Solving a Random Asymmetric TSP Exactly in Quasi-Polynomial Time w.h.p.

Tolson Bell* and Alan M. Frieze†

Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh, PA, 15213, USA

June 29, 2024

Abstract

Let the costs $C(i,j)$ for an instance of the Asymmetric Traveling Salesperson Problem (ATSP) be independent copies of an absolutely continuous random variable C that (i) satisfies $\mathbb{P}(C \leq x) = x + O(x^2)$ as $x \to 0$ and (ii) has an exponential tail. We describe an algorithm that solves ATSP exactly in time $e^{\log^{2+o(1)} n}$, w.h.p.

*thbell@cmu.edu. Research supported in part by NSF Graduate Research Fellowship grant DGE 2140739.

†frieze@cmu.edu. Research supported in part by NSF grant DMS1952285

1 Introduction

Given an $n \times n$ matrix $(C(i,j))_{i,j\in[n]}$, the Asymmetric Traveling Salesperson Problem (ATSP) asks for the cyclic permutation π on n elements that minimizes $\sum_{i=1}^n C(i,\pi(i))$. We consider the case where the costs $C(i,j)$ are drawn iid from a distribution C that satisfies

(i) $\mathbb{P}(C \le x) = x + O(x^2)$ as $x \to 0$ and (ii) $\mathbb{P}(C \ge x) \le \alpha e^{-\beta x}$ for constants $\alpha, \beta > 0$.

For instance, C could be uniform in $[0,1]$ or exponential with mean 1. The main result of this paper is as follows:

Theorem 1 *Let the costs for ATSP, $(C(i,j))$, be independent copies of C. There is an algorithm that solves ATSP exactly in $e^{\log^{2+o(1)} n}$ time, with high probability over the choices of random costs.*

1.1 Background

Given $(C(i,j))_{i,j\in[n]}$, we can define two discrete optimization problems. Let S_n denote the set of permutations of $[n] = \{1,2,\ldots,n\}$. Let $T_n \subseteq S_n$ denote the set of *cyclic* permutations i.e. those permutations whose cycle structure consists of a single cycle. The *Assignment Problem* (AP) is the problem of minimising $C(\pi) = \sum_{i=1}^n C(i,\pi(i))$ over all permutations $\pi \in S_n$. We let $Z_{\text{AP}} = Z_{\text{AP}}^{(C)}$ denote the optimal cost for AP. The *Asymmetric Traveling-Salesperson Problem* (ATSP) is the problem of minimising $C(\pi) = \sum_{i=1}^n C(i,\pi(i))$ over all permutations $\pi \in T_n$. We let $Z_{\text{ATSP}} = Z_{\text{ATSP}}^{(C)}$ denote the optimal cost for ATSP.

Another view of the assignment problem is that it is the problem of finding a minimum cost perfect matching in the complete bipartite graph $K_{A,B}$ where $A = \{a_1, a_2, \ldots, a_n\}$ and $B = \{b_1, b_2, \ldots, b_n\}$ and the cost of edge (a_i, b_j) is $C(i,j)$.

It is evident that $Z_{\text{AP}}^{(C)} \le Z_{\text{ATSP}}^{(C)}$. The ATSP is NP-hard, whereas the AP is solvable in time $O(n^3)$ [Tom71, EK72]. In 1971, Bellmore and Malone [BM71] conjectured that using the AP in a branch and bound algorithm would give a polynomial expected time algorithm for the ATSP. Lenstra and Rinnooy Kan [LR79] and Zhang [Zha97] found errors in the argument of [BM71].

Several authors, e.g. Balas and Toth [BT86], Kalczynski [Kal05], Miller and Pekny [MP91], Zhang [Zha04] have investigated using the AP in a branch-and-bound algorithm to solve the ATSP and have observed that the AP gives extremely good bounds on random instances. Experiments suggest that if the costs $C(i,j)$ are independently and uniformly generated as integers in the range $[0,L]$, then as L gets larger, the problem gets harder to solve. Rigorous analysis supporting this thesis was given by Frieze, Karp and Reed [FKR95]. They showed that if $L(n) = o(n)$ then $Z_{\text{ATSP}} = Z_{\text{AP}}$ w.h.p. and that w.h.p. $Z_{\text{ATSP}} > Z_{\text{AP}}$ if $L(n)/n \to \infty$.

We implicitly study a case where $L(n)/n \to \infty$. Historically, researchers have considered the case where the costs $C(i,j)$ are independent copies of the uniform $[0,1]$ random variable

$U[0,1]$. This model was first considered by Karp [Kar79]. He proved the surprising result that

$$Z_{\text{ATSP}} - Z_{\text{AP}} = o(1) \text{ w.h.p.} \tag{1}$$

Since w.h.p. $Z_{\text{AP}} > 1$ we see that this rigorously explained the observed quality of the assignment bound. Karp [Kar79] proved (1) constructively, analysing an $O(n^3)$ *patching* heuristic that transformed an optimal AP solution into a good ATSP solution. Karp and Steele [KS85] simplified and sharpened this analysis, and Dyer and Frieze [DF90] improved the error bound of through the analysis of a related more elaborate algorithm to $O\left(\frac{\log^4 n}{n \log\log n}\right)$. Frieze and Sorkin [FS07] reduced the error bound to

$$Z_{\text{ATSP}} - Z_{\text{AP}} \leq \frac{\zeta \log^2 n}{n} \text{ w.h.p.} \tag{2}$$

and gave an $e^{n^{.5+o(1)}}$ time algorithm to solve the ATSP w.h.p. Our Theorem 1 improves the exponent from $n^{.5+o(1)}$ to $\log^{2+o(1)} n$ and allows a larger class of random distributions for C.

One might think that with such a small gap between Z_{AP} and Z_{ATSP}, that branch and bound might run in polynomial time w.h.p. Indeed one is encouraged by the recent results of Dey, Dubey and Molinaro [DDM21] and Borst, Dadush, Huiberts and Tiwari [BDHT23] that with a similar integrality gap, branch and bound with LP based bounds solves random multi-dimensional knapsack problems in polynomial time w.h.p. Given Theorem 1, one is tempted to side with [BM71] and conjecture that branch and bound can be made to run in polynomial time w.h.p.

2 Proof Outline of Theorem 1

Let $M^* = \{(a_i, b_{\phi(i)}), i \in [n]\}$ denote the optimum matching that solves AP in terms of the permutaton ϕ on $[n]$. Any other perfect matching of $K_{A,B}$ can be obtained from M^* by choosing a set of vertex disjoint *alternating cycles* $C_1, C_2, \ldots, C_m$ in $K_{A,B}$ and replacing M^* by $M^* \oplus C_1 \cdots \oplus C_m$. Here an alternating cycle is one whose edges alternate between being in M^* and not in M^*. We use the notation $S \oplus T = (S \setminus T) \cup (T \setminus S)$.

For a matching M we let $C(M) = \sum_{e \in M} C(e)$. The basic idea of the proof is to show that if a matching M is "too far" from M^*, then w.h.p. $C(M) - C(M^*) > \frac{\zeta \log^2 n}{n}$ (where ζ is from (2)), and thus M cannot be the optimal ATSP solution. Given this, it does not take too long to check all possible M that are "close to" M^*, to see if M defines a tour and then determine its total cost.

More specifically, we find in polynomial time a particular $T^* \supseteq M^*$ such that T^* is a spanning tree of $K_{A,B}$ (and thus $|T^*| = 2|M^*| - 1 = 2n - 1$). Our definition of "too far" is then that M contains more than $\log^{2+o(1)} n$ edges outside T^*. T^* arises as the spanning tree representing the optimal basis in the LP formulation of AP.

In Sections 3 and 4, we will assume that the distribution of the costs $C(i,j)$ is exponential mean one, $EXP(1)$, i.e. $\mathbb{P}(C \geq x) = e^{-x}$. We need to make this assumption for the proof of

Lemma 6, which itself is needed for the proof of Lemma 11. We will subsequently generalise this to our more general class of distribution in Section 6.

