OPERATIONS RESEARCH
Vol. 62, No. 5, September–October 2014, pp. 1107–1125
ISSN 0030-364X (print) | ISSN 1526-5463 (online)
http://dx.doi.org/10.1287/opre.2014.1301

A Dynamic Traveling Salesman Problem with Stochastic Arc Costs

Alejandro Toriello
H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332,
atoriello@isye.gatech.edu

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

We propose a dynamic traveling salesman problem (TSP) with stochastic arc costs motivated by applications, such as dynamic vehicle routing, in which the cost of a decision is known only probabilistically beforehand but is revealed dynamically before the decision is executed. We formulate this as a dynamic program (DP) and compare it to static counterparts to demonstrate the advantage of the dynamic paradigm over an a priori approach. We then apply approximate linear programming (ALP) to overcome the DP's curse of dimensionality, obtain a semi-infinite linear programming lower bound, and discuss its tractability. We also analyze a rollout version of the price-directed policy implied by our ALP and derive worst-case guarantees for its performance. Our computational study demonstrates the quality of a heuristically modified rollout policy using a computationally effective a posteriori bound.

Subject classifications: traveling salesman problem; stochastic vehicle routing; approximate dynamic program; semi-infinite linear program.
Area of review: Transportation.
History: Received August 2012; revisions received June 2013, September 2013, February 2014; accepted May 2014. Published online in *Articles in Advance* July 16, 2014.

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 (Applegate et al. 2006).

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 (1985, 1988) assumed that only the probability that a city must be visited is available beforehand, and that the decision maker must choose 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 still popular modeling and solution paradigm (Campbell and Thomas 2008a).

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 selecting the next place to visit. 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 selecting the next location to visit. The intelligent use of such real-time information within routing offers transportation companies the opportunity for differentiation and a competitive advantage (Cheong and White 2012, Larsen et al. 2008, Psaraftis 1995). 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; Cheong and White (2012) mention that such dynamic routing is informally implemented by urban pickup and delivery companies in Tokyo.

One way to approach problems such as 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 structure (Applegate et al. 2006). 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 even lead to an uncountable state space depending on the costs' support set. To circumvent this difficulty, we use approximate linear programming (ALP); see e.g., de Farias and van Roy (2003). 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 (Adelman 2003, 2004, 2007b; Adelman and Nemhauser 1999). ALP offers a tractable way to study problems such as our dynamic TSP, yielding bounds with theoretical guarantees. Moreover, although unmodified price-directed policies do not always produce good solutions, they give rise to high quality heuristically modified policies that can be efficiently computed. Our computational study verifies the empirical quality of one such policy, using an effective a posteriori bound.

1.1. Our Contribution

We consider the following to be our main contributions:

Section 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.

Section 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 polytopes with polynomially many extreme points or hyper-rectangles.

Section 4. We analyze a lookahead or rollout 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.

Section 5. We propose a heuristic version of the rollout policy and adapt a computationally effective a posteriori bound from Secomandi and Margot (2009). Our computational study verifies the empirical quality of both.

1.2. Literature Review

Although models such as the one we propose appear rarely, the literature on stochastic and/or dynamic TSP and more general vehicle routing is vast. We briefly review some salient topics; interested readers may refer to Cordeau et al. (2007), Gendreau et al. (1996), Golden et al. (2008) for more details. These texts Applegate et al. (2006), Gutin and Punnen (2002), Schrijver (2003) provide comprehensive coverage of the deterministic TSP.

The literature includes various stochastic or probabilistic TSP models, usually assuming that 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 (Leipälä 1978), city locations in a Euclidean instance (Goemans and Bertsimas 1991, Jaillet 1985), or the subset of cities to visit from a ground set (Jaillet 1985, 1988).

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., Bertsimas (1992), Campbell and Thomas (2008b), Jaillet (1985, 1988), Sungur et al. (2008); see Campbell and Thomas (2008a) 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 (Bertsimas and van Ryzin 1991, 1993; Chen and Xu 2006; Ghiani et al. 2012; Goodson et al. 2013; Larsen 2000; Larsen et al. 2004; Secomandi 2001; Secomandi and Margot 2009). The surveys Larsen et al. (2008), Psaraftis (1995) cover dynamic routing issues.

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 Jaillet and Wagner (2008) 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 (Jula et al. 2006, Kenyon and Morton 2003, Laporte et al. 1992). Our approach to model uncertain arc costs is similar to many stochastic shortest path problems, e.g., Bertsekas and Tsitsiklis (1991), Murthy and Sarkar (1998), Patek and Bertsekas (1999), Polychronopoulos and Tsitsiklis (1996), Provan (2003), and sometimes appears in real-time shortest path applications (Kim et al. 2005a, b; Thomas and White 2007).

In the literature, the most similar models to ours are perhaps Cheong and White (2012), Secomandi (2003). The model in Cheong and White (2012) 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 Secomandi (2003) also has stochastic costs and allows dynamic decisions, though the paper focuses more on minimizing tardiness. The author proposes a rollout policy (Bertsekas and Castañón 1999) and demonstrates its effectiveness with computational experiments. Our work can be viewed in this context as well; we discuss similarities below in §§4 and 5. Other work applying similar techniques in routing and related topics includes Bertsimas and Demir (2002), Bertsimas and Popescu (2003), Goodson et al. (2013), Secomandi (2001).

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 (Toriello 2014) is the only other use of ALP in routing models. The concept was first studied as early as a quarter-century ago (Schweitzer and Seidmann 1985; Trick and Zin 1993, 1997), and has received growing attention in the past decade (de Farias and van Roy 2003, 2004; Desai et al. 2012). Specific applications of ALP include commodity valuation (Nadarajah et al. 2011), economic lot scheduling (Adelman and Barz 2014), inventory routing (Adelman 2003, 2004), joint replenishment (Adelman and Klabjan 2005, 2011; Klabjan and Adelman 2007), revenue management (Adelman 2007a), and stochastic games (Farias et al. 2012).

However, the general toolkit of approximate dynamic programming (ADP) has been used extensively in routing and fleet management, e.g., Bertsimas and Weismantel (2005, Chap. 11.4), Godfrey and Powell (2002a, b), Toriello et al. (2010); the text Powell (2010) 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 (Bertsekas 2012), approximate value iteration (Powell 2010), approximate bilinear programming (Petrik 2010, Petrik and Zilberstein 2011), as well as various statistical methods, e.g., parametric and nonparametric regression (Hastie et al. 2009, Powell 2010).

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 for a rollout version of the policy. Section 5 discusses the heuristically modified policy, the computational a posteriori bound technique, and outlines our computational study. Section 6 presents concluding remarks and suggested avenues for future research. The appendix contains additional discussion, 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 returning to a distinguished city 0, sometimes called the home city or depot. In our model, each arc cost is random and realized on arrival at the arc's tail. The objective is to minimize the expected total cost of the tour. 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 $C_i = (C_{ij}\colon j \in N \cup 0 \setminus i) \in \mathbb{R}^n$ 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 $\mathscr{C}_i \subseteq \mathbb{R}^n$ 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 $\mathsf{E}[C_{ij}]$, $\min_{c_i \in \mathscr{C}_i} c_{ij}$, and $\max_{c_i \in \mathscr{C}_i} c_{ij}$.

We base our DP formulation on the classical formulation for the deterministic TSP (Bellman 1962, Held and Karp 1962), 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

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

States $\{(0, N, c_0)\colon c_0 \in \mathscr{C}_0\}$ correspond to the start of the tour. In state $(0, N, c_0)$, the salesman is at city 0, and has all cities in N left to visit; the outgoing arc costs are given by c_0. The states $\{(i, U, c_i)\colon i \in N,\ U \subseteq N \setminus i,\ c_i \in \mathscr{C}_i\}$ 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 $c_j \in \mathscr{C}_j$, 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; 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 $(i, U, c_i) \in \mathscr{S}$; the collection $(y^*_{i,U}(c_i)\colon c_i \in \mathscr{C}_i)$ 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 1. *Each $y^*_{i,U}: \mathscr{C}_i \to \mathbb{R}$ is nondecreasing, 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 $\pi^*: \mathscr{S}\backslash(0,\varnothing) \to N \cup 0$ that maps from the current state $(i, U, c_i) \in \mathscr{S}$ to an action $j \in U$,

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

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

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

$$\text{s.t.} \quad y_{0,N}(c_0) - \mathsf{E}[y_{i,N\backslash i}(C_i)] \leqslant c_{0i}, \quad \forall\, i \in N,\ c_0 \in \mathscr{C}_0, \tag{3b}$$

$$y_{i,U\cup j}(c_i) - \mathsf{E}[y_{j,U}(C_j)] \leqslant c_{ij}, \quad \forall\, i \in N,\ j \in N\backslash i,\ U \subseteq N\backslash\{i,j\},\ c_i \in \mathscr{C}_i, \tag{3c}$$

$$y_{i,\varnothing}(c_i) \leqslant c_{i0}, \quad \forall\, i \in N,\ c_i \in \mathscr{C}_i. \tag{3d}$$

This LP is potentially doubly infinite, since each variable $y_{i,U}$ is a function from the possibly infinite set $\mathscr{C}_i$ to $\mathbb{R}$, and the constraints range over the state-action pairs, also indexed partly by the sets $\mathscr{C}_i$. In light of Proposition 1, we can restrict the feasible region of each $y_{i,U}$ to an appropriately chosen, well behaved functional space on $\mathscr{C}_i$, such as the space of continuous functions on $\mathscr{C}_i$; 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.

PROPOSITION 2. *Let $y_{i,U}: \mathscr{C}_i \to \mathbb{R}$ for $i \in N \cup 0$ and $U \subseteq N\backslash i$ be feasible for (3). Then $y_{i,U}(c_i) \leqslant y^*_{i,U}(c_i)$ for all $(i, U, c_i) \in \mathscr{S}$. In particular, $\mathsf{E}[y_{0,N}(C_0)] \leqslant \mathsf{E}[y^*_{0,N}(C_0)]$.*

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

Note that analogous results hold for the LP formulation of any DP.

