Probabilistic analysis of the Traveling Salesman Problem to appear as a chapter in a forthcoming book on the TSP

Alan Frieze and Joe Yukich

June 21, 2000

1 Introduction

In this chapter we study the Hamiltonian cycle and Traveling Salesman problem from a probabilistic point of view. Here we try to elucidate the properties of typical rather than worst-case examples. Structurally, one hopes to bring out the surprising properties of typical instances. Algorithmically the hope is that one can in some way explain the successful solution of large problems, much larger than that predicted by worst-case analysis. This of course raises the question of what do we mean by typical? The mathematical view of this is to define a probability space Ω of instances and study the expected properties of ω drawn from Ω with a given probability measure.

Our discussion falls naturally into two parts: the independent case and the Euclidean case. The independent case will include a discussion of the existence of Hamiltonian cycles in various classes of random graphs and digraphs. We will then discuss algorithms for finding Hamiltonian cycles which are both fast and likely to succeed. We include a discussion of extension-rotation constructions and the variance reduction technique of Robinson and Wormald. Following this we consider Traveling Salesman Problems where the coefficients are drawn independently. We describe both exact and approximate algorithms. We include a section on open problems.

After this we survey stochastic results for the total edge length of the Euclidean TSP. Here the cities are assumed to be points in $\mathbb{R}^d$ and the distance between points is the usual Euclidean distance. We describe ways to prove a.s. limit theorems and concentration inequalities for the total edge length of the shortest tour on a random sample of size n, as $n \to \infty$. The focus is on presenting probabilistic techniques which not only describe the behavior of the random Euclidean TSP, but which also describe (or have the potential to describe) the behavior of heuristics. The approach centers around the boundary functional method as well as the isoperimetric methods of Rhee and Talagrand.

1.1 Probabilistic preliminaries

We first collect some basic probabilistic tools which will be needed in the sequel.

Jensen's inequality. A function $f : \mathbb{R} \to \mathbb{R}$ is called *convex* if for all $x,\ y \in \mathbb{R}$ and all $0 \le \lambda \le 1$ we have

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda)f(y).$$

In geometric terms this says that each point on the chord between $(x, f(x))$ and $(y, f(y))$ is above the graph of f. Jensen's inequality says that if X is a random variable with finite mean $\mathbb{E}[X]$ and if f is convex, then

$$\mathbb{E}[f(X)] \ge f(\mathbb{E}[X]). \tag{1.1}$$

Generalized Chebyshev inequality. Let $f : \mathbb{R}^+ \to \mathbb{R}^+$ be an increasing function and let X be a positive random variable. Then for all $t > 0$

$$\mathbb{P}[X \ge t] \le \mathbb{E}[f(X)]/f(t). \tag{1.2}$$

If we replace X by $|X - \mathbb{E}X|$ and let $f(x) = x^2$ we obtain a special case of (1.2), namely

$$\mathbb{P}[|X - \mathbb{E}X| > t] \le \mathrm{Var}X/t^2. \tag{1.3}$$

With high probability. A sequence of events $\mathcal{E}_n,\ n \ge 1$, is said to occur with high probability (whp) if $\lim_{n\to\infty} \mathbb{P}[\mathcal{E}_n] = 1$.

Stochastic convergence. Let $X_1, X_2, ..., X$ be random variables on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We say that

(a) $X_n \to X$ *almost surely*, written $\lim_{n\to\infty} X_n = X$ *a.s.*, if

$$\mathbb{P}[\omega \in \Omega : X_n(\omega) \to X(\omega) \text{ as } n \to \infty] = 1.$$

(b) $X_n \to X$ in L^1 or *in mean* if

$$\mathbb{E}|X_n - X| \to 0 \text{ as } n \to \infty.$$

(c) $X_n \to X$ *completely*, written $\lim_{n\to\infty} X_n = X$ *c.c.*, if for all $\epsilon > 0$,

$$\sum_{n=1}^{\infty} \mathbb{P}[|X_n - X| > \epsilon] < \infty.$$

Binomial random variable. For all $n \in \mathbf{N}$ and $0 < p < 1$, $B(n,p)$ denotes a binomial random variable with parameters n and p. Chernoff's bound for the binomial says that for all $0 \le \epsilon \le 1$

$$\mathbb{P}[|B(n,p) - np| \ge \epsilon np] \le 2\exp^{-\epsilon^2 np/3}. \tag{1.4}$$

See e.g. Alon and Spencer [7].

Cardinality of a set. If A is a set, then $|A|$ denotes its cardinality.

2 Hamiltonian Cycles in Random Graphs

2.1 Models of random graphs

In this section we describe two simple models of a random graph and their relationship. The random graph $G_{n,m}$ is chosen uniformly at random from the set $\mathcal{G}_{n,m}$: the set of graphs with vertex set $[n] = \{1, 2, \ldots, n\}$ and exactly m edges. Since $|\mathcal{G}_{n,m}| = \binom{N}{m}$, $N = \binom{n}{2}$, each graph in $\mathcal{G}_{n,m}$ has probability $1/\binom{N}{m}$ of being selected. This model was intensely studied in a seminal sequence of papers by Erdős and Rényi [29, 30] and has since become a well established branch of Combinatorics, see e.g. the book of Bollobás [11]. The related model $G_{n,p}$ (here $0 \leq p \leq 1$) is defined as follows: It has vertex set $[n]$ and each of the N possible edges is independently included with probability p and excluded with probability $1 - p$. Thus if $G = ([n], E)$,

$$\mathbb{P}[G_{n,p} = G] = p^{|E|}(1-p)^{N-|E|}.$$

Thus for example if $p = \frac{1}{2}$ then all graphs with vertex set $[n]$ are equally likely.

The next thing to observe is that *conditional on having* m *edges*, $G_{n,p}$ is distributed as $G_{n,m}$. Thus if $\mathcal{P}_n$ is any graph property for the set of graphs with vertex set $[n]$ then

$$\mathbb{P}[G_{n,p} \in \mathcal{P}_n] = \sum_{m=0}^{N} \binom{N}{m} p^m (1-p)^{N-m} \mathbb{P}[G_{n,m} \in \mathcal{P}_n]. \tag{2.1}$$

The number of edges of $G_{n,p}$ is the binomial random variable $B(N, p)$ and if its mean Np is large, $B(N, p)$ is concentrated around it. More precisely, Chernoff's bound on the binomial (1.4) implies that if $Np/\log n \to \infty$ then

$$\mathbb{P}[|B(N,p) - Np| \geq \sqrt{KNp \log n}] \leq 2n^{-K/3} \tag{2.2}$$

which tends to zero as $n \to \infty$. Thus, plausibly, with (2.1) and (2.2), one can argue simultaneously about properties of $G_{n,p}$ and $G_{n,m}$. This is not true all the time, but it is in many cases. Much of the interest in the theory of random graphs concerns properties that occur *with high probability* (whp). Erdős and Rényi proved many beautiful results about random graphs from this viewpoint. Suppose $p = p(n)$ depends on n:

- If $np - \log n \to \infty$ then $G_{n,p}$ is connected whp.
- If $np = c > 1$, c constant, then $G_{n,p}$ has a unique (giant) component of size order n, whp.

They also observed that for many properties $\mathcal{P}_n$ there is a threshold probability $p_0 = p_0(n)$ such that if $p/p_0 \to 0$ then $G_{n,p} \notin \mathcal{P}_n$ whp and if $p/p_0 \to \infty$ then $G_{n,p} \in \mathcal{P}_n$ whp. For example if $\mathcal{A}_n = \{G \text{ contains a copy of } K_4\}$ then

$$\begin{aligned} \mathbb{P}[G_{n,p} \in \mathcal{A}_n] &= o(1) && \text{if } n^{2/3}p \to 0. \\ \mathbb{P}[G_{n,p} \in \mathcal{A}_n] &= 1 - o(1) && \text{if } n^{2/3}p \to \infty. \end{aligned}$$

Thus one line of research in random graphs is the determination of thresholds for graph properties.

2.2 Existence Results

Erdős and Rényi did not manage to establish the threshold for the existence of a Hamiltonian cycle in a random graph. It took about twenty years to solve this problem.

2.2.1 The threshold for Hamiltonicity

A breakthrough came when Posa [65] proved the following theorem.

Theorem 2.1 *If $K > 16$ is constant and $p \geq \frac{K \log n}{n}$ then $G_{n,p}$ is Hamiltonian whp.*

Proof We give the details of the proof because (i) it is very elegant, (ii) it is not technically difficult, and (iii) and it illustrates well the most basic methods of proof in this area.

We first show that $G_{n,p}$ has a Hamiltonian path whp. For a graph $G = (V, E)$ we let $\lambda(G)$ denote the length of the longest path in G. Observe that

$$G \text{ has a Hamiltonian path iff } \lambda(G) > \lambda(G - v) \text{ for all } v \in V.$$

For all $1 \leq i \leq n$, let $\mathcal{E}_i$ be the event $\{\lambda(G_{n,p}) = \lambda(G_{n,p} - i)\}$. Then

$$\mathbb{P}[G_{n,p} \text{ has no Hamiltonian path}] \leq \sum_{i=1}^{n} \mathbb{P}[\mathcal{E}_i] = n\mathbb{P}[\mathcal{E}_n]. \tag{2.3}$$

We will show that

$$\mathbb{P}[\mathcal{E}_n] \leq n^{3-K/4} + n^{-K/4}. \tag{2.4}$$

Let $H = G_{n,p} - n$ and notice that H is distributed as $G_{n-1,p}$. Let $P = (x_0, x_1, \ldots, x_k)$ be any longest path of H and let

$$X = \{x : \ H \text{ contains a path of length } k \text{ from } x_0 \text{ to } x\}.$$

We will show that

$$\mathbb{P}[|X| \leq \tfrac{n}{4}] \leq n^{3-K/4}. \tag{2.5}$$

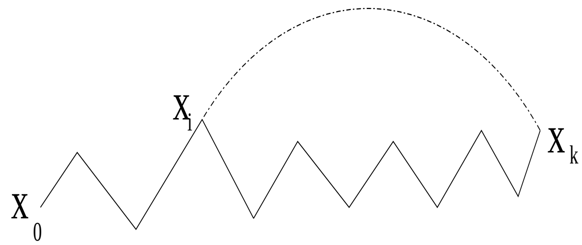
Assuming (2.5) and noting that $\mathcal{E}_n$ implies there is no edge joining X to n in $G_{n,p}$, it follows that

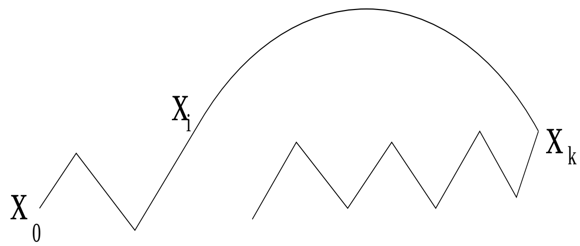
$$\begin{aligned} \mathbb{P}[\mathcal{E}_n] &\leq \mathbb{P}[|X| \leq \tfrac{n}{4}] + \mathbb{P}[\mathcal{E}_n \mid |X| > \tfrac{n}{4}] \\ &\leq n^{3-K/4} + (1-p)^{n/4} \\ &\leq n^{3-K/4} + n^{-K/4}, \end{aligned}$$

which is (2.4). We now have to prove (2.5). For each $i \in \{0, 1, \ldots, k-2\}$ such that (x_k, x_i) is an edge of H we can define a new path (see Figure 2.2.1)

$$P' = \text{ROTATE}(P; x_k, x_i) = (x_0, x_1, \ldots, x_i, x_k, x_{k-1}, \ldots, x_{i+1}).$$

We say that P' is obtained from P by a *rotation* with x_0 as the fixed endpoint. Note that P' is also of length k and so $x_{i+1} \in X$. Let $X' \subseteq X$ be the set of endpoints obtainable by doing a sequence of rotations, starting at P, and always using x_0 as the fixed endpoint.





For $S \subseteq V(H)$ let

$$N_H(S) = \{w \in V(H) \setminus S : \ \exists v \in S \text{such that} (v, w) \in E(H)\}.$$

The following lemma is perhaps the key idea behind Posá's result.

Lemma 2.2

$$|N_H(X')| < 2|X'|.$$

Proof We show that if $y \in N_H(X') \setminus \{x_0\}$ then there exists $z \in X'$ such that (y, z) is an edge of P. The lemma follows immediately.

Since $y \in N_H(X')$ there exists $x \in X'$ such that $(x, y) \in E(H)$. So there exists a path Q of length k from x_0 to x with y as an internal vertex, obtainable from P by a sequence of rotations. Let z_1 be the neighbor of y on Q between y and x. Then $z_1 \in X'$ because it is the endpoint of $\textsc{rotate}(Q; x, y)$. So if (y, z_1) is an edge of P we are done. If (y, z) is not an edge of P then one of the edges $e = (y, z_2)$ incident with y in P has been deleted in the sequence of rotations which produced Q. But when an edge e is deleted by a rotation, one of its endpoints is placed in X'. Since $y \notin X'$ we have $z_2 \in X'$ and the lemma follows. $\square$

Now for a calculation:

$$\begin{aligned}
\mathbb{P}[\exists S, 1 \leq |S| \leq \tfrac{n}{4}, |N_H(S)| < 2|S|] &\leq \sum_{s=1}^{n/4} \binom{n-1}{s}\binom{n-1}{2s}(1-p)^{s(n-3s)} \\
&\leq \sum_{s=1}^{n/4} n^s n^{2s} n^{-Ks/4} \\
&\leq 2n^{3-K/4}.
\end{aligned}$$

So (2.5) follows from this and Lemma 2.2. Applying (2.3) we see that $G_{n,p}$ has a Hamiltonian path whp.

To finish the proof consider the graph $\Gamma = G_{n,p_1=5p/6} \cup G_{n,p_2=p/6}$. Each of the N possible edges appears independently with probability $p_1 + p_2 - p_1 p_2 < p$ and so

$$\mathbb{P}[\Gamma \text{ is Hamiltonian}] \leq \mathbb{P}[G_{n,p} \text{ is Hamiltonian}] \tag{2.6}$$

Now by the previous analysis, whp G_{n,p_1} has a Hamiltonian path, P say. Fix one of its endpoints x_0 and let X be the set of endpoints of Hamiltonian paths with one endpoint x_0. By the above analysis, $|X| \geq \frac{n}{4}$ whp. Now given X, the set of edges of G_{n,p_2} which join x_0 and X is independent of G_{n,p_1} and if there is one, Γ is Hamiltonian. So

$$\mathbb{P}[\Gamma \text{ is not Hamiltonian}] \leq$$
$$\mathbb{P}[G_{n,p_1} \text{ has no Hamiltonian path}] + \mathbb{P}[|X| \leq \tfrac{n}{4}] + (1-p_2)^{n/4} = o(1).$$

This together with (2.6) proves Theorem 2.1. □

Now a graph with a vertex of degree 0 or 1 is not Hamiltonian. Let $m = \frac{n}{2}(\log n + \log\log n + c_n n)$. Erdős and Rényi [31]) showed that

$$\lim_{n\to\infty} \mathbb{P}[\delta(G_{n,m}) \geq 2] = \begin{cases} 0 & c_n \to -\infty \\ e^{-e^{-c}} & c_n \to c \\ 1 & c_n \to +\infty \end{cases} \tag{2.7}$$

(There is no mystery to the right hand side of the above when $c_n = c$. The expected number of vertices of degree 0 or 1 is $\approx e^{-c}$ and the distribution of this number is asymptotically Poisson.)

