Optimal Toll Design: A Lower Bound Framework for the Asymmetric Traveling Salesman Problem

Alejandro Toriello
Daniel J. Epstein Department of Industrial and Systems Engineering
University of Southern California
3715 McClintock Avenue GER 240, Los Angeles, California 90089
toriello *at* usc *dot* edu
Phone: (213)740-4893 Fax: (213)740-1120

December 24, 2012

Abstract

We propose a framework of lower bounds for the asymmetric traveling salesman problem (TSP) based on approximating the dynamic programming formulation with different basis vector sets. We discuss how several well-known TSP lower bounds correspond to intuitive basis vector choices and give an economic interpretation wherein the salesman must pay tolls as he travels between cities. We then introduce an exact reformulation that generates a family of successively tighter lower bounds, all solvable in polynomial time. We show that the base member of this family yields a bound greater than or equal to the well-known Held-Karp bound, obtained by solving the linear programming relaxation of the TSP's integer programming arc-based formulation.

Keywords: traveling salesman problem, dynamic program, approximate linear program, integer program, lower bound technique

1 Introduction

The application of *dynamic programming* (DP) to routing problems dates back at least half a century. The *traveling salesman problem* (TSP) was already a notoriously difficult and well-studied discrete optimization problem in the early 1960's when three different articles [11, 24, 26] proposed dynamic programming formulations to solve it. DP still gives the best worst-case running time of any exact algorithm for the TSP; however, this running time is exponential because of the *curse of dimensionality*. In their text on the TSP [7], Applegate, et al, summarize their section on DP by noting:

> We will not treat further this dynamic programming approach to the TSP. Despite the nice worst-case time bound obtained by Held and Karp, the inherent growth in the practical running time of the method restricts its use, even today, to tiny instances of the problem.

Nevertheless, DP formulations have since been proposed for many other routing problems, such as the *vehicle routing problem* [13], the *dial-a-ride* problem [35] and various other extensions of the

TSP [18, 33]. Though not useful as exact algorithms except in special cases [10], these formulations serve as starting points for many heuristic algorithms [20, 32] and lower bounding procedures [13, 33].

The DP formulation of the TSP is a shortest path problem on a network where the number of nodes and arcs is exponential with respect to the number of cities. The *linear programming* (LP) dual of the shortest path formulation then yields the famous Bellman recursion. The exponential size of this LP prevents exact solutions, but one can compute valid dual bounds by solving a tractable restriction. Furthermore, any feasible dual solution represents an approximate *cost-to-go* which can also be exploited to generate primal solutions. This *approximate linear programming* (ALP) method for DP was proposed as early as the mid 1980's [37] and early 1990's [39, 40], and gained traction within the operations research community around the turn of the century, e.g. [1, 2, 4, 6, 15, 17, 21]. ALP has had success in various areas within operations management, such as inventory control [1, 2, 4, 5], commodity valuation [34] and revenue management [3, 21]. Some of these applications, e.g. inventory routing [1, 2], require the solution of many TSP sub-problems to calculate approximations, but in every application we have seen this computation is treated in a black-box fashion. To our knowledge, in fact, ALP has never been applied in a routing context.

The modeling, solution, application and analysis of the TSP form a vast body of work within operations research and mathematical programming that we will not attempt to treat here; the recent texts [7, 25, 36] cover many of these topics in detail. The TSP is defined by a finite set $\{0,\ldots,n\}$ of cities and a cost or distance vector $c_{ij} \in \mathbb{R}$ for $i,j = 0,\ldots,n$ with $i \neq j$. If $c_{ij} = c_{ji}$ for every pair of cities, the problem is *symmetric*; otherwise, it is *asymmetric*. We do not assume symmetry or any other restriction on c. For notational convenience, we single out the city 0 and use $N = \{1,\ldots,n\}$. The TSP's objective is to find a permutation $\sigma : N \to N$ that minimizes $c_{0,\sigma(1)} + \sum_{i=1}^{n-1} c_{\sigma(i),\sigma(i+1)} + c_{\sigma(n),0}$; i.e. the cost of starting at city 0, visiting the cities N in the order specified by σ, and returning to 0. The TSP can be modeled as a binary integer program:

$$\min \sum_{i\in N\cup 0} \sum_{j\in (N\cup 0)\setminus i} c_{ij}x_{ij} \tag{1a}$$

$$\text{s.t.} \sum_{j\in (N\cup 0)\setminus i} x_{ij} = 1,\ \forall\, i \in N\cup 0 \tag{1b}$$

$$\sum_{j\in (N\cup 0)\setminus i} x_{ji} = 1,\ \forall\, i \in N\cup 0 \tag{1c}$$

$$\sum_{i\in U} \sum_{j\in (N\cup 0)\setminus U} x_{ji} \geq 1,\ \forall\, \varnothing \neq U \subsetneq N\cup 0 \tag{1d}$$

$$x_{ij} \geq 0,\ x_{ij} \in \mathbb{Z},\ \forall\, i \in N\cup 0,\ j \in (N\cup 0)\setminus i. \tag{1e}$$

Here and elsewhere we identify an element i with its singleton set $\{i\}$. This formulation is sometimes called *arc-based*, to emphasize that the solution is represented by a set of arcs that form a *tour*. Constraints (1b) require that the tour depart each city exactly once, while constraints (1c) require the tour to enter each city exactly once. The *subtour elimination* constraints (1d) require the tour to leave each proper, non-empty subset of cities at least once, ensuring that the solution is connected.

The DP formulation for the TSP [11, 24, 26] is based on the following simple observation: Given that one is at city i, the only additional information needed to choose the next city to visit is the subset of N that has not been visited yet. (Equivalently, one can consider the complementary subset of cities already visited, which includes i.) The set $S = \{(i,U) : i \in N, U \subseteq N \setminus i\} \cup \{(0,N),(0,\varnothing)\}$ denotes all possible states, and

$$\begin{aligned} A = \{((0,N),(i,N\setminus i)) : i \in N\} \cup \{((i,\varnothing),(0,\varnothing)) : i \in N\} \\ \cup \{((i,U\cup j),(j,U)) : i \in N,\ j \in N\setminus i,\ U \subseteq N \setminus \{i,j\}\} \end{aligned}$$

denotes all possible transitions from one state to another. The cost of any action $((i,U\cup j),(j,U)) \in A$ is then c_{ij}, and the solution of the TSP is given by the shortest path between $(0,N)$ and $(0,\varnothing)$ in the directed network (S,A):

$$\min \sum_{i\in N}\Bigg(c_{0i}x_{(0,N),(i,N\setminus i)} + \sum_{\varnothing\neq U\subseteq N\setminus i}\sum_{j\in U} c_{ij}x_{(i,U),(j,U\setminus j)} + c_{i0}x_{(i,\varnothing),(0,\varnothing)}\Bigg) \tag{2a}$$

$$\text{s.t.} \sum_{i\in N} x_{(0,N),(i,N\setminus i)} = 1 \tag{2b}$$

$$x_{(0,N),(i,N\setminus i)} - \sum_{j\in N\setminus i} x_{(i,N\setminus i),(j,N\setminus\{i,j\})} = 0,\ \forall\, i \in N \tag{2c}$$

$$\sum_{k\in N\setminus(U\cup i)} x_{(k,U\cup i),(i,U)} - \sum_{j\in U} x_{(i,U),(j,U\setminus j)} = 0,\ \forall i \in N,\ \varnothing \neq U \subsetneq N \setminus i \tag{2d}$$

$$\sum_{k\in N\setminus i} x_{(k,i),(i,\varnothing)} - x_{(i,\varnothing),(0,\varnothing)} = 0,\ \forall\, i \in N \tag{2e}$$

$$\sum_{i\in N} x_{(i,\varnothing),(0,\varnothing)} = 1 \tag{2f}$$

$$x_a \geq 0,\ \forall\, a \in A. \tag{2g}$$

The LP dual of the shortest path formulation (2) is

$$\max\ y_{0,N} - y_{0,\varnothing} \tag{3a}$$

$$\text{s.t.}\ y_{0,N} - y_{i,N\setminus i} \leq c_{0i},\ \forall\, i \in N \tag{3b}$$

$$y_{i,U\cup j} - y_{j,U} \leq c_{ij},\ \forall\, i \in N,\ j \in N \setminus i,\ U \subseteq N \setminus \{i,j\} \tag{3c}$$

$$y_{i,\varnothing} - y_{0,\varnothing} \leq c_{i0},\ \forall\, i \in N \tag{3d}$$

$$y_{0,N}, y_{0,\varnothing} \in \mathbb{R};\quad y_{i,U} \in \mathbb{R},\ \forall i \in N,\ U \subseteq N \setminus i. \tag{3e}$$