2.1. Comparison to Static TSP

An important question about the dynamic TSP model is whether we gain by allowing dynamic decisions; i.e., by being adaptive (Dean et al. 2008). 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 1 (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 $t = 0, \ldots, n-1$, 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 Dean et al. (2008), 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 $\mathscr{C}_i$ 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

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

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

PROPOSITION 3. *For any $(i, U, c_i) \in \mathscr{S}$,*

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

and

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

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

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

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

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

where the last inequality follows from induction. □

The norm in the definition of $\mathbb{D}$ may be any l_p norm, i.e., $\|x\|_p = (\sum_i |x_i|^p)^{1/p}$. Also, from the definition of $\bar{c}$ this proof only uses $\bar{c}_i \in \operatorname{conv}(\mathscr{C}_i)$; 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 $\underline{c}_{ij} = \min_{c_i \in \mathscr{C}_i} c_{ij}$. 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., Shapiro et al. (2009, Chap. 7).

Proposition 4. *For any* $(i, U, c_i) \in \mathscr{S}$,

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

and

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

Proof. The left-hand inequality is proved in a similar fashion to Proposition 3. For the right-hand inequality, if we let $y_{i,U}(c_i) = \underline{y}_{i,U}$, $\forall c_i \in \mathscr{C}_i$, then y is feasible for (3); 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 4 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 $m \lll |\mathscr{S}|$ we define a collection of basis vectors or functions $b_{i,U}\colon \mathscr{C}_i \to \mathbb{R}^m$, 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

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

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

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

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

Problem (7) is a semi-infinite LP; when the support sets $\mathscr{C}_i$ 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 2 (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 $\underline{c}_{ij} = \min_{c_i\in\mathscr{C}_i} c_{ij}$; the left-hand sides of the constraints (7b) through (7d) do not vary with c, and therefore (7) becomes

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

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 4.

Though we do not pursue it here, it is possible to apply ALP to the deterministic TSP in this way to obtain lower bounds (Toriello 2014). Our focus instead is solving (7) for a more complex basis that accounts for different arc costs to improve on the optimistic bound.

3.1. Affine Cost-to-Go Approximation

Consider the cost-to-go approximation

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

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

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

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 before 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

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

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

$$\lambda_{i0} - \lambda_{j0} + \lambda_{ij} + c_{ij}\eta_{ij} + \sum_{k \in U} (\lambda_{ik} + c_{ik}\eta_{ik} - \lambda_{jk} - \bar{c}_{jk}\eta_{jk}) \leqslant c_{ij}, \quad \forall i \in N,\ j \in N\setminus i,\ \varnothing \neq U \subseteq N \setminus \{i,j\},\ c_i \in \mathscr{C}_i, \tag{9c}$$

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

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

As with many ALPs, 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 $\mathscr{C}_i$.

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

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

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

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

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

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

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

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

$$x \geqslant 0, \quad x \text{ has finite support.} \tag{10h}$$

In this relaxed primal model, each variable $x_{ij,U}^{c_i}$ corresponds to the probability of visiting state $(i, U \cup j, c_i)$ and choosing action j. Similarly, variables $x_{0i}^{c_0}$ and $x_{i0}^{c_i}$ 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 $\mathscr{C}_i$, but in any of these cases only a finite number must be positive at any optimal solution.

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 $x_{0i}^{c_0}$ 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 on 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 remains, either by going to j itself or to another remaining city. Here again we have a relaxation because 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 on visiting i before j, to equal $\bar{c}_{ij}$. The proof of the following proposition is included in the appendix.

Proposition 5. *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 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 1. *In* (9), *we can impose* $\eta \geqslant 0$ *without loss of optimality.*

Proof. Adding $\eta \geqslant 0$ to (9) and using an argument identical to the proof of Proposition 5, 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 $\underline{c}$,

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

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

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

$$\sum_{i\in U} \sum_{j\in N\cup 0\backslash U} z_{ij} \geqslant 1, \quad \forall\, \varnothing \neq U \subseteq N\cup 0, \tag{11d}$$

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

We use z variables instead of x to distinguish the formulations. This LP is solvable in polynomial time (Applegate et al. 2006) 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 4.

THEOREM 1. *The optimal value of* (9) *provides a lower bound greater than or equal to the bound provided by* (11).

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

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

It follows from Theorem 18 in Toriello (2014) 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 (EXAMPLE 1 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

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

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 $\eta \geqslant 0$, the first and second constraint classes collapse to

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

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

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

where V represents the second LP's optimal value. It is simple to show that $V = \max\{0, p-(n-2)(1-p)\}$. 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 nonzero for $p > (n-2)/(n-1)$, 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 Adelman (2003, 2004), Adelman and Klabjan (2011), Klabjan and Adelman (2007), Toriello (2014). For constraint classes (9b, 9d, 9e), the separation problem is equivalent to maximizing a linear function over one of the sets $\mathscr{C}_i$, 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 efficiently accounted for 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 $j \in N\backslash i$ the separation problem is equivalent to

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

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

LEMMA 1. *The separation problem* (12) *is NP-hard even when the sets* $\mathscr{C}_i$ *are* l_2 *balls.*

The proof of this lemma is in the appendix. In this general case, exact separation is inefficient; constraint sampling (de Farias and van Roy 2003, 2004; Desai et al. 2012) is likely the only viable choice, but requires 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 6. *Suppose each set* $\mathscr{C}_i$ *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 $c_i \in \mathscr{C}_i$, the maximization in (12) can be solved greedily: For each $k \in N\backslash\{i, j\}$, include it in the set U only if the term in the parenthesis is positive. When none of these terms is positive, include the nonpositive term of smallest absolute value to ensure $U \neq \varnothing$. If $\mathscr{C}_i$ has $O(p(n))$ extreme points, we can use this greedy optimization procedure at each extreme point and then choose the overall maximizer, all in $O(np(n))$ time. □

As two examples, Proposition 6 covers the cases where each set $\mathscr{C}_i$ is a simplex or an l_1 ball.

PROPOSITION 7. *Suppose the convex hull of each set* $\mathscr{C}_i$ *for* $i \in N$ *is a hyper-rectangle*: $\operatorname{conv}(\mathscr{C}_i) = [\underline{c}_i, \hat{c}_i]$, *with* $\underline{c}_i, \hat{c}_i \in \mathbb{R}^n$ *representing lower- and upper-bound vectors, respectively. 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 $k \in N\backslash\{i, j\}$, by Corollary 1 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 nonpositive 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 $\mathscr{C}_i$: Solve (12) over the hyper-rectangles defined by $\underline{c}_{ij} = \min_{c_i \in \mathscr{C}_i} c_{ij}$ and $\hat{c}_{ij} = \max_{c_i \in \mathscr{C}_i} c_{ij}$ 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 (de Farias and van Roy 2003, Desai et al. 2012), or price-directed (Adelman 2003, 2004, 2007b; Adelman and Nemhauser 1999). 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

$$\pi_{\lambda,\eta}(i, U, c_i) = \begin{cases} \operatorname*{arg\,min}_{j\in U}\left\{c_{ij} + \lambda_{j0} + \sum_{k\in U\backslash j}(\lambda_{jk} + \bar{c}_{jk}\eta_{jk})\right\}, \\ \qquad |U| \geqslant 2, \\ j, \quad U = j \in N\backslash i, \\ 0, \quad U = \varnothing, \end{cases} \quad (13)$$

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 nondecreasing with respect to the set of remaining cities:

$$\mathsf{E}[y^*_{i,U}(C_i)] \leqslant \mathsf{E}[y^*_{i,U\cup j}(C_i)], \quad \forall\, i \in N,\ j \in N\backslash i,\ U \subseteq N\backslash\{i, j\}, \quad (14a)$$

$$\mathsf{E}[y^*_{i,N\backslash i}(C_i)] \leqslant \mathsf{E}[y^*_{0,N}(C_0)], \quad \forall\, i \in N. \quad (14b)$$

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 nonnegative 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 2. *Let* $\tilde{y}_{i,U}\colon \mathscr{C}_i \to \mathbb{R}$ *be an approximate cost-to-go function for each* $i \in N \cup 0$, $U \subseteq N\backslash i$ *satisfying* $\mathsf{E}[\tilde{y}_{i,U}(C_i)] \leqslant \mathsf{E}[y^*_{i,U}(C_i)]$. *Assume* (14) *holds, and suppose in addition that* $\mathsf{E}[y^*_{i,U}(C_i)] \geqslant 0$ *and*

$$\alpha\,\mathsf{E}[y^*_{i,U}(C_i)] \leqslant \mathsf{E}[\tilde{y}_{i,U}(C_i)], \quad \forall\, i \in N \cup 0,\ U \subseteq N\backslash i, \quad (15)$$

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., de Farias and van Roy (2003, Theorem 1) or Desai et al. (2012, 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., Hochbaum (1997).

