A Dynamic Traveling Salesman Problem with Stochastic Arc Costs

Alejandro Toriello William B. Haskell Michael Poremba
Daniel J. Epstein Department of Industrial and Systems Engineering
University of Southern California
Los Angeles, California 90089
{toriello, wbhaskel, poremba} *at* usc *dot* edu

June 7, 2013

Abstract

We propose a dynamic *traveling salesman problem* (TSP) with stochastic arc costs motivated by applications, such as dynamic vehicle routing, in which a decision's cost is known only probabilistically beforehand but is revealed dynamically before the decision is executed. We formulate the problem as a *dynamic program* (DP) and compare it to static counterparts to demonstrate the dynamic paradigm's advantage over an a priori approach. We then apply *approximate linear programming* (ALP) to overcome the DP's curse of dimensionality and obtain a semi-infinite linear programming lower bound. The bound requires only expected arc costs and knowledge of the uncertainty support set, and is valid for any distribution with these parameters. Though NP-hard for arbitrary compact uncertainty sets, we show that it can be solved in polynomial time for two important classes, polytopes with polynomially many extreme points and hyper-rectangles. We also analyze the price-directed policy implied by our ALP and derive worst-case guarantees for its performance. Our computational experiments demonstrate the advantage of both the ALP bound and a related heuristic policy.

1 Introduction

The *traveling salesman problem* (TSP) is a fundamental combinatorial optimization model studied in the operations research community for the past three quarters of a century. The TSP encapsulates the basic structure at the heart of important vehicle routing applications, and also appears in a variety of other contexts, such as genetics, manufacturing and scheduling [11].

In many of these applications, some or all of the TSP's parameters are not known with certainty ahead of time. The last three decades have seen a variety of work on TSP and other routing models with stochastic parameters. For instance, Jaillet [43, 44] assumed that only the probability that a city must be visited is available beforehand, and the decision maker must decide on an *a priori* or *static* order to visit the cities, omitting the cities that do not require a visit in a particular realization. This work introduced the notion of a priori optimization for routing and general decision making under uncertainty, a popular modeling and solution paradigm to this day [20].

Though a priori optimization offers many benefits, by definition it also restricts the decision maker's options. Soon after its appearance, other authors began considering *dynamic* or *adaptive* settings for stochastic TSP and other routing models. This paradigm offers a more flexible solution space and the potential for cost savings over an a priori solution, at the expense of (usually) more complicated models and a heavier computational burden.

Within the stochastic routing context, in this paper we propose a TSP variant in which arc costs are unknown ahead of time except in distribution, and the objective is to minimize the expected cost of the tour. From a static perspective, the problem reduces to a deterministic TSP using expected arc costs. However, we assume the salesman is allowed to observe outgoing arc realizations at each city before deciding what place to visit next. This modeling choice is relevant to many practical settings in which a decision's cost is unknown a priori but revealed before the decision maker must execute it. For instance, in real-time routing it may be possible for the driver to observe outgoing traffic on different routes before deciding what location to visit next. The intelligent use of such real-time information within routing offers transportation companies the opportunity for differentiation and a competitive advantage [23, 54, 65]. One specific example is urban pickup and delivery, where traffic congestion plays a major role in a route's duration and dynamic routing coupled with real-time traffic information can significantly reduce travel times; [23] mention that such dynamic routing is informally implemented by urban pickup and delivery companies in Tokyo.

One way to approach problems like ours is via *dynamic programming* (DP). Unfortunately, the well known curse of dimensionality severely limits the applicability of traditional DP methodology for any model with the TSP's structure [11]. Furthermore, allowing the salesman to consider arc cost realizations before choosing his next destination implies extending the deterministic DP states by all possible outgoing arc cost realizations, which could lead even to an uncountable state space depending on the costs' support set. To circumvent this difficulty, we employ *approximate linear programming* (ALP); see e.g. [26]. This method involves approximating the true DP cost-to-go function (with an affine function class in our case), and choosing the particular function within this class that yields the best possible lower bound on the optimal expected cost-to-go. Once computed, the approximate cost-to-go can also be used within the traditional Bellman recursion to derive a policy, often called *price-directed* for its dependence on dual multipliers [1, 2, 4, 8]. ALP offers a tractable way to study problems like our dynamic TSP, yielding bounds with theoretical guarantees and empirically verified quality. Moreover, although unmodified price-directed policies do not always produce good solutions, they give rise to high-quality heuristically modified policies that can be computed efficiently.

1.1 Our Contribution

We consider the following to be our main contributions:

§2 We formally propose a dynamic TSP with stochastic arc costs, and demonstrate the advantage of the dynamic decision paradigm over a static tour solution.

§3 We derive a semi-infinite *linear programming* (LP) lower bound from an affine approximation of the cost-to-go function, which requires only expected costs and a description of each random cost vector's support set, and is robust with respect to any distribution having these parameters. We give a worst-case guarantee for the bound's quality, and show that it is polynomially solvable when the support sets are either polytopes with polynomially many extreme points or hyper-rectangles.

§4 We analyze a lookahead version of the price-directed policy implied by our cost-to-go approximation and derive worst-case performance guarantees that depend on the approximation's fidelity to the optimal cost-to-go.

§5 We propose a heuristic version of the lookahead policy and benchmark it and the ALP bound on various instances to demonstrate their effectiveness.

1.2 Literature Review

Models like the one we propose appear only rarely, but the literature on stochastic and/or dynamic TSP and more general vehicle routing is vast. We briefly review some salient topics; the interested reader may refer to [25, 32, 38] for more details. The texts [11, 39, 67] cover the deterministic TSP comprehensively.

The literature includes various stochastic or *probabilistic* TSP models, usually assuming a known distribution governs some of the problem's parameters, and then analyzing the expected cost of the optimal tour, a heuristic and/or a lower bound. Different authors consider uncertainty in the arc costs [55], city locations in a Euclidean instance [37, 43], or the subset of cities to visit from a ground set [43, 44].

In general routing problems, many authors have studied models in which demand, i.e. the requirement to visit a particular city or customer, is uncertain. The a priori approach fixes a customer order or a route, generating different particular solutions based on demand realization, e.g. [17, 21, 43, 44, 73]; see [20] for a recent survey. However, as technological advances enable more real-time computation, the focus has shifted towards models that dynamically respond to demand realization [18, 19, 22, 33, 52, 53, 71]. The surveys [54, 65] cover issues in dynamic routing.

Another paradigm to study dynamic routing is *online* optimization, where instead of minimizing expected costs, the objective is to benchmark a solution against an omniscient algorithm that knows all uncertainty a priori; this benchmark is referred to as a solution's *competitive ratio*. The survey [45] and its references cover many such models.

Routing models in which arc costs are uncertain have also been studied, though perhaps not to the extent of models with uncertain demand. In many cases, costs represent time and the objective is to minimize expected tardiness, the probability of tardiness, or to find a minimum-cost route that meets an acceptable tardiness service level [46, 47, 51]. Our model's approach to uncertain arc costs is also similar to many stochastic shortest path problems, e.g. [15, 56, 59, 62, 64], and sometimes appears in real-time shortest path applications [48, 49, 74].

The most similar models to ours in the literature are perhaps [23, 70]. The model in [23] is also a TSP with random arc costs, but the entire network is visible to the salesman at all times and costs evolve according to an underlying Markov chain. The authors propose an algorithm that generates an optimal policy and computationally demonstrate the benefit of dynamic policies over a fixed tour. Their analysis also indicates that even instances with as few as ten or twelve cities are already computationally challenging and of practical significance. The TSP model in [70] also has stochastic costs and allows dynamic decisions, though the paper focuses more on minimizing tardiness. The author proposes a *rollout* policy [14] and demonstrates its effectiveness with computational experiments. Such policies share several traits with our work; we discuss differences and similarities below.

To our knowledge, this paper is the first application of ALP in a stochastic routing context, and the first author's previous work for the deterministic TSP [75] is the only other use of ALP in routing models. The concept was first studied as early as a quarter-century ago [68, 77, 78], and has received growing attention in the past decade [26, 27, 29]. Specific applications of ALP include commodity valuation [57], economic lot scheduling [5], inventory routing [1, 2], joint replenishment [6, 7, 50], revenue management [3], and stochastic games [30].

However, the general toolkit of *approximate dynamic programming* (ADP) has been used extensively in routing and fleet management, e.g. [35, 36, 76] or [16, Chapter 11.4]; the text [63] includes many such applications and surveys general ADP methodology. There are also many ADP methodologies in addition to ALP for approximating value functions, such as approximate policy iteration [13], approximate value iteration [63], approximate bilinear programming [60, 61], as well as various statistical methods, e.g. parametric and non-parametric regression [40, 63].

The remainder of the paper is organized as follows: Section 2 introduces our notation, formulates the dynamic TSP with stochastic arc costs and provides preliminary results. Section 3 discusses ALP formulations for the problem, gives our cost-to-go approximation, and explains issues related to obtaining the approximate cost-to-go and resulting bound. Section 4 covers how our approximate cost-to-go determines a price-directed policy, and discusses worst-case performance results. Section 5 outlines our computational experiments, and Section 6 concludes and provides future research avenues. The Appendix contains some technical proofs and the experimental results not included in the body of the article.

2 Problem Formulation and Preliminaries

In the TSP, the salesman visits each city in $N := \{1, \dots, n\}$ exactly once starting from and finally returning to a distinguished city 0, sometimes called the *home city* or *depot*. In our model, each arc cost is random and realized upon arrival at the arc's tail. The objective is to minimize the expected total cost of the tour, and the desired solution is not simply a tour, but rather a policy that chooses the next city to visit based on the current location, the remaining cities to visit, as well as the realized vector of outgoing arc costs.

Let $C_i = (C_{ij} : j \in N \cup 0 \setminus i) \in \mathbb{R}^n$ be the random vector of outgoing costs at city i; C_i is realized upon arrival at i. The distribution of each C_i may differ among cities, and all C_i are pairwise independent (though arc costs sharing the same tail are not necessarily independent). Furthermore, we assume C_i only depends on the current city i and not on the remaining set of cities to visit. Let $\mathcal{C}_i \subseteq \mathbb{R}^n$ be the support of C_i; we assume this set is compact. For notational convenience, we use $\bar{c}_{ij}$, $\underline{c}_{ij}$ and $\hat{c}_{ij}$ respectively to denote $\mathsf{E}[C_{ij}]$, $\min_{c_i \in \mathcal{C}_i} c_{ij}$ and $\max_{c_i \in \mathcal{C}_i} c_{ij}$.

We base our DP formulation on the classical formulation for the deterministic TSP [12, 41], augmented to include outgoing costs. A state indicates the current city, the remaining cities to visit, and the realized vector of outgoing travel costs. The state space of the problem is

$$\mathcal{S} := \{(0, N, c_0) : c_0 \in \mathcal{C}_0\} \cup \{(i, U, c_i) : i \in N, U \subseteq N \setminus i,\ c_i \in \mathcal{C}_i\} \cup \{(0, \varnothing)\}.$$

States $\{(0, N, c_0) : c_0 \in \mathcal{C}_0\}$ correspond to the start of the tour. In state $(0, N, c_0)$, the salesman is at city 0, has all cities in N left to visit, and the outgoing arc costs are given by c_0. The states $\{(i, U, c_i) : i \in N,\ U \subseteq N \setminus i,\ c_i \in \mathcal{C}_i\}$ correspond to intermediate steps of the tour. In state (i, U, c_i), the salesman is at city $i \in N$, has the cities in U left to visit, and the costs are given by c_i. The terminal state $(0, \varnothing)$ corresponds to the end of the tour. In each state (i, U, c_i) with $U \neq \varnothing$, the salesman must choose a city $j \in U$ to visit next. Then, the salesman will transition to some state $(j, U \setminus j, c_j)$ with $c_j \in \mathcal{C}_j$, according to the distribution of C_j. The transition from (i, U) to $(j, U \setminus j)$ is deterministic, while outgoing arc costs at $(j, U \setminus j)$ are dictated by the distribution of C_j. The cardinality $|U|$ decreases by one with each transition, and we sometimes use $t := n - |U|$ as a time index in the problem. At the start of the tour we have $|U| = n$, and the time period t is zero. After the first transition, $t = n - |U| = 1$, and so forth.

Let $y^*_{i,U}(c_i) \in \mathbb{R}$ represent the optimal expected cost-to-go from state $(i, U, c_i) \in \mathcal{S}$; the collection $(y^*_{i,U}(c_i) : c_i \in \mathcal{C}_i)$ can be interpreted as a random variable or as a function of c_i. The terminal cost is zero, $y^*_{0,\varnothing} = 0$, and we obtain the DP recursion

$$y^*_{i,U}(c_i) := \begin{cases} \min_{j \in U}\{c_{ij} + \mathsf{E}[y^*_{j,U\setminus j}(C_j)]\}, & U \neq \varnothing \\ c_{i0}, & i \in N, U = \varnothing \\ 0, & i = 0, U = \varnothing \end{cases}. \tag{1}$$

Proposition 2.1. *Each $y^*_{i,U} : \mathcal{C}_i \to \mathbb{R}$ is non-decreasing, piecewise linear and concave as a function of c_i.*

Proof. From the independence of the random vectors C_i and C_j it follows that the expectation in (1) is constant with respect to c_i, and therefore $y^*_{i,U}$ is the minimum over a set of affine functions. □

A solution of (1) induces a policy $\pi^* : \mathcal{S} \setminus (0, \varnothing) \to N \cup 0$ that maps from the current state $(i, U, c_i) \in \mathcal{S}$ to an action $j \in U$,

$$\pi^*(i, U, c_i) := \begin{cases} \arg\min_{j \in U}\{c_{ij} + \mathsf{E}[y^*_{U\setminus j}(C_j)]\}, & U \neq \varnothing \\ 0, & U = \varnothing \end{cases}, \tag{2}$$

breaking ties arbitrarily. The LP formulation of (1) is

$$\max_y \; \mathsf{E}[y_{0,N}(C_0)] \tag{3a}$$

$$\text{s.t. } y_{0,N}(c_0) - \mathsf{E}[y_{i,N\setminus i}(C_i)] \le c_{0i}, \qquad \forall\, i \in N, c_0 \in \mathcal{C}_0 \tag{3b}$$

$$y_{i,U\cup j}(c_i) - \mathsf{E}[y_{j,U}(C_j)] \le c_{ij}, \qquad \forall\, i \in N, j \in N \setminus i, U \subseteq N \setminus \{i, j\}, c_i \in \mathcal{C}_i \tag{3c}$$

$$y_{i,\varnothing}(c_i) \le c_{i0}, \qquad \forall\, i \in N, c_i \in \mathcal{C}_i. \tag{3d}$$

This LP is potentially doubly infinite, since each variable $y_{i,U}$ is a function from the possibly infinite set $\mathcal{C}_i$ to $\mathbb{R}$, and the constraints range over the state-action pairs, also indexed partly by the sets $\mathcal{C}_i$. In light of Proposition 2.1, we can restrict the feasible region of each $y_{i,U}$ to an appropriately chosen, well-behaved functional space on $\mathcal{C}_i$, such as the space of continuous functions on $\mathcal{C}_i$; the exact choice is not important since (3) is usually intractable and we do not intend to solve it directly. The formulation is important, however, because any feasible solution provides a lower bound on the optimal expected cost-to-go.