The above provides a lower bound for the asymptotic probability of a graph being Hamiltonian. Komlós and Szemerédi [61] proved that this was tight, essentially saying that the best constant in Theorem 2.1 is $K = 1$.

Theorem 2.3

$$\lim_{n\to\infty} \mathbb{P}[G_{n,m} \textit{ is Hamiltonian}] = \lim_{n\to\infty} \mathbb{P}[\delta(G_{n,m}) \geq 2].$$

2.2.2 Graph Process

There is an alternative stronger version of Theorem 2.3 which is quite remarkable at first sight. The graph process is a random sequence $G_0, G_1, \ldots, G_i = ([n], E_i), \ldots, G_N$ where $E_i = \{e_1, e_2, \ldots, e_i\}$ and e_i is chosen randomly from $\bar{E}_{i-1}$. Thus starting from the empty graph G_0, we randomly add new edges until we have a complete graph. Note the G_m has the same distribution as $G_{n,m}$.

For a graph property $\mathcal{G}$, the *hitting time* $\tau(\mathcal{G})$ is given by

$$\tau(\mathcal{G}) = \min\{i : \ G_i \in \mathcal{G}\}.$$

Let $\mathcal{H} = \{G \text{ is Hamiltonian}\}$ and $\mathcal{D}_k = \{\delta(G) \geq k\}$. Clearly $\tau(\mathcal{H}) \geq \tau(\mathcal{D}_2)$. Bollobás [12] and Ajtai, Komlós and Szemerédi [1] showed

Theorem 2.4

$$\tau(\mathcal{H}) = \tau(\mathcal{D}_2) \qquad\qquad whp.$$

In other words, if we randomly add edges one by one, whp the first edge that raises the minimum degree to 2, also makes the graph Hamiltonian!

Other properties

The results of Theorems 2.3 and 2.4 have been generalized in a number of ways: All statements are claimed to hold whp.

- $G_{\tau(\mathcal{D}_k)}$ has $\lfloor k/2 \rfloor$ edge disjoint cycles plus a further edge disjoint (near) perfect matching if k is odd, $k = O(1)$ [18].
- $G_{\tau(\mathcal{D}_2)}$ has $(\log n)^{n-o(n)}$ distinct Hamiltonian cycles [21].
- $G_{\tau(\mathcal{D}_2)}$ contains k vertex disjoint cycles of size n/k, $k = O(1)$ [35].
- $G_{\tau(\mathcal{D}_2)}$ contains a cycle of every size, $3 \leq k \leq n$, [22, 62, 20]

See also [23] and the surveys [38], [44].

2.2.3 Regular Graphs

In this section we discuss the existence of Hamiltonian cycles in random regular graphs. We use $G_{n,r}$ to denote a graph chosen uniformly at random from the set of r-regular graphs with vertex set $[n]$.

It is easy to show that whp $G_{n,2}$ is a collection of $O(\log n)$ vertex disjoint cycles. Thus $G_{n,3}$ or random cubic graphs is where the real interest starts. There was some success in applying the extension-rotation methods of Section 2.2, [14],[34] and [39]. In the last mentioned paper it was shown that $G_{n,r}$ is Hamiltonian whp for all *constant* $r \geq 85$.

A breakthrough came with the work of Robinson and Wormald [77, 78] who used a completely different approach to solve the problem. They proved that $G_{n,r}$ is Hamiltonian whp for all *constant* $r \geq 3$.

2.2.4 Model of random regular graphs

We describe the *configuration model* of Bollobás [13]. This is a probabilistic interpretation of a counting formula of Bender and Canfield [10].

Thus let $W = [n] \times [r]$ ($W_v = v \times [r]$ represents r *half edges* incident with vertex $v \in [n]$.) The elements of W are called *points* and a 2-element subset of W is called a *pairing*. A *configuration* F is a partition of W into $rn/2$ pairings. We associate with F a multigraph $\mu(F) = ([n], E(F))$ where, as a multi-set,

$$E(F) = \{(v, w) : \{(v, i), (w, j)\} \in F \text{ for some } 1 \leq i, j \leq r\}.$$

(Note that $v = w$ is possible here.)

Let Ω denote the set of possible configurations. Thus

$$|\Omega| = \Lambda(rn)$$

where

$$\Lambda(2m) = \frac{(2m)!}{m!2^m}.$$

We say that F is *simple* if the multigraph $\mu(F)$ has no loops or multiple edges. Let Ω_0 denote the set of simple configurations.

We turn Ω into a probability space by giving each element the same probability. The main properties that we need of this model are:

P1 Each $G \in \mathcal{G}(n,r)$ is the image (under μ) of exactly $(r!)^n$ simple configurations.

P2 $\mathbb{P}[F \in \Omega_0] \approx e^{-(r^2-1)/4}$.

(Here $\alpha \approx \beta$ means that $\alpha/\beta \to 1$ as $n \to \infty$.)

Suppose now that $\mathcal{A}^*$ is a property of configurations and $\mathcal{A}$ is a property of graphs such that when $F \in \Omega_0, \mu(F) \in \mathcal{A}$ implies $F \in \mathcal{A}^*$. Then P1 and P2 imply

$$\mathbb{P}[G \in \mathcal{A}] \leq (1+o(1))e^{(r^2-1)/4}\mathbb{P}[F \in \mathcal{A}^*]$$

where G is chosen randomly from $\mathcal{G}$ and F is chosen randomly from Ω. So if r is a constant independent of n then we can use this to see that

$$\mathbb{P}[F \in \mathcal{A}^*] = o(1) \text{ implies } \mathbb{P}[G \in \mathcal{A}] = o(1).$$

2.2.5 The Robinson-Wormald approach

Suppose we wanted to use Chebyshev's inequality (1.3) to prove that the existence of a Hamiltonian cycle! For $F \in \Omega$ let

$$Z_H = Z_H(F) = \text{the number of Hamiltonian cycles in } H.$$

A reasonably straightforward calculation shows that

$$\mathbb{E}[Z_H] \approx \sqrt{\frac{\pi}{2n}}\left(\left(\frac{r-2}{r}\right)^{(r-2)/2}(r-1)\right)^n.$$

A much more difficult calculation [42] gives

$$\mathbb{E}[Z_H^2] \approx \frac{\pi r}{2(r-2)n}\left(\left(\frac{r-2}{r}\right)^{r-2}(r-1)^2\right)^n.$$

So we immediately get

$$\mathbb{P}[Z_H \neq 0] \geq \frac{\mathbb{E}(Z_H)^2}{\mathbb{E}(Z_H^2)} \approx \frac{r-2}{r}.$$

So when $r = 3$ we see already that $\mathbb{P}[Z_H \neq 0] \geq 1/3$. We need to boost this to 1-o(1).

Let C_l denote the number of ℓ-cycles of $\mu(F)$ for $\ell \geq 1$. We will be concerned mainly with C_l where l is odd. For $\mathbf{c} = (c_1, c_2, \ldots, c_b) \in N^b$, where $N = \{0, 1, 2, \ldots\}$, let group $\Omega_{\mathbf{c}} = \{F \in \Omega : C_{2k-1} = c_k, 1 \leq k \leq b\}$. Surprisingly,there is much less variation in the number of Hamiltonian cycles *within* groups, than there is *between* groups. In fact, almost all of the variance can be "explained" by variation between group means.

For $\mathbf{c} \in N^b$ let

$$\pi_{\mathbf{c}} = \mathbb{P}[F \in \Omega_{\mathbf{c}}],\ \mathbb{E}_{\mathbf{c}} = \mathbb{E}[Z_H \mid F \in \Omega_{\mathbf{c}}] \text{ and } V_{\mathbf{c}} = \mathrm{Var}[Z_H \mid F \in \Omega_{\mathbf{c}}].$$

Then by conditioning on $\mathbf{c}$ we have

$$\mathbb{E}[Z_H^2] = \sum_{\mathbf{c} \in N^b} \pi_{\mathbf{c}} \mathbb{E}[Z_H^2 \mid F \in \Omega_{\mathbf{c}}] = \sum_{\mathbf{c} \in N^b} \pi_{\mathbf{c}} V_{\mathbf{c}} + \sum_{\mathbf{c} \in N^b} \pi_{\mathbf{c}} \mathbb{E}_{\mathbf{c}}^2.$$

Robinson and Wormald [77, 78] prove that

$$\sum_{\mathbf{c} \in N^b} \pi_{\mathbf{c}} V_{\mathbf{c}} \leq \delta\, \mathbb{E}[Z_H]^2, \tag{2.8}$$

for some small value δ.

The rest is an application of Chebyshev's inequality (1.3). Define the random variable $\hat{Z}_H$ by

$$\hat{Z}_H = \mathbb{E}_{\mathbf{c}},\ \text{if } F \in \Omega_{\mathbf{c}}.$$

Then for any $t > 0$

$$\begin{aligned} \mathbb{P}[|Z_H - \hat{Z}_H| \geq t] &\leq \mathbb{E}[(Z_H - \hat{Z}_H)^2/t^2] \\ &= \sum_{\mathbf{c} \in N^b} \pi_{\mathbf{c}} V_{\mathbf{c}}/t^2 \\ &\leq \delta \mathbb{E}[Z_H]^2/t^2 \end{aligned} \tag{2.9}$$

where the last inequality follows from (2.8).

Robinson and Wormald argue that for likely values of $\mathbf{c}$ we find that $\mathbb{E}_{\mathbf{c}} > \epsilon \mathbb{E}(Z_H)$ where δ/ϵ^2 can be chosen arbitrarily small, though fixed as $n \to \infty$. Putting $t = \epsilon \mathbb{E}(Z_H)$ in (2.9) gives the required result

$$\mathbb{P}[Z_H = 0] \leq \delta/\epsilon^2 + \theta$$

where θ is the small probability that $G_{n,r} \in \mathbb{E}_{\mathbf{c}}$ where $\mathbb{E}_{\mathbf{c}} \leq \epsilon \mathbb{E}(Z_H)$.

2.3 Polynomial time algorithms

Here we turn our attention to algorithms for finding Hamiltonian cycles which work whp.

2.3.1 Sparse case

Angluin and Valiant [8] described a randomized algorithm which whp finds a Hamiltonian cycle in $G_{n,p}$ provided $p \geq K \log n/n$ for sufficiently large K. Shamir [80] improved this to $p \geq (\log n + (2+\epsilon)\log\log n)/n$ which is almost best possible.

We discuss the algorithm HAM of Bollobás, Fenner and Frieze [17] which yields a constructive proof of Theorem 2.3. We work in $G_{n,m}$ where $m = \frac{n}{2}(\log n + \log\log n + c_n)$ and either $c_n \to c$ or $c_n \to \infty$. Let $d = 2m/n$ be the average degree of G. We note that G is connected whp and we check for for minimum degree at least 2 before running HAM.

Algorithm HAM works in stages. At the start of stage k we have a path P_k of length k. Suppose that its endpoints are w_0, w_1. We start the algorithm in Stage 0 with $P_0 = \{1\}$. We first grow a tree of paths of length k, all with endpoints w_0. (There are places in this procedure where we can jump to stage $k+1$. When k is small we are likely to do this immediately, see below.) The tree is grown in breadth first fashion to a depth T, where $T = \lceil \log n/(\log d - \log\log d)\rceil + 1$. The children of a node P are all those paths which can be obtained from P by a rotation with w_0 as the fixed endpoint.

Let $END(w_0, G)$ denote the set of endpoints, other than w_0, of the paths produced in this manner. Then for each $x \in END(w_0, G)$ we start with the first path Q_x produced with endpoints w_0, x and grow a tree of depth T, with root Q_x and this time use rotations with x as fixed endpoint. Let $END(x, G)$ denote the set of endpoints, other than x, of the paths produced in this manner.

We may not need to do this much work in a stage. If ever we find a path Q with endpoints x, y such that $(y, z) \in E$ and $z \notin Q$ then clearly we have a path Q, z of length $k+1$ and we go immediately to the next stage by a *simple extension.* If $(x, y) \in E$ then because G is connected, there is an edge (u, v) joining the cycle $Q + (x, y)$ to the rest of G. We then have a path $Q + (x, y) + (u, v) - (z, u)$ of length $k+1$ where z is adjacent to u on Q. This is called a *cycle extension* and we call (x, y) the *closing edge.* If neither of these two possibilities occurs during stage k then HAM fails.

HAM stops successfully in stage $n-1$ if it manages to close one of the Hamiltonian paths that it creates.

In the event that HAM fails, if $END = END(G) = \{w_0\} \cup END(w_0, G)$ then we know that

$$x \in END,\ y \in END(x, G) \text{ implies } (x, y) \notin G. \tag{2.10}$$

We will show that $|END|$ is of order n whp and so if there were no conditioning, this would be unlikely to happen.

Suppose now that HAM fails in stage k. Let $W = W(G)$ denote the edges of the paths $P^{(0)}, P^{(1)}, \ldots, P^{(M)} = P_k$ where $P^{(i+1)}$ is obtained from $P^{(i)}$ by a rotation, simple or cycle extension plus the closing edges of the cycle extensions in the sequence. Clearly

$$|W| \leq nT. \tag{2.11}$$

For $X \subset E$ let $G_X = ([n], E \setminus X)$. The following should be clear.

Lemma 2.5 *Suppose that HAM terminates unsuccessfully in stage k on input G. If $X \subseteq E \setminus W$ then HAM will terminate unsuccessfully on input G_X. Furthermore, on G_X, HAM will generate P_k at the start of stage k through the same sequence of rotations and extensions.*

A vertex of G is *small* if its degree is at most $d/20$ and *large* otherwise. The following lemma is proved in [17].

Lemma 2.6 *Assume $c_n \not\to \infty$. Then*

(a) *G contains no more than $n^{1/2}$ small vertices.*

(b) *G does not contain 2 small vertices at a distance of 4 or less apart.*

(c) *G contains no vertex of degree $5d$ or more.*

(d) *There does not exist a set of large vertices S, $|S| \leq n/d$ and $|N_G(S)| \leq d|S|/300$.*

Now let $\mathcal{G}_0 = \mathcal{G}_{n,m}$. Let $\mathcal{G}_1 = \{G :\ G$ is connected, has minimum degree at least 2 and satisfies the conditions of Lemma 2.6$\}$. Then, by (2.7),

$$|\mathcal{G}_1| \approx e^{-e^{-c_n}} \binom{N}{m}. \tag{2.12}$$

Note that the running time of HAM on a member of $\mathcal{G}_1$ is $O(n^2(5d)^T \log n) = O(n^{3+o(1)})$ assuming [8] that we can do a rotation in $O(\log n)$ time. We can make HAM always run in polynomial time by not trying on graphs with maximum degree $\geq 5d$. Note also that HAM is a *deterministic* algorithm.

Suppose that HAM terminates unsuccessfully in stage k on input G. Let $X \subseteq E$ be *deletable* if the following three properties hold:

1. No edge of X is incident with a small vertex,
2. No large vertex meets more than $d/1000$ edges of X, and
3. $X \cap W = \emptyset$.

