than  $|\mathcal{L}|$ . Therefore, the computational complexity of the method proposed by van Ryzin and Vulcano (2006) is more favorable than the computational complexity of our method.

We note that  $\partial_i^X \mathbb{E}\{R_1(x_1, \omega, \alpha, \lambda)\}$  captures the change in the total expected revenue of the policy characterized by bid prices  $\lambda$  when we perturb the remaining capacity on flight leg  $i$  by an infinitesimal amount. Therefore,  $\partial_i^X \mathbb{E}\{R_1(x_1, \omega, \alpha, \lambda)\}$  captures the opportunity cost of a unit of capacity and it is tempting to use  $\partial_i^X \mathbb{E}\{R_1(x_1, \omega, \alpha, \lambda)\}$  as the bid price associated with flight leg  $i$ . Nevertheless, there is an important problem associated with this approach. The derivative  $\partial_i^X \mathbb{E}\{R_1(x_1, \omega, \alpha, \lambda)\}$  captures the opportunity cost of a unit of capacity on flight leg  $i$  under the assumption that the itinerary acceptance decisions are made by the bid price policy characterized by bid prices  $\lambda$ . However, if the itinerary acceptance decisions are made by using  $\{\partial_i^X \mathbb{E}\{R_1(x_1, \omega, \alpha, \lambda)\} : i \in \mathcal{L}\}$  as the bid prices, then  $\partial_i^X \mathbb{E}\{R_1(x_1, \omega, \alpha, \lambda)\}$  does not capture the opportunity cost of a unit of capacity on flight

leg  $i$  under this new bid price policy. As a result, there does not exist a concrete basis for using  $\{\partial_i^X \mathbb{E}\{R_1(x_1, \omega, \alpha, \lambda)\} : i \in \mathcal{L}\}$  as the bid prices. Similarly, van Ryzin and Vulcano (2006) use a cumulative revenue function of the form  $R_t(x_t, \omega, \alpha, y)$ , where the policy parameter  $y$  corresponds to the protection levels, and develop a recursion to compute the sample path-based derivative  $\partial_i^X R_t(x_t, \omega, \alpha, y)$ . It is again tempting to use  $\partial_i^X \mathbb{E}\{R_1(x_1, \omega, \alpha, y)\}$  as the bid price associated with flight leg  $i$ . However, the derivative  $\partial_i^X \mathbb{E}\{R_1(x_1, \omega, \alpha, y)\}$  captures the opportunity cost of a unit of capacity on flight leg  $i$  under the assumption that the itinerary acceptance decisions are made by the protection level policy characterized by protection levels  $y$ . In contrast, if the itinerary acceptance decisions are made by using  $\{\partial_i^X \mathbb{E}\{R_1(x_1, \omega, \alpha, y)\} : i \in \mathcal{L}\}$  as the bid prices, then  $\partial_i^X \mathbb{E}\{R_1(x_1, \omega, \alpha, y)\}$  does not capture the opportunity cost of a unit of capacity on flight leg  $i$  under this new bid price policy. As a result, there does not exist a concrete basis for using  $\{\partial_i^X \mathbb{E}\{R_1(x_1, \omega, \alpha, y)\} : i \in \mathcal{L}\}$  as the bid prices either. Despite the lack of a concrete basis, one can, of course, still check the performance of the bid price policy characterized by bid prices  $\{\partial_i^X \mathbb{E}\{R_1(x_1, \omega, \alpha, y)\} : i \in \mathcal{L}\}$ , but this is not the focus of our paper.

Closing this section, we emphasize that there does not exist a choice for  $\theta(\cdot)$  that makes the decision functions in (1) and (5) equivalent. Intuitively, we expect the policy described by the decision function in (5) to get closer to a conventional bid price policy as  $\theta(\cdot)$  gets closer to the step function  $\mathbf{1}(\cdot \geq 0)$ . However, the decision functions in (1) and (5) are different even if we choose  $\theta(\cdot)$  as the step function. Theoretically characterizing the implications of the difference in the decision functions in (1) and (5) is a difficult research question and we limit our evaluation to computational experiments.

In the next section, we describe an algorithm that uses the sample path-based derivatives given by (8), (9), (11) and (13) to search for a stationary point of the objective function of problem (7).

### 3 SOLUTION ALGORITHM AND CONVERGENCE

We propose the following algorithm to solve problem (7).

#### Algorithm 1

**Step 1.** Initialize the bid prices  $\lambda^1 = \{\lambda_i^1 : i \in \mathcal{L}\}$  arbitrarily and initialize the iteration counter by letting  $k = 1$ .

**Step 2.** Letting  $\omega^k$  be the itinerary requests and  $\alpha^k$  be the perturbation random variables at iteration  $k$ , compute  $\partial_i^\Lambda R_1(x_1, \omega^k, \alpha^k, \lambda^k)$  for all  $i \in \mathcal{L}$  by using (8), (9), (11) and (13).

**Step 3.** Letting  $\sigma^k$  be a step size parameter, compute the bid prices  $\lambda^{k+1} = \{\lambda_i^{k+1} : i \in \mathcal{L}\}$  at the next iteration as  $\lambda_i^{k+1} = \lambda_i^k + \sigma^k \partial_i^\Lambda R_1(x_1, \omega^k, \alpha^k, \lambda^k)$  for all  $i \in \mathcal{L}$ .

**Step 4.** Increase  $k$  by 1 and go to Step 2.

We let  $\mathcal{F}^k$  be the filtration generated by the random variables  $\{\lambda^1, \omega^1, \dots, \omega^{k-1}, \alpha^1, \dots, \alpha^{k-1}\}$  in this algorithm and assume that the joint distribution of  $(\omega^k, \alpha^k)$  conditional on  $\mathcal{F}^k$  is the same as the joint distribution of  $(\omega, \alpha)$ . We have the next convergence result for Algorithm 1.

**Proposition 1** *Assume that the sequence of step size parameters  $\{\sigma^k\}_k$  are  $\mathcal{F}^k$ -measurable and satisfy  $\sigma^k \geq 0$  for all  $k = 1, 2, \dots$ ,  $\sum_{k=1}^\infty \sigma^k = \infty$  and  $\sum_{k=1}^\infty [\sigma^k]^2 < \infty$  w.p.1. If the sequence of bid prices  $\{\lambda^k\}_k$*

are generated by Algorithm 1, then we have  $\lim_{k \rightarrow \infty} \mathbb{E}\{\partial_i^\Lambda R_1(x_1, \omega, \alpha, \lambda^k)\} = 0$  w.p.1 for all  $i \in \mathcal{L}$  and every limit point of the sequence of bid prices  $\{\lambda^k\}_k$  is a stationary point of the objective function of problem (7) w.p.1.

**Proof** Propositions 1, 2 and 3 in Section A of the online supplement show that the following statements hold for all  $\lambda, \gamma \in \mathbb{R}^{|\mathcal{L}|}$ ,  $i \in \mathcal{L}$ .

**(A.1)** Using  $\partial_i^\Lambda \mathbb{E}\{R_1(x_1, \omega, \alpha, \lambda)\}$  to denote the derivative of  $\mathbb{E}\{R_1(x_1, \omega, \alpha, \cdot)\}$  with respect to the bid price of flight leg  $i$  evaluated at bid prices  $\lambda$ , we have  $\mathbb{E}\{\partial_i^\Lambda R_1(x_1, \omega, \alpha, \lambda)\} = \partial_i^\Lambda \mathbb{E}\{R_1(x_1, \omega, \alpha, \lambda)\}$ .

**(A.2)** There exists a finite scalar  $B_R^\Lambda$  such that we have  $|\partial_i^\Lambda R_1(x_1, \omega, \alpha, \lambda)| \leq B_R^\Lambda$  w.p.1.

**(A.3)** Using  $\|\cdot\|$  to denote the Euclidean norm, there exists a finite scalar  $L_R^\Lambda$  such that we have  $\mathbb{E}\{|\partial_i^\Lambda R_1(x_1, \omega, \alpha, \lambda) - \partial_i^\Lambda R_1(x_1, \omega, \alpha, \gamma)|\} \leq L_R^\Lambda \|\lambda - \gamma\|$ .

In this case, the result follows from Proposition 4.1 in Bertsekas and Tsitsiklis (1996), which we briefly state in Section D of the online supplement for completeness. In particular, since the joint distribution of  $(\omega^k, \alpha^k)$  conditional on  $\mathcal{F}^k$  is the same as the joint distribution of  $(\omega, \alpha)$ , (A.1), (A.2) and (A.3) respectively show that (B.1), (B.2) and (B.3) in Proposition 4.1 in Bertsekas and Tsitsiklis (1996) are satisfied.  $\square$

(A.1) implies that the expected value of the step direction that we use in Step 3 of Algorithm 1 is an ascent direction of the objective function of problem (7). (A.2) implies that the norm of the step direction is uniformly bounded. (A.3) implies that the expected value of the step direction is Lipschitz when viewed as a function of the bid prices.

We emphasize that the objective function of problem (7) is not necessarily concave and the stationary point mentioned in Proposition 1 can be a local maximum, a saddle point, or even a local minimum. To illustrate, we consider a problem instance with large leg capacities (larger than  $B_a$  is enough) and  $\tau = 1$ . In this case, it is easy to see that  $R_1(x_1, \omega, \alpha, \lambda) = r_1 \theta(r_1 - \sum_{i \in \mathcal{L}} a_{i1} \lambda_i)$  and the objective function of problem (7) inherits the properties of  $\theta(\cdot)$  when, for example,  $(r_1, a_1)$  is deterministic. Some practical choices for  $\theta(\cdot)$  may have local minima or inflection points.

## 4 COMPUTATIONAL EXPERIMENTS

In this section, we compare the performances of the bid prices obtained by solving problem (7) with the performances of the bid prices obtained by using other benchmark strategies.

### 4.1 EXPERIMENTAL SETUP

In our test problems, we consider two types of airline networks that serve  $N$  spokes out of a single hub. In the first airline network, there are two flights associated with each spoke. One of these flights is out of the hub and the other one is to the hub. There is a high-fare and a low-fare itinerary that connect every possible origin-destination pair. Therefore, the first airline network involves  $2N$  flight legs and  $2N(N+1)$  itineraries,  $4N$  of which include one flight leg and  $2N(N-1)$  of which include two flight legs. Figure 1.a shows the first airline network for the case where  $N = 6$ . In the second airline network, there

is one flight associated with each spoke. For the first half of the spokes, this flight is to the hub and for the second half of the spokes, this flight is from the hub. Similar to the first airline network, there is a high-fare and a low-fare itinerary that connect every possible origin-destination pair. Therefore, the second airline network involves  $N$  flight legs and  $N(N/2 + 2)$  itineraries,  $2N$  of which include one flight leg and  $2(N/2)^2$  of which include two flight legs. Figure 1.b shows the second airline network for the case where  $N = 6$ . For both airline networks, the revenues associated with the high-fare itineraries are  $\rho$  times larger than the revenues associated with the low-fare itineraries. The probability of having a request for a high-fare itinerary increases over time, whereas the probability of having a request for a low-fare itinerary decreases over time. Since the total expected demand for the capacity on flight leg  $i$  is  $\sum_{t=1}^{\tau} \mathbb{E}\{a_{it}\}$ , we measure the tightness of the leg capacities by

$$\kappa = \frac{\sum_{i \in \mathcal{L}} \sum_{t=1}^{\tau} \mathbb{E}\{a_{it}\}}{\sum_{i \in \mathcal{L}} x_{i1}}.$$

We vary  $N$ ,  $\kappa$  and  $\rho$  in our test problems and label them by  $(T, N, \kappa, \rho) \in \{\text{I, II}\} \times \{6, 12\} \times \{1.0, 1.2, 1.6\} \times \{2, 4, 8\}$ , where the first component describes whether the problem takes place over the first or second airline network and the other three components are as described above. In our test problems, the initial capacities on the flight legs range over the interval  $[25, 50]$ , the total expected numbers of itinerary requests range over the interval  $[250, 500]$  and the fares range over the interval  $[50, 750]$ .

## 4.2 BENCHMARK STRATEGIES

We compare the performances of the bid prices obtained by the following five strategies.

**Sample path-based derivatives with randomized decision rule (SDR)** This is the strategy that we develop in this paper but our implementation deviates from the previous discussion in two important aspects. First, since it is not realistic to assume that we can accept a fraction of an itinerary request, our implementation assumes that  $\theta(\cdot)$  characterizes the probability of accepting an itinerary request. In particular, SDR computes the bid prices by solving problem (7). Letting  $\lambda^*$  be these bid prices, if there is enough capacity, then SDR accepts itinerary request  $t$  with probability  $\theta(r_t - \sum_{i \in \mathcal{L}} a_{it} \lambda_i^*)$ . Otherwise, it rejects the itinerary request. In our computational experiments, we use

$$\theta(p) = \begin{cases} 1 - a e^{-\frac{(1-a)}{b} p} & \text{if } p \geq 0 \\ (1-a) e^{\frac{a}{b} p} & \text{if } p < 0, \end{cases} \quad (16)$$

with  $a \in (0, 1)$  and  $b > 0$ . Figure 2 plots the function in (16) and shows that this function looks like the step function as  $b$  approaches to zero. Therefore, the distinction between assuming that  $\theta(\cdot)$  characterizes a fraction or a probability diminishes as  $b$  approaches to zero. It turns out that our results are relatively insensitive to the choice of  $(a, b)$  and we use  $(a, b) = (1/2, 20/3)$  in our computational experiments. Nevertheless, we emphasize that choosing  $b$  too small may create numerical difficulties. Specifically, as  $b$  approaches to zero, the derivative of  $\theta(\cdot)$  at any point away from zero gets arbitrarily small. Noting (8) and (11), this implies that the sample path-based derivative of the cumulative revenue function also gets arbitrarily small. This slows down the progress of Algorithm 1 extensively.

