# 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

December 14, 2025

**Abstract**

Let the costs $C(i,j)$ for an instance of the Asymmetric Traveling Salesperson Problem (ATSP) be independent copies of a non-negative random variable $C$ from a class of distributions that include the uniform $[0,1]$ distribution and the exponential mean 1 distribution with mean 1. We describe an algorithm that solves ATSP exactly in time $e^{\log^{2+o(1)} n}$, w.h.p.

## 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 $\pi$ on $n$ elements that minimizes $\sum_{i=1}^{n} C(i,\pi(i))$. We let $Z_{\text{ATSP}} = Z_{\text{ATSP}}^{(C)}$ denote the optimal cost for ATSP. As the ATSP is NP-hard in general, the goal of this paper is to analyze whether the ATSP can be solved efficiently when $(C(i,j))_{i,j\in[n]}$ is an average-case, rather than worst-case, matrix. We consider the situation where the costs $C(i,j)$ are independent copies of a continuous random variable $C$. We assume that $C$ has a density $f$ and satisfies

(i) $f(x) = a + bx + O(x^2)$ for $0 \le x \le L$, where $a, b$ are constants and $aL \ge 1$.

(ii) $\mathbb{P}(C \ge x) \le \alpha e^{-\beta x}$ for constants $\alpha, \beta > 0$.

*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

(iii) To avoid some pathologies, we will also assume that there is a constant $M$ such that $f(x) \leq aM$ for $x \geq 0$.

The prime examples are the uniform $[0,1]$ distribution $U[0,1]$ $(a=1, b=0, M=1)$ and the exponential mean 1 distribution $EXP(1)$ $(a=1, b=-1, \alpha=\beta=M=1)$. The main result of this paper is as follows:

**Theorem 1** *Let the costs for ATSP, $(C(i,j))$, be independent copies of $C$, where $C$ has a distribution satisfying (i),(ii),(iii) above. 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 another discrete optimization problem. Let $S_n$ denote the set of permutations of $[n] = \{1, 2, \ldots, n\}$ . The *Assignment Problem* (AP) is the problem of minimising $C(\pi) = \sum_{i=1}^{n} C(i, \pi(i))$ over all permutations $\pi \in S_n$, (while the ATSP only optimizes over cyclic permutations). We let $Z_{\text{AP}} = Z_{\text{AP}}^{(C)}$ denote the optimal cost for the AP.

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 $A = \{a_1, a_2, \ldots, a_n\}$, $B = \{b_1, b_2, \ldots, b_n\}$, and the cost of edge $(a_i, b_j)$ is $C(i,j)$.

It is evident that $Z_{\text{AP}}^{(C)} \leq Z_{\text{ATSP}}^{(C)}$. 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 gives a rigorous explanation for the previous observation that the AP often gives extremely good bounds on the ATSP. 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 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

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

Frieze and Sorkin also used this result to give an $e^{n^{1/2+o(1)}}$ time algorithm to solve the ATSP exactly w.h.p. The main result of this paper is that we significantly improve this run-time, reducing the exponent from $n^{1/2+o(1)}$ to $\log^{2+o(1)} n$. Our result also allows for a larger class of random distributions for $C$.

One might think that with such a small gap between $Z_{\text{AP}}$ and $Z_{\text{ATSP}}$, 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.

### 1.2 Chernoff bounds

We use the following Chernoff bounds on the tails of the binomial distribution $Bin(n,p)$: here $0 \leq \epsilon \leq 1$.

$$\mathbb{P}(Bin(n,p) \leq (1-\epsilon)np) \leq e^{-\epsilon^2 np/2}.$$
$$\mathbb{P}(Bin(n,p) \geq (1+\epsilon)np) \leq e^{-\epsilon^2 np/3}.$$

We also use McDiarmid's inequality: let $Z = Z(Y_1, Y_2, \ldots, Y_n)$ where $Y_1, Y_2, \ldots, Y_n$ are independent random variables. Suppose that $|Z(Y_1, Y_2, \ldots, Y_n) - Z(\widehat{Y}_1, \widehat{Y}_2, \ldots, \widehat{Y}_n)| \leq c$ when $Y_i = \widehat{Y}_i$ except for exactly one index. Then

$$\mathbb{P}(|Z - \mathbf{E}(Z)| \geq t) \leq \exp\left\{-\frac{2t^2}{c^2 n}\right\}. \tag{3}$$

## 2 The Assignment Problem and Nearby Permutations

Frieze and Sorkin [FS07] proved that when the distribution of costs is $U[0,1]$, 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** $Z_{\text{ATSP}} - Z_{\text{AP}} \leq \frac{\zeta \log^2 n}{n}$ *for some absolute constant* $\zeta > 0$.

In Section 6, 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. We will use $\gamma^*$ to denote $\frac{\gamma \log n}{n}$ and $\zeta^*$ to denote $\frac{\zeta \log^2 n}{n}$ for the corresponding $\gamma$ and $\zeta$ under which we prove Lemmas 2 and 3.

## 2.1 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.

$$\begin{aligned} \mathcal{LP} \qquad \text{Minimise } & \sum_{(i,j)\in[n]^2} C(i,j) z_{i,j} \\ \text{subject to } & \sum_{j=1}^{n} z_{i,j} = 1, \text{ for } i = 1,2,\ldots,n. \\ & \sum_{i=1}^{n} z_{i,j} = 1, \text{ for } j = 1,2,\ldots,n. \\ & z_{i,j} \geq 0, \text{ for } (i,j) \in [n]^2. \end{aligned} \tag{4}$$

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 $2n-1$ edges of $T^*$ are referred to as *basic* edges and the edges not in $T^*$ are referred to as non-basic edges. In the dual formulation, w.h.p., $T^*$ is consists of those $(i,j) \in [n]^2$ for which $u_i + v_j = C(i,j)$.

