An Approximate Dynamic Programming Approach to Network Revenue Management

Vivek F. Farias *   Benjamin Van Roy †

April 23, 2007

Abstract

We develop an approximation algorithm for a dynamic capacity allocation problem with Markov modulated customer arrival rates. For each time period and each state of the modulating process, the algorithm approximates the dynamic programming value function using a concave function that is separable across resource inventory levels. We establish via computational experiments that our algorithm increases expected revenue, in some cases by close to 8%, relative to a deterministic linear program that is widely used for bid-price control.

Keywords: Stochastic Control, Approximate Dynamic Programming, Network Revenue Management

*Electrical Engineering, Stanford University, email: vivekf@stanford.edu.

†Management Science & Engineering and Electrical Engineering, Stanford University, email: bvr@stanford.edu.

1 Introduction

Network revenue management refers to the activity of a vendor who is endowed with limited quantities of multiple resources and sells products, each composed of a bundle of resources, controlling their availability and/or prices over time with an aim to maximize revenue. The airline industry is perhaps the most notable source for such problems. An airline typically operates flights on each leg of a network of cities and offers for sale 'fare products' composed of seats on one or more of these legs. Each fare product is associated, among other things, with some fixed price which the airline receives upon its sale. Since demand for fare products is stochastic and capacity on each leg limited, the airline's problem becomes one of deciding which of its fare products to offer for sale at each point in time over a finite sales period so as to maximize expected revenues. This paper presents a new algorithm for this widely studied problem.

For most models of interest, the dynamic capacity allocation problem we have described can be cast as a dynamic program, albeit one with a computationally intractable state-space even for networks of moderate size. As such, revenue management techniques have typically resorted to heuristic control strategies. Early heuristics for the problem were based primarily on the solutions to a set of single resource problems solved for each leg. Today's state of the art techniques involve 'bid-price' control. A generic bid-price control scheme might work as follows: At each point in time the scheme generates a bid-price for a seat or unit of capacity on each leg of the network. A request for a particular fare product at that point in time is then accepted if and only if the revenue garnered from the sale is no smaller than the sum of the bid prices of the resources or seats that constitute that fare product. There is a vast array of available algorithms that may be used in the generation of bid-prices. There are two important dimensions along which such an algorithm must be evaluated. One, of course, is revenues generated from the strategy. Since bid-prices must be generated in real time, a second important dimension is the efficiency of the procedure used to generate them. A simple approach to this problem which has found wide-spread acceptance involves the solution of a single linear program referred to as the 'deterministic' LP (DLP). This approach and associated bid-price techniques have found widespread use in modern revenue management systems and are believed to have generated incremental revenues on the order of 1-2% greater than previously used 'fare-class' level heuristics (see P.P.Belobaba and Lee (2000), P.P.Belobaba (2001)).

The algorithm we present applies to models with Markov-modulated customer arrival rates. This represents a substantial generalization of the deterministic arrival rate arrival process models generally considered in the literature and accommodates a broad class of demand forecasting models. We demonstrate via a sequence of computational examples that our algorithm consistently produces higher revenues than a strategy using bid-prices computed via re-solution of the DLP *at each time step*. While the performance gain relative to the DLP is modest ($\sim 1\%$) for a model with time homogeneous arrivals, this gain increases significantly when arrival rates vary stochastically. Even for a simple arrival process in which the modulating process has three states, we report relative performance gains of up to about 8% over a DLP approach suitably modified to account for the stochasticity in arrival rates.

Our algorithm is based on a linear programming approach to approximate dynamic programming (de Farias and Van Roy (2003), de Farias and Van Roy (2004)). A linear program is solved to produce for each modulating process state and each time an approximation to the optimal value function that is separable across resource inventory levels. A heuristic is then given by the greedy policy with respect to this approximate value function. This policy can be interpreted in terms of bid-price control for which bid prices are generated at each point in time via a table look-up, which takes far less time than solving the DLP.

The ALP has as many constraints as the size of the state space and practical solution requires a constraint sampling procedure. We exploit the structural properties afforded by our specific approximation architecture to derive a significantly simpler alternative (the rALP) for which the number of constraints grows linearly with maximal capacity on each network leg. The rALP generates a feasible solution to the ALP. We show that this solution is optimal for affine approximations. While we aren't able to prove that this solution is optimal for concave approximations, the rALP generates optimal ALP solutions in all of our computational experiments with that architecture as well. The rALP thus significantly enhances the scalability of our approach.

The literature on both general dynamic capacity allocation heuristics, as well as bid-price controls is vast and predominantly computational; Talluri and van Ryzin (2004) provides an excellent review. Closest to this work is the paper by Adelman (2005), which also proposes an approximate DP approach to computing bid prices via an *affine* approximation to the value function. The rALP we propose allows an exponential reduction in the number of constraints for the ALP with affine approximation. However, in spite of affine approximation being a computationally attractive approximation architecture, our computational experiments suggest that affine approximations are not competitive with an approach that uses bid-prices computed via re-solution of the DLP at each time step.

Our approach might be viewed as a means of generating bid-prices. There have been a number of algorithms and heuristics proposed for this purpose. One class of schemes is based on mathematical programming formulations of essentially static versions of the problem that make the simplifying assumption that demand is deterministic and equal to its mean. The DLP approach is representative of this class and apparently the method of choice in practical applications (Talluri and van Ryzin (2004)). We compare the performance of our approach to such a scheme. Highly realistic simulations in P.P.Belobaba (2001) suggest that this class of approaches generates incremental revenues of approximately 1-2% over earlier leg-based RM techniques. There are alternatives to the use of bid price controls, the most prominent among them being 'virtual nesting' schemes such as the displacement adjusted (DAVN) scheme (see Talluri and van Ryzin (2004)). We do not consider our performance relative to such schemes; a subjective view (E. A. Boyd (2005)) is that these schemes are consistently outperformed by bid-price based schemes in practice.

An important thrust of our work is the incorporation of Markov-modulated customer arrival processes. There is an emerging literature on optimization techniques for models that incorporate demand processes where arrival rates are correlated in time. A recent example is the paper by Miguel and Mishra (2006), that evaluates various multi-stage stochastic programming techniques for a linear (with additive noise) model of demand evolution. These approaches rely on building 'scenario-trees' based on simulations of demand trajectories. While they can be applied to Markov-modulated arrival processes, scenario trees and their associated computational requirements typically grow exponentially in the horizon.

The remainder of this paper is organized as follows: In section 2, we formally specify a model for the dynamic capacity allocation problem. In section 3 we review the benchmark DLP heuristic. Section 4 presents an ADP approach to the dynamic capacity allocation problem and specifies our approximation architecture. That section also discusses some simple structural properties possessed by our approximation to the value function. Section 5 presents a series of computational examples comparing the performance of our algorithm with the DLP approach as also an approach based on an affine approximation to the value function. Section 6 studies a simple scalable alternative to the ALP, the rALP, and discusses computational experience with that program. Section 7 concludes.