3 The Assignment Problem and Nearby Permutations

Frieze and Sorkin [FS07] proved that the following two lemmas hold w.h.p.:

Lemma 2 $\max_{e\in M^*} C(e) \leq \frac{\gamma \log n}{n}$ *for some absolute constant* $\gamma > 0$.

Lemma 3 $Z_{\mathrm{ATSP}} - Z_{\mathrm{AP}} \leq \frac{\zeta \log^2 n}{n}$ *for some absolute constant* $\zeta > 0$.

We will let γ^* be the upper bound on costs given in the first lemma and ζ^* be the upper bound given in the second. Frieze and Sorkin's results were only proved for the random distribution being $U[0,1]$. In Section 5, we will prove that these lemmas still hold for our more general class of distributions. For Lemma 3, this comes via a short coupling argument, but the analysis to generalize Lemma 2 is more involved.

3.1 AP as a linear program

The assignment problem AP has a linear programming formulation $\mathcal{LP}$. In the following $z_{i,j}$ indicates whether or not (a_i, b_j) is an edge of the optimal solution.

$$\begin{aligned} \mathcal{LP} \qquad & \text{Minimise} \sum_{(i,j)\in[n]^2} C(i,j) z_{i,j} \\ & \text{subject to} \sum_{j=1}^{n} z_{i,j} = 1, \text{ for } i = 1,2,\ldots,n. \\ & \sum_{i=1}^{n} z_{i,j} = 1, \text{ for } j = 1,2,\ldots,n. \\ & 0 \leq z_{i,j} \leq 1, \text{ for } (i,j)\in[n]^2. \end{aligned} \tag{3}$$

This has the dual linear program:

$$\begin{aligned} \mathcal{DLP} \qquad & \text{Maximise} \sum_{i=1}^{n} u_i + \sum_{j=1}^{n} v_j \\ & \text{subject to } u_i + v_j \leq C(i,j), \text{ for } (i,j)\in[n]^2. \end{aligned} \tag{4}$$

Proposition 4 *Condition on an optimal basis for* (3). *We may w.l.o.g. take* $u_1 = 0$ *in* (4), *whereupon with probability 1 the other dual variables are uniquely determined. Furthermore, the reduced costs of the* non-basic *variables* $\bar{C}(i,j) = C(i,j) - u_i - v_j$ *are independently distributed as either (i)* $C - u_i - v_j$ *if* $u_i + v_j < 0$ *or (ii)* $C - u_i - v_j$ *conditional on* $C \geq u_i + v_j$, *if* $u_i + v_j \geq 0$.

Proof. The $2n-1$ dual variables are unique with probability 1 because they satisfy $2n-1$ full rank linear equations. The only conditions on the non-basic edge costs are that $C(i,j) \geq (u_i + v_j)^+$, where $x^+ = \max\{x, 0\}$. □

3.2 Trees and bases

An optimal basis of $\mathcal{LP}$ can be represented by a spanning tree T^* of $K_{A,B}$ that contains the perfect matching M^*, see for example Ahuja, Magnanti and Orlin [AMO93], Chapter 11. The edges of such a tree are referred to as *basic* edges, when the tree in question is T^*. We have that for every optimal basis T^*,

$$C(i,j) = u_i + v_j \text{ for } (a_i, b_j) \in E(T^*) \tag{5}$$

and

$$C(i,j) \geq u_i + v_j \text{ for } (a_i, b_j) \notin E(T^*). \tag{6}$$

Lemma 2 implies that we can ignore any edge (a_i, b_j) for which $C(i,j) > \gamma^*$ in our search for an optimal matching. Note also that the subgraph induced by the edges (a_i, b_j) with $C(i,j) \leq \gamma^*$ is connected w.h.p. So, we can assume w.h.p. that all the edges of T^* have cost at most γ^*.

For any tree T, let $\mathcal{B} = \mathcal{B}(T)$ be the event that $C(i,j) \geq u_i + v_j$ for all $(a_i, b_j) \notin E(T)$. Note that if $\mathcal{B}$ occurs, then $T^* = T$.

Lemma 5

$$|u_i|, |v_i| \leq 2\gamma^* \textit{ for } i \in [n], \textit{ w.h.p.} \tag{7}$$

Proof. For each $i \in [n]$ there is some $j \in [n]$ such that $u_i + v_j = C(i,j) \leq \gamma^*$. This is because of the fact that a_i meets at least one edge of T and we assume that (5) holds. We also know that if $\mathcal{B}$ occurs then $u_{i'} + v_j \leq C(i',j)$ for all $i' \neq i$. It follows that $u_i - u_{i'} \geq C(i,j) - C(i',j) \geq -\gamma^*$ for all $i' \neq i$. Since i is arbitrary, we deduce that $|u_i - u_{i'}| \leq \gamma^*$ for all $i, i' \in [n]$. This implies that $|u_i| \leq \gamma^*$ for $i \in [n]$. We deduce by a similar argument that $|v_j - v_{j'}| \leq \gamma^*$ for all $j, j' \in [n]$. Now because for the optimal matching edges $(i, \phi(i)), i \in [n]$ we have $u_i + v_{\phi(i)} = C(i, \phi(i))$, we see that $|v_j| \leq 2\gamma^*$ for $j \in [n]$. □

Condition on M^* and let G_+ denote the subgraph of $K_{A,B}$ induced by the edges (a_i, b_j) for which $u_i + v_j \geq 0$, where $\mathbf{u}, \mathbf{v}$ are optimal dual variables. Let $\mathcal{T}_+$ denote the set of spanning trees of G_+ that contain the edges of M^*.

Lemma 6 *If $T \in \mathcal{T}_+$ and (7) holds then*

$$\mathbb{P}(T^* = T \mid \mathbf{u}, \mathbf{v}) = \prod_{\substack{(a_i,b_j) \in G_+ \\ C(i,j) \leq 4\gamma^*}} e^{-u_i - v_j}, \tag{8}$$

which is independent of T.

Proof. Fixing $\mathbf{u}$,$\mathbf{v}$ and T fixes the lengths of the edges in T. If $(a_i, b_j) \notin E(T)$ then $\mathbb{P}(C(i,j) \geq u_i + v_j) = 1$ if $u_i + v_j < 0$ and $e^{-(u_i+v_j)}$ otherwise. Remember that we are assuming that C is exponential mean one at the moment. Thus,

$$\begin{aligned} \mathbb{P}(T^* = T \mid \mathbf{u}, \mathbf{v}) &= \prod_{(a_i,b_j)\notin E(T)} e^{-(u_i+v_j)^+} \prod_{(a_i,b_j)\in E(T)} e^{-(u_i+v_j)} \\ &= \prod_{(a_i,b_j)\in G_+} e^{-(u_i+v_j)}. \end{aligned} \tag{9}$$

In the first product we use (6), and the second product comes from (5) and from the density function of the costs. □

Thus

$$T^* \text{ is a uniform random member of } \mathcal{T}_+. \tag{10}$$

Now let Γ_+ be the multi-graph obtained from G_+ by contracting the edges of M^* and let $\widehat{T}^*$ be the corresponding contraction of T^*.

Lemma 7 *The distribution of the tree $\widehat{T}^*$ is asymptotically equal to that of a random spanning tree of $K_n + \widehat{M}$ where $\widehat{M}$ is a matching of size at most $\lambda^* = \lambda \log^4 n$ for some constant $\lambda > 0$ and K_n is the complete graph on vertex set $[n]$. ($\widehat{M}$ yields double edges, other edges occur once.)*

Proof. We have that for all $i, j \in [n]$,

$$(u_i + v_{\phi(j)}) + (u_j + v_{\phi(i)}) = (u_i + v_{\phi(i)}) + (u_j + v_{\phi(j)}) = C(i, \phi(i)) + C(j, \phi(j)) > 0.$$

So, either $u_i + v_{\phi(j)} > 0$ or $u_j + v_{\phi(i)} > 0$ which implies that Γ_+ contains the edge $\{a_i, a_j\}$. So, Γ_+ contains K_A as a subgraph.