$\mathcal{LP}$ has the dual linear program:

$$\begin{aligned} \mathcal{DLP} \qquad \text{Maximise } & \sum_{i=1}^{n} u_i + \sum_{j=1}^{n} v_j \\ \text{subject to } & u_i + v_j \leq C(i,j), \text{ for } (i,j) \in [n]^2. \end{aligned} \tag{5}$$

**Remark 4** *Note that replacing $u_i, v_j$ by $u_i+\lambda, v_j-\lambda$ for all $i,j$ does not affect the constraints or the objective value. We can therefore, when necessary, choose an $s$ and assume that $u_s = 0$ for some $s \in [n]$.*

**Proposition 5** *Condition on an optimal basis for (4), that is, fix $T^*$ (which is unique with probability 1) and fix $C(i,j)$ for all $(a_i, b_j) \in E(T^*)$, but do not yet fix $C(i,j)$ for $(a_i, b_j) \notin E(T^*)$. We may w.l.o.g. take $u_1 = 0$ in (5), 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 Outline Proof of Theorem 1

Recall the complete bipartite graph $K_{A,B}$ where $A = \{a_1, a_2, \ldots, a_n\}$, $B = \{b_1, b_2, \ldots, b_n\}$, and the cost of edge $(a_i, b_j)$ is $C(i, j)$. Let $M^*$ denote the minimum-cost perfect matching in $K_{A,B}$, which is the solution to the AP. 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 different" from $M^*$, then w.h.p. $C(M) - C(M^*) > \frac{\zeta \log^2 n}{n}$ (where $\zeta$ is from (2)), and thus by (2), $M$ cannot be the optimal ATSP solution. Once we have shown this, it does not take too long to check all possible $M$ that are "similar to" $M^*$, to see if $M$ defines a tour and then determine its total cost.

More specifically, let $T^* \supseteq M^*$ be the spanning tree representing the optimal basis in the LP formulation of AP (see Section 2.1). Our definition of "too different" is that $M$ contains more than $\log^{2+o(1)} n$ edges outside $T^*$. Each such edge corresponds to a non-basic variable. These reduced costs turn out to be conditionally independent. This makes it easy to show that the sum of "many" of them is greater than $\frac{\zeta \log^2 n}{n}$ w.h.p. Lemma 15 shows that this will w.h.p. preclude $M$ from corresponding to the optimum to ATSP. This is because the sum of the reduced costs associated with these non-$T^*$ edges will be greater than the upper bound in (2) w.h.p.

In Sections 4 and 5, we will prove Theorem 1 in the special case where 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 7. In Section 7, we will subsequently generalise our proof to the full class of distributions in Theorem 1.

## 4 Trees and bases

We have that for every optimal basis $T^*$,

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

and

$$C(i, j) > u_i + v_j \text{ for } (a_i, b_j) \notin E(T^*). \tag{7}$$

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 $S \subseteq A, T \subseteq B$,

$$N_k(S) = \{b \in B : \exists s \in S \text{ s.t. } (s, b) \text{ is one of the } k \text{ least cost edges incident with } s\}, \tag{8}$$
$$N_k(T) = \{a \in A : \exists t \in T \text{ s.t. } (a, t) \text{ is one of the } k \text{ least cost edges incident with } t\}. \tag{9}$$

Given the complete bipartite graph $K_{A,B}$ and a perfect matching $M$, we define a directed graph $\vec{D}_M$ as follows: it has *backwards* matching edges $\vec{E}_M$ and forward "short" edges $\vec{E}_{\bar{M}}$, where

$$\vec{E}_M = \{(b,a) : \ b \in B, \ a \in A, \ \{a,b\} \in M\}.$$
$$\vec{E}_{\bar{M}} = \{(a,b) : \ a \in A, \ b \in N_{40}(a)\} \cup \ \{(a,b) : \ b \in B, \ a \in N_{40}(b)\}.$$

Paths in $\vec{D}_M$ are necessarily *alternating*, i.e. the edges alternate between being in $M$ and not in $M$. We assign a cost $C(i,j)$ to edge $(a_i,b_j) \in \vec{E}_{\bar{M}}$ and a cost $-C(i,j)$ to an edge $(b_j,a_i) \in \vec{E}_M$. The cost of a path $P$ being the sum of the costs of the edges in the path equals the change in the cost of a matching due to replacing $P \cap \vec{E}_M$ by $P \cap \vec{E}_{\bar{M}}$. Lemma 23 below proves that w.h.p. for every perfect matching $M$, $a \in A, b \in B$,

$$\vec{D}_M \text{ contains a path } P \text{ from } a \text{ to } b \text{ of cost at most } \gamma^*. \tag{10}$$

Given this we can prove a high probability upper bound on the $u_i, v_j$.