Second, SDR recomputes the bid prices  $n$  times over the decision horizon by resolving problem (7) after every  $\tau/n$  itinerary requests. Given the remaining leg capacities just before making the decision for

itinerary request  $t$ , we compute the bid prices by solving the problem  $\min_{\lambda} \mathbb{E}\{R_t(x_t, \omega, \alpha, \lambda)\}$  and use these bid prices until we resolve problem (7). We use  $n \in \{1, 3, 6\}$  in our computational experiments.

In all of our computational experiments, we use the step size parameter  $\sigma^k = 20/(40+k)$  in Algorithm 1 and terminate the algorithm after 20,000 iterations.

**Sample path-based derivatives with deterministic decision rule (SDD)** In certain practical settings, it may not be realistic to use the randomized decision rule prescribed by SDR. In particular, the managers may not be comfortable with the idea of flipping a coin to decide whether to accept or reject an itinerary request or to decide which itineraries should remain open. As a remedy, SDD simply uses the bid prices obtained by solving problem (7) in the decision function in (1). In other words, SDD computes the bid prices by solving problem (7). Letting  $\lambda^*$  be these bid prices, if there is enough capacity and we have  $r_t \geq \sum_{i \in \mathcal{L}} a_{it} \lambda_i^*$ , then SDD accepts itinerary request  $t$ . Otherwise, it rejects the itinerary request. Similar to SDR, SDD recomputes the bid prices  $n$  times over the decision horizon.

We emphasize that the results in Section 3 are under the assumption that we can accept a fraction of an itinerary request. Therefore, SDR and SDD should be visualized only as practical extensions of the results in Section 3.

**Deterministic linear program (DLP)** Using the notation at the beginning of Section 1 and noting that  $\sum_{t=1}^{\tau} \mathbf{1}(J_t = j)$  is the number of requests for itinerary  $j$  over the decision horizon, DLP computes the bid prices by solving the linear program

$$\max \quad \sum_{j \in \mathcal{J}} \tilde{r}_j z_j \tag{17}$$

$$\text{subject to} \quad \sum_{j \in \mathcal{J}} \tilde{a}_{ij} z_j \leq x_{i1} \quad \text{for all } i \in \mathcal{L} \tag{18}$$

$$0 \leq z_j \leq \sum_{t=1}^{\tau} \mathbb{E}\{\mathbf{1}(J_t = j)\} \quad \text{for all } j \in \mathcal{J}; \tag{19}$$

see Talluri and van Ryzin (2004). The decision variable  $z_j$  in the problem above is the number of requests for itinerary  $j$  that we plan to accept over the decision horizon. Constraints (18) ensure that the itinerary requests that we plan to accept do not violate the leg capacities, whereas constraints (19) ensure that the itinerary requests that we plan to accept do not exceed the expected numbers of the itinerary requests. Letting  $\{\lambda_i^* : i \in \mathcal{L}\}$  be the optimal values of the dual variables associated with constraints (18), if there is enough capacity and we have  $r_t \geq \sum_{i \in \mathcal{L}} a_{it} \lambda_i^*$ , then DLP accepts itinerary request  $t$ . Otherwise, it rejects the itinerary request. It is also possible to show that the optimal objective value of problem (17)-(19) provides an upper bound on the total expected revenue obtained by the optimal policy. This information becomes useful when assessing the optimality gap of a benchmark strategy.

Similar to SDR, DLP recomputes the bid prices  $n$  times over the decision horizon by resolving problem (17)-(19) after every  $\tau/n$  itinerary requests. Given the remaining leg capacities just before making the decision for itinerary request  $t$ , we replace the right side of constraints (18) with  $\{x_{it} : i \in \mathcal{L}\}$  and the right side of constraints (19) with  $\{\sum_{t'=t}^{\tau} \mathbb{E}\{\mathbf{1}(J_{t'} = j)\} : j \in \mathcal{J}\}$ , and solve problem (17)-(19).

We use the optimal values of the dual variables associated with constraints (18) as the bid prices until we resolve problem (17)-(19).

**Randomized linear program (RLP)** DLP uses only the expected numbers of the itinerary requests and RLP tries to make up for this deficiency. The idea behind RLP is to replace the right side of constraints (19) with the samples of  $\{\sum_{t=1}^{\tau} \mathbf{1}(J_t = j) : j \in \mathcal{J}\}$ . We generate  $K$  samples of  $\{J_t : t = 1, \dots, \tau\}$  and denote these samples by  $\{\hat{J}_t^k : t = 1, \dots, \tau\}$  for  $k = 1, \dots, K$ . We replace the right side of constraints (19) with  $\{\sum_{t=1}^{\tau} \mathbf{1}(\hat{J}_t^k = j) : j \in \mathcal{J}\}$  and solve problem (17)-(19). Letting  $\{\lambda_i^{k*} : i \in \mathcal{L}\}$  be the optimal values of the dual variables associated with constraints (18), RLP uses  $\{\sum_{k=1}^K \lambda_i^{k*} / K : i \in \mathcal{L}\}$  as the bid prices; see Talluri and van Ryzin (1999). We use  $K = 25$  in our computational experiments.

**Finite differences on the deterministic linear program (FD)** FD tries to improve on DLP by capturing the total opportunity cost of the leg capacities consumed by an itinerary request more accurately. Letting  $L_1(x_1)$  be the optimal objective value of problem (17)-(19), we replace the right side of constraints (18) with  $\{x_{i1} - \tilde{a}_{ij} : i \in \mathcal{L}\}$  and solve problem (17)-(19) to obtain the optimal objective value  $L_{j1}^-(x_1)$ . If there is enough capacity and we have  $r_t \geq L_1(x_1) - L_{j1}^-(x_1)$ , then FD accepts itinerary request  $t$ . Otherwise, it rejects the itinerary request; see Bertsimas and Popescu (2003). Both RLP and FD recompute the bid prices  $n$  times over the decision horizon by using an approach similar to the one used by DLP.

### 4.3 COMPUTATIONAL RESULTS

Our main computational results are summarized in six tables. For the first airline network, Tables 1, 2 and 3 respectively show the results for the cases where we recompute the bid prices once, three times and six times over the decision horizon. Tables 4, 5 and 6 do the same for the second airline network. The first five columns in these tables show the total expected revenues obtained by SDR, SDD, DLP, RLP and FD. We estimate these total expected revenues by simulating 250 different trajectories of itinerary requests. We use common random numbers when simulating the performances of the bid prices obtained by different strategies. The next four columns show the percent difference between the total expected revenues obtained by SDD and the other four strategies. SDD turns out to be one of the better strategies and we use it as a reference point. The last four columns compare the performance of SDD with the performances of SDR, DLP, RLP and FD. In particular, the tenth column includes a “✓” if SDD performs better than SDR, a “×” if SDR performs better than SDD and a “o” if there does not exist a statistically significant difference between the performances of SDD and SDR at 95% significance level. The interpretations of the eleventh, twelfth and thirteenth columns are similar but they respectively compare the performance of SDD with the performances of DLP, RLP and FD. Table 7 shows the optimal objective value of problem (17)-(19) for all of our test problems. As mentioned in the previous section, this is useful to get a feel for the optimality gap of different strategies.

Tables 1 and 4 indicate that if we compute the bid prices only once at the beginning of the decision horizon, then SDD performs significantly better than DLP, RLP and FD. The performance gap is especially large when there is a large difference between the revenues associated with the high-fare and

low-fare itineraries. Tables 2, 3, 5 and 6, on the other hand, indicate that the performance gap between SDD and the other three strategies gets smaller as we recompute the bid prices more frequently. For example, in Table 1, the total expected revenues obtained by SDD exceed those obtained by RLP by 10% on the average. If we recompute the bid prices three times over the decision horizon, then the total expected revenues obtained by SDD exceed those obtained by RLP by 3% on the average. If we recompute the bid prices six times over the decision horizon, then the same performance gap reduces to 1%. Nevertheless, we emphasize that such seemingly small performance gaps are quite significant in the network revenue management setting. Similar observations can be made from Tables 4, 5 and 6.

A general observation from the tables in this section is that SDD performs better than DLP, RLP and FD for a majority of the test problems. Specifically, SDD performs worse than RLP in only three cases and SDD performs worse than FD in only one case. SDD never performs worse than DLP. It is also interesting to note that the performances of DLP, RLP and FD improve significantly when we recompute the bid prices three or six times over the decision horizon. On the other hand, the performance of SDD is fairly satisfactory even when we compute the bid prices only once at the beginning of the decision horizon. It is possible to find test problems in Tables 1 and 2 for which the performance of SDD with  $n = 1$  is better than the performance of DLP, RLP or FD with  $n = 3$ .

If we compute the bid prices only once at the beginning of the decision horizon, then SDR performs better than SDD. This is not surprising, since the results in Section 3 are under the assumption that we can accept a fraction of an itinerary request and SDR partially accommodates this assumption by using a randomized decision rule. It is, however, surprising that if we recompute the bid prices three or six times over the decision horizon, then SDD performs noticeably better than SDR. Fortunately, this is welcome news since the bid prices in practical settings are usually recomputed several times over the decision horizon, and as mentioned in Section 4.2, SDD may be more appropriate in practical settings. Given that both SDR and SDD should be visualized only as practical extensions of the results in Section 3 and neither of them exactly matches the decision function in (5), it is difficult to justify why one strategy performs better than the other one. Our findings suggest that a choice between SDR and SDD should be made by considering the number of times that we recompute the bid prices and by comparing the performances of the two strategies in a specific problem context.

For our test problems, the performance of Algorithm 1 is relatively insensitive to the choice of the initial bid prices. In the computational experiments that we present in this section, we choose the initial bid prices as  $\{\sum_{j \in \mathcal{J}} \tilde{a}_{ij} \tilde{r}_j / \sum_{j \in \mathcal{J}} \tilde{a}_{ij} : i \in \mathcal{L}\}$ . However, choosing the initial bid prices in a different manner does not yield drastically different results. For example, Figure 3 plots  $\mathbb{E}\{R_1(x_1, \omega, \alpha, \lambda^k)\}$  for test problem (I, 12, 1.6, 8) as a function of the iteration counter  $k$  in Algorithm 1. The three data series correspond to the cases where we choose the initial bid prices as  $\{\sum_{j \in \mathcal{J}} \tilde{a}_{ij} \tilde{r}_j / \sum_{j \in \mathcal{J}} \tilde{a}_{ij} : i \in \mathcal{L}\}$ , as zero and as the bid prices obtained by the deterministic linear program. The figure indicates that the differences in the objective function values that we obtain after 10,000 iterations are less than 1%. Despite these encouraging empirical results, we emphasize the objective function of problem (7) is not concave and the performance of Algorithm 1 may potentially depend on the choice of the initial bid prices. Figure 3 also indicates that the performance of Algorithm 1 stabilizes after about 5,000 iterations.

Nevertheless, to compensate for the lack of good stopping criteria for stochastic approximation methods and to be on the safe side, we terminate the algorithm after 20,000 iterations.

For test problem (I, 12, 1.6, 8), letting  $\{\lambda_i^{1*} : i \in \mathcal{L}\}$ ,  $\{\lambda_i^{2*} : i \in \mathcal{L}\}$  and  $\{\lambda_i^{3*} : i \in \mathcal{L}\}$  respectively be the bid prices obtained by Algorithm 1 when we choose the initial bid prices as  $\{\sum_{j \in \mathcal{J}} \tilde{a}_{ij} \tilde{r}_j / \sum_{j \in \mathcal{J}} \tilde{a}_{ij} : i \in \mathcal{L}\}$ , as zero and as the bid prices obtained by the deterministic linear program, Figure 4 gives scatter plots of  $\{(\lambda_i^{1*}, \lambda_i^{2*}) : i \in \mathcal{L}\}$  and  $\{(\lambda_i^{1*}, \lambda_i^{3*}) : i \in \mathcal{L}\}$ . The plots indicate that the bid prices obtained by starting from different initial bid prices are close to each other, but may not be exactly the same.