We know from (5) and Lemma 7 that $\widehat{T}^*$ only contains edges of cost at most $2\gamma^*$. So from (10), $\widehat{T}^*$ is a random spanning tree of a graph distributed as G_{n,γ^*} plus a set of edges $\widehat{M}$. The edges $\widehat{M}$ arise from 4-cycles (C_4) where each edge has cost at most γ^*. The expected number of such cycles is $O((n\gamma^*)^4)$ and so by standard results on the number of copies of balanced graphs, we see that $|\widehat{M}| = O(\log^4 n)$ w.h.p. At this density, any copies of C_4 will be vertex disjoint w.h.p., as can easily be verified by a first moment calculation.

A random spanning tree of $G_{n,p} + \widehat{M}$, where $\widehat{M}$ is a random matching, is by symmetry, a random spanning tree of $K_n + \widehat{M}$. □

We need to know that w.h.p., for each a_i, there are many b_j for which $u_i + v_j \geq 0$. We fix a tree T and condition on $T^* = T$. For $i = 1, 2, \ldots, r$ let $L_{i,+} = \{j : u_i + v_j \geq 0\}$ and let $L_{j,-} = \{i : u_i + v_j \geq 0\}$. Then let $\mathcal{A}_{i,+}$ be the event that $|L_{i,+}| \leq \eta n$ and let $\mathcal{A}_{j,-}$ be the event that $|L_{j,-}| \leq \eta n$ where η will be some small positive constant.

Lemma 8 *Fix a spanning tree T of G_+.*

$$\mathbb{P}\left(\bigcup_{i=1}^{n} (\mathcal{A}_{i,+} \cup \mathcal{A}_{j,-}) \mid T^* = T\right) = o(1). \tag{11}$$

Proof. We assume that $C(i,j) \leq \gamma^*$ for $(a_i, b_j) \in T$. The justification for this is Lemma 2, in that we can solve the assignment problem only using edges of cost at most γ^*.

The number of edges in G_+ of cost at most γ^* incident with a fixed vertex is dominated by $Bin(n, \gamma^*)$ and so w.h.p. the maximum degree of the trees we consider can be bounded by $2\gamma \log n$.

Let $\mathcal{E} \supseteq \mathcal{B}$ be the event that $|u_i|, |v_j| \leq 2\gamma^*$ for all i, j. (See the proof of Lemma 5 for the explanation of the set inclusion.)

We now condition on the set E_T of edges (and the associated costs) of $\{(a_i, b_j) \notin E(T)\}$ such that $C(i,j) \geq 2\gamma^*$. Let $F_T = \{(a_i, b_j) \notin E(T)\} \setminus E_T$. Note that $|F_T|$ is dominated by $Bin(n^2, 1 - e^{-2\gamma^*})$ and so $|F_T| \leq 3n^2\gamma^*$ with probability $1 - o(n^{-2})$.

Let $Y = \{C(i,j) : (a_i, b_j) \in E(T)\}$ and let $\delta_1(Y)$ be the indicator for $\mathcal{A}_{s,+} \wedge \mathcal{E}$. We write,

$$\mathbb{P}(\mathcal{A}_{s,+} \mid \mathcal{B}) = \mathbb{P}(\mathcal{A}_{s,+} \wedge \mathcal{E} \mid \mathcal{B}) = \frac{\int \delta_1(Y)\mathbb{P}(\mathcal{B} \mid Y)dC}{\int \mathbb{P}(\mathcal{B} \mid Y)dC} \tag{12}$$

Then we note that since $(a_i, b_j) \notin F_T \cup E(T)$ satisfies the condition (6),

$$\mathbb{P}(\mathcal{B} \mid Y) = \prod_{(a_i,b_j)\in F_T} e^{-(u_i(Y)+v_j(Y))^+} = e^{-W},$$

where $W = W(Y) = \sum_{(a_i,b_j)\in F_T}(u_i(Y) + v_j(Y))^+ \leq 12n^2(\gamma^*)^2 = 12\gamma^2 \log^2 n$. Then we have

$$\begin{aligned}
\int_Y \delta_1(Y)\mathbb{P}(\mathcal{B} \mid Y)\, dC &= \int_Y e^{-W}\delta_1(Y)\, dC \\
&\leq \left(\int_Y e^{-2W}\, dC\right)^{1/2} \times \left(\int_Y \delta_1(Y)^2\, dC\right)^{1/2} \\
&= e^{-\mathbf{E}(W)}\left(\int_Y e^{-2(W-\mathbf{E}(W))} dC\right)^{1/2} \times \mathbb{P}(\mathcal{A}_{s,+} \mid \mathcal{E})^{1/2} \\
&\leq e^{-\mathbf{E}(W)} e^{12\gamma^2\log^2 n}\mathbb{P}(\mathcal{A}_{s,+} \mid \mathcal{E})^{1/2}.
\end{aligned} \tag{13}$$

We also have

$$\int_Y \mathbb{P}(\mathcal{B} \mid Y)dC = \mathbf{E}(e^{-W}) \geq e^{-\mathbf{E}(W)}. \tag{14}$$

It then follows from (12),(13) and (14) that

$$\mathbb{P}(\mathcal{A}_{s,+} \mid \mathcal{B}) \leq e^{12\gamma^2\log^2 n}\mathbb{P}(\mathcal{A}_{s,+} \mid \mathcal{E})^{1/2}. \tag{15}$$

Equation (15) removes the conditioning on $\mathcal{B}$ at the expense of inflating our probability estimate by $e^{12\gamma^2\log^2 n}$. Note that if $\mathcal{B}$ occurs and (5) holds then $T^* = T$. For the remainder of the lemma we assume that the $C(i,j)$ for $(a_i, b_j) \in T$ satisfy $C \leq \gamma^*$ and that $\mathcal{E}$ holds. Denote this conditioning by $\mathcal{F}$. Let b_j be a neighbor of a_s in G_+ and let $P_j = (i_1 = s, j_1, i_2, j_2, \ldots, i_k, j_k = j)$ define the path from a_s to b_j in T. Then it follows from (5) that $v_{j_l} = v_{j_{l-1}} - C(i_l, j_{l-1}) + C(i_l, j_l))$. Thus v_j is the final value S_k of a random walk $S_t = X_0 + X_1 + \cdots + X_t, t = 0, 1, \ldots, k$, where $X_0 \geq 0$ and each $X_t, t \geq 1$ is the difference between two copies of C subject only to $\mathcal{F}$. Given $\mathcal{E}$ we can assume that the partial sums S_i satisfy $|S_i| \leq 2\gamma^*$ for $i = 1, 2, \ldots, k-1$. Assume for the moment that $k \geq 4$ and let $x = u_{i_{k-3}} \in [-2\gamma^*, 2\gamma^*]$. Given x we see that there is some positive probability $p_0 = p_0(x)$ that $S_k > 0$. Indeed,

$$p_0 = \mathbb{P}(S_k > 0 \mid \mathcal{E}) \geq \mathbb{P}(S_k > 0) - \mathbb{P}(\bar{\mathcal{E}}) = \mathbb{P}(x + Z_1 - Z_2 > 0) - \mathbb{P}(\bar{\mathcal{E}}), \tag{16}$$

where $Z_1 = Z_{1,1} + Z_{1,2} + Z_{1,3}$ and $Z_2 = Z_{2,1} + Z_{2,2}$ are the sums of independent copies of C, each conditioned on being bounded above by γ^* and such that $|x + \sum_{j=1}^{t}(Z_{1,j} - Z_{2,j})| \leq 2\gamma^*$ for $t = 1, 2$ and that $|x + Z_1 - Z_2| \leq 2\gamma^*$. The absolute constant $\eta_1 = p_0(-2\gamma^*) > 0$ is such that $\min\{x \geq -2\gamma^* : p_0(x)\} \geq \eta_0$.