Lemma 2.2. *Let $y_{i,U} : \mathcal{C}_i \to \mathbb{R}$ for $i \in N \cup 0$ and $U \subseteq N \setminus i$ be feasible for (3). Then $y_{i,U}(c_i) \le y^*_{i,U}(c_i)$ for all $(i, U, c_i) \in \mathcal{S}$. In particular, $\mathsf{E}[y_{0,N}(C_0)] \le \mathsf{E}[y^*_{0,N}(C_0)]$.*

Proof. The proof follows inductively from the definition of y^*. □

2.1 Allowing Returns

One important issue is the impact of requiring exactly one visit to each city. In certain settings it could be reasonable to allow returns to previously visited cities or the depot, if this could reduce overall costs. In a deterministic TSP, this requirement can be made without loss of generality if costs satisfy the triangle inequality, but the situation is more complicated in our case. To begin, allowing returns transforms our model from a finite-stage to an infinite-stage DP with walks of arbitrary length. However, we can ensure a finite objective and walks of bounded length if we assume costs are almost surely positive and bounded away from zero, which is reasonable in most practical settings. The more complex issue involves how information is used in repeat visits. Each of two possible extensions offers its own challenges and advantages, which we briefly outline:

i) Costs are fixed upon their first realization: Suppose the costs the salesman observes upon his first arrival at a city are fixed for the remainder of the problem. This variant is a TSP version of the stochastic shortest path problem with recourse [62], or a stochastic version of the *Canadian traveler's problem* [58]. Even shortest path versions of this problem are quite difficult: The adversarial version is PSPACE-complete and the stochastic version is #P-hard [58, 62]. The state space $\mathcal{S}$ must be augmented to include previously observed costs from past visits, increasing its dimension by an order of magnitude and drastically increasing the number of possible states. Though the overall approach outlined in this work could be applied to such a problem, it would require additional analysis and is beyond our current scope.

ii) Costs are re-sampled for each visit to a city: Now suppose costs are re-sampled from the same distribution every time the salesman arrives at a city. In this case the structure of the state space $\mathcal{S}$ remains the same, but we add intermediate states $(0, U, c_0)$ representing a mid-way stop at the depot, and we augment the action space by always allowing the salesman to visit any city. This extension is also quite challenging, since after visiting the last city the problem is still a stochastic shortest path problem [15]. The DP recursion is now

$$y^*_{i,U}(c_i) = \min\Big\{\min_{j\in U}\{c_{ij} + \mathsf{E}[y^*_{j,U\setminus j}(C_j)]\}, \min_{j\in N\cup 0\setminus(U\cup i)}\{c_{ij} + \mathsf{E}[y^*_{j,U}(C_j)]\}\Big\},\ i \in N\cup 0, U \subseteq N\setminus i,$$

where we again assume $y^*_{0,\varnothing} = 0$, and the LP formulation is

$$\begin{aligned}
\max_y\ & \mathsf{E}[y_{0,N}(C_0)] \\
\text{s.t. } & y_{i,U\cup j}(c_i) - \mathsf{E}[y_{j,U}(C_j)] \le c_{ij}, && \forall\, i \in N\cup 0, j \in N\setminus i, U \subseteq N\setminus\{i,j\}, c_i \in \mathcal{C}_i \\
& y_{i,U}(c_i) - \mathsf{E}[y_{j,U}(C_j)] \le c_{ij}, && \forall\, i \in N\cup 0, j \in N\cup 0\setminus i, U \subseteq N\setminus\{i,j\}, c_i \in \mathcal{C}_i \\
& y_{0,\varnothing}(c_i) \le 0, && \forall\, c_0 \in \mathcal{C}_0.
\end{aligned}$$

Some of our subsequent analysis and approximation of (3) applies to this LP at the expense of slightly more complex bases and notation; this includes the tractability of an ALP bound and related policies. However, our theoretical worst-case guarantees on the bound and resulting policies do not carry through, as they depend on structural properties that are lost in this variant of the problem.

2.2 Comparison to Static TSP

Another question regarding the dynamic TSP model is whether we gain by allowing dynamic decisions; i.e. by being adaptive [28]. An alternative for the problem would be to solve a deterministic TSP, perhaps with arc costs given by $\bar{c}_{ij} = \mathsf{E}[C_{ij}]$, and implement this solution regardless of the actual cost realizations.

Example 2.3 (Adaptivity Gap)**.** Consider an instance in which arc costs entering cities in N are i.i.d. Bernoulli random variables with parameter $p \in (0,1)$, and arc costs entering the depot are zero with probability one. The expected cost of any fixed tour is then pn. In contrast, a greedy dynamic policy that at every city chooses any outgoing arc of minimum cost fares much better. At period $t = 0, \dots, n-1$, the policy incurs zero cost unless all outgoing arcs have unit cost, which occurs with probability p^{n-t}. At the final period, the policy incurs no cost. Therefore, the expected total cost of the greedy policy is

$$\sum_{t=0}^{n-1} p^{n-t} = \frac{p(1-p^n)}{1-p}.$$

The ratio of the two costs is

$$\frac{(1-p)n}{1-p^n},$$

and this ratio goes to infinity as $n \to \infty$. In the terminology introduced in [28], our problem has an infinite *adaptivity gap*; that is, there exist problem instances for which a dynamic policy performs arbitrarily better than a fixed route. In other words, allowing dynamic updating of decisions may significantly decrease expected cost over any fixed tour. Note that we invert the ratio from the original definition because our model is a minimization problem.

On the other hand, if each $\mathcal{C}_i$ is sufficiently "small" the difference between the dynamic TSP and its deterministic counterpart given by $\bar{c}$ may be small. Define the optimal cost-to-go function for this deterministic TSP as

$$\bar{y}_{i,U} := \begin{cases} \min_{j \in U}\{\bar{c}_{ij} + \bar{y}_{j,U\setminus i}\}, & U \neq \varnothing \\ \bar{c}_{i0}, & i \in N, U = \varnothing \\ 0, & i = 0, U = \varnothing \end{cases}. \tag{4}$$

For $A \subseteq \mathbb{R}^n$, let $\mathbb{D}(A) := \sup_{x,y \in A} \|x - y\|$ be A's *diameter*.

Proposition 2.4. *For any* $(i, U, c_i) \in \mathcal{S}$,

$$|y^*_{i,U}(c_i) - \bar{y}_{i,U}| \le \sum_{j \in U \cup i} \mathbb{D}(\mathcal{C}_j), \tag{5a}$$

and

$$|\mathsf{E}[y^*_{0,N}(c_0)] - \bar{y}_{0,N}| \le \sum_{i \in N \cup 0} \mathbb{D}(\mathcal{C}_i). \tag{5b}$$

Proof. Let $i \in N$ and $c_i \in \mathcal{C}_i$. Starting with $U = \varnothing$,

$$|y^*_{i,\varnothing}(c_i) - \bar{y}_{i,\varnothing}| = |c_{i0} - \bar{c}_{i0}| \le \sup_{c_i \in \mathcal{C}_i} \{|c_{i0} - \bar{c}_{i0}|\} \le \mathbb{D}(\mathcal{C}_i),$$

using the fact that $\bar{c}_i \in \text{conv}(\mathcal{C}_i)$. Subsequently, the difference between the stochastic and deterministic cost-to-go functions is

$$\begin{aligned}
|y^*_{i,U}(c_i) - \bar{y}_{i,U}| &= \left|\min_{j\in U}\{c_{ij} + \mathsf{E}[y^*_{j,U\setminus j}(C_j)]\} - \min_{j\in U}\{\bar{c}_{ij} + \bar{y}_{j,U\setminus j}\}\right| \\
&\le \max_{j\in U}\{|c_{ij} + \mathsf{E}[y^*_{j,U\setminus j}(C_j)] - \bar{c}_{ij} - \bar{y}_{j,U\setminus j}|\} \\
&\le \max_{j\in U}\{|c_{ij} - \bar{c}_{ij}|\} + \max_{j\in U}\{|\mathsf{E}[y^*_{j,U\setminus j}(C_j)] - \bar{y}_{j,U\setminus j}|\} \\
&\le \mathbb{D}(\mathcal{C}_i) + \max_{j\in U}\{|\mathsf{E}[y^*_{j,U\setminus j}(C_j)] - \bar{y}_{j,U\setminus j}|\} \le \sum_{j\in U\cup i} \mathbb{D}(\mathcal{C}_j),
\end{aligned}$$

where the last inequality follows from induction. □

The norm in the definition of $\mathbb{D}$ may be any ℓ_p norm, i.e. $\|x\|_p = \left(\sum_i |x_i|^p\right)^{\frac{1}{p}}$. Also, from the definition of $\bar{c}$ this proof only uses $\bar{c}_i \in \text{conv}(\mathcal{C}_i)$, and thus a similar result holds for any fixed set of costs in the convex hull of each support set. Since the difference between the optimal expected cost-to-go in our model and the deterministic cost-to-go is bounded by the diameters of each support set, if these diameters are small enough (e.g. constant with respect to n), we can approximate the dynamic TSP deterministically with *any* possible realization of arc costs and obtain a close approximation.

Another useful deterministic option for our problem is to consider the *optimistic* deterministic TSP with arc costs $\underline{c}_{ij} = \min_{c_i\in\mathcal{C}_i} c_{ij}$. Let $\underline{y}_{i,U}$ be the optimistic cost-to-go function defined by (4) with costs $\underline{c}$ instead of $\bar{c}$. For two sets $A, B \subseteq \mathbb{R}^n$, let $\mathbb{D}(A,B) := \sup_{x\in A}\inf_{y\in B}\|x-y\|$ be the sets' *deviation*; see e.g. [72, Chapter 7].

Proposition 2.5. *For any* $(i, U, c_i) \in \mathcal{S}$,

$$y^*_{i,U}(c_i) - \sum_{j\in U\cup i} \mathbb{D}(\mathcal{C}_j, \underline{c}_j) \le \underline{y}_{i,U} \le y^*_{i,U}(c_i), \tag{6a}$$

and

$$\mathsf{E}[y^*_{0,N}(C_0)] - \sum_{i\in N\cup 0} \mathbb{D}(\mathcal{C}_i, \underline{c}_i) \le \underline{y}_{0,N} \le \mathsf{E}[y^*_{0,N}(C_0)]. \tag{6b}$$

Proof. The left-hand inequality is proved in a similar fashion to Proposition 2.4. For the right-hand inequality, if we let $y_{i,U}(c_i) = \underline{y}_{i,U}$, $\forall\ c_i \in \mathcal{C}_i$, then y is feasible for (3), and thus $\underline{y}_{i,U}$ is a lower bound for the cost-to-go at any state (i, U, c_i); in particular, $\mathsf{E}[y_{0,N}(C_0)] = \underline{y}_{0,N}$ is a lower bound on the optimal value of (3a). □

Proposition 2.5 implies that any bound for a deterministic TSP cost-to-go function with costs $\underline{c}$ also provides a lower bound for the dynamic TSP cost-to-go, and the bound's quality partly depends on how much the actual costs can vary from the optimistic prediction $\underline{c}$. We compare our approach to one such bound in the next section.

3 Approximate Linear Program

The exact solution of (3) is intractable for even moderately sized instances. In ALP, for $m \lll |\mathcal{S}|$ we define a collection of *basis vectors* or *functions* $b_{i,U} : \mathcal{C}_i \to \mathbb{R}^m$, for each $i \in N \cup 0$, $U \subseteq N \setminus i$. Then for any $\lambda \in \mathbb{R}^m$, we can approximate the cost-to-go as $y_{i,U}(c_i) \approx \langle \lambda, b_{i,U}(c_i) \rangle$, where $\langle \cdot, \cdot \rangle$ represents the inner product. The corresponding *approximate* LP is

$$\max_{\lambda} \ \mathsf{E}[\langle \lambda, b_{0,N}(C_0) \rangle] \tag{7a}$$

$$\text{s.t.} \ \langle \lambda, b_{0,N}(c_0) \rangle - \mathsf{E}[\langle \lambda, b_{i,N\setminus i}(C_i) \rangle] \leq c_{0i}, \qquad \forall\, i \in N, c_0 \in \mathcal{C}_0 \tag{7b}$$

$$\langle \lambda, b_{i,U\cup j}(c_i) \rangle - \mathsf{E}[\langle \lambda, b_{j,U}(C_j) \rangle] \leq c_{ij}, \qquad \forall\, i \in N, j \in N \setminus i, U \subseteq N \setminus \{i,j\}, c_i \in \mathcal{C}_i \tag{7c}$$

$$\langle \lambda, b_{i,\varnothing}(c_i) \rangle \leq c_{i0}, \qquad \forall\, i \in N, c_i \in \mathcal{C}_i. \tag{7d}$$

Problem (7) is a semi-infinite LP; when the support sets $\mathcal{C}_i$ are finite, it is a finite-dimensional LP. Because it defines a feasible solution for (3), any feasible solution of (7) immediately implies a lower bound on (3a) with (7a).

Example 3.1 (Deterministic Basis)**.** As an initial example, suppose the basis b does not differ based on arc costs, so that $b_{i,U}(c_i) = b_{i,U}$ only depends on the current city i and the set of remaining cities U. Recall that $\underline{c}_{ij} = \min_{c_i \in \mathcal{C}_i} c_{ij}$; the left-hand sides of the constraints (7b) through (7d) do not vary with c, and therefore (7) becomes

$$\begin{aligned}
\max_{\lambda} \ & \langle \lambda, b_{0,N} \rangle \\
\text{s.t.} \ & \langle \lambda, b_{0,N} - b_{i,N\setminus i} \rangle \leq \underline{c}_{0i}, && \forall\, i \in N \\
& \langle \lambda, b_{i,U\cup j} - b_{j,U} \rangle \leq \underline{c}_{ij}, && \forall\, i \in N, j \in N \setminus i, U \subseteq N \setminus \{i,j\} \\
& \langle \lambda, b_{i,\varnothing} \rangle \leq \underline{c}_{i0}, && \forall\, i \in N;
\end{aligned}$$

Any feasible solution to this LP provides a lower bound on the optimal value of the deterministic TSP with optimistic costs $\underline{c}$, and thus a lower bound for (3) through Proposition 2.5.

Though we do not pursue it here, it is possible to apply ALP to the deterministic TSP in this fashion to obtain lower bounds [75]. Instead, our focus is to solve (7) for a more complex basis that accounts for different arc costs.