This polyhedron contains the line defined by the vector of all ones; the optimal solution y^* with $y^*_{0,\varnothing} = 0$ yields the familiar DP *backwards recursion*:

$$y^*_{i,U} = \begin{cases} \min_{j\in U}\{c_{ij} + y^*_{j,U\setminus j}\}, & U \neq \varnothing \\ c_{i0}, & U = \varnothing,\ i \neq 0, \end{cases}$$

where $y^*_{i,U}$ represents the cost-to-go from state (i,U) to the end, $(0,\varnothing)$. A formulation that zeroes the $y_{0,N}$ variable instead yields the *forward recursion.* Throughout the remainder of the paper, we assume without loss of generality that $y_{0,\varnothing} = 0$.

The goal of this paper is to apply ALP techniques to (3) and obtain tractable lower bounds for the asymmetric TSP. We see the main contributions in two categories:

i) We formulate a general lower bound framework for the TSP that approximates the formulation (3). We show how several well-known lower bounds for the TSP occur as special cases within this framework, how they correspond to intuitive choices of the approximating basis vectors, and how the approximation gives a new "tolling" economic interpretation of the bounds.

ii) We give an exact reformulation of (3) that generates a new family of successively tighter polynomially-solvable approximations, and show that the base member of this family produces a bound at least as good as the well-known *Held-Karp bound* [27, 28] given by the LP relaxation of (1).

From an integer programming perspective, (2) and (3) can be viewed as extended formulations. In this context, our results concern *approximate extended formulations* [41], but our approximate formulations operate in the dual space, use only continuous variables, and do not converge to an exact formulation. Our framework does not include semidefinite lower bounding techniques for the symmetric TSP. We refer the reader to [14, 16] and references therein for details on the subject.

The remainder of the paper is organized as follows: Section 2 details the general ALP lower bound framework and the previously studied lower bounds that occur as special cases. Section 3 gives an exact reformulation of (3), introduces the new bound family, and compares it to the Held-Karp bound. Section 4 concludes and outlines future research avenues. Throughout the paper, we use x variables to represent primal solutions; e.g. $x_{ij} = 1$ if city i immediately precedes city j in a solution. We use y as dual variables for x. We use λ, μ as basis vector multipliers and b to represent a basis. The letters U, W represent subsets of N, and $e_U \in \{0,1\}^N$ and 2^U represent the characteristic vector and power set of U respectively. We use $e_{i,j}$ to denote a matrix of all zeroes except for a one in the (i, j)-th entry, where dimensions are specified or clear from the context.

2 Approximate Linear Program

In ALP, the restriction of the dual LP (3) is achieved by means of a set of basis vectors. Specifically, let $\lambda \in \mathbb{R}^m$, where usually $m \lll |S|$, and define a collection $b_{i,U} \in \mathbb{R}^m$ for $i \in N, U \subseteq N \setminus i$. Then we can set $y_{i,U} = \langle b_{i,U}, \lambda \rangle$ for every dual variable corresponding to an intermediate node in the network (S, A), and the resulting dual solution is feasible provided it satisfies the constraints in (3). For state (i, U), $\langle b_{i,U}, \lambda \rangle$ is also an approximate cost-to-go. If we define a matrix b with rows $b_{i,U}^{\mathsf{T}}$ indexed by pairs (i, U) for each $i \in N$ and $U \subseteq N \setminus i$, its column space is a subspace of the dual space of variables $y_{i,U}$, and its columns are basis vectors. The best lower bound achievable with this choice of m and b is given by the solution to the approximate LP

$$\max \; y_{0,N} \tag{4a}$$
$$\text{s.t. } y_{0,N} - \langle b_{i,N\setminus i}, \lambda \rangle \le c_{0i}, \; \forall \, i \in N \tag{4b}$$
$$\langle (b_{i,U\cup j} - b_{j,U}), \lambda \rangle \le c_{ij}, \; \forall \, i \in N, \; j \in N \setminus i, \; U \subseteq N \setminus \{i, j\} \tag{4c}$$
$$\langle b_{i,\varnothing}, \lambda \rangle \le c_{i0}, \; \forall \, i \in N \tag{4d}$$
$$y_{0,N} \in \mathbb{R}; \quad \lambda \in \mathbb{R}^m. \tag{4e}$$

Sometimes it is convenient to consider λ and $b_{i,U}$ as matrices; in this case, $\langle b_{i,U}, \lambda \rangle$ represents the Frobenius inner product. In particular, it is sometimes natural to consider $\lambda \in \mathbb{R}^{m \times 2}$; for this situation we instead use (λ, μ), where $\lambda, \mu \in \mathbb{R}^m$.

The ALP (4) has an intuitive interpretation that motivates particular basis vector choices: Suppose a company operates a tolled shuttle service between the cities $N \cup 0$ that the traveling salesman can utilize in lieu of his own personal transportation. The toll between two cities, $y_{i,U\cup j} - y_{j,U}$, can depend not only on the origin and destination, but also on the remaining cities the salesman needs to visit before returning to city 0. However, no toll can exceed the salesman's own travel cost for a pair of cities, because otherwise he would not utilize the service for this leg of the journey. The company would ideally like to solve (3); however, for computational reasons they consider instead approximations given by m and b. The next sections detail how this choice of m and b can yield well-known TSP lower bounds.

2.1 State Space Relaxation

In [13], Christofides, et al, introduced the notion of a *state space relaxation* for general DP, and explored various specific relaxations for the TSP and other routing variants (see also [33]). We summarize their pertinent results for the TSP below.

Lemma 1 ([13]). *Let $g : 2^N \to G$, where $|G| \leq 2^n$. For $\gamma \in G$, let $E(i, \gamma) = \{j \in N \setminus i : \exists\, U \ni j,\ U \not\ni i,\ g(U) = \gamma\}$. The DP with states $(i, g(U))$ and action sets $E(i, g(U))$ provides a lower bound $f(0, g(N))$ for* (3) *via the recursion*

$$f(i, g(U)) = \begin{cases} \min_{j \in E(i,g(U))}\{c_{ij} + f(j, g(U \setminus j))\}, & U \neq \varnothing \\ c_{i0}, & U = \varnothing,\ i \neq 0. \end{cases} \tag{5}$$

A similar lower bound is achievable using the forward recursion instead. Because the state space relaxation is itself a DP, it fits into the framework provided by (4) in a straightforward manner.