Table 8 shows the CPU seconds required to compute one set of bid prices on a Pentium IV PC running Windows XP with 2.4 GHz of CPU and 1 GB of RAM. Since the structure of the airline network and the number of spokes are the primary factors that affect the CPU seconds, we only show the average CPU seconds over different test problems. Since both SDR and SDD compute the bid prices by solving problem (7), they have the same CPU seconds in Table 8. DLP is quite fast as it requires solving one linear program to compute the bid prices. RLP and FD are a bit slower than DLP as they require solving multiple linear programs. Running Algorithm 1 for 20,000 iterations for SDR and SDD takes a few minutes, and the CPU seconds for SDR and SDD are larger than those for the other strategies by orders of magnitude. Nevertheless, we emphasize that these CPU seconds increase by about a factor of 5 when the number of spokes increases by a factor of 2, the number of flight legs increases by a factor of 2 and the number of possible itineraries increases by a factor of 4. Also, the test problems with  $N = 12$  have approximately twice as many itinerary requests as those with  $N = 6$ . Given that the bid prices are usually computed through overnight runs, these CPU seconds are acceptable from a practical standpoint. In addition, due to the administrative overhead associated with opening and closing the itineraries, a method that requires recomputing the bid prices fewer times but takes longer to run may be preferable to a method that requires recomputing the bid prices more frequently but takes shorter to run.

## 5 CONCLUSIONS

In this paper, we developed a convergent stochastic approximation method to compute bid prices in network revenue management problems. To facilitate the convergence proof, we worked with a smoothed version of the problem, which assumes that the leg capacities are continuous and we can accept a fraction of an itinerary request. SDR used the bid prices obtained by our stochastic approximation method through a randomized decision rule. Since such a randomized decision rule may not be realistic in certain practical settings, SDD used the bid prices obtained by our stochastic approximation method through the decision function in (1). Computational experiments demonstrated that the bid prices obtained by our stochastic approximation method are especially advantageous when there are multiple fare classes with large differences in the fares and the bid prices are not recomputed frequently.

There are several directions for further research. First, it is possible to use different sets of bid prices to make the decisions for different itinerary requests. Specifically, we can use the bid prices  $\{\lambda_{it} : i \in \mathcal{L}\}$  to make the decision for itinerary request  $t$ , in which case the cumulative revenue function in (6) becomes a function of  $\{\lambda_{it} : i \in \mathcal{L}, t = 1, \dots, \tau\}$ . The difficulty with this approach is that

the sample path-based derivative of the cumulative revenue function with respect to  $\lambda_{it}$  is zero when  $a_{it} = 0$  and it is not always possible to obtain a meaningful sample path-based derivative with respect to all bid prices. Our preliminary computational experiments indicate that using different sets of bid prices to make the decisions for different itinerary requests does not provide a noticeable advantage over using the same set of bid prices to make the decisions for all itinerary requests. More work is needed to make this approach work. Second, as  $b$  approaches to zero, the function in (16) looks like the step function and the distinction between assuming that  $\theta(\cdot)$  characterizes a fraction or a probability diminishes. It is possible to visualize a version of Algorithm 1 where we decrease  $b$  at each iteration. It would be interesting to get a convergence result for this version as  $b$  approaches to zero and the number of iterations approaches to infinity.

#### ACKNOWLEDGEMENTS

The author thanks two anonymous referees for their useful comments that especially improved the exposition and the computational results. This work was supported in part by National Science Foundation grant DMI-0422133.

#### REFERENCES

- Adelman, D. (2007), ‘Dynamic bid-prices in revenue management’, *Operations Research* **55**(4), 647–661.
- Bashyam, S. and Fu, M. C. (1998), ‘Optimization of  $(s, S)$  inventory systems with random lead times and a service level constraint’, *Management Science* **44**(12), 243–256.
- Bertsekas, D. P. and Tsitsiklis, J. N. (1996), *Neuro-Dynamic Programming*, Athena Scientific, Belmont, MA.
- Bertsimas, D. and de Boer, S. (2005), ‘Simulation-based booking limits for airline revenue management’, *Operations Research* **53**(1), 90–106.
- Bertsimas, D. and Popescu, I. (2003), ‘Revenue management in a dynamic network environment’, *Transportation Science* **37**, 257–277.
- Fu, M. (1994), ‘Sample path derivatives for  $(s, S)$  inventory systems’, *Operations Research* **42**(2), 351–363.
- Glasserman, P. (1994), Perturbation analysis of production networks, in D. D. Yao, ed., ‘*Stochastic Modeling and Analysis of Manufacturing Systems*’, Springer, New York, NY.
- Glasserman, P. and Tayur, S. (1995), ‘Sensitivity analysis for base-stock levels in multiechelon production-inventory systems’, *Management Science* **41**(2), 263–281.
- Karaesmen, I. and van Ryzin, G. (2004), ‘Overbooking with substitutable inventory classes’, *Operations Research* **52**(1), 83–104.
- Kunnumkal, S. and Topaloglu, H. (2007), ‘Using stochastic approximation algorithms to compute optimal base-stock levels in inventory control problems’, *Operations Research* (to appear).
- Kushner, H. J. and Clark, D. S. (1978), *Stochastic Approximation Methods for Constrained and Unconstrained Systems*, Springer-Verlag, Berlin.
- L’Ecuyer, P. and Glynn, P. (1994), ‘Stochastic optimization by simulation: Convergence proofs for the GI/G/1 queue in steady state’, *Management Science* **40**, 1245–1261.

- Mahajan, S. and van Ryzin, G. (2001), ‘Stocking retail assortments under dynamic customer substitution’, *Operations Research* **49**(3), 334–351.
- Simpson, R. W. (1989), Using network flow techniques to find shadow prices for market and seat inventory control, Technical report, MIT Flight Transportation Laboratory Memorandum M89-1, Cambridge, MA.
- Talluri, K. T. and van Ryzin, G. J. (2004), *The Theory and Practice of Revenue Management*, Kluwer Academic Publishers.
- Talluri, K. and van Ryzin, G. (1998), ‘An analysis of bid-price controls for network revenue management’, *Management Science* **44**(11), 1577–1593.
- Talluri, K. and van Ryzin, G. (1999), ‘A randomized linear programming method for computing network bid prices’, *Transportation Science* **33**(2), 207–216.
- Topaloglu, H. (2006), Using Lagrangian relaxation to compute capacity-dependent bid-prices in network revenue management, Technical report, Cornell University, School of Operations Research and Information Engineering.  
Available at <http://legacy.orie.cornell.edu/~huseyin/publications/publications.html>.
- van Ryzin, G. and McGill, J. (2000), ‘Revenue management without forecasting or optimization: An adaptive algorithm for determining airline seat protection levels’, *Management Science* **46**(6), 760–775.
- van Ryzin, G. and Vulcano, G. (2004), Computing virtual nesting controls for network revenue management under customer choice behavior, Technical Report DRO-2004-09, Columbia Business School.
- van Ryzin, G. and Vulcano, G. (2006), ‘Simulation-based optimization of virtual nesting controls for network revenue management’, *Operations Research* (to appear).
- Williamson, E. L. (1992), Airline Network Seat Control, PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.

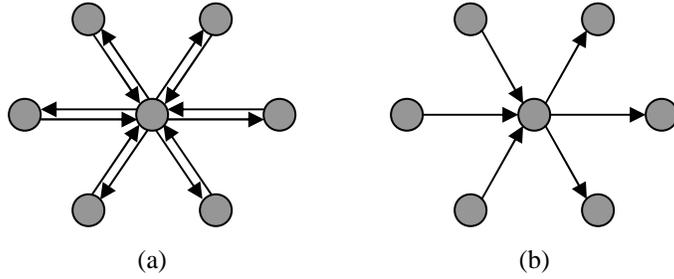


Figure 1: Structures of the airline networks for the case where  $N = 6$ .

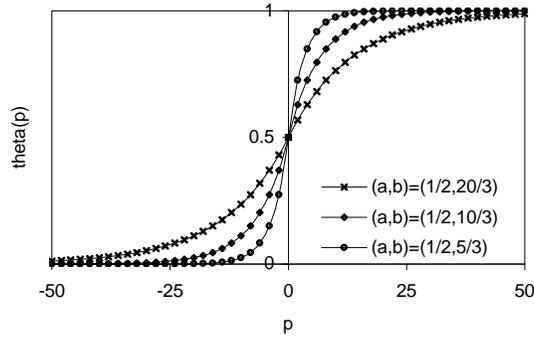


Figure 2: The function in (16) for different values of  $b$ .

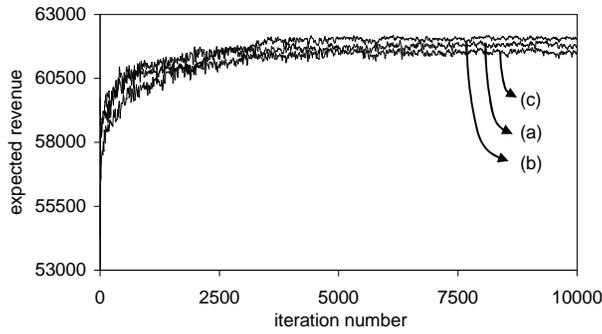


Figure 3: Trajectory of Algorithm 1 for test problem  $(I, 12, 1.6, 8)$ . We choose the initial bid prices (a) as  $\{\sum_{j \in \mathcal{J}} \tilde{a}_{ij} \tilde{r}_j / \sum_{j \in \mathcal{J}} \tilde{a}_{ij} : i \in \mathcal{L}\}$ , (b) as zero and (c) as the bid prices obtained by the deterministic linear program.

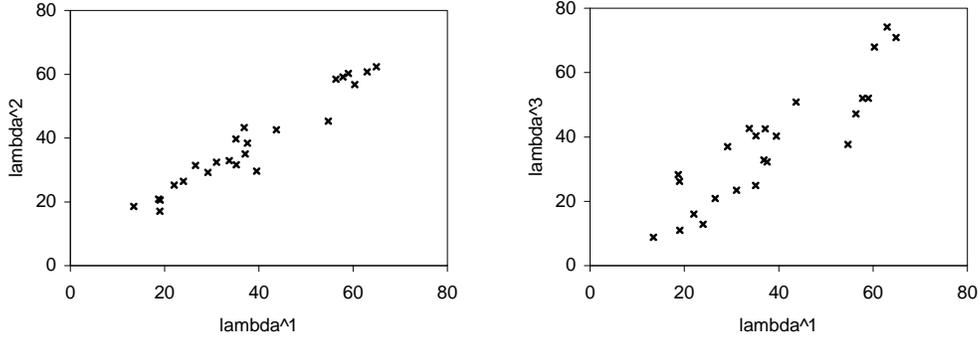


Figure 4: Comparison of the bid prices obtained by Algorithm 1 for test problem (I, 12, 1.6, 8) when we use different initial bid prices.

prob.	total expected revenue obtained by					% diff. btwn. SDD and				sig. diff. btwn. SDD and			
	SDR	SDD	DLP	RLP	FD	SDR	DLP	RLP	FD	SDR	DLP	RLP	FD
(I, 6, 1.0, 2)	16,956	16,948	16,538	16,887	16,616	-0.05	2.42	0.36	1.96	○	✓	✓	✓
(I, 6, 1.0, 4)	24,052	23,552	22,295	23,054	22,459	-2.12	5.34	2.12	4.64	×	✓	✓	✓
(I, 6, 1.0, 8)	38,945	38,669	33,809	35,359	34,156	-0.71	12.57	8.56	11.67	×	✓	✓	✓
(I, 6, 1.2, 2)	14,941	14,786	14,304	14,363	14,873	-1.05	3.26	2.86	-0.59	×	✓	✓	○
(I, 6, 1.2, 4)	21,904	21,055	18,563	18,467	21,413	-4.03	11.84	12.29	-1.70	×	✓	✓	×
(I, 6, 1.2, 8)	36,023	34,803	27,110	26,735	28,641	-3.51	22.10	23.18	17.70	×	✓	✓	✓
(I, 6, 1.6, 2)	12,061	12,006	11,313	11,572	11,782	-0.45	5.77	3.62	1.87	○	✓	✓	✓
(I, 6, 1.6, 4)	18,889	18,687	14,867	15,867	17,080	-1.08	20.44	15.09	8.60	×	✓	✓	✓
(I, 6, 1.6, 8)	33,221	32,677	21,976	24,450	24,102	-1.67	32.75	25.18	26.24	×	✓	✓	✓
(I, 12, 1.0, 2)	30,611	30,491	29,725	29,949	30,095	-0.39	2.51	1.78	1.30	×	✓	✓	✓
(I, 12, 1.0, 4)	43,908	43,740	39,936	40,452	41,451	-0.38	8.70	7.52	5.23	○	✓	✓	✓
(I, 12, 1.0, 8)	71,718	71,370	60,358	61,417	61,506	-0.49	15.43	13.95	13.82	×	✓	✓	✓
(I, 12, 1.2, 2)	27,390	27,697	26,481	26,607	27,644	1.11	4.39	3.94	0.19	✓	✓	✓	○
(I, 12, 1.2, 4)	40,436	40,494	35,106	35,423	39,738	0.14	13.31	12.52	1.87	○	✓	✓	✓
(I, 12, 1.2, 8)	67,189	66,602	52,353	52,973	57,590	-0.88	21.39	20.46	13.53	×	✓	✓	✓
(I, 12, 1.6, 2)	22,447	22,905	21,993	22,302	22,736	2.00	3.98	2.63	0.74	✓	✓	✓	✓
(I, 12, 1.6, 4)	35,084	34,577	30,214	31,223	34,346	-1.47	12.62	9.70	0.67	×	✓	✓	○
(I, 12, 1.6, 8)	61,809	60,147	46,363	48,834	52,515	-2.76	22.92	18.81	12.69	×	✓	✓	✓