The following lemma is proved in [17].

Lemma 2.7 *Suppose that HAM terminates unsuccessfully in stage k on input $G \in \mathcal{G}_1$. Suppose $X \subseteq E$ is deletable. Then for large n,*

$$|END(G_X)| \geq n/1000 \tag{2.13}$$
$$|END(x, G_X)| \geq n/1000 \quad \textit{for } x \in END(G_X). \tag{2.14}$$

(Idea of proof: Let S_t denote the set of large endpoints of paths at depth t in a tree. The conditions in Lemma 2.6 imply that $|S_t| \leq n/d$ implies $|S_{t+1}| \geq d|S_t|/1000$. This is basically because $|S_{t+1}| \geq (\frac{1}{2} - o(1))|N(S_t)|$. The o(1) accounts for small neighbors and the $\frac{1}{2}$ comes from two neighbors giving the same new endpoint.)

Now let $\mathcal{G}_2 = \{G : \ G \in \mathcal{G}_1$ and HAM terminates unsuccessfully on $G\}$. We use a "coloring" argument of Fenner and Frieze [34] to prove

$$\frac{|\mathcal{G}_2|}{|\mathcal{G}_0|} \to 0 \text{ as } n \to \infty. \tag{2.15}$$

Combining (2.12), our assumption on c_n, and (2.15) we obtain:

Theorem 2.8

$$\lim_{n\to\infty} \mathbb{P}[\textit{HAM finds a Hamiltonian cycle}] = \lim_{n\to\infty} \mathbb{P}[G_{n,m} \textit{ is Hamiltonian}].$$

Proof of (2.15): Let $\omega = \lceil \lambda d \rceil$ for some (arbitrary) positive constant λ. For $G \in \mathcal{G}_1$ let $(G,j), j = 1,2,\ldots,\nu = \binom{m}{\omega}$ enumerate all the possible ways of choosing ω edges of G and coloring them green and the remaining $m-\omega$ edges blue. Let $X = X(G,j)$ denote the set of green edges. Let

$$a(G,j) = \begin{cases} 1 & \begin{array}{ll} (i) & \text{HAM terminates unsuccessfully on } G \text{ and } G_X. \\ (ii) & \text{There does not exist } e=(x,y) \in X \text{such that} \\ & x \in END(G_X) \text{ and } y \in END(x, G_X) \\ (iii) & |END(G_X)| \geq n/1000 \text{ and } |END(x,G_X)| \geq n/1000 \\ & \text{for all } x \in END(G_X). \end{array} \\ 0 & \text{otherwise} \end{cases}$$

We show next that for $G \in \mathcal{G}_2$

$$\sum_{j=1}^{\nu} a(G,j) \geq (1-o(1))\binom{m-nT}{\omega}. \tag{2.16}$$

To see this let $G \in \mathcal{G}_2$ and let HAM terminate unsuccessfully in stage k on G. It follows from (2.10) and Lemmas 2.5 and 2.7 that if $X = X(G,j)$ is deletable then $a(G,j) = 1$. Let $G' = (V', E')$ be the subgraph of G induced by the large vertices and the edges not in $W(G)$. Then $|V'| \geq n - n^{1/2}$ and $|E'| \geq m - nT$. The number of deletable sets is the number of ways of choosing ω edges from E' subject to the condition that no vertex in V' is incident with more than $d/1000$ edges. Almost all choices of ω edges satisfy this and (2.16) follows.

On the other hand let H be a fixed graph with vertex set $[n]$ and $m-\omega$ edges. Let $b(H) = |\{(G,j) : \ H = G_X,\ G \in \mathcal{G}_1 \text{ and } a(G,j) = 1\}|$. Then

$$b(H) \leq \binom{N'-m+\omega}{\omega} \text{ where } N' = N - \binom{\lceil n/1000 \rceil}{2}. \tag{2.17}$$

Then, by (2.16),

$$
\begin{aligned}
(1-o(1))\binom{m-nT}{\omega}|\mathcal{G}_2| &\leq \sum_{G\in\mathcal{G}_2}\sum_{j=1}^{\nu} a(G,j) \\
&\leq \sum_{G\in\mathcal{G}_0}\sum_{j=1}^{\nu} a(G,j) = \sum_{H} b(H) \\
&\leq \binom{N'-m+\omega}{\omega}\binom{N}{m-\omega}
\end{aligned}
$$

on using (2.17).

It follows that

$$
\frac{|\mathcal{G}_2|}{|\mathcal{G}_0|} \leq (1+o(1))\frac{\binom{N'-m+\omega}{\omega}\binom{N}{m-\omega}}{\binom{m_1}{\omega}\binom{N}{\omega}} \leq e^{-\lambda d/1000001},
$$

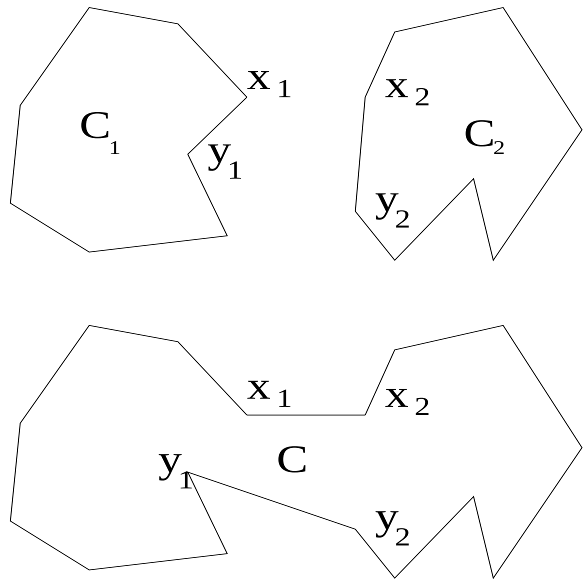
which proves (2.15).

2.3.2 Dense case

Algorithm HAM was designed to work at the threshold for the existence of Hamiltonian cycles. At the other end of the scale we have the dense case $G_{n,p}$, p constant, where our graph has $\Omega(n^2)$ edges whp. For the case $p = 1/2$ (i.e. each graph equally likely), one can show that HAM fails with probability $o(2^{-n})$ and if in the rare occasions where it fails to find a Hamiltonian cycle we use the $O(n^2 2^n)$ dynamic programming algorithm of Held and Karp [50] then we have a deterministic algorithm which (i) runs in polynomial expected time and (ii) *always* determines whether or not a graph has a Hamiltonian cycle.

Gurevich and Shelah [49] and Thomason [91] constructed randomized algorithms which run in (i.e. $O(n^2)$) expected time which is *linear* in the number of edges whp.

It is in fact quite easy to construct an algorithm which finds a Hamiltonian cycle in $G_{n,1/2}$ whp. It is more difficult to make the failure probability small enough to make it run in polynomial expected time. Here is a very simple algorithm for the former case. It is based on the idea of *patching* cycles. Given two cycles C_1, C_2 we look for edges $e_i = (x_i, y_i)$, $i = 1, 2$ such that G contains the edges $f_1 = (x_1, x_2)$ and $f_2 = (y_1, y_2)$. We then create a new cycle $C = C_1 + C_2 + f_1 + f_2 - e_1 - e_2$ which covers the vertices of C_1 and C_2 – see Figure 2.3.2.



Patching Algorithm

(a) Divide $[n]$ into $s = \lfloor n/\omega \rfloor$ sets $X_1, X_2, \ldots, X_s$ of size ω or $\omega + 1$ where $\omega = \lceil 2 \log_2 n \rceil$.

(b) Use the algorithm of [50] to find a Hamiltonian cycle C_i through the vertices of each X_i, $i = 1, 2, \ldots, s$.
Step (b) takes $O(n\omega 2^{\omega}) = O(n^3 \log n)$ time and it fails with probability $O(2^{-\omega} n)$ $= o(1)$.

(c) For $i = 1, 2, \ldots, s$ divide each cycle into 2 paths $P_{i,1}, P_{i,2}$ of lengths $\approx \omega/2$ and then for $i = 1, 2, \ldots, s-1$ try to patch C_i into C_{i+1} using edges from $P_{i,2}$ and $P_{i+1,1}$ which do not contain endpoints of the paths (to avoid dependencies).
The probability that Step (c) fails is $O(s 2^{-\Omega(\omega^2)}) = o(1)$.

The method is easily parallelizable and a more complicated parallel version runs in $O(\log \log n)^2$ time [40]. Mackenzie and Stout [63] and Van Wieren and Stout [92] found $O(\log^* n)$ parallel expected time algorithms. Parallel algorithms for random graphs with $Kn \log n$ edges were considered in Coppersmith, Raghavan, and Tompa [26].

2.3.3 Regular graphs

In this section we consider the problem of finding a Hamiltonian cycle in a random regular graph. The paper [39] provides an $O(n^{3+o(1)})$ time extension-rotation algorithm for finding a Hamiltonian cycle in $G_{n,r}$ whp provided $r \geq 85$, r constant. Frieze, Jerrum, Molloy, Robinson and Wormald [42] used a completely different approach for $r \geq 3$, r constant.

The idea is very simple: Use the algorithm of Jerrum and Sinclair [51] to generate a (near) random 2-factor of $G_{n,r}$. The paper [39] then argues that whp the number of Hamiltonian cycles in $G_{n,r}$ is at least a proportion $1/(2n^{5/2})$ of the number of 2-factors. Thus after $O(n^{5/2}\log n)$ applications of the Jerrum-Sinclair algorithm we will whp produce a Hamiltonian cycle.

A similar idea was used in Frieze and Suen [47] for finding Hamiltonian cycles in a random digraph with m edges, $m/n^{3/2} \to \infty$.

2.3.4 Digraphs

Analogously to $G_{n,m}$ the random digraph $D_{n,m}$ has vertex set $[n]$ and m random edges. The model $D_{n,p}$ is defined similarly. A natural conjecture would be that $D_{n,m}$ is Hamiltonian whp when it has enough edges to ensure that both the minimum in-degree and out-degree are both at least one whp. This was proved in Frieze [37]. Let $m = n(\log n + c_n)$. Then

Theorem 2.9

$$\lim_{n\to\infty} \mathbb{P}[D_{n,m} \text{ is Hamiltonian}] = \begin{cases} 0 & c_n \to -\infty \\ e^{-2e^{-c}} & c_n \to c \\ 1 & c_n \to +\infty \end{cases}$$

This was proved algorithmically. The cycle is found whp by an $O(n^{3/2})$ time algorithm. In Section 3.2 we will discuss a result, in some detail, which implies Theorem 2.9.

Earlier, Angluin and Valiant [8] had given an $O(n(\log n)^2)$ time algorithm which works whp for $m \geq Kn\log n$, K sufficiently large. McDiarmid [64] proved the interesting inequality

$$\mathbb{P}[D_{n,p} \text{ is Hamiltonian}] \geq \mathbb{P}[G_{n,p} \text{ is Hamiltonian}]. \tag{2.18}$$

So putting $p = (\log n + \log\log n + \omega)/n$, $\omega \to \infty$ (2.18) and Theorem 2.3 implies that $D_{n,p}$ is Hamiltonian whp. This is not quite as strong as Theorem 2.9.

2.4 Other models

We briefly consider some results pertaining to the existence of Hamiltonians in other models of a random graph.

Random regular digraphs. Let $D_{n,r}$ be chosen uniformly from the set of digraphs with vertex set $[n]$ in which each vertex has in-degree and out-degree r.

Theorem 2.10 *[25] Assume r is constant independent of n.*

$$\lim_{n\to\infty} \mathbb{P}[D_{n,r} \text{ is Hamiltonian}] = \begin{cases} 0 & r \leq 2 \\ 1 & r \geq 3 \end{cases}$$

k-out model. In this model G_{k-out} the vertex set is $[n]$ and then each vertex $v \in [n]$ independently chooses k neighbors. This is a graph, not a digraph, the average degree is $2k$ and multiple edges are possible. What is known is summarized in the following. The case $k = 3$ is an important open question.

Theorem 2.11 *[24]*

$$\lim_{n\to\infty} \mathbb{P}[G_{k-out} \text{ is Hamiltonian}] = \begin{cases} 0 & k \leq 2 \\ 1 & k \geq 4 \end{cases}$$

k-in, k-out model. In this model $D_{k-in,k-out}$ the vertex set is $[n]$ and then each vertex $v \in [n]$ independently chooses k in-neighbors and k out-neighbors.

Theorem 2.12 *[24]*

$$\lim_{n\to\infty} \mathbb{P}[D_{k-in,k-out} \text{ is Hamiltonian}] = \begin{cases} 0 & k = 1 \\ 1 & k \geq 2 \end{cases}$$

Hidden Hamiltonian cycles. In this model we start with a cycle of size n and add either (i) cn random edges to it, c constant or (ii) a random perfect matching. This has some cryptographic significance in relation to authentication schemes [19]. [19] shows that in case (i), if c is sufficiently large then one can find a Hamiltonian cycle whp in polynomial time. This may not be the original cycle, but it nevertheless kills the authentication scheme. Similarly, [42] shows that whp we can find a Hamiltonian cycle in case (ii). See Wormald [94] for a recent survey which explains the relation between graphs generated as in (ii) and random regular graphs.

This ends our discussion of the Hamiltonian cycle problem in random graphs, except for open problems, see Section 3.5. We turn to the Traveling Salesman Problem (TSP).

3 Traveling Salesman Problem: independent model

We consider problems where the coefficients of the cost matrix $C = [c_{i,j}]$ are either completely independent (asymmetric model) or constrained by $c_{i,j} = c_{j,i}$ (symmetric model). Sections 3.1, 3.2 consider exact solutions and then Section 3.3 considers approximate solutions.

3.1 Symmetric case: exact solution

Let us first consider a symmetric model in which the coefficients $c_{i,j}$ are random integers in the range $[0..B]$. Frieze [36] described an algorithm which finds an exact solution whp provided $B = o(n/\log\log n)$.

The strategy of the algorithm is to first find a set X_0 of *troublesome* vertices and then to find a set of vertex disjoint paths $\mathcal{P} = \{P_1, P_2, \ldots, P_t\}$ which cover X_0 as cheaply as possible. By cover we mean that each vertex of X_0 is an interior vertex of one of the paths of $\mathcal{P}$. We choose $\mathcal{P}$ to minimize $c(\mathcal{P}) = \sum_{i=1}^{t} c(P_i)$ where $c(P_i)$ is the weight of the edges of P_i.

Having found $\mathcal{P}$ we then find a Hamiltonian cycle H which contains the paths of $\mathcal{P}$ as sub-paths and which otherwise only contains zero length edges. It is easy to see that H solves the associated traveling salesman problem.

For $B = o(n/(\log n)^{1/2})$ it suffices to take $X_0 = \{v \in [n] : d_0(v) \leq d/2\}$ where $d_0(v)$ is the number of zero length edges incident with v and $d = n/B$ is (close to) the expected number of zero length edges incident with a vertex. So in the construction of H we essentially have to find a Hamiltonian cycle in a random graph with minimum degree $\geq d/2$. A modification of the algorithm HAM of Section 2.3.1 will suffice. For larger B we have to augment X_0 by other vertices. There are many details best left to the interested reader of [36].