Proposition 2. *Let $\lambda \in \mathbb{R}^{N\times G}$, and let $b_{i,U} = e_{i,g(U)} \in \{0,1\}^{N\times G}$. Then the optimal value of* (4) *is the lower bound $f(0, g(N))$.*

Proof. Substituting this b into (4), it is simple to check that the resulting LP is in fact the dual LP formulation of the state space relaxation DP with backward recursion (5). □

This proposition shows that the state space relaxation bound $f(0, g(N))$ results from assuming that all subsets $U \subseteq N \setminus i$ with equal $g(U)$ should be valued equally when the salesman is at city i; this approximate cost-to-go is $\lambda_{i,g(U)}$. The next example gives one intuitive choice of g.

Example 3 ([29]). Let $g(U) = |U| + 1$ and $E(i, U) = N \setminus i$, so $b_{i,U} = e_{i,|U|+1} \in \{0,1\}^{N\times n}$. The approximate cost-to-go is $\lambda_{i,|U|+1}$, implying that from a particular city i, the basis assigns the same cost to all subsets of the same cardinality. In this case, the formulation (4) is

$$\begin{aligned} \max\ & y_{0,N} \\ \text{s.t. } & y_{0,N} - \lambda_{i,n} \leq c_{0i},\ \forall\, i \in N \\ & \lambda_{i,k+1} - \lambda_{j,k} \leq c_{ij},\ \forall\, i \in N,\ j \in N \setminus i,\ k = 1, \dots, n-1 \\ & \lambda_{i,1} \leq c_{i0},\ \forall\, i \in N \\ & \lambda \in \mathbb{R}^{N\times n}, \end{aligned}$$

and the LP dual is

$$\begin{aligned}
\min\ & \sum_{i\in N}\Bigg(c_{0i}x_{0i} + \sum_{j\in N\setminus i} c_{ij}\sum_{k=1}^{n-1} x^k_{ij} + c_{i0}x_{i0}\Bigg)\\
\text{s.t. } & \sum_{i\in N} x_{0i} = 1\\
& -x_{0i} + \sum_{j\in N\setminus i} x^{n-1}_{ij} = 0,\ \forall\, i\in N\\
& \sum_{j\in N\setminus i}(x^k_{ij} - x^{k+1}_{ji}) = 0,\ \forall\, i\in N,\ k=1,\dots,n-2\\
& -\sum_{j\in N\setminus i} x^1_{ji} + x_{i0} = 0,\ \forall\, i\in N\\
& x_{0i}, x_{i0} \ge 0,\ \forall\, i\in N; \quad x^k_{ij}\ge 0,\ \forall\, i\in N,\ j\in N\setminus i,\ k=1,\dots,n-1.
\end{aligned}$$

The optimal solution of this LP is the minimum cost $(n+1)$-step closed walk that starts and ends at city 0 and doesn't visit it otherwise. Returning to the tolling interpretation, the toll given by this basis, $\lambda_{i,k+1} - \lambda_{j,k}$, depends on the starting and ending city, and also on the number of cities the salesman has left to visit.

Of course, tighter lower bounds result from more complex choices of g. We refer the reader to [13] for details.

2.2 Matroid Bounds and the LP Relaxation

The state space relaxation bounds stem from assigning a single equal value to different states; i.e. from a particular city, we value different remaining sets of cities equally via g. We next consider some additive approximations.

Example 4. Let $\lambda\in\mathbb{R}^N$, $b_{i,U} = e_{U\cup i}\in\{0,1\}^N$. This basis choice assigns $y_{i,U} = \sum_{j\in U\cup i}\lambda_j$; that is, the cost-to-go assigns individual costs to each remaining city plus the current location. Thus, when the salesman is at city i, there is an exit toll λ_i, regardless of the remaining cities to visit or the chosen destination.

Letting $\lambda_0 = y_{0,N} - \sum_{i\in N}\lambda_i$, the formulation (4) under this choice of b becomes

$$\begin{aligned}
\max\ & \sum_{i\in N\cup 0}\lambda_i\\
\text{s.t. } & \lambda_i \le c_{ij},\ \forall\, i\in N\cup 0,\ j\in (N\cup 0)\setminus i\\
& \lambda\in\mathbb{R}^{N\cup 0}.
\end{aligned}$$

The optimal solution is $\lambda^*_i = \min_{j\in(N\cup 0)\setminus i} c_{ij}$, the minimum-cost basis of the partition matroid given by the forward stars of $N\cup 0$. The dual is the LP relaxation of (1) keeping only constraints (1b).

The bound given by this example is very weak. One reason is the basis' inability to account for the salesman's current location. The next proposition resolves this issue.

Proposition 5. *Let* $\lambda \in \mathbb{R}^N$*,* $\mu \in \mathbb{R}^{N\cup 0}$*, and* $b_{i,U} = (e_{U\cup i}, e_{U\cup 0})$*, where* $e_{U\cup i} \in \{0,1\}^N$*,* $e_{U\cup 0} \in \{0,1\}^{N\cup 0}$*. The optimal value of* (4) *is then the minimum cost of a basis in the intersection of the two partition matroids given by the forward and backward stars of* $N \cup 0$*; i.e. the* degree relaxation*, or* assignment *bound.*

The approximate cost-to-go is now $y_{i,U} = \lambda_i + \mu_0 + \sum_{j\in U}(\lambda_j + \mu_j)$, and the toll to go from i to j is $\lambda_i + \mu_j$, regardless of other remaining cities. In other words, the salesman must pay an entrance toll μ_i and exit toll λ_i at every city i.

Proof. Rewriting $y_{0,N} = \sum_{i\in N\cup 0}(\lambda_i + \mu_i)$ with the new variable λ_0, formulation (4) becomes

$$\begin{aligned} \max \quad & \sum_{i\in N\cup 0} (\lambda_i + \mu_i) \\ \text{s.t.} \quad & \lambda_i + \mu_j \leq c_{ij}, \ \forall \ i \in N \cup 0, \ j \in (N \cup 0) \setminus i \\ & \lambda, \mu \in \mathbb{R}^{N\cup 0}, \end{aligned}$$

The dual of this formulation is the LP relaxation of (1) keeping only constraints (1b) and (1c), precisely the degree relaxation. □

This proposition's basis reduces to a well-known special case when the TSP instance satisfies additional conditions, explained next.

Definition 6 ([8])**.** A TSP instance is *geometric* if each city i corresponds to a point $z_i \in \mathbb{R}^q$, and the travel costs are symmetric and satisfy $c_{ij} = c_{ji} = \|z_i - z_j\|_p$, where $\|\cdot\|_p$ is the ℓ_p norm for some $p \in \mathbb{N} \cup \infty$.

Jünger and Pulleyblank [30] introduced the notion of *geometric duality* for combinatorial optimization problems. For the geometric TSP, Proposition 5's formulation is their *control zone* bound: Because the instance is symmetric, we take $\lambda_i = \mu_i, \forall \ i \in N \cup 0$, and the optimal solution represents the radii of ℓ_p-balls centered at each z_i with pairwise non-intersecting interiors and maximum sum. The ALP (4) gives an additional tolling interpretation of their bound that extends to the non-geometric and non-symmetric case. Figure 1 gives an example with $p = q = 2$ and $n = 3$.

Another way to strengthen the bound given by Example 4 is to consider tolls for sets of cities in addition to individual city tolls. The next basis explores this idea; however, we first need an additional definition.

Definition 7 ([27])**.** A 0-*arborescence* is a directed graph with vertex set $N \cup 0$ where exactly one arc is directed into each city and there is exactly one directed cycle, which includes 0.

As Held and Karp note in their classical papers [27, 28], any TSP tour is a 0-arborescence, and therefore the minimum cost of a 0-arborescence is a lower bound on (1).