Table 1: Computational results for the first airline network for the case where  $n = 1$ .

prob.	total expected revenue obtained by					% diff. btwn. SDD and				sig. diff. btwn. SDD and			
	SDR	SDD	DLP	RLP	FD	SDR	DLP	RLP	FD	SDR	DLP	RLP	FD
(I, 6, 1.0, 2)	17,185	17,343	17,082	17,210	17,205	0.91	1.51	0.77	0.79	✓	✓	✓	✓
(I, 6, 1.0, 4)	24,570	24,715	23,766	24,131	24,250	0.59	3.84	2.36	1.88	✓	✓	✓	✓
(I, 6, 1.0, 8)	39,301	39,282	37,133	38,312	38,292	-0.05	5.47	2.47	2.52	○	✓	✓	✓
(I, 6, 1.2, 2)	15,312	15,517	15,209	15,261	15,330	1.32	1.98	1.65	1.20	✓	✓	✓	✓
(I, 6, 1.2, 4)	22,556	22,711	21,170	21,761	22,099	0.68	6.79	4.18	2.69	✓	✓	✓	✓
(I, 6, 1.2, 8)	35,585	35,603	33,211	35,297	34,785	0.05	6.72	0.86	2.30	○	✓	○	✓
(I, 6, 1.6, 2)	12,495	12,667	12,266	12,468	12,433	1.36	3.16	1.57	1.84	✓	✓	✓	✓
(I, 6, 1.6, 4)	19,673	19,712	17,569	18,956	19,048	0.20	10.87	3.84	3.37	○	✓	✓	✓
(I, 6, 1.6, 8)	33,361	32,974	28,206	32,714	30,957	-1.17	14.46	0.79	6.11	×	✓	○	✓
(I, 12, 1.0, 2)	31,114	31,508	30,930	31,064	31,173	1.25	1.84	1.41	1.06	✓	✓	✓	✓
(I, 12, 1.0, 4)	44,647	45,187	43,291	43,666	44,019	1.20	4.20	3.37	2.59	✓	✓	✓	✓
(I, 12, 1.0, 8)	72,586	73,041	68,053	69,547	69,620	0.62	6.83	4.78	4.68	✓	✓	✓	✓
(I, 12, 1.2, 2)	28,017	28,510	27,971	28,008	28,216	1.73	1.89	1.76	1.03	✓	✓	✓	✓
(I, 12, 1.2, 4)	41,261	41,754	39,279	39,869	40,611	1.18	5.93	4.51	2.74	✓	✓	✓	✓
(I, 12, 1.2, 8)	68,775	69,199	62,101	64,441	65,598	0.61	10.26	6.88	5.20	✓	✓	✓	✓
(I, 12, 1.6, 2)	23,073	23,658	23,217	23,108	23,476	2.47	1.87	2.32	0.77	✓	✓	✓	✓
(I, 12, 1.6, 4)	35,948	36,290	34,008	34,586	35,435	0.94	6.29	4.70	2.36	✓	✓	✓	✓
(I, 12, 1.6, 8)	63,199	63,387	55,894	59,254	60,066	0.30	11.82	6.52	5.24	○	✓	✓	✓

Table 2: Computational results for the first airline network for the case where  $n = 3$ .

prob.	total expected revenue obtained by					% diff. btwn. SDD and				sig. diff. btwn. SDD and			
	SDR	SDD	DLP	RLP	FD	SDR	DLP	RLP	FD	SDR	DLP	RLP	FD
(I, 6, 1.0, 2)	17,277	17,442	17,232	17,322	17,282	0.95	1.21	0.69	0.92	✓	✓	✓	✓
(I, 6, 1.0, 4)	24,671	24,923	24,132	24,500	24,408	1.01	3.17	1.70	2.07	✓	✓	✓	✓
(I, 6, 1.0, 8)	39,158	39,159	37,933	39,289	38,635	0.00	3.13	-0.33	1.34	○	✓	○	✓
(I, 6, 1.2, 2)	15,440	15,683	15,433	15,465	15,509	1.55	1.59	1.39	1.11	✓	✓	✓	✓
(I, 6, 1.2, 4)	22,711	23,000	21,807	22,381	22,325	1.26	5.19	2.69	2.93	✓	✓	✓	✓
(I, 6, 1.2, 8)	36,504	36,415	34,646	36,934	35,983	-0.25	4.86	-1.43	1.19	○	✓	×	✓
(I, 6, 1.6, 2)	12,597	12,824	12,571	12,666	12,570	1.77	1.97	1.23	1.98	✓	✓	✓	✓
(I, 6, 1.6, 4)	19,844	20,031	18,519	19,450	19,220	0.94	7.55	2.90	4.05	✓	✓	✓	✓
(I, 6, 1.6, 8)	33,023	33,267	30,503	33,916	32,281	0.74	8.31	-1.95	2.96	✓	✓	×	✓
(I, 12, 1.0, 2)	31,240	31,666	31,231	31,314	31,333	1.35	1.38	1.11	1.05	✓	✓	✓	✓
(I, 12, 1.0, 4)	44,856	45,351	43,928	44,379	44,411	1.09	3.14	2.14	2.07	✓	✓	✓	✓
(I, 12, 1.0, 8)	72,692	73,265	69,435	71,373	70,472	0.78	5.23	2.58	3.81	✓	✓	✓	✓
(I, 12, 1.2, 2)	28,204	28,733	28,294	28,325	28,361	1.84	1.53	1.42	1.30	✓	✓	✓	✓
(I, 12, 1.2, 4)	41,483	42,001	40,108	40,731	40,981	1.23	4.51	3.02	2.43	✓	✓	✓	✓
(I, 12, 1.2, 8)	69,006	69,593	64,069	67,063	66,434	0.84	7.94	3.64	4.54	✓	✓	✓	✓
(I, 12, 1.6, 2)	23,303	23,897	23,541	23,583	23,562	2.48	1.49	1.31	1.40	✓	✓	✓	✓
(I, 12, 1.6, 4)	36,236	36,724	34,852	35,624	35,632	1.33	5.10	3.00	2.97	✓	✓	✓	✓
(I, 12, 1.6, 8)	63,642	63,741	57,730	61,827	60,663	0.15	9.43	3.00	4.83	○	✓	✓	✓

Table 3: Computational results for the first airline network for the case where  $n = 6$ .

prob.	total expected revenue obtained by					% diff. btwn. SDD and				sig. diff. btwn. SDD and			
	SDR	SDD	DLP	RLP	FD	SDR	DLP	RLP	FD	SDR	DLP	RLP	FD
(II, 6, 1.0, 2)	17,665	17,614	17,604	17,609	17,275	-0.29	0.06	0.03	1.93	○	✓	○	✓
(II, 6, 1.0, 4)	25,471	24,491	24,468	24,479	24,603	-4.00	0.09	0.05	-0.46	×	✓	○	○
(II, 6, 1.0, 8)	38,891	38,844	38,195	38,221	38,467	-0.12	1.67	1.60	0.97	○	✓	✓	✓
(II, 6, 1.2, 2)	15,393	15,018	14,223	14,222	14,938	-2.50	5.29	5.30	0.53	×	✓	✓	○
(II, 6, 1.2, 4)	23,216	22,201	18,085	18,621	20,683	-4.57	18.54	16.13	6.84	×	✓	✓	✓
(II, 6, 1.2, 8)	35,788	35,651	25,815	27,417	25,818	-0.38	27.59	23.10	27.58	○	✓	✓	✓
(II, 6, 1.6, 2)	12,462	12,497	10,763	11,268	10,763	0.28	13.88	9.84	13.88	○	✓	✓	✓
(II, 6, 1.6, 4)	20,379	19,331	13,446	16,485	13,446	-5.42	30.44	14.72	30.44	×	✓	✓	✓
(II, 6, 1.6, 8)	33,711	33,641	18,811	26,931	18,810	-0.21	44.08	19.95	44.09	×	✓	✓	✓
(II, 12, 1.0, 2)	30,731	30,435	30,319	30,396	30,545	-0.97	0.38	0.13	-0.36	×	✓	✓	○
(II, 12, 1.0, 4)	43,742	42,761	41,236	41,437	42,520	-2.29	3.56	3.10	0.56	×	✓	✓	○
(II, 12, 1.0, 8)	70,388	68,788	63,071	63,511	63,223	-2.33	8.31	7.67	8.09	×	✓	✓	✓
(II, 12, 1.2, 2)	27,598	27,439	26,438	27,204	27,397	-0.58	3.65	0.86	0.15	×	✓	✓	○
(II, 12, 1.2, 4)	40,321	39,646	35,213	36,930	35,218	-1.70	11.18	6.85	11.17	×	✓	✓	✓
(II, 12, 1.2, 8)	66,970	64,001	52,771	56,425	52,776	-4.64	17.55	11.84	17.54	×	✓	✓	✓
(II, 12, 1.6, 2)	22,474	22,201	22,227	22,220	21,477	-1.23	-0.12	-0.09	3.26	×	○	○	✓
(II, 12, 1.6, 4)	34,952	33,468	29,421	30,196	29,429	-4.43	12.09	9.77	12.07	×	✓	✓	✓
(II, 12, 1.6, 8)	58,093	56,140	44,566	46,437	44,568	-3.48	20.62	17.28	20.61	×	✓	✓	✓

Table 4: Computational results for the second airline network for the case where  $n = 1$ .

prob.	total expected revenue obtained by					% diff. btwn. SDD and				sig. diff. btwn. SDD and			
	SDR	SDD	DLP	RLP	FD	SDR	DLP	RLP	FD	SDR	DLP	RLP	FD
(II, 6, 1.0, 2)	17,849	17,955	17,863	17,891	17,747	0.60	0.51	0.36	1.16	✓	✓	✓	✓
(II, 6, 1.0, 4)	25,802	25,826	25,334	25,559	25,450	0.09	1.90	1.04	1.46	○	✓	✓	✓
(II, 6, 1.0, 8)	39,644	41,676	40,280	41,078	40,710	4.88	3.35	1.43	2.32	✓	✓	✓	✓
(II, 6, 1.2, 2)	15,664	15,656	15,133	15,501	15,382	-0.05	3.34	0.99	1.75	○	✓	✓	✓
(II, 6, 1.2, 4)	23,568	23,452	21,006	22,627	22,243	-0.49	10.43	3.52	5.16	○	✓	✓	✓
(II, 6, 1.2, 8)	36,867	38,314	32,736	37,204	34,204	3.78	14.56	2.90	10.73	✓	✓	✓	✓
(II, 6, 1.6, 2)	12,721	12,778	12,083	12,599	12,299	0.44	5.44	1.40	3.75	○	✓	✓	✓
(II, 6, 1.6, 4)	20,645	20,297	17,556	19,874	18,182	-1.71	13.50	2.08	10.42	×	✓	✓	✓
(II, 6, 1.6, 8)	34,627	34,725	28,546	34,711	29,883	0.28	17.80	0.04	13.94	○	✓	○	✓
(II, 12, 1.0, 2)	31,130	31,401	31,105	31,183	31,206	0.86	0.94	0.69	0.62	✓	✓	✓	✓
(II, 12, 1.0, 4)	44,333	44,667	43,429	43,908	44,000	0.75	2.77	1.70	1.49	✓	✓	✓	✓
(II, 12, 1.0, 8)	70,749	71,117	68,132	69,608	69,161	0.52	4.20	2.12	2.75	✓	✓	✓	✓
(II, 12, 1.2, 2)	28,126	28,625	27,954	28,250	28,194	1.74	2.34	1.31	1.50	✓	✓	✓	✓
(II, 12, 1.2, 4)	41,134	41,658	39,370	40,397	40,109	1.26	5.49	3.03	3.72	✓	✓	✓	✓
(II, 12, 1.2, 8)	65,967	67,177	62,305	65,109	64,059	1.80	7.25	3.08	4.64	✓	✓	✓	✓
(II, 12, 1.6, 2)	23,160	23,556	23,154	23,359	23,272	1.68	1.71	0.84	1.21	✓	✓	✓	✓
(II, 12, 1.6, 4)	35,819	35,946	33,651	35,142	34,450	0.35	6.38	2.24	4.16	○	✓	✓	✓
(II, 12, 1.6, 8)	60,517	61,447	54,747	59,635	57,439	1.51	10.90	2.95	6.52	✓	✓	✓	✓

Table 5: Computational results for the second airline network for the case where  $n = 3$ .