3.2 Asymmetric case: exact solution

In this section we consider randomly generated asymmetric problems. In particular we discuss a result of Frieze, Karp and Reed [43]. The main result of [43] is the following: For an $n \times n$ matrix C let $AP(C)$ be the minimum value for the assignment problem with matrix C and let $ATSP(C)$ be the minimum cost of a traveling salesman tour with costs C. We always have $AP(C) \leq ATSP(C)$. The following provides sufficient conditions for equality whp.

Theorem 3.1 *Let $\{X_n\}$ be a sequence of random variables over the nonnegative reals. Let $p = p_n = \mathbb{P}[X_n = 0]$ and let $\omega = \omega(n) = np$. Let $C = C(n)$ be an $n \times n$ matrix whose entries are drawn independently from the same distribution as X_n. If $\omega \to \infty$ as $n \to \infty$ then $AP(C) = ATSP(C)$ whp.*

We see easily that this implies Theorem 2.9 for the case $c_n \to +\infty$. Let $X_n = 0$ or 1. Now $np = \log n + c_n \to \infty$ and it is known from Erdős and Rényi [32] that $AP(C) = 0$ whp. Thus $ATSP(C) = 0$ whp i.e. whp there is a Hamiltonian cycle in the graph induced by zero length edges.

Let H be the weighted bipartite graph with vertex set $X \cup Y$, where $X = \{x_1, x_2, \ldots, x_n\}$ and $Y = \{y_1, y_2, \ldots, y_n\}$, and with an edge of weight $c_{i,j}$ between x_i and y_j. Let D be the complete digraph on vertex set $[n]$, in which each edge (i, j) has weight $c_{i,j}$. A *cycle cover* is a subgraph of D in which each of the n vertices has in-degree 1 and out-degree 1. The Assignment Problem can be stated in any of the following equivalent forms:

- Find a perfect matching of minimum weight in H.
- Find a cycle cover of minimum weight in D.
- Find a permutation σ (of $[n]$) to minimize $\sum_{i=1}^{n} c_{i,\sigma(i)}$.

Let the indicator variable $z_{i,j}$ be 1 if $c_{i,j} = 0$ and 0 otherwise. Then the $z_{i,j}$ are independent, and each $z_{i,j}$ is equal to 1 with probability p. Emulating a useful trick due to Walkup [93] we view the $z_{i,j}$ as being generated in the following way. Let $h = h(n)$ be defined by the equation $1 - p = (1 - h)^5$ and let $z_{i,j}^k$, for $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, n$ and $k = 1, 2, 3, 4, 5$, be independent indicator variables, each of which is equal to 1 with probability h. Let $z_{i,j} = \max_{k=1}^{5} z_{i,j}^k$. Then the $z_{i,j}$ are independent, and each is equal to 1 with probability p. For $k = 1, 2$, let H_k be the bipartite graph with vertex set $X \cup Y$, and with an edge between x_i and y_j if and only if $z_{i,j}^k = 1$.

For $k = 3, 4, 5$, let D_k be the digraph with vertex set $[n]$ and an edge from i to j if and only if $z^k_{i,j} = 1$. The edges of D_3, D_4 and D_5, respectively, will be called *out-edges*, *in-edges* and *patch edges*. Each type of edge will play a special role in the construction of a Hamiltonian circuit of weight $AP(C)$. It will be important that the random graphs H_1 and H_2, and the random digraphs D_3, D_4 and D_5, are completely independent. Also, let $s = s(n) = nh$; s is the expected degree of a vertex in H_1 or H_2, and the expected out-degree of a vertex in D_3, D_4 or D_5. Clearly, $s \geq \omega/5$, and thus s tends to infinity if ω does.

The construction of the desired Hamiltonian circuit proceeds in the following stages:

(a) (Identification of *troublesome vertices*). By considering the edges of $H_1 \cup H_2$ identify a set $A \subset X$ and a set $B \subset Y$, $|A| = |B|$. The cardinality of $A \cup B$ is small whp.

The set $A \cup B$ contains the vertices of exceptionally small degree plus certain other vertices that are likely to be incident with edges of nonzero weight in an optimal assignment. At the same time construct a matching in H which is of minimum weight, subject to the condition that it covers the vertices in $A \cup B$ and no other vertices. (Compare with the algorithm of the previous section.)

(b) Consider the subgraph of $H_1 \cup H_2$ induced by $(X \setminus A) \cup (Y \setminus B)$. This bipartite graph has a perfect matching whp. Combining that perfect matching with the matching constructed in the previous step, obtain an optimal assignment for H in which every non-zero-weight edge is incident with a vertex in $A \cup B$.

(c) The optimal assignment just constructed has the properties of a random permutation. In particular it has $O(\log n)$ cycles.

(d) Using the out-edges and in-edges, attempt to convert the original optimal assignment into a permutation with no short cycles. This process succeeds whp.

(e) Using the patch edges, patch the long cycles together into a single cycle, thus solving the $ATSP$, much as in the patching algorithm of Section 2.3.2. The patching process succeeds whp.

The overall strategy of the proof is to construct an optimal assignment while keeping the in-edges, out-edges and patch edges (except those incident with $A \cup B$) in reserve for use in converting the optimal assignment to a tour without increasing cost.

In this summary we focus on Step (d). The reason we want to get large cycles is that if we have two cycles C_1 and C_2, then we get $\Omega(|C_1||C_2|)$ opportunities to create patches and we need to be sure that this is significantly greater than the inverse of the probability s^2/n^2 of making a single patch in D_5. So if C_1 and C_2 are large then $|C_1||C_2| \geq n^2/\omega$ and this is large compared with n^2/s^2.

3.2.1 Elimination of Small Cycles

Call a cycle in a permutation *small* if it contains fewer than $n/\sqrt{\omega}$ vertices. We now discuss how the out-edges and in-edges are used to convert the original optimal assignment into an optimal assignment in which no cycle is small. Our procedure is to take each small cycle of the original optimal assignment σ in turn and try to remove it without creating any new small cycles.

We now describe the *rotation-closure algorithm* that is used to eliminate one small cycle. Let C be a small cycle. We make a number of separate attempts to remove C. The ith attempt consists of an Out Phase and an In Phase. We will ignore some technical points which are necessary to maintain some independence.

The Out Phase

Define a *near-cycle-cover* as a digraph θ consisting of a directed path P_θ plus a set of vertex-disjoint directed cycles covering the vertices not in P_θ. We obtain an initial near-cycle-cover by deleting an edge of C from the current (optimal) assignment, thus converting the small cycle C into a path. We then attempt to obtain many near-cycle-covers by a rotation process. The state of this process is described by a rooted tree whose nodes are near-cycle-covers, with the original near-cycle-cover at the root. (Compare with HAM of Section 2.3.1.) Consider a typical node θ consisting of a path P_θ directed from a_θ to b_θ plus a cycle cover of the remaining vertices. We obtain descendants of θ by looking at out-edges directed from b_θ. Consider an edge that is directed from b_θ to a vertex y with predecessor x. Such an edge is *successful* if either y lies on a large cycle or y lies on P_θ and the sub-paths of P_θ from a_θ to x and from y to b_θ are both of length at least $n/\sqrt{\omega}$. In such cases a descendant of θ is created by deleting (x, y) and inserting (b_θ, y). The tree of near-cycle-covers is grown in a breadth-first manner until the number of leaves reaches $m = \sqrt{n \ln n}$.

Assuming that the number of vertices on short cycles is less than $n/\omega^{1/3}$ (this is true whp) and ignoring some technicalities, the number of descendants of node θ is a is a binomial random variable $B(n - o(n), s/n)$, and the random variables associated with distinct nodes are independent. Suppose that level t of the rooted tree describing the Out Phase has a vertices. Then, applying the Chernoff bound (1.4) on the tails of the binomial, the number of nodes at level $t + 1$ lies between $as/2$ and $2as$, with probability greater than or equal to $1 - e^{-as/10}$. Hence the probability that the Out Phase fails to produce m leaves is (quite conservatively) at most $\sum_{k=1}^{\infty} e^{-ks/10} \leq e^{-s/20}$.

The In Phase

The tree produced by an Out Phase has m terminal nodes. Each of these is a near-cycle-cover in which the directed path begins at the same vertex v. Let the jth terminal node be denoted G_j, and let the directed path in G_j run from v to x_j. During the In Phase we grow rooted trees independently from all the G_j, $j = 1, 2, \ldots, m$. The process is like the Out Phase, except that, in computing the descendants of a node θ, we fan backwards along in-edges, rather than forwards along out-edges.

If all goes according to plan, we end the In Phase with at least $m^2/2$ near-cycle-covers. The conditional probability that we cannot close any of the paths of these covers into large cycles is at most $(1 - s/n)^{m^2/2} \leq n^{-s/2}$. Closing a path creates a new optimal assignment with at least one less small cycle.

3.3 Asymmetric case: approximation algorithm

We now consider approximation algorithms for the case where the entries $c_{i,j}$ of our $n \times n$ matrix C are independent uniform [0,1] random variables. Karp [54] was the first to prove the rather surprising fact that

$$ATSP(C) = AP(C) + o(1) \geq 1 - o(1) \qquad whp. \tag{3.1}$$

He constructed an $O(n^3)$ patching algorithm to do this. We see then that Karp's algorithm is *asymptotically optimal.* (Note that the lower bound in (3.1) has been increased to more than 1.5 and it is conjectured that $AP(C) \approx \frac{\pi^2}{6}$ whp.) Later Karp and Steele [55] improved this to

$$ATSP(C) = AP(C) + O(n^{-1/2}) \qquad whp$$

and Dyer and Frieze [27] made further improvements showing $ATSP(C) = AP(C) + O((\log n)^4/(n \log\log n))$. More recently, Frieze and Sorkin [46] have made a small improvement, showing

$$ATSP(C) - AP(C) \leq c_1 \frac{(\log n)^2}{n} \qquad whp \tag{3.2}$$

for some constant $c_1 > 0$.

We now outline the proof of (3.2). It was shown in [46] that whp the optimum assignment solution (cycle cover) only has edges of length $\leq c_2 \frac{\log n}{n}$ whp for some constant $c_2 > 0$. We let an edge be coloured
Red: $c(i,j) \in [0, c_2 \frac{\log n}{n}]$; Blue : $c(i,j) \in [c_2 \frac{\log n}{n}, 2c_2 \frac{\log n}{n}]$;
Green: $c(i,j) \in [2c_2 \frac{\log n}{n}, 3c_2 \frac{\log n}{n}]$; Black otherwise.

Then whp the optimum assignment solution (cycle cover) $\mathcal{C}$ will consist entirely of Red edges. By symmetry it will have the cycle structure of a random cycle cover, i.e. have $\leq 2 \log n$ cycles whp.

Let a cycle of $\mathcal{C}$ be small if its length is $\leq n/(\log n)^{1/3}$. We use the rotation-closure algorithm of the previous section to remove small cycles. To grow the trees we are only allowed to use edges (i,j) which either one of the 10 shortest out of i or one of the 10 shortest into j. One can show that whp the trees grow at a rate of at least 3 per level and so they only need to be grown to a depth of $O(\log n)$. The expected length of these edges is $O(1/n)$. When we have grown enough paths, we can whp close one using a Blue edge. Thus whp it costs $O(\frac{\log n}{n})$ to remove a small cycle, the cost of the added Red/Blue edges. There are $\leq 2 \log n$ cycles altogether and so it costs $O((\log n)/n)$ to make all cycles of length at least $n/(\log n)^{1/3}$.

We can then whp patch all these cycles into one at the extra cost of $O((\log n)^{2/3} \times (\log n/n))$ proving (3.2).

3.4 Asymmetric case: enumeration algorithm

The AP can be expressed as a *linear program*:

$$\text{LP}: \text{ Maximise } \sum_{i,j} c(i,j)x_{i,j} \text{ subject to } \sum_i x_{i,k} = \sum_j x_{k,j} = 1, \forall k, 0 \leq x_{i,j} \leq 1, \forall i,j.$$

This has the dual

$$\text{DLP}: \text{ Maximise } \sum_i u_i + \sum_j v_j \text{ subject to } u_i + v_j \leq c(i,j), \forall i,j.$$

Suppose now that we condition on an optimal basis for LP and insist that $u_1 = 0$. It follows that the remaining dual variables are uniquely determined, with probability 1 – they satisfy $2n-1$ linear equations. Furthermore the reduced costs $\bar{c}(i,j) = c(i,j) - u_i - v_j$ of the *non-basic* variables N are independently and uniformly distributed in the interval $[\max\{0, u_i + v_j\}, 1 - u_i - v_j]$. (Within this interval $c(i,j) \in [0,1]$ and $\bar{c}(i,j) \geq 0$.) It is shown in [46] that whp

$$\max_{i,j}\{|u_i|, |v_j|\} = O\left(\frac{\log n}{n}\right). \tag{3.3}$$

Let I_k denote the interval $[2^{-k}c_1\frac{(\log n)^2}{n}, 2^{-(k-1)}c_1\frac{(\log n)^2}{n}]$ for $k \geq 1$. It follows from (3.3) and the distribution of the reduced costs $\bar{c}(i,j)$ of the non-basic variables, that whp (i) there are $\leq c_1 2^{-(k-1)} n \log n$ non-basic variables $x_{i,j}$ whose reduced cost is in $I_k, 1 \leq k \leq k_0 = \frac{1}{2}\log_2 n$ and (ii) there are $\leq 2c_1\sqrt{n}\log n$ non-basic variables $x_{i,j}$ whose reduced cost is $\leq c_1 \frac{(\log n)^2}{n^{3/2}}$.

We can search for the optimal solution to ATSP by choosing a set of non-basic variables, setting them to 1 and then resolving the assignment problem. If we try all sets and choose the best tour we find, then we will clearly solve the problem exactly. However, it follows from (3.2) that we need only consider sets which contain $\leq 2^k$ variables with reduced costs in I_k and none with reduced cost $\geq c_1\frac{(\log n)^2}{n}$. Thus **whp** we need only check at most

$$2^{2c_1\sqrt{n}\log n} \prod_{k=1}^{k_0} \sum_{t=1}^{2^k} \binom{c_1 2^{-(k-1)} n \log n}{t} = e^{\tilde{O}(\sqrt{n})}$$

sets.

We have thus shown

Theorem 3.2 *ATSP can be solved exactly whp in* $e^{\tilde{O}(\sqrt{n})}$ *time.*

3.5 Open problems

Problem 1 *Is* G_{3-out} *Hamiltonian whp?*

It is unthinkable that the answer is no!

Problem 2 *Is there an extension-rotation algorithmic proof of the fact that random cubic graphs are Hamiltonian?*

Some empirical work suggests that there is such a proof.

Problem 3 *Show that $G_{n,r}$ is Hamiltonian whp in the case where $r = r(n) \to \infty$.*

In unpublished work, [41], it was shown to be true for $r = o(n^{1/5})$.

Problem 4 *For what range of values of m is it true that $G = G_{n,m}$ contains $\lfloor \delta(G)/2 \rfloor$ edge disjoint Hamiltonian cycles whp?*

The paper [18] deals with the case where $m \approx \frac{1}{2} n \log n$.

Problem 5 *Let Q_n denote the n-cube – vertex set $\{0,1\}^n$ and an edge joining two vectors of Hamming distance 1. Let $Q_{n,p}$ be the random subgraph of Q_n obtained by independently deleting edges with probability $1-p$. Find the threshold for the existence of a Hamiltonian cycle in $Q_{n,p}$.*

The approximate threshold for connectivity is $p = \frac{1}{2}$ (Erdős and Spencer [33]); $p = \frac{1}{2}$ is also the approximate threshold for the existence of a perfect matching (Bollobás [15]).