Proposition 8. *Let* $\mu \in \mathbb{R}^{2^N\setminus N}$ *and* $b_{i,U} = e_{2^U} \in \{0,1\}^{2^N\setminus N}$*. If we impose the additional constraints* $\mu_U \geq 0$ *for* $|U| \geq 2$*, the optimal value of* (4) *is the minimum cost of a* 0*-arborescence.*

This basis approximates the cost-to-go using only the set of remaining cities: $y_{i,U} = \sum_{W\subseteq U} \mu_W$, implying there is an entrance toll not only for cities, but also for sets of cities (we use μ here to maintain notational consistency with the other bases).

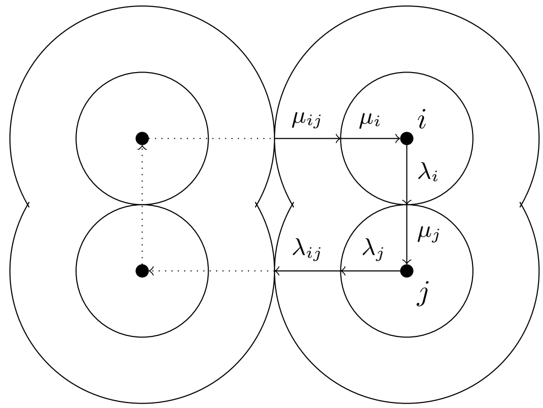

Figure 1: An example of geometric duality for the TSP with dual variables representing tolls.

Proof. Rewriting $y_{0,N} = \sum_{U \subseteq N} \mu_U$, where we include a new unrestricted variable μ_N, (4) becomes

$$\begin{aligned} \max \quad & \sum_{U \subseteq N} \mu_U \\ \text{s.t.} \quad & \sum_{U \subseteq N \setminus i} \mu_{U \cup i} \leq c_{0i}, \ \forall\, i \in N \\ & \sum_{U \subseteq N \setminus \{i,j\}} \mu_{U \cup j} \leq c_{ij}, \ \forall\, i \in N, \ j \in N \setminus i \\ & \mu_\varnothing \leq c_{i0}, \ \forall\, i \in N \\ & \mu_U \geq 0, \ \forall\, 2 \leq |U| \leq n-1 \\ & \mu \in \mathbb{R}^{2^N}, \end{aligned}$$

where the second set of constraints implies all constraints (4c) because of non-negativity. The dual is the LP relaxation of (1) keeping all constraints (1c), constraint (1b) only for city 0, and constraints (1d) for $U \subseteq N$. The integer vectors satisfying these constraints are precisely the characteristic vectors of 0-arborescences.

Moreover, the set of 0-arborescences is the set of bases in the intersection of two matroids: The partition matroid given by backward stars of $N \cup 0$, and the modified graphic matroid on $(N \cup 0)^2$ that allows at most one cycle that must include 0. By the Matroid Intersection Theorem [19], the polyhedron defined by these constraints is integral, and therefore we can take the optimal solution to be the characteristic vector of a 0-arborescence. □

We can combine the concepts from Propositions 5 and 8 to yield the following result.

Theorem 9. *Let $\lambda \in \mathbb{R}^{2^N \setminus \varnothing}$ and $\mu \in \mathbb{R}^{2^N \setminus N}$. Let $b_{i,U} = (e_{2^N \setminus (2^{N \setminus (U \cup i)})}, e_{2^U})$, with $e_{2^N \setminus (2^{N \setminus (U \cup i)})} \in \{0,1\}^{2^N \setminus \varnothing}$ and $e_{2^U} \in \{0,1\}^{2^N \setminus N}$. Suppose we impose the additional constraints $\lambda_U, \mu_U \geq 0$, $\forall\, 2 \leq |U| \leq n-1$ and replace constraints (4c) with*

$$\sum_{U \subseteq N \setminus \{i,j\}} (\lambda_{U \cup i} + \mu_{U \cup j}) \leq c_{ij}, \ \forall\, i \in N, \ j \in N \setminus i \tag{6}$$

in the formulation (4). *Then the optimal solution is feasible for* (4) *and the optimal value is that of the LP relaxation of the arc-based formulation* (1), *the Held-Karp bound.*

The approximate cost-to-go is

$$y_{i,U} = \sum_{(U \cup i) \not\subseteq W \subseteq N} \lambda_{N \setminus W} + \sum_{W \subseteq U} \mu_W;$$

there is an entrance toll μ_U and exit toll λ_U for every subset of cities $U \subseteq N$.

Proof. For this basis, constraints (4c) become

$$\sum_{W \subseteq N \setminus (U \cup \{i,j\})} \lambda_{W \cup i} + \sum_{W \subseteq U} \mu_{W \cup j} \leq c_{ij}, \forall\, i \in N, j \in N \setminus i, U \subseteq N \setminus \{i,j\},$$

and they are implied by constraints (6) together with non-negativity, so any feasible solution of this modified LP is feasible for the original (4). As in previous proofs, define additional unrestricted decision variables $\lambda_\varnothing$ and μ_N, and set $y_{0,N} = \sum_{U \subseteq N}(\lambda_U + \mu_U)$. The formulation is then rewritten as

$$\max \sum_{U \subseteq N} (\lambda_U + \mu_U) \tag{7a}$$

$$\text{s.t. } \lambda_\varnothing + \sum_{U \subseteq N \setminus i} \mu_{U \cup i} \leq c_{0i}, \ \forall\, i \in N \tag{7b}$$

$$\sum_{U \subseteq N \setminus \{i,j\}} (\lambda_{U \cup i} + \mu_{U \cup j}) \leq c_{ij}, \ \forall\, i \in N, \ j \in N \setminus i \tag{7c}$$

$$\sum_{U \subseteq N \setminus i} \lambda_{U \cup i} + \mu_\varnothing \leq c_{i0}, \ \forall\, i \in N \tag{7d}$$

$$\lambda_U, \mu_U \geq 0, \ \forall\, 2 \leq |U| \leq n-1 \tag{7e}$$

$$\lambda, \mu \in \mathbb{R}^{2^N}. \tag{7f}$$

This is exactly the dual of the LP relaxation of (1): Constraints (1b) correspond to λ_U with $|U| \leq 1$, (1c) correspond to μ_U with $|U| \leq 1$, and constraints (1d) with $U \ni 0$ and $U \not\ni 0$ correspond to $\lambda_{N \setminus U}$ and μ_U respectively. □

As with the optimal 0-arborescence, it is well known that the LP relaxation of (1) (and thus (4) with this basis) can be solved in polynomial time [7, 36].

When the TSP instance is geometric, this basis and formulation correspond to Jünger and Pulleyblank's [30] control zone and *moat* bound: By symmetry, we take $\lambda_U = \mu_U$, and the optimal solution represents the radii of ℓ_p-balls around each city (for U empty or a singleton) and the radii of ℓ_p moats around sets of cities (for $|U| \geq 2$). Figure 1 shows an example of a two-city moat with its associated dual variables.

Table 1 summarizes this section's results. For each basis, it indicates the basis dimension, the actual basis values, and the corresponding bounds for both the general TSP as well as the geometric TSP.

m	$b_{i,U}$	General TSP Bound	Geometric TSP Bound
n	$e_{U \cup i}$	forward star	
$2n+1$	$(e_{U \cup i}, e_{U \cup 0})$	degree relaxation	control zone [30]
n^2	$e_{i,\lvert U \rvert+1}$	$(n+1)$-step walk [29]	
$n \times \lvert G \rvert$	$e_{i,g(U)}$	state space relaxation [13]	
$2^n - 1$	$e_{2^U}{}^\star$	0-arborescence [27, 28]	
$2^{n+1} - 2$	$(e_{2^N \setminus (2^{N \setminus (U \cup i)})}, e_{2^U})^\star$	LP relaxation of (1) [27]	control zone and moat [30]

Table 1: TSP bounds for various basis choices. A star next to the basis indicates that additional constraints are applied to the ALP (4).