2 Model

We consider an airline operating L flight legs. The airline may offer up to F fare products for sale at each point in time. Each fare product f is associated with a price p_f and requires seats on one or more legs. A matrix $A \in \mathbb{Z}_+^{L\times F}$ encodes the capacity on each leg consumed by each fare product: $A_{l,f} = k$ if and only if fare product f requires k seats on leg l. For concreteness we will restrict attention to the situation wherein a given fare product can consume at most 1 seat on any given leg although our discussion and algorithms carry over without any change to the more general case. Initial capacity on each leg is given by a vector $x_0 \in \mathbb{Z}_+^L$. Time is discrete. We assume an N period horizon with at most one customer arrival in a single period. A customer for fare product f arrives in the nth period with probability $\lambda_f(m_n)$. Here $m_n \in \mathcal{M}$ (a finite set) and represents the current demand 'mode'. m_n evolves according to a discrete time Markov process on $\mathcal{M}$ with transition kernel P_n. We note that the discrete time arrival process model we have described may be viewed as a uniformization of an appropriately defined continuous time arrival process. At the start of the nth period the airline must decide which subset of fare products from the set $\{f : A_f \preceq x_n\}$ it will offer for sale; an arriving customer for fare product f is assigned that fare product should it be available, the airline receives p_f, and $x_{n+1} = x_n - A_f$.

We define the *state-space* $\mathcal{S} = \{x : x \in \mathbb{Z}_+^L, x \preceq x_0\} \times \{0, 1, 2, \ldots, N\} \times \mathcal{M}$. Encoding the fare products offered for sale at time n by a vector in $\{0, 1\}^F \equiv \mathcal{A}$, a control policy is a mapping $\pi : \mathcal{S} \to \mathcal{A}$ satisfying $A\pi(s) \le x(s)$ for all $s \in \mathcal{S}$. Let Π represent the set of all such policies. Let $R(s, a)$ be a random variable representing revenue generated by the airline in state $s \in \mathcal{S}$ when fare products $a \in \mathcal{A}$ are offered for sale, and define for $s \in \mathcal{S}$,

$$J^\pi(s) = E_\pi\left[\sum_{n'=n(s)}^{N-1} R(s_{n'}, \pi(s_{n'}))\big|s_{n(s)} = s\right].$$

We let $J^*(s) = \max_{\pi\in\Pi} J^\pi(s)$, denote the expected revenue under the optimal policy π^* upon starting in state s.

J^* and π^* can, in principle, be computed via Dynamic Programming. In particular, define the dynamic programming operator T for $s \in \{s' : n(s') < N - 1\}$ according to

$$\begin{aligned}(TJ)(s) = &\sum_{f:A_f \le x(s)} \lambda_f(m(s)) \max\Big[p_f + E\left[J\left(S'_f\right)\right], E\left[J\left(S'\right)\right]\Big] \\ &+ \Big(1 - \sum_{f:A_f \le x(s)} \lambda_f(m(s))\Big) E\left[J(S')\right].\end{aligned} \tag{1}$$

where $S'_f = (x(s) - A_f, n(s) + 1, m_{n(s)+1})$ and $S' = (x(s), n(s) + 1, m_{n(s)+1})$. For $s \in \{s' : n(s') = N - 1\}$ we define $(TJ)(s) = \sum_{f:A_f \le x(s)} \lambda_f(m(s)) p_f$. We define $(TJ)(s) = 0$ for all $s \in \{s' : n(s') = N\}$. J^* may then be identified as the unique solution to the fixed point equation $TJ = J$. π^* is then the policy that achieves the maximum in (1); in particular, $\pi^*(s)_f = 0$ iff $p_f + E\big[J(S'_f)\big] < E\left[J\left(S'\right)\right]$ and $n(s) < N - 1$.

We will focus on three special cases of the above model, with N assumed even for notational convenience:

- (M1) Time homogeneous arrivals: Here we have $|\mathcal{M}| = 1$. That is the arrival rate of customers for the various fare products is constant over time and the arrival process is un-correlated in time.

- (M2) Multiple demand modes, deterministic transition time: Here we consider a model with $\mathcal{M} = \{\text{med,hi,lo}\}$. We have $m_n = \text{med}$ for $n \le N/2$. With probability $\tilde{p}$, $m_n = \text{lo}$ for all $n > N/2$ and with probability $1 - \tilde{p}$ and $m_n = \text{hi}$ for all $n > N/2$. This is representative of a situation where there is likely to be a change in arrival rates at some known point during the sales season. The revenue manager has a probabilistic model of what the new arrival rates are likely to be.

- (M3) Multiple demand modes, random transition time: Here we consider a model with $\mathcal{M} = \{\text{med, hi,lo}\}$, with the transition kernel P_t defined according to

$$P_n(m_{n+1} = y | m_n = x) = \begin{bmatrix} 1-q & q\tilde{p} & q(1-\tilde{p}) \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}_{xy}.$$

  where $q, \tilde{p} \in (0, 1)$. This arrival model is similar to the second with the exception that instead of a change in demand modes occurring at precisely $n = N/2$, there is now uncertainty in when this transition will occur. In particular, the transition time is now a geometric random variable with expectation $1/q$.

The above models were chosen since they are simple and yet serve to illustrate the relative merits of our approach for Markov-modulated demand processes.

3 Benchmark Heuristic: The Deterministic LP (DLP)

The Dynamic Programming problem we have formulated is computationally intractable and so one must resort to various sub-optimal control strategies. We review the DLP-heuristic for generating bid prices. This heuristic makes the simplifying assumption that demand is deterministic and equal to its expectation. In doing so, the resulting control problem reduces to the solution of a simple LP (the DLP) and the optimal control policy is static. In particular, if demand for fare product f over a $N-n$ period sales season, $D_{n,f}$, were deterministic and equal to expected demand, $\mathbb{E}[D_{n,f}|m_n]$, the maximal revenue that one may generate with an initial capacity $x(s)$ is given by the optimal solution to the DLP:

$$\begin{array}{lll} DLP(s): & \max & p'z \\ & \text{s.t.} & Az \leq x(s) \\ & & 0 \leq z \leq \mathbb{E}[D_{n(s)}|m_{n(s)} = m(s)] \end{array}$$

Denote by $r^*(s)$ a vector of optimal shadow prices corresponding to the constraint $Az \leq x(s)$ in $DLP(s)$. The bid price control policy based on the DLP solution is then given by:

$$\pi^{\text{DLP}}(s)_f = \begin{cases} 1 & \text{if } A_f' r^*(s) \leq p_f \text{ and } A_f \leq x(s) \\ 0 & \text{otherwise} \end{cases}$$

The above description of the DLP heuristic assumes that the shadow prices r^* are recomputed at each time step. While this many not always be the case, a general computational observation according to Talluri and van Ryzin (2004) is that frequent re-computation of r^* improves performance. This is consistent with our computational experience.