We now partition (most of) the neighbors of a_s into N_0, N_1, N_2 where
$N_t = \{b_j : k \geq 3, k \mod 3 = t\}$, k being the number of edges in the path P_j from a_s to b_j. Now because T has maximum degree $2\gamma \log n$, as observed at the beginning of the proof of this lemma, we know that there exists t such that $|N_t| \geq (n - (2\gamma \log n)^3)/3 \geq n/4$. It then follows from (16) that $|L_{s,+}|$ dominates $Bin(n, \eta_0)$ and then $\mathbb{P}(|L_{s,+}| \leq \eta_0 n/10) = O(e^{-\Omega(n)})$ follows from the Chernoff bounds. Similarly for $L_{1,-}$. Applying the union bound over n choices for s and applying (15) gives the lemma with $\eta = \eta_0/10$. □

3.3 Alternating paths

We now consider the the number of edges in alternating paths that consist only of basic edges. We call these *basic alternating paths.*

Lemma 9 *The expected number of basic alternating paths with k edges is at most $n^2\left(1 - \frac{\eta}{1+\eta}\right)^k$, where η is as in Lemma 8.*

Proof. Let $P = (b_{\phi(i_1)}, a_{i_1}, b_{\phi(i_2)}, a_{i_2}, \ldots, b_{\phi(i_k)}, a_{i_k})$ be a prospective basic alternating path. Then $Q = (a_{i_1}, a_{i_2}, \ldots, a_{i_k})$ must be a path in $\widehat{T}^*$. When we uncontract M^*, the edge $\{a_{i_t}, a_{i_{t+1}}\}$ arises either (i) from $(b_{\phi(i_t)}, a_{i_{t+1}}, b_{\phi(i_{t+1})})$ or (ii) from $(a_{i_t}, b_{\phi(i_{t+1})}, a_{i_{t+1}})$ and we get an alternating path only if we have the former case for $t = 1, 2, \ldots, k$.

Consider the random walk struction of a spanning tree as described in Aldous [Ald90] and Broder [Bro89]. We have to modify the walk so that the tree contains M^*. We do this by giving the edges of M^* a large weight $W \gg n$. This will mean that when the walk arrives at some a_i it is very likely to move to $b_{\phi(i)}$ and then back to a_i and so on. It will however eventually leave the edge $(a_i, b_{\phi(i)})$ and either leave from a_i or from $b_{\phi(i)}$. We can model this via a sequence of independent experiments where the probability of success is at most $n/(W+n)$ at odd steps at least $\eta n/(W + \eta n)$ at even steps. Here odd steps correspond to being at a_i and being in case (i) and even steps correspond to being at $b_{\phi(i)}$ and being in case (ii). Lemma 8 implies that when the walk adds an edge to the tree there is a probability of at least η that the edge arises from case (ii) above. The probability of an even success is therefore at least

$$\sum_{k\geq 1}\left(1-\frac{n}{W+n}\right)^k\left(1-\frac{\eta n}{W+n}\right)^{k-1}\cdot\frac{\eta n}{W+n}=$$
$$\left(1-\frac{n}{W+n}\right)\cdot\frac{\eta n}{W+n}\cdot\frac{1}{1-\left(1-\frac{n}{W+n}\right)\left(1-\frac{\eta n}{W+n}\right)}\sim\frac{\eta}{1+\eta}. \quad (17)$$

This will be independent of the addition of previous edges and so the expected number of basic alternating path with k edges can be bounded by $n^2\left(1-\frac{\eta}{1+\eta}\right)^k$ and the lemma follows. □

Corollary 10 *W.h.p. the maximum length of a basic alternating path is at most* $3\eta^{-1}\log n$.

Fix $\omega=\omega(n)$ to be an arbitrary function such that $\omega\to\infty,\omega=\log^{o(1)} n$.

Lemma 11 *W.h.p there are at most $m=\omega n$ basic alternating paths, each using $O(\log n)$ edges.*

Proof. Let Z_1 denote the number of basic alternating paths. We would like to use the following result of Meir and Moon [MM70]: if T is a uniform random spanning tree of the complete graph K_n and $d_T(i,j)$ is the distance between $i\neq j\in[n]$ in T, then

$$\mathbb{P}(d_T(i,j)=k)=\frac{k}{n-1}\cdot\frac{n(n-1)\cdots(n-k+1)}{n^k},\qquad \text{for } 1\leq k\leq n-1.$$

The problem is that if a tree of K_A contains ℓ edges of $\widehat{M}$ (see Lemma 7) then its probability of occurring in Γ_+ is inflated by 2^ℓ. On the other hand, the probability that a random tree in K_A contains ℓ given edges is at most $(2/n)^\ell$. ($2/n$ for $\ell=1$ and at most $(2/n)^\ell$ in general using negative correlation, see [LP17].) So, assuming $|\widehat{M}|\leq\lambda\log^4 n$ (see Lemma 7), the expected number of basic alternating paths can be bounded by

$$n^2\sum_{k=1}^{n}\frac{k}{n-1}\cdot\left(1-\frac{\eta}{1+\eta}\right)^k\cdot\sum_{\ell\geq 0}\binom{\lambda\log^4 n}{\ell}\left(\frac{4}{n}\right)^\ell\leq\frac{2n}{\eta}.$$

This finishes the proof when combined with the Markov inequality and Corollary 10. □

So, w.h.p. the matching M corresponding to the ATSP solution is derived from a collection of short basic alternating paths $P_1,P_2,\ldots,P_m$ joined by non-basic edges to create alternating cycles $Q_1,Q_2,\ldots,Q_\ell$. Now consider an alternating cycle $Q=(a_{i_1},b_{j_1},\ldots,b_{j_t},a_{i_1})$ made up from such paths by adding non-basic edges joining up the endpoints. Putting $\tilde{C}(i,j)=$

$C(i,j) - u_i - v_j$, we have that where $j_t = \phi(i_t)$ for $t = 1, 2, \ldots, k$,

$$
\begin{aligned}
C(Q \oplus M^*) - C(M^*) &= \sum_{k=1}^{t} (C(i_{k+1}, j_k) - C(i_k, j_k)) \\
&= \sum_{k=1}^{t} ((\tilde{C}(i_{k+1}, j_k)) + u_{i_{k+1}} + v_{j_k}) - (\tilde{C}(i_k, j_k) + u_{i_k} + v_{j_k}) \\
&= \sum_{k=1}^{t} \tilde{C}(i_{k+1}, j_k) \\
&= \sum_{\substack{k=1 \\ (i_{k+1}, j_k) \text{ non-basic}}}^{t} \tilde{C}(i_{k+1}, j_k).
\end{aligned}
$$

Lemma 12 *The optimal solution to the ATSP uses at most* $\log^{2+o(1)} n$ *non-basic edges w.h.p.*

Proof. For any $k \in \mathbb{N}$, let Z_k be the number of perfect matchings M in $K_{A,B}$ with $|M \setminus T^*| = k$ and $C(M) - C(M^*) \leq \zeta^*$ (see Lemma 3).

For a perfect matching M with $|M \setminus T^*| = k$, we have that $M \oplus M^*$ consists of ℓ cycles $Q_1, Q_2, \ldots, Q_\ell$ where each Q_i has k_i edges outside T^* and $k_1 + \ldots + k_\ell = k$. For each $i \in [\ell]$, we specify Q_i by choosing k_i alternating paths in T^*, then ordering and orienting them in at most $k_i! 2^{k_i}$ ways. M then satisfies the criteria only if the $\tilde{C}$ of the k corresponding non-basic edges sum to at most ζ^*. Thus, if $\tilde{C}_1, \tilde{C}_2, \ldots, \tilde{C}_k$ are independent random variables distributed as (i), (ii) of Proposition 4, we have

$$
\begin{aligned}
\mathbf{E}(Z_k) &\leq \sum_{\ell=1}^{k} \sum_{k_1+\cdots+k_\ell=k} \prod_{i=1}^{\ell} \binom{m}{k_i} k_i! 2^{k_i} \mathbb{P}\left(\tilde{C}_1 + \tilde{C}_2 + \cdots + \tilde{C}_k \leq \zeta^*\right) && (18) \\
&\lesssim (2m)^k \frac{(2\zeta^*)^k}{k!} \sum_{\ell=1}^{k} \binom{k}{\ell} && (19) \\
&\leq \left(\frac{8\zeta e \omega \log^2 n}{k}\right)^k. && (20)
\end{aligned}
$$

Where, to go from (18) to (19), we used

$$
\begin{aligned}
\mathbb{P}\left(\tilde{C}_1 + \cdots + \tilde{C}_k \leq \zeta^*\right) &\leq \int_{z_1+\cdots+z_k \leq \zeta^*} \prod_{i=1}^{k} e^{-z_i + 4\gamma^*} d\mathbf{z} \\
&\lesssim 2^k \int_{z_1+\cdots+z_k \leq \zeta^*} 1 d\mathbf{z} = \frac{(2\zeta^*)^k}{k!}.
\end{aligned}
$$

The proof is completed by noting that $\mathbf{E}(\sum_{\log^{2+o(1)} n}^{\infty} Z_k) = o(1)$. (The $4\gamma^*$ in the first line comes from the $u_i + v_j$ in (i) of Proposition 4.) $\square$

4 Finishing the proof of Theorem 1

Now that we know the solution to the ATSP satisfies Lemma 12, we will give an algorithm to iterate over the possible ATSP solutions.