3 A New Lower Bound Family

Our exact reformulation of (3) is based on the following basic linear algebra fact.

Lemma 10. *The column space of the matrix* $(e_{2^U}^\intercal)_{U \subseteq N}$ *is* $\mathbb{R}^{2^N}$*.*

Proof. Let $y \in \mathbb{R}^{2^N}$ and define $\lambda \in \mathbb{R}^{2^N}$ recursively as $\lambda_\varnothing = y_\varnothing$ and $\lambda_U = y_U - \sum_{W \subsetneq U} \lambda_W$, $\forall\, \varnothing \neq U \subseteq N$. Then $y_U = \sum_{W \subseteq U} \lambda_U = \langle e_{2^U}, \lambda \rangle$, $\forall\, U \subseteq N$. □

Corollary 11. *Let* $\lambda_i \in \mathbb{R}^{2^{N \setminus i}}$*,* $\forall\, i \in N$ *and consider* $\lambda = (\lambda_i^\intercal)_{i \in N}$ *as an* $n \times 2^{n-1}$ *matrix. Let* $b_{i,U} \in \{0,1\}^{n \times 2^{n-1}}$ *be a matrix with all zero rows, except for the* i*-th row, equal to* $e_{2^U}^\intercal$*. Then* (4) *is equivalent to* (3)*; that is, the problems have equal optimal values and a one-to-one correspondence between feasible solutions.*

Remark 12. Another equivalent formulation arises from rewriting a variable $y \in \mathbb{R}^{2^N}$ in terms of supersets rather than subsets: $y_U = \sum_{W \supseteq U} \mu_{N \setminus W} = \sum_{W \subseteq N \setminus U} \mu_W = \langle e_{2^{N \setminus U}}, \mu \rangle$. In this case the basis $b_{i,U} \in \{0,1\}^{n \times 2^{n-1}}$ consists of a matrix with all zero rows except for the i-th one, equal to $e_{2^{N \setminus (U \cup i)}}^\intercal$.

Neither basis' formulation is useful in its entirety, but they generate a family of approximations.

Definition 13. For each $i \in N$, let $\mathcal{U}_i^+, \mathcal{U}_i^- \subseteq 2^{N \setminus i}$ and $\mathcal{U} = (\mathcal{U}_i^+, \mathcal{U}_i^-)_{i \in N}$. Define as $P(\mathcal{U})$ the problem (4) with decision variables $\lambda, \mu \in \mathbb{R}^{n \times 2^{n-1}}$ (considered as matrices) and $b_{i,U} \in \{0,1\}^{n \times 2^n}$ equal to a matrix with all zero rows except for the i-th row, which is $\left(e_{2^U \cap \mathcal{U}_i^+}^\intercal, e_{2^{N \setminus (U \cup i)} \cap \mathcal{U}_i^-}^\intercal\right)$.

This basis yields the approximation

$$y_{i,U} = \sum_{\substack{W \subseteq U \\ W \in \mathcal{U}_i^+}} \lambda_{i,W} + \sum_{\substack{W \subseteq N \setminus (U \cup i) \\ W \in \mathcal{U}_i^-}} \mu_{i,W}.$$

The state (i, U) is valued using subsets of U (the remaining cities) in the set $\mathcal{U}_i^+$, and subsets of $N \setminus (U \cup i)$ (the previously visited cities) in $\mathcal{U}_i^-$.

Example 14. Let $\mathcal{U}_i^+ = \{\varnothing\}$, $\mathcal{U}_i^- = \varnothing$, $\forall\ i \in N$. Then $P(\mathcal{U})$ is

$$\begin{aligned}\max\ & y_{0,N}\\ \text{s.t. } & y_{0,N} - \lambda_{i,\varnothing} \le c_{0i},\ \forall\ i \in N\\ & \lambda_{i,\varnothing} - \lambda_{j,\varnothing} \le c_{ij},\ \forall\ i \in N,\ j \in N \setminus i\\ & \lambda_{i,\varnothing} \le c_{i0},\ \forall\ i \in N\\ & y_{0,N} \in \mathbb{R}; \quad \lambda \in \mathbb{R}^N.\end{aligned}$$

The optimal value of this LP is the minimum cost of a circulation that forces one unit of flow through city 0.

The next theorem gives our first main result.

Theorem 15. *Fix $t \in \mathbb{Z}$, $0 \le t \le \frac{n-1}{2}$, and define*

$$\begin{aligned}\mathcal{U}_{i,t}^+ = \mathcal{U}_{i,t}^- &= \{U \subseteq N \setminus i : |U| \ge n-t\} \cup \{U \subseteq N \setminus i : |U| \le 1\}\\ \mathcal{U}_t &= (\mathcal{U}_{i,t}^+, \mathcal{U}_{i,t}^-)_{i \in N}.\end{aligned}$$

Then the separation problem for $P(\mathcal{U}_t)$ is solvable using $O(n^{t+2} + n^3)$ arithmetic operations, and therefore $P(\mathcal{U}_t)$ is solvable in polynomial time.

For this basis, the approximate cost-to-go of state (i, U) is

$$y_{i,U} = \lambda_{i,\varnothing} + \mu_{i,\varnothing} + \sum_{k \in U} \lambda_{i,k} + \sum_{k \in N \setminus (U \cup i)} \mu_{i,k} + \sum_{\substack{W \subseteq U\\ |W| \ge n-t}} \lambda_{i,W} + \sum_{\substack{W \subseteq N \setminus (U \cup i)\\ |W| \ge n-t}} \mu_{i,W},$$

For a city i, every state (i, U) is valued using $\lambda_{i,\varnothing} + \mu_{i,\varnothing}$ and $n-1$ variables of the form $\lambda_{i,k}$ and $\mu_{i,k}$. In addition, states with $|U| \ge n-t$ or $|U| \le t-1$ are valued more finely with the additional, higher-order variables. Before proving the result, we show that several of the variables defining the basis can be dropped without any loss.

Proposition 16. *If we reduce $\mathcal{U}_{i,t}^-$ to $\{U \subseteq N \setminus i : |U| \ge n-t\}$ the optimal value of $P(\mathcal{U}_t)$ remains unchanged. That is, variables $\mu_{i,U}$ for $|U| \le 1$ are redundant.*

Proof. Let $i \in N$, and let $\lambda_{i,U}, \mu_{i,U} \in \mathbb{R}$ for $|U| \le 1$. Our aim is to define new variables $\pi_{i,U}$, $|U| \le 1$ to replace the λ variables, and show that their approximate cost-to-go is equal to the one given by λ and μ together. Let $\pi_{i,\varnothing} = \lambda_{i,\varnothing} + \mu_{i,\varnothing} + \sum_{k \in N \setminus i} \mu_{i,k}$, and $\pi_{i,j} = \lambda_{i,j} - \mu_{i,j}$. Then

$$\begin{aligned}\pi_{i,\varnothing} + \sum_{k \in U} \pi_{i,k} &= \lambda_{i,\varnothing} + \mu_{i,\varnothing} + \sum_{k \in N \setminus i} \mu_{i,k} + \sum_{k \in U} (\lambda_{i,k} - \mu_{i,k})\\ &= \lambda_{i,\varnothing} + \mu_{i,\varnothing} + \sum_{k \in U} \lambda_{i,k} + \sum_{k \in N \setminus (U \cup i)} \mu_{i,k}.\end{aligned}$$