prob.	total expected revenue obtained by					% diff. btwn. SDD and				sig. diff. btwn. SDD and			
	SDR	SDD	DLP	RLP	FD	SDR	DLP	RLP	FD	SDR	DLP	RLP	FD
(II, 6, 1.0, 2)	17,924	18,065	17,991	18,002	17,913	0.78	0.41	0.35	0.84	✓	✓	✓	✓
(II, 6, 1.0, 4)	25,857	26,118	25,646	25,894	25,792	1.00	1.81	0.86	1.25	✓	✓	✓	✓
(II, 6, 1.0, 8)	41,331	42,003	40,961	41,992	41,392	1.60	2.48	0.03	1.45	✓	✓	○	✓
(II, 6, 1.2, 2)	15,783	15,926	15,699	15,842	15,754	0.90	1.43	0.53	1.08	✓	✓	✓	✓
(II, 6, 1.2, 4)	23,713	23,918	22,647	23,495	23,143	0.86	5.32	1.77	3.24	✓	✓	✓	✓
(II, 6, 1.2, 8)	38,900	39,857	36,518	39,291	37,481	2.40	8.38	1.42	5.96	✓	✓	✓	✓
(II, 6, 1.6, 2)	12,803	12,980	12,565	12,950	12,641	1.36	3.19	0.23	2.61	✓	✓	○	✓
(II, 6, 1.6, 4)	20,739	20,835	19,143	20,596	19,514	0.46	8.12	1.15	6.34	✓	✓	✓	✓
(II, 6, 1.6, 8)	36,012	35,694	32,280	36,315	33,091	-0.89	9.57	-1.74	7.29	○	✓	×	✓
(II, 12, 1.0, 2)	31,237	31,572	31,294	31,350	31,363	1.06	0.88	0.70	0.66	✓	✓	✓	✓
(II, 12, 1.0, 4)	44,432	44,839	43,932	44,293	44,275	0.91	2.02	1.22	1.26	✓	✓	✓	✓
(II, 12, 1.0, 8)	70,918	71,140	69,275	70,552	69,967	0.31	2.62	0.83	1.65	○	✓	✓	✓
(II, 12, 1.2, 2)	28,297	28,808	28,358	28,529	28,421	1.77	1.56	0.97	1.34	✓	✓	✓	✓
(II, 12, 1.2, 4)	41,329	41,892	40,363	41,123	40,899	1.34	3.65	1.84	2.37	✓	✓	✓	✓
(II, 12, 1.2, 8)	66,829	66,968	64,456	67,033	65,847	0.21	3.75	-0.10	1.67	○	✓	○	✓
(II, 12, 1.6, 2)	23,345	23,845	23,541	23,679	23,533	2.10	1.27	0.69	1.31	✓	✓	✓	✓
(II, 12, 1.6, 4)	36,013	36,331	34,940	35,866	35,717	0.88	3.83	1.28	1.69	✓	✓	✓	✓
(II, 12, 1.6, 8)	60,937	61,261	58,036	61,627	59,643	0.53	5.26	-0.60	2.64	○	✓	○	✓

Table 6: Computational results for the second airline network for the case where  $n = 6$ .

prob.	up. bnd.	prob.	up. bnd.
(I, 6, 1.0, 2)	18,859	(II, 6, 1.0, 2)	19,188
(I, 6, 1.0, 4)	26,817	(II, 6, 1.0, 4)	27,603
(I, 6, 1.0, 8)	42,735	(II, 6, 1.0, 8)	44,432
(I, 6, 1.2, 2)	17,200	(II, 6, 1.2, 2)	16,944
(I, 6, 1.2, 4)	25,129	(II, 6, 1.2, 4)	25,357
(I, 6, 1.2, 8)	41,036	(II, 6, 1.2, 8)	42,186
(I, 6, 1.6, 2)	14,294	(II, 6, 1.6, 2)	13,935
(I, 6, 1.6, 4)	22,219	(II, 6, 1.6, 4)	22,346
(I, 6, 1.6, 8)	38,122	(II, 6, 1.6, 8)	39,175
(I, 12, 1.0, 2)	34,115	(II, 12, 1.0, 2)	33,372
(I, 12, 1.0, 4)	48,455	(II, 12, 1.0, 4)	47,271
(I, 12, 1.0, 8)	77,136	(II, 12, 1.0, 8)	75,068
(I, 12, 1.2, 2)	31,700	(II, 12, 1.2, 2)	30,640
(I, 12, 1.2, 4)	46,035	(II, 12, 1.2, 4)	44,535
(I, 12, 1.2, 8)	74,710	(II, 12, 1.2, 8)	72,332
(I, 12, 1.6, 2)	26,798	(II, 12, 1.6, 2)	25,661
(I, 12, 1.6, 4)	40,943	(II, 12, 1.6, 4)	39,354
(I, 12, 1.6, 8)	69,618	(II, 12, 1.6, 8)	67,151

Table 7: Optimal objective value of problem (17)-(19) for all of our test problems.

benchmark strategy	first airline network		second airline network	
	$N = 6$	$N = 12$	$N = 6$	$N = 12$
SDR/SDD	34.843	169.734	26.734	130.921
DLP	0.002	0.004	0.002	0.003
RLP	0.050	0.162	0.043	0.116
FD	0.069	0.461	0.026	0.125

Table 8: CPU seconds for SDR, SDD, DLP, RLP and FD to compute one set of bid prices.

ONLINE SUPPLEMENT

A VERIFYING (A.1)-(A.3) IN PROPOSITION 1

In this section, we establish that the cumulative revenue function in (6) satisfies (A.1)-(A.3) in the proof of Proposition 1. The next proposition shows that (A.1) holds.

**Proposition 1** *We have  $\mathbb{E}\{\partial_i^\Lambda R_1(x_1, \omega, \alpha, \lambda)\} = \partial_i^\Lambda \mathbb{E}\{R_1(x_1, \omega, \alpha, \lambda)\}$  for all  $\lambda \in \mathbb{R}^{|\mathcal{L}|}$ ,  $i \in \mathcal{L}$ .*

**Proof** Since we have  $|\min\{p_1, p_2\} - \min\{q_1, q_2\}|^2 \leq |p_1 - q_1|^2 + |p_2 - q_2|^2$ ,  $\min\{\cdot, \cdot\} : \mathbb{R}^2 \rightarrow \mathbb{R}$  is Lipschitz. Since  $\theta(\cdot)$  is also Lipschitz and the composition of Lipschitz functions is Lipschitz, the decision function in (5) is Lipschitz when viewed as a function of the bid prices and leg capacities. Moving backwards through the itinerary requests and using the fact that the composition of Lipschitz functions is Lipschitz, one can show that the cumulative revenue function in (6) is Lipschitz when viewed as a function of the bid prices and leg capacities. Therefore,  $R_1(x_1, \omega, \alpha, \cdot)$  is Lipschitz. By the discussion in Section 2.1, the derivative of  $R_1(x_1, \omega, \alpha, \cdot)$  with respect to the bid price of flight leg  $i$  evaluated at bid prices  $\lambda$  exists w.p.1. Finally, the cumulative revenue function is bounded by  $\tau B_r$  w.p.1. In this case, the result follows from Lemma 6.3.1 in Glasserman (1994), which we briefly state in Section E of the online supplement for completeness.  $\square$

The next proposition shows that (A.2) holds.

**Proposition 2** *We have  $|\partial_i^\Lambda R_1(x_1, \omega, \alpha, \lambda)| \leq B_R^\Lambda$  w.p.1 for a finite scalar  $B_R^\Lambda$ .*

**Proof** All statements in the proof are in w.p.1 sense. We first show that

$$|\partial_i^X R_t(x_t, \omega, \alpha, \lambda)| \leq B_r + B_r (1 + B_a) |\mathcal{L}| + \dots + B_r (1 + B_a)^{\tau-t} |\mathcal{L}|^{\tau-t} \quad (\text{E.1})$$

for all  $x_t \in \mathbb{R}_+^{|\mathcal{L}|}$ ,  $i \in \mathcal{L}$ ,  $t = 1, \dots, \tau$ . Since  $a_{it} \geq 1$  for all  $i \in \mathcal{L}_t^+$ , we have  $|\partial_i^X u_t(x_t, \omega, \alpha, \lambda)| \leq 1$  for all  $i \in \mathcal{L}$  by (13) and (9) implies that

$$|\partial_i^X R_t(x_t, \omega, \alpha, \lambda)| \leq B_r + \sum_{j \in \mathcal{L}} (1 + B_a) |\partial_j^X R_{t+1}(x_t + \alpha_t - u_t(x_t, \omega, \alpha, \lambda) a_t, \omega, \alpha, \lambda)|$$

for all  $x_t \in \mathbb{R}_+^{|\mathcal{L}|}$ ,  $i \in \mathcal{L}$ . Using the inequality above and moving backwards through the itinerary requests, it is easy to show that (E.1) holds. Therefore, if we let  $B_R^X = B_r + B_r (1 + B_a) |\mathcal{L}| + \dots + B_r (1 + B_a)^{\tau-1} |\mathcal{L}|^{\tau-1}$ , then we have  $|\partial_i^X R_t(x_t, \omega, \alpha, \lambda)| \leq B_R^X$  for all  $x_t \in \mathbb{R}_+^{|\mathcal{L}|}$ ,  $i \in \mathcal{L}$ ,  $t = 1, \dots, \tau$ .

Since  $L_\theta$  is the Lipschitz modulus of  $\theta(\cdot)$ , we have  $|\dot{\theta}(p)| \leq L_\theta$  for all  $p \in \mathbb{R}$  and (11) implies that  $|\partial_i^\Lambda u_t(x_t, \omega, \alpha, \lambda)| \leq B_a L_\theta$  for all  $x_t \in \mathbb{R}_+^{|\mathcal{L}|}$ ,  $i \in \mathcal{L}$ ,  $t = 1, \dots, \tau$ . Since  $|\partial_i^X R_t(x_t, \omega, \alpha, \lambda)| \leq B_R^X$ , (8) implies that  $|\partial_i^\Lambda R_t(x_t, \omega, \alpha, \lambda)| \leq B_r B_a L_\theta + |\partial_i^\Lambda R_{t+1}(x_t + \alpha_t - u_t(x_t, \omega, \alpha, \lambda) a_t, \omega, \alpha, \lambda)| + B_a^2 B_R^X |\mathcal{L}| L_\theta$  for all  $x_t \in \mathbb{R}_+^{|\mathcal{L}|}$ ,  $i \in \mathcal{L}$ . Using this inequality and moving backwards through the itinerary requests, it is easy to show that  $|\partial_i^\Lambda R_1(x_1, \omega, \alpha, \lambda)| \leq \tau [B_r B_a L_\theta + B_a^2 B_R^X |\mathcal{L}| L_\theta]$  and the result follows.  $\square$

We introduce some new notation to show that (A.3) holds. We let  $x_t^\lambda$  be the leg capacities just before making the decision for itinerary request  $t$  when we use the policy characterized by bid prices  $\lambda$ . That is, the random variables  $\{x_t^\lambda : t = 1, \dots, \tau\}$  are given recursively by

$$x_{t+1}^\lambda = x_t^\lambda + \alpha_t - u_t(x_t^\lambda, \omega, \alpha, \lambda) a_t, \quad (\text{E.2})$$

with  $x_1^\lambda = x_1$ . The next two lemmas are preliminary results that are useful to show that (A.3) holds. We provide detailed proofs for these lemmas in Sections B and C of the online supplement. Lemma 1 shows that the expected value of the derivative of the decision function with respect to the bid price of flight leg  $i$  is Lipschitz when viewed as a function of the bid prices.

**Lemma 1** *We have  $\mathbb{E}\{|\partial_i^\Lambda u_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^\Lambda u_t(x_t^\gamma, \omega, \alpha, \gamma)|\} \leq L_u^\Lambda \|\lambda - \gamma\|$  for a finite scalar  $L_u^\Lambda$ .*

Lemma 2 shows that the expected value of the derivative of the cumulative revenue function with respect to the remaining capacity on flight leg  $i$  is Lipschitz when viewed as a function of the bid prices.

**Lemma 2** *We have  $\mathbb{E}\{|\partial_i^X R_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^X R_t(x_t^\gamma, \omega, \alpha, \gamma)|\} \leq L_R^X \|\lambda - \gamma\|$  for a finite scalar  $L_R^X$ .*

We are now ready to show that (A.3) holds.