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

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

and therefore

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

Let $c \in \mathscr{C}_0 \times \cdots \times \mathscr{C}_n$ 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

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

Let $\mathscr{U}_i$ 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,

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

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

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

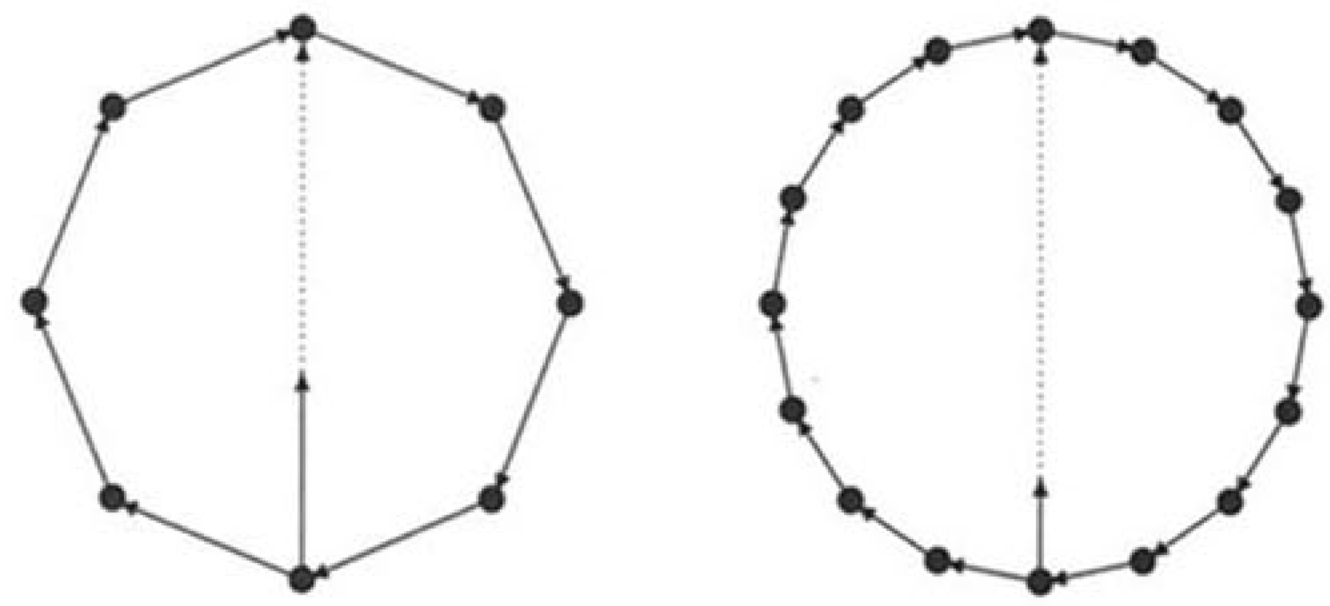
where the last inequality follows from (14). □

Unfortunately, this result cannot be used directly with an optimal solution of (9). As the next example shows, even in the deterministic case the ALP could 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. Consider a deterministic TSP instance with cities arrayed around a unit circle at equal intervals from 0 to n, and arc costs given by Euclidean distances; see Figure 1. The distance between consecutive cities is $d_n = 2\sin(\pi/(n+1))$, an optimal tour with cost $(n+1)d_n$ follows the cities 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 $((n+1)/2, \varnothing)$, i.e., when the salesman is at city $(n+1)/2$ 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 $n \to \infty$.

To take advantage of Theorem 2, we must therefore recompute the bound at every step for every available action

Figure 1. Instances from Example 4 with $n=7$ and $n=15$.



Note. 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) underapproximates the cost-to-go. The true cost is the vertical dotted arrow, and the approximate cost is the overlying solid arrow.

in a rollout framework (Bertsekas and Castañón 1999); this involves $O(n^2)$ total solutions of LPs 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. In the next section, we discuss a heuristic modification of this idea that yields practical computing times and high quality solutions.

Theorem 3. *Suppose the following conditions hold:*

(i) $\underline{c}_{ij} > 0$ *and* $\hat{c}_{ij}/\underline{c}_{ij} \leqslant \Gamma$, *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* $\underline{c}_{ij} \leqslant \underline{c}_{ik} + \underline{c}_{kj}$, *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 2 applies with* $\alpha = 5/(8\Gamma)$, *so that the total expected cost of this policy is bounded above by* $(1 + (1 - 5/(8\Gamma))n)\, \mathsf{E}[y^*_{0,N}(C_0)]$.

For the conditions given in Theorem 3, if we apply Christofides' heuristic (1976) using arc costs $\underline{c}$, we are guaranteed a $(3\Gamma)/2$-approximation of the optimal policy's expected cost. However, in contrast to the policies we introduce in this section, this solution is nonadaptive and may perform poorly in practice, as Example 1 illustrates.

Proof. From (i) we know that $\underline{y}_{j,U\setminus j} \geqslant \Gamma\, \mathsf{E}[y^*_{j,U\setminus j}(C_j)]$ 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 1, $\mathsf{E}[\tilde{y}_{j,U\setminus j}(C_j)]$ 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 Sebő (2013). □

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 (An et al. 2012, Sebő 2013).

Any improvement in this bound guarantee would immediately imply a corresponding improved guarantee for Theorem 3's policy via Theorem 1.

5. Computational Experiments

We next discuss a series of computational experiments designed to test the efficacy of various bounds and policies, including a heuristic rollout version of the price-directed policy. To generate test instances, we used deterministic asymmetric TSP instances from TSPLIB (Reinelt 1991) 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, Cheong and White's (2012) computational work on a related model suggests 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 $\mathscr{C}_i$ 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 $\beta_H = \beta_L$. Yet 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 $\mathscr{C}_i$ 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 $\underline{y}_{0,N}$.

In addition, we include an a posteriori bound (Secomandi and Margot 2009), 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 (Leipälä 1978) 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. However, for the instance size we consider these deterministic TSP's can be solved efficiently, and thus the a posteriori technique is computationally feasible.

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 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 rollout policy to test the recomputing idea presented in Theorem 3. At any encountered state (i, U, c_i) with $|U| \geqslant 2$, we use

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

where $\bar{y}^{\mathrm{LP}}_{j, U\setminus j}$ 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 other rollout policies (Bertsekas and Castañón 1999, Secomandi 2003), in which a 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 Adelman and Klabjan 2011), 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 other rollout policies, our heuristic uses $\bar{y}^{\mathrm{LP}}_{j, U\setminus j}$, 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; full results

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

Instance: ftv33 (%)						
H	1.05	1.10	1.15	1.2	1.25	1.30
$\mathsf{P}(H) = 0.5$						
Optimistic bound	95.12	90.76	87.28	84.49	82.02	79.91
ALP bound	98.59	95.89	93.69	91.98	90.41	89.08
A posteriori bound	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Heuristic policy cost	**100.04**	**100.55**	**101.73**	**104.25**	**107.04**	**110.04**
Fixed tour cost	100.13	100.85	102.69	105.61	109.36	114.16
$\mathsf{P}(H) = 0.6$						
Optimistic bound	92.53	86.21	81.26	77.13	73.74	70.86
ALP bound	97.68	94.14	91.31	88.84	86.72	84.90
A posteriori bound	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Heuristic policy cost	**100.09**	**101.24**	**104.10**	**108.76**	**114.02**	**121.57**
Fixed tour cost	100.03	101.42	104.85	110.19	117.99	128.83
$\mathsf{P}(H) = 0.675$						
Optimistic bound	89.86	81.33	74.61	69.23	64.89	61.51
ALP bound	97.17	92.61	89.08	85.49	82.61	80.21
A posteriori bound	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Heuristic policy cost	**100.15**	**101.93**	**106.85**	**114.18**	**124.30**	**141.58**
Fixed tour cost	100.27	102.65	108.38	118.42	134.97	163.19
$\mathsf{P}(H) = 0.75$						
Optimistic bound	85.76	74.16	65.11	57.74	52.48	50.46
ALP bound	96.54	91.25	85.86	80.46	75.47	72.72
A posteriori bound	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Heuristic policy cost	**100.27**	**104.24**	**113.64**	**131.36**	**170.11**	**291.17**
Fixed tour cost	100.89	105.94	118.39	144.35	209.92	504.56

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

Instance: ftv33 (%)						
H	1.05	1.10	1.15	1.2	1.25	1.30
$\mathsf{P}(H) = 0.5$						
Optimistic bound	95.51	90.97	86.39	81.75	77.09	72.49
ALP bound	100.00	99.04	97.48	95.62	93.63	91.66
A posteriori bound	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Heuristic policy cost	**100.00**	**100.00**	**100.00**	**100.01**	**100.04**	**100.24**
Fixed tour cost	100.54	101.08	101.63	102.19	102.79	103.56
$\mathsf{P}(H) = 0.6$						
Optimistic bound	93.14	86.19	80.72	72.06	65.09	58.14
ALP bound	99.78	98.04	97.44	92.70	89.74	86.59
A posteriori bound	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Heuristic policy cost	**100.00**	**100.00**	**102.00**	**100.10**	**100.56**	**101.49**
Fixed tour cost	100.69	101.40	104.15	102.94	104.15	105.71
$\mathsf{P}(H) = 0.675$						
Optimistic bound	90.27	80.39	70.39	60.60	50.79	40.83
ALP bound	99.30	96.58	92.95	88.84	83.71	76.44
A posteriori bound	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Heuristic policy cost	**100.00**	**100.00**	**100.04**	**100.67**	**101.94**	**103.34**
Fixed tour cost	100.73	101.46	102.25	103.66	105.63	108.32
$\mathsf{P}(H) = 0.75$						
Optimistic bound	85.73	71.21	56.78	42.35	27.48	11.54
ALP bound	98.67	94.54	88.76	79.70	66.05	43.94
A posteriori bound	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Heuristic policy cost	**100.00**	**100.00**	**100.64**	**101.90**	**104.64**	**108.35**
Fixed tour cost	100.86	101.73	103.23	105.87	109.91	115.41

for all 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:

$$\text{best optimistic bound, } \underline{y}_{0,N} \leqslant \text{ALP bound, } \lambda_0^* + \sum_{k\in N} \bar{c}_{0k}\eta_{0k}^* \leqslant \textit{a posteriori} \text{ bound} \leqslant \text{heuristic policy} \leqslant \text{best fixed tour, } \bar{y}_{0,N}. \quad (16)$$

Because the a posteriori bound is the tightest, we report all other quantities as percentages of this bound; tables in the appendix also contain absolute quantities. 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 a posteriori bound is quite tight in most cases, though its quality decreases in extreme cases (the lower-right corner of the table). The ALP is reasonably close to the *a posteriori* bound, within approximately 80%, in all but the most extreme cases as well. 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 any two bounds is increasing with respect to both H and $\mathsf{P}(H)$. 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 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 shows results for ftv33 with high/low correlated arc costs, again reported as percentages of the a posteriori bound, with full results in the appendix. We do not include results for ftv44 because we were unable 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 3 suggests $\bar{y}^{\mathrm{LP}}_{j,U\setminus j}$ 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. The design of more sophisticated bounds, using a posteriori techniques, ALP or different methods altogether, is therefore 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 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 bound the problem, construct policies with theoretical worst-case performance guarantees, and derive a high-quality heuristic. Our experimental results verify the empirical quality of the heuristic and also of the a posteriori bounding technique adapted from Secomandi and Margot (2009). Our work leads to several questions.