In the case of model M2, one might correctly point out that a simple modification of the DLP is likely to have superior performance. In particular, one may consider retaining the probabilistic structure of the demand mode transition model and solving a multi-stage stochastic program with recourse variables for capacity allocation in the event of a transition to the hi and lo demand modes respectively. We do not consider such a stochastic programming approach as it is intractable except for very simple models (such as M2); for a general Markov-modulated demand model with at least two demand modes, the number of recourse variables grows exponentially with horizon length.

4 Bid Price Heuristics via Approximate DP

Given a component-wise positive vector c, the optimal value function J^* may be identified as the optimal solution to the following LP:

$$\begin{array}{ll} \min & c'J \\ \text{s.t.} & (TJ)(s) \leq J(s) \quad \forall s \in \mathcal{S} \end{array}$$

The linear programming approach to approximate DP entails adding to the above LP, the further constraint that the value function J lie in the linear span of some set of basis functions $\phi_i : \mathcal{S} \to \mathbb{R},\ i = 1, 2, \ldots, k$. Encoding these functions as a matrix $\Phi \in \mathbb{R}^{|\mathcal{S}| \times k}$, the approximate LP (ALP) computes a vector of weights $r \in \mathbb{R}^k$ that optimally solve:

$$\begin{array}{ll} \min & c'\Phi r \\ \text{s.t.} & (T\Phi r)(s) \leq (\Phi r)(s) \quad \forall s \in \mathcal{S} \end{array}$$

Given a solution r^* to the ALP (assuming it is feasible), one then uses a policy that is greedy with respect to Φr^*. Of course, the success of this approach depends crucially upon the choice of the set of basis functions Φ. In the next two subsections we examine affine and concave approximation architectures. The affine approximation architecture for the network RM problem was proposed by Adelman (2005) in the context of the M1 model. The concave architecture is the focus of this paper. In the sequel we assume that $c_{s_0} = 1$ and that all other components of c are 0.

4.1 Separable Affine Approximation

Adelman (2005) considers the use of affine basis functions in the M1 model. In particular, Adelman (2005) explores the use of the following set of $(L+1)N$ basis functions defined according to

$$\phi_{l,n}(x, n') = \begin{cases} x_l & \text{if } l \leq L \text{ and } n = n' \\ 1 & \text{if } l = L+1 \text{ and } n = n' \\ 0 & \text{otherwise} \end{cases}$$

The ALP here consequently has $\Theta(LN)$ variables but $\Theta(\bar{x}^L NF)$ constraints. Adelman (2005) proposes the use of a column generation procedure to solve the ALP. We show in Section 6 that the ALP can be reduced to an LP with $\Theta(LN)$ variables and $\Theta(\bar{x}2^L LNF)$ constraints making practical solution of the ALP to optimality possible for relatively large networks (including, for instance, the largest examples in Adelman (2005)). In spite of being a computationally attractive approximation architecture, affine approximations have an obvious weakness: the greedy policy with respect to an affine approximation to the value function is insensitive to intermediate capacity levels so that the set of fare products offered for sale at any intermediate point in time depends only upon the time left until the sales season ends. In particular the greedy policy with respect to an affine approximation, π^{aff}, will satisfy $\pi^{\text{aff}}(x, n) = \pi^{\text{aff}}(\tilde{x}, n)$ provided x and $\tilde{x}$ are positive in identical components. We observe in computational experiments that a policy that is greedy with respect to an affine approximation to the value function is in fact not competitive with a policy based on re-computation of bid-prices at each time step via the DLP. While one possible approach to consider is frequent re-solution of the ALP with affine approximation, this is not a feasible option given that bid-prices must often be generated in real time. It is simple to show (using, for example, the monotonicity of the T operator) for any vector $e \in \{0,1\}^L$ that is positive in a single component, that $J^*(x+e, n) - J^*(x, n)$ is non-increasing in x. Affine approximations are incapable of capturing this concavity of J^* in inventory level. This motivates us to consider a separable concave approximation architecture which is the focus of this paper.

4.2 Separable Concave Approximation

Consider the following set of basis functions, $\phi_{l,n,i,m}$, defined for integers $l \in [1, L]; n \in [0, N], i \in [0, (x_0)_l]$, and $m \in \mathcal{M}$ according to:

$$\phi_{l,n,i,m}(x', n', m') = \begin{cases} 1 & \text{if } x'_l \geq i \text{ , } n = n' \text{ and } m = m' \\ 0 & \text{otherwise} \end{cases}$$

The ALP in this case will have $\Theta(\bar{x}LN|\mathcal{M}|)$ variables and $\Theta(\bar{x}^L NF|\mathcal{M}|)$ constraints. Note that optimal solution is intractable since $|\mathcal{S}|$ is exponentially large. One remedy is the constraint sampling procedure in de Farias and Van Roy (2004) which suggests sampling constraints from $\mathcal{S}$ according to the state-distribution induced by an optimal policy. Assuming a sales season of N periods and an initial inventory of x_0, we propose using the following procedure with parameter K:

1. Simulate a bid price control policy starting at state $s_0 = (x_0, 0, m_0)$, using bid prices generated by re-solving the DLP at each time step. Let $\mathcal{X}$ be the set of states visited over the course of several simulations. We generate a set with $|\mathcal{X}| = K$

2. Solve the following Relaxed LP (RLP):

$$
\begin{array}{lll}
\min & (\Phi r)(s_0) & \\
\text{s. t.} & (T\Phi r)(s) \leq (\Phi r)(s) & \text{for } s \in \mathcal{X} \\
 & r_{l,n,i,m} \geq r_{l,n,i+1,m} & \forall i > 0, l, n, m
\end{array}
$$

3. Given a solution r^* to the RLP, use the following control policy over the actual sales season:

$$
\pi^{\text{con}}(s)_f = \begin{cases} 1 & \text{if } \sum_{l:A_{l,f}=1} r^*_{l,n(s),x(s)_l,m(s)} \leq p_f \text{ and } A_f \leq x(s) \\ 0 & \text{otherwise} \end{cases}
$$

Several comments on the above procedure are in order. Step 1 in the procedure entails choosing a suitable number of samples K; de Farias and Van Roy (2004) provides some guidance on this choice. Our choice of K was heuristic and is described in the next section. Step 2 of the procedure entails solving the RLP whose constraints are samples of the original ALP. We will shortly mention several simple structural properties that an optimal solution to the ALP must posses. Adding these constraints to the RLP strengthens the quality of our solution. Also, note that the inequality constraints on the weights enforce concavity of the approximation. Finally note that the greedy policy with respect to the our approximation to J^* takes the form of a bid price policy as in the case of affine approximation. However, unlike affine approximation the resulting policy decisions depend on available capacity as well as time.

4.3 ALP Solution Properties

The optimal solution to the DLP provides an upper bound to the true value function J^*, i.e. $DLP(s) \geq J^*(s)$. There are several proofs of this fact for the time homogeneous model M1. For example, see Gallego and van Ryzin (1997) or Adelman (2005). The DLP continues to be an upper bound to the true value function for the more general model we study here (via a simple concavity argument and the use of Jensen's inequality). We can show that the ALP with separable concave approximation provides a tighter upper bound than does the DLP for model M2, and generalizations to M2 which allow for more than a single branching time. The same result for time homogeneous arrival rates (i.e. for model M1) follows as a corollary. We are at present unable to establish such a result for the general model.