3.1 Affine Cost-to-Go Approximation

Consider the cost-to-go approximation

$$y_{0,N}(c_0) \approx \lambda_0 + \sum_{i \in N} c_{0i}\eta_{0i}, \qquad \forall\, c_0 \in \mathcal{C}_0 \tag{8a}$$

$$y_{i,U}(c_i) \approx \lambda_{i0} + \sum_{k \in U} (\lambda_{ik} + c_{ik}\eta_{ik}), \qquad \forall\, i \in N, \varnothing \neq U \subseteq N \setminus i, c_i \in \mathcal{C}_i \tag{8b}$$

$$y_{i,\varnothing}(c_i) \approx \lambda_{i0} + c_{i0}\eta_{i0}, \qquad \forall\, i \in N, c_i \in \mathcal{C}_i, \tag{8c}$$

where $\lambda \in \mathbb{R}^{n^2+1}$, $\eta \in \mathbb{R}^{n^2+n}$. Intuitively, from a city $i \in N$ basis (8) assigns a nominal cost λ_{i0} for returning to the depot, and additional nominal costs λ_{ik} for each city that must be visited prior to the return. The cost is then adjusted by η_{ik} based on the realization c_{ik} of each outgoing cost to a city k that could be visited next. The approximation uses only one nominal cost, λ_0, for the initial states at the home city 0, because the salesman always has the same remaining cities from that location. With this approximation, (7) becomes

$$\max_{\lambda,\eta} \ \lambda_0 + \sum_{k\in N} \bar{c}_{0k}\eta_{0k} \tag{9a}$$

$$\text{s.t. } \lambda_0 - \lambda_{i0} + c_{0i}\eta_{0i} + \sum_{k\in N\setminus i} (c_{0k}\eta_{0k} - \lambda_{ik} - \bar{c}_{ik}\eta_{ik}) \le c_{0i}, \quad \forall\, i \in N, c_0 \in \mathcal{C}_0 \tag{9b}$$

$$\begin{aligned}&\lambda_{i0} - \lambda_{j0} + \lambda_{ij} + c_{ij}\eta_{ij} + \sum_{k\in U}(\lambda_{ik} + c_{ik}\eta_{ik} - \lambda_{jk} - \bar{c}_{jk}\eta_{jk}) \le c_{ij},\\ &\quad \forall\, i \in N, j \in N\setminus i, \varnothing \ne U \subseteq N \setminus \{i,j\}, c_i \in \mathcal{C}_i\end{aligned} \tag{9c}$$

$$\lambda_{i0} - \lambda_{j0} + \lambda_{ij} + c_{ij}\eta_{ij} - \bar{c}_{j0}\eta_{j0} \le c_{ij}, \quad \forall\, i \in N, j \in N\setminus i, c_i \in \mathcal{C}_i \tag{9d}$$

$$\lambda_{i0} + c_{i0}\eta_{i0} \le c_{i0}, \quad \forall\, i \in N, c_i \in \mathcal{C}_i. \tag{9e}$$

As with many ALP's, even though we have reduced the decision variables to a manageable number, the constraint set remains very large – at least exponential in n and possibly uncountable depending on the support sets $\mathcal{C}_i$.

Before discussing the model's optimization, we formulate its dual to derive further insight into the approximation given by (8). The dual is

$$\min_x \ \sum_{i\in N}\sum_{c_0\in\mathcal{C}_0} c_{0i}x_{0i}^{c_0} + \sum_{i\in N}\sum_{j\in N\setminus i}\sum_{U\subseteq N\setminus\{i,j\}}\sum_{c_i\in\mathcal{C}_i} c_{ij}x_{ij,U}^{c_i} + \sum_{i\in N}\sum_{c_i\in\mathcal{C}_i} c_{i0}x_{i0}^{c_i} \tag{10a}$$

$$\text{s.t. } \sum_{i\in N}\sum_{c_0\in\mathcal{C}_0} x_{0i}^{c_0} = 1 \tag{10b}$$

$$\sum_{i\in N}\sum_{c_0\in\mathcal{C}_0} c_0 x_{0i}^{c_0} = \bar{c}_0 \tag{10c}$$

$$-\sum_{c_0\in\mathcal{C}_0} x_{0i}^{c_0} + \sum_{j\in N\setminus i}\sum_{U\subseteq N\setminus\{i,j\}}\left[\sum_{c_i\in\mathcal{C}_i} x_{ij,U}^{c_i} - \sum_{c_j\in\mathcal{C}_j} x_{ji,U}^{c_j}\right] + \sum_{c_i\in\mathcal{C}_i} x_{i0}^{c_i} = 0, \quad \forall\, i \in N \tag{10d}$$

$$-\bar{c}_{i0}\sum_{j\in N\setminus i}\sum_{c_j\in\mathcal{C}_j} x_{ji,\varnothing}^{c_j} + \sum_{c_i\in\mathcal{C}_i} c_{i0}x_{i0}^{c_i} = 0, \quad \forall\, i \in N \tag{10e}$$

$$\begin{aligned}&-\sum_{c_0\in\mathcal{C}_0} x_{0i}^{c_0} + \sum_{U\subseteq N\setminus\{i,j\}}\sum_{c_i\in\mathcal{C}_i} x_{ij,U}^{c_i}\\ &\quad + \sum_{k\in N\setminus\{i,j\}}\sum_{U\subseteq N\setminus\{i,j,k\}}\left[\sum_{c_i\in\mathcal{C}_i} x_{ik,U\cup j}^{c_i} - \sum_{c_k\in\mathcal{C}_k} x_{ki,U\cup j}^{c_k}\right] = 0, \quad \forall\, i \in N, j \in N\setminus i\end{aligned} \tag{10f}$$

$$\begin{aligned}&-\bar{c}_{ij}\sum_{c_0\in\mathcal{C}_0} x_{0i}^{c_0} + \sum_{U\subseteq N\setminus\{i,j\}}\sum_{c_i\in\mathcal{C}_i} c_{ij}x_{ij,U}^{c_i}\\ &\quad + \sum_{k\in N\setminus\{i,j\}}\sum_{U\subseteq N\setminus\{i,j,k\}}\left[\sum_{c_i\in\mathcal{C}_i} c_{ij}x_{ik,U\cup j}^{c_i} - \bar{c}_{ij}\sum_{c_k\in\mathcal{C}_k} x_{ki,U\cup j}^{c_k}\right] = 0, \quad \forall\, i \in N, j \in N\setminus i\end{aligned} \tag{10g}$$

$$x \ge 0; \quad x \text{ has finite support.} \tag{10h}$$

In this relaxed primal model, each variable $x^{c_i}_{ij,U}$ corresponds to the probability of visiting state $(i, U \cup j, c_i)$ and choosing action j. Similarly, variables $x^{c_0}_{0i}$ and $x^{c_i}_{i0}$ respectively correspond to the probability of choosing i from initial state $(0, N, c_0)$ and of visiting state $(i, \varnothing, c_i)$ immediately before the terminal state. This set of decision variables may be finite, countably or uncountably infinite depending on the sets $\mathcal{C}_i$, but in any of these cases only a finite number must be positive at any optimal solution. The proof of the following lemma is included in the Appendix.

Lemma 3.2. *Problems* (9) *and* (10) *are strong duals. That is, the two problems are weak duals, both attain their optimal values, and these optimal values are equal.*

The objective (10a) minimizes the expected cost given by the probabilities x. Constraint (10b), corresponding to λ_0, requires the total probability of visiting state-action pairs $(0, N, c_0, i)$ at the depot to be one. Similarly, the vector equation (10c), which corresponds to η_{0i} for each $i \in N$, requires the expected cost vector implied by probabilities $x^{c_0}_{0i}$ to equal $\bar{c}_0$. This is indeed a relaxation, since the probabilities implied by an actual policy must exactly match C_0's distribution, and not simply match its expectation. Constraint (10d), corresponding to λ_{i0}, is a probability flow balance requiring the probability of entering city i to equal the probability of exiting it. The constraint (10e) stemming from η_{i0} requires the expected cost of returning to the depot 0 from i, conditioned upon visiting i last, to be equal to $\bar{c}_{i0}$. Constraint (10f), corresponding to λ_{ij}, is a probability flow balance on the ordered pair (i, j): The probability of visiting i before j must equal the probability of exiting i when j is still remaining, either by going to j itself or to another remaining city. Here again we have a relaxation, since in a policy this must hold not only for individual cities j, but for sets of cities as well. Finally, (10g), which corresponds to η_{ij}, requires the expectation of arc (i, j)'s cost, conditioned upon visiting i before j, to equal $\bar{c}_{ij}$.

The relaxed primal also allows us to make an additional common sense observation about the cost-to-go approximation (8). Intuitively, we expect higher realized costs in any state to imply a higher cost-to-go; the next result confirms this intuition.

Corollary 3.3. *In* (9)*, we can impose* $\eta \geq 0$ *without loss of optimality.*

Proof. Adding $\eta \geq 0$ to (9) and using an argument identical to the proof of Lemma 3.2, we obtain a strong dual optimization problem similar to (10), except with constraints (10c), (10e) and (10g) relaxed to greater-than-or-equal. However, any optimal solution to the modified model must satisfy these constraints at equality, because an optimal solution would not have outgoing probability flow in any of these constraints that is more expensive than necessary. □

Another question about (9) is how the bound it provides for (3) compares to other tractable bounds. One possible comparison is with the LP relaxation of the arc-based formulation for the deterministic TSP with optimistic costs $\underline{c}$,

$$\min_{z} \sum_{i \in N \cup 0} \sum_{j \in N \cup 0 \setminus i} \underline{c}_{ij} z_{ij} \tag{11a}$$

$$\text{s.t.} \sum_{j \in N \cup 0 \setminus i} z_{ij} = 1, \qquad \forall\, i \in N \cup 0 \tag{11b}$$

$$\sum_{j \in N \cup 0 \setminus i} z_{ji} = 1, \qquad \forall\, i \in N \cup 0 \tag{11c}$$

$$\sum_{i \in U} \sum_{j \in N \cup 0 \setminus U} z_{ij} \geq 1, \qquad \forall\, \varnothing \neq U \subseteq N \cup 0 \tag{11d}$$

$$z \geq 0. \tag{11e}$$

We use z variables instead of x to distinguish the formulations. This LP is solvable in polynomial time [11] and is a lower bound for the optimal expected cost $\mathsf{E}[y^*_{0,N}(C_0)]$, since it is a lower bound on $\underline{y}_{0,N}$, the deterministic TSP with arc costs $\underline{c}$, which in turn bounds our problem by Proposition 2.5.

Theorem 3.4. *The optimal value of* (9) *provides a lower bound greater than or equal to the bound provided by* (11)*.*

Proof. By setting $\eta_{ij} = 0$, $\forall\, i \in N \cup 0$, $j \in N \cup 0 \setminus i$ in (9), we obtain the LP

$$\begin{aligned}
\max_{\lambda}\ & \lambda_0 \\
\text{s.t. } & \lambda_0 - \lambda_{i0} - \sum_{k \in N \setminus i} \lambda_{ik} \leq \underline{c}_{0i}, && \forall\, i \in N \\
& \lambda_{i0} - \lambda_{j0} + \lambda_{ij} + \sum_{k \in U} (\lambda_{ik} - \lambda_{jk}) \leq \underline{c}_{ij}, && \forall\, i \in N, j \in N \setminus i, U \subseteq N \setminus \{i,j\} \\
& \lambda_{i0} \leq \underline{c}_{i0}, && \forall\, i \in N.
\end{aligned}$$

It follows from Theorem 18 in [75] that this LP's optimal value is greater than or equal to the optimal value of (11). □

The next example shows that the bound provided by our affine cost-to-go approximation can indeed exceed the best possible deterministic bound.

Example 3.5 (Example 2.3 Continued)**.** Consider again the instance where the C_{ij} with $j \in N$ are i.i.d. Bernoulli random variables with parameter $p \in (0,1)$, and the C_{i0} are zero with probability one. Because $\underline{c} = 0$, no deterministic bound can improve on the trivial zero bound. The ALP (9) is now

$$\begin{aligned}
\max_{\lambda,\eta}\ & \lambda_0 + p \sum_{k \in N} \eta_{0k} \\
\text{s.t. } & \lambda_0 - \lambda_{i0} + c_{0i}\eta_{0i} + \sum_{k \in N \setminus i} (c_{0k}\eta_{0k} - \lambda_{ik} - p\eta_{ik}) \leq c_{0i}, \quad \forall\, i \in N, c_0 \in \{0,1\}^N \\
& \lambda_{i0} - \lambda_{j0} + \lambda_{ij} + c_{ij}\eta_{ij} + \sum_{k \in U} (\lambda_{ik} + c_{ik}\eta_{ik} - \lambda_{jk} - p\eta_{jk}) \leq c_{ij}, \\
& \quad \forall\, i \in N, j \in N \setminus i, U \subseteq N \setminus \{i,j\}, c_i \in \{0,1\}^{U \cup j} \\
& \lambda_{i0} \leq 0, \quad \forall\, i \in N; \quad \eta \geq 0.
\end{aligned}$$

By the instance's symmetry, we can take $\lambda_{i0} = \lambda_{j0}$, $\lambda_{ik} = \lambda_{jk}$ for distinct $i, j, k \in N$, with similar equalities holding for η. The last constraint class immediately implies $\lambda_{i0} = 0$, and η_{i0} may be ignored since it does not appear in the LP. After applying these derived equations and using $\eta \geq 0$, the first and second constraint classes collapse to

$$\begin{aligned}
\lambda_0 + (c_{i0} + (n-1))\eta_{0i} &\leq c_{0i} + (n-2)(\lambda_{ij} + p\eta_{ij}), && \forall\, c_{0i} \in \{0,1\} \\
\lambda_{ij} + (c_{ij} + (n-2)(1-p))\eta_{ij} &\leq c_{ij}, && \forall\, c_{ij} \in \{0,1\}.
\end{aligned}$$

Therefore, the ALP reduces to two decoupled two-variable LP's,

$$\begin{aligned} \max_{\lambda_0,\eta_{0i}} \quad & \lambda_0 + pn\eta_{0i} \\ \text{s.t.} \quad & \lambda_0 + (n-1)\eta_{0i} \le (n-2)V \\ & \lambda_0 + n\eta_{0i} \le 1 + (n-2)V \\ & \eta_{0i} \ge 0 \end{aligned} \qquad \begin{aligned} \max_{\lambda_{ij},\eta_{ij}} \quad & \lambda_{ij} + p\eta_{ij} \\ \text{s.t.} \quad & \lambda_{ij} + (n-2)(1-p)\eta_{ij} \le 0 \\ & \lambda_{ij} + (1 + (n-2)(1-p))\eta_{ij} \le 1 \\ & \eta_{ij} \ge 0, \end{aligned}$$

where V represents the second LP's optimal value. It is simple to show that $V = \max\{0, p - (n-2)(1-p)\}$, and using this, the first LP's optimal value is then

$$\max\{0, (n-1)(p(n-1) - (n-2)), pn - (n-1)^2(1-p)\}.$$

This bound is non-zero for $p > \frac{n-2}{n-1}$, and approaches the optimal expected cost-to-go as $p \to 1$.