For example, we do not address how to obtain an ALP bound when the separation problem for (9) is NP-hard. As we indicate, constraint sampling may be an option (de Farias and van Roy 2003, 2004; Desai et al. 2012), though it requires 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 successfully applied in other contexts, e.g., Adelman and Klabjan (2011).

The empirical success of the a posteriori bound compared to the ALP bound suggests questions as well. On one hand, a theoretical analysis of the a posteriori technique could shed light on its effectiveness for this problem and related transportation models, and perhaps motivate different heuristic policies. On the other, one could also design better bases to more accurately approximate the cost-to-go with ALP. This is a challenging question as any basis must trade off approximation fidelity with computational tractability, and the trade-off is not always obvious. For the TSP, the results in Toriello (2014) suggest that only small improvements are possible when considering the combinatorial aspect of the state space; whether the same is true for the cost component is unclear.

Finally, the wide gaps for instances with high variability suggest that significant work remains to obtain bounds with theoretical guarantees or empirical success. More generally, though ALP offers one approach to model, bound and solve these problems, it is clear that many challenges remain.

Acknowledgments

The authors thank the associate editor and two anonymous referees for their prompt, thorough reviews and many helpful comments, which helped greatly improve the paper.

Appendix

Allowing Returns

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 on their first realization: Suppose the costs the salesman observes on 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 (Polychronopoulos and Tsitsiklis 1996), or a stochastic version of the *Canadian traveler's problem* (Papadimitriou and Yannakakis 1991). Even shortest path versions of this problem are quite difficult: The adversarial version is PSPACE-complete and the stochastic version is #P-hard (Papadimitriou and Yannakakis 1991, Polychronopoulos and Tsitsiklis 1996). The state space $\mathcal{S}$ must be augmented to include previously observed costs from past visits, increasing its dimension by one 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 resampled for each visit to a city: Suppose costs are resampled 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, this is still a stochastic shortest path problem (Bertsekas and Tsitsiklis 1991). The DP recursion is now

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

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

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

Some of our 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.

Remaining Proofs

Proof of Proposition 5. 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 Goberna and López (1998, Theorem 5.3), the semi-infinite constraint system is Farkas-Minkowski, and thus the Haar dual (10) is strong, with an optimal solution (Goberna and López 1998, Theorem 8.4). □

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

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

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

Instance: ftv33						
H	1.05	1.10	1.15	1.2	1.25	1.30
			$\mathsf{P}(H)=0.5$			
Optimistic bound	1,221.70 (95.12%)	1,157.40 (90.76%)	1,093.10 (87.28%)	1,028.80 (84.49%)	964.50 (82.02%)	900.20 (79.91%)
ALP bound	1,266.19 (98.59%)	1,222.82 (95.89%)	1,173.34 (93.69%)	1,120.05 (91.98%)	1,063.11 (90.41%)	1,003.44 (89.08%)
A posteriori bound	**1,284.35**	**1,275.20**	**1,252.36**	**1,217.73**	**1,175.90**	**1,126.48**
Heuristic policy cost	**1,284.85 (100.04%)**	**1,282.17 (100.55%)**	**1,274.07 (101.73%)**	**1,269.47 (104.25%)**	**1,258.63 (107.04%)**	**1,239.55 (110.04%)**
Fixed tour cost	1,286.00 (100.13%)	1,286.00 (100.85%)	1,286.00 (102.69%)	1,286.00 (105.61%)	1,286.00 (109.36%)	1,286.00 (114.16%)
			$\mathsf{P}(H)=0.6$			
Optimistic bound	1,189.55 (92.53%)	1,093.10 (86.21%)	996.65 (81.26%)	900.20 (77.13%)	803.75 (73.74%)	707.30 (70.86%)
ALP bound	1,255.72 (97.68%)	1,193.63 (94.14%)	1,119.86 (91.31%)	1,036.81 (88.84%)	945.19 (86.72%)	847.46 (84.90%)
A posteriori bound	**1,285.57**	**1,267.96**	**1,226.48**	**1,167.06**	**1,089.93**	**998.23**
Heuristic policy cost	**1,286.72 (100.09%)**	**1,283.71 (101.24%)**	**1,276.75 (104.10%)**	**1,269.34 (108.76%)**	**1,242.72 (114.02%)**	**1,213.57 (121.57%)**
Fixed tour cost	1,286.00 (100.03%)	1,286.00 (101.42%)	1,286.00 (104.85%)	1,286.00 (110.19%)	1,286.00 (117.99%)	1,286.00 (128.83%)
			$\mathsf{P}(H)=0.675$			
Optimistic bound	1,152.45 (89.86%)	1,018.91 (81.33%)	885.36 (74.61%)	751.82 (69.23%)	618.27 (64.89%)	484.72 (61.51%)
ALP bound	1,246.28 (97.17%)	1,160.22 (92.61%)	1,057.06 (89.08%)	928.39 (85.49%)	787.13 (82.61%)	632.08 (80.21%)
A posteriori bound	**1,282.54**	**1,252.83**	**1,186.60**	**1,085.92**	**952.81**	**788.04**
Heuristic policy cost	**1,284.43 (100.15%)**	**1,277.00 (101.93%)**	**1,267.91 (106.85%)**	**1,239.88 (114.18%)**	**1,184.30 (124.30%)**	**1,115.72 (141.58%)**
Fixed tour cost	1,286.00 (100.27%)	1,286.00 (102.65%)	1,286.00 (108.38%)	1,286.00 (118.42%)	1,286.00 (134.97%)	1,286.00 (163.19%)
			$\mathsf{P}(H)=0.75$			
Optimistic bound	1,093.10 (85.76%)	900.20 (74.16%)	707.30 (65.11%)	514.40 (57.74%)	321.50 (52.48%)	128.60 (50.46%)
ALP bound	1,230.60 (96.54%)	1,107.64 (91.25%)	932.65 (85.86%)	716.83 (80.46%)	462.33 (75.47%)	185.35 (72.72%)
A posteriori bound	**1,274.67**	**1,213.88**	**1,086.25**	**890.92**	**612.62**	**254.87**
Heuristic policy cost	**1,278.12 (100.27%)**	**1,265.30 (104.24%)**	**1,234.38 (113.64%)**	**1,170.34 (131.36%)**	**1,042.11 (170.11%)**	**742.13 (291.17%)**
Fixed tour cost	1,286.00 (100.89%)	1,286.00 (105.94%)	1,286.00 (118.39%)	1,286.00 (144.35%)	1,286.00 (209.92%)	1,286.00 (504.56%)

Table A.2. Experiment results for `ftv35` ($n = 35$) with independently distributed costs.

Instance: ftv35						
H	1.05	1.10	1.15	1.2	1.25	1.30
			$\mathrm{P}(H) = 0.5$			
Optimistic bound	1,399.35 (95.72%)	1,325.70 (91.95%)	1,252.05 (88.74%)	1,178.40 (86.03%)	1,104.75 (83.73%)	1,031.10 (81.76%)
ALP bound	1,425.15 (97.48%)	1,378.83 (95.64%)	1,325.46 (93.94%)	1,263.20 (92.22%)	1,194.44 (90.53%)	1,122.58 (89.02%)
A posteriori bound	**1,461.94**	**1,441.69**	**1,410.98**	**1,369.72**	**1,319.35**	**1,261.11**
Heuristic policy cost	**1,470.69 (100.60%)**	**1,463.15 (101.49%)**	**1,456.03 (103.19%)**	**1,451.35 (105.96%)**	**1,427.15 (108.17%)**	**1,415.95 (112.28%)**
Fixed tour cost	1,473.00 (100.76%)	1,473.00 (102.17%)	1,473.00 (104.40%)	1,473.00 (107.54%)	1,473.00 (111.65%)	1,473.00 (116.80%)
			$\mathrm{P}(H) = 0.6$			
Optimistic bound	1,362.53 (93.44%)	1,252.05 (87.64%)	1,141.58 (82.67%)	1,031.10 (78.72%)	920.63 (74.86%)	810.15 (73.15%)
ALP bound	1,414.33 (96.99%)	1,345.68 (94.19%)	1,262.01 (91.39%)	1,162.56 (88.76%)	1,052.29 (85.56%)	934.76 (84.40%)
A posteriori bound	**1,458.16**	**1,428.61**	**1,380.93**	**1,309.84**	**1,229.83**	**1,107.59**
Heuristic policy cost	**1,470.85 (100.87%)**	**1,461.08 (102.27%)**	**1,449.94 (105.00%)**	**1,426.23 (108.89%)**	**1,416.45 (115.17%)**	**1,351.20 (121.99%)**
Fixed tour cost	1,473.00 (101.02%)	1,473.00 (103.11%)	1,473.00 (106.67%)	1,473.00 (112.46%)	1,473.00 (119.77%)	1,473.00 (132.99%)
			$\mathrm{P}(H) = 0.675$			
Optimistic bound	1,320.03 (90.74%)	1,167.07 (82.66%)	1,014.10 (76.06%)	861.14 (70.95%)	708.17 (66.99%)	555.21 (63.91%)
ALP bound	1,402.88 (96.43%)	1,310.57 (92.82%)	1,187.00 (89.03%)	1,032.84 (85.10%)	864.19 (81.74%)	687.51 (79.14%)
A posteriori bound	**1,454.78**	**1,411.89**	**1,333.27**	**1,213.72**	**1,057.18**	**868.76**
Heuristic policy cost	**1,468.67 (100.95%)**	**1,457.82 (103.25%)**	**1,439.96 (108.00%)**	**1,406.56 (115.89%)**	**1,346.17 (127.34%)**	**1,259.09 (144.93%)**
Fixed tour cost	1,473.00 (101.25%)	1,473.00 (104.33%)	1,473.00 (110.48%)	1,473.00 (121.36%)	1,473.00 (139.33%)	1,473.00 (169.55%)
			$\mathrm{P}(H) = 0.75$			
Optimistic bound	1,252.05 (86.57%)	1,031.10 (75.14%)	810.15 (66.31%)	589.20 (59.43%)	368.25 (54.92%)	147.30 (52.75%)
ALP bound	1,388.88 (96.03%)	1,249.72 (91.07%)	1,036.51 (84.83%)	780.51 (78.73%)	498.32 (74.32%)	201.04 (72.00%)
A posteriori bound	**1,446.31**	**1,372.23**	**1,221.81**	**991.42**	**670.48**	**279.22**
Heuristic policy cost	**1,465.39 (101.32%)**	**1,453.57 (105.93%)**	**1,411.82 (115.55%)**	**1,315.46 (132.68%)**	**1,131.87 (168.81%)**	**827.07 (296.21%)**
Fixed tour cost	1,473.00 (101.85%)	1,473.00 (107.34%)	1,473.00 (120.56%)	1,473.00 (148.58%)	1,473.00 (219.69%)	1,473.00 (527.54%)