**Lemma 6** *Let* $\mathbf{u}$,$\mathbf{v}$ *be optimal dual variables and suppose that* $u_1 = 0$. *Then,*

$$|u_i|, |v_i| \leq 2\gamma^* \text{ for } i \in [n], \text{ w.h.p.} \tag{11}$$

**Proof.** Fix $a_i, b_j$ and let $P = (a_{i_1}, b_{j_1}, \ldots, a_{i_k}, b_{j_k})$ be the alternating path from $a_i$ to $b_j$ promised by Lemma 23. Then, using (6) and (7), we have

$$\gamma^* \geq C(P) = \sum_{l=1}^{k} C(i_l, j_l) - \sum_{l=1}^{k-1} C(i_{l+1}, j_l) \geq \sum_{l=1}^{k}(u_{i_l} + v_{j_l}) - \sum_{l=1}^{k-1}(u_{i_{l+1}} + v_{i_l}) = u_i + v_j. \tag{12}$$

For each $i \in [n]$ there is some $j \in [n]$ such that $u_i + v_j = C(i,j)$. This is because of the fact that $a_i$ meets at least one edge of $T^*$ and we assume that (6) holds. We also know that from (12) that $u_{i'} + v_j \leq \gamma^*$ for all $i' \neq i$. It follows that $u_i - u_{i'} > C(i,j) - \gamma^* \geq -\gamma^*$ for all $i' \neq i$. Since $i$ is arbitrary, we deduce that $|u_i - u_{i'}| \leq \gamma^*$ for all $i, i' \in [n]$. Since $u_1 = 0$, 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 $j, j' \in [n]$. Suppose that the optimal matching

$$M^* = \left\{ \left\{ a_i, b_{\phi(i)} \right\} : i \in [n] \right\}.$$

Then we have $u_i + v_{\phi(i)} = C(i, \phi(i))$, and because (as will be shown in Section 6) $C(i,\phi(i)) \leq \gamma^*$, we have $|v_j| \leq 2\gamma^*$ for $j \in [n]$. □

Condition on the edges of the matching $M^*$ 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 $M^*$. Note that $\mathbf{u}$,$\mathbf{v}$ do not determine $T^*$. Let $f(\mathbf{u}, \mathbf{v})$ denote the joint density of $\mathbf{u}, \mathbf{v}$ then

**Lemma 7** *If* $T \in \mathcal{T}_+$ *then*

$$\mathbb{P}(T^* = T \mid \mathbf{u}, \mathbf{v}) f(\mathbf{u}, \mathbf{v}) = \prod_{(a_i, b_j) \in G_+} e^{-u_i - v_j}, \tag{13}$$

*which is independent of* $T$.

**Proof.** Fixing **u**,**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. If $(a_i, b_j) \in E(T)$ then the density of $C(i,j)$ is $e^{-(u_i+v_j)}$. Thus, Proposition 5 and the "memoryless property" of the exponential distribution imply

$$\mathbb{P}(T^* = T \mid \mathbf{u}, \mathbf{v}) f(\mathbf{u}, \mathbf{v}) = \prod_{(a_i,b_j) \in G_+ \setminus E(T)} e^{-(u_i+v_j)} \prod_{(a_i,b_j) \in E(T)} e^{-(u_i+v_j)}$$
$$= \prod_{(a_i,b_j) \in G_+} e^{-(u_i+v_j)}. \tag{14}$$

In the first product we use (7), and the second product comes from (6) and from the density function of the costs. □

Thus

$$T^* \text{ is a uniform random member of } \mathcal{T}_+. \tag{15}$$

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

**Lemma 8** *The distribution of the tree $\widehat{T}^*$ is equal to that of a random spanning tree of the complete graph $K_n$, plus $\widehat{M}$ where w.h.p. $\widehat{M}$ is a matching of size at most $\lambda^* = \log^5 n$. ($\widehat{M}$ 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 $\widehat{G}_+$ contains the edge $\{a_i, a_j\}$ and so $\widehat{G}_+$ contains $K_A$, the complete graph on $A$, as a subgraph. So from (15), $\widehat{T}^*$ consists of a random spanning tree of $K_A$ plus a set of edges $\widehat{M}$. The edges $\widehat{M}$ arise when both $u_i + v_{\phi(j)} > 0$ and $u_j + v_{\phi(i)} > 0$ and where both edges have cost at most $4\gamma^*$.

We know from (6) and Lemma 6 that $\widehat{T}^*$ only contains edges of cost at most $4\gamma^*$. Thus each repeated edge arises from a cycle of length 4 $(a_i, b_{\phi(j)}, a_j, b_{\phi(i)}, a_i)$ in which each edge has length at most $4\gamma^*$. The expected number of such cycles is $O((n\gamma^*)^4)$ and so by the Markov inequality, $|\widehat{M}| \leq \log^5 n$ 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. □

We need the following lemma:

**Lemma 9** *Let $T$ be a random spanning tree of $K_n$. Let $\mathcal{P}$ be the set of paths in $T$ of the form $x_1, x_2, x_3, x_4$ such that $x_2, x_3, x_4$ are of degree 2 and let $X = |\mathcal{P}|$. Then w.h.p. $X \geq 2e^{-3}n/3$.*

**Proof.** We use the following relationship between random spanning trees and random mappings from Joyal [J82] (this paper is quite long and so we give details of the exact relationship). We begin with a random mapping $\psi : [n] \to [n]$ and construct the digraph $D_\psi$ with vertex set $[n]$ and edges $(i, \psi(i)) : i \in [n]$. Given the cycles $C_1, C_2, \ldots, C_r$ of $D_\psi$ we interpret them as the cycles of a permutation $x_1, x_2, \ldots, x_s$ on $V_C = \bigcup_{i=1}^r V(C_i)$. We then let $P$ be the path $x_1, x_2, \ldots, x_s$ and we let $T_i$ be the sub-tree pointing into $x_i$ in $D_\psi$. We then construct a tree $T_\psi$ by adding $T_i$ rooted at $x_i, i = 1, 2, \ldots, s$ to $P$. Ignoring orientation, $T_\psi$ is a uniform random spanning tree of $K_n$.