Lemma 1. *For model M2 with initial state s, $J^*(s) \leq ALP(s) \leq DLP(s)$*

The proof of the lemma can be found in the appendix. The above result is not entirely conclusive. In particular, while it is indeed desirable to have a good approximation to the true value function, a tighter approximation does not guarantee an improved policy. Nonetheless, stronger approximations to the true value function imply stronger bounds on policy performance. Finally, solutions to the ALP must satisfy simple structural properties. For example, in the case of model M1 it is clear that we must have $\sum_{i=0}^{x} r_{l,n,i} \geq 0$ for all $x, l, n < N$ and further $\sum_{i=0}^{x} r_{l,n,i} \leq \sum_{i=0}^{x} r_{l,n-1,i}$ for all $l, x, 0 < n < N$. We explicitly enforce these constraints in our computational experiments.

5 Computational Results

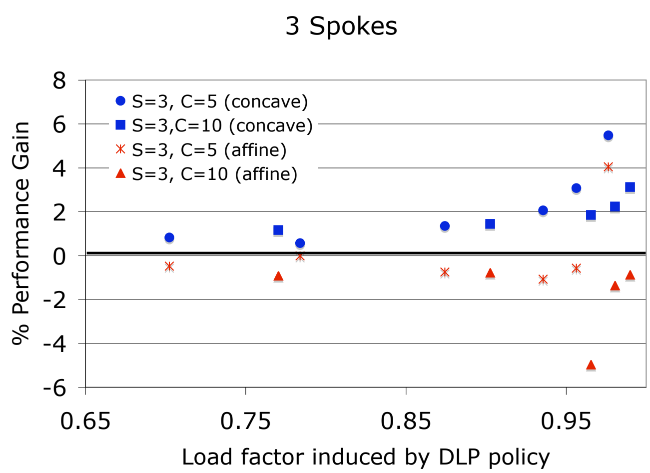

Figure 1: Performance relative to the DLP for model M1

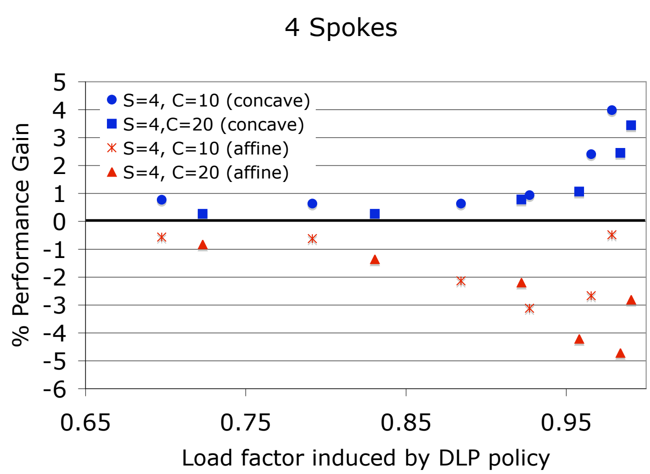

Figure 2: Performance relative to the DLP for model M1

It is difficult to establish theoretical performance guarantees for our algorithm. Indeed, we are unaware of any algorithm for the dynamic capacity allocation problem for which non-asymptotic theoretical per-

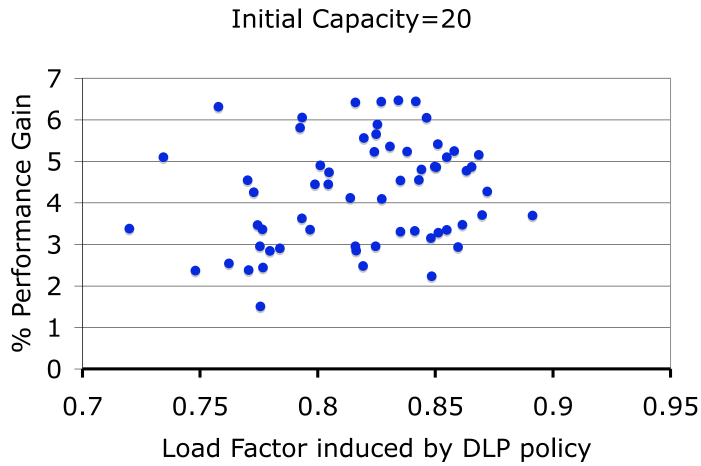

Figure 3: Performance relative to the DLP for model M2

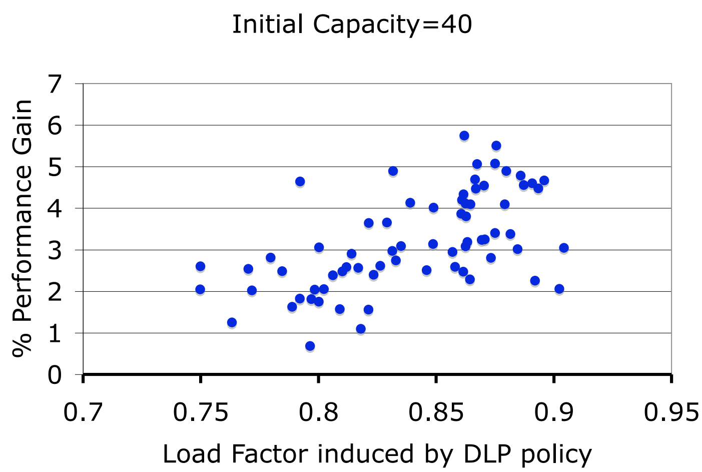

Figure 4: Performance relative to the DLP for model M2

formance guarantees are available. As such, we will establish performance merits for our algorithm via a computational study. We will consider two simple test networks each with a single 'hub' and either three or four spoke cities. This topology is representative of actual airline network topologies. Each leg in our network represents two separate aircraft (one in each direction) making for a total of $f = 15$ itineraries on the 3 spoke network and $f = 24$ itineraries on the 4 spoke network. Arrival rates for each itinerary, demand mode i.e. (f, m) pair were picked randomly from the unit f-dimensional simplex and suitably normalized. Route prices were generated uniformly in the interval $[50, 150]$ for single leg routes and $[50, 250]$ for two leg routes. We consider a random instantiation of arrival rates and probabilities for each network topology and for each instantiation measure policy performance upon varying initial capacity levels and sales horizon. We compare performance against the DLP with re-solution at each time step. In the case of model M1, we also include policies generated via the separable affine approximation architecture in our experiments. We solve RLPs with $50,000$ sampled states, this number being determined by memory constraints. We now describe

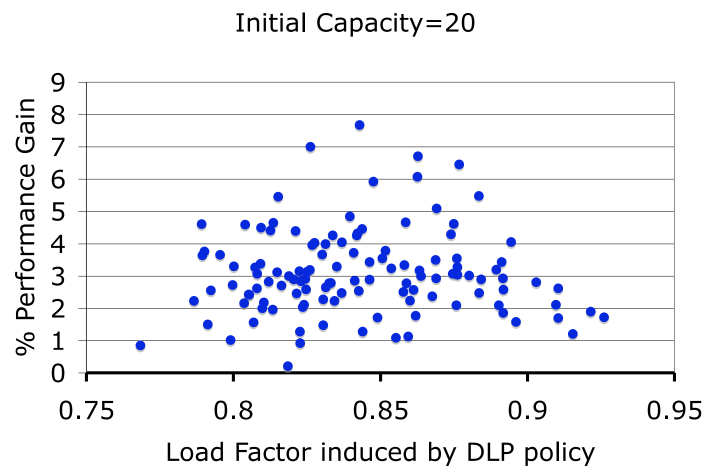

Figure 5: Performance relative to the DLP for model M3

in detail our experiments and results for each of the three models.