□

Proof of Theorem 15. After the variable replacement of Proposition 16, $P(\mathcal{U}_t)$ is

$$\max \; y_{0,N} \tag{8a}$$

$$\text{s.t. } y_{0,N} - \pi_{i,\varnothing} - \sum_{k \in N \setminus i} \pi_{i,k} - \sum_{\substack{U \subseteq N \setminus i \\ |U| \geq n-t}} \lambda_{i,U} \leq c_{0i}, \; \forall \, i \in N \tag{8b}$$

$$\begin{aligned} &\pi_{i,\varnothing} - \pi_{j,\varnothing} + \pi_{i,j} + \sum_{k \in U} (\pi_{i,k} - \pi_{j,k}) + \sum_{\substack{W \subseteq U \cup j \\ |W| \geq n-t}} \lambda_{i,W} - \sum_{\substack{W \subseteq U \\ |W| \geq n-t}} \lambda_{j,W} \leq c_{ij}, \\ &\qquad \forall \, i \in N, \; j \in N \setminus i, \; U \subseteq N \setminus \{i,j\}, \; |U| \geq n-t \end{aligned} \tag{8c}$$

$$\begin{aligned} &\pi_{i,\varnothing} - \pi_{j,\varnothing} + \pi_{i,j} + \sum_{k \in U} (\pi_{i,k} - \pi_{j,k}) + \lambda_{i,U \cup j} \leq c_{ij}, \\ &\qquad \forall \, i \in N, \; j \in N \setminus i, \; U \subseteq N \setminus \{i,j\}, \; |U| = n-t-1 \end{aligned} \tag{8d}$$

$$\begin{aligned} &\pi_{i,\varnothing} - \pi_{j,\varnothing} + \pi_{i,j} + \sum_{k \in U} (\pi_{i,k} - \pi_{j,k}) \leq c_{ij}, \\ &\qquad \forall \, i \in N, \; j \in N \setminus i, \; U \subseteq N \setminus \{i,j\}, \; t \leq |U| \leq n-t-2 \end{aligned} \tag{8e}$$

$$\begin{aligned} &\pi_{i,\varnothing} - \pi_{j,\varnothing} + \pi_{i,j} + \sum_{k \in U} (\pi_{i,k} - \pi_{j,k}) - \mu_{j,N \setminus (U \cup j)} \leq c_{ij}, \\ &\qquad \forall \, i \in N, \; j \in N \setminus i, \; U \subseteq N \setminus \{i,j\}, \; |U| = t-1 \end{aligned} \tag{8f}$$

$$\begin{aligned} &\pi_{i,\varnothing} - \pi_{j,\varnothing} + \pi_{i,j} + \sum_{k \in U} (\pi_{i,k} - \pi_{j,k}) + \sum_{\substack{W \subseteq N \setminus (U \cup \{i,j\}) \\ |W| \geq n-t}} \mu_{i,W} - \sum_{\substack{W \subseteq N \setminus (U \cup j) \\ |W| \geq n-t}} \mu_{j,W} \leq c_{ij}, \\ &\qquad \forall \, i \in N, \; j \in N \setminus i, \; U \subseteq N \setminus \{i,j\}, \; |U| \leq t-2 \end{aligned} \tag{8g}$$

$$\pi_{i,\varnothing} + \sum_{\substack{U \subseteq N \setminus i \\ |U| \geq n-t}} \mu_{i,U} \leq c_{i0}, \; \forall \, i \in N \tag{8h}$$

$$y_{0,N} \in \mathbb{R}; \quad \lambda, \mu \in \mathbb{R}^{n \times 2^{n-1}}. \tag{8i}$$

Fix $(\hat{\pi}, \hat{\lambda}, \hat{\mu})$; we consider the constraint classes in pairs.

(8b), (8h) There are $O(n)$ constraints, each with $O(n + n^{\max\{t-1,0\}})$ variables, requiring a total of $O(n^2 + n^t)$ operations.

(8c) If $t \leq 1$, this class is empty. Otherwise, there are $O(n^2)$ ordered city pairs (i, j), and for each $s = 2, \ldots, t$, there are $O(n^{s-2})$ subsets $U \subseteq N \setminus \{i, j\}$ with $|U| = n - s$, and hence $O(n^{s-2})$ constraints per ordered city pair. Each constraint has $O(n^{t-s+1})$ variables, and so the total number of operations is $O(n^{t+1})$. Constraints (8g) are handled analogously.

(8d) If $t = 0$, this class is empty. When $t \geq 1$, for each city pair there are $O(n^{t-1})$ subsets $U \subseteq N \setminus \{i, j\}$ with $|U| = n - t - 1$, and each constraint has $O(n)$ variables, yielding $O(n^t)$ operations per pair, and $O(n^{t+2})$ total operations. Constraints (8f) are analogous.

(8e) For any ordered pair (i, j), the most violated constraint in this class is given by the optimal solution of

$$\max_{\substack{U\subseteq N\setminus\{i,j\}\\ t\le|U|\le n-t-2}} \sum_{k\in U}(\hat{\pi}_{i,k}-\hat{\pi}_{j,k}).$$

To find the optimal set, it suffices to sort the values $\hat{\pi}_{i,k}-\hat{\pi}_{j,k}$ in non-increasing order, add the first t elements to U and thereafter add elements greedily until we either encounter the first non-positive value or add $n-t-2$ elements.

However, the optimization can actually be carried out in linear time because t is constant: For every $k\in N\setminus\{i,j\}$, add the element to U if $\hat{\pi}_{i,k}-\hat{\pi}_{j,k}>0$, but also keep a count of the number of positive elements and two arrays of size t, one with the t smallest positive elements added, the other with the t non-positive elements of smallest absolute value. After checking every element k, if we have added more than $n-t-2$, delete enough elements to reduce the set's cardinality to this number, deleting elements greedily based on the values stored in the first array; if we have added less than t, add enough elements from the second array in a similar fashion. Thus, the total number of operations is $O(n^3)$. □

The choice of basis in the approximation is unbalanced, since there are only $O(n)$ "lower order" multipliers (corresponding to empty or singleton sets) but $O(n^{\max\{t-1,0\}})$ "higher order" terms. The following negative complexity result shows that we cannot substantially increase the fineness of the approximation at the lower end without sacrificing efficiency.

Lemma 17. *Suppose for some $i\in N$ that either $\hat{\mathcal{U}}_i^+$ or $\hat{\mathcal{U}}_i^-$ contain $\Omega(n^2)$ sets of cardinality two. Then the separation problem for $P(\hat{\mathcal{U}})$ is NP-hard.*

Proof. For simplicity, assume $\hat{\mathcal{U}}_i^+=\hat{\mathcal{U}}_i^-=\{U\subseteq N\setminus i : |U|\le 2\},\ \forall\ i\in N$. After variable replacement, the constraint class (4c) is

$$\begin{aligned}&\pi_{i,\varnothing}-\pi_{j,\varnothing}+\pi_{i,j}+\sum_{k\in U}(\pi_{i,k}-\pi_{j,k}+\lambda_{i,\{j,k\}})-\sum_{k\in N\setminus(U\cup\{i,j\})}\mu_{j,\{i,k\}}\\ &\quad+\sum_{\substack{W\subseteq U\\|W|=2}}(\lambda_{i,W}-\lambda_{j,W})+\sum_{\substack{W\subseteq N\setminus(U\cup\{i,j\})\\|W|=2}}(\mu_{i,W}-\mu_{j,W})\le c_{ij},\ \forall\ i\in N,\ j\in N\setminus i,\ U\subseteq N\setminus\{i,j\}.\end{aligned}$$