3.2 Constraint Generation

To efficiently model the constraints in problem (9), we use constraint generation; see also [1, 2, 7, 50, 75]. For constraint classes (9b, 9d, 9e), the separation problem is equivalent to maximizing a linear function over one of the sets $\mathcal{C}_i$, for $i \in N \cup 0$. As long as this maximization can be carried out efficiently (i.e. in polynomial time), these constraint classes can be accounted for efficiently as well. However, for constraints (9c), the situation is more complex. Fix λ and η; for an ordered pair of cities $i \in N$ and $j \in N \setminus i$ the separation problem is equivalent to

$$\max_{\substack{c_i \in \mathcal{C}_i \\ \varnothing \neq U \subseteq N \setminus \{i,j\}}} \quad c_{ij}(\eta_{ij} - 1) + \sum_{k \in U} (\lambda_{ik} + c_{ik}\eta_{ik} - \lambda_{jk} - \bar{c}_{jk}\eta_{jk}). \tag{12}$$

This problem is a bilinear mixed-binary optimization problem over the compact set $\mathcal{C}_i \times \{0,1\}^{N \setminus \{i,j\}}$, and thus usually intractable.

Lemma 3.6. *The separation problem* (12) *is NP-hard, even when the sets $\mathcal{C}_i$ are ℓ_2 balls.*

The proof of this lemma is in the Appendix. In this general case, exact separation is inefficient; constraint sampling [26, 27, 29] is likely the only viable choice, but suffers from requiring access to an idealized distribution over the constraints. In our current research we are exploring possible solutions to this difficulty. Nonetheless, there are tractable special cases of the separation problem, which we discuss next.

Proposition 3.7. *Suppose each set $\mathcal{C}_i$ for $i \in N$ has polynomially many extreme points, say $O(p(n))$. Then the separation problem* (12) *is solvable in $O(np(n))$ time for each ordered pair (i, j), and thus* (9) *is solvable in polynomial time via the ellipsoid algorithm.*

Proof. For a fixed $c_i \in \mathcal{C}_i$, the maximization in (12) can be solved greedily: For each $k \in N \setminus \{i, j\}$, include it in the set U only if the term in the parenthesis is positive. When none of these terms is positive, include the non-positive term of smallest absolute value to ensure $U \neq \varnothing$. If $\mathcal{C}_i$ has $O(p(n))$ extreme points, we can carry out this greedy optimization procedure at each extreme point and then choose the overall maximizer, all in $O(np(n))$ time. □

As two examples, Proposition 3.7 covers the cases where each set $\mathcal{C}_i$ is a simplex or an ℓ_1 ball.

Proposition 3.8. *Suppose the convex hull of each set $\mathcal{C}_i$ for $i \in N$ is a hyper-rectangle:* $\mathrm{conv}(\mathcal{C}_i) = [\underline{c}_i, \hat{c}_i]$*, with $\underline{c}_i, \hat{c}_i \in \mathbb{R}^n$ respectively representing lower- and upper-bound vectors. Then the separation problem* (12) *is solvable in $O(n)$ time for each ordered pair (i,j), and thus* (9) *is solvable in polynomial time.*

Proof. With hyper-rectangles, (12) also has a simple greedy algorithm: For each $k \in N \setminus \{i,j\}$, by Corollary 3.3 we have $\max\{\underline{c}_{ik}\eta_{ik}, \hat{c}_{ik}\eta_{ik}\} = \hat{c}_{ik}\eta_{ik}$. So if $\lambda_{ik} + \hat{c}_{ik}\eta_{ik} - \lambda_{jk} - \bar{c}_{jk}\eta_{jk} > 0$, add $k \in U$; otherwise $k \notin U$. If no term is positive, include the non-positive term with smallest absolute value to ensure $U \neq \varnothing$. Finally, if $\eta_{ij} - 1 > 0$, use $\hat{c}_{ij}$ as a coefficient for it; otherwise use $\underline{c}_{ij}$. □

This last result also suggests a tractable option when dealing with more general arc cost support sets $\mathcal{C}_i$: Solve (12) over the hyper-rectangles defined by $\underline{c}_{ij} = \min_{c_i \in \mathcal{C}_i} c_{ij}$ and $\hat{c}_{ij} = \max_{c_i \in \mathcal{C}_i} c_{ij}$ respectively. This approach yields a more conservative but computationally efficient bound.

4 Price-Directed Policies

Any candidate solution y to (3), regardless of feasibility or optimality, determines a policy via (2) by substituting it for y^*. Such policies are called *greedy* with respect to y [26, 29], or *price-directed* [1, 2, 4, 8]. In the specific case of the approximation given by (8), for any $\lambda \in \mathbb{R}^{n^2+1}$, $\eta \in \mathbb{R}^{n^2+n}$ we obtain

$$\pi_{\lambda,\eta}(i, U, c_i) = \begin{cases} \arg\min_{j \in U}\left\{c_{ij} + \lambda_{j0} + \sum_{k \in U \setminus j}(\lambda_{jk} + \bar{c}_{jk}\eta_{jk})\right\}, & |U| \geq 2 \\ j, & U = j \in N \setminus i\ , \\ 0, & U = \varnothing \end{cases} \tag{13}$$

with ties again broken arbitrarily.

To analyze these policies for our problem, it is necessary to assume additional structure on the arc costs and the optimal expected cost-to-go. In particular, we assume in this section that the optimal expected cost-to-go is non-decreasing with respect to the set of remaining cities:

$$\mathsf{E}[y^*_{i,U}(C_i)] \leq \mathsf{E}[y^*_{i,U \cup j}(C_i)], \qquad \forall\ i \in N, j \in N \setminus i, U \subseteq N \setminus \{i,j\} \tag{14a}$$

$$\mathsf{E}[y^*_{i,N \setminus i}(C_i)] \leq \mathsf{E}[y^*_{0,N}(C_0)], \qquad \forall\ i \in N. \tag{14b}$$

This assumption is natural in most real-world situations; the more cities the salesman has left to visit, the higher the cost he should expect to incur. In the deterministic case, (14a) holds provided costs are non-negative and satisfy the triangle inequality. However, in our stochastic model we cannot assume the triangle inequality always holds without violating the independence of arc costs with different tails.

Our first approximation result for price-directed policies is generic. While all of our approximations are pointwise lower bounds on the optimal expected cost-to-go, this theorem requires only that the approximation be a lower bound in expectation.

Theorem 4.1. *Let $\tilde{y}_{i,U} : \mathcal{C}_i \to \mathbb{R}$ be an approximate cost-to-go function for each $i \in N \cup 0$, $U \subseteq N \setminus i$ satisfying $\mathsf{E}[\tilde{y}_{i,U}(C_i)] \leq \mathsf{E}[y^*_{i,U}(C_i)]$. Assume* (14) *holds, and suppose in addition that $\mathsf{E}[y^*_{i,U}(C_i)] \geq 0$ and*

$$\alpha \mathsf{E}[y^*_{i,U}(C_i)] \leq \mathsf{E}[\tilde{y}_{i,U}(C_i)], \qquad \forall\ i \in N \cup 0, U \subseteq N \setminus i, \tag{15}$$

for some $\alpha \in (0,1]$*. Then the expected cost of using the price-directed policy with approximation* $\tilde{y}$ *is bounded above by* $(1+(1-\alpha)n)\mathsf{E}[y^*_{0,N}(C_0)]$.

As with other ALP performance guarantees on price-directed policies, e.g. [26, Theorem 1] or [29, Theorem 3], when $\tilde{y} = y^*$ we recover the optimal cost, and the performance guarantee decreases with the approximation quality. Unlike those results, however, our performance guarantee depends on a multiplicative factor α, which is more common in approximation algorithms; see e.g. [42].

Proof. Consider a state (i, U, c_i), and suppose the policy chooses $j \in U$, whereas $k \in U$ is an optimal choice. Then

$$c_{ij} + \mathsf{E}[\tilde{y}_{j,U\setminus j}(C_j)] \le c_{ik} + \mathsf{E}[\tilde{y}_{k,U\setminus k}(C_k)] \le c_{ik} + \mathsf{E}[y^*_{k,U\setminus k}(C_k)] = y^*_{i,U}(c_i),$$

and therefore

$$c_{ij} \le y^*_{i,U}(c_i) - \mathsf{E}[\tilde{y}_{j,U\setminus j}(C_j)] \le y^*_{i,U}(c_i) - \alpha\mathsf{E}[y^*_{j,U\setminus j}(C_j)].$$

Let $c \in \mathcal{C}_0 \times \cdots \times \mathcal{C}_n$ be a realization of all arc costs, and relabel the cities in the order that the price-directed policy visits them under this realization. Summing over traversed arcs, the total cost of the price-directed tour satisfies

$$\sum_{i=0}^{n-1} c_{i,i+1} + c_{n0} \le y^*_{0,N}(c_0) + \sum_{i\in N}\big(y^*_{i,\{i+1,\dots,n\}}(c_i) - \alpha\mathsf{E}[y^*_{i,\{i+1,\dots,n\}}(C_i)]\big).$$

Let $\mathcal{U}_i$ denote the random variable representing the set of remaining cities when i is visited under the price-directed policy. Taking the expectation on the previous inequality,

$$\mathsf{E}[\text{total cost}] \le \mathsf{E}[y^*_{0,N}(C_0)] + \sum_{i\in N}\mathsf{E}\big[y^*_{i,\mathcal{U}_i}(C_i) - \alpha\mathsf{E}[y^*_{i,\mathcal{U}_i}(C_i)]\big].$$

Using the pairwise independence of the different C_i random cost vectors, it follows that $\mathcal{U}_i$ and C_i are independent, and hence

$$\mathsf{E}[\text{total cost}] \le \mathsf{E}[y^*_{0,N}(C_0)] + (1-\alpha)\sum_{i\in N}\sum_{U\subseteq N\setminus i}\mathsf{P}(\mathcal{U}_i = U)\mathsf{E}[y^*_{i,U}(C_i)] \le (1+(1-\alpha)n)\mathsf{E}[y^*_{0,N}(C_0)],$$

where the last inequality follows from (14). □

Unfortunately, this result cannot be used directly with an optimal solution of (9). As the next example shows, even in the deterministic case it is possible for the ALP to provide a tight bound for $\mathsf{E}[y^*_{0,N}(C_0)]$ while still giving an arbitrarily bad approximation of the optimal expected cost-to-go in particular states.

Example 4.2. Consider a deterministic TSP instance with cities arrayed around a unit circle at equal intervals in order from 0 to n, and arc costs given by Euclidean distances; see Figure 1. The distance between consecutive cities is $d_n = 2\sin(\frac{\pi}{n+1})$, an optimal tour with cost $(n+1)d_n$ follows the cities in order from 0 to n and back to 0, and a tight optimal solution (λ^*, η^*) of (9) is $\lambda^*_{i0} = \lambda^*_{ij} = d_n$, $\lambda^*_0 = (n+1)d_n$ and $\eta^* = 0$. However, when n is odd the optimal cost-to-go from state $(\frac{n+1}{2}, \varnothing)$, i.e. when the salesman is at city $\frac{n+1}{2}$ and must only return to 0, is 2, while the approximate cost-to-go given by this dual solution is $\lambda^*_{(n+1)/2,0} = d_n$, which goes to zero as $n \to \infty$.

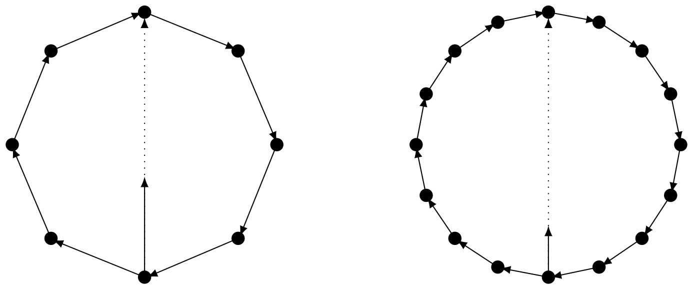

Figure 1: Instances from Example 4.2 with $n = 7$ and $n = 15$. The optimal tour travels around the circle, but if the salesman is in the bottom city and only needs to return to the top, the optimal solution of (9) under-approximates the cost-to-go: The true cost is the vertical dotted arrow, and the approximate cost is the overlying solid arrow.

In order to take advantage of Theorem 4.1, we must therefore recompute the bound at every step for every available action; this involves $O(n^2)$ total solutions of LP's of the form (9). If the ALP is solvable in polynomial time, we retain polynomial solvability, although the high computational burden may be unrealistic in some settings. The ideas of re-solving the ALP and "looking ahead" have previously been applied with empirical success, e.g. [7].

Theorem 4.3. *Suppose the following conditions hold:*

i) $\underline{c}_{ij} > 0$ *and* $\frac{\hat{c}_{ij}}{\underline{c}_{ij}} \leq \Gamma$, *for some* $\Gamma > 0$ *and all distinct* $i, j \in N \cup 0$.

ii) $\underline{c}$ *is* symmetric *and satisfies the triangle inequality; i.e.* $\underline{c}_{ij} = \underline{c}_{ji}$ *and* $\underline{c}_{ij} \leq \underline{c}_{ik} + \underline{c}_{kj}$, *for all distinct* $i, j, k \in N \cup 0$.

iii) Condition (14) *holds.*

Consider the following price-directed policy: At any encountered state (i, U, c_i) *with* $U \neq \varnothing$, *let* $\mathsf{E}[\tilde{y}_{j,U\setminus j}(C_j)]$ *for* $j \in U$ *be given by recomputing* (9) $|U|$ *times, using every remaining city* $j \in U$ *instead of* 0 *as start city and* $U \setminus j$ *as cities to visit. Then Theorem 4.1 applies with* $\alpha = \frac{5}{8\Gamma}$, *so that the total expected cost of using this policy is bounded above by* $\left(1 + \left(1 - \frac{5}{8\Gamma}\right)n\right)\mathsf{E}[y^*_{0,N}(C_0)]$.

For the conditions given in Theorem 4.3, if we apply Christofides' heuristic [24] using arc costs $\underline{c}$, we are guaranteed a $\frac{3\Gamma}{2}$-approximation of the optimal policy's expected cost. However, in contrast to the policies we introduce in this section, this solution is non-adaptive and may perform poorly in practice, as Example 2.3 illustrates.

Proof. From (i) we know that $\underline{y}_{j,U\setminus j} \geq \Gamma\mathsf{E}[y^*_{j,U\setminus j}(C_j)]$ for any $\varnothing \neq U \subseteq N$, $j \in U$. This number $\underline{y}_{j,U\setminus j}$ is the cost of a shortest Hamiltonian path starting at j, visiting cities $U \setminus j$ and ending at 0. By Theorem 3.4, $\mathsf{E}[\tilde{y}_{j,U\setminus j}(C_j)]$ provides a bound on $\underline{y}_{j,U\setminus j}$ at least as good as its LP relaxation. It remains to prove using (ii) that the LP relaxation gives a bound within a factor of 5/8 for $\underline{y}_{j,U\setminus j}$. This last component of the argument follows from [69]. □

There is a conjecture in the algorithms community that the LP relaxation of the symmetric shortest Hamiltonian path problem, also called the *s*-*t* TSP or TSP path, actually gives a bound within a factor of 2/3 [10, 69]. Any improvement in this bound guarantee would immediately imply a corresponding improved guarantee for Theorem 4.3's policy via Theorem 3.4.