5.1 Time homogeneous arrivals (M1)

We consider three and four spoke models. The arrival probabilities for each fare product were drawn uniformly at random on the unit simplex and normalized so that the probability of no customer arrival in each period was 0.7. For both models, we consider fixed capacities (of 5 and 10 for 3-spoke networks, 10 and 20 for four spoke networks) on each network leg and vary the sales horizon N. For each value of N we record the average load-factor (i.e. the average fraction of seats sold) under the DLP policy; we select values of N so that this induced load factor is > 0.7. We plot in Figures 1 and 2 the performance of the the ADP based approaches with affine and separable approximations relative to the DLP heuristic for two different initial capacity levels. The x-intercept for a data point in both plots is the average load factor induced by the DLP heuristic for the problem data in question at that point.

The plots suggest a few broad trends. The affine approximation architecture is almost uniformly dominated by the DLP heuristic when the DLP is re-solved at every time step, while the separable architecture uniformly dominates both heuristics in every problem instance. We note that since a bid price computation in the ADP approach is simply a lookup it is far quicker than solving the DLP, so that together these facts support the plausibility of using an ADP approach with separable approximation. Another trend is performance gain. This is actually quite low at low induced load factors ($< 0.5\%$) but can be as high as 5% at high load factors. At moderate load factors (that are at least nominally representative) the performance gain is on the order of 1%. We anticipate the gain to be larger for more complex networks.

It is difficult to expect higher performance gains than we have observed for the M1 demand model. In particular, at low load factors, the problem is trivialized (since it is optimal to accept all requests). Moreover, it is well known (see Gallego and van Ryzin (1997)) that in a certain fluid scaling (which involves scaling both initial capacity x_0 and sales horizon N by some scaling factor $\tilde{N}$), the DLP heuristic is optimal as $\tilde{N}$ gets large. The purpose of our experiments with this model is to illustrate the fact that the separable concave approximations we employ are robust in this simple demand setting.

5.2 Multiple demand modes (M2, M3)

Model M1 is potentially a poor representation of reality. This leads us to consider incorporating a demand forecasting model such as that in models M2 and M3. In our experiments, the arrival probabilities for each demand mode were drawn uniformly at random on the unit 24-dimensional simplex and normalized so that the probability of no customer arrival in each period was 0.55 for the 'med' demand mode, 0.7 for the 'lo' mode, and 0.1 in the 'hi' mode. The probability of transitioning from the med to lo demand mode, p, was set to 0.5 in both models, and we set m_0 = med.The probability of transitioning out of the med demand state, q, was set to $2/N$ in model M3. The sales horizon N was varied so that the load-factor induced by the DLP policy was approximately between 0.8 and 0.9. We generate a random ensemble of 40 such problems for a network with 4 spokes and consider initial capacity levels of 20 seats and 40 seats. We measure the performance gain of our ADP with separable concave approximation derived bid price control over the DLP. The DLP is resolved at every time step so that it may recompute expected total remaining demand for each fare product conditioned on the current demand mode.

For model M2, we plot in Figures 3 and 4 the performance of the the ADP based approach with separable concave approximation relative to the DLP heuristic with initial capacity levels of 20 and 40 respectively. We note that the relative performance gain here is significant (up to about 8%) in a realistic operating regime. In the case of model M3, Figure 5 illustrates similar performance trends.

We see that the approximate DP approach with concave approximation offers substantiative gains over the use of the DLP even with very simple stochastic variation in arrival rates. We anticipate that these gains will be further amplified for more complex models of arrival rate variability (for example in models with a larger number of demand modes etc.).

6 Towards scalability: A simpler ALP

Assuming maximal capacities of $\bar{x}$ on each of L legs, a time horizon N, and F fare products, the ALP with separable concave approximations has $\Theta(\bar{x}^L NF)$ constraints. In this section we will demonstrate a program - the relaxed ALP (rALP) - with $O(\bar{x}NLF2^L)$ constraints that generates a feasible solution to the ALP. The rALP has the same decision variables as the ALP, and a small number of additional auxiliary variables. The rALP is consequently a significantly simpler program than the ALP. In the case of affine approximation, the rALP generates the optimal solution to the ALP. In the case of separable concave approximations, the rALP generates a feasible solution to the ALP whose quality we demonstrate through computational experiments to be excellent. The rALP solution in fact coincides with the ALP solution in all of our experiments. Our presentation will assume that an itinerary can consist of at most 2 flight legs and will be in the context of model M1 for simplicity; extending the program to more general arrival process model is straightforward.

6.1 The rALP

In what follows, we understand that for a state $s \in \mathcal{S}$, $s \equiv (x(s), n(s))$. Let us partition $\mathcal{S}$ into sets of the form $\mathcal{S}_y = \{s : s \in \mathcal{S}, x(s)_i = 0 \iff y_i = 0\}$ for all $y \in \{0,1\}^L$. Clearly, $\mathcal{S}$ can be expressed as the disjoint union of all such sets $\mathcal{S}_y$. Also define the subset of fare products $\mathcal{F}_y$ according to $\mathcal{F}_y = \{f : A_f \leq y\}$ and assume $\mathcal{F}_y \neq \emptyset$ for all y. For some $y \in \{0,1\}^L$, consider the set of constraints

$$(T\Phi r)(s) \leq (\Phi r)(s) \text{ for all } s \in \mathcal{S}_y. \tag{2}$$

We will approximate the feasible region specified by this set of constraints by the following set of constraints in r and the auxiliary variables m:

$$\begin{array}{ll} LP_{y,n}(r,m) \leq 0 & \forall n < N \\ m^f_{l,n,i} \geq r_{l,n,i} & \forall\, l, n \leq N, i \in \{1,\ldots,\bar{x}\}, f \in \mathcal{F}_y \\ \sum_{l:A_{l,f}=1} m^f_{l,n,\bar{x}} \geq p_f & \forall n \leq N, f \in \mathcal{F}_y \\ m^f_{l,n,i+1} \leq m^f_{l,n,i} & \forall l, n \leq N, i \in \{1,\ldots,\bar{x}-1\}, f \in \mathcal{F}_y \end{array} \tag{3}$$

where $LP_{y,n}(r,m)$ refers to a certain linear program with decision variables $x \in \mathbb{R}^{LN(\bar{x}+1)}$. We will now proceed to describe this linear program as also discuss how the constraint $LP_{y,n}(r,m) \leq 0$ may itself be described by a set of linear constraints in r, m and certain additional auxiliary variables.