Table A.3. Experiment results for `ftv38` ($n = 38$) with independently distributed costs.

Instance: ftv38						
H	1.05	1.10	1.15	1.2	1.25	1.30
			$\mathrm{P}(H) = 0.5$			
Optimistic bound	1,453.50 (95.65%)	1,377.00 (91.94%)	1,300.50 (88.76%)	1,224.00 (86.04%)	1,147.50 (83.77%)	1,071.00 (81.90%)
ALP bound	1,477.39 (97.22%)	1,426.16 (95.22%)	1,363.75 (93.08%)	1,295.80 (91.09%)	1,225.92 (89.50%)	1,152.47 (88.13%)
A posteriori bound	**1,519.67**	**1,497.78**	**1,465.15**	**1,422.60**	**1,369.76**	**1,307.63**
Heuristic policy cost	**1,526.46 (100.45%)**	**1,522.46 (101.65%)**	**1,510.92 (103.12%)**	**1,494.98 (105.09%)**	**1,477.46 (107.86%)**	**1,441.38 (110.23%)**
Fixed tour cost	1,530.00 (100.68%)	1,530.00 (102.15%)	1,530.00 (104.43%)	1,530.00 (107.55%)	1,530.00 (111.70%)	1,530.00 (117.01%)
			$\mathrm{P}(H) = 0.6$			
Optimistic bound	1,415.25 (93.42%)	1,300.50 (87.73%)	1,185.75 (82.92%)	1,071.00 (79.02%)	956.25 (75.90%)	841.50 (73.45%)
ALP bound	1,466.04 (96.78%)	1,389.87 (93.76%)	1,299.12 (90.85%)	1,194.60 (88.14%)	1,080.20 (85.74%)	959.50 (83.75%)
A posteriori bound	**1,514.894**	**1,482.37**	**1,429.93**	**1,355.38**	**1,259.83**	**1,145.67**
Heuristic policy cost	**1,524.69 (100.65%)**	**1,524.49 (102.84%)**	**1,510.47 (105.63%)**	**1,489.46 (109.89%)**	**1,432.64 (113.72%)**	**1,393.09 (121.60%)**
Fixed tour cost	1,530.00 (101.00%)	1,530.00 (103.21%)	1,530.00 (107.00%)	1,530.00 (112.88%)	1,530.00 (121.45%)	1,530.00 (133.55%)
			$\mathrm{P}(H) = 0.675$			
Optimistic bound	1,371.12 (90.65%)	1,212.23 (82.75%)	1,053.35 (76.09%)	894.46 (71.03%)	735.58 (67.06%)	576.69 (64.20%)
ALP bound	1,454.12 (96.13%)	1,350.42 (92.18%)	1,221.03 (88.21%)	1,059.72 (84.15%)	886.83 (80.85%)	705.43 (78.53%)
A posteriori bound	**1,512.61**	**1,465.00**	**1,384.27**	**1,259.30**	**1,096.94**	**898.25**
Heuristic policy cost	**1,527.32 (100.97%)**	**1,519.09 (103.69%)**	**1,493.90 (107.92%)**	**1,455.99 (115.62%)**	**1,378.30 (125.65%)**	**1,265.95 (140.93%)**
Fixed tour cost	1,530.00 (101.15%)	1,530.00 (104.44%)	1,530.00 (110.53%)	1,530.00 (121.50%)	1,530.00 (139.48%)	1,530.00 (170.33%)
			$\mathrm{P}(H) = 0.75$			
Optimistic bound	1,300.50 (86.46%)	1,071.00 (74.77%)	841.50 (65.71%)	612.00 (59.38%)	382.50 (55.32%)	153.00 (53.60%)
ALP bound	1,437.29 (95.56%)	1,286.90 (89.84%)	1,064.47 (83.12%)	800.80 (77.69%)	511.68 (74.00%)	205.98 (72.16%)
A posteriori bound	**1,504.11**	**1,432.42**	**1,280.70**	**1,030.71**	**691.46**	**285.44**
Heuristic policy cost	**1,526.59 (101.49%)**	**1,507.67 (105.25%)**	**1,458.24 (113.86%)**	**1,354.56 (131.42%)**	**1,163.57 (168.28%)**	**797.46 (279.38%)**
Fixed tour cost	1,530.00 (101.72%)	1,530.00 (106.81%)	1,530.00 (119.47%)	1,530.00 (148.44%)	1,530.00 (221.27%)	1,530.00 (536.02%)

Table A.4. Experiment results for ftv44 ($n = 44$) 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	1,532.35 (95.42%)	1,451.70 (91.39%)	1,371.05 (87.78%)	1,290.40 (84.68%)	1,209.75 (82.06%)	1,129.10 (79.88%)
ALP bound	1,550.82 (96.57%)	1,503.80 (94.67%)	1,442.55 (92.35%)	1,376.98 (90.36%)	1,307.02 (88.66%)	1,230.08 (87.02%)
A posteriori bound	**1,605.83**	**1,588.52**	**1,561.97**	**1,523.85**	**1,474.21**	**1,413.48**
Heuristic policy cost	1,634.46 (101.78%)	1,635.27 (102.94%)	**1,619.47 (103.68%)**	**1,609.62 (105.63%)**	**1,589.83 (107.84%)**	**1,557.18 (110.17%)**
Fixed tour cost	**1,631.00 (101.57%)**	**1,631.00 (102.67%)**	1,631.00 (104.42%)	1,631.00 (107.03%)	1,631.00 (110.64%)	1,631.00 (115.39%)
			$\mathsf{P}(H) = 0.6$			
Optimistic bound	1,492.03 (93.13%)	1,371.05 (87.13%)	1,250.08 (81.97%)	1,129.10 (77.47%)	1,008.13 (73.94%)	887.15 (71.14%)
ALP bound	1,542.13 (96.26%)	1,471.02 (93.49%)	1,380.95 (90.55%)	1,278.63 (87.73%)	1,159.59 (85.05%)	1,032.47 (82.80%)
A posteriori bound	**1,602.06**	**1,573.50**	**1,525.05**	**1,457.41**	**1,363.35**	**1,246.97**
Heuristic policy cost	1,635.20 (102.07%)	**1,625.96 (103.33%)**	**1,606.05 (105.31%)**	**1,570.25 (107.74%)**	**1,531.64 (112.34%)**	**1,485.02 (119.09%)**
Fixed tour cost	**1,631.00 (101.81%)**	1,631.00 (103.65%)	1,631.00 (106.95%)	1,631.00 (111.91%)	1,631.00 (119.63%)	1,631.00 (130.80%)
			$\mathsf{P}(H) = 0.675$			
Optimistic bound	1,445.50 (90.45%)	1,277.99 (82.05%)	1,110.49 (75.07%)	942.98 (69.46%)	775.48 (65.27%)	607.98 (62.28%)
ALP bound	1,533.22 (95.94%)	1,434.97 (92.13%)	1,307.76 (88.40%)	1,145.77 (84.40%)	961.63 (80.94%)	764.14 (78.27%)
A posteriori bound	**1,598.05**	**1,557.49**	**1,479.33**	**1,357.50**	**1,188.04**	**976.25**
Heuristic policy cost	1,636.57 (102.41%)	**1,624.08 (104.28%)**	**1,596.03 (107.89%)**	**1,554.67 (114.52%)**	**1,485.94 (125.07%)**	**1,371.46 (140.48%)**
Fixed tour cost	**1,631.00 (102.06%)**	1,631.00 (104.72%)	1,631.00 (110.25%)	1,631.00 (120.15%)	1,631.00 (137.28%)	1,631.00 (167.07%)
			$\mathsf{P}(H) = 0.75$			
Optimistic bound	1,371.05 (86.10%)	1,129.10 (74.17%)	887.15 (64.73%)	645.20 (57.96%)	403.25 (53.67%)	161.30 (51.86%)
ALP bound	1,517.40 (95.29%)	1,375.90 (90.38%)	1,152.58 (84.10%)	870.12 (78.17%)	554.52 (73.80%)	223.20 (71.76%)
A posteriori bound	**1,592.44**	**1,522.33**	**1,370.45**	**1,113.16**	**751.39**	**311.01**
Heuristic policy cost	1,637.68 (102.84%)	**1,614.17 (106.03%)**	**1,565.60 (114.24%)**	**1,447.39 (130.03%)**	**1,281.98 (170.62%)**	**923.55 (296.95%)**
Fixed tour cost	**1,631.00 (102.42%)**	1,631.00 (107.14%)	1,631.00 (119.01%)	1,631.00 (146.52%)	1,631.00 (217.07%)	1,631.00 (524.41%)

Table A.5. Experiment results for ftv33 ($n = 33$) with high/low correlated costs.