5 Computational Experiments

We next discuss a series of computational experiments designed to test the efficacy of the bound given by (9) and the related price-directed policies. To generate test instances, we used deterministic asymmetric TSP instances from `TSPLIB` [66] and created two types of stochastic instance from each deterministic one. The instance set includes all `ftv` instances with 44 cities or fewer; though these instances appear small, [23] suggest that even instances with as few as ten or twelve cities are computationally difficult and practically relevant. However, the instance size is worth highlighting in particular because the performance of policies may depend on problem size. We carried out all experiments on a Dell workstation with dual Intel Xeon 3.2 GHz processor and 2 GB RDRAM, using CPLEX 9.0 as an LP solver.

The first instance type has independently distributed arc costs with two possible realizations, so that each support set $\mathcal{C}_i$ is composed of the vertices of a hyper-rectangle. Each arc is either *high*, where the deterministic cost is multiplied by a factor $H = 1 + \beta_H$, or *low*, where the deterministic cost is multiplied by a factor $L = 1 - \beta_L$. The experiment's two input parameters are β_H, the increment factor for high arc costs, and $\mathsf{P}(H)$, the probability of the arc cost being high. The probability of a low arc cost is of course the complementary probability, $\mathsf{P}(L) = 1 - \mathsf{P}(H)$, and we calculate β_L so that the arc's expected cost matches the deterministic instance's arc cost:

$$\beta_L = \frac{\beta_H \mathsf{P}(H)}{1 - \mathsf{P}(H)}.$$

For example, if high and low arc costs are equally likely, then clearly $\beta_H = \beta_L$. But if there is a 60% probability of a high cost, then $\beta_L = 3\beta_H/2$. By choosing the parameters in this way, the optimal expected cost of a fixed tour equals the deterministic instance's optimal cost, which we can use to benchmark our results.

The arcs in the second instance type also have two possible realizations, high or low. However, in this case a city's outgoing arc costs are all either high or low, simulating a high-traffic versus low-traffic possibility, so that the support set $\mathcal{C}_i$ is comprised of two points. The parameters are otherwise defined exactly as in the independent case.

For a given instance, recall that $\bar{y}_{0,N}$ denotes the optimal expected cost of a fixed tour, equal in this case to its optimal cost in the deterministic model and available from `TSPLIB`. By the definition of the arc cost distribution, we have $\underline{c}_{ij} = L\bar{c}_{ij}$ for every pair of cities i, j, and thus the best possible optimistic bound is $\underline{y}_{0,N} = L\bar{y}_{0,N}$. Our experiments compare the ALP bound (9) to this optimistic bound $\underline{y}_{0,N}$.

As an additional bound benchmark, we include an *a posteriori* bound [71], whose rationale is the following. Suppose the salesman had access to all arc costs at the start of the tour; then he could solve a deterministic TSP on whatever realization of arc costs he observes. The expected cost of an optimal TSP tour in this setting [55] is an anticipative version of our problem, and thus a lower bound. However, computing this expectation exactly is difficult, if not impossible, so the *a posteriori* procedure repeatedly samples a full realization of all arcs and solves a deterministic

TSP to optimality on this realization; the average of all realizations' deterministic optimal values is then an estimate of the bound. We note that, unlike the ALP bound, the *a posteriori* procedure solves many NP-hard optimization problems instead of a single polynomially-solvable LP, and has access through simulation to the full arc cost distribution, instead of requiring only expected costs and a description of the costs' support sets, as the ALP does. Therefore, the *a posteriori* bound should be considered a best-possible benchmark rather than a competing methodology.

The experiments then compare the optimal expected fixed tour cost $\bar{y}_{0,N}$ to a policy. In preliminary experiments, we used the price-directed policy given by the optimal solution of (9) within (13). However, this policy's performance was quite weak, confirming the indication from Example 4.2 that a fixed set of dual multipliers do not necessarily provide an accurate estimate of a state's cost-to-go, even when the bound they provide is tight. Instead, we designed a heuristic policy to test the recomputing idea presented in Theorem 4.3. At any encountered state (i, U, c_i) with $|U| \geq 2$, we use

$$\bar{\pi}^{\text{LP}}(i, U, c_i) := \arg\min_{j \in U} \{ c_{ij} + \bar{y}^{\text{LP}}_{j,U \setminus j} \},$$

where $\bar{y}^{\text{LP}}_{j,U \setminus j}$ is the optimal value of the LP relaxation of a j-0 shortest Hamiltonian path problem with deterministic costs $\bar{c}$, a formulation analogous to (11) with expected costs instead of optimistic costs. This approximation is not necessarily a lower bound, but is also computationally much simpler to solve than the ALP (9). We used 50 simulations of each instance's arc costs to estimate both the *a posteriori* bound and the heuristic policy's expected cost.

Our heuristic policy shares many traits with *rollout* policies [14, 70], in which a deterministic heuristic is used within a simulation framework to design a more sophisticated policy. In particular, our heuristic uses a lookahead at each city, estimating the cost-to-go of each possible next state (see also [7]), and also uses *certainty equivalence*, the idea that potential future states' cost-to-go can be estimated by replacing unknown parameters with their expectations. However, unlike rollout policies, our heuristic uses $\bar{y}^{\text{LP}}_{j,U \setminus j}$, a cost-to-go estimate derived from a relaxation.

Our experiments use every combination of $H \in \{1.05, 1.1, 1.15, 1.2, 1.25, 1.3\}$ and $\mathsf{P}(H) \in \{0.5, 0.6, 0.675, 0.75\}$ for every one of the four selected instances from `TSPLIB`, yielding 96 instances each for the two kinds of arc cost distributions. The motivation behind this choice of instance parameters is to ensure enough distance between $\underline{y}_{0,N}$ and $\bar{y}_{0,N}$.

Table 1 contains the results of our experiments for instance `ftv33` with independently distributed arc costs; the results for the remaining instances with independently distributed costs are included in the Appendix. Our results indicate the following relationship between each of the examined bounds and solution expected costs holds in virtually all cases:

$$\begin{aligned} \text{best optimistic bound, } \underline{y}_{0,N} \leq \text{ALP bound, } \lambda_0^* + \sum_{k \in N} \bar{c}_{0k} \eta_{0k}^* \leq \textit{a posteriori} \text{ bound} \\ \leq \text{heuristic policy} \leq \text{best fixed tour, } \bar{y}_{0,N}. \end{aligned} \tag{16}$$

Because the *a posteriori* bound is the tightest, in addition to reporting all quantities in absolutes, we report all other quantities as percentages of this bound. The gaps between successive quantities in (16) uniformly grow with H or $\mathsf{P}(H)$, i.e. either as the costs' support sets grow larger or as high costs become more likely.

Our results indicate that the ALP bound is always greater than the optimistic bound $\underline{y}_{0,N}$, the best possible bound achieved with deterministic costs. The relative difference between the two is increasing with respect to both H and $\mathsf{P}(H)$. In addition, the ALP is reasonably close to the *a*

Instance: ftv33						
H	1.05	1.10	1.15	1.2	1.25	1.30
$\mathsf{P}(H) = 0.5$						
Optimistic Bound	1221.70 (95.12%)	1157.40 (90.76%)	1093.10 (87.28%)	1028.80 (84.49%)	964.50 (82.02%)	900.20 (79.91%)
ALP Bound	1266.19 (98.59%)	1222.82 (95.89%)	1173.34 (93.69%)	1120.05 (91.98%)	1063.11 (90.41%)	1003.44 (89.08%)
A Posteriori Bound	**1284.35**	**1275.20**	**1252.36**	**1217.73**	**1175.90**	**1126.48**
Heuristic Policy Cost	**1284.85 (100.04%)**	**1282.17 (100.55%)**	**1274.07 (101.73%)**	**1269.47 (104.25%)**	**1258.63 (107.04%)**	**1239.55 (110.04%)**
Fixed Tour Cost	1286.00 (100.13%)	1286.00 (100.85%)	1286.00 (102.69%)	1286.00 (105.61%)	1286.00 (109.36%)	1286.00 (114.16%)
$\mathsf{P}(H) = 0.6$						
Optimistic Bound	1189.55 (92.53%)	1093.10 (86.21%)	996.65 (81.26%)	900.20 (77.13%)	803.75 (73.74%)	707.30 (70.86%)
ALP Bound	1255.72 (97.68%)	1193.63 (94.14%)	1119.86 (91.31%)	1036.81 (88.84%)	945.19 (86.72%)	847.46 (84.90%)
A Posteriori Bound	**1285.57**	**1267.96**	**1226.48**	**1167.06**	**1089.93**	**998.23**
Heuristic Policy Cost	**1286.72 (100.09%)**	**1283.71 (101.24%)**	**1276.75 (104.10%)**	**1269.34 (108.76%)**	**1242.72 (114.02%)**	**1213.57 (121.57%)**
Fixed Tour Cost	1286.00 (100.03%)	1286.00 (101.42%)	1286.00 (104.85%)	1286.00 (110.19%)	1286.00 (117.99%)	1286.00 (128.83%)
$\mathsf{P}(H) = 0.675$						
Optimistic Bound	1152.45 (89.86%)	1018.91 (81.33%)	885.36 (74.61%)	751.82 (69.23%)	618.27 (64.89%)	484.72 (61.51%)
ALP Bound	1246.28 (97.17%)	1160.22 (92.61%)	1057.06 (89.08%)	928.39 (85.49%)	787.13 (82.61%)	632.08 (80.21%)
A Posteriori Bound	**1282.54**	**1252.83**	**1186.60**	**1085.92**	**952.81**	**788.04**
Heuristic Policy Cost	**1284.43 (100.15%)**	**1277.00 (101.93%)**	**1267.91 (106.85%)**	**1239.88 (114.18%)**	**1184.30 (124.30%)**	**1115.72 (141.58%)**
Fixed Tour Cost	1286.00 (100.27%)	1286.00 (102.65%)	1286.00 (108.38%)	1286.00 (118.42%)	1286.00 (134.97%)	1286.00 (163.19%)
$\mathsf{P}(H) = 0.75$						
Optimistic Bound	1093.10 (85.76%)	900.20 (74.16%)	707.30 (65.11%)	514.40 (57.74%)	321.50 (52.48%)	128.60 (50.46%)
ALP Bound	1230.60 (96.54%)	1107.64 (91.25%)	932.65 (85.86%)	716.83 (80.46%)	462.33 (75.47%)	185.35 (72.72%)
A Posteriori Bound	**1274.67**	**1213.88**	**1086.25**	**890.92**	**612.62**	**254.87**
Heuristic Policy Cost	**1278.12 (100.27%)**	**1265.30 (104.24%)**	**1234.38 (113.64%)**	**1170.34 (131.36%)**	**1042.11 (170.11%)**	**742.13 (291.17%)**
Fixed Tour Cost	1286.00 (100.89%)	1286.00 (105.94%)	1286.00 (118.39%)	1286.00 (144.35%)	1286.00 (209.92%)	1286.00 (504.56%)

Table 1: Experiment results for `ftv33` ($n = 33$) with independently distributed costs.

posteriori benchmark – within approximately 80% – in all but the most extreme cases ($\mathsf{P}(H) = 0.75$ and $H \in \{1.25, 1.3\}$). We observe a similar pattern for all other instances with independently distributed costs. The heuristic policy also performs quite well, consistently outperforming the best fixed tour and lying within approximately 20% of the *a posteriori* benchmark except in extreme cases (particularly when H is large). However, in the extreme cases the policy significantly outperforms the fixed tour. For instance, for the most extreme case we tested where $\mathsf{P}(H) = 0.75$ and $H = 1.3$, the policy is about half as expensive, 57% of the fixed tour's expected cost. Similar results hold for the other instances we tested with independently distributed costs.

Table 2 has results for `ftv33` with high/low correlated arc costs, with remaining results in the Appendix. We do not include results for `ftv44` because we were not able to solve the ALP for several of the parameter values within 24 hours. The results for this instance type are similar to the independent case, with some key differences. The gaps between the *a posteriori* bound and the other two bounds seem to depend more heavily on the parameters, with smaller gaps for smaller parameters (i.e. results closer to the top-left corner) and larger gaps for the more extreme parameter settings (i.e. the bottom-right corner). However, the gaps between the *a posteriori* bound and the two solutions, particularly the heuristic policy, are significantly smaller. For `ftv33` the heuristic policy is within 9% of optimality in all parameter settings and within 2% in all but the most extreme cases; similar results hold for the other instances. The gap between the heuristic and the bound is also consistently about half as large (or less) as the gap between the fixed tour and the bound.

In both cases we can ask whether the bound or the policy are responsible for the remaining gap. We suspect the bound is farther away from the optimal expected cost than the heuristic policy, especially in the extreme cases. Specifically, Proposition 2.4 suggests $\bar{y}^{\text{LP}}_{j,U\setminus j}$ is a reasonable proxy for each action's future cost-to-go $\mathsf{E}[y^*_{j,U\setminus j}(C_j)]$, because computationally it is a close approximation of $\bar{y}_{j,U\setminus j}$. Thus the heuristic policy may closely mimic an optimal policy. As for the ALP bound, other authors have observed similar bound gaps when using affine cost-to-go approximations computationally, e.g. [7]. The design of more sophisticated approximations that maintain polynomial-time solvability in (7) is thus an interesting and important research question, particularly in the presence of high variability, as exemplified by the more extreme parameter settings in the experiments where