Let $Y$ be the number of paths of the form $x_1, x_2, x_3, x_4$ such that $x_1, x_2, x_3 \in U$ where $U = \{v : |\psi^{-1}(\psi(v))| = 1\}$ i.e. where $x_1, x_2, x_3$ have unique pre-images under $\psi$. Now $\mathbf{E}(Y) \sim e^{-3}n$ and changing $\psi(w)$ for some $w \in [v]$ changes $Y$ by at most 4. It follows from McDiarmid's inequality (3) with $c = 4$ and $t = n^{2/3}$ that $Y \geq e^{-3}n/2$ w.h.p. (Here $A \sim B$ stands for $A = (1 + o(1))B$, assuming that $A, B$ are functions of $n$ and that $n \to \infty$.)

Now when we delete the edges of the cycles of $D_\psi$, we delete $o(n)$ edges w.h.p. (see for example [FK15], Chapter 14), and so $Y \geq e^{-3}n/3$ w.h.p. □

**Corollary 10** *Let $T$ be a random spanning tree of $K_n$. Wh.p. for every $v \in [n]$ there are at least $n/400$ edge disjoint paths $P = (x_1, x_2, x_3, x_4)$ of length 3 such that there is a path from $v$ in $T$ that finishes with $P$ or its reversal. Denote such a path as being* useful.

**Proof.** Let $\mathcal{P}$ be as in Lemma 9. A path $P \in \mathcal{P}$ can share edges with at most 4 other members of $\mathcal{P}$ and so we can find $|\mathcal{P}|/5$ edge disjoint members of $\mathcal{P}$. Given $v$, all but at most one of these paths will contribute to the set of paths described in the statement of the corollary. Giving us at least $e^{-3}n/15 - 1 \geq n/400$ paths. □

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 $i = 1, 2, \ldots, n$ let $L_{i,+} = \{j : u_i + v_j \geq 0\}$ and let $L_{j,-} = \{i : u_i + v_j \geq 0\}$. Then let $\mathcal{A}_{i,+} = \mathcal{A}_{i,+}(\eta)$ be the event that $|L_{i,+}| \leq \eta n$ and let $\mathcal{A}_{j,-}(\eta)$ be the event that $|L_{j,-}| \leq \eta n$, where $\eta > 0$ is some small positive constant. For any tree $T$, let $\mathcal{B} = \mathcal{B}(T)$ be the event that $C(i, j) > u_i + v_j$ for $(a_i, b_j) \notin E(T)$. Note that if $\mathcal{B}$ occurs and (6) holds then $T^* = T$.

The following lemma shows that the minimum degree in $G_+, \widehat{G}_+$ is $\Omega(n)$. This enables us to prove that there are no long alternating paths made up entirely of basic edges, see Lemma 12.

**Lemma 11** *Fix a spanning tree $T$ of $G_+$. Let $\mathcal{E}$ be the event $\{|u_i|, |v_i| \leq 2\gamma^*$ for $i \in [n]\}$ as shown to hold w.h.p. in Lemma 6. Then there exists $\eta > 0$ such that*

$$\mathbb{P}\left(\exists i : (\mathcal{A}_{i,+} \vee \mathcal{A}_{i,-}) \wedge \mathcal{E} \mid T^* = T\right) = o(1). \tag{16}$$

**Proof.** We can assume as in Lemma 8 that $C(i, j) \leq 4\gamma^*$ for $(a_i, b_j) \in T$. 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 fix $u_s = 0$ and write,

$$\mathbb{P}(\mathcal{A}_{s,+} \wedge \mathcal{E} \mid \mathcal{B}) = \frac{\int \delta_1(Y)\mathbb{P}(\mathcal{B} \mid Y)dY}{\int \mathbb{P}(\mathcal{B} \mid Y)dY} \tag{17}$$

Then we note that since $(a_i, b_j) \notin E(T)$ satisfies the condition (7),

$$\mathbb{P}(\mathcal{B} \mid Y) = \prod_{(a_i,b_j)} (\mathbb{P}(C(i,j) \geq (u_i(Y) + v_j(Y))^+)) \\ = e^{-W}, \tag{18}$$

where $W = W(Y) = \sum_{(a_i,b_j)} (u_i(Y) + v_j(Y))^+$.

We first observe that McDiarmid's inequality implies that

$$\mathbb{P}(|W - \mathbf{E}(W)| \geq t) \leq 2\exp\left\{-\frac{2t^2}{(2n-1)(4n\gamma^*)^2}\right\} \leq 2\exp\left\{-\frac{t^2}{16n\gamma^2 \log^2 n}\right\}. \tag{19}$$

To see this, we view the random variable $W$ as a function of $2n-1$ random variables, each independently distributed as $EXP(1)$ conditioned on being at most $4\gamma^*$. (The variables are the costs of the tree edges.) If we change the value of one variable then we change $W$ by at most $4n\gamma^*$. To see this, suppose that in this change the cost of edge $e = \{a_{i_1}, b_{j_1}\}$ goes from $C(e)$ to $C(e) + \xi, |\xi| \leq 2\gamma^*$. The effect on $\mathbf{u}$,$\mathbf{v}$, under the assumption that $u_{i_1}$ does not change is as follows: (i) $v_j \leftarrow v_j + \xi$ for all $j \in [n]$ and (ii) $u_i \leftarrow u_i - \xi$ for all $i \in [n] \setminus \{i_1\}$. So, $u_i + v_j$ changes only for $i = i_1$.