Let us define:

$$\begin{aligned} c_{y,n}(r,m)'x = &\sum_l \sum_{i=0}^{\bar{x}} (r_{l,n+1,i} - r_{l,n,i}) x_{l,n,i} \quad + \\ &\sum_{f\in\mathcal{F}_y} \lambda_f \sum_{l:A_{l,f}=1} \left(\sum_{i=1}^{\bar{x}-1} (m^f_{l,n+1,i} - r_{l,n+1,i})(x_{l,n,i} - x_{l,n,i+1}) + (m^f_{l,n+1,\bar{x}} - r_{l,n+1,\bar{x}}) x_{l,n,\bar{x}} \right) \end{aligned}$$

Implicit in this definition, the vector $c_{y,n}(r,m)$ has components that are themselves linear functions of r and m. Delaying a precise description for a moment, our goal is to employ the approximation

$$\sum_{l:A_{l,f}=1} m^f_{l,n,\tilde{x}_l} \sim \max((\Phi r)(\tilde{x}, n) - (\Phi r)(\tilde{x} - A_f, t), p_f),$$

for all $f \in \mathcal{F}_y$, so that $c_{y,n}(r,m)'x$ will serve as our approximation to $(T\Phi r)(s) - (\Phi r)(s)$ when $s \in \mathcal{S}_y, n(s) = n$ and $x(s)_l = \sum_i x_{l,n,i}$.

We next define the linear program $LP_{y,n}(r,m)$.

$$
\begin{array}{llll}
LP_{y,n}(r,m): & \max & c_{y,n}(r,m)'x & \\
 & \text{s.t.} & x_{l,n,0} = 1 \quad \forall l & \\
 & & x_{l,n,1} = 1 \quad \forall l \text{ s.t. } y_l = 1 & \\
 & & x_{l,n,1} = 0 \quad \forall l \text{ s.t. } y_l = 0 & \\
 & & x_{l,n,i+1} \le x_{l,n,i} \quad \forall l,n,i \ge 1 & \\
 & & 0 \le x_{l,n,i} \quad \forall l,n,i \ge 1 &
\end{array}
$$

The constraint set for $LP_{y,n}(r,m)$ may be written in the form $\{x : Cx \le b, x \ge 0\}$ where C and b have entries in $\{0,1,-1\}$. The dual to $LP_{y,n}(r,m)$ is then given by:

$$
\begin{array}{ll}
\min & b'z_{y,n} \\
\text{s.t.} & C'z_{y,n} \ge c_{y,n}(r,m) \\
 & z_{y,n} \ge 0
\end{array}
$$

so that by strong duality, our approximation to the set of constraints (2), i.e. (3), may equivalently be written as the following set of linear constraints in the variables r, m and z_y:

$$
\begin{array}{ll}
b'z_{y,n} \le 0 & \forall n < N \\
C'z_{y,n} \ge c_{y,n}(r,m) & \forall n < N \\
z_{y,n} \ge 0 & \forall n < N \\
m^f_{l,n,i} \ge r_{l,n,i} & \forall\, l, n \le N, i \in \{1,\ldots,\bar{x}\}, f \in \mathcal{F}_y \\
\sum_{l:A_{l,f}=1} m^f_{l,n,\bar{x}} \ge p_f & \forall n \le N, f \in \mathcal{F}_y \\
m^f_{l,n,i+1} \le m^f_{l,n,i} & \forall l, n \le N, i \in \{1,\ldots,\bar{x}-1\}, f \in \mathcal{F}_y
\end{array}
\tag{4}
$$

Assuming a starting state $s_0 = (\bar{x}, 0)$, we thus propose to minimize $(\Phi r)(s_0)$ subject to the set of constraints (4) for all $y \in \{0,1\}^L$ and

$$
\begin{array}{ll}
r_{l,n,i,m} \ge r_{l,n,i+1,m} & \forall l,n,i,m \\
r_{l,n,i,m} = 0 & \forall i,l,m; n = N
\end{array}
$$

in order to compute our approximation to the value function. We will refer to this program as $rALP(s_0)$.

6.2 Quality of Approximation

We have proposed approximating the feasible region specified by the set of constraints

$$(T\Phi r)(s) \le (\Phi r)(s) \text{ for all } s \in \mathcal{S}$$

which has size that is $\Theta(\bar{x}^L NF)$ by a set of linear constraints of size $O(\bar{x}NLF2^L)$. There are two potential sources of error for this approximation: For one, we would ideally like to enforce the constraint $c_{y,n}(r,m)'x \le 0$ only for $x_{\cdot,n,\cdot}$ in $\{0,1\}^{(\bar{x}+1)L}$, whereas in fact we allow $x_{\cdot,n,\cdot}$ to take values in $[0,1]^{(\bar{x}+1)L}$.

It turns out that this relaxation introduces no error to the approximation, simply because the vertices of $LP_{y,n}(r,m)$ are integral. That is, the optimal solutions always satisfy $x^*_{l,n,i} \in \{0,1\}$. This is simple to verify; $LP_{y,n}(r,m)$ may be rewritten as a min-cost flow problem on a certain graph with integral supplies at the sources and sinks.

The second source of approximation arises from the fact that we approximate $\max((\Phi r)(x,n)-(\Phi r)(x-A_f,n),p_f)$ by $\sum_{l:A_{l,f}=1} m^f_{l,n,x_l}$. In particular, we have:

$$\sum_{l:A_{l,f}=1} m^f_{l,n(s),x(s)_l} \geq \max((\Phi r)(s) - (\Phi r)(x(s) - A_f, n(s)), p_f) \tag{5}$$

This yields the following Lemma. A proof may be found in the appendix.

Lemma 2. $rALP(s_0) \geq ALP(s_0)$. *Moreover if* (r^{rALP}, m) *is a feasible solution to the rALP then* r^{rALP} *is a feasible solution to the ALP.*

In the case of affine approximations the reverse is true as well. That is, we have:

Lemma 3. *For affine approximations,* $rALP(s_0) \leq ALP(s_0)$. *Moreover if* r^{ALP} *is a feasible solution to the ALP then there exists a feasible solution to the rALP,* (r^{rALP}, m) *satisfying* $r^{rALP} = r^{ALP}$.

Consequently, the rALP yields the optimal solution to the ALP for affine approximations. In the case of separable concave approximations, the rALP will in general yield suboptimal solutions to the ALP. One may however show that there exists an optimal solution to the rALP satisfying for all $s \in \mathcal{S}$:

$$\sum_{l:A_{l,f}=1} m^{*,f}_{l,n(s),x(s)_l} \geq \max((\Phi r^*)(s) - (\Phi r^*)(x(s) - A_f, n(s)), p_f) \geq \frac{1}{2} \sum_{l:A_{l,f}=1} m^{*,f}_{l,n(s),x(s)_l}, \tag{6}$$

so that heuristically we might expect the rALP to provide solutions to the ALP that are of reasonable quality. In fact, as our computational experiments in the next subsection illustrate, the rALP appears to yield the optimal ALP solution in the case of separable concave approximations as well.