Lemma 13 *W.h.p., every non-basic edge in the optimal solution to the ATSP that is not in T^* has cost at most $2\zeta^*$.*

Proof. Let $\tilde{C}_1, \tilde{C}_2, \ldots, \tilde{C}_k$ be the reduced costs of the non-basic edges that are used to go from the optimal solution to AP to the optimal solution to the ATSP. We have $\tilde{C}_1 + \cdots + \tilde{C}_k \leq \zeta^*$, and we have $\tilde{C}_e \geq 0$ for every edge e, so in particular, every $\tilde{C}_e \leq \zeta^*$. Additionally, for $e = (i, j)$, we have $C_e = \tilde{C}_e + u_i + v_j \leq \zeta^* + 2\gamma^* + 2\gamma^* \leq 2\zeta^*$. □

4.1 Specifying a possible ATSP

A *valid* cycle or path is a cycle or path respectively in $K_{A,B}$ whose edges outside of T^* have length at most $2\zeta^*$. We know from Lemma 13 that the ATSP solution is formed from M^* by XORing it with valid cycles. For a valid cycle or path, let its *non-basic cardinality* be the number of edges that are in the cycle or path but not in T^*.

By Lemma 12, all sets of cycles that XOR with M^* to form the ATSP solution have non-basic cardinalities that total to at most $\log^{2+o(1)}(n)$.

Lemma 14 *With high probability, for all $1 \leq k \leq \log^3 n$, there are at most $\log^{3k+1} n$ valid cycles with non-basic cardinality k.*

Proof. Let N_k be the number of valid cycles with non-basic cardinality k. Let $\tilde{C}_j, j = 1, 2, \ldots, k$ be the non-basic costs in a generic valid cycle. Then,

$$\mathbf{E}(N_k) \leq \binom{m}{k} k! 2^k \mathbb{P}(\tilde{C}_j \leq 2\zeta^*, 1 \leq j \leq k) \leq \binom{m}{k} k! 2^k (3\zeta^*)^k \leq (6\zeta\omega \log^2 n)^k \leq \log^{3k} n.$$

We choose k alternating paths and order them in $k!2^k$ ways and multiply by the probability that the reduced cost of the non-basic edge joining them into a cycle is at most $2\zeta^*$.

The result then follows from the Markov inequality. □

We also have an analogous statement for paths instead of cycles:

Lemma 15 *With high probability, for all $1 \leq k \leq \log^3 n$, there are at most $n \log^{3k+2} n$ valid paths with non-basic cardinality k.*

Proof. This is proven just like Lemma 14, except that instead of choosing k alternating paths, we can now choose $k + 1$ alternating paths, giving an extra factor of $2\omega n \leq n \log n$ throughout. □

Now, the previous two lemmas were not algorithmic. To prove Theorem 1, we need to actually compute the valid cycles.

Lemma 16 *With high probability, we can find and store all valid cycles and paths with non-basic cardinality at most k in time $O(n^2 \log^{3k+6} n)$.*

Proof. We will prove this inductively. Our base case is that for $k = 1$, we can take all of the ωn (see Lemma 11) basic alternating paths and see which pairs form a valid path or cycle with non-basic cardinality 1. This takes $O(\omega^2 n^2 \log n)$ time.

Assume that we have already found and stored all valid paths and cycles of non-basic cardinality less than k. Now, every valid path of non-basic cardinality k is the concatenation of a valid path with non-basic cardinality $\lfloor \frac{k}{2} \rfloor$ and a valid path with non-basic cardinality $\lceil \frac{k}{2} \rceil$. So, we iterate over all pairs of valid paths, where the first has non-basic cardinality $\lfloor \frac{k}{2} \rfloor$ and the second has non-basic cardinality $\lceil \frac{k}{2} \rceil$. By Lemma 15, the total number of such pairs is at most

$$\left(n \log^{1.5k+2} n\right) \left(n \log^{1.5k+3.5} n\right) \le n^2 \log^{3k+6} n.$$

We store the pairs where the end-point of the first equals the start point of the second, as these are exactly the valid cycles with non-basic cardinality k.

Finally, we iterate through the valid paths with non-basic cardinality k to check which are cycles. □

4.2 Iterating through Possible ATSP Solutions

First, we precompute and store all valid cycles with non-basic cardinality at most the value given by Lemma 12, which by Lemma 16 takes time at most

$$O\left(\sum_{k=1}^{\log^{2+o(1)} n} n^2 \log^{3k+6} n\right) = O(n^2 \log^{3 \log^{2+o(1)} n + 6} n) = O(e^{\log^{2+o(1)} n}).$$

Then, we run through the possibilities for k, the number of edges in the ATSP but not in T^*. By Lemma 12, $k \le \log^{2+o(1)} n$ w.h.p.

Next, we run through the possibilities for ℓ, the number of distinct valid cycles made by adding the k edges into T^*, and XORing the number of times a tree edge is in one of these cycles. We then specify the non-basic cardinalities $(k_1, \ldots, k_\ell)$ of the ℓ valid cycles. Because $k_1 + \ldots + k_\ell = k$, we have that this step selects a partition of k, and thus we have (crudely) that there are at most $2^{2k} \le e^{\log^{2+o(1)} n}$ choices for $(k_1, \ldots, k_\ell)$, and thus at most $e^{\log^{2+o(1)} n}$ possibilities to iterate through in the outer loop.

Now, for a fixed $(k_1, \ldots, k_\ell)$, for each $1 \le i \le \ell$, we specify the ith valid cycle by choosing one of the at most $\log n^{3k_i+1} n$ pre-computed and stored valid cycles with non-basic cardinality k_i. Thus, the total amount of possibilities needed to iterate through for this particular

$(k_1, \dots, k_\ell)$ is at most

$$\prod_{j=1}^{k} \left(\log^{3j+1} n\right)^{|\{i:k_i=j\}|} \le (\log n)^{\sum_{j=1}^{k}(3j+1)|\{i:k_i=j\}|} \le (\log n)^{3\log^{2+o(1)} n + \log^{2+o(1)} n} \le e^{\log^{2+o(1)} n}$$

as desired. (We use the fact that $\sum_{j=1}^{\log^{2+o(1)} n} j|\{i : k_i = j\}| \le \sum_{j=1}^{\log^{2+o(1)} n} j(k/j) \le \log^{2+o(1)} n$.)

So we can iterate through each of these possible selections, check in polynomial time whether this gives a Hamilton cycle and if so evaluate its cost, and then remember the Hamilton cycle of minimum cost.

This finishes the proof of Theorem 1, subject to the generalized proofs of Lemmas 2 and 3 in Section 5.

4.3 Reducing the Space Complexity

A downside of the previous algorithm is that of storing all $e^{\log^{2+o(1)} n}$ valid paths and cycles with non-basic cardinality up to $\log^{2+o(1)} n$ uses $e^{\log^{2+o(1)} n}$ memory. The previous algorithm can be amended to use only the optimal amount of space, $O(n^2)$ (the amount of space needed to store the costs) without significantly increasing the time complexity:

Theorem 17 *With high probability, we can find the ATSP within $e^{\log^{2+o(1)} n}$ time and $O(n^2)$ space.*

Proof. Instead of precomputing and storing all valid cycles with non-basic cardinality up to $\log^{2+o(1)} n$, we instead only precompute and store all valid paths and cycles with non-basic cardinality up to $\frac{\log n}{6 \log\log n}$. By Lemma 14, there are at most

$$\left(\sum_{k=1}^{\log n/(6\log\log n)} n \log^{3k+2} n \right) \le 2n \log^{\log n/(2\log\log n)+2} n \le 2n^{1.5} \log^2 n$$

of these paths, each of which has length at most $\log^2 n$ (as alternating paths in T^* have length $O(\log n)$).