Given $\hat{\pi},\hat{\lambda},\hat{\mu}$ and an ordered pair (i,j), the most violated constraint corresponds to the optimal solution of

$$\begin{aligned}\max_{U\subseteq N\setminus\{i,j\}}\Bigg\{&\sum_{k\in U}(\hat{\pi}_{i,k}-\hat{\pi}_{j,k}+\hat{\lambda}_{i,\{j,k\}})-\sum_{k\in N\setminus(U\cup\{i,j\})}\hat{\mu}_{j,\{i,k\}}\\ &+\sum_{\substack{W\subseteq U\\|W|=2}}(\hat{\lambda}_{i,W}-\hat{\lambda}_{j,W})+\sum_{\substack{W\subseteq N\setminus U\cup\{i,j\}\\|W|=2}}(\hat{\mu}_{i,W}-\hat{\mu}_{j,W})\Bigg\}.\end{aligned}$$

Even if $\hat{\mu}=0$ and $\hat{\lambda}_{i,W}-\hat{\lambda}_{j,W}\le 0$ for $W\subseteq N\setminus\{i,j\}$ with $|W|=2$, this is an arbitrary instance of the *quadratic cost partitioning* problem [31] on $N\setminus\{i,j\}$, itself a generalization of the max-cut problem. □

We next address our bound's relation to the Held-Karp bound.

Theorem 18. *The optimal value of $P(\mathcal{U}_0)$ is greater than or equal to the optimal value of the LP relaxation of* (1)*, the Held-Karp bound.*

Proof. Let $(\hat{\lambda}, \hat{\mu})$ be feasible for (7), the dual of the LP relaxation of (1). We may assume $\hat{\lambda}_\varnothing = \hat{\mu}_\varnothing = 0$, since these variables are made redundant by μ_N and λ_N respectively. Define $\hat{y}_{0,N} = \sum_{U\subseteq N}(\hat{\lambda}_U + \hat{\mu}_U)$ and

$$\hat{\pi}_{i,\varnothing} = \sum_{U\subseteq N\setminus i} \hat{\lambda}_{U\cup i}, \quad \hat{\pi}_{i,j} = \sum_{U\subseteq N\setminus\{i,j\}} \frac{\hat{\lambda}_{U\cup j} + \hat{\mu}_{U\cup j}}{|U|+1}.$$

Intuitively, for each city $i \in N$ this construction attempts to give each remaining city $j \in N \setminus i$ equal share in each of its moats. By definition, the two solutions' objectives are equal in their respective problems, so it remains to check the feasibility of $(\hat{y}_{0,N}, \hat{\pi})$ for $P(\mathcal{U}_0)$; since $t = 0$, only constraints (8b), (8e) and (8h) must be checked.

For (8b), let $i \in N$:

$$\hat{y}_{0,N} - \hat{\pi}_{i,\varnothing} - \sum_{k\in N\setminus i} \hat{\pi}_{i,k} = \sum_{U\subseteq N}(\hat{\lambda}_U + \hat{\mu}_U) - \sum_{U\subseteq N}\hat{\lambda}_U - \sum_{U\subseteq N\setminus i}\hat{\mu}_U = \sum_{U\subseteq N\setminus i}\hat{\mu}_{U\cup i} \le c_{0i}.$$

Similarly, for (8h) and $i \in N$, we have

$$\hat{\pi}_{i,\varnothing} = \sum_{U\subseteq N\setminus i} \hat{\lambda}_{U\cup i} \le c_{i0}.$$

Finally, for (8e), let $i \in N$, $j \in N \setminus i$ and $U \subseteq N \setminus \{i,j\}$. Then

$$\begin{aligned}
\hat{\pi}_{i,\varnothing} - \hat{\pi}_{j,\varnothing} + \hat{\pi}_{i,j} + \sum_{k\in U}(\hat{\pi}_{i,k} - \hat{\pi}_{j,k}) &= \sum_{W\subseteq N\setminus i}\hat{\lambda}_{W\cup i} - \sum_{W\subseteq N\setminus j}\hat{\lambda}_{W\cup j} + \sum_{W\subseteq N\setminus\{i,j\}}\frac{\hat{\lambda}_{W\cup j} + \hat{\mu}_{W\cup j}}{|W|+1} \\
&\quad + \sum_{k\in U}\left[\sum_{W\subseteq N\setminus\{i,k\}}\frac{\hat{\lambda}_{W\cup k} + \hat{\mu}_{W\cup k}}{|W|+1} - \sum_{W\subseteq N\setminus\{j,k\}}\frac{\hat{\lambda}_{W\cup k} + \hat{\mu}_{W\cup k}}{|W|+1}\right] \\
= \hat{\lambda}_i + \hat{\mu}_j + \sum_{\varnothing\neq W\subseteq N\setminus\{i,j\}} &\left[\frac{|W\setminus U|}{|W|+1}\hat{\lambda}_{W\cup i} - \frac{|U\cap W|}{|W|+1}\hat{\mu}_{W\cup i} - \frac{|W\setminus U|}{|W|+1}\hat{\lambda}_{W\cup j} + \frac{|U\cap W|+1}{|W|+1}\hat{\mu}_{W\cup j}\right] \\
&\le \sum_{W\subseteq N\setminus\{i,j\}}(\hat{\lambda}_{W\cup i} + \hat{\mu}_{W\cup j}) \le c_{ij},
\end{aligned}$$

where for the first inequality we use $\hat{\lambda}_U, \hat{\mu}_U \ge 0, \forall\ 2 \le |U| \le n-1$. □

Corollary 19. *Suppose $c \ge 0$ and it satisfies the triangle inequality: $c_{ij} \le c_{ik} + c_{kj}, \forall\ i,j,k \in N$. Let $y^*_{0,N}$ be the cost of an optimal tour, and let $y^0_{0,N}$ be the optimal value of $P(\mathcal{U}_0)$. If c is symmetric, then $y^*_{0,N} \le \frac{3}{2}y^0_{0,N}$. Otherwise, $y^*_{0,N} \le \min\{\log(n+1), 2 + \frac{8\log(n+1)}{\log\log(n+1)}\}y^0_{0,N}$.*

Proof. This follows directly from integrality gap results for the Held-Karp bound [9, 22, 38, 42, 43]. □

There is a well-known conjecture that the integrality gap of the Held-Karp bound in the symmetric case is actually $\frac{3}{4}$ [12, 23]. If this conjecture is proved, Theorem 18 would then imply the same bound guarantee for $P(\mathcal{U}_0)$.

Although we have found instances for which the optimal value of $P(\mathcal{U}_1)$ is greater than the Held-Karp bound, we have thus far been unable to find an instance for which $P(\mathcal{U}_0)$'s objective is greater. This motivates the following conjecture.

Conjecture 20. The optimal value of $P(\mathcal{U}_0)$ is equal to the Held-Karp bound.

Our inability may also stem from the difficulty we have encountered in solving $P(\mathcal{U}_0)$. We are so far able to solve only modestly sized instances, under 50 cities. For the instances we can solve, the computation times are an order of magnitude greater than the Held-Karp computation times on the same machine. There is also an additional inherent difficulty in solving $P(\mathcal{U}_0)$: Unlike the Held-Karp bound, our constraint generation operates in the dual space. Therefore, a relaxed dual solution does not provide a valid bound, so if we attempt to solve the problem in textbook fashion, only the optimal solution actually yields a valid bound. Our attempts to mitigate this difficulty have only met with partial success.

4 Conclusions

We have introduced a framework to generate lower bounds for the asymmetric TSP via an ALP approach to the DP formulation. This approach includes several existing bound techniques for the TSP, and also generates a new family of bounds by restricting an exact reformulation of the DP formulation. Our results motivate further research questions.