Instance: ftv33						
H	1.05	1.10	1.15	1.2	1.25	1.30
$\mathsf{P}(H) = 0.5$						
Optimistic Bound	1221.70 (95.51%)	1157.40 (90.97%)	1093.10 (86.39%)	1028.80 (81.75%)	964.50 (77.09%)	900.20 (72.49%)
ALP Bound	1279.12 (100.00%)	1259.98 (99.04%)	1233.54 (97.48%)	1203.24 (95.62%)	1171.33 (93.63%)	1138.21 (91.66%)
A Posteriori Bound	**1279.12**	**1272.25**	**1265.37**	**1258.42**	**1251.09**	**1241.80**
Heuristic Policy Cost	**1279.12** (**100.00**%)	**1272.25** (**100.00**%)	**1265.37** (**100.00**%)	**1258.50** (**100.01**%)	**1251.62** (**100.04**%)	**1244.74** (**100.24**%)
Fixed Tour Cost	1286.00 (100.54%)	1286.00 (101.08%)	1286.00 (101.63%)	1286.00 (102.19%)	1286.00 (102.79%)	1286.00 (103.56%)
$\mathsf{P}(H) = 0.6$						
Optimistic Bound	1189.55 (93.14%)	1093.10 (86.19%)	996.65 (80.72%)	900.20 (72.06%)	803.75 (65.09%)	707.30 (58.14%)
ALP Bound	1274.37 (99.78%)	1243.42 (98.04%)	1203.07 (97.44%)	1158.01 (92.70%)	1108.08 (89.74%)	1053.48 (86.59%)
A Posteriori Bound	**1277.13**	**1268.26**	**1234.7375**	**1249.22**	**1234.74**	**1216.56**
Heuristic Policy Cost	**1277.13** (**100.00**%)	**1268.26** (**100.00**%)	**1259.39** (**102.00**%)	**1250.52** (**100.10**%)	**1241.65** (**100.56**%)	**1234.66** (**101.49**%)
Fixed Tour Cost	1286.00 (100.69%)	1286.00 (101.40%)	1286.00 (104.15%)	1286.00 (102.94%)	1286.00 (104.15%)	1286.00 (105.71%)
$\mathsf{P}(H) = 0.675$						
Optimistic Bound	1152.45 (90.27%)	1018.91 (80.39%)	885.36 (70.39%)	751.82 (60.60%)	618.27 (50.79%)	484.72 (40.83%)
ALP Bound	1267.82 (99.30%)	1224.09 (96.58%)	1169.13 (92.95%)	1102.11 (88.84%)	1019.07 (83.71%)	907.50 (76.44%)
A Posteriori Bound	**1276.74**	**1267.48**	**1257.74**	**1240.59**	**1217.41**	**1187.19**
Heuristic Policy Cost	**1276.74** (**100.00**%)	**1267.48** (**100.00**%)	**1258.22** (**100.04**%)	**1248.96** (**100.67**%)	**1241.08** (**101.94**%)	**1226.79** (**103.34**%)
Fixed Tour Cost	1286.00 (100.73%)	1286.00 (101.46%)	1286.00 (102.25%)	1286.00 (103.66%)	1286.00 (105.63%)	1286.00 (108.32%)
$\mathsf{P}(H) = 0.75$						
Optimistic Bound	1093.10 (85.73%)	900.20 (71.21%)	707.30 (56.78%)	514.40 (42.35%)	321.50 (27.48%)	128.60 (11.54%)
ALP Bound	1258.09 (98.67%)	1195.09 (94.54%)	1105.75 (88.76%)	968.11 (79.70%)	772.77 (66.05%)	489.57 (43.94%)
A Posteriori Bound	**1275.05**	**1264.09**	**1245.78**	**1214.68**	**1170.04**	**1114.28**
Heuristic Policy Cost	**1275.05** (**100.00**%)	**1264.10** (**100.00**%)	**1253.80** (**100.64**%)	**1237.74** (**101.90**%)	**1224.35** (**104.64**%)	**1207.28** (**108.35**%)
Fixed Tour Cost	1286.00 (100.86%)	1286.00 (101.73%)	1286.00 (103.23%)	1286.00 (105.87%)	1286.00 (109.91%)	1286.00 (115.41%)

Table 2: Experiment results for `ftv33` ($n = 33$) with high/low correlated costs.

the gaps are quite large.

It is worth noting that the procedures we outline in this section are computationally intensive, as evidenced by our difficulty with the ALP bound for the `ftv44` instances with correlated costs. As a sample, in the Appendix we include two tables outlining experiment running times for the `ftv33` instances. The tables contain total running times required to compute the ALP bound, as well as average time per simulated sample to compute both the *a posteriori* bound and the heuristic policy. (Recall that the optimistic bound and the fixed tour cost are obtained directly from `TSPLIB`.) In the case of the *a posteriori* bound, the time corresponds to solving a single TSP with the sampled arc costs; for the policy, it corresponds to an entire run of the policy for the same sampled arc costs, i.e. $O(n^2)$ LP solves. For instances with independently distributed costs, we observe ALP solution times on the order of one hour, *a posteriori* averages of one to two minutes, and policy averages on the order of three to five minutes for most parameter settings. The latter two averages are similar for instances with correlated costs; however, ALP solution times were much higher for these instances, with running times stretching to many hours. This difference surprised us, since the independently distributed cost instances have many more constraints than the correlated cost instances with all other things being equal: For a given pair (i, U) the former instance class has $2^{|U|}$ constraints and the latter only two. However, this may indicate an additional challenge for this problem and others like it when high correlation among uncertain parameters is present.

6 Conclusions

We have presented a dynamic TSP model with stochastic arc costs and applied ALP to tractably bound the problem, construct policies with theoretical worst-case performance guarantees, and derive high-quality heuristics. Our work leads to several questions. For example, we do not address how to obtain a bound when the separation problem for (9) is NP-hard. As we indicate, constraint sampling may be an option [26, 27, 29], though it suffers from requiring idealized access to the TSP's optimal policy to sample constraints. Using a heuristic policy instead could suffice for practical purposes. Another option in this case is exact separation, which has been applied successfully in other contexts, e.g. [2, 7].

Another issue concerns the design of better bases to more accurately approximate the cost-to-go. This is a challenging question since any basis must trade off approximation fidelity with computational tractability, and the trade-off is not always obvious. For the TSP, the results in [75] suggest that only small improvements are possible when considering the combinatorial aspect of the state space, but whether the same is true for the cost component is unclear.

More generally, we hope our work verifies ALP as an effective methodology for dynamic routing models under uncertainty, though it is clear that many challenges remain. These models are some of the most challenging problems currently studied in transportation and discrete optimization, and the ALP approach offers an approach to tractably model, bound and solve them.

Appendix

Remaining Proofs

Proof of Lemma 3.2. In problem (9), the semi-infinite constraint system's index set ranges over a compact space, and all variable coefficients and the right-hand side are continuous functions of the index. The system also has an easily constructed Slater point: $\lambda_0 = \lambda_{ij} = -M$, for large enough $M > 0$, and all other variables set to zero. By [34, Theorem 5.3], the semi-infinite constraint system is Farkas-Minkowski, and thus the Haar dual (10) is a strong dual with an optimal solution [34, Theorem 8.4]. □

Proof of Lemma 3.6. Fix $i \in N$, $j \in N \setminus i$, λ and η, and suppose $\mathcal{C}_i = \{c \in \mathbb{R}^n : c = \gamma + v, \|v\|_2 \leq 1\}$; i.e. $\mathcal{C}_i$ is an ℓ_2 unit ball centered at $\gamma \in \mathbb{R}^n$. Using

$$\max_{c_i \in \mathcal{C}_i} c_{ij}(\eta_{ij} - 1) + \sum_{k \in U} c_{ik}\eta_{ik} = \gamma_{ij}(\eta_{ij} - 1) + \sum_{k \in U} \gamma_{ik}\eta_{ik} + \sqrt{(\eta_{ij} - 1)^2 + \sum_{k \in U} \eta_{ik}^2},$$

problem (12) is equivalent to

$$\max_{\varnothing \neq U \subseteq N \setminus \{i,j\}} \sqrt{a_j + \sum_{k \in U} a_k} + \sum_{k \in U} b_k, \tag{17}$$

for appropriately chosen $a_j, a_k \geq 0$ and b_k. This is a special case of a *submodular utility maximization problem* [9]; following Proposition 1 of this reference, we show (17) is NP-hard by a reduction from the *partition problem* [31]. Given a collection of numbers $a_k > 0$, *partition* asks whether there exists a set U with $\sum_{k \in U} a_k = \sum_{k \notin U} a_k$. Let $a_j = 0$, and by rescaling if necessary, assume $\sum a_k = 2$; then setting $b_k = -\frac{1}{2}a_k$, a partition exists if and only if the optimal value of (17) is $\frac{1}{2}$. □

Remaining Experiment Results

This section contains the remaining results for experiments outlined in Section 5. Tables 3 through 5 have results for instances `ftv35`, `ftv38` and `ftv44` with independently distributed costs. Tables 6 and 7 have results for instances `ftv35` and `ftv38` with high/low correlated costs. Tables 8 and 9 give running times for the `ftv33` instances.

Instance: ftv35						
H	1.05	1.10	1.15	1.2	1.25	1.30
$\mathsf{P}(H) = 0.5$						
Optimistic Bound	1399.35 (95.72%)	1325.70 (91.95%)	1252.05 (88.74%)	1178.40 (86.03%)	1104.75 (83.73%)	1031.10 (81.76%)
ALP Bound	1425.15 (97.48%)	1378.83 (95.64%)	1325.46 (93.94%)	1263.20 (92.22%)	1194.44 (90.53%)	1122.58 (89.02%)
A Posteriori Bound	**1461.94**	**1441.69**	**1410.98**	**1369.72**	**1319.35**	**1261.11**
Heuristic Policy Cost	**1470.69 (100.60%)**	**1463.15 (101.49%)**	**1456.03 (103.19%)**	**1451.35 (105.96%)**	**1427.15 (108.17%)**	**1415.95 (112.28%)**
Fixed Tour Cost	1473.00 (100.76%)	1473.00 (102.17%)	1473.00 (104.40%)	1473.00 (107.54%)	1473.00 (111.65%)	1473.00 (116.80%)
$\mathsf{P}(H) = 0.6$						
Optimistic Bound	1362.53 (93.44%)	1252.05 (87.64%)	1141.58 (82.67%)	1031.10 (78.72%)	920.63 (74.86%)	810.15 (73.15%)
ALP Bound	1414.33 (96.99%)	1345.68 (94.19%)	1262.01 (91.39%)	1162.56 (88.76%)	1052.29 (85.56%)	934.76 (84.40%)
A Posteriori Bound	**1458.16**	**1428.61**	**1380.93**	**1309.84**	**1229.83**	**1107.59**
Heuristic Policy Cost	**1470.85 (100.87%)**	**1461.08 (102.27%)**	**1449.94 (105.00%)**	**1426.23 (108.89%)**	**1416.45 (115.17%)**	**1351.20 (121.99%)**
Fixed Tour Cost	1473.00 (101.02%)	1473.00 (103.11%)	1473.00 (106.67%)	1473.00 (112.46%)	1473.00 (119.77%)	1473.00 (132.99%)
$\mathsf{P}(H) = 0.675$						
Optimistic Bound	1320.03 (90.74%)	1167.07 (82.66%)	1014.10 (76.06%)	861.14 (70.95%)	708.17 (66.99%)	555.21 (63.91%)
ALP Bound	1402.88 (96.43%)	1310.57 (92.82%)	1187.00 (89.03%)	1032.84 (85.10%)	864.19 (81.74%)	687.51 (79.14%)
A Posteriori Bound	**1454.78**	**1411.89**	**1333.27**	**1213.72**	**1057.18**	**868.76**
Heuristic Policy Cost	**1468.67 (100.95%)**	**1457.82 (103.25%)**	**1439.96 (108.00%)**	**1406.56 (115.89%)**	**1346.17 (127.34%)**	**1259.09 (144.93%)**
Fixed Tour Cost	1473.00 (101.25%)	1473.00 (104.33%)	1473.00 (110.48%)	1473.00 (121.36%)	1473.00 (139.33%)	1473.00 (169.55%)
$\mathsf{P}(H) = 0.75$						
Optimistic Bound	1252.05 (86.57%)	1031.10 (75.14%)	810.15 (66.31%)	589.20 (59.43%)	368.25 (54.92%)	147.30 (52.75%)
ALP Bound	1388.88 (96.03%)	1249.72 (91.07%)	1036.51 (84.83%)	780.51 (78.73%)	498.32 (74.32%)	201.04 (72.00%)
A Posteriori Bound	**1446.31**	**1372.23**	**1221.81**	**991.42**	**670.48**	**279.22**
Heuristic Policy Cost	**1465.39 (101.32%)**	**1453.57 (105.93%)**	**1411.82 (115.55%)**	**1315.46 (132.68%)**	**1131.87 (168.81%)**	**827.07 (296.21%)**
Fixed Tour Cost	1473.00 (101.85%)	1473.00 (107.34%)	1473.00 (120.56%)	1473.00 (148.58%)	1473.00 (219.69%)	1473.00 (527.54%)

Table 3: Experiment results for `ftv35` ($n = 35$) with independently distributed costs.

Instance: ftv38						
H	1.05	1.10	1.15	1.2	1.25	1.30
$\mathsf{P}(H) = 0.5$						
Optimistic Bound	1453.50 (95.65%)	1377.00 (91.94%)	1300.50 (88.76%)	1224.00 (86.04%)	1147.50 (83.77%)	1071.00 (81.90%)
ALP Bound	1477.39 (97.22%)	1426.16 (95.22%)	1363.75 (93.08%)	1295.80 (91.09%)	1225.92 (89.50%)	1152.47 (88.13%)
A Posteriori Bound	**1519.67**	**1497.78**	**1465.15**	**1422.60**	**1369.76**	**1307.63**
Heuristic Policy Cost	**1526.46 (100.45%)**	**1522.46 (101.65%)**	**1510.92 (103.12%)**	**1494.98 (105.09%)**	**1477.46 (107.86%)**	**1441.38 (110.23%)**
Fixed Tour Cost	1530.00 (100.68%)	1530.00 (102.15%)	1530.00 (104.43%)	1530.00 (107.55%)	1530.00 (111.70%)	1530.00 (117.01%)
$\mathsf{P}(H) = 0.6$						
Optimistic Bound	1415.25 (93.42%)	1300.50 (87.73%)	1185.75 (82.92%)	1071.00 (79.02%)	956.25 (75.90%)	841.50 (73.45%)
ALP Bound	1466.04 (96.78%)	1389.87 (93.76%)	1299.12 (90.85%)	1194.60 (88.14%)	1080.20 (85.74%)	959.50 (83.75%)
A Posteriori Bound	**1514.894**	**1482.37**	**1429.93**	**1355.38**	**1259.83**	**1145.67**
Heuristic Policy Cost	**1524.69 (100.65%)**	**1524.49 (102.84%)**	**1510.47 (105.63%)**	**1489.46 (109.89%)**	**1432.64 (113.72%)**	**1393.09 (121.60%)**
Fixed Tour Cost	1530.00 (101.00%)	1530.00 (103.21%)	1530.00 (107.00%)	1530.00 (112.88%)	1530.00 (121.45%)	1530.00 (133.55%)
$\mathsf{P}(H) = 0.675$						
Optimistic Bound	1371.12 (90.65%)	1212.23 (82.75%)	1053.35 (76.09%)	894.46 (71.03%)	735.58 (67.06%)	576.69 (64.20%)
ALP Bound	1454.12 (96.13%)	1350.42 (92.18%)	1221.03 (88.21%)	1059.72 (84.15%)	886.83 (80.85%)	705.43 (78.53%)
A Posteriori Bound	**1512.61**	**1465.00**	**1384.27**	**1259.30**	**1096.94**	**898.25**
Heuristic Policy Cost	**1527.32 (100.97%)**	**1519.09 (103.69%)**	**1493.90 (107.92%)**	**1455.99 (115.62%)**	**1378.30 (125.65%)**	**1265.95 (140.93%)**
Fixed Tour Cost	1530.00 (101.15%)	1530.00 (104.44%)	1530.00 (110.53%)	1530.00 (121.50%)	1530.00 (139.48%)	1530.00 (170.33%)
$\mathsf{P}(H) = 0.75$						
Optimistic Bound	1300.50 (86.46%)	1071.00 (74.77%)	841.50 (65.71%)	612.00 (59.38%)	382.50 (55.32%)	153.00 (53.60%)
ALP Bound	1437.29 (95.56%)	1286.90 (89.84%)	1064.47 (83.12%)	800.80 (77.69%)	511.68 (74.00%)	205.98 (72.16%)
A Posteriori Bound	**1504.11**	**1432.42**	**1280.70**	**1030.71**	**691.46**	**285.44**
Heuristic Policy Cost	**1526.59 (101.49%)**	**1507.67 (105.25%)**	**1458.24 (113.86%)**	**1354.56 (131.42%)**	**1163.57 (168.28%)**	**797.46 (279.38%)**
Fixed Tour Cost	1530.00 (101.72%)	1530.00 (106.81%)	1530.00 (119.47%)	1530.00 (148.44%)	1530.00 (221.27%)	1530.00 (536.02%)

Table 4: Experiment results for `ftv38` ($n = 38$) with independently distributed costs.