Problem 6 *Is there a polynomial expected time algorithm for checking Hamiltonicity (??) of $G_{n,m}$ for all m?*

Problem 7 *Suppose $c \geq 3/2$ is constant and $p = c/n$ and $G = G_{n,p}$. Is it true that*

$$\lim_{n\to\infty} \mathbb{P}[G_{n,p} \text{ is Hamiltonian} \mid \delta(G) \geq 3] = 1?$$

Bollobás, Cooper, Fenner and Frieze [16] prove this result for c sufficiently large.

Problem 8 *Can you improve the result of Section 3.1 to $B = o(n)$?*

Problem 9 *Is there an analogous algorithm to that of Section 3.3 for the symmetric case where $c_{i,j} = c_{j,i}$ are i.i.d. uniform [0,1]?*

A natural approach is to find a minimum weight 2-factor first and then try to patch the cycles together. The problem with this is proving that such a 2-factor does not have many cycles.

Problem 10 *Suppose that $np = c$ in Theorem 3.1. Determine the limiting probability that $AP(C) = ATSP(C)$.*

It is shown in [43] that this probability is not 1.

Problem 11 *Determine the precise order of magnitude of $|AP(C) - ATSP(C)|$ under the assumptions of Section 3.3. If this is small, can you find a polynomial time algorithm that solves $ATSP(C)$ exactly whp?*

Problem 12 *Suppose the edges of $G_{n,m}$ are randomly colored using n colors. What is the threshold for the existence of a Hamiltonian cycle in which each edge has a different color?*

The same question for spanning trees was solved in Frieze and McKay [45].

4 Euclidean Traveling Salesman Problem

Most of what follows has appeared elsewhere in the literature and our aim is to bring together the various methods and results. The focus is on describing those probabilistic methods used to analyze the TSP which have the potential to describe the behavior of heuristics as well. The methods described here also treat the probabilistic behavior of prototypical problems of Euclidean combinatorial optimization, including the minimal spanning tree problem, the minimal matching problem, and the Steiner minimal spanning tree problem. See the monographs of Steele [85] and Yukich [97] for details. Here we describe two key tools used in the probabilistic analysis of the total edge length of the Euclidean TSP on random samples of large size. The first tool is the boundary functional method and the second tool involves the isoperimetric inequalities of Rhee and Talagrand.

4.1 Basic properties of the Euclidean TSP

If $F \subset \mathbb{R}^d$ is a point set and $R \subset \mathbb{R}^d$ a d-dimensional rectangle, then we let $T(F, R)$ denote the total edge length of the shortest tour through $F \cap R$. We view T as a function on pairs of the form (F, R), where F is a finite set and R is a rectangle. When $R = [0,1]^d$ we write $T(F)$ for $T(F, [0,1]^d)$. This notation emphasizes that T is a function of two arguments and helps draw out the sub-additivity and super-additivity intrinsic to T. Elementary but essential properties of T include:

(a) *Monotonicity.* If $F \subseteq G$, then $T(F, R) \leq T(G, R)$ for all rectangles R.
(b) *Scaling* (homogeneity). For all $\alpha > 0$, for all rectangles R, and for all $F \subset R$

$$T(\alpha F, \alpha R) = \alpha T(F, R).$$

(c) *Translation invariance.* For all $y \in \mathbb{R}^d$, for all rectangles R, and for all $F \subset R$ we have

$$T(F, R) = T(F + y, R + y).$$

(d) *Geometric sub-additivity.* Subdivide $[0,1]^d$ into m^d sub-cubes $Q_1, Q_2, ..., Q_{m^d}$ of edge length m^{-1}. Given $F \subset [0,1]^d$, let $T(F \cap Q_i)$ denote the length of the shortest tour through $F \cap Q_i$. By adding and deleting edges it is easy to see that the length of the shortest tour through F is bounded above by the sum of the tour lengths $T(F \cap Q_i)$, $1 \leq i \leq m^d$, plus at most $2m^d$ edges each of length at most twice the diagonal of any sub-cube Q_i. This last set of edges is used to connect the minimal tours on each set $F \cap Q_i$, $1 \leq i \leq m^d$, into a feasible grand tour through F. Thus we have shown that T satisfies geometric sub-additivity with an error term:

$$T(F, [0,1]^d) \leq \sum_{i=1}^{m^d} T(F \cap Q_i, Q_i) + C_1 m^{d-1}, \tag{4.1}$$

where C_1 is a finite constant.

(e) *Growth bounds.* For all $d = 2, 3, \ldots$ there is a constant $C_2 := C_2(d)$ such that for all $F \subset [0,1]^d$ we have

$$T(F, [0,1]^d) \le C_2 |F|^{(d-1)/d}. \tag{4.2}$$

These growth bounds follow from an easy application of the pigeonhole principle and, as noted by Rhee [72], also follow from geometric sub-additivity [97].

(f) *Smoothness* (Hölder continuity). For all $d = 2, 3, \ldots$ there is a constant $C_3 := C_3(d)$ such that for all sets $F,\ G \subset [0,1]^d$, T satisfies the smoothness condition:

$$|T(F \cup G) - T(F)| \le C_3 |G|^{(d-1)/d}.$$

Since T is monotonic, the proof of smoothness follows once we show

$$T(F \cup G) \le T(F) + C_3 |G|^{(d-1)/d}.$$

This last estimate is immediate since the length of the shortest tour through $F \cup G$ is bounded above by the sum of the length $T(F)$ of the shortest tour through F, the length $T(G)$ of the shortest tour through G, and the length of two edges connecting the sets F and G. Since $T(G) \le C_2 |G|^{(d-1)/d}$, smoothness follows.

If a functional such as the TSP satisfies conditions (b),(c),(d), and (f), then we say that it is a *smooth sub-additive Euclidean functional* [81], [97].

The probabilistic analysis of the Euclidean TSP is simplified conceptually and technically by two key ideas, which are developed in the next sections. The first idea involves the concentration of the shortest tour length around its average value. By showing that the Euclidean TSP is tightly concentrated around its mean value, the probabilistic analysis of the TSP often reduces to an analysis of the behavior of its average value. This idea lies at the heart of asymptotic analysis. The second key idea involves a modified TSP functional, called the *boundary functional*, which closely approximates the standard TSP, and which also has a natural super-additive structure. By combining super-additivity together with sub-additivity, the TSP functional becomes "nearly additive" in the sense that

$$T(F \cap R) \approx T(F \cap R_1) + T(F \cap R_2),$$

where R_1 and R_2 form a partition of the rectangle R. Relations of this sort are crucial in showing that the global TSP tour length can be approximately expressed as a sum of the lengths of local components.

4.2 The concentration of the TSP around its mean

The following basic result provides the a.s. asymptotics for $T(U_1, \ldots, U_n)$, where $U_1, \ldots, U_n$ are i.i.d. uniform random variables in $[0,1]^d$, $d \ge 2$. These asymptotics were first proved by Beardwood, Halton, and Hammersley [9].

Theorem 4.1 *For all $d = 2, 3, \ldots$ there is a finite positive constant $\beta(d)$ such that*

$$\lim_{n\to\infty} \frac{T(U_1, \ldots, U_n)}{n^{(d-1)/d}} = \beta(d) \quad a.s.$$

Thus, the length of the average edge is roughly $\beta(d)n^{-1/d}$. In dimension 2 Rhee and Talagrand [76] show that the length of the longest edge in the TSP tour on $U_1, \ldots, U_n$ is bounded above by $C(\log n/n)^{1/2}$ with high probability. Theorem 4.1 shows that there are not many "long" edges.

There are many ways to prove Theorem 4.1 and we refer to the classic papers of Steele [81], [84] for simple proofs as well as non-trivial generalizations. Perhaps the easiest way to prove Theorem 4.1 involves showing that

$$\lim_{n\to\infty} \frac{\mathbb{E}T(U_1, \ldots, U_n)}{n^{(d-1)/d}} = \beta(d), \tag{4.3}$$

and then showing that the total tour length $T(U_1, \ldots, U_n)$ is closely approximated by the mean tour length $\mathbb{E}T(U_1, \ldots, U_n)$.

The isoperimetric method lies at the heart of this approach. Loosely speaking, isoperimetric methods show that suitably regular functions are "close" to their average values [88]. For a full appreciation of the power of these methods, the reader should consult Section 4.8 below. For the moment we will consider only the following isoperimetric inequality. It is stated in a generality which lends itself to the study of both heuristics and a panoply of related problems in Euclidean combinatorial optimization.

Theorem 4.2 *(Rhee's isoperimetric inequality [72]) Let $X_i,\ i \geq 1$, be independent random variables with values in $[0,1]^d,\ d \geq 2$. Let L be a smooth, sub-additive Euclidean functional. Then there is a constant $C_4 := C_4(d)$ such that for all $t > 0$*

$$\mathbb{P}[|L(X_1, \ldots, X_n) - \mathbb{E}L(X_1, \ldots, X_n)| > t] \leq C_4 \exp\left(-\frac{(t/C_3)^{2d/(d-1)}}{C_4 n}\right) \tag{4.4}$$

The estimate (4.4) is an example of a deviation inequality: it tells us how L deviates from its average value. We call (4.4) an isoperimetric inequality since its proof [72],[97] rests upon an isoperimetric inequality for the Hamming distance on $([0,1]^d)^n$. The upshot of (4.4) is that, excepting a set with polynomially small probability, the functional $L(X_1, \ldots, X_n)$ and its mean $\mathbb{E}L(X_1, \ldots, X_n)$ do not differ by more than $C(n \log n)^{(d-1)/2d}$. The most useful consequence of Rhee's concentration estimate (4.4) is that it reduces the problem of showing complete convergence of L to one of showing the convergence of the mean of L. The mean of L is a scalar and showing convergence of scalars is usually easier than showing convergence of random variables.

Corollary 4.3 *(convergence of means implies complete convergence) Let $X_i,\ i \geq 1$, be i.i.d. random variables with values in $[0,1]^d,\ d \geq 2$. Let L be a functional which is homogeneous, translation invariant, and smooth. If the mean of L converges in the sense that*

$$\lim_{n\to\infty} \mathbb{E}L(X_1, \ldots, X_n)/n^{(d-1)/d} = \alpha(L, d)$$

then

$$\lim_{n\to\infty} L(X_1,...,X_n)/n^{(d-1)/d} = \alpha(L,d) \quad c.c.$$

Proof The deviation estimate (4.4) implies for all $\epsilon > 0$

$$\sum_{n=1}^{\infty} \mathbb{P}\left\{\left|\frac{L(X_1,...,X_n) - \mathbb{E}L(X_1,...,X_n)}{n^{(d-1)/d}}\right| > \epsilon\right\} \le C\sum_{n=1}^{\infty} \exp\left(-\left(\frac{\epsilon}{C_3}\right)^{2d/(d-1)} \frac{n}{C_4}\right).$$

Thus, by definition, $\left|\frac{L(X_1,...,X_n)-\mathbb{E}L(X_1,...,X_n)}{n^{(d-1)/d}}\right|$ converges completely to zero and the proof is complete. □

Thus to prove Theorem 4.1 it will be enough to show the limit (4.3). The next section describes a method of proof involving the boundary TSP. This approach helps prove probabilistic statements that go considerably beyond Theorem 4.1. For estimates of $\beta(d)$ we refer to [70] and pages 50-51 of [97].

4.3 The boundary TSP

Given $F \subset [0,1]^d$, the *boundary TSP functional* $T_B(F)$ is, loosely speaking, the cost of the least expensive tour through F, where the cost of travel within $[0,1]^d$ is the usual Euclidean distance, but where travel along any path on the boundary $\partial[0,1]^d$ of the unit cube is free.

We provide a more precise and slightly more general definition as follows. For all rectangles R, finite sets $F \subset R$, and pairs $\{a,b\} \subset \partial R$, let $T(F,R,\{a,b\})$ denote the length of the shortest path through $F \cup \{a,b\}$ with endpoints a and b. The boundary TSP functional is given by

$$T_B(F,R) := \min\left(T(F,R),\ \inf \sum_i T(F_i,R,\{a_i,b_i\})\right),$$

where the infimum ranges over all partitions $(F_i)_{i\ge 1}$ of F and all sequences of pairs of points $(a_i,b_i)_{i\ge 1}$ belonging to ∂R.

Boundary functionals were first used by Redmond [66] and Redmond and Yukich [67], [68] in the analysis of general Euclidean functionals. They are reminiscent of the "wired boundary condition" and the "wired spanning forest" used in the study of percolation and random trees, respectively. T_B may be interpreted as the length of the shortest tour through F which may repeatedly exit to the boundary of R at one point and re-enter at another, incurring no cost for travel along the boundary.

The boundary TSP functional T_B satisfies geometric super-additivity: if the rectangle R is partitioned into rectangles R_1 and R_2 then

$$T_B(F,R) \ge T_B(F,R_1) + T_B(F,R_2). \tag{4.5}$$

To see this, note simply that the restriction of the global tour realizing $T_B(F,R)$ to rectangle R_i, $1 \le i \le 2$, produces a feasible boundary tour of $F \cap R_i$, $1 \le i \le 2$, which by minimality has a length which is at least as large as $T_B(F, R_i)$. The absence of an error term in super-additivity contrasts sharply with the geometric sub-additivity of T. This distinction, which has telling consequences, often makes the analysis of T_B far easier than the analysis of T.

By definition $T_B \leq T$. It is easily checked that T_B satisfies all of the basic properties of the usual Euclidean TSP, excepting geometric sub-additivity.

To prove convergence of the mean of T (which by Corollary 4.3 implies c.c. convergence) over independent random variables $U_1, U_2, ..., U_n$ with the uniform distribution on $[0,1]^d$, namely to prove the limit

$$\lim_{n\to\infty} \frac{\mathbb{E}T(U_1, ..., U_n)}{n^{(d-1)/d}} = \beta(d),$$

it is enough to prove the following two basic lemmas:

Lemma 4.4 *(Asymptotics for the boundary functional) For all $d = 2, 3, ...$ there is a constant $\beta(d)$ such that*

$$\lim_{n\to\infty} \frac{\mathbb{E}T_B(U_1, ..., U_n)}{n^{(d-1)/d}} = \beta(d).$$

Lemma 4.5 *(T_B approximates T) The following approximation holds:*

$$|\mathbb{E}T_B(U_1, ..., U_n) - \mathbb{E}T(U_1, ..., U_n)| = o(n^{(d-1)/d}).$$

The proof of Lemma 4.5 is deferred to the next section. There are at least two ways to prove Lemma 4.4. The first way, which bears a likeness to the original bare hands approach of Beardwood et al. [9], relies on straightforward analysis of super-additive sequences of real numbers. The second way draws heavily on a multi-parameter super-additive ergodic theorem.