Now, as in the algorithm described in Section 4.2, we still iterate through all possible $(k_1, \dots, k_\ell)$. For the ith cycle, we specify it in one of two ways, depending on whether $k_i \le \frac{\log n}{6\log\log n}$ or whether $k_i > \frac{\log n}{6\log\log n}$.

Case 1: $k_i \le \frac{\log n}{6\log\log n}$. Then just as before we iterate through the $O(\log^{3k_i+1} n)$ stored cycles.

Case 2: $k_i > \frac{\log n}{6\log\log n}$. Now, we have not stored the valid cycles with non-basic cardinality k_i. Instead, we specify these cycles in a similar way to Lemma 16. In particular, we can specify any cycle with non-basic cardinality k_i by specifying $\lceil \frac{k_i}{\log n/6\log\log n} \rceil$ valid paths of non-basic cardinalilty at most $\frac{\log n}{6\log\log n}$ that concatenate to form this cycle. Using this process, as we had at most $2n^{1.5}\log^2 n \le n^2$ valid paths of non-basic cardinality at most $\frac{\log n}{6\log\log n}$, we have at most

$$n^{2(1+k_i/(\log n/6\log\log n))} = n^2 e^{12k_i \log\log n}$$

possibilities to iterate through in order to iterate through every valid cycle of non-basic cardinality k_i.

Then the total amount of time this new lower-space algorithm takes on a given $(k_1, \dots, k_\ell)$ is asymptotically at most

$$\begin{aligned}
&\prod_{j=1}^{\frac{\log n}{6\log\log n}} ((\log^{3j+1} n)^{|\{i:k_i=j\}|} \prod_{j=\frac{\log n}{6\log\log n}}^{\log^{2+o(1)} n} \left(n^2 e^{12k_i \log\log n}\right)^{|\{i:k_i=j\}|} \\
&\le (\log n)^{\sum_{j=1}^{k}(3j+1)|\{i:k_i=j\}|} \left(n^{2\left(\{\sum_{j=\frac{\log n}{6\log\log n}}^{\log^{2+o(1)} n} |\{i:k_i=j\}|\right)} \right) \left(e^{12\log\log n\left(\sum_{j=\frac{\log n}{6\log\log n}}^{\log^{2+o(1)} n} j|\{i:k_i=j\}|\right)} \right) \\
&\le (\log n)^{4\log^{2+o(1)} n} \left(n^{2\left(\left|\left\{i:k_i \ge \frac{\log n}{6\log\log n}\right\}\right)} \right) \left(e^{12\log\log n \log^{2+o(1)} n} \right) \\
&\le e^{\log^{2+o(1)} n} \left(n^{2\frac{\log^{2+o(1)} n}{\log n/6\log\log n}} \right) \left(e^{\log^{2+o(1)} n} \right) \\
&\le e^{\log^{2+o(1)} n}
\end{aligned}$$

as desired. □

5 Properties of the assignment problem

In this section, we verify Lemmas 2 and 3 under our more general distribution of costs.

5.1 M^* only has low cost edges

In this section we prove that w.h.p.,

$$\max_{e\in M^*} \{C(e)\} \le \gamma^* = \frac{\gamma \log n}{n} \text{ for some absolute constant } \gamma > 0. \tag{21}$$

Define the k-neighborhood of a vertex to be the k vertices nearest it, where distance is given by the matrix C. Let the k-neighborhood of a set be the union of the k-neighborhoods of its vertices. In particular, for a complete bipartite graph $K_{A,B}$ and any $S \subseteq A, T \subseteq B$,

$$N_k(S) = \{b \in B : \exists s \in S \text{ s.b. } (s,b) \text{ is one of the } k \text{ least cost edges incident with } s\}, \tag{22}$$
$$N_k(T) = \{a \in A : \exists t \in T \text{ s.t. } (a,t) \text{ is one of the } k \text{ least cost edges incident with } t\}. \tag{23}$$

Given the complete bipartite graph $K_{A,B}$, any permutation $\pi : A \to B$ has an associated matching $M_\pi = \{(a,b) : \ a \in A, \ b \in B, \ a = \pi(b)\}$. Given a cost matrix C and permutation π, define the digraph

$$\vec{D} = \vec{D}_{C,\pi} = (A \cup B, \ \vec{E}) \tag{24}$$

consisting of *backwards* matching edges and forward "short" edges:

$$\vec{E} = \{(b,a) : \ b \in B, \ a \in A, \ b = \pi(a)\} \cup \ \{(a,b) : \ a \in A, \ b \in N_{40}(a)\} \\ \cup \ \{(a,b) : \ b \in B, \ a \in N_{40}(b)\}. \tag{25}$$

The edges of directed paths in $\vec{D}$ are alternately forwards $X \to Y$ and backwards $Y \to X$ and so they correspond to *alternating paths* with respect to the perfect matching defined by π. Since "adding" an alternating cycle to a matching produces a new matching, finding low-cost alternating paths is key to our constructions. In particular, an alternating path's backward edges (from the old matching) will be replaced by its forward ones, and so it helps to know (Lemma 18, next) that given $x \in X, y \in Y$ we can find an alternating path from x to y with $O(\log n)$ edges. The forward edges have expected length $O(1/n)$ and we will be able to show (Lemma 20, below) that we can w.h.p. be guaranteed to find an alternating path from x to y in which the *difference* in weight between forward and backward edges is $O(\log n/n)$. It is then simple to prove the upper bound in Lemma 2. A long edge can be removed by the use of such an alternating path.

Lemma 18 *W.h.p. over random cost matrices C, for every permutation π, the (unweighted) diameter of $\vec{D} = \vec{D}_{C,\pi}$ is at most $k_0 = \lceil 3 \log_4 n \rceil$.*

Proof. This is Lemma 5 of [FS07]. □

If we ignore the savings from edge deletions in traversing an alternating path then it follows fairly easily that

$$\max \{C(i, \phi(i))\} \leq \frac{\gamma_1 \log^2 n}{n} \text{ for some absolute constant } \gamma_1 > 0. \tag{26}$$

For a fixed i we have

$$\mathbb{P}\left(C(i,j) \geq \frac{6 \log n}{n} \text{ for } j \in [n/2]\right) \leq \left(1 - \frac{6 \log n}{n} + O\left(\frac{\log^2 n}{n^2}\right)\right)^{n/2} = n^{-(3-o(1))}.$$

It follows that w.h.p. all of the forward edges in the paths alluded to in Lemma 18 have cost at most $\frac{6 \log n}{n}$. If $x \in A$ and $y \in B$ then Lemma 18 implies that w.h.p. there is a path from x to y for which the sum of the costs of the forward edges is at most $\frac{6k_0 \log n}{n}$. So if there is a matching edge of cost greater than $\frac{6k_0 \log n}{n}$ then there is an alternating path of using at most k_0 edges that can be used to give a matching of lower cost, contradiction. This verifies (26).

We now take account of the edges removed in an alternating path and thereby remove an extra $\log n$ factor. We will need the following inequality, analogous to Lemma 4.2(b) of [FG85], which deals with uniform $[0,1]$ random variables.

Lemma 19 *Suppose that* $k_1 + k_2 + \cdots + k_M = K \leq a \log N, a = O(1)$, *and* $Y_1, Y_2, \ldots, Y_M$ *are independent random variables with* Y_i *distributed as the* k_i*th minimum of* N *independent copies of* C. *If* $\lambda > 1, \lambda = O(1)$ *and* N *is large, then*

$$\mathbb{P}\left(Y_1 + \cdots + Y_M \geq \frac{\lambda a \log N}{N}\right) \leq N^{a(\alpha + \log \lambda - \theta a^{-1}\lambda)},$$

where $\theta = \frac{1}{2}\min\{1, \beta, L^{-1}\}$ *where* L *is the hidden constant in* $\mathbb{P}(C \leq x) = x + O(x^2)$ *for* $x \leq 1$.