We have, using Holder's inequality with $p = n^{3/4}$, that

$$\begin{aligned} \int_Y \delta_1(Y)\mathbb{P}(\mathcal{B} \mid Y)\, dY &= \int_Y e^{-W}\delta_1(Y)\, dY \\ &\leq \left(\int_Y e^{-pW/(p-1)}\, dY\right)^{(p-1)/p} \left(\int_Y \delta_1(Y)^p\, dY\right)^{1/p} \\ &= e^{-\mathbf{E}(W)} \left(\int_Y e^{-p(W-\mathbf{E}(W))/(p-1)}\, dY\right)^{(p-1)/p} \left(\int_Y \delta_1(Y)^p\, dY\right)^{1/p}. \end{aligned} \tag{20}$$

We also have

$$\int_Y \mathbb{P}(\mathcal{B} \mid Y) dY = e^{-\mathbf{E}(W)} \int_Y e^{-(W-\mathbf{E}(W))} dY \tag{21}$$

Putting $t = n^{2/3}$ in (19), we see that if $\Omega_1 = \left\{Y : |W - \mathbf{E}(W)| \leq n^{2/3}\right\}$ then $Y \in \Omega_1$ w.h.p. Conditioning on $Y \in \Omega_1$ we have that since $p = n^{3/4}$,

$$\int_{Y\in\Omega_1} e^{-p(W-\mathbf{E}(W))/(p-1)}\, dY \sim \int_{Y\in\Omega_1} e^{-(W-\mathbf{E}(W))}\, dY.$$

Combining this with (17), (20) and (21) we see that

$$\mathbb{P}(\mathcal{A}_{s,+} \wedge \mathcal{E} \mid \mathcal{B}) \lesssim \mathbb{P}(\mathcal{A}_{s,+} \wedge \mathcal{E})^{1/p} \sim \mathbb{P}(\mathcal{A}_{s,+} \mid \mathcal{E})^{1/p}. \tag{22}$$

Due to equation (22), for the remainder of this proof, we will now only take our edges to be selected conditioned on $\mathcal{E}$ holding and that the $C(i,j)$ for $(a_i, b_j) \in T$ satisfy $C \leq \gamma^*$. Denote this conditioning by $\mathcal{F}$.

Fix $j$ 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 (6) 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 $\mathcal{F}$. Assume for the moment that $k \geq 3$ and let $x = u_{i_{k-3}} \in [-2\gamma^*, 2\gamma^*]$. Given $x$ we see that there is some positive probability $p_0 = p_0(x) = \mathbb{P}(x + X_{k-2} + X_{k-1} + C(i_k, j_k) > 0 \mid \mathcal{F})$. Let $\eta_0 = \min\{x \geq -2\gamma^* : p_0(x)\}$ and note that $\eta_0 > 0$ and bounded away from 0.

The number of edges in $G_+$ of cost at most $\gamma^*$ 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 $2\gamma \log n$. So the number of vertices in $T$ at distance at most 3 from $a_s$ in $T$ is $O(\log^4 n)$. This justifies assuming $k \geq 3$ in the context of the next claim. Corollary 10 implies that there are $n/400 - O(\log^4 n)$ choices of $b_j$ giving rise to edge disjoint useful paths. We know that each such $j$ belongs to $L_{i,+}$ with probability at least $\eta_0$, conditional on $\mathcal{E}$, but conditionally independent of the other useful $j$'s. Using equation (22) and the Chernoff bounds we see that

$$\mathbb{P}(\mathcal{A}_{s,+} \wedge \mathcal{E} \mid \mathcal{B}) \lesssim \mathbb{P}(Bin(n/401, \eta_0) \leq \eta_0 n/500)^{1/p} \leq e^{-\Omega(n-o(n)/p)}.$$

It follows that by taking the union bound over $i \in [n]$ and +,-, we have

$$\mathbb{P}\left(\exists i : (\mathcal{A}_{i,+} \vee \mathcal{A}_{i,-}) \wedge \mathcal{E} \mid T^* = T\right) \leq o(1) + 2ne^{-\Omega(n-o(n)/p)} = o(1).$$

This proves the lemma with $\eta = \eta_0/500$. □

So, from now on we assume that $\mathcal{E}$ occurs and that $\mathcal{A}_{s,\pm}$ does not occur for any $s = 1, 2, \ldots, n$.

## 4.1 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 12** *The expected number of basic alternating paths with $k$ edges is at most $n^2 \left(1 - \frac{\eta}{1+\eta}\right)^k$, where $\eta$ is as in Lemma 11.*

**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 $\widehat{T}^*$, which is defined in Lemma 8 above. The edge $\left\{a_{i_t}, a_{i_{t+1}}\right\}$ 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 in $T^*$ only if we have the former case for $t = 1, 2, \ldots, k$.