First proof of Lemma 4.4. Set $\phi(n) = \mathbb{E}T_B(U_1, ..., U_n)$. Partition $[0,1]^d$ into m^d sub-cubes $Q_1, Q_2, ..., Q_{m^d}$ of edge length m^{-1}. The number of points from the sample $(U_1, ..., U_n)$ which fall in a given sub-cube of $[0,1]^d$ of volume m^{-d} is a binomial random variable $B(n, m^{-d})$ with parameters n and m^{-d}. It follows from scaling and the super-additivity (4.5) of T_B that

$$\phi(n) \geq m^{-1} \sum_{i=1}^{m^d} \phi(B(n, m^{-d})).$$

By smoothness and Jensen's inequality (1.1) in this order we have

$$\begin{aligned}\phi(n) &\geq m^{-1} \sum_{i=1}^{m^d} \left(\phi(nm^{-d}) - C_3\mathbb{E}(|B(n, m^{-d}) - nm^{-d}|^{(d-1)/d})\right) \\ &\geq m^{-1} \sum_{i=1}^{m^d} \left(\phi(nm^{-d}) - C_3(nm^{-d})^{(d-1)/2d}\right).\end{aligned}$$

Simplifying, we get

$$\phi(n) \geq m^{d-1}\phi(nm^{-d}) - C_3 m^{(d-1)/2} n^{(d-1)/2d}.$$

Dividing by $n^{(d-1)/d}$ and replacing n by nm^d yields the homogenized relation

$$\frac{\phi(nm^d)}{(nm^d)^{(d-1)/d}} \geq \frac{\phi(n)}{n^{(d-1)/d}} - \frac{C_3}{n^{(d-1)/2d}}.$$

Set $\beta := \beta(d) := \limsup_{n\to\infty} \frac{\phi(n)}{n^{(d-1)/d}}$ and note that $\beta \leq C_2$ by growth bounds (4.2). (Here and elsewhere the symbol ":=" means "equal to by definition".) For all $\epsilon > 0$, choose n_o such that for all $n \geq n_o$, we have $C_3/n^{(d-1)/2d} \leq \epsilon$ and $\phi(n_o)/n_o^{(d-1)/d} \geq \beta - \epsilon$. Thus, for all $m = 1, 2, ...$ it follows that

$$\frac{\phi(n_o m^d)}{(n_o m^d)^{(d-1)/d}} \geq \beta - 2\epsilon.$$

To now obtain Lemma 4.4 we use the smoothness of T_B and a simple interpolation argument. For an arbitrary integer $k \geq 1$ find the unique integer m such that

$$n_o m^d < k \leq n_o(m+1)^d.$$

Then $|n_o m^d - k| \leq C n_o m^{d-1}$, where here and elsewhere C denotes a finite positive constant whose value may change from line to line. By smoothness we therefore obtain

$$\begin{aligned}
\frac{\phi(k)}{k^{(d-1)/d}} &\geq \frac{\phi(n_o m^d)}{(n_o(m+1)^d)^{(d-1)/d}} - \frac{C_3(C n_o m^{d-1})^{(d-1)/d}}{(m+1)^{d-1} n_o^{(d-1)/d}} \\
&\geq (\beta - 2\epsilon)(\frac{m}{m+1})^{d-1} - \frac{C_3(C n_o m^{d-1})^{(d-1)/d}}{(m+1)^{d-1} n_o^{(d-1)/d}}.
\end{aligned}$$

Since the last term in the above goes to zero as m goes to infinity, it follows that

$$\liminf_{k\to\infty} \phi(k)/k^{(d-1)/d} \geq \beta - 2\epsilon.$$

Now let ϵ tend to zero to see that the lim inf and the lim sup of the sequence $\frac{\phi(k)}{k^{(d-1)/d}}$, $k \geq 1$, coincide, that is

$$\lim_{k\to\infty} \frac{\phi(k)}{k^{(d-1)/d}} = \beta.$$

We have thus shown

$$\lim_{n\to\infty} \frac{\mathbb{E}T_B(U_1, ..., U_n)}{n^{(d-1)/d}} = \beta$$

as desired. In the next few paragraphs we will see that β is positive. This concludes the first proof of Lemma 4.4. □

Before providing the second proof of Lemma 4.4 we recall a general super-additive ergodic theorem. Let $\mathcal{R}(d)$ denote the d-dimensional rectangles of $\mathbb{R}^d$. Let $L =$

$\{L(R),\ R \in \mathcal{R}(d)\}$ be a multi-parameter functional defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. L is *stationary* if for all $m \geq 1$, $R_1, ..., R_m \in \mathcal{R}(d)$, and $u \in (\mathbb{R}^+)^d$, the joint distributions of $L(R_1), ..., L(R_m)$ and $L(R_1 + u), ..., L(R_m + u)$ are the same. L is *bounded* if $\sup_n \mathbb{E}L([0,n]^d)/n^d < \infty$ and L is *strongly super-additive* if $L(R) \geq \sum_{i=1}^m L(R_i)$, where the rectangles $R_1, R_2, ..., R_m$ form a partition of R. Notice that strong super-additivity is stronger than the usual super- additivity.

The following strong law of large numbers of Akcoglu and Krengel [2] generalizes Kingman's sub-additive ergodic theorem.

Theorem 4.6 *(Super-additive ergodic theorem) Let* $L := \{L(R) : R \in \mathcal{R}(d)\}$ *be a stationary, bounded, super-additive functional defined on* $(\Omega, \mathcal{A}, \mathbb{P})$. *Then*

$$\lim_{n\to\infty} \frac{L([0,n]^d)}{n^d} = f(L,d)$$

a.s. and in L^1, *where* $f(L,d) \in L^1(\Omega, \mathcal{A}, \mathbb{P})$. *Moreover,*

$$\mathbb{E}f(L,d) = \alpha(L,d) = \sup_R \frac{\mathbb{E}L(R)}{volume R}.$$

$\alpha(L,d)$ is the *spatial constant* for the process L. It is a generalization of the *time* constant in the theory of one-dimensional sub-additive processes. It is now an easy matter to provide an alternate proof of Lemma 3.4.

Second proof of Lemma 4.4 [96] . Set

$$T_B(R) = T_B(\Pi \cap R, R), \quad R \in \mathcal{R}(d),$$

where Π is a Poisson point process on $(\mathbb{R}^+)^d$ with intensity 1. T_B is stationary and strongly super-additive. Since $\Pi \cap [0,n]^d = n(U_1, ..., U_N)$, where N is an independent Poisson random variable with parameter n^d, we see that T_B is bounded:

$$\begin{aligned}
\mathbb{E}T_B([0,n]^d) &= \mathbb{E}T_B(n(U_1, ..., U_N), [0,n]^d) \\
&= n\mathbb{E}T_B(\{U_1, ..., U_N\}, [0,1]^d) \\
&\leq C_2 n \mathbb{E}N^{(d-1)/d} \\
&\leq C_2 n^d,
\end{aligned}$$

by Jensen's inequality (1.1). The L^1 convergence given by the super-additive ergodic theorem implies

$$\lim_{n\to\infty} \frac{\mathbb{E}T_B([0,n]^d)}{n^d} = \alpha(d) = \sup_{R \in \mathcal{R}(d)} \frac{\mathbb{E}T(R)}{\mathrm{volume} R},$$

where we note that the spatial constant $\alpha(d)$ is clearly positive. Since $\mathbb{E}T_B([0,n]^d) = n\mathbb{E}T_B(U_1, .., U_N)$ we o btain

$$\lim_{n\to\infty} \frac{\mathbb{E}T_B(U_1, ..., U_N)}{n^{d-1}} = \alpha(d).$$

By straightforward smoothness arguments, the above convergence is unaffected if N is replaced by its mean n^d. Moreover, interpolation arguments show that n^d may be replaced by n, which thus yields Lemma 4.4 with $\beta(d) = \alpha(d)$. Notice that this proof shows that $\beta(d)$ is strictly positive. This concludes the second proof of Lemma 4.4. $\square$

4.4 The boundary TSP approximates the standard TSP

To obtain laws of large numbers, rates of convergence of means, and large deviation principles, it is extremely useful to know that the boundary TSP functional closely approximates the standard TSP functional. The following estimate, which establishes Lemma 4.5, is a start in this direction:

$$|\mathbb{E}T(U_1, ..., U_n) - \mathbb{E}T_B(U_1, ..., U_n)| \le Cn^{(d-2)/d}. \tag{4.6}$$

We will show (4.6) by following the approach of Redmond and Yukich [67], [68]. Since $T_B \le T$, in order to prove (4.6) it suffices to show

$$\mathbb{E}T(U_1, ..., U_n) \le \mathbb{E}T_B(U_1, ..., U_n) + Cn^{(d-2)/d}.$$

To show this, we first estimate the cardinality of points which are joined directly to the boundary. As in [97], let F denote one of the faces of $[0,1]^d$. Letting $\mathcal{U}_F \subset \{U_1, ..., U_n\}$ be the set of points that are joined directly to F by the graph realizing T_B, we first show that $\mathbb{E}|\mathcal{U}_F| \le Cn^{(d-1)/d}$. For all $\epsilon > 0$ and $x \in F$, let $C(\epsilon, x)$ denote the cylinder in $[0,1]^d$ determined by the ϵ disk in F centered at x. We make the critical observation that in the part of $C(\epsilon, x)$ which is at a distance greater than ϵ from F, there are at most two points which are joined to F. Were there three or more points, then two of these points could be joined with an edge, which would result in a cost savings, contradicting optimality. Since F can be covered with $O(\epsilon^{-(d-1)})$ disks of radius ϵ, we have the bound

$$\mathbb{E}|\mathcal{U}_F| \le \mathbb{E}\,|\{x \in (U_i)_{i \le n} : \quad d(x,F) \le \epsilon\}| \ + \ C\epsilon^{-(d-1)},$$

where $d(x, F)$ denotes the distance between the point x and the set F. The above is bounded by $n\epsilon \ + \ C\epsilon^{-(d-1)}$ and so putting $\epsilon = n^{-1/d}$ gives the desired estimate $\mathbb{E}|\mathcal{U}_F| \le Cn^{(d-1)/d}$. If $\mathcal{U} \subset \{U_1, ..., U_n\}$ denotes the set of points which are joined directly to any face of the boundary, then

$$\mathbb{E}|\mathcal{U}| \le Cn^{(d-1)/d}.$$

We are now positioned to show (4.6). Let $\mathcal{U}' \subset \partial[0,1]^d$ be the set of points on the boundary of $[0,1]^d$ which are joined to points in $\mathcal{U}$. Let $T(\mathcal{U}')$ (respectively, $MM(\mathcal{U}')$ denote the total edge length of the graph of the minimal tour (respectively, minimal matching) through $\mathcal{U}'$ whose edges lie on $\partial[0,1]^d$. Note that the union of the minimal tour graph through $\mathcal{U}'$, the minimal tour graph through $\mathcal{U}'$, and the boundary TSP graph through $U_1, ..., U_n$ defines an Eulerian path through $\{U_1, ..., U_n\} \cup \mathcal{U}'$, where all vertices have even degree. Deleting some of the edges in this path yields a feasible tour through $\{U_1, ..., U_n\} \cup \mathcal{U}'$. It follows that

$$\begin{aligned}\mathbb{E}T(U_1, ..., U_n) &\leq \mathbb{E}T((U_1, ..., U_n) \cup \mathcal{U}') \\ &\leq \mathbb{E}T_B(U_1, ..., U_n) + \mathbb{E}T(\mathcal{U}') + \mathbb{E}MM(\mathcal{U}').\end{aligned}$$

Now $\mathbb{E}T(\mathcal{U}') \leq C\mathbb{E}|\mathcal{U}'|^{(d-2)/(d-1)}$ since $\mathcal{U}'$ lies in the union of sets of dimension $d-1$. By Jensen's inequality (1.1), we find $\mathbb{E}T(\mathcal{U}') \leq Cn^{(d-2)/d}$. Since $\mathbb{E}MM(\mathcal{U}') \leq \mathbb{E}T(U')$ we have shown (4.6) as desired. □

The approximation (4.6) is probabilistic. With a little extra effort one can show the following deterministic estimate. Let $F \subset [0,1]^d$ and $|F| = n$. If $m^d = n/\sigma(n)$, where $\sigma(n)$ is an unbounded increasing function of n, then as explained in detail in [97], T satisfies the following near additivity condition

$$|T(F) - \sum_{i=1}^{m^d} T(F \cap Q_i, Q_i)| \leq C\sigma^{-1/d(d-1)} n^{(d-1)/d}. \tag{4.7}$$

The first term on the right hand side of (4.7) arises from approximating $T(F \cap Q_i)$, $1 \leq i \leq m^d$, by the length of the restriction of the global tour to sub-cube Q_i and the last term in (4.7) is just the sub-additive error m^{d-1} from (4.1).

In other words,

$$|T(F) - \sum_{i=1}^{m^d} T(F \cap Q_i, Q_i)| = o(n^{(d-1)/d}). \tag{4.8}$$

The approximation (4.6) readily shows that asymptotics for $\mathbb{E}T(U_1, ..., U_n)$ follow from the asymptotics for $\mathbb{E}T_B(U_1, ..., U_n)$. However, these approximations, together with approximations (4.7) and (4.8) carry many additional benefits and provide:

- a straightforward probabilistic analysis of partitioning heuristics,
- estimates for the *rate of convergence* of $\mathbb{E}T(U_1, ..., U_n)$,
- asymptotics for $T(X_1, ..., X_n)$, where X_i, $i \geq 1$, are i.i.d. with an arbitrary distribution, and
- large deviation principles for $T(U_1, ..., U_n)$.

The following sections explore these benefits.

4.5 Analysis of heuristics

The a.s. asymptotics of Theorem 4.1 led Karp [52], [53] to find efficient methods for approximating the length $T(U_1, ..., U_n)$ of the shortest path through i.i.d. uniformly distributed random variables $U_1, ..., U_n$ on the unit square. In his seminal work [52], [53], Karp developed the "fixed dissection algorithm" which provides a simple heuristic $T_H(U_1, ..., U_n)$ having the property that $T_H(U_1, ..., U_n)/T(U_1, ..., U_n)$ converges completely to 1 and which moreover has polynomial mean execution time.

Karp's fixed dissection algorithm consists of dividing the unit cube $[0,1]^d$ into m^d congruent sub-cubes $Q_1, ..., Q_{m^d}$, finding the shortest tour T_i of length $T(\{U_1, ..., U_n\} \cap Q_i)$ on each of the sub-cubes, constructing a tour T which links representatives from each T_i, and then deleting excess edges to generate a grand (heuristic) tour through $U_1, ..., U_n$ having length $T_H(U_1, ..., U_n)$.

Karp and Steele [55] show via elementary methods that the partitioning heuristic T_H is ϵ-optimal with probability one:

Theorem 4.7 *(Karp and Steele, [55]) If $m^d := n/\sigma$, where σ is an unbounded increasing function of n, then for all $\epsilon > 0$*

$$\sum_{n=1}^{\infty} \mathbb{P}\left\{ \frac{T_H(U_1, ..., U_n)}{T(U_1, ..., U_n)} \geq 1 + \epsilon \right\} < \infty.$$

Theorem 4.7 shows that the ratio of the lengths of the heuristic tour and the optimal tour converges completely to 1. Given the computational complexity of the TSP, it is remarkable that the optimal tour length is so well approximated by a sum of individual tour lengths, where the sum has polynomial mean execution time.