Instance: ftv33						
H	1.05	1.10	1.15	1.2	1.25	1.30
			$\mathsf{P}(H) = 0.5$			
Optimistic bound	1,221.70 (95.51%)	1,157.40 (90.97%)	1,093.10 (86.39%)	1,028.80 (81.75%)	964.50 (77.09%)	900.20 (72.49%)
ALP bound	1,279.12 (100.00%)	1,259.98 (99.04%)	1,233.54 (97.48%)	1,203.24 (95.62%)	1,171.33 (93.63%)	1,138.21 (91.66%)
A posteriori bound	**1,279.12**	**1,272.25**	**1,265.37**	**1,258.42**	**1,251.09**	**1,241.80**
Heuristic policy cost	**1,279.12 (100.00%)**	**1,272.25 (100.00%)**	**1,265.37 (100.00%)**	**1,258.50 (100.01%)**	**1,251.62 (100.04%)**	**1,244.74 (100.24%)**
Fixed tour cost	1,286.00 (100.54%)	1,286.00 (101.08%)	1,286.00 (101.63%)	1,286.00 (102.19%)	1,286.00 (102.79%)	1,286.00 (103.56%)
			$\mathsf{P}(H) = 0.6$			
Optimistic bound	1,189.55 (93.14%)	1,093.10 (86.19%)	996.65 (80.72%)	900.20 (72.06%)	803.75 (65.09%)	707.30 (58.14%)
ALP bound	1,274.37 (99.78%)	1,243.42 (98.04%)	1,203.07 (97.44%)	1,158.01 (92.70%)	1,108.08 (89.74%)	1,053.48 (86.59%)
A posteriori bound	**1,277.13**	**1,268.26**	**1,234.7375**	**1,249.22**	**1,234.74**	**1,216.56**
Heuristic policy cost	**1,277.13 (100.00%)**	**1,268.26 (100.00%)**	**1,259.39 (102.00%)**	**1,250.52 (100.10%)**	**1,241.65 (100.56%)**	**1,234.66 (101.49%)**
Fixed tour cost	1,286.00 (100.69%)	1,286.00 (101.40%)	1,286.00 (104.15%)	1,286.00 (102.94%)	1,286.00 (104.15%)	1,286.00 (105.71%)
			$\mathsf{P}(H) = 0.675$			
Optimistic bound	1,152.45 (90.27%)	1,018.91 (80.39%)	885.36 (70.39%)	751.82 (60.60%)	618.27 (50.79%)	484.72 (40.83%)
ALP bound	1,267.82 (99.30%)	1,224.09 (96.58%)	1,169.13 (92.95%)	1,102.11 (88.84%)	1,019.07 (83.71%)	907.50 (76.44%)
A posteriori bound	**1,276.74**	**1,267.48**	**1,257.74**	**1,240.59**	**1,217.41**	**1,187.19**
Heuristic policy cost	**1,276.74 (100.00%)**	**1,267.48 (100.00%)**	**1,258.22 (100.04%)**	**1,248.96 (100.67%)**	**1,241.08 (101.94%)**	**1,226.79 (103.34%)**
Fixed tour cost	1,286.00 (100.73%)	1,286.00 (101.46%)	1,286.00 (102.25%)	1,286.00 (103.66%)	1,286.00 (105.63%)	1,286.00 (108.32%)
			$\mathsf{P}(H) = 0.75$			
Optimistic bound	1,093.10 (85.73%)	900.20 (71.21%)	707.30 (56.78%)	514.40 (42.35%)	321.50 (27.48%)	128.60 (11.54%)
ALP bound	1,258.09 (98.67%)	1,195.09 (94.54%)	1,105.75 (88.76%)	968.11 (79.70%)	772.77 (66.05%)	489.57 (43.94%)
A posteriori bound	**1,275.05**	**1,264.09**	**1,245.78**	**1,214.68**	**1,170.04**	**1,114.28**
Heuristic policy cost	**1,275.05 (100.00%)**	**1,264.10 (100.00%)**	**1,253.80 (100.64%)**	**1,237.74 (101.90%)**	**1,224.35 (104.64%)**	**1,207.28 (108.35%)**
Fixed tour cost	1,286.00 (100.86%)	1,286.00 (101.73%)	1,286.00 (103.23%)	1,286.00 (105.87%)	1,286.00 (109.91%)	1,286.00 (115.41%)

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

Instance: ftv35						
H	1.05	1.10	1.15	1.2	1.25	1.30
			$P(H) = 0.5$			
Optimistic bound	1,399.35 (95.34%)	1,325.70 (90.72%)	1,252.05 (86.10%)	1,178.40 (81.48%)	1,104.75 (75.82%)	1,031.10 (72.26%)
ALP bound	1,441.72 (98.22%)	1,419.30 (97.13%)	1,394.35 (95.89%)	1,366.42 (94.48%)	1,333.07 (91.49%)	1,294.93 (90.74%)
A posteriori bound	**1,467.82**	**1,461.29**	**1,454.13**	**1,446.23**	**1,457.06**	**1,427.00**
Heuristic policy cost	**1,473.80 (100.41%)**	**1,466.46 (100.35%)**	**1,461.63 (100.52%)**	**1,456.80 (100.73%)**	**1,472.86 (101.08%)**	**1,447.23 (101.42%)**
Fixed tour cost	1,473.00 (100.35%)	1,473.00 (100.80%)	1,473.00 (101.30%)	1,473.00 (101.85%)	1,473.00 (101.09%)	1,473.00 (103.22%)
			$P(H) = 0.6$			
Optimistic bound	1,362.53 (92.94%)	1,252.05 (85.92%)	1,141.58 (78.92%)	1,031.10 (71.90%)	920.63 (64.84%)	810.15 (57.70%)
ALP bound	1,435.05 (97.89%)	1,403.20 (96.29%)	1,364.31 (94.32%)	1,315.44 (91.73%)	1,253.42 (88.28%)	1,182.27 (84.21%)
A posteriori bound	**1,465.99**	**1,457.25**	**1,446.48**	**1,434.09**	**1,419.76**	**1,404.02**
Heuristic policy cost	**1,470.45 (100.30%)**	**1,464.46 (100.50%)**	**1,458.48 (100.83%)**	**1,452.49 (101.28%)**	**1,446.50 (101.88%)**	**1,441.36 (102.66%)**
Fixed tour cost	1,473.00 (100.48%)	1,473.00 (101.08%)	1,473.00 (101.83%)	1,473.00 (102.71%)	1,473.00 (103.75%)	1,473.00 (104.91%)
			$P(H) = 0.675$			
Optimistic bound	1,320.03 (90.17%)	1,167.07 (80.35%)	1,014.10 (70.54%)	861.14 (60.67%)	708.17 (50.65%)	555.21 (40.44%)
ALP bound	1,428.30 (97.56%)	1,384.12 (95.30%)	1,325.64 (92.21%)	1,241.04 (87.44%)	1,137.29 (81.33%)	1,009.28 (73.51%)
A posteriori bound	**1,463.98**	**1,452.41**	**1,437.63**	**1,419.36**	**1,398.31**	**1,372.90**
Heuristic policy cost	**1,469.01 (100.34%)**	**1,460.94 (100.59%)**	**1,452.87 (101.06%)**	**1,444.80 (101.79%)**	**1,436.73 (102.75%)**	**1,427.80 (104.00%)**
Fixed tour cost	1,473.00 (100.62%)	1,473.00 (101.42%)	1,473.00 (102.46%)	1,473.00 (103.78%)	1,473.00 (105.34%)	1,473.00 (107.29%)
			$P(H) = 0.75$			
Optimistic bound	1,252.05 (85.78%)	1,031.10 (71.52%)	810.15 (57.15%)	589.20 (42.45%)	368.25 (27.28%)	147.30 (11.32%)
ALP bound	1,418.38 (97.18%)	1,352.28 (93.80%)	1,240.80 (87.53%)	1,080.29 (77.83%)	859.99 (63.70%)	539.08 (41.44%)
A posteriori bound	**1,459.55**	**1,441.67**	**1,417.64**	**1,388.10**	**1,350.13**	**1,300.72**
Heuristic policy cost	**1,465.81 (100.43%)**	**1,454.30 (100.88%)**	**1,442.79 (101.77%)**	**1,431.28 (103.11%)**	**1,417.01 (104.95%)**	**1,391.31 (106.96%)**
Fixed tour cost	1,473.00 (100.92%)	1,473.00 (102.17%)	1,473.00 (103.91%)	1,473.00 (106.12%)	1,473.00 (109.10%)	1,473.00 (113.24%)