Consider the random walk construction of a spanning tree as described in Aldous [Ald90] and Broder [Bro89]. Here a spanning tree is constructed by including the edges of a random walk that are used to first visit each vertex. 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. ($\eta$ from Lemma 11.) 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). The degree bound in Lemma 11 implies that when the walk adds an edge to the tree there is a probability of at least $\eta$ that the edge arises from case (ii) above. The probability of an even success is therefore at least

$$\sum_{k\geq 1}\left(1-\frac{n}{W+n}\right)^k\left(1-\frac{\eta n}{W+n}\right)^{k-1}\cdot\frac{\eta n}{W+n} = \left(1-\frac{n}{W+n}\right)\cdot\frac{\eta n}{W+n}\cdot\frac{1}{1-\left(1-\frac{n}{W+n}\right)\left(1-\frac{\eta n}{W+n}\right)}\sim\frac{\eta}{1+\eta}. \tag{23}$$

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 13** *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 14** *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 $\ell$ edges of $\widehat{M}$ (see Lemma 8) then its probability of occurring in $\widehat{G}_+$ is inflated by $2^\ell$. On the other hand, the probability that a random tree in $K_A$ contains $\ell$ given edges is at most $(2/n)^\ell$. (This is $2/n$ for $\ell=1$ and at most $(2/n)^\ell$ in general using negative correlation, see [LP17].) So, assuming $|\widehat{M}|\leq\lambda\log^5 n$ (see Lemma 8), the expected number of basic alternating paths can be bounded by

$$n^2\sum_{k=1}^n\frac{k}{n-1}\cdot\left(1-\frac{\eta}{1+\eta}\right)^k\cdot\sum_{\ell\geq 0}\binom{\lambda\log^5 n}{\ell}\left(\frac{4}{n}\right)^\ell\leq\frac{2n}{\eta}.$$

This finishes the proof when combined with the Markov inequality and Corollary 13. □

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 15** *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 $\ell$ cycles $Q_1, Q_2, \ldots, Q_\ell$ where each $Q_i$ has $k_i$ edges outside $T^*$ where $k_1 + \ldots + k_\ell = k$. 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. $M$ can correspond to the ATSP solution only if the $\tilde{C}$ of the $k$ corresponding non-basic edges sum to at most $\zeta^*$. Thus, if $\tilde{C}_1, \tilde{C}_2, \ldots, \tilde{C}_k$ are independent random variables distributed as (i), (ii) of Proposition 5, we have

$$\begin{aligned} \mathbf{E}(Z_k) &\leq \sum_{\ell=1}^{k} \sum_{k_1+\cdots+k_\ell=k} \prod_{i=1}^{\ell} \binom{m}{k_i} k_i! 2^{k_i} \mathbb{P}\left(\tilde{C}_1 + \tilde{C}_2 + \cdots + \tilde{C}_k \leq \zeta^*\right) && (24) \\ &\leq (2m)^k \frac{(2\zeta^*)^k}{k!} \sum_{\ell=1}^{k} |\{k_1, \ldots, k_\ell \geq 2 : k_1 + \cdots + k_\ell = k\}| && (25) \\ &\lesssim (2m)^k \frac{(2\zeta^*)^k}{k!} \sum_{\ell=1}^{k} \binom{k}{\ell} \\ &\leq \left(\frac{8\zeta e\omega \log^2 n}{k}\right)^k. && (26) \end{aligned}$$

Where, to go from (24) to (25), 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 5.) $\square$

# 5 Finishing the proof of Theorem 1

Now that we know the solution to the ATSP satisfies Lemma 15, we will give an algorithm to iterate over the possible ATSP solutions. As w.h.p. every potential ATSP solution uses at most $\log^{2+o(1)} n$ non-basic edges, and there are $n^2$ edges in total, we can quickly see that there are at most

$$\binom{n^2}{\log^{2+o(1)} n} \leq (n^2)^{\log^{2+o(1)} n} \leq e^{\log^{3+o(1)} n}$$

potential ATSP solutions to iterate through. The goal of the technical work in this section is to describe an algorithm that works with a run-time reduced from $e^{\log^{3+o(1)} n}$ to $e^{\log^{2+o(1)} n}$ The intuitive reason that we can obtain run-time $e^{\log^{2+o(1)} n}$ is the following lemma, recalling $\zeta^*$ defined at the start of Section 2:

**Lemma 16** *W.h.p., every non-basic edge in the optimal solution to the ATSP that is not in $T^*$ has reduced cost at most $\zeta^*$.*

**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 $\tilde{C}_e \leq \zeta^*$. $\square$

Now, though there are $n^2$ total edges, intuitively each vertex is on average likely incident to approximately $n\zeta^* = \zeta \log^2 n$ edges that have reduced cost at most $\zeta^*$, giving a "branching factor" that is polylogarithmic in $n$ for the number of possible non-basic edges to choose. We will formalize this intuition in the following subsections.

## 5.1 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 16 that the ATSP solution is formed from $M^*$ by taking its symmetric difference 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 15, all sets of cycles $Q_1, \ldots, Q_\ell$ that we need to consider along with $M^*$ to form the ATSP solution have non-basic cardinalities that total to at most $\log^{2+o(1)}(n)$.

**Lemma 17** *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, if $m$ is as in Lemma 14,

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

We choose $k$ alternating paths, 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 $\zeta^*$.

The result then follows from the Markov inequality. □

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

**Lemma 18** *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 17, 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 19** *With high probability, we can find and store all valid cycles and paths with non-basic cardinality at most $k$ in time $O(n^2 \log^{3k+4} n)$.*

**Proof.** We will prove this inductively. Our base case is that for $k = 1$, we can take all of the $\omega n$ (see Lemma 14) 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 18, the total number of such pairs is at most

$$\left(n \log^{3\lfloor k/2 \rfloor + 2} n\right) \left(n \log^{3\lceil k/2 \rceil + 2} n\right) \leq n^2 \log^{3k+4} n.$$

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. □

## 5.2 Iterating through Possible ATSP Solutions

First, we precompute and store all valid cycles with non-basic cardinality at most the value given by Lemma 15, which by Lemma 19 takes time at most

$$O\left(\sum_{k=1}^{\log^{2+o(1)} n} n^2 \log^{3k+4} n\right) = O(n^2 \log^{3\log^{2+o(1)} n+4} n) = O(e^{\log^{2+o(1)} n}).$$

Then, we run through the possibilities for $k$, the number of edges in the ATSP but not in $T^*$. By Lemma 15, $k \leq \log^{2+o(1)} n$ w.h.p.