**Proposition 3** *We have  $\mathbb{E}\{|\partial_i^\Lambda R_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^\Lambda R_t(x_t^\gamma, \omega, \alpha, \gamma)|\} \leq L_R^\Lambda \|\lambda - \gamma\|$  for a finite scalar  $L_R^\Lambda$ .*

**Proof** All statements in the proof are in w.p.1 sense. Using (E.2), we can write (8) as

$$\begin{aligned} \partial_i^\Lambda R_t(x_t^\lambda, \omega, \alpha, \lambda) &= r_t \partial_i^\Lambda u_t(x_t^\lambda, \omega, \alpha, \lambda) + \partial_i^\Lambda R_{t+1}(x_{t+1}^\lambda, \omega, \alpha, \lambda) \\ &\quad - \sum_{j \in \mathcal{L}} a_{jt} \partial_i^\Lambda u_t(x_t^\lambda, \omega, \alpha, \lambda) \partial_j^X R_{t+1}(x_{t+1}^\lambda, \omega, \alpha, \lambda). \end{aligned} \quad (\text{E.3})$$

On the other hand, using the fact that  $|p_1 q_1 - p_2 q_2| \leq |p_1| |q_1 - q_2| + |p_1 - p_2| |q_2|$ , we have

$$\begin{aligned} &|\partial_i^\Lambda u_t(x_t^\lambda, \omega, \alpha, \lambda) \partial_j^X R_{t+1}(x_{t+1}^\lambda, \omega, \alpha, \lambda) - \partial_i^\Lambda u_t(x_t^\gamma, \omega, \alpha, \gamma) \partial_j^X R_{t+1}(x_{t+1}^\gamma, \omega, \alpha, \gamma)| \\ &\leq |\partial_i^\Lambda u_t(x_t^\lambda, \omega, \alpha, \lambda)| |\partial_j^X R_{t+1}(x_{t+1}^\lambda, \omega, \alpha, \lambda) - \partial_j^X R_{t+1}(x_{t+1}^\gamma, \omega, \alpha, \gamma)| \\ &\quad + |\partial_i^\Lambda u_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^\Lambda u_t(x_t^\gamma, \omega, \alpha, \gamma)| |\partial_j^X R_{t+1}(x_{t+1}^\gamma, \omega, \alpha, \gamma)|. \end{aligned}$$

Using the fact that  $|\partial_i^\Lambda u_t(x_t^\lambda, \omega, \alpha, \lambda)| \leq B_a L_\theta$  and noting  $B_R^X$  in the proof of Proposition 2, the inequality above implies that

$$\begin{aligned} &|\partial_i^\Lambda u_t(x_t^\lambda, \omega, \alpha, \lambda) \partial_j^X R_{t+1}(x_{t+1}^\lambda, \omega, \alpha, \lambda) - \partial_i^\Lambda u_t(x_t^\gamma, \omega, \alpha, \gamma) \partial_j^X R_{t+1}(x_{t+1}^\gamma, \omega, \alpha, \gamma)| \\ &\leq B_a L_\theta |\partial_j^X R_{t+1}(x_{t+1}^\lambda, \omega, \alpha, \lambda) - \partial_j^X R_{t+1}(x_{t+1}^\gamma, \omega, \alpha, \gamma)| \\ &\quad + B_R^X |\partial_i^\Lambda u_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^\Lambda u_t(x_t^\gamma, \omega, \alpha, \gamma)|. \end{aligned} \quad (\text{E.4})$$

Therefore, by (E.3) and (E.4), we obtain

$$\begin{aligned}
& \mathbb{E}\{|\partial_i^\Lambda R_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^\Lambda R_t(x_t^\gamma, \omega, \alpha, \gamma)|\} \\
& \leq B_r \mathbb{E}\{|\partial_i^\Lambda u_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^\Lambda u_t(x_t^\gamma, \omega, \alpha, \gamma)|\} \\
& \quad + \mathbb{E}\{|\partial_i^\Lambda R_{t+1}(x_{t+1}^\lambda, \omega, \alpha, \lambda) - \partial_i^\Lambda R_{t+1}(x_{t+1}^\gamma, \omega, \alpha, \gamma)|\} \\
& \quad + \sum_{j \in \mathcal{L}} B_a^2 L_\theta \mathbb{E}\{|\partial_j^X R_{t+1}(x_{t+1}^\lambda, \omega, \alpha, \lambda) - \partial_j^X R_{t+1}(x_{t+1}^\gamma, \omega, \alpha, \gamma)|\} \\
& \quad + \sum_{j \in \mathcal{L}} B_a B_R^X \mathbb{E}\{|\partial_i^\Lambda u_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^\Lambda u_t(x_t^\gamma, \omega, \alpha, \gamma)|\},
\end{aligned}$$

in which case Lemmas 1 and 2 imply that

$$\begin{aligned}
& \mathbb{E}\{|\partial_i^\Lambda R_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^\Lambda R_t(x_t^\gamma, \omega, \alpha, \gamma)|\} \\
& \leq B_r L_u^\Lambda \|\lambda - \gamma\| + \mathbb{E}\{|\partial_i^\Lambda R_{t+1}(x_{t+1}^\lambda, \omega, \alpha, \lambda) - \partial_i^\Lambda R_{t+1}(x_{t+1}^\gamma, \omega, \alpha, \gamma)|\} \\
& \quad + B_a^2 |\mathcal{L}| L_\theta L_R^X \|\lambda - \gamma\| + B_a B_R^X |\mathcal{L}| L_u^\Lambda \|\lambda - \gamma\|.
\end{aligned}$$

Using the inequality above and moving backwards through the itinerary requests, it is easy to show that  $\mathbb{E}\{|\partial_i^\Lambda R_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^\Lambda R_t(x_t^\gamma, \omega, \alpha, \gamma)|\} \leq (\tau - t + 1) [B_r L_u^\Lambda + B_a^2 |\mathcal{L}| L_\theta L_R^X + B_a B_R^X |\mathcal{L}| L_u^\Lambda] \|\lambda - \gamma\|$  and the result follows by letting  $L_R^\Lambda = \tau [B_r L_u^\Lambda + B_a^2 |\mathcal{L}| L_\theta L_R^X + B_a B_R^X |\mathcal{L}| L_u^\Lambda]$ .  $\square$

## B PROOF OF LEMMA 1

The next two results are useful when showing Lemma 1.

**Lemma 3** *If  $\|x_t - z_t\| \leq M \|\lambda - \gamma\|$  for some  $M \in \mathbb{R}_+$ , then we have  $|u_t(x_t, \omega, \alpha, \lambda) - u_t(z_t, \omega, \alpha, \gamma)| \leq (L_u + M) \|\lambda - \gamma\|$  w.p.1 for a finite scalar  $L_u$ .*

**Proof** We consider four cases.

**Case 1.** We assume that  $\theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) \leq \min_{j \in \mathcal{L}_t^+} \{[x_{jt} + \alpha_{jt}]/a_{jt}\}$  and  $\theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \gamma_j) \leq \min_{j \in \mathcal{L}_t^+} \{[z_{jt} + \alpha_{jt}]/a_{jt}\}$ . Using (5), we have  $|u_t(x_t, \omega, \alpha, \lambda) - u_t(z_t, \omega, \alpha, \gamma)| = |\theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) - \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \gamma_j)| \leq L_\theta |\sum_{j \in \mathcal{L}} a_{jt} [\lambda_j - \gamma_j]| \leq B_a |\mathcal{L}| L_\theta \|\lambda - \gamma\|$ .

**Case 2.** We assume that  $\theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) \leq \min_{j \in \mathcal{L}_t^+} \{[x_{jt} + \alpha_{jt}]/a_{jt}\}$  and  $\theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \gamma_j) > \min_{j \in \mathcal{L}_t^+} \{[z_{jt} + \alpha_{jt}]/a_{jt}\}$ . Using (5), we have  $|u_t(x_t, \omega, \alpha, \lambda) - u_t(z_t, \omega, \alpha, \gamma)| = |\theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) - \min_{j \in \mathcal{L}_t^+} \{[z_{jt} + \alpha_{jt}]/a_{jt}\}|$  and we consider two subcases.

**Case 2.a.** We assume that  $\theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) \geq \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \gamma_j)$ , in which case we have  $\min_{j \in \mathcal{L}_t^+} \{[z_{jt} + \alpha_{jt}]/a_{jt}\} < \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \gamma_j) \leq \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) \leq \min_{j \in \mathcal{L}_t^+} \{[x_{jt} + \alpha_{jt}]/a_{jt}\}$ . Since  $z_{jt} - M \|\lambda - \gamma\| \leq x_{jt} \leq z_{jt} + M \|\lambda - \gamma\|$  for all  $j \in \mathcal{L}$  and  $M \|\lambda - \gamma\|/a_{jt} \leq M \|\lambda - \gamma\|$  for all  $j \in \mathcal{L}_t^+$ , we obtain

$$\begin{aligned}
|u_t(x_t, \omega, \alpha, \lambda) - u_t(z_t, \omega, \alpha, \gamma)| & \leq \min_{j \in \mathcal{L}_t^+} \{[x_{jt} + \alpha_{jt}]/a_{jt}\} - \min_{j \in \mathcal{L}_t^+} \{[z_{jt} + \alpha_{jt}]/a_{jt}\} \\
& \leq \min_{j \in \mathcal{L}_t^+} \{[z_{jt} + M \|\lambda - \gamma\| + \alpha_{jt}]/a_{jt}\} - \min_{j \in \mathcal{L}_t^+} \{[z_{jt} + \alpha_{jt}]/a_{jt}\} \\
& \leq M \|\lambda - \gamma\|.
\end{aligned}$$

**Case 2.b.** We assume that  $\theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) < \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \gamma_j)$  and consider two (sub)subcases.

**Case 2.b.i.** We assume that  $\theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) \geq \min_{j \in \mathcal{L}_t^+} \{[z_{jt} + \alpha_{jt}]/a_{jt}\}$ , in which case we have  $\min_{j \in \mathcal{L}_t^+} \{[z_{jt} + \alpha_{jt}]/a_{jt}\} \leq \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) \leq \min_{j \in \mathcal{L}_t^+} \{[x_{jt} + \alpha_{jt}]/a_{jt}\}$ . Therefore, we obtain  $|u_t(x_t, \omega, \alpha, \lambda) - u_t(z_t, \omega, \alpha, \gamma)| \leq \min_{j \in \mathcal{L}_t^+} \{[x_{jt} + \alpha_{jt}]/a_{jt}\} - \min_{j \in \mathcal{L}_t^+} \{[z_{jt} + \alpha_{jt}]/a_{jt}\} \leq M \|\lambda - \gamma\|$ , where the second inequality follows from the same argument in Case 2.a.

**Case 2.b.ii.** Assume that  $\theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) < \min_{j \in \mathcal{L}_t^+} \{[z_{jt} + \alpha_{jt}]/a_{jt}\}$ , in which case we have  $\theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) < \min_{j \in \mathcal{L}_t^+} \{[z_{jt} + \alpha_{jt}]/a_{jt}\} < \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \gamma_j)$ . Therefore, we obtain  $|u_t(x_t, \omega, \alpha, \lambda) - u_t(z_t, \omega, \alpha, \gamma)| < \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \gamma_j) - \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) \leq B_a |\mathcal{L}| L_\theta \|\lambda - \gamma\|$ .

The other cases that we do not cover above can be handled in a similar manner. If we combine all cases, then it is easy to see that letting  $L_u = B_a |\mathcal{L}| L_\theta$  suffices.  $\square$

**Lemma 4** We have  $\|x_t^\lambda - x_t^\gamma\| \leq L_X \|\lambda - \gamma\|$  w.p.1 for a finite scalar  $L_X$ .

**Proof** All statements in the proof are in w.p.1 sense. We show by induction that  $\|x_t^\lambda - x_t^\gamma\| \leq [2(1+B_a)L_u + \dots + 2^{t-1}(1+B_a)^{t-1}L_u] \|\lambda - \gamma\|$  for all  $t = 2, \dots, \tau$ , in which case the result follows by letting  $L_X = 2(1+B_a)L_u + \dots + 2^{\tau-1}(1+B_a)^{\tau-1}L_u$  and noting that  $x_1^\lambda = x_1^\gamma = x_1$ . Assuming that the result holds for itinerary request  $t$  and using (E.2), we have

$$\begin{aligned} \|x_{t+1}^\lambda - x_{t+1}^\gamma\| &\leq \|x_t^\lambda - x_t^\gamma\| + B_a |u_t(x_t^\lambda, \omega, \alpha, \lambda) - u_t(x_t^\gamma, \omega, \alpha, \gamma)| \\ &\leq (1+B_a) \|x_t^\lambda - x_t^\gamma\| + (1+B_a) |u_t(x_t^\lambda, \omega, \alpha, \lambda) - u_t(x_t^\gamma, \omega, \alpha, \gamma)| \\ &\leq (1+B_a) [2(1+B_a)L_u + \dots + 2^{t-1}(1+B_a)^{t-1}L_u] \|\lambda - \gamma\| \\ &\quad + (1+B_a) [L_u + 2(1+B_a)L_u + \dots + 2^{t-1}(1+B_a)^{t-1}L_u] \|\lambda - \gamma\| \\ &\leq [2^2(1+B_a)L_u + 2^2(1+B_a)^2L_u + \dots + 2^t(1+B_a)^tL_u] \|\lambda - \gamma\|, \end{aligned}$$

where the third inequality follows from the induction hypothesis and Lemma 3. Therefore, the result holds for itinerary request  $t+1$ . We complete the induction argument by noting that

$$\|x_2^\lambda - x_2^\gamma\| \leq \|x_1^\lambda - x_1^\gamma\| + B_a |u_1(x_1^\lambda, \omega, \alpha, \lambda) - u_1(x_1^\gamma, \omega, \alpha, \gamma)| \leq B_a L_u \|\lambda - \gamma\| \leq 2(1+B_a)L_u \|\lambda - \gamma\|,$$

where we use Lemma 3 and the fact that  $\|x_1^\lambda - x_1^\gamma\| \leq 0 \|\lambda - \gamma\|$ .  $\square$

We are now ready to show Lemma 1. We begin by considering four cases.