Instance: ftv44						
H	1.05	1.10	1.15	1.2	1.25	1.30
$\mathsf{P}(H) = 0.5$						
Optimistic Bound	1532.35 (95.42%)	1451.70 (91.39%)	1371.05 (87.78%)	1290.40 (84.68%)	1209.75 (82.06%)	1129.10 (79.88%)
ALP Bound	1550.82 (96.57%)	1503.80 (94.67%)	1442.55 (92.35%)	1376.98 (90.36%)	1307.02 (88.66%)	1230.08 (87.02%)
A Posteriori Bound	**1605.83**	**1588.52**	**1561.97**	**1523.85**	**1474.21**	**1413.48**
Heuristic Policy Cost	1634.46 (101.78%)	1635.27 (102.94%)	**1619.47 (103.68%)**	**1609.62 (105.63%)**	**1589.83 (107.84%)**	**1557.18 (110.17%)**
Fixed Tour Cost	**1631.00 (101.57%)**	**1631.00 (102.67%)**	1631.00 (104.42%)	1631.00 (107.03%)	1631.00 (110.64%)	1631.00 (115.39%)
$\mathsf{P}(H) = 0.6$						
Optimistic Bound	1492.03 (93.13%)	1371.05 (87.13%)	1250.08 (81.97%)	1129.10 (77.47%)	1008.13 (73.94%)	887.15 (71.14%)
ALP Bound	1542.13 (96.26%)	1471.02 (93.49%)	1380.95 (90.55%)	1278.63 (87.73%)	1159.59 (85.05%)	1032.47 (82.80%)
A Posteriori Bound	**1602.06**	**1573.50**	**1525.05**	**1457.41**	**1363.35**	**1246.97**
Heuristic Policy Cost	1635.20 (102.07%)	**1625.96 (103.33%)**	**1606.05 (105.31%)**	**1570.25 (107.74%)**	**1531.64 (112.34%)**	**1485.02 (119.09%)**
Fixed Tour Cost	**1631.00 (101.81%)**	1631.00 (103.65%)	1631.00 (106.95%)	1631.00 (111.91%)	1631.00 (119.63%)	1631.00 (130.80%)
$\mathsf{P}(H) = 0.675$						
Optimistic Bound	1445.50 (90.45%)	1277.99 (82.05%)	1110.49 (75.07%)	942.98 (69.46%)	775.48 (65.27%)	607.98 (62.28%)
ALP Bound	1533.22 (95.94%)	1434.97 (92.13%)	1307.76 (88.40%)	1145.77 (84.40%)	961.63 (80.94%)	764.14 (78.27%)
A Posteriori Bound	**1598.05**	**1557.49**	**1479.33**	**1357.50**	**1188.04**	**976.25**
Heuristic Policy Cost	1636.57 (102.41%)	**1624.08 (104.28%)**	**1596.03 (107.89%)**	**1554.67 (114.52%)**	**1485.94 (125.07%)**	**1371.46 (140.48%)**
Fixed Tour Cost	**1631.00 (102.06%)**	1631.00 (104.72%)	1631.00 (110.25%)	1631.00 (120.15%)	1631.00 (137.28%)	1631.00 (167.07%)
$\mathsf{P}(H) = 0.75$						
Optimistic Bound	1371.05 (86.10%)	1129.10 (74.17%)	887.15 (64.73%)	645.20 (57.96%)	403.25 (53.67%)	161.30 (51.86%)
ALP Bound	1517.40 (95.29%)	1375.90 (90.38%)	1152.58 (84.10%)	870.12 (78.17%)	554.52 (73.80%)	223.20 (71.76%)
A Posteriori Bound	**1592.44**	**1522.33**	**1370.45**	**1113.16**	**751.39**	**311.01**
Heuristic Policy Cost	1637.68 (102.84%)	**1614.17 (106.03%)**	**1565.60 (114.24%)**	**1447.39 (130.03%)**	**1281.98 (170.62%)**	**923.55 (296.95%)**
Fixed Tour Cost	**1631.00 (102.42%)**	1631.00 (107.14%)	1631.00 (119.01%)	1631.00 (146.52%)	1631.00 (217.07%)	1631.00 (524.41%)

Table 5: Experiment results for `ftv44` ($n = 44$) with independently distributed costs.

Instance: ftv35						
H	1.05	1.10	1.15	1.2	1.25	1.30
$\mathsf{P}(H) = 0.5$						
Optimistic Bound	1399.35 (95.34%)	1325.70 (90.72%)	1252.05 (86.10%)	1178.40 (81.48%)	1104.75 (75.82%)	1031.10 (72.26%)
ALP Bound	1441.72 (98.22%)	1419.30 (97.13%)	1394.35 (95.89%)	1366.42 (94.48%)	1333.07 (91.49%)	1294.93 (90.74%)
A Posteriori Bound	**1467.82**	**1461.29**	**1454.13**	**1446.23**	**1457.06**	**1427.00**
Heuristic Policy Cost	**1473.80 (100.41%)**	**1466.46 (100.35%)**	**1461.63 (100.52%)**	**1456.80 (100.73%)**	**1472.86 (101.08%)**	**1447.23 (101.42%)**
Fixed Tour Cost	1473.00 (100.35%)	1473.00 (100.80%)	1473.00 (101.30%)	1473.00 (101.85%)	1473.00 (101.09%)	1473.00 (103.22%)
$\mathsf{P}(H) = 0.6$						
Optimistic Bound	1362.53 (92.94%)	1252.05 (85.92%)	1141.58 (78.92%)	1031.10 (71.90%)	920.63 (64.84%)	810.15 (57.70%)
ALP Bound	1435.05 (97.89%)	1403.20 (96.29%)	1364.31 (94.32%)	1315.44 (91.73%)	1253.42 (88.28%)	1182.27 (84.21%)
A Posteriori Bound	**1465.99**	**1457.25**	**1446.48**	**1434.09**	**1419.76**	**1404.02**
Heuristic Policy Cost	**1470.45 (100.30%)**	**1464.46 (100.50%)**	**1458.48 (100.83%)**	**1452.49 (101.28%)**	**1446.50 (101.88%)**	**1441.36 (102.66%)**
Fixed Tour Cost	1473.00 (100.48%)	1473.00 (101.08%)	1473.00 (101.83%)	1473.00 (102.71%)	1473.00 (103.75%)	1473.00 (104.91%)
$\mathsf{P}(H) = 0.675$						
Optimistic Bound	1320.03 (90.17%)	1167.07 (80.35%)	1014.10 (70.54%)	861.14 (60.67%)	708.17 (50.65%)	555.21 (40.44%)
ALP Bound	1428.30 (97.56%)	1384.12 (95.30%)	1325.64 (92.21%)	1241.04 (87.44%)	1137.29 (81.33%)	1009.28 (73.51%)
A Posteriori Bound	**1463.98**	**1452.41**	**1437.63**	**1419.36**	**1398.31**	**1372.90**
Heuristic Policy Cost	**1469.01 (100.34%)**	**1460.94 (100.59%)**	**1452.87 (101.06%)**	**1444.80 (101.79%)**	**1436.73 (102.75%)**	**1427.80 (104.00%)**
Fixed Tour Cost	1473.00 (100.62%)	1473.00 (101.42%)	1473.00 (102.46%)	1473.00 (103.78%)	1473.00 (105.34%)	1473.00 (107.29%)
$\mathsf{P}(H) = 0.75$						
Optimistic Bound	1252.05 (85.78%)	1031.10 (71.52%)	810.15 (57.15%)	589.20 (42.45%)	368.25 (27.28%)	147.30 (11.32%)
ALP Bound	1418.38 (97.18%)	1352.28 (93.80%)	1240.80 (87.53%)	1080.29 (77.83%)	859.99 (63.70%)	539.08 (41.44%)
A Posteriori Bound	**1459.55**	**1441.67**	**1417.64**	**1388.10**	**1350.13**	**1300.72**
Heuristic Policy Cost	**1465.81 (100.43%)**	**1454.30 (100.88%)**	**1442.79 (101.77%)**	**1431.28 (103.11%)**	**1417.01 (104.95%)**	**1391.31 (106.96%)**
Fixed Tour Cost	1473.00 (100.92%)	1473.00 (102.17%)	1473.00 (103.91%)	1473.00 (106.12%)	1473.00 (109.10%)	1473.00 (113.24%)

Table 6: Experiment results for `ftv35` ($n = 35$) with high/low correlated costs.

Instance: ftv38						
H	1.05	1.10	1.15	1.2	1.25	1.30
$\mathsf{P}(H) = 0.5$						
Optimistic Bound	1453.50 (95.34%)	1377.00 (90.70%)	1300.50 (86.06%)	1224.00 (81.42%)	1147.50 (76.78%)	1071.00 (72.16%)
ALP Bound	1499.04 (98.32%)	1473.81 (97.08%)	1445.99 (95.69%)	1415.49 (94.16%)	1379.98 (92.34%)	1340.45 (90.31%)
A Posteriori Bound	**1524.61**	**1518.17**	**1511.08**	**1503.35**	**1494.52**	**1484.21**
Heuristic Policy Cost	**1526.60 (100.13%)**	**1521.20 (100.20%)**	**1515.81 (100.31%)**	**1509.53 (100.41%)**	**1503.71 (100.62%)**	**1497.89 (100.92%)**
Fixed Tour Cost	1530.00 (100.35%)	1530.00 (100.78%)	1530.00 (101.25%)	1530.00 (101.77%)	1530.00 (102.37%)	1530.00 (103.09%)
$\mathsf{P}(H) = 0.6$						
Optimistic Bound	1415.25 (92.88%)	1300.50 (85.78%)	1185.75 (78.68%)	1071.00 (71.58%)	956.25 (64.47%)	841.50 (57.31%)
ALP Bound	1491.37 (97.88%)	1455.65 (96.02%)	1413.24 (93.77%)	1361.93 (91.02%)	1300.43 (87.67%)	1227.81 (83.62%)
A Posteriori Bound	**1523.70**	**1516.04**	**1507.12**	**1496.32**	**1483.26**	**1468.32**
Heuristic Policy Cost	**1525.70 (100.13%)**	**1519.40 (100.22%)**	**1513.10 (100.40%)**	**1508.64 (100.82%)**	**1502.40 (101.29%)**	**1496.16 (101.90%)**
Fixed Tour Cost	1530.00 (100.41%)	1530.00 (100.92%)	1530.00 (101.52%)	1530.00 (102.25%)	1530.00 (103.15%)	1530.00 (104.20%)
$\mathsf{P}(H) = 0.675$						
Optimistic Bound	1371.12 (90.12%)	1212.23 (80.23%)	1053.35 (70.33%)	894.46 (60.37%)	735.58 (50.33%)	576.69 (40.17%)
ALP Bound	1483.36 (97.50%)	1434.95 (94.97%)	1372.87 (91.66%)	1289.13 (87.01%)	1183.59 (80.98%)	1050.56 (73.17%)
A Posteriori Bound	**1521.36**	**1510.96**	**1497.80**	**1481.52**	**1461.53**	**1435.70**
Heuristic Policy Cost	**1524.18 (100.19%)**	**1516.37 (100.36%)**	**1508.55 (100.72%)**	**1501.92 (101.38%)**	**1493.90 (102.22%)**	**1486.50 (103.54%)**
Fixed Tour Cost	1530.00 (100.57%)	1530.00 (101.26%)	1530.00 (102.15%)	1530.00 (103.27%)	1530.00 (104.69%)	1530.00 (106.57%)
$\mathsf{P}(H) = 0.75$						
Optimistic Bound	1300.50 (85.65%)	1071.00 (71.24%)	841.50 (56.73%)	612.00 (42.05%)	382.50 (27.00%)	153.00 (11.20%)
ALP Bound	1472.09 (96.95%)	1402.12 (93.27%)	1290.13 (86.98%)	1124.61 (77.26%)	890.58 (62.86%)	554.86 (40.63%)
A Posteriori Bound	**1518.36**	**1503.33**	**1483.24**	**1455.54**	**1416.67**	**1365.78**
Heuristic Policy Cost	**1521.70 (100.22%)**	**1511.41 (100.54%)**	**1501.11 (101.20%)**	**1492.64 (102.55%)**	**1480.63 (104.52%)**	**1456.11 (106.61%)**
Fixed Tour Cost	1530.00 (100.77%)	1530.00 (101.77%)	1530.00 (103.15%)	1530.00 (105.12%)	1530.00 (108.00%)	1530.00 (112.02%)

Table 7: Experiment results for `ftv38` ($n = 38$) with high/low correlated costs.

Instance: ftv33						
H	1.05	1.10	1.15	1.2	1.25	1.30
$\mathsf{P}(H) = 0.5$						
ALP Bound (total)	4776.00	2764.00	2832.00	3099.00	2975.00	3337.00
A Posteriori Bound (avg.)	60.82	72.26	74.98	105.30	132.38	162.32
Heuristic Policy (avg.)	101.68	113.76	122.94	159.94	182.22	205.28
$\mathsf{P}(H) = 0.6$						
ALP Bound (total)	3779.00	2654.00	3243.00	3386.00	3268.00	3843.00
A Posteriori Bound (avg.)	71.98	93.38	110.24	117.98	71.60	40.60
Heuristic Policy (avg.)	102.66	198.02	146.98	181.36	210.22	276.30
$\mathsf{P}(H) = 0.675$						
ALP Bound (total)	2769.00	2599.00	2747.00	2997.00	3337.00	3559.00
A Posteriori Bound (avg.)	69.66	97.36	97.16	63.74	30.44	10.60
Heuristic Policy (avg.)	103.76	118.54	305.36	213.06	481.50	315.82
$\mathsf{P}(H) = 0.75$						
ALP Bound (total)	2145.00	2146.00	2740.00	3421.00	3975.00	4382.00
A Posteriori Bound (avg.)	81.70	152.20	101.04	35.38	6.54	3.16
Heuristic Policy (avg.)	183.66	142.56	353.54	268.42	301.92	320.72

Table 8: Experiment running times in seconds for `ftv33` ($n = 33$) with independently distributed costs.