6.3 Computational experience with the rALP

We consider problems with 3, 4, and 8 flights with problem data generated as in the computational experiments in Section 5. Table 1 illustrates the solution objective and solution time for the rALP and ALP for each of these problems. For the 3 and 4 dimensional problems, we consider instances small enough so that it is possible to solve the ALP exactly. We see in these instances that the rALP delivers the same solution as the ALP in a far shorter time. The ALP for the 8 dimensional instance cannot be stored - let alone solved - on most conventional computers; the rALP for that problem on the other hand is relatively easy to solve and yields a near optimal solution (the comparison here being with the optimal solution of an RLP with 100,000 sampled constraints; recall that the RLP solution is a lower bound on the ALP).

In practice we envision the rALP being used in conjunction with constraint sampling. In particular, consider the following alternative to the RLP of Section 4: Let $\mathcal{X}$ be the set of sampled states one might use for the RLP. We then include in the rALP the set of constraints (3) for only those (y,n) such that there

Dimension	$\bar{x}$	T	$ALP(s_0)$	$rALP(s_0)$	t_{ALP}	t_{rALP}
3	10	50	925.46	925.46	99.33	2.674
	20	50	1035.23	1035.23	2611.89	11.46
4	10	10	218.28	218.28	161.67	0.41
	10	30	653.37	653.37	2771.0	2.45
8	10	100	5019.90*	5028.15	849.2*	1251.92
	10	100	5019.90*	5028.12**	849.2*	177.77 **

Table 1: Solution quality and computation time for the rALP and ALP. * indicates values for an ***RLP*** with 100,000 constraints (recall that the RLP provides a lower bound on the ALP). ** indicates values for the sampled rALP described in section 6.3 using the same sample set as that in the computation of the corresponding RLP. Computation time reported in seconds for the CPLEX barrier LP optimizer running on a workstation with a 64 bit AMD processor and 8GB of RAM.

exists a sampled state $(x, n) \in \mathcal{X}$ with $x(s) \in \mathcal{S}_y$. The sampled rALP will have $O(\bar{x}LFK(\mathcal{X}))$ constraints where $K(\mathcal{X}) = |\{(y, n) : \exists s \in \mathcal{X} \text{ s.t. } s \in \mathcal{S}_y, n(s) = n\}|$. Since a majority of sampled states are likely to be in $\mathcal{S}_e$ where e is the vector of all ones (indicating that all fare products can potentially be serviced), one may expect $K(\mathcal{X})$ to be *far* smaller than $|\mathcal{X}|$, making the sampled rALP a significantly simpler program than the RLP. Moreover, since the sampled rALP attempts to enforce $(T\Phi r)(s) \leq (\Phi r)(s)$ for a collection states that are a superset of the states in $\mathcal{X}$, we might expect it to provide a stronger approximation as well. While a thorough exploration of the sampled rALP is beyond the scope of this paper, the last row of Table 1 provides encouraging supporting evidence.

7 Conclusion

We have explored the use of separable concave functions for the approximation of the optimal value function for the dynamic capacity allocation problem. The approximation architecture is quite flexible and we have illustrated how it might be employed in the context of a general arrival process model wherein arrival rates vary stochastically according to a Markov process. Our computational experiments indicate that the use of the LP approach to Approximate DP along with this approximation architecture can yield significant performance gains over the DLP (of up to about 8%) , even when re-computation of DLP bid prices is allowed at every time step. Moreover, our control policy is a bid price policy where policy execution requires a table look-up at each epoch making the methodology ideally suited to real time implementation. State of the art heuristics for the dynamic capacity allocation problem typically resort to using point estimates of demand in conjunction with a model that assumes simple time homogeneous arrival processes in order to make capacity allocation decisions dynamically. As such, our algorithm may be viewed as a viable approach to moving beyond the use of point estimates and instead integrating forecasting and optimization. The approach we propose is also scalable. For example, the sampled rALP proposed in section 6 may be solved in a few minutes for quite large problems.

Several issues remain to be resolved. For example, in the interest of very large-scale implementations, it would be useful to explore the use of simpler basis functions that are nonetheless capable of capturing the concavity of the true value function. In computational experiments, the rALP produced optimal solutions to the ALP; it would be interesting to establish that the programs are equivalent (as we have in the case of affine approximations). The ALP produces a tighter approximation to the true value function than does the DLP but it remains to show that the ALP policy dominates the DLP policy as well if this is at all true. Finally, a computational exploration of our approach with a highly realistic simulator such as that used by P.P.Belobaba (2001) would give a better sense for the gains that one may hope to achieve via the use of this approach in practice.

Acknowledgment

The authors would like to thank Mike Harrison for useful discussions that led to our consideration of some of the models here, as well as his suggestions and pointers to the literature.

References

Adelman, D. 2005. Dynamic bid price in revenue management. To appear in Operations Research.

de Farias, D.P., B. Van Roy. 2003. The linear programming approach to approximate dynamic programming. *Operations Research* **51:6** 850–865.

de Farias, D.P., B. Van Roy. 2004. On constraint sampling in the linear programming approach to approximate dynamic programming. *Mathematics of Operations Research* **29:3** 462–478.

E. A. Boyd, I. C. Bilegan. 2005. Revenue management and e-commerce. Pros-RM white paper. URL: `http://www.prospricing.com`.

Gallego, G., G. van Ryzin. 1997. A multiproduct dynamic pricing problem and its application to network yield management. *Operations Research* **45**(1) 24–41.

Miguel, Victor De, Nishant Mishra. 2006. What multistage stochastic programming can do for network revenue management. London Business School working paper.

P.P.Belobaba. 2001. Revenue and competitive impacts of o-d control: Summary of pods results. *First Annual INFORMS Revenue Management Section Meeting*. New York, NY.

P.P.Belobaba, S. Lee. 2000. Pods update: Large network o-d control results. *Proc. 2000 AGIFORS Reservations and Yield Management Study Group Symposium*. New York, NY.

Talluri, Kalyan T., Garrett J. van Ryzin. 2004. *The Theory and Practice of Revenue Management*. Springer Science+Business Media.