Proof. The density function $f_k(x)$ of the kth order statistic $Y_{(k)}$ satisfies

$$f_k(x) = \binom{N}{k}(x + O(x^2))^{k-1}\left(1 - \frac{x}{2}\right)^{N-k} \qquad \text{for } x \leq \frac{1}{2L}.$$
$$f_k(x) \leq \alpha\binom{N}{k}(x + O(x^2))^{k-1}e^{-\beta(N-k)x} \qquad \text{for } x > \frac{1}{2L}.$$

Let $\widehat{\beta} = \min\{\beta, 1\}$. The moment generating function of $Y_{(k)}$ is given by

$$\begin{aligned}
\mathbf{E}(e^{tY_{(k)}}) &\leq \alpha^k \binom{N}{k} \int_{x \geq 0} e^{tx}(x + O(x^2))^{k-1} e^{-(N-k)\theta x} dx \\
&\leq \alpha^k \binom{N}{k} \int_{x \geq 0} x^{k-1}(1 + O(x))^{k-1} e^{-((N-k)\theta - t)x} dx \\
&\leq \alpha^k \binom{N}{k} \int_{x \geq 0} x^{k-1} e^{-((N-k)\theta - t - Lk)x} dx \\
&\leq \frac{(\alpha N)^k}{k((N-k)\theta - t - L)^k}.
\end{aligned}$$

So, if $Y = Y_1 + \cdots + Y_M$ then

$$\begin{aligned}
\mathbf{E}(e^{tY}) \leq \prod_{i=1}^{M}\left(\frac{(\alpha N)^{k_i}}{((N-k_i)\theta - t - L)^{k_i}}\right) &= \left(\frac{\alpha N}{\theta N - t}\right)^K \prod_{i=1}^{M}\left(1 + \frac{\theta k_i + L}{(N-k_i)\theta - t - c_i}\right) \\
&\sim \left(\frac{\alpha N}{\theta N - t}\right)^K = \lambda^K,
\end{aligned}$$

if we take $t = \theta N - \alpha\lambda^{-1}$.

So,

$$\mathbb{P}\left(Y \geq \frac{\lambda a \log N}{N}\right) \leq \mathbb{P}\left(e^{tY} \geq \exp\left\{\frac{t\lambda a \log N}{N}\right\}\right) \lesssim \frac{\lambda^K}{N^{\theta\lambda - \alpha a}}.$$

□

Given this lemma we can verify (21).

Lemma 20 *Equation* (2) *holds w.h.p.*

Proof. Let

$$Z_1 = \max\left\{\sum_{i=0}^{k} C(x_i,y_i) - \sum_{i=0}^{k-1} C(y_i,x_{i+1})\right\}, \tag{27}$$

where the maximum is over sequences $x_0, y_0, x_1, \ldots, x_k, y_k$ where (x_i, y_i) is one of the 40 shortest edges leaving x_i for $i = 0, 1, \ldots, k \leq k_0 = \lceil 3 \log_4 n \rceil$, and (y_i, x_{i+1}) is a backwards matching edge. Also, in the maximum we assume that all $C(\cdot,\cdot)$ are bounded above by $L = \frac{\gamma_1 \log^2 n}{n}$, see (26). We compute an upper bound on the probability that Z_1 is large. For any constant $\zeta > 0$ we have

$$\mathbb{P}\left(Z_1 \geq \frac{\xi \log n}{n}\right) \lesssim \sum_{k=0}^{k_0} n^{2k+2} \frac{1}{(n-1)^{k+1}} \times$$
$$\int_{y=0}^{L} \left[\frac{1}{(k-1)!}\left(\frac{y \log n}{n}\right)^{k-1} \sum_{\rho_0+\rho_1+\cdots+\rho_k \leq 40(k+1)} q(\rho_0, \rho_1, \ldots, \rho_k; \xi + y)\right] dy$$

where

$$q(\rho_0, \rho_1, \ldots, \rho_k; \eta) = \mathbb{P}\left(X_0 + X_1 + \cdots + X_k \geq \frac{\eta \log n}{n}\right),$$

$X_0, X_1, \ldots, X_k$ are independent and X_j is distributed as the ρ_jth minimum of $n-1$ copies of C. (When $k = 0$ there is no term $\frac{1}{k!}\left(\frac{y \log n}{n}\right)^k$).

Explanation: We have $\leq n^{2k+2}$ choices for the sequence $x_0, y_0, x_1, \ldots, x_k, y_k$. The term $\frac{1}{(k-1)!}\left(\frac{y \log n}{n}\right)^{k-1} dy$ asymptotically bounds the probability that the sum $\Sigma = C(y_0, x_1) + \cdots + C(y_{k-1}, x_k)$, is in $\frac{\log n}{n}[y, y+dy]$. Indeed, if $C_1, C_2, \ldots, C_k$ are independent copies of C then since $y \leq L$,

$$\mathbb{P}\left(C_1 + \cdots + C_k \in \frac{\log n}{n}[y, y+dy]\right) = \int_{z_1+\cdots+z_k \in \frac{\log n}{n}[y,y+dy]} \prod_{i=1}^{k}\left(1 + O\left(\frac{\log^2 n}{n}\right)\right) d\mathbf{z}$$
$$\sim \int_{z_1+\cdots+z_k \in \frac{\log n}{n}[y,y+dy]} 1 d\mathbf{z} = \frac{1}{(k-1)!}\left(\frac{y \log n}{n}\right)^{k-1} dy.$$

We integrate over y. $\frac{1}{n-1}$ is the probability that (x_i, y_i) is the ρ_ith shortest edge leaving x_i, and these events are independent for $0 \leq i \leq k$. The final summation bounds the probability that the associated edge lengths sum to at least $\frac{(\xi+y)\log n}{n}$.

It follows from Lemma 19 that if ξ is sufficiently large then, for all $y \geq 0$, $q(\rho_1, \ldots, \rho_k; \xi+y) \leq n^{-(\xi+y)/2}$ and since the number of choices for $\rho_0, \rho_1, \ldots, \rho_k$ is at most $\binom{41k+40}{k}$ (the number of non-negative integral solutions to $x_0 + x_1 + \ldots + x_{k+1} = 40(k+1)$) we have

$$\begin{aligned}
\mathbb{P}\left(Z_1 \geq \xi \frac{\log n}{n}\right) &\leq 2n^{2-\xi/2} \sum_{k=0}^{k_0} \frac{\log^{k-1} n}{(k-1)!} \binom{42k}{k} \int_{y=0}^{\infty} y^{k-1} n^{-y/2} dy \\
&\leq 2n^{2-\xi/2} \sum_{k=0}^{k_0} \frac{\log^{k-1} n}{(k-1)!} \left(\frac{42e}{\log n}\right)^k \Gamma(k) \\
&\leq 2n^{2-\xi/2} (42e)^{k_0+1} \\
&= o(n^{-2}).
\end{aligned}$$

If $a \in A$ and $b \in B$ then Lemma 18 implies that w.h.p. there is a path of length at most k_0 from a to b and by the above, it will w.h.p. have length at most $\frac{\xi \log n}{n}$. So if there is a matching edge of cost greater than $\frac{\xi \log n}{n}$ there is an alternating path of length at most k_0 that can be used to give a matching of lower cost, contradiction. □

5.2 A high probability bound on $Z_{\text{ATSP}} - Z_{\text{AP}}$

We now verify (2) with our more general distribution for costs. We let the $\widehat{C}(i,j)$ be independent copies of a uniform $[0,1]$ random variable and then let $C(i,j) = F^{-1}(\widehat{C}(i,j))$. Then we have

$$\begin{aligned}
C(ATSP) &\leq \left(1 + O\left(\frac{\log n}{n}\right)\right) \widehat{C}(ATSP) \\
&\leq \left(1 + O\left(\frac{\log n}{n}\right)\right) \left(\widehat{C}(AP) + O\left(\frac{\log^2 n}{n}\right)\right), \qquad \text{from (2)}, \\
&\leq \left(1 + O\left(\frac{\log^2 n}{n}\right)\right) \widehat{C}(AP) \\
&\leq C(AP) + O\left(\frac{\log^2 n}{n}\right).
\end{aligned}$$

6 From exponential mean one to more general distributions

Lemma 6 is only valid for the costs C being exponential with mean 1, as in general we only get that there is uniformity up to a polynomial factor in n, which is not good enough for the proof of Lemmas 9 or 11. However, we will now show that our result still holds on a more general class of distributions.