Next, we run through the possibilities for $\ell$, the number of distinct valid cycles made by adding the $k$ edges into $T^*$, and accounting for the parity of 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 $\ell$ 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 $2^{2k} \leq e^{\log^{2+o(1)} n}$ 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 $1 \leq i \leq \ell$, we specify the $i$th valid cycle by choosing one of the at most $\log^{3k_i+1} n$ 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 $(k_1, \ldots, k_\ell)$ is at most

$$\prod_{i=1}^{\ell} \log^{3k_i+1} n = \log^{3k+\ell} n \leq e^{\log^{2+o(1)} n}$$

as desired.

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 6.

## 5.3 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 20** *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 $\ell_0 = \frac{\log n}{6 \log\log n}$. By Lemma 17, there are at most

$$O\left(\sum_{k=1}^{\ell_0} n \log^{3k+2} n\right) \leq 2n \log^{\log n/(2\log\log n)+2} n \leq 2n^{1.5} \log^2 n$$

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 5.2, we still iterate through all possible $(k_1, \ldots, k_\ell)$. For the $i$th cycle, we specify it in one of two ways, depending on whether $k_i \leq \ell_0$ or whether $k_i > \ell_0$.

**Case 1:** $k_i \leq \ell_0$. Then just as before we iterate through the $O(\log^{3k_i+1} n)$ stored cycles.

**Case 2:** $k_i > \ell_0$. 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 19. In particular, we can specify any cycle with non-basic cardinality $k_i$ by specifying $\lceil \frac{k_i}{\ell_0} \rceil$ valid paths of non-basic cardinalilty at most $\ell_0$ that concatenate to form this cycle. Using this process, as we had at most $2n^{1.5}\log^2 n \leq n^2$ valid paths of non-basic cardinality at most $\ell_0$, we have at most

$$n^{2(1+k_i/\ell_0)} = n^2 e^{12k_i \log\log n}$$

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 $(k_1, \ldots, k_\ell)$ is asymptotically at most

$$\begin{aligned}
&\prod_{i:k_i \leq \ell_0 = 1} \log^{3k_i+1} n \prod_{i:\ell_0 < k_i \leq \log^{2+o(1)} n} n^2 e^{12k_i \log\log n} \\
&\leq (\log^{3k+\ell} n)(n^{\log^{2+o(1)} n/\ell_0}) e^{12k \log\log n} \\
&\leq e^{\log^{2+o(1)} n}
\end{aligned}$$

as desired. □

# 6 Properties of the assignment problem

In this section, we verify Lemmas 2 and 3 under our more general distribution of costs. We assume that $C$ has a density $f$ and satisfies

(i) $f(x) = a + bx + O(x^2)$ for $0 \leq x \leq L$, where $a, b$ are constants and $aL \geq 1$.

(ii) (a) $\mathbb{P}(C \geq x) \leq \alpha e^{-\beta x}$ for constants $\alpha, \beta > 0$, or (b) $f(x) = 0$ for $x > L$.

(iii) To avoid some pathologies, we will also assume that there is a constant $M$ such that $f(x) \leq aM$ for $x \geq 0$.

## 6.1 $M^*$ only has low cost edges

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

$$\max_{e \in M^*} \{C(e)\} \leq \gamma^* = \frac{\gamma \log n}{n} \text{ for some absolute constant } \gamma > 0. \tag{27}$$

Much of this section is a direct adaptation of the proof used by Frieze and Sorkin to show the same lemma in the particular case where the cost distribution is $U[0,1]$ [FS07].

Let $\vec{D}_M$ etc. be as defined prior to (10).

**Lemma 21** *W.h.p. over random cost matrices $C$, for every perfect matching $M$, the (unweighted) diameter of $\vec{D}_M$ 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

$$\max \{C(i,j) : \{a_i, b_j\} \in M\} \leq \frac{\gamma_1 \log^2 n}{n} \text{ for some absolute constant } \gamma_1 > 0. \tag{28}$$

Indeed, for a fixed $i$ we have, where the density of $C$ is $a + bx + O(x^2)$ as $x \to 0$,

$$\mathbb{P}\left(C(i,j) \geq \frac{10 \log n}{an} \text{ for } j \in [n/2]\right) \leq \left(1 - \int_{x=0}^{10 \log n/(an)} (a + bx + O(x^2))dx\right)^{n/2}$$
$$\leq \left(1 - \frac{10 \log n}{n} + O\left(\frac{\log^2 n}{n^2}\right)\right)^{n/2} = o(n^{-9}).$$

It follows that w.h.p. all of the forward edges in the paths alluded to in Lemma 21 have cost at most $\frac{10 \log n}{an}$. If $x \in A$ and $y \in B$ then Lemma 21 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 $\frac{10 k_0 \log n}{an}$. So if there is a matching edge of cost greater than $\frac{10 k_0 \log n}{an}$ then there is an alternating path using at most $k_0$ edges that can be used to give a matching of lower cost, contradiction. This verifies (28).