A Proofs for Section 3

Lemma 1. *For model M2 with initial state* s, $J^*(s) \leq ALP(s) \leq DLP(s)$

Proof: We assume for notational convenience that N is even. Consider the following linear program:

$$\begin{array}{lll} sDLP(s): & \max & p'z_0 + \Pr(s_{N/2} = \text{lo})p'z_1 + \Pr(s_{N/2} = \text{hi})p'z_2 \\ & \text{s.t.} & A(z_0 + z_1) \leq x(s) \\ & & A(z_0 + z_2) \leq x(s) \\ & & 0 \leq z_0 \leq \mathbb{E}[D_0] - E[D_{N/2}] \\ & & 0 \leq z_1 \leq \mathbb{E}[D_{N/2}|s_{N/2} = \text{lo}] \\ & & 0 \leq z_2 \leq \mathbb{E}[D_{N/2}|s_{N/2} = \text{hi}] \end{array}$$

It is clear that $sDLP(s) \leq DLP(s)$. This is because $z_0 + \Pr(s_{N/2} = \text{lo})z_1 + \Pr(s_{N/2} = \text{hi})z_2$ is a feasible solution to $DLP(s)$ of the same value as $sDLP(s)$. We will first show that $ALP(s) \leq sDLP(s)$. The dual to $sDLP(s)$ is given by:

$$\begin{array}{ll} \min & x(s)'y_{1,1} + x(s)'y_{1,2} + \tilde{D}_0'y_{2,0} + \tilde{D}_1'y_{2,1} + \tilde{D}_2'y_{2,2} \\ \text{s.t.} & A'(y_{1,1} + y_{1,2}) + y_{2,0} \geq p \\ & A'y_{1,1} + y_{2,1} \geq p\Pr(s_{N/2} = \text{lo}) \\ & A'y_{1,2} + y_{2,2} \geq p\Pr(s_{N/2} = \text{hi}) \\ & y_{1,1}, y_{1,2}, y_{2,0}, y_{2,1}, y_{2,2} \geq 0 \end{array}$$

where $\tilde{D}_0 = \mathbb{E}[D_0] - E[D_{N/2}], \tilde{D}_1 = \mathbb{E}[D_{N/2}|s_{N/2} = \text{lo}]$ and $\tilde{D}_2 = \mathbb{E}[D_{N/2}|s_{N/2} = \text{hi}]$. Consider the following solution to the ALP for M2: Set

$$r^*_{l,n,i,med} \begin{cases} = (y^*_{1,1})_l + (y^*_{1,2})_l \\ \qquad \text{for } i > 0, n < N/2 \\ = \tilde{D}'_{0,n}y^*_{2,0} + \Pr(s_{N/2} = \text{lo})r^*_{1,N/2,0,lo} + \Pr(s_{N/2} = \text{hi})r^*_{1,N/2,0,hi} \\ \qquad \text{for } i = 0, l = 1, n < N/2 \\ = 0 \qquad \text{otherwise} \end{cases}$$

$$r^*_{l,n,i,lo} \begin{cases} = (y^*_{1,1})_l/\Pr(s_{N/2} = \text{lo}) & \text{for } i > 0, N > n \geq N/2 \\ = \tilde{D}'_{1,n}y^*_{2,1}/\Pr(s_{N/2} = \text{lo}) & \text{for } i = 0, l = 1, N > n \geq N/2 \\ = 0 & \text{otherwise} \end{cases}$$

$$r^*_{l,n,i,hi} \begin{cases} = (y^*_{1,2})_l/\Pr(s_{N/2} = \text{hi}) & \text{for } i > 0, N > n \geq N/2 \\ = \tilde{D}'_{2,n}y^*_{2,2}/\Pr(s_{N/2} = \text{hi}) & \text{for } i = 0, l = 1, N > n \geq N/2 \\ = 0 & \text{otherwise} \end{cases}$$

where $\tilde{D}_{0,n} = \mathbb{E}[D_n] - E[D_{N/2}]$ for $n < N/2$, $\tilde{D}_{1,n} = \mathbb{E}[D_n|s_{N/2} = \text{lo}]$ and $\tilde{D}_{2,n} = \mathbb{E}[D_n|s_{N/2} = \text{hi}]$ for $n \geq N/2$. It is routinely verified that this solution is in fact feasible for the ALP and has value equal to $sDLP(s)$. The fact that $ALP(s) \geq J^*(s)$ follows from the monotonicity of the T operator and the fact that J^* is the unique fixed point of T. This completes the proof.

B Proofs for Section 6

Lemma 2. $rALP(s_0) \geq ALP(s_0)$. *Moreover if* (r^{rALP}, m) *is a feasible solution to the rALP then* r^{rALP} *is a feasible solution to the ALP.*

Proof: Let r^{rALP} be optimal weights from a solution to the rALP, and consider an arbitrary state $s \in \mathcal{S}$. We have

$$\begin{aligned}
& (\Phi r^{\text{rALP}})(s) \\
\geq \quad & \sum_{f: A_f \leq x(s)} \lambda_f \left((\Phi r^{\text{rALP}})(x(s) - A_f, n(s)+1) + \sum_{l: A_{l,f}=1} m^f_{l, n(s)+1, x(s)_l} \right) \\
& + \left(1 - \sum_{f: A_f \leq x(s)} \lambda_f \right) (\Phi r^{\text{dALP}})(x(s), n(s)+1) \\
\geq \quad & \sum_{f: A_f \leq x(s)} \lambda_f \Big((\Phi r^{\text{rALP}})(x(s) - A_f, n(s)+1) \\
& \quad + \max\left((\Phi r^{\text{rALP}})(x(s), n(s)+1) - (\Phi r^{\text{rALP}})(x(s) - A_f, n(s)+1), p_f \right) \Big) \\
& + \left(1 - \sum_{f: A_f \leq x(s)} \lambda_f \right) (\Phi r^{\text{rALP}})(x(s), n(s)+1) \\
= \quad & (T\Phi r^{\text{rALP}})(s)
\end{aligned}$$

where the first inequality is by the feasibility of r^{rALP}, m^* for the rALP and the second inequality is enforced by the fourth through sixth constraints in (4). This yields the result. □

Lemma 3. *For affine approximations, $rALP(s_0) \leq ALP(s_0)$. Moreover if (r^{ALP}) is a feasible solution to the ALP then there exists a feasible solution to the rALP, (r^{rALP}, m) satisfying $r^{rALP} = r^{ALP}$.*

Proof: Let r^* be the optimal solution to the ALP. For each $i \geq 1, l, f, n \leq N$, define

$$m^{*,f}_{l,n,i} = r_{l,n,i} + \left(\max\left(\sum_{l: A_{l,f}=1} r^*_{l,n,1}, p_f \right) - \sum_{l: A_{l,f}=1} r^*_{l,n,1} \right) / L(f)$$

where $L(f) = |\{l : A_{l,f} = 1\}|$. Since we are considering affine approximations, $r^*_{l,n,i'} = r^*_{l,n,i''}$ for $i', i'' > 0$. Consequently, our definition implies that for every f, n,

$$\sum_{l: A_{l,f}=1} m^{*,f}_{l,n,i} = \max\left(\sum_{l: A_{l,f}=1} r^*_{l,n,i}, p_f \right)$$

so that the feasibility of r^* for the ALP implies that $LP_{y,n}(r^*, m^*) \leq 0$ for each $y \in \{0,1\}^L, n < N$. Moreover, m^* clearly satisfies the second through fourth constraints of (3). This completes the proof. □