The first such question involves efficient computational methods to solve $P(\mathcal{U}_t)$. The inadequacy of a typical constraint generation algorithm for the problem suggests using an approach that maintains dual feasibility, and thus a valid bound, at every step. However, in this case one has to contend with the constraint class (8e) of exponential size. Another alternative could be to use constraint generation but somehow shift the relaxed dual solutions to enforce dual feasibility; this approach is viable only if the shift does not decrease the objective excessively.

A second important issue is the use of solutions to (4) as approximate cost-to-go functions for price-directed tours. In particular, there is potential to use lower bound guarantees like those given in Corollary 19 to derive worst-case performance ratios for the price-directed tours. This approach would reverse the usual direction of many TSP integrality gap proofs, which depend on an approximation algorithm guarantee [9, 22, 43].

Acknowledgements

The author would like to thank two anonymous referees for many helpful comments.

References

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

[2] D. Adelman. A Price-Directed Approach to Stochastic Inventory/Routing. *Operations Research*, 52:499–514, 2004.

[3] D. Adelman. Dynamic Bid Prices in Revenue Management. *Operations Research*, 55:647–661, 2007.

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

[5] D. Adelman and D. Klabjan. Computing Near-Optimal Policies in Generalized Joint Replenishment. *INFORMS Journal on Computing*, 24:148–164, 2011.

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

[7] D.L. Applegate, R.E. Bixby, V. Chvátal, and W.J. Cook. *The Traveling Salesman Problem: A Computational Study*. Princeton University Press, Princeton, New Jersey, 2006.

[8] S. Arora. Approximation Algorithms for Geometric TSP. In Gutin and Punnen [25], pages 207–222.

[9] A. Asadpour, M.X. Goemans, A. Madry, S. Oveis Gharan, and A. Saberi. An $O(\log n/\log\log n)$-approximation algorithm for the asymmetric traveling salesman problem. In *Proceedings of the 21st ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 379–389. SIAM, 2010.

[10] E. Balas and N. Simonetti. Linear Time Dynamic-Programming Algorithms for New Classes of Restricted TSPs: A Computational Study. *INFORMS Journal on Computing*, 13:56–75, 2001.

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

[12] R. Carr and S. Vempala. On the Held-Karp relaxation for the asymmetric and symmetric traveling salesman problems. *Mathematical Programming*, 100:569–587, 2004.

[13] N. Christofides, A. Mingozzi, and P. Toth. State space relaxation procedures for the computation of bounds to routing problems. *Networks*, 11:145–164, 1981.

[14] D. Cvetković, M. Cangalović, and V. Kovačević-Vujčić. Semidefinite Programming Methods for the Symmetric Traveling Salesman Problem. In G. Cornuéjols, R.E. Burkard, and G.J. Woeginger, editors, *Proceedings of the 7th International Integer Programming and Combinatorial Optimization (IPCO) Conference, Graz, Austria, June 9-11, 1999*, pages 126–136, 1999.

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

[16] E. de Klerk, D. Pasechnik, and R. Sotirov. On semidefinite programming relaxations of the traveling salesman problem. *SIAM Journal on Optimization*, 19:1559–1573, 2008.

[17] V.V. Desai, V.F. Farias, and C.C. Moallemi. The Smoothed Approximate Linear Program. To appear in *Operations Research*. Preprint available online at `arXiv:0908.0506v2`, 2009.

[18] Y. Dumas, J. Desrosiers, E. Gelinas, and M. Solomon. An Optimal Algorithm for the Traveling Salesman Problem with Time Windows. *Operations Research*, 43:367–371, 1995.

[19] J. Edmonds. Submodular functions, matroids and certain polyhedra. In R. Guy, H. Hanani, N. Sauer, and J. Schönheim, editors, *Combinatorial Structures and Their Applications*, pages 69–87. Gordon and Breach, New York, 1970.

[20] Ö. Ergun and J.B. Orlin. A dynamic programming methodology in very large scale neighborhood search applied to the traveling salesman problem. *Discrete Optimization*, 3:78–85, 2006.

[21] V.F. Farias and B. van Roy. An Approximate Dynamic Programming Approach to Network Revenue Management. Preprint available online at `http://web.mit.edu/~vivekf/www/mypapers2.html`, 2007.

[22] A.M. Frieze, G. Galbiati, and F. Maffioli. On the Worst-Case Performance of some Algorithms for the Asymmetric Traveling Salesman Problem. *Networks*, 12:23–39, 1982.

[23] M.X. Goemans. Worst-case Comparison of Valid Inequalities for the TSP. *Mathematical Programming*, 69:335–349, 1995.

[24] R.H. Gonzales. Solution to the traveling salesman problem by dynamic programming on the hypercube. Technical Report 18, Operations Research Center, Massachusetts Institute of Technology, 1962.

[25] G. Gutin and A.P. Punnen, editors. *The Traveling Salesman Problem and Its Variations*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.

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

[27] M. Held and R.M. Karp. The traveling-salesman problem and minimum spanning trees. *Operations Research*, 18:1138–1162, 1970.

[28] M. Held and R.M. Karp. The traveling-salesman problem and minimum spanning trees: Part II. *Mathematical Programming*, 1:6–25, 1971.

[29] D. Houck, J.C. Picard, M. Queyranne, and R.R. Vemuganti. The Traveling Salesman Problem and Shortest n-paths. Technical report, University of Maryland, 1977.

[30] M. Jünger and W.R. Pulleyblank. Geometric duality and combinatorial optimization. In S.D. Chatterji, B. Fuchssteiner, U. Kulisch, and R. Liedl, editors, *Jahrbuch Überblicke Mathematik*, pages 1–24. Vieweg, Braunschweig/Wiesbaden, 1993.

[31] H. Lee, G.L. Nemhauser, and Y. Wang. Maximizing a submodular function by integer programming: Polyhedral results for the quadratic case. *European Journal of Operational Research*, 94:154–166, 1996.

[32] C. Malandraki and R.B. Dial. A restricted dynamic programming heuristic algorithm for the time dependent traveling salesman problem. *European Journal of Operational Research*, 30:45–55, 1996.

[33] A. Mingozzi, L. Bianco, and S. Ricciardelli. Dynamic programming strategies for the traveling salesman problem with time window and precedence constraints. *Operations Research*, 45:365–377, 1997.

[34] S. Nadarajah, F. Margot, and N. Secomandi. 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, 2011.

[35] H.N. Psaraftis. A Dynamic Programming Solution to the Single Vehicle Many-to-Many Immediate Request Dial-a-Ride Problem. *Transportation Science*, 14:130–154, 1980.

[36] A. Schrijver. The traveling salesman problem. In *Combinatorial Optimization: Polyhedra and Efficiency*, volume B, chapter 58, pages 981–1004. Springer, Berlin, 2003.

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

[38] D.B. Shmoys and D.P. Williamson. Analyzing the Held-Karp TSP Bound: A Monotonicity Property with Application. *Information Processing Letters*, 35:281–285, 1990.

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

[40] M.A. Trick and S.E. Zin. Spline Approximations to Value Functions: A Linear Programming Approach. *Macroeconomic Dynamics*, 1:255–277, 1997.

[41] M. van Vyve and L.A. Wolsey. Approximate extended formulations. *Mathematical Programming*, 105:501–522, 2006.

[42] D.P. Williamson. Analysis of the Held-Karp Heuristic for the Traveling Salesman Problem. Master's thesis, Massachusetts Institute of Technology, 1989.

[43] L.A. Wolsey. Heuristic Analysis, Linear Programming and Branch and Bound. *Mathematical Programming Study*, 13:121–134, 1980.