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

Instance: ftv38						
H	1.05	1.10	1.15	1.2	1.25	1.30
			$P(H) = 0.5$			
Optimistic bound	1,453.50 (95.34%)	1,377.00 (90.70%)	1,300.50 (86.06%)	1,224.00 (81.42%)	1,147.50 (76.78%)	1,071.00 (72.16%)
ALP bound	1,499.04 (98.32%)	1,473.81 (97.08%)	1,445.99 (95.69%)	1,415.49 (94.16%)	1,379.98 (92.34%)	1,340.45 (90.31%)
A posteriori bound	**1,524.61**	**1,518.17**	**1,511.08**	**1,503.35**	**1,494.52**	**1,484.21**
Heuristic policy cost	**1,526.60 (100.13%)**	**1,521.20 (100.20%)**	**1,515.81 (100.31%)**	**1,509.53 (100.41%)**	**1,503.71 (100.62%)**	**1,497.89 (100.92%)**
Fixed tour cost	1,530.00 (100.35%)	1,530.00 (100.78%)	1,530.00 (101.25%)	1,530.00 (101.77%)	1,530.00 (102.37%)	1,530.00 (103.09%)
			$P(H) = 0.6$			
Optimistic bound	1,415.25 (92.88%)	1,300.50 (85.78%)	1,185.75 (78.68%)	1,071.00 (71.58%)	956.25 (64.47%)	841.50 (57.31%)
ALP bound	1,491.37 (97.88%)	1,455.65 (96.02%)	1,413.24 (93.77%)	1,361.93 (91.02%)	1,300.43 (87.67%)	1,227.81 (83.62%)
A posteriori bound	**1,523.70**	**1,516.04**	**1,507.12**	**1,496.32**	**1,483.26**	**1,468.32**
Heuristic policy cost	**1,525.70 (100.13%)**	**1,519.40 (100.22%)**	**1,513.10 (100.40%)**	**1,508.64 (100.82%)**	**1,502.40 (101.29%)**	**1,496.16 (101.90%)**
Fixed tour cost	1,530.00 (100.41%)	1,530.00 (100.92%)	1,530.00 (101.52%)	1,530.00 (102.25%)	1,530.00 (103.15%)	1,530.00 (104.20%)
			$P(H) = 0.675$			
Optimistic bound	1,371.12 (90.12%)	1,212.23 (80.23%)	1,053.35 (70.33%)	894.46 (60.37%)	735.58 (50.33%)	576.69 (40.17%)
ALP bound	1,483.36 (97.50%)	1,434.95 (94.97%)	1,372.87 (91.66%)	1,289.13 (87.01%)	1,183.59 (80.98%)	1,050.56 (73.17%)
A posteriori bound	**1,521.36**	**1,510.96**	**1,497.80**	**1,481.52**	**1,461.53**	**1,435.70**
Heuristic policy cost	**1,524.18 (100.19%)**	**1,516.37 (100.36%)**	**1,508.55 (100.72%)**	**1,501.92 (101.38%)**	**1,493.90 (102.22%)**	**1,486.50 (103.54%)**
Fixed tour cost	1,530.00 (100.57%)	1,530.00 (101.26%)	1,530.00 (102.15%)	1,530.00 (103.27%)	1,530.00 (104.69%)	1,530.00 (106.57%)
			$P(H) = 0.75$			
Optimistic bound	1,300.50 (85.65%)	1,071.00 (71.24%)	841.50 (56.73%)	612.00 (42.05%)	382.50 (27.00%)	153.00 (11.20%)
ALP bound	1,472.09 (96.95%)	1,402.12 (93.27%)	1,290.13 (86.98%)	1,124.61 (77.26%)	890.58 (62.86%)	554.86 (40.63%)
A posteriori bound	**1,518.36**	**1,503.33**	**1,483.24**	**1,455.54**	**1,416.67**	**1,365.78**
Heuristic policy cost	**1,521.70 (100.22%)**	**1,511.41 (100.54%)**	**1,501.11 (101.20%)**	**1,492.64 (102.55%)**	**1,480.63 (104.52%)**	**1,456.11 (106.61%)**
Fixed tour cost	1,530.00 (100.77%)	1,530.00 (101.77%)	1,530.00 (103.15%)	1,530.00 (105.12%)	1,530.00 (108.00%)	1,530.00 (112.02%)

Table A.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
			$\mathrm{P}(H)=0.5$			
ALP bound (total)	4,776.00	2,764.00	2,832.00	3,099.00	2,975.00	3,337.00
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
			$\mathrm{P}(H)=0.6$			
ALP bound (total)	3,779.00	2,654.00	3,243.00	3,386.00	3,268.00	3,843.00
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
			$\mathrm{P}(H)=0.675$			
ALP bound (total)	2,769.00	2,599.00	2,747.00	2,997.00	3,337.00	3,559.00
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
			$\mathrm{P}(H)=0.75$			
ALP bound (total)	2,145.00	2,146.00	2,740.00	3,421.00	3,975.00	4,382.00
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 A.9. Experiment running times in seconds for `ftv33` ($n=33$) with high/low correlated costs.

	Instance: ftv33					
H	1.05	1.10	1.15	1.2	1.25	1.30
			$\mathrm{P}(H)=0.5$			
ALP bound (total)	37,861.00	20,634.00	17,173.00	25,279.00	19,361.00	19,645.00
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
			$\mathrm{P}(H)=0.6$			
ALP bound (total)	37,979.00	32,703.00	27,398.00	13,857.00	19,631.00	21,297.00
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
			$\mathrm{P}(H)=0.675$			
ALP bound (total)	89,388.00	24,058.00	24,235.00	18,119.00	32,521.00	34,236.00
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
			$\mathrm{P}(H)=0.75$			
ALP bound (total)	29,677.00	29,147.00	31,555.00	33,493.00	39,247.00	28,277.00
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

problem (12) is equivalent to

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

for appropriately chosen $a_j, a_k \geqslant 0$ and b_k. This is a special case of a *submodular utility maximization problem* (Ahmed and Atamtürk 2011): Following Proposition 1 of this reference, we show that (17) is NP-hard by a reduction from the *partition problem* (Garey and Johnson 1979). 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 Experimental Results

This section contains full results for the experiments outlined in §5. Tables A.1 through A.4 have results for instances `ftv33`, `ftv35`, `ftv38` and `ftv44` with independently distributed costs. Tables A.5 through A.7 have results for instances `ftv33`, `ftv35` and `ftv38` with high/low correlated costs. Tables A.8 and A.9 give running times for the `ftv33` instances.

References

Adelman D (2003) Price-directed replenishment of subsets: Methodology and its application to inventory routing. *Manufacturing Service Oper. Management* 5(4):348–371.

Adelman D (2004) A price-directed approach to stochastic inventory/routing. *Oper. Res.* 52(4):499–514.

Adelman D (2007a) Dynamic bid prices in revenue management. *Oper. Res.* 55(4):647–661.

Adelman D (2007b) Price-directed control of a closed logistics queueing network. *Oper. Res.* 55(6):1022–1038.

Adelman D, Barz C (2014) A unifying approximate dynamic programming model for the economic lot scheduling problem. *Math. Oper. Res.* 39(2):374–402.

Adelman D, Klabjan D (2005) Duality and existence of optimal policies in generalized joint replenishment. *Math. Oper. Res.* 30(1):28–50.

Adelman D, Klabjan D (2011) Computing near-optimal policies in generalized joint replenishment. *INFORMS J. Comput.* 24(1):148–164.

Adelman D, Nemhauser GL (1999) Price-directed control of remnant inventory systems. *Oper. Res.* 47(6):889–898.

Ahmed S, Atamtürk A (2011) Maximizing a class of submodular utility functions. *Math. Programming* 128:149–169.

An H-C, Kleinberg R, Shmoys DB (2012) Improving christofides' algorithm for the *s-t* path TSP. *Proc. 44th ACM Sympos. Theory Comput. (STOC)* (Association for Computing Machinery), 875–886.

Applegate DL, Bixby RE, Chvátal V, Cook WJ (2006) *The Traveling Salesman Problem: A Computational Study* (Princeton University Press, Princeton, NJ).

Bellman R (1962) Dynamic programming treatment of the travelling salesman problem. *J. Assoc. Comput. Machinery* 9:61–63.

Bertsimas D, Demir R (2002) An approximate dynamic programming approach to multidimensional knapsack problems. *Management Sci.* 48(4):550–565.

Bertsimas D, Popescu I (2003) Revenue management in a dynamic network environment. *Transportation Sci.* 37(3):257–277.

Bertsimas D, Weismantel R (2005) *Optimization Over Integers* (Dynamic Ideas, Belmont, MA).

Bertsimas DJ (1992) A vehicle routing problem with stochastic demand. *Oper. Res.* 40(3):574–585.

Bertsimas DJ, van Ryzin G (1991) A stochastic and dynamic vehicle routing problem in the euclidean plane. *Oper. Res.* 39(4):601–615.

Bertsimas DJ, van Ryzin G (1993) Stochastic and dynamic vehicle routing in the euclidean plane with multiple capacitated vehicles. *Oper. Res.* 41(1):60–76.

Bertsekas DP (2012) *Dynamic Programming and Optimal Control: Approximate Dynamic Programming*, 4th ed., Vol. II (Athena Scientific, Belmont, MA).

Bertsekas DP, Castañón DA (1999) Rollout algorithms for stochastic scheduling problems. *J. Heuristics* 5:89–108.

Bertsekas DP, Tsitsiklis JN (1991) An analysis of stochastic shortest path problems. *Math. Oper. Res.* 16(3):580–595.

Campbell AM, Thomas BW (2008a) Challenges and advances in a priori routing. Golden B, Raghavan S, Wasil E, eds. *The Vehicle Routing Problem: Latest Advances and New Challenges*, (Springer, New York), 123–142.

Campbell AM, Thomas BW (2008b) Probabilistic traveling salesman problem with deadlines. *Transportation Sci.* 42(1):1–21.

Chen Z-L, Xu H (2006) Dynamic column generation for dynamic vehicle routing with time windows. *Transportation Sci.* 40(1):74–88.

Cheong T, White CC III (2012) Dynamic traveling salesman problem: Value of real-time traffic information. *IEEE Trans. Intelligent Transportation Systems* 13:619–630.

Christofides N (1976) Worst-case analysis of a new heuristic for the travelling salesman problem. Technical Report ADA025602, Graduate School of Industrial Administration, Carnegie Mellon University, Pittsburgh.