Since Karp's work [52], [53], considerable attention has been given to developing the probability theory of heuristics. Goemans and Bertsimas [48] develop the a.s. asymptotics for the Held-Karp heuristic [50]. They do this by essentially showing that the Held-Karp heuristic is a smooth sub-additive Euclidean functional and therefore is amenable to the methods discussed here. To develop the probability theory of other heuristics, such at Christofides' heuristic, one could hope to apply modifications of the methods discussed here. This could involve approximating the heuristic by a super-additive "boundary heuristic". Here we limit ourselves to a discussion of the proof of Theorem 4.7 using the approximations (4.7) and (4.8).

Given $F \subset [0,1]^d$ consider the feasible tour through F obtained by solving the optimization problem on the sub-cubes Q_i, $1 \leq i \leq m^d$, and then adding and deleting edges in the resulting graph to obtain a global solution on the set F. This feasible solution, which we call the canonical heuristic H, has a total edge length denoted by $T_H(F, m^d)$. $T_H(F, m^d)$ is the sum of the lengths $T(F \cap Q_i)$, $1 \leq i \leq m^d$, plus a correction term which is bounded by $C_1 m^{d-1}$. Thus T_H satisfies

$$T(F) \leq T_H(F, m^d) \leq \sum_{i=1}^{m^d} T(F \cap Q_i) + C_1 m^{d-1}.$$

As in the previous section, we let the number of sub-cubes depend on the cardinality of F, denoted by $|F|$ for brevity. To make this precise, let $\sigma := \sigma(n)$ denote a function of n such that $\sigma(n)$ and $n/\sigma(n)$ increase up to infinity. Such functions σ define heuristics $H := H(\sigma)$ having length

$$T_{H(\sigma)}\left(F, \frac{|F|}{\sigma(|F|)}\right).$$

Thus the heuristic $H(\sigma)$ subdivides the unit cube into $\frac{|F|}{\sigma(|F|)}$ sub-cubes. If we let $m^d := \frac{|F|}{\sigma(|F|)}$ we obtain from the near additivity condition (4.7)

$$|T(F) - T_{H(\sigma)}(F)| = o(n^{(d-1)/d}). \tag{4.9}$$

Thus the length $T_{H(\sigma)}(F)$ is larger than $T(F)$ by a quantity which is deterministically small compared to $n^{(d-1)/d}$. Therefore the asymptotic behavior of the scaled heuristic

$$T_{H(\sigma)}(X_1, ..., X_n)/n^{(d-1)/d} \tag{4.10}$$

coincides with the asymptotic behavior of

$$T(X_1, ..., X_n)/n^{(d-1)/d}, \tag{4.11}$$

where $X_i,\ i \geq 1$, are i.i.d. random variables with values in the unit cube $[0,1]^d$. We will see shortly (see Theorem 4.10 below) that the ratio (4.11) converges completely to a positive constant whenever the law of X_1 has a continuous part. By (4.9) it thus follows that (4.10) also converges completely to a constant. It therefore follows by standard arguments that the ratio of (4.10) to (4.11) converges completely to 1. We have thus extended Karp and Steele's result to general sequences of random variables:

Theorem 4.8 *(the heuristic T_H is ϵ-optimal over general sequences) For all $\epsilon > 0$ and all i.i.d. sequences $X_i,\ i \geq 1$, of random variables with a continuous part, the heuristic $T_{H(\sigma)}$ is ϵ-optimal:*

$$\sum_{n=1}^{\infty} \mathbb{P}\left\{ \frac{T_H(X_1, ..., X_n)}{T(X_1, ..., X_n)} \geq 1 + \epsilon \right\} < \infty. \tag{4.12}$$

We conclude the discussion of heuristics by verifying that the expected execution time for $T_H(U_1, ..., U_n)$ is polynomially bounded. The required computing time is bounded by

$$T_n := \sum_{i=1}^{n/\sigma(n)} f(N_i),$$

where $N_i := |\{Q_i \cap \{U_1, ..., U_n\}\}|,\ 1 \leq i \leq n/\sigma(n)$, and where $f(N)$ denotes a bound on the time needed to compute $T(F),\ |F| = N$. It is well-known that we may take f to have the form $f(x) = Ax^B 2^x$, for some constants A and B. Since the $N_i,\ 1 \leq i \leq n/\sigma(n)$, are binomial random variables, straightforward calculations [55] show that

$$\mathbb{E}T_n \leq 4An(\sigma(n))^{B-1} \exp(\sigma(n)).$$

Now choose $\sigma(n) = \log n$ to conclude that the expected execution time for the heuristic T_H is $O(n^2 \log^{B-1} n)$, as desired.

4.6 Rates of convergence of mean values

The sub-additivity of the TSP functional is not enough to give rates of convergence. Sub-additivity only yields one sided estimates whereas rate results require two sided estimates. However, since the TSP functional can be made super-additive through use of the boundary TSP functional this will be enough to yield rates of convergence. Similar methods apply for other problems in Euclidean combinatorial optimization [97]. Alexander [5] obtains rate results without the use of boundary functionals. Rhee [73] shows that the following rates are optimal on the unit square provided that $\mathbb{E}T(U_1, ..., U_n)$ is replaced by the Poissonized version $\mathbb{E}T(U_1, ..., U_{N(n)})$.

Theorem 4.9 *(rates of convergence of means) For all $d = 2, 3, ...,$ there is a constant C such that*

$$|\mathbb{E}T(U_1, ..., U_n) - \beta(d)n^{(d-1)/d}| \leq Cn^{(d-2)/d}.$$

Proof We will first prove

$$|\mathbb{E}T(U_1, ..., U_{N(n)}) - \beta(d)n^{(d-1)/d}| \leq Cn^{(d-2)/d}, \tag{4.13}$$

where we adhere to the convention that $N(n)$ denotes an independent Poisson random variable with parameter n. The proof of (4.13) involves simple but useful sub-additive techniques. We set $\phi(n) = \mathbb{E}T(U_1, ..., U_{N(n)})$. It follows from translation invariance, homogeneity, and sub-additivity, that

$$\phi(nm^d) \leq m^{-1} \sum_{i=1}^{m^d} \phi(n) + Cm^{d-1} = m^{d-1}\phi(n) + cm^{d-1}.$$

Dividing by $(nm^d)^{(d-1)/d}$ yields the homogenized relation

$$\frac{\phi(nm^d)}{(nm^d)^{(d-1)/d}} \leq \frac{\phi(n)}{n^{(d-1)/d}} + \frac{C}{n^{(d-1)/d}}.$$

The limit of the left side as m tends to infinity exists and equals $\beta(d)$. Thus

$$\frac{\phi(n)}{n^{(d-1)/d}} - \beta(d) \geq \frac{-C}{n^{(d-1)/d}}$$

or simply

$$\phi(n) - \beta(d)n^{(d-1)/d} \geq -C. \tag{4.14}$$

Setting $\phi_B(n) := \mathbb{E}T_B(U_1, ..., U_{N(n)})$ and exploiting the super-additivity of T_B in the same way that we exploited the sub-additivity of T, we obtain the companion estimate to (4.14) where we may now let $C = 0$:

$$\phi_B(n) - \beta(d)n^{(d-1)/d} \leq 0. \tag{4.15}$$

By closeness (4.6) of T and T_B , Fubini's theorem, and independence, we have

$$\begin{aligned}|\phi_B(n) - \phi(n)| &\leq \mathbb{E}_N |\mathbb{E}_U T_B(U_1, ..., U_N) - \mathbb{E}_U T(U_1, ..., U_N)| \\ &\leq \mathbb{E}_N (C N^{(d-2)/d}) \\ &\leq C(n^{(d-2)/d})\end{aligned}$$

where $\mathbb{E}_N$ and $\mathbb{E}_U$ denote the expectation with respect to the random variables N and U, respectively. Now (4.13) follows from (4.14) and (4.15). It remains to de-Poissonize (4.13) to obtain Theorem 4.9. Notice by smoothness that

$$|\mathbb{E}T(U_1, ..., U_N) - \mathbb{E}T(U_1, ..., U_n)| \leq C\mathbb{E}|N - n|^{(d-1)/d} \leq Cn^{(d-1)/2d},$$

and thus when $d \geq 3$ we easily obtain Theorem 4.9. For $d = 2$ this simple method does not work and to obtain Theorem 4.9 we need slightly more work (see [97]).

This completes the proof of Theorem (4.9). □

4.7 Asymptotics over non-uniform samples

Let $X_i,\ i \geq 1$, be i.i.d. random variables with values in $[0,1]^d,\ d \geq 2$, and let f denote the density of the absolutely continuous part of the law μ of X_1. In their seminal paper, Beardwood, Halton, and Hammersley [9] established the following asymptotics for the length of the shortest tour $T(X_1, ..., X_n)$.

Theorem 4.10 *(asymptotics for the TSP non-uniform samples) For all $d = 2, 3, ...$ we have*

$$\lim_{n\to\infty} \frac{T(X_1, ..., X_n)}{n^{(d-1)/d}} = \beta(d) \int_{[0,1]^d} f(x)^{(d-1)/d} dx \quad c.c. \tag{4.16}$$

The limit (4.16) is proved in two stages. One first proves (4.16) when the density of μ is a step function of the form $\sum_{i=1}^{m^d} \alpha_i 1_{Q_i},\ \alpha_i \in \mathbb{R}^+,$ where $Q_i,\ 1 \leq i \leq m^d$, is a partition of $[0,1]^d$ into sub-cubes of edge length m^{-1}. Then (4.16) is deduced by approximating general densities by step densities. There are several ways to carry out this program. The easiest arguments [81], [85] make clever use of the monotonicity of T, but such arguments are not easily generalized to treat other Euclidean functionals. We now outline a simple approach to (4.16) which is general enough to deliver asymptotics for a wide variety of problems in combinatorial optimization and computational geometry and which has the potential to describe asymptotics for heuristics as well. The approach yields c.c. asymptotics and it can be extended to treat the case when the random variables $X_i,\ i \geq 1$, have unbounded support (Rhee [71]). The methods here are part of the "umbrella approach" described in detail in [97].

The proof of (4.16) depends upon two observations which greatly simplify the analysis. The first is that by Corollary 4.3, it is enough to show that (4.16) holds in expectation, i.e. to show

$$\lim_{n\to\infty} \frac{\mathbb{E}T(X_1, ..., X_n)}{n^{(d-1)/d}} = \beta(d) \int_{[0,1]^d} f(x)^{(d-1)/d} dx. \tag{4.17}$$

The limit (4.17) is a statement about a sequence of scalars and, in principle, is easier to prove than the limit (4.16).

The second observation is that in the presence of smoothness of T, it is enough to establish (4.17) for a special class of distributions which we call *blocked* distributions. These are distributions μ on $[0,1]^d$ with the form $\phi(x)dx + \mu_s$, where $\phi(x)$ is a simple non-negative function of the form $\sum_{i=1}^{m^d} \alpha_i 1_{Q_i}$, where the measure μ_s is purely singular and $Q_i,\ i \geq 1$, are the usual sub-cubes. More precisely, we have the following lemma which is due to Steele [83].

Lemma 4.11 *Suppose that for every sequence of i.i.d. random variables $X_i,\ i \geq 1$, distributed with a blocked distribution $\mu := \phi(x)dx + \mu_s$ we have that*

$$\lim_{n\to\infty} \frac{\mathbb{E}T(X_1,...,X_n)}{n^{(d-1)/d}} = \beta(d) \int_{[0,1]^d} f(x)^{(d-1)/d} dx. \tag{4.18}$$

We then have that

$$\lim_{n\to\infty} \frac{\mathbb{E}T(Y_1,...,Y_n)}{n^{(d-1)/d}} = \beta(d) \int_{[0,1]^d} f(x)^{(d-1)/d} dx \tag{4.19}$$

where $Y_i,\ i \geq 1$, are i.i.d. random variables whose law has an absolutely continuous part given by $f(x)dx$.

Proof Assume that the distribution of Y has the form $\mu_Y := f(x)dx + \mu_s$, where μ_s is singular. For all $\epsilon > 0$ we may find a blocked approximation to μ_Y of the form $\mu_X := \phi(x) + \mu_s$, where $\phi := \phi_\epsilon$ approximates f in the L^1 sense:

$$\int_{[0,1]^d} |\phi(x) - f(x)| dx < \epsilon.$$

By standard coupling arguments there is a joint distribution for the pair of random variables (X, Y) such that $\mathbb{P}[X \neq Y] \leq 2\epsilon$. Thus it follows that

$$\begin{aligned} |\mathbb{E}T(X_1,...,X_n) - \mathbb{E}T(Y_1,...,Y_n)| &\leq C\mathbb{E}|\{i \leq n\ : X_i \neq Y_i\}|^{(d-1)/d} \\ &\leq C(\epsilon n)^{(d-1)/d}. \end{aligned}$$

Thus, by (4.18) we obtain

$$\lim_{n\to\infty} \left| \frac{\mathbb{E}T(Y_1,...,Y_n)}{n^{(d-1)/d}} - \beta(d) \int_{[0,1]^d} \phi(x)^{(d-1)/d} dx \right| \leq C(\epsilon^{(d-1)/d}). \tag{4.20}$$

For all $a,\ b \geq 0$ we have

$$|a^{(d-1)/d} - b^{(d-1)/d}| \leq |a - b|^{(d-1)/d}$$

and therefore by the L^1 approximation

$$\left| \int f(x)^{(d-1)/d} dx - \int \phi(x)^{(d-1)/d} dx \right| \leq \int |f(x) - \phi(x)|^{(d-1)/d} dx \tag{4.21}$$

$$\leq \epsilon^{(d-1)/d}. \tag{4.22}$$

Combining (4.20) and (4.21) and letting ϵ tend to zero gives (4.19) as desired. $\square$

We are now ready to prove Theorem 4.10. By the above lemma, the proof of Theorem 4.10 is reduced to showing (4.18). By simple smoothness arguments it it enough to show

$$\lim_{n\to\infty} \frac{\mathbb{E}T(X_1,...,X_{N(n)})}{n^{(d-1)/d}} = \beta(d)\int_{[0,1]^d}\phi(x)^{(d-1)/d}dx, \tag{4.23}$$

where $N(n)$ is a Poisson random variable with parameter n. For simplicity we will assume $\mu_s = 0$.

Notice that $|\{i \leq N(n) : Y_i \in Q_j\}|$, $1 \leq j \leq m^d$, is a Poisson random variable $N(n\alpha_j m^{-d})$. Let U_i, $i \geq 1$, be i.i.d. uniform random variables with values in $[0,1]^d$. By geometric sub-additivity and scaling we have

$$\mathbb{E}T(X_1,...,X_{N(n)}) \leq m^{-1}\sum_{j=1}^{m^d}\mathbb{E}T(\{U_i\}_{i=1}^{N(n\alpha_j m^{-d})}) + Cm^{d-1}.$$

By Theorem 4.1 it follows that

$$\begin{aligned}\limsup_{n\to\infty}\frac{\mathbb{E}T(X_1,...,X_{N(n)})}{n^{(d-1)/d}} &\leq \sum_{j=1}^{m^d}\limsup_{n\to\infty}\left(\frac{\mathbb{E}T(U_1,...,U_{N(n\alpha_j m^{-d})})}{n^{(d-1)/d}\alpha_j^{(d-1)/d}m^{-(d-1)}}\cdot\alpha_j^{(d-1)/d}m^{-d}\right)\\ &= \beta(d)\int_{[0,1]^d}\phi(x)^{(d-1)/d}dx.\end{aligned}$$

Similarly, by geometric super-additivity and by Lemma 4.4 we see that

$$\liminf_{n\to\infty}\frac{\mathbb{E}T_B(X_1,...,X_{N(n)})}{n^{(d-1)/d}} \geq \beta(d)\int_{[0,1]^d}\phi(x)^{(d-1)/d}dx$$

and the desired limit (4.23) follows since $T \geq T_B$.