**Case 1.** We assume that  $\theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) \leq \min_{j \in \mathcal{L}_t^+} \{[x_{jt}^\lambda + \alpha_{jt}]/a_{jt}\}$  and  $\theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \gamma_j) \leq \min_{j \in \mathcal{L}_t^+} \{[x_{jt}^\gamma + \alpha_{jt}]/a_{jt}\}$ . By (11), we have  $|\partial_i^\Lambda u_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^\Lambda u_t(x_t^\gamma, \omega, \alpha, \gamma)| = a_{it} |\dot{\theta}(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) - \dot{\theta}(r_t - \sum_{j \in \mathcal{L}} a_{jt} \gamma_j)| \leq B_a L_\theta |\sum_{j \in \mathcal{L}} a_{jt} [\lambda_j - \gamma_j]| \leq B_a^2 |\mathcal{L}| L_\theta \|\lambda - \gamma\|$ .

**Case 2.** We assume that  $\theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) \leq \min_{j \in \mathcal{L}_t^+} \{[x_{jt}^\lambda + \alpha_{jt}]/a_{jt}\}$  and  $\theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \gamma_j) > \min_{j \in \mathcal{L}_t^+} \{[x_{jt}^\gamma + \alpha_{jt}]/a_{jt}\}$ . We have  $|\partial_i^\Lambda u_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^\Lambda u_t(x_t^\gamma, \omega, \alpha, \gamma)| = a_{it} |\dot{\theta}(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j)| \leq B_a L_\theta$  by (11).

**Case 3.** We assume that  $\theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) > \min_{j \in \mathcal{L}_t^+} \{[x_{jt}^\lambda + \alpha_{jt}]/a_{jt}\}$  and  $\theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \gamma_j) \leq \min_{j \in \mathcal{L}_t^+} \{[x_{jt}^\gamma + \alpha_{jt}]/a_{jt}\}$ , in which case the same argument in the second case yields  $|\partial_i^\Lambda u_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^\Lambda u_t(x_t^\gamma, \omega, \alpha, \gamma)| \leq B_a L_\theta$ .

**Case 4.** We assume that  $\theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) > \min_{j \in \mathcal{L}_t^+} \{[x_{jt}^\lambda + \alpha_{jt}]/a_{jt}\}$  and  $\theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \gamma_j) > \min_{j \in \mathcal{L}_t^+} \{[x_{jt}^\gamma + \alpha_{jt}]/a_{jt}\}$ . In this case, (11) implies that  $|\partial_i^\Lambda u_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^\Lambda u_t(x_t^\gamma, \omega, \alpha, \gamma)| = 0$ .

We now obtain a bound on the probability of the second case. Lemma 4 shows that  $\|x_t^\lambda - x_t^\gamma\| \leq L_X \|\lambda - \gamma\|$  w.p.1 for a finite scalar  $L_X$ . Therefore, we have

$$\begin{aligned} & |[a_{it} \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \gamma_j) - x_{it}^\gamma] - [a_{it} \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) - x_{it}^\lambda]| \\ & \leq B_a |\theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \gamma_j) - \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j)| + \|x_t^\lambda - x_t^\gamma\| \\ & \leq B_a^2 |\mathcal{L}| L_\theta \|\lambda - \gamma\| + L_X \|\lambda - \gamma\| \quad (\text{E.5}) \end{aligned}$$

w.p.1. On the other hand, for two sets of random variables  $\{P_i : i \in \mathcal{A}\}$  and  $\{Q_i : i \in \mathcal{A}\}$ , we have  $\mathbb{P}\{p \leq \min_{j \in \mathcal{A}} \{P_j\}, q > \min_{j \in \mathcal{A}} \{Q_j\}\} \leq \sum_{i \in \mathcal{A}} \mathbb{P}\{p \leq \min_{j \in \mathcal{A}} \{P_j\}, q > Q_i\} \leq \sum_{i \in \mathcal{A}} \mathbb{P}\{p \leq P_i, q > Q_i\}$ . Therefore, we obtain a bound on the probability of the second case by noting that

$$\begin{aligned} & \mathbb{P}\left\{\theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) \leq \min_{j \in \mathcal{L}_t^+} \{[x_{jt}^\lambda + \alpha_{jt}]/a_{jt}\}, \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \gamma_j) > \min_{j \in \mathcal{L}_t^+} \{[x_{jt}^\gamma + \alpha_{jt}]/a_{jt}\}\right\} \\ & \leq \sum_{i \in \mathcal{L}} \mathbb{P}\left\{\theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) \leq [x_{it}^\lambda + \alpha_{it}]/a_{it}, \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \gamma_j) > [x_{it}^\gamma + \alpha_{it}]/a_{it}\right\} \\ & = \sum_{i \in \mathcal{L}} \mathbb{P}\left\{a_{it} \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) - x_{it}^\lambda \leq \alpha_{it} < a_{it} \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \gamma_j) - x_{it}^\gamma\right\} \\ & \leq [B_a^2 |\mathcal{L}|^2 L_\theta + |\mathcal{L}| L_X] \|\lambda - \gamma\|/\epsilon, \end{aligned}$$

where the second inequality follows from (E.5), and the fact that  $\alpha_{it}$  is uniformly distributed over the interval  $[0, \epsilon]$  and it is independent of  $x_t^\lambda$  and  $a_t$ . The same bound applies to the probability of the third case. Combining the four cases at the beginning of the proof and using the trivial bound of one for the probability of the first case, we obtain

$$\mathbb{E}\{|\partial_i^\Lambda u_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^\Lambda u_t(x_t^\gamma, \omega, \alpha, \gamma)|\} \leq B_a^2 |\mathcal{L}| L_\theta \|\lambda - \gamma\| + 2 [B_a^2 |\mathcal{L}|^2 L_\theta + |\mathcal{L}| L_X] \|\lambda - \gamma\| B_a L_\theta/\epsilon$$

and the result follows.

## C PROOF OF LEMMA 2

The next result is useful when showing Lemma 2.

**Lemma 5** *We have  $\mathbb{E}\{|\partial_i^X u_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^X u_t(x_t^\gamma, \omega, \alpha, \gamma)|\} \leq L_u^X \|\lambda - \gamma\|$  for a finite scalar  $L_u^X$ .*

**Proof** By (15),  $|\partial_i^X u_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^X u_t(x_t^\gamma, \omega, \alpha, \gamma)|$  is equal to  $1/a_{it}$  for the following four cases and is equal to zero otherwise.

**Case 1.**  $i \in \operatorname{argmin}_{j \in \mathcal{L}_t^+} \{[x_{jt}^\lambda + \alpha_{jt}]/a_{jt}\}$ ,  $[x_{it}^\lambda + \alpha_{it}]/a_{it} \leq \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j)$ ,  
 $i \in \operatorname{argmin}_{j \in \mathcal{L}_t^+} \{[x_{jt}^\gamma + \alpha_{jt}]/a_{jt}\}$ ,  $[x_{it}^\gamma + \alpha_{it}]/a_{it} > \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \gamma_j)$ .

**Case 2.**  $i \in \operatorname{argmin}_{j \in \mathcal{L}_t^+} \{[x_{jt}^\lambda + \alpha_{jt}]/a_{jt}\}, [x_{it}^\lambda + \alpha_{it}]/a_{it} > \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j),$   
 $i \in \operatorname{argmin}_{j \in \mathcal{L}_t^+} \{[x_{jt}^\gamma + \alpha_{jt}]/a_{jt}\}, [x_{it}^\gamma + \alpha_{it}]/a_{it} \leq \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \gamma_j).$

**Case 3.**  $i \in \operatorname{argmin}_{j \in \mathcal{L}_t^+} \{[x_{jt}^\lambda + \alpha_{jt}]/a_{jt}\}, [x_{it}^\lambda + \alpha_{it}]/a_{it} \leq \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j),$   
 $i \notin \operatorname{argmin}_{j \in \mathcal{L}_t^+} \{[x_{jt}^\gamma + \alpha_{jt}]/a_{jt}\}.$

**Case 4.**  $i \notin \operatorname{argmin}_{j \in \mathcal{L}_t^+} \{[x_{jt}^\lambda + \alpha_{jt}]/a_{jt}\},$   
 $i \in \operatorname{argmin}_{j \in \mathcal{L}_t^+} \{[x_{jt}^\gamma + \alpha_{jt}]/a_{jt}\}, [x_{it}^\gamma + \alpha_{it}]/a_{it} \leq \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \gamma_j).$

Since we have either  $i \in \operatorname{argmin}_{j \in \mathcal{L}_t^+} \{[x_{jt}^\lambda + \alpha_{jt}]/a_{jt}\}$  or  $i \in \operatorname{argmin}_{j \in \mathcal{L}_t^+} \{[x_{jt}^\gamma + \alpha_{jt}]/a_{jt}\}$  for these four cases and  $a_{it} \geq 1$  for all  $i \in \mathcal{L}_t^+$ , we also have  $|\partial_i^X u_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^X u_t(x_t^\gamma, \omega, \alpha, \gamma)| \leq 1.$

We obtain a bound on the probability of the first case by noting that

$$\begin{aligned} \mathbb{P} \left\{ i \in \operatorname{argmin}_{j \in \mathcal{L}_t^+} \{[x_{jt}^\lambda + \alpha_{jt}]/a_{jt}\}, [x_{it}^\lambda + \alpha_{it}]/a_{it} \leq \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j), \right. \\ \left. i \in \operatorname{argmin}_{j \in \mathcal{L}_t^+} \{[x_{jt}^\gamma + \alpha_{jt}]/a_{jt}\}, [x_{it}^\gamma + \alpha_{it}]/a_{it} > \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \gamma_j) \right\} \\ \leq \mathbb{P} \left\{ a_{it} \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \gamma_j) - x_{it}^\gamma < \alpha_{it} \leq a_{it} \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j) - x_{it}^\lambda \right\}. \end{aligned}$$

By (E.5), the probability on right side above is bounded by  $[B_a^2 |\mathcal{L}| L_\theta + L_X] \|\lambda - \gamma\|/\epsilon.$  The same bound applies to the probability of the second case. Using  $\mathbb{P}_\omega \{ \cdot \}$  to denote probability conditional on the filtration generated by  $\omega$ , we obtain a bound on the probability of the third case by noting that

$$\begin{aligned} \mathbb{P}_\omega \left\{ i \in \operatorname{argmin}_{j \in \mathcal{L}_t^+} \{[x_{jt}^\lambda + \alpha_{jt}]/a_{jt}\}, [x_{it}^\lambda + \alpha_{it}]/a_{it} \leq \theta(r_t - \sum_{j \in \mathcal{L}} a_{jt} \lambda_j), \right. \\ \left. i \notin \operatorname{argmin}_{j \in \mathcal{L}_t^+} \{[x_{jt}^\gamma + \alpha_{jt}]/a_{jt}\} \right\} \\ \leq \mathbb{P}_\omega \left\{ i \in \operatorname{argmin}_{j \in \mathcal{L}_t^+} \{[x_{jt}^\lambda + \alpha_{jt}]/a_{jt}\}, i \notin \operatorname{argmin}_{j \in \mathcal{L}_t^+} \{[x_{jt}^\gamma + \alpha_{jt}]/a_{jt}\} \right\} \\ \leq \sum_{i' \in \mathcal{L}_t^+ \setminus \{i\}} \mathbb{P}_\omega \left\{ i \in \operatorname{argmin}_{j \in \mathcal{L}_t^+} \{[x_{jt}^\lambda + \alpha_{jt}]/a_{jt}\}, [x_{i't}^\gamma + \alpha_{i't}]/a_{i't} < [x_{it}^\gamma + \alpha_{it}]/a_{it} \right\} \\ \leq \sum_{i' \in \mathcal{L}_t^+ \setminus \{i\}} \mathbb{P}_\omega \left\{ [x_{i't}^\lambda + \alpha_{i't}]/a_{i't} \leq [x_{i't}^\gamma + \alpha_{i't}]/a_{i't}, [x_{i't}^\gamma + \alpha_{i't}]/a_{i't} < [x_{it}^\gamma + \alpha_{it}]/a_{it} \right\} \\ = \sum_{i' \in \mathcal{L}_t^+ \setminus \{i\}} \mathbb{P}_\omega \left\{ \frac{x_{i't}^\gamma}{a_{i't}} - \frac{x_{it}^\gamma}{a_{it}} < \frac{\alpha_{it}}{a_{it}} - \frac{\alpha_{i't}}{a_{i't}} \leq \frac{x_{i't}^\lambda}{a_{i't}} - \frac{x_{it}^\lambda}{a_{it}} \right\}. \end{aligned} \quad (\text{E.6})$$