Suppose that the edge costs $C = C(i,j)$ are distributed as claimed in Theorem 1. We can naturally couple these edge costs to EXP(1) edge costs, which we will call X, as follows:

first sample a cost c from C, then find $p = \mathbb{P}(C \leq c)$, and set the cost x under X such that $1 - e^{-x} = p$ (in other words, drawing from the same places on the respective cumulative distribution functions).

We know $\mathbb{P}(C \leq z) = z + O(z^2)$ and $\mathbb{P}(X \leq z) = z + O(z^2)$ as $z \to 0$. We also know from Lemma 13 that under X, w.h.p. all edges in the AP solution have cost at most $2\zeta^* = O\left(\frac{\log^2 n}{n}\right)$. Suppose now that the costs of the edges in the optimal ATSP under X are U_i, $i = 1, 2, \ldots, n$ where each $U_i \leq 2\zeta^*$. Now we know that $\sum_{i=1}^n U_i = O(1)$ w.h.p. [FS07], which then implies that $\sum_{i=1}^n U_i^2 = O\left(\frac{\log^2 n}{n}\right)$. (We maximise the sum by putting $U_i = 2\zeta^*$ for at most $O(n/\log^2 n)$ indices and putting $U_i = 0$ for the remaining indices.) It follows that $C(ATSP) - X(ATSP) \leq O\left(\frac{\log^2 n}{n}\right)$. The same argument (noting Lemma 2 holds for both X and C) gives $|X(AP) - C(AP)| \leq O\left(\frac{\log n}{n}\right)$. Therefore, $|C(ATSP) - X(AP)| \leq O\left(\frac{\log^2 n}{n}\right)$, so our enumeration above (starting from the optimal AP under X and replacing ζ^* by $\frac{K \log^2 n}{n}$ for sufficiently large K) will also w.h.p. find the optimal ATSP under C.

7 Summary and open questions

One can easily put the enumerative algorithm in the framework of branch and bound. At each node of the B&B tree one branches by excluding edges of M^*. So, at the top of the tree the branching factor is n and in general, at level k, it is $n - k$. W.h.p. the tree will have depth at most $e^{\log^{4+o(1)} n}$.

The result of Theorem 1 does not resolve the question as to whether or not there is a branch and bound algorithm that solves ATSP w.h.p. in polynomial time. This remains an open question.

Less is known probabilistically about the symmetric TSP. Frieze [Fri04] proved that if the costs $C(i, j) = C(j, i)$ are independent uniform $[0, 1]$ then the asymptotic cost of the TSP and the cost 2F of the related 2-factor relaxation are asymptotically the same. The probabilistic bounds on $|TSP - 2F|$ are inferior to those given in [FS07]. Still, it is conceivable that the 2-factor relaxation or the subtour elimination constraints are sufficient for branch and bound to run in polynomial time w.h.p.

Yatharth Dubey [D23] pointed out that combining the above analysis with arguments from [DDM21] shows that using the subtour elimination LP relaxation in a branch and bound algorithm will also lead to a quasi-polynomial time algorithm w.h.p.

7.1 Acknowledgement

We thank an anonymous referee for helping to correct a previous version.

References

[AMO93] R.K. Ahuja, T.L. Magnanti and J.B. Orlin, *Network Flows Theory, Algorithms and Applications*, Prentice-Hall, 1993.

[Ald90] D. Aldous, The Random Walk Construction of Uniform Spanning Trees and Uniform Labelled Trees, *SIAM Journal on Discrete Mathematics* 3 (1990) 450-465.

[BT86] E. Balas and P. Toth, *Branch and Bound Methods*, in The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization (E.L. Lawler, J.K. Lenstra, A.H. Rinnooy Kan and D.B. Shmoys (editors)), Wiley, 1986, 361–400.

[BM71] M. Bellmore and J. Malone, Pathology of Traveling-Salesman subtour elimination algorithms, *Operations Research* 19 (1971) 278-307.

[BDHT23] S. Borst, D. Dadush, S. Huiberts and S. Tiwari, On the Integrality Gap of Binary Integer Programs with Gaussian Data, *Mathematical Programming* 197 (2023) 1221-1263.

[Bro89] A. Broder, Generating random spanning trees, *Proceedings of the 30th Annual Symposium on Foundations of Computer Science* (1989) 442-447.

[DDM21] S. Dey, Y. Dubey and M. Molinaro, Branch-and-Bound Solves Random Binary Packing IPs in Polytime, *Proceedings of the 32nd Annual ACM-SIAM Symposium on Discrete Algorithms* (2021) 579–591.

[D23] Y. Dubey, private communication.

[DF90] M.E. Dyer and A.M. Frieze, On patching algorithms for random asymmetric traveling salesman problems, *Mathematical Programming* 46 (1990) 361–378.

[EK72] J. Edmonds and R.M. Karp, Theoretical Improvements in Algorithmic Efficiency for Network Flow Problems, *Journal of the ACM* 19 (1972) 248–264.

[Fri04] A.M. Frieze, On random symmetric travelling salesman problems, *Mathematics of Operations Research* 29 (2004) 878-890.

[FG85] A.M. Frieze and G. Grimmett, The shortest path problem for graphs with random arc-lengths, *Discrete Applied Mathematics* 10 (1985) 57-77.

[FKR95] A.M. Frieze, R.M. Karp and B. Reed, When is the assignment bound asymptotically tight for the asymmetric traveling-salesman problem?, *SIAM Journal on Computing* 24 (1995) 484–493.

[FS07] A.M. Frieze and G. Sorkin, The probabilistic relationship between the assignment and asymmetric traveling salesman problems, *SIAM Journal on Computing* 36 (2007) 1435-1452.

[HK62] M. Held and R.M. Karp, A dynamic programming approach to sequencing problems, *SIAM Journal of Applied Mathematics* 10 (1962) 196–210.

[Kal05] P. Kalczynski, A Java implementation of the Branch and Bound Algorithm: The Asymmetric Traveling Salesman Problem, *Journal of Object Technology* 4 (2005).

[Kar79] R.M. Karp, A patching algorithm for the non-symmetric traveling salesman problem, *SIAM Journal on Computing* 8 (1979) 561–573.

[KS85] R.M. Karp and J.M. Steele, Probabilistic analysis of heuristics, in *The traveling salesman problem: a guided tour of combinatorial optimization*, E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan and D.B. Shmoys Eds. (1985) 181–206.

[LR79] J. Lenstra and A. Rinnooy Kan, On the expected performance of branch and bound algorithms, *Operations Research* 26 (1979) 347-349.

[LP17] R. Lyons and Y. Peres, Probability on trees and networks, Cambridge University Press, 2017.

[MM70] A. Meir and J. Moon, The distance between points in random trees, *Journal of Combinatorial Theory* 8 (1970) 99-103.

[MP91] D. Miller and J. Pekny, Exact solution of large asymmetric traveling salesman problems, *Science* 251 (1991) 754-761.

[Tom71] N. Tomizawa, On some techniques useful for solution of transportation network problems, *Networks* 1 (1971) 173–194.

[Zha97] W. Zhang, A note on the complexity of the Asymmetric Traveling Salesman Problem, *Operations Research Letters* 20 (1997) 31-38.

[Zha04] W. Zhang, Phase Transitions and Backbones of the Asymmetric Traveling Salesman problem, Journal of Artificial Intelligence 21 (2004) 471-497.