Instance: ftv33						
H	1.05	1.10	1.15	1.2	1.25	1.30
$\mathsf{P}(H) = 0.5$						
ALP Bound (total)	37861.00	20634.00	17173.00	25279.00	19361.00	19645.00
A Posteriori Bound (avg.)	35.80	43.28	46.12	50.86	58.44	56.86
Heuristic Policy (avg.)	99.74	110.92	114.64	135.86	142.58	144.90
$\mathsf{P}(H) = 0.6$						
ALP Bound (total)	37979.00	32703.00	27398.00	13857.00	19631.00	21297.00
A Posteriori Bound (avg.)	41.80	47.60	52.68	59.68	55.80	55.10
Heuristic Policy (avg.)	96.44	182.64	119.80	121.22	139.30	178.08
$\mathsf{P}(H) = 0.675$						
ALP Bound (total)	89388.00	24058.00	24235.00	18119.00	32521.00	34236.00
A Posteriori Bound (avg.)	45.14	48.48	55.78	55.20	51.10	43.84
Heuristic Policy (avg.)	96.18	189.60	199.86	147.68	190.60	234.16
$\mathsf{P}(H) = 0.75$						
ALP Bound (total)	29677.00	29147.00	31555.00	33493.00	39247.00	28277.00
A Posteriori Bound (avg.)	45.50	51.76	55.64	46.14	47.20	37.22
Heuristic Policy (avg.)	168.10	107.28	139.40	195.56	275.00	248.62

Table 9: Experiment running times in seconds for `ftv33` ($n = 33$) with high/low correlated costs.

References

[1] D. Adelman, *Price-Directed Replenishment of Subsets: Methodology and its Application to Inventory Routing*, Manufacturing and Service Operations Management **5** (2003), 348–371.

[2] ———, *A Price-Directed Approach to Stochastic Inventory/Routing*, Operations Research **52** (2004), 499–514.

[3] ———, *Dynamic Bid Prices in Revenue Management*, Operations Research **55** (2007), 647–661.

[4] ———, *Price-Directed Control of a Closed Logistics Queueing Network*, Operations Research **55** (2007), 1022–1038.

[5] D. Adelman and C. Barz, *A Unifying Approximate Dynamic Programming Model for the Economic Lot Scheduling Problem*, Mathematics of Operations Research (2013), Forthcoming.

[6] D. Adelman and D. Klabjan, *Duality and Existence of Optimal Policies in Generalized Joint Replenishment*, Mathematics of Operations Research **30** (2005), 28–50.

[7] ———, *Computing Near-Optimal Policies in Generalized Joint Replenishment*, INFORMS Journal on Computing **24** (2011), 148–164.

[8] D. Adelman and G.L. Nemhauser, *Price-Directed Control of Remnant Inventory Systems*, Operations Research **47** (1999), 889–898.

[9] S. Ahmed and A. Atamtürk, *Maximizing a class of submodular utility functions*, Mathematical Programming **128** (2011), 149–169.

[10] H.-C. An, R. Kleinberg, and D.B. Shmoys, *Improving Christofides' algorithm for the s-t path TSP*, Proceedings of the 44th ACM Symposium on Theory of Computing (STOC), Association for Computing Machinery, 2012, pp. 875–886.

[11] D.L. Applegate, R.E. Bixby, V. Chvátal, and W.J. Cook, *The traveling salesman problem: A computational study*, Princeton University Press, Princeton, New Jersey, 2006.

[12] R. Bellman, *Dynamic Programming Treatment of the Travelling Salesman Problem*, Journal of the Association for Computing Machinery **9** (1962), 61–63.

[13] D.P. Bertsekas, *Dynamic Programming and Optimal Control: Approximate Dynamic Programming*, fourth ed., vol. II, Athena Scientific, Belmont, Massachusetts, 2012.

[14] D.P. Bertsekas and D.A. Castañón, *Rollout Algorithms for Stochastic Scheduling Problems*, Journal of Heuristics **5** (1999), 89–108.

[15] D.P. Bertsekas and J.N. Tsitsiklis, *An Analysis of Stochastic Shortest Path Problems*, Mathematics of Operations Research **16** (1991), 580–595.

[16] D. Bertsimas and R. Weismantel, *Optimization over Integers*, Dynamic Ideas, Belmont, Massachusetts, 2005.

[17] D.J. Bertsimas, *A Vehicle Routing Problem with Stochastic Demand*, Operations Research **40** (1992), 574–585.

[18] D.J. Bertsimas and G. van Ryzin, *A Stochastic and Dynamic Vehicle Routing Problem in the Euclidean Plane*, Operations Research **39** (1991), 601–615.

[19] ———, *Stochastic and Dynamic Vehicle Routing in the Euclidean Plane with Multiple Capacitated Vehicles*, Operations Research **41** (1993), 60–76.

[20] A.M. Campbell and B.W. Thomas, *Challenges and Advances in A Priori Routing*, in Golden et al. [38], pp. 123–142.

[21] ———, *Probabilistic Traveling Salesman Problem with Deadlines*, Transportation Science **42** (2008), 1–21.

[22] Z.-L. Chen and H. Xu, *Dynamic Column Generation for Dynamic Vehicle Routing with Time Windows*, Transportation Science **40** (2006), 74–88.

[23] T. Cheong and C.C. White, III, *Dynamic Traveling Salesman Problem: Value of Real-Time Traffic Information*, IEEE Transactions on Intelligent Transportation Systems **13** (2012), 619–630.

[24] N. Christofides, *Worst-Case Analysis of a New Heuristic for the Travelling Salesman Problem*, Tech. Report ADA025602, Graduate School of Industrial Administration, Carnegie Mellon University, 1976.

[25] J.-F. Cordeau, G. Laporte, M.W.P. Savelsbergh, and D. Vigo, *Vehicle Routing*, Handbook in Operations Research and Management Science: Transportation, Volume 14 (C. Barnhart and G. Laporte, eds.), Elsevier, 2007, pp. 367–428.

[26] D.P. de Farias and B. van Roy, *The Linear Programming Approach to Approximate Dynamic Programming*, Operations Research **51** (2003), 850–865.

[27] ———, *On Constraint Sampling in the Linear Programming Approach to Approximate Dynamic Programming*, Mathematics of Operations Research **29** (2004), 462–478.

[28] B.C. Dean, M.X. Goemans, and J. Vondrák, *Approximating the Stochastic Knapsack Problem: The Benefit of Adaptivity*, Mathematics of Operations Research **33** (2008), 945–964.

[29] V.V. Desai, V.F. Farias, and C.C. Moallemi, *Approximate Dynamic Programming via a Smoothed Linear Program*, Operations Research **60** (2012), 655–674.

[30] V.F. Farias, D. Sauré, and G.Y. Weintraub, *An Approximate Dynamic Programming Approach to Solving Dynamic Oligopoly Models*, RAND Journal of Economics **43** (2012), 253–282.

[31] M.R. Garey and D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W.H. Freeman and Company, New York, 1979.

[32] M. Gendreau, G. Laporte, and R. Séguin, *Stochastic vehicle routing*, European Journal of Operational Research **88** (1996), 3–12.

[33] G. Ghiani, E. Manni, and B.W. Thomas, *A Comparison of Anticipatory Algorithms for the Dynamic and Stochastic Traveling Salesman Problem*, Transportation Science (2011), Published online before print, doi: 10.1287/trsc.1110.0374.

[34] M.A. Goberna and M.A. López, *Linear Semi-Infinite Optimization*, Wiley Series in Mathematical Methods in Practice, John Wiley & Sons, Chichester, England, 1998.

[35] G.A. Godfrey and W.B. Powell, *An Adaptive Dynamic Programming Algorithm for Dynamic Fleet Management, I: Single Period Travel Times*, Transportation Science **36** (2002), 21–39.

[36] ———, *An Adaptive Dynamic Programming Algorithm for Dynamic Fleet Management, II: Multiperiod Travel Times*, Transportation Science **36** (2002), 40–54.

[37] M.X. Goemans and D.J. Bertsimas, *Probabilistic Analysis of the Held and Karp Lower Bound for the Euclidean Traveling Salesman Problem*, Mathematics of Operations Research **16** (1991), 72–89.

[38] B.L. Golden, S. Raghavan, and E.A. Wasil (eds.), *The Vehicle Routing Problem: Latest Advances and New Challenges*, Springer, 2008.

[39] G. Gutin and A.P. Punnen (eds.), *The Traveling Salesman Problem and Its Variations*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.

[40] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, second ed., Springer – Verlag, 2009.

[41] M. Held and R.M. Karp, *A Dynamic Programming Approach to Sequencing Problems*, Journal of the Society of Industrial and Applied Mathematics **10** (1962), 196–210.

[42] D.S. Hochbaum (ed.), *Approximation Algorithms for NP-Hard Problems*, PWS Publishing Company, Boston, Massachusetts, 1997.

[43] P. Jaillet, *Probabilistic Traveling Salesman Problems*, Ph.D. thesis, Massachusetts Institute of Technology, 1985.

[44] ———, *A Priori Solution of a Traveling Salesman Problem in Which a Random Subset of the Customers are Visited*, Operations Research **36** (1988), 929–936.

[45] P. Jaillet and M.R. Wagner, *Online Vehicle Routing Problems: A Survey*, in Golden et al. [38], pp. 221–237.

[46] H. Jula, M.M. Dessouky, and P. Ioannu, *Truck Route Planning in Non-Stationary Stochastic Networks with Time-Windows at Customer Locations*, IEEE Transactions on Intelligent Transportation Systems **37** (2006), 51–63.

[47] A.S. Kenyon and D.P. Morton, *Stochastic Vehicle Routing with Random Travel Times*, Transportation Science **37** (2003), 69–82.

[48] S. Kim, M.E. Lewis, and C.C. White, III, *Optimal vehicle routing with real-time traffic information*, IEEE Transactions on Intelligent Transportation Systems **6** (2005), 178–188.

[49] ———, *State Space Reduction for Non-stationary Stochastic Shortest Path Problems with Real-Time Traffic Information*, IEEE Transactions on Intelligent Transportation Systems **6** (2005), 273–284.

[50] D. Klabjan and D. Adelman, *An Infinite-Dimensional Linear Programming Algorithm for Deterministic Semi-Markov Decision Processes on Borel Spaces*, Mathematics of Operations Research **32** (2007), 528–550.

[51] G. Laporte, F. Louveaux, and H. Mercure, *The vehicle routing problem with stochastic travel times*, Transportation Science **26** (1992), 161–170.

[52] A. Larsen, *The Dynamic Vehicle Routing Problem*, Ph.D. thesis, Technical University of Denmark, 2000.

[53] A. Larsen, O.B.G. Madsen, and M.M. Solomon, *The A Priori Dynamic Traveling Salesman Problem with Time Windows*, Transportation Science **38** (2004), 459–472.

[54] ———, *Recent Developments in Dynamic Vehicle Routing Systems*, in Golden et al. [38], pp. 199–218.

[55] T. Leipälä, *On the solutions of stochastic traveling salesman problems*, European Journal of Operational Research **2** (1978), 291–297.

[56] I. Murthy and S. Sarkar, *Stochastic Shortest Path Problems with Piecewise-Linear Concave Utility Functions*, Management Science **44** (1998), S125–S136.

[57] S. Nadarajah, F. Margot, and N. Secomandi, *Approximate Dynamic Programs for Natural Gas Storage Valuation Based on Approximate Linear Programming Relaxations*, Tech. Report 2011-E5, Tepper School of Business, Carnegie Mellon University, 2011.

[58] C.H. Papadimitriou and M. Yannakakis, *Shortest paths without a map*, Theoretical Computer Science **84** (1991), 127–150.

[59] S.D. Patek and D.P. Bertsekas, *Stochastic Shortest Path Games*, SIAM Journal on Control and Optimization **37** (1999), 804–824.

[60] M. Petrik, *Optimization-Based Approximate Dynamic Programming*, Ph.D. thesis, University of Massachusetts, Amherst, 2010.

[61] M. Petrik and S. Zilberstein, *Robust Approximate Bilinear Programming for Value Function Approximation*, Journal of Machine Learning Research **12** (2011), 3027–3063.

[62] G.H. Polychronopoulos and J.N. Tsitsiklis, *Stochastic Shortest Path Problems with Recourse*, Networks **27** (1996), 133–143.

[63] W.B. Powell, *Approximate dynamic programming: Solving the curses of dimensionality*, second ed., John Wiley & Sons, Inc., Hoboken, New Jersey, 2010.

[64] J.S. Provan, *A polynomial-time algorithm to find shortest paths with recourse*, Networks **41** (2003), 115–125.

[65] H.N. Psaraftis, *Dynamic vehicle routing: Status and prospects*, Annals of Operations Research **61** (1995), 143–164.

[66] G. Reinelt, *TSPLIB - A Traveling Salesman Problem Library*, ORSA Journal on Computing **3** (1991), 376–384, Updated archive available online at `http://comopt.ifi.uni-heidelberg.de/software/TSPLIB95/`.

[67] A. Schrijver, *The traveling salesman problem*, Combinatorial Optimization: Polyhedra and Efficiency, vol. B, Springer, Berlin, 2003, pp. 981–1004.

[68] P.J. Schweitzer and A. Seidmann, *Generalized Polynomial Approximations in Markovian Decision Processes*, Journal of Mathematical Analysis and Applications **110** (1985), 568–582.

[69] A. Sebő, *Eight-Fifth Approximation for TSP Paths*, To appear in the Proceedings of the 16th Conference on Integer Programming and Combinatorial Optimization, March 18–20, 2013, Valparaíso - Chile. Available on-line at `http://arxiv.org/abs/1209.3523`, 2013.

[70] N. Secomandi, *Analysis of a Rollout Approach to Sequencing Problems with Stochastic Routing Applications*, Journal of Heuristics **9** (2003), 321–352.

[71] N. Secomandi and F. Margot, *Reoptimization Approaches for the Vehicle-Routing Problem with Stochastic Demands*, Operations Research **57** (2009), 214–230.

[72] A. Shapiro, D. Dentcheva, and A.P. Ruszczyński, *Lectures on stochastic programming: Modeling and theory*, Society for Industrial and Applied Mathematics and Mathematical Programming Society, Philadelphia, 2009.

[73] I. Sungur, F. Ordóñez, and M.M. Dessouky, *A Robust Optimization Approach for the Capacitated Vehicle Routing Problem with Demand Uncertainty*, IIE Transactions **40** (2008), 509–523.

[74] B. Thomas and C.C. White, III, *Dynamic Shortest Path Problem with Anticipation*, European Journal of Operational Research **176** (2007), 836–854.

[75] A. Toriello, *Optimal Toll Design: A Lower Bound Framework for the Asymmetric Traveling Salesman Problem*, Mathematical Programming (2013), DOI 10.1007/s10107-013-0631-6. Forthcoming.

[76] A. Toriello, G. Nemhauser, and M. Savelsbergh, *Decomposing inventory routing problems with approximate value functions*, Naval Research Logistics **57** (2010), 718–727.

[77] M.A. Trick and S.E. Zin, *A Linear Programming Approach to Solving Stochastic Dynamic Programs*, Unpublished manuscript available online at `http://mat.gsia.cmu.edu/trick/`, 1993.

[78] ———, *Spline Approximations to Value Functions: A Linear Programming Approach*, Macroeconomic Dynamics **1** (1997), 255–277.