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 22** *Suppose that $k_1 + k_2 + \cdots + k_P = K \leq \kappa \log N, \kappa = O(1)$, and $Y_1, Y_2, \ldots, Y_P$ 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 sufficiently large, then*

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

*where $\theta = \beta/2$.*

**Proof.** For $0 \leq x \leq L$ we have $\mathbb{P}(C \leq x) = ax + O(x^2) \leq ax(1 + Ax)$ for some $A > 0$. Therefore, the density function $f_k(x)$ of the $k$th order statistic $Y_{(k)}$ satisfies

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

Therefore the moment generating function of $Y_{(k)}$ satisfies

$$\begin{aligned}\mathbf{E}(e^{tY_{(k)}}) &\leq a^k M \binom{N}{k} \int_{x \geq 0} e^{tx}(x + Ax^2))^{k-1} e^{-(N-k)\theta x} dx \\ &\leq a^k M \binom{N}{k} \int_{x \geq 0} x^{k-1} e^{-((N-k)\theta - t - Ak)x} dx \\ &\leq \frac{a^k M N^k}{k((N-k)\theta - t - A)^k}.\end{aligned}$$

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

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

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

So,

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

□

We go along way towards proving (27) by proving

**Lemma 23** *The following holds with probability $1 - o(n^{-2})$. For all $a \in A, b \in B$, $\vec{D}$ contains an alternating path from $a$ to $b$ of total cost less than $\gamma^*$*

**Proof.** Let

$$Z_1 = \max\left\{ \sum_{i=0}^{k} C(x_i, y_i) - \sum_{i=0}^{k-1} C(y_i, x_{i+1}) \right\}, \tag{29}$$

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 (28). 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 $\rho_j$th 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 $\rho_i$th 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 22 that if $\gamma$ is sufficiently large then, for all $y \geq 0$, $q(\rho_1, \ldots, \rho_k; \gamma+y) \leq n^{-(\gamma+y)/2}$ 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

$$\begin{aligned}
\mathbb{P}(Z_1 \geq \gamma^*) &\leq 2n^{2-\gamma/2} \sum_{k=0}^{k_0} \frac{\log^{k-1} n}{(k-1)!} \binom{84k}{k} \int_{y=0}^{\infty} y^{k-1} n^{-y/2} dy \\
&\leq 2n^{2-\gamma/2} \sum_{k=0}^{k_0} \frac{\log^{k-1} n}{(k-1)!} \left(\frac{168e}{\log n}\right)^k \Gamma(k) \\
&\leq 2n^{2-\gamma/2}(168e)^{k_0+1} \\
&= o(n^{-2}).
\end{aligned}$$

□

**Proof of Lemma 2**

Suppose $e = \{a_i, b_j\} \in M^*$ and $C(e) > \gamma^*$. It follows from Lemma 23 that there is an alternating path $P=(a_1,\dots,b_j)$ of cost at most $\gamma^*$. But then deleting $e$ and the $M^*$-edges of $P$ and adding the non-$M^*$ edges of $P$ to $M^*$ creates a matching from $A$ to $B$ of lower cost than $M^*$, contradiction.

# 7 From exponential mean one to more general distributions

## 7.1 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 $C(i,j)=F^{-1}(\widehat{C}(i,j))$, where $F(x)=\mathbb{P}(C\le x)$.
(Note that $\mathbb{P}(C(i,j)\le x)=\mathbb{P}(F^{-1}(\widehat{C}(i,j))\le x)=\mathbb{P}(\widehat{C}(i,j)\le F(x))=F(x)$ and $C(i,j)=(1+O(x))\widehat{C}(i,j)$ for $x=o(1)$.)

Then, using Lemma 2, we have

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

## 7.2 Dealing with Lemma 7

Lemma 7 is only valid for the costs $C$ being exponential with mean 1. However, we will now show that our result still holds on a more general class of distributions. If we replace the costs $C(i,j)$ by $\widehat{C}(i,j)=aC(i,j)$ then our proof will go through almost unchanged. This yields a distribution as in the introduction where $a$ is now 1. We assume then that $a=1$.

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 $p=\mathbb{P}(C\le c)=F(c)$, and set the cost $x$ under $X$ such that $1-e^{-x}=p$ or $x=\log\frac{1}{1-F(c)}$ (in other words, drawing from the same places on the respective cumulative distribution functions).

We know $\mathbb{P}(C\le z)=z+O(z^2)$ and $\mathbb{P}(X\le z)=z+O(z^2)$ as $z\to 0$. We also know from Lemma 16 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 $\zeta^*$ by $\frac{K \log^2 n}{n}$ for sufficiently large $K$) will also w.h.p. find the optimal ATSP under $C$.

# 8 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 $e^{\log^{2+o(1)} n}$.

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.

# 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.

[DFJ54] G. Dantzig, R. Fulkerson, S. Johnson, Solution of a Large-Scale Traveling-Salesman Problem, *Journal of the Operations Research Society of America* 2 (1954) 393–410.

[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.

[FK15] A.M. Frieze and M. Karoński, Introduction to Random Graphs, Cambridge University Press, 2015.

[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.

[J82] A. Joyal, Une théorie combinatoire des séries formelles, *Advances in Mathematics* 42 (1981) 1-82.

[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.

[MTZ60] C. Miller, A. Tucker, and R. Zemlin, Integer Programming Formulation of Traveling Salesman Problems, *Journal of the ACM* 7 (1960), 326–329.

[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.