Using the fact that the random variables  $\alpha_{it}$  and  $\alpha_{i't}$  are uniformly distributed over the interval  $[0, \epsilon],$  and they are independent of each other and  $\omega$ , a straightforward computation shows that

$$\mathbb{P}_\omega \left\{ \frac{\alpha_{it}}{a_{it}} - \frac{\alpha_{i't}}{a_{i't}} \leq p \right\} = \begin{cases} 0 & \text{if } p \leq -\frac{\epsilon}{a_{i't}} \\ \frac{a_{it}}{2a_{i't}} + \frac{a_{it}}{\epsilon} p + \frac{a_{i't}}{2a_{it}} \left[ \frac{a_{it} p}{\epsilon} \right]^2 & \text{if } -\frac{\epsilon}{a_{i't}} < p \leq 0 \text{ and } \frac{a_{it}}{a_{i't}} + \frac{a_{it} p}{\epsilon} \leq 1 \\ 1 + \frac{a_{i't}}{2a_{it}} \left[ \frac{a_{it} p}{\epsilon} \right]^2 - \frac{a_{i't}}{2a_{it}} \left[ 1 - \frac{a_{it} p}{\epsilon} \right]^2 & \text{if } -\frac{\epsilon}{a_{i't}} < p \leq 0 \text{ and } \frac{a_{it}}{a_{i't}} + \frac{a_{it} p}{\epsilon} > 1 \\ \frac{a_{i't}}{2a_{it}} + \frac{a_{it} p}{\epsilon} & \text{if } 0 < p \leq \frac{\epsilon}{a_{it}} \text{ and } \frac{a_{it}}{a_{i't}} + \frac{a_{it} p}{\epsilon} \leq 1 \\ 1 - \frac{a_{i't}}{2a_{it}} \left[ 1 - \frac{a_{it} p}{\epsilon} \right]^2 & \text{if } 0 < p \leq \frac{\epsilon}{a_{it}} \text{ and } \frac{a_{it}}{a_{i't}} + \frac{a_{it} p}{\epsilon} > 1 \\ 1 & \text{if } p > \frac{\epsilon}{a_{it}}. \end{cases}$$

Using this expression, it is easy to check that the cumulative distribution function of  $[\alpha_{it}/a_{it}] - [\alpha_{i't}/a_{i't}]$  conditional on  $\omega$  is Lipschitz with modulus  $B_a/\epsilon$  for all  $i, i' \in \mathcal{L}_t^+$ . Therefore, we have

$$\mathbb{P}_\omega \left\{ q \leq \frac{\alpha_{it}}{a_{it}} - \frac{\alpha_{i't}}{a_{i't}} \leq p \right\} \leq \frac{B_a}{\epsilon} |p - q|$$

w.p.1 for all  $i, i' \in \mathcal{L}_t^+$ . On the other hand, since we have  $a_{it} \geq 1$  for all  $i \in \mathcal{L}_t^+$ , Lemma 4 implies that

$$\left| \left[ \frac{x_{i't}^\lambda}{a_{i't}} - \frac{x_{it}^\lambda}{a_{it}} \right] - \left[ \frac{x_{i't}^\gamma}{a_{i't}} - \frac{x_{it}^\gamma}{a_{it}} \right] \right| \leq [|x_{it}^\lambda - x_{it}^\gamma| + |x_{i't}^\lambda - x_{i't}^\gamma|] \leq 2 L_X \|\lambda - \gamma\|$$

w.p.1 for all  $i, i' \in \mathcal{L}_t^+$ . By the last two inequalities, we obtain

$$\mathbb{P}_\omega \left\{ \frac{x_{i't}^\gamma}{a_{i't}} - \frac{x_{it}^\gamma}{a_{it}} < \frac{\alpha_{it}}{a_{it}} - \frac{\alpha_{i't}}{a_{i't}} \leq \frac{x_{i't}^\lambda}{a_{i't}} - \frac{x_{it}^\lambda}{a_{it}} \right\} \leq \frac{2 B_a}{\epsilon} L_X \|\lambda - \gamma\| \quad (\text{E.7})$$

w.p.1 for all  $i, i' \in \mathcal{L}_t^+$ . If we combine (E.6) and (E.7), and integrate out the conditional expectation, then it is easy to see that the probability of the third case is bounded by  $2 B_a |\mathcal{L}| L_X \|\lambda - \gamma\|/\epsilon$ . The same bound applies to the probability of the fourth case. Combining the bounds on the probabilities of the four cases and noting that  $|\partial_i^X u_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^X u_t(x_t^\gamma, \omega, \alpha, \gamma)| \leq 1$  for these four cases, we obtain  $\mathbb{E}\{|\partial_i^X u_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^X u_t(x_t^\gamma, \omega, \alpha, \gamma)|\} \leq 2[B_a^2 |\mathcal{L}| L_\theta + L_X] \|\lambda - \gamma\|/\epsilon + 4 B_a |\mathcal{L}| L_X \|\lambda - \gamma\|/\epsilon$  and the result follows.  $\square$

We are now ready to show Lemma 2. All statements in the proof are in w.p.1 sense. Using (E.2), we write (9) as

$$\begin{aligned} \partial_i^X R_t(x_t^\lambda, \omega, \alpha, \lambda) &= r_t \partial_i^X u_t(x_t^\lambda, \omega, \alpha, \lambda) \\ &\quad + \sum_{j \in \mathcal{L}} [\mathbf{1}(j = i) - a_{jt} \partial_i^X u_t(x_t^\lambda, \omega, \alpha, \lambda)] \partial_j^X R_{t+1}(x_{t+1}^\lambda, \omega, \alpha, \lambda). \end{aligned} \quad (\text{E.8})$$

Since  $a_{it} \geq 1$  for all  $i \in \mathcal{L}_t^+$ , we have  $|\partial_i^X u_t(x_t^\lambda, \omega, \alpha, \lambda)| \leq 1$  for all  $i \in \mathcal{L}$  by (13) and we can use an argument similar to the one in (E.4) to obtain

$$\begin{aligned} &|\partial_i^X u_t(x_t^\lambda, \omega, \alpha, \lambda) \partial_j^X R_{t+1}(x_{t+1}^\lambda, \omega, \alpha, \lambda) - \partial_i^X u_t(x_t^\gamma, \omega, \alpha, \gamma) \partial_j^X R_{t+1}(x_{t+1}^\gamma, \omega, \alpha, \gamma)| \\ &\leq B_R^X |\partial_i^X u_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^X u_t(x_t^\gamma, \omega, \alpha, \gamma)| + |\partial_j^X R_{t+1}(x_{t+1}^\lambda, \omega, \alpha, \lambda) - \partial_j^X R_{t+1}(x_{t+1}^\gamma, \omega, \alpha, \gamma)|, \end{aligned}$$

where  $B_R^X$  is as in the proof of Proposition 2. Therefore, (E.8) and the inequality above imply that

$$\begin{aligned} &\mathbb{E}\{|\partial_i^X R_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^X R_t(x_t^\gamma, \omega, \alpha, \gamma)|\} \\ &\leq B_r \mathbb{E}\{|\partial_i^X u_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^X u_t(x_t^\gamma, \omega, \alpha, \gamma)|\} \\ &\quad + \sum_{j \in \mathcal{L}} \mathbb{E}\{|\partial_j^X R_{t+1}(x_{t+1}^\lambda, \omega, \alpha, \lambda) - \partial_j^X R_{t+1}(x_{t+1}^\gamma, \omega, \alpha, \gamma)|\} \\ &\quad + \sum_{j \in \mathcal{L}} B_a B_R^X \mathbb{E}\{|\partial_i^X u_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^X u_t(x_t^\gamma, \omega, \alpha, \gamma)|\} \\ &\quad + \sum_{j \in \mathcal{L}} B_a \mathbb{E}\{|\partial_j^X R_{t+1}(x_{t+1}^\lambda, \omega, \alpha, \lambda) - \partial_j^X R_{t+1}(x_{t+1}^\gamma, \omega, \alpha, \gamma)|\}, \end{aligned}$$

in which case Lemma 5 implies that

$$\begin{aligned} &\mathbb{E}\{|\partial_i^X R_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^X R_t(x_t^\gamma, \omega, \alpha, \gamma)|\} \\ &\leq B_r L_u^X \|\lambda - \gamma\| + B_a B_R^X |\mathcal{L}| L_u^X \|\lambda - \gamma\| \\ &\quad + \sum_{j \in \mathcal{L}} (1 + B_a) \mathbb{E}\{|\partial_j^X R_{t+1}(x_{t+1}^\lambda, \omega, \alpha, \lambda) - \partial_j^X R_{t+1}(x_{t+1}^\gamma, \omega, \alpha, \gamma)|\}. \end{aligned} \quad (\text{E.9})$$

Letting  $M = B_r L_u^X + B_a B_R^X |\mathcal{L}|$ , we now use (E.9) to show by induction that  $\mathbb{E}\{|\partial_i^X R_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^X R_t(x_t^\gamma, \omega, \alpha, \gamma)|\} \leq [M + (1 + B_a) |\mathcal{L}| M + \dots + (1 + B_a)^{\tau-t} |\mathcal{L}|^{\tau-t} M] \|\lambda - \gamma\|$  for all  $i \in \mathcal{L}$ ,  $t = 1, \dots, \tau$ , in which case the result follows by letting  $L_R^X = M + (1 + B_a) |\mathcal{L}| M + \dots + (1 + B_a)^{\tau-1} |\mathcal{L}|^{\tau-1} M$ . Assuming that the result holds for itinerary request  $t + 1$  and using (E.9), we have

$$\begin{aligned} & \mathbb{E}\{|\partial_i^X R_t(x_t^\lambda, \omega, \alpha, \lambda) - \partial_i^X R_t(x_t^\gamma, \omega, \alpha, \gamma)|\} \\ & \leq M \|\lambda - \gamma\| + (1 + B_a) |\mathcal{L}| [M + (1 + B_a) |\mathcal{L}| M + \dots + (1 + B_a)^{\tau-t-1} |\mathcal{L}|^{\tau-t-1} M] \|\lambda - \gamma\| \\ & = [M + (1 + B_a) |\mathcal{L}| M + \dots + (1 + B_a)^{\tau-t} |\mathcal{L}|^{\tau-t} M] \|\lambda - \gamma\|. \end{aligned}$$

We complete the induction argument by noting that  $\mathbb{E}\{|\partial_i^X R_\tau(x_\tau^\lambda, \omega, \alpha, \lambda) - \partial_i^X R_\tau(x_\tau^\gamma, \omega, \alpha, \gamma)|\} \leq M \|\lambda - \gamma\|$  by (E.9).

#### D PROPOSITION 4.1 IN BERTSEKAS AND TSITSIKLIS (1996)

For a function  $f(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}$ , we consider the algorithm

$$\lambda^{k+1} = \lambda^k + \sigma^k s^k$$

to solve the problem  $\max_\lambda f(\lambda)$ , where  $\{\sigma^k\}_k$  is a sequence of step size parameters and  $\{s^k\}_k$  is a sequence of step directions. We let  $\mathcal{F}^k$  be the filtration generated by the random variables  $\{\lambda^1, s^1, \dots, s^{k-1}\}$  in this algorithm and assume that the following statements hold for all  $\lambda, \gamma \in \mathbb{R}^n$ .

- (B.0) We have  $f(\lambda) \geq 0$ .
- (B.1) We have  $\mathbb{E}\{s^k | \mathcal{F}^k\} = \nabla f(\lambda^k)$  w.p.1 for all  $k = 1, 2, \dots$
- (B.2) There exists a finite scalar  $M_s$  such that we have  $\|s^k\| \leq M_s$  w.p.1 for all  $k = 1, 2, \dots$
- (B.3) There exists a finite scalar  $L_f$  such that we have  $\|\nabla f(\lambda) - \nabla f(\gamma)\| \leq L_f \|\lambda - \gamma\|$ .

In this case, the next convergence result is from Proposition 4.1 in Bertsekas and Tsitsiklis (1996).

**Proposition 4** *Assume that the sequence of step size parameters  $\{\sigma^k\}_k$  are  $\mathcal{F}^k$ -measurable and satisfy  $\sigma^k \geq 0$  for all  $k = 1, 2, \dots$ ,  $\sum_{k=1}^\infty \sigma^k = \infty$  and  $\sum_{k=1}^\infty [\sigma^k]^2 < \infty$  w.p.1. If the sequence  $\{\lambda^k\}_k$  is generated by the algorithm above and (B.0)-(B.3) hold, then we have  $\lim_{k \rightarrow \infty} \nabla f(\lambda^k) = 0$  w.p.1 and every limit point  $\lambda^*$  of the sequence  $\{\lambda^k\}_k$  satisfies  $\nabla f(\lambda^*) = 0$  w.p.1.*

#### E LEMMA 6.3.1 IN GLASSERMAN (1994)

For a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and a function  $f(\cdot, \cdot) : \mathbb{R}^n \times \Omega \rightarrow \mathbb{R}$ , we assume that the following statements hold for all  $\lambda, \gamma \in \mathbb{R}^n$ .

- (C.1) The function  $f(\cdot, \omega)$  is differentiable at  $\lambda$  for  $\mathbb{P}$ -almost all values of  $\omega$ .
- (C.2) There exists a finite scalar  $L_f$  such that we have  $\|f(\lambda, \omega) - f(\gamma, \omega)\| \leq L_f \|\lambda - \gamma\|$  for  $\mathbb{P}$ -almost all values of  $\omega$ .

In this case, the next result is from Lemma 6.3.1 in Glasserman (1994).

**Lemma 6** *Assume that there exists a finite scalar  $M_f$  that satisfies  $\mathbb{E}\{|f(\lambda, \omega)|\} \leq M_f$  for all  $\lambda \in \mathbb{R}^n$ . In this case, if (C.1) and (C.2) hold, then  $\nabla \mathbb{E}\{f(\lambda, \omega)\}$  exists and we have  $\nabla \mathbb{E}\{f(\lambda, \omega)\} = \mathbb{E}\{\nabla f(\lambda, \omega)\}$  for all  $\lambda \in \mathbb{R}^n$ .*