Cordeau J-F, Laporte G, Savelsbergh MWP, Vigo D (2007) Vehicle routing. Barnhart C, Laporte G, eds. *Handbook in Operations Research and Management Science: Transportation*, Chap. 6, Vol. 14 (Elsevier, Amsterdam), 367–428.

de Farias DP, van Roy B (2003) The linear programming approach to approximate dynamic programming. *Oper. Res.* 51(6):850–865.

de Farias DP, van Roy B (2004) On constraint sampling in the linear programming approach to approximate dynamic programming. *Math. Oper. Res.* 29(3):462–478.

Dean BC, Goemans MX, Vondrák J (2008) Approximating the stochastic knapsack problem: The benefit of adaptivity. *Math. Oper. Res.* 33(4):945–964.

Desai VV, Farias VF, Moallemi CC (2012) Approximate dynamic programming via a smoothed linear program. *Oper. Res.* 60(3):655–674.

Farias VF, Sauré D, Weintraub GY (2012) An approximate dynamic programming approach to solving dynamic oligopoly models. *RAND J. Econom.* 43:253–282.

Garey MR, Johnson DS (1979) *Computers and Intractability: A Guide to the Theory of NP-Completeness* (W. H. Freeman and Company, New York).

Gendreau M, Laporte G, Séguin R (1996) Stochastic vehicle routing. *Eur. J. Oper. Res.* 88:3–12.

Ghiani G, Manni E, Thomas BW (2012) A comparison of anticipatory algorithms for the dynamic and stochastic traveling salesman problem. *Transportation Sci.* 46(3):374–387.

Goberna MA, López MA (1998) *Linear Semi-Infinite Optimization*, Wiley Series in Mathematical Methods in Practice (John Wiley & Sons, Chichester, UK).

Godfrey GA, Powell WB (2002a) An adaptive dynamic programming algorithm for dynamic fleet management, I: Single period travel times. *Transportation Sci.* 36(1):21–39.

Godfrey GA, Powell WB (2002b) An adaptive dynamic programming algorithm for dynamic fleet management, II: Multiperiod travel times. *Transportation Sci.* 36(1):40–54.

Goemans MX, Bertsimas DJ (1991) Probabilistic analysis of the held and karp lower bound for the euclidean traveling salesman problem. *Math. Oper. Res.* 16(1):72–89.

Golden BL, Raghavan S, Wasil EA, eds. (2008) *The Vehicle Routing Problem: Latest Advances and New Challenges* (Springer, New York).

Goodson JC, Ohlmann JW, Thomas BW (2013) Rollout policies for dynamic solutions to the multi-vehicle routing problem with stochastic demand and duration limits. *Oper. Res.* 61(1):138–154.

Gutin G, Punnen AP, eds. (2002) *The Traveling Salesman Problem and Its Variations* (Kluwer Academic Publishers, Dordrecht, The Netherlands).

Hastie T, Tibshirani R, Friedman J (2009) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. (Springer, Berlin).

Held M, Karp RM (1962) A dynamic programming approach to sequencing problems. *J. Soc. Indust. Appl. Math.* 10:196–210.

Hochbaum DS, ed. (1997) *Approximation Algorithms for NP-Hard Problems* (PWS Publishing, Boston).

Jaillet P (1985) Probabilistic traveling salesman problems. Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Jaillet P (1988) A priori solution of a traveling salesman problem in which a random subset of the customers are visited. *Oper. Res.* 36(6):929–936.

Jaillet P, Wagner MR (2008) Online vehicle routing problems: A survey. Golden B, Raghavan S, Wasil E, eds. *The Vehicle Routing Problem: Latest Advances and New Challenges*, (Springer, New York), 221–237.

Jula H, Dessouky MM, Ioannu P (2006) Truck route planning in nonstationary stochastic networks with time-windows at customer locations. *IEEE Trans. Intelligent Transportation Systems* 37:51–63.

Kenyon AS, Morton DP (2003) Stochastic vehicle routing with random travel times. *Transportation Sci.* 37(1):69–82.

Kim S, Lewis ME, White CC III (2005a) Optimal vehicle routing with real-time traffic information. *IEEE Trans. Intelligent Transportation Systems* 6:178–188.

Kim S, Lewis ME, White CC III (2005b) State space reduction for nonstationary stochastic shortest path problems with real-time traffic information. *IEEE Trans. Intelligent Transportation Systems* 6:273–284.

Klabjan D, Adelman D (2007) An infinite-dimensional linear programming algorithm for deterministic semi-markov decision processes on borel spaces. *Math. Oper. Res.* 32(3):528–550.

Laporte G, Louveaux F, Mercure H (1992) The vehicle routing problem with stochastic travel times. *Transportation Sci.* 26(3):161–170.

Larsen A (2000) The dynamic vehicle routing problem. Unpublished doctoral dissertation, Technical University of Denmark, Kongens, Lyngby, Denmark.

Larsen A, Madsen OBG, Solomon MM (2004) The a priori dynamic traveling salesman problem with time windows. *Transportation Sci.* 38(4):459–472.

Larsen A, Madsen OBG, Solomon MM (2008) Recent developments in dynamic vehicle routing systems. Golden B, Raghavan S, Wasil E, eds.

The Vehicle Routing Problem: Latest Advances and New Challenges, (Springer, New York), 199–218.

Leipälä T (1978) On the solutions of stochastic traveling salesman problems. *Eur. J. Oper. Res.* 2:291–297.

Murthy I, Sarkar S (1998) Stochastic shortest path problems with piecewise-linear concave utility functions. *Management Sci.* 44(11, Part 2): S125–S136.

Nadarajah S, Margot F, Secomandi N (2011) Approximate dynamic programs for natural gas storage valuation based on approximate linear programming relaxations. Technical Report 2011-E5, Tepper School of Business, Carnegie Mellon University, Pittsburgh.

Papadimitriou CH, Yannakakis M (1991) Shortest paths without a map. *Theoret. Comput. Sci.* 84:127–150.

Patek SD, Bertsekas DP (1999) Stochastic shortest path games. *SIAM J. Control Optim.* 37:804–824.

Petrik M (2010) Optimization-based approximate dynamic programming. Ph.D. thesis, University of Massachusetts, Amherst, MA.

Petrik M, Zilberstein S (2011) Robust approximate bilinear programming for value function approximation. *J. Machine Learn. Res.* 12:3027–3063.

Polychronopoulos GH, Tsitsiklis JN (1996) Stochastic shortest path problems with recourse. *Networks* 27:133–143.

Powell WB (2010) *Approximate Dynamic Programming: Solving the Curses of Dimensionality*, 2nd ed. (John Wiley & Sons, Hoboken, NJ).

Provan JS (2003) A polynomial-time algorithm to find shortest paths with recourse. *Networks* 41:115–125.

Psaraftis HN (1995) Dynamic vehicle routing: Status and prospects. *Ann. Oper. Res.* 61:143–164.

Reinelt G (1991) TSPLIB—A traveling salesman problem library. *ORSA J. Comput.* 3:376–384. Updated archive available online at http://comopt.ifi.uni-heidelberg.de/software/TSPLIB95/.

Schrijver A (2003) The traveling salesman problem. *Combinatorial Optimization: Polyhedra and Efficiency*, Chap. 58, Vol. B (Springer, Berlin), 981–1004.

Schweitzer PJ, Seidmann A (1985) Generalized polynomial approximations in markovian decision processes. *J. Math. Anal. Appl.* 110:568–582.

Sebő A (2013) Eight-fifth approximation for TSP paths. Goemans M, Correa J, eds. *Proc. 16th Conf. Integer Programming and Combinatorial Optim.*, Lecture Notes in Computer Science, Vol. 7801 (Springer, Berlin, Heidelberg), 362–374.

Secomandi N (2001) A rollout policy for the vehicle routing problem with stochastic demands. *Oper. Res.* 49(5):796–802.

Secomandi N (2003) Analysis of a rollout approach to sequencing problems with stochastic routing applications. *J. Heuristics* 9:321–352.

Secomandi N, Margot F (2009) Reoptimization approaches for the vehicle-routing problem with stochastic demands. *Oper. Res.* 57(1):214–230.

Shapiro A, Dentcheva D, Ruszczyński AP (2009) *Lectures on Stochastic Programming: Modeling and Theory* (SIAM-MPS, Philadelphia).

Sungur I, Ordóñez F, Dessouky MM (2008) A robust optimization approach for the capacitated vehicle routing problem with demand uncertainty. *IIE Trans.* 40:509–523.

Thomas B, White CC III (2007) Dynamic shortest path problem with anticipation. *Eur. J. Oper. Res.* 176:836–854.

Toriello A (2014) Optimal toll design: A lower bound framework for the asymmetric traveling salesman problem. *Math. Programming*. 144:247–264.

Toriello A, Nemhauser G, Savelsbergh M (2010) Decomposing inventory routing problems with approximate value functions. *Naval Res. Logist.* 57:718–727.

Trick MA, Zin SE (1993) A linear programming approach to solving stochastic dynamic programs. Unpublished manuscript available online at http://mat.gsia.cmu.edu/trick/.

Trick MA, Zin SE (1997) Spline approximations to value functions: Alinear programming approach. *Macroeconomic Dynam.* 1:255–277.

Alejandro Toriello is an assistant professor in the H. Milton Stewart School of Industrial and Systems Engineering at Georgia Institute of Technology. His research interests include the theory and applications of supply chain and logistics models primarily involving transportation and inventory decisions, and in related optimization topics.

William B. Haskell earned his Ph.D. in Operations Research from the University of California, Berkeley. He is a post-doctoral research associate at the University of Southern California in the Computer Science and Electrical Engineering departments.

Michael Poremba is a Ph.D. candidate in the Daniel J. Epstein Department of Industrial and Systems Engineering at the University of Southern California. His research interests are in the area of supply chain management and transportation, including scheduling, routing, and inventory applications.