4.8 Talagrand's isoperimetric theory

Section 4.2 described how the TSP deviates from its average value. The concentration estimates of Theorem 4.2 are general and apply to functionals which are homogeneous, translation invariant, and smooth. In the case of the TSP, however, the deviation estimates of Theorem 4.2 are far from optimal. In this section we will see that Talagrand's [88], [89] deep isoperimetric methods for product spaces yields improved tail estimates for the TSP. These estimates resemble the tail estimates of a standard normal random variable and suggest that perhaps the TSP is asymptotically normally distributed. Isoperimetric methods have the potential to describe the deviations of heuristics as well.

Talagrand's isoperimetric theory is substantial and has had a profound impact in the probability theory of random graphs. This section illustrates how one piece of Talagrand's theory may be used to study the tail behavior of the TSP. Our presentation is motivated by the fine exposition of Steele [85]. Naturally, we will not present all the technical aspects of the theory and refer the reader to [85] and [88] for complete details.

To study the tail behavior of $T(U_1, ..., U_n)$, where $U_i,\ i \geq 1$, are i.i.d. with the uniform distribution on the unit cube $[0,1]^d$, it is useful to view T as a functional defined on the n-fold product space $\Omega^n := ([0,1]^d)^n$. We adopt this point of view and will thus consider isoperimetric methods on the product space Ω^n, where the triple $(\Omega, \mathcal{A}, \mathcal{P})$ is a probability space. The following discussion makes no use of the fact that Ω is the unit cube and indeed, the value of isoperimetric theory lies precisely in the fact that $(\Omega, \mathcal{A}, \mathbb{P})$ can be any abstract probability space.

Given $A \subset \Omega^n$, Talagrand's isoperimetric theory provides estimates of the measure of the set of points in Ω^n which are within a specified distance of A. There are many ways to define a distance on Ω^n and perhaps the simplest is the Hamming distance $d_H(x,y) := |\{1 \leq i \leq n;\ x_i \neq y_i\}|$. The Hamming distance from a point $x \in \Omega^n$ to a set $A \subset \Omega^n$ is given by

$$d_H(x, A) := \inf_{y \in A} \sum_{i=1}^{n} I_{\{x_i \neq y_i\}}.$$

Rhee's isoperimetric inequality (4.4) is actually based on a simple isoperimetric inequality involving the Hamming distance d_H.

Talagrand [88] shows for every $\lambda > 0$ and every probability measure $\mathbb{P}$ on Ω that the Hamming distance satisfies the following exponential integrability condition:

$$\int_{\Omega^n} \exp(\lambda d_H(x, A)) d\mathbb{P}^n \leq \frac{1}{\mathbb{P}^n[A]} \exp(-n\lambda^2/4). \tag{4.24}$$

We will see shortly how to use this sort of condition to derive concentration estimates for the TSP. Talagrand's method of proof actually applies to all of the Hamming metrics

$$d_a(x,y) := \sum_{i=1}^{n} a_i I_{\{x_i \neq y_i\}}, \quad a = (a_1, ..., a_n) \in (\mathbb{R}^+)^n,$$

provided that the factor of n in the exponent of (4.24) is replaced by $\|a\|^2 = \sum_{i=1}^n a_i^2$. We let $d_a(x, A) = \inf_{y \in A} d_a(x,y)$.

By considering a still different and slightly more complicated method for measuring distance we will be able to improve upon the isoperimetric inequality (4.4) when the functional L is the TSP. Talagrand's *convex hull control* distance or simply *convex* distance is given by

$$d_c(x, A) := \sup_{|a|=1} d_a(x, A).$$

The convex distance uniformly controls the Hamming metrics $d_a,\ |a| = 1$. The present definition of d_c, while the simplest for our purposes, somewhat obscures its convexity properties.

There is a second way to formulate the definition of d_c:

$$d_c(x, A) := \min \left\{ t : \forall\ \{a_i\}\ \exists y \in A \text{ such that } \sum_{i=1}^{n} a_i I_{\{x_i \neq y_i\}} \leq t (\sum_{i=1}^{n} a_i^2)^{1/2} \right\}.$$

It is straightforward to see that these two definitions of d_c are equivalent. Notice that by letting $a_i = n^{-1/2}$ for all $1 \leq i \leq n$ we see that $d_c(x, A)$ is always at least as large as $n^{-1/2} d_H(x, A)$.

The convex distance satisfies an exponential square integrability condition, as described by the next theorem, also due to Talagrand [88]. The proof [88] of this theorem uses induction on n and while it is surprisingly short, we will not reproduce it here. Notice that this integrability condition is in general stronger than (4.24).

Theorem 4.12 *For every set $A \subset \Omega^n$ we have*

$$\int_{\Omega^n} \exp\left(\frac{1}{4} d_c^2(x, A)\right) d\mathbb{P}^n(x) \leq \frac{1}{\mathbb{P}^n[A]}.$$

It follows by the generalized Chebyshev inequality (1.2) with $f(x) = \exp(\frac{1}{4}x^2)$ that

$$\mathbb{P}^n[d_c(x, A) > t] \leq \frac{e^{-t^2/4}}{\mathbb{P}^n[A]}. \tag{4.25}$$

This is an abstract and general deviation bound for the convex distance. It is far from obvious that this general bound has relevance to the tail behavior of the TSP. We will now see shortly how to put this to good use. We first need one technical lemma which helps us compare the behavior of T on $x := \{x_1, ..., x_n\}$ and $y := \{y_1, ..., y_n\}$. For details concerning the proof see Steele [85].

Lemma 4.13 *There is a non-negative weight function $a(x) = (a_1(x), ..., a_n(x))$ such that for all $x := \{x_1, ..., x_n\}$ and $y = \{y_1, ..., y_n\}$ the TSP functional satisfies*

$$T(x_1, ..., x_n) \leq T(y_1, ..., y_n) + \sum_{i=1}^{n} a_i(x) I_{\{x_i \neq y_i\}}, \tag{4.26}$$

where $\|a(x)\|^2 \leq C^2$ uniformly in x.

Combining the above lemma and Theorem 4.12 we can now show that T does not deviate much from its median. We recall that U_i, $i \geq 1$, are i.i.d. random variables with the uniform distribution on $[0,1]^d$. We let M_n denote a median of $T(U_1, ..., U_n)$ and we follow Steele [85] closely. The next statement, which improves upon the martingale estimates of [74] and [75], shows that $T(U_1, ..., U_n) - M_n$ exhibits sub-Gaussian tail behavior. For $d = 2$ and t in the range $0 \leq t \leq Cn^{1/2}$, this tail behavior is sharp [69].

Corollary 4.14 *(concentration for the TSP) We have*

$$\mathbb{P}[|T(U_1, ..., U_n) - M_n| \geq t] \leq 4 \exp(-t^2/4C^2).$$

Proof Following Steele [85], we use Theorem 4.12 for the parameterized family of sets

$$A(b) := \{\{y_1, y_2, ..., y_n\} : T(y_1, ..., y_n) \le b\}.$$

We know by inequality (4.26) for all $\{x_1, x_2, ..., x_n\}$ and $\{y_1, y_2, ..., y_n\}$ that

$$T(\{x_1, ..., x_n\}) \le T(\{y_1, ..., y_n\}) + \sum_{i=1}^{n} a_i I_{\{x_i \neq y_i\}}.$$

Minimizing over $y \in A(b)$ gives

$$\begin{aligned} T(x_1, x_2, ..., x_n) &\le b + \min_{y \in A(b)} \sum_{i=1}^{n} a_i(x) I_{\{x_i \neq y_i\}} \\ &= b + \min_{y \in A(b)} \|a\|_2 \sum_{i=1}^{n} (a_i(x)/\|a\|_2) I_{\{x_i \neq y_i\}} \\ &\le b + C d_c(x, A(b)) \end{aligned}$$

by the definition of $d_c(x, A(b))$ and by hypothesis (4.13). We obtain for $x = \{U_1, U_2, ..., U_n\}$ that $d_c(x, A(b)) \ge C^{-1}(T(U_1, U_2, ..., U_n) - b)$.

By Theorem 4.12 we have

$$\mathbb{P}^n[d_c(x, A(b)) > t] \le \frac{1}{\mathbb{P}^n[A(b)]} \exp(-t^2/4),$$

and therefore if we write $T_n := T(U_1, ..., U_n)$ then

$$\mathbb{P}^n[T_n \ge b + Ct] \le \frac{1}{\mathbb{P}^n[A(b)]} \exp(-t^2/4).$$

Thus, letting $u = Ct$, we have

$$\mathbb{P}^n[T_n \le b]\mathbb{P}^n[T_n \ge b + u] \le \exp(-u^2/4C^2)$$

and by letting $b = M_n$ and then $b = M - n - u$ we complete the proof of Corollary 4.14.

□

4.9 Further applications of boundary functionals

(a) Large Deviation Principles. Letting $F = \{U_1, ..., U_n\}$, the near additivity condition (4.7) expresses the fact that the global tour length through F is roughly the sum of i.i.d. local tour lengths. Together with smoothness, this condition shows that if $X(n) = n^{1/d} T(U_1, ..., U_{N(n)})$, where $N(n)$ is the usual independent Poisson random variable with parameter n, then $X(n)$ satisfies the following Donsker-Varadhan large deviation principle. This principle quantifies the precise deviations of $X(n)/n$ and refines the strong law of large numbers expressed in Theorem 4.1. For $t \in \mathbb{R}$ we let

$$\Lambda(t) := \lim n^{-1} \log \mathbb{E}[\exp tX(n)]$$

be the logarithmic moment generating function for $X(n)$. The convex dual is

$$\Lambda^*(x) := \sup_{t\in\mathbb{R}}\{tx - \Lambda(t)\}.$$

Theorem 4.15 *(Large deviation principle for the TSP [79]) For all closed sets $F \subset \mathbb{R}$ we have*

$$\limsup_{n\to\infty} \log \mathbb{P}[n^{-1}X(n) \in F] \le - \inf_{x\in F} \Lambda^*(x)$$

and for all open sets $O \subset \mathbb{R}$ we have

$$\liminf_{n\to\infty} \log \mathbb{P}[n^{-1}X(n) \in O] \ge - \inf_{x\in O} \Lambda^*(x).$$

Moreover, Λ^* has its unique zero at $\beta(d)$.

The proof of Theorem 4.15 is long and involved. Complete details may be found in Seppäläinen and Yukich [79], which also gives an entropic characterization of Λ^*.

(b) Directed TSP Consider the random directed graph G_n whose vertices are independent and uniformly distributed random variables $U_1, ..., U_n$ on the unit square. For $1 \le i \le j \le n$, the orientation of the edge X_iX_j is selected at random, independently for each edge and independently of the $U_i,\ i \ge 1$. The edges in G_n are given a direction by flipping fair coins. The directed TSP involves finding the shortest directed path through the random vertex set. If $D(n) = D(U_1, ..., U_n)$ denotes the length of the shortest directed path through the sample $U_1, ..., U_n$, then Steele [82] shows that

$$\lim_{n\to\infty} ED(n)/n = \alpha$$

for some constant α. Talagrand [87] shows that this convergence can be improved to complete convergence. It is not clear whether boundary fuctionals can be used to obtain asymptotics over more general point sets as in (4.16).

(c) Power Weighted Edges. For all $p > 0$, consider the length $T^p(F)$ of the shortest tour through F with pth power weighted edges. Thus

$$T^p(F) := \min_T \sum_{e\in T} |e|^p,$$

where the minimum is over all tours T and where $|e|$ denotes the Euclidean edge length of the edge e. The method of boundary functionals [58], [97] shows that for all $0 < p < d$ we have the following generalization of (4.16):

$$\lim_{n\to\infty} \frac{T^p(X_1, ..., X_n)}{n^{(d-p)/d}} = \beta(d,p) \int_{[0,1]^d} f(x)^{(d-p)/d} dx \quad c.c. \tag{4.27}$$

where $\beta(d,p)$ is a constant depending only on d and p. For the case $p = d$, a more delicate use of boundary functionals [95] shows that if $U_i,\ i \ge 1$, are i.i.d. with the uniform distribution on $[0,1]^d$ then

$$\lim_{n\to\infty} T^d(U_1, ..., U_n) = C(d) \quad a.s.,$$

where $C(d)$ is some finite constant.

(d) Worst Case Tour Length. Let the largest possible length of a minimal tour with pth power weighted edges through n points in $[0,1]^d$ be denoted by

$$\tau^p(n) := \max_{F \subset [0,1]^d, |F|=n} T^p(F).$$

By considering boundary TSP functionals (with power weighted edges) it is particularly easy to show the asymptotics

$$\lim_{n \to \infty} \frac{\tau^p(n)}{n^{(d-p)/d}} = \gamma(d,p)$$

where $\gamma(d,p)$ is some positive constant. Steele and Snyder [86] were the first to prove these asymptotics, although they restricted attention to the case $p = 1$. Using boundary functionals, Yukich [97] treats the case $1 \leq p < d$ and Lee [59] treats the cases $0 < p < 1$ and $p \geq d$.

4.10 Open problems

Problem 13 *Theorem 4.15 and Talagrand's concentration inequality for the TSP provide circumstantial evidence that the TSP functional satisfies asymptotic normality in the following sense:*

$$\frac{T(U_1, ..., U_n) - \mathbb{E}T(U_1, ..., U_n)}{(VarT(U_1, ..., U_n))^{1/2}} \to N(0,1).$$

Proving or disproving the above central limit theorem remains a difficult open problem and would add to the central limit theorem for the length of the Euclidean minimal spanning tree over a random sample [6], [56], [57].

Problem 14 *Develop the a.s. limit theory for Christofides' heuristic. In particular, does the Christofides' heuristic satisfy the limit given by Theorem 4.10?*

The methods described here can possibly be modified to treat the Christofides' heuristic. This may involve defining a superadditive boundary heuristic.

Problem 15 *Develop the a.s. limit theory for the directed TSP: investigate whether the directed TSP satisfies the limit in Theorem 4.10.*

Problem 16 *Establish that the variance of the TSP converges, i.e., show that in dimension 2 we have $VarT(U_1, ..., U_n) \to C$, where C is some positive constant.*

Problem 17 *Find theoretical values for the limiting constants $\beta(d)$ and $\gamma(d,p)$. Perhaps the problem is simplified by using a metric on $\mathbb{R}^d$ other than the Euclidean metric. Rhee [70] shows that in high dimensions $\beta(d)$ is close to $(d/2\pi e)^{1/2}$.*

Problem 18 *Use the Aldous - Steele objective method [4] to show the a.s. and the* L^1 *convergence of* $T^p(U_1, ..., U_n)$ *when* p *equals the dimension* d.

This would provide a second way to identify the limiting constant $\beta(d, d)$ and it would complement the Aldous-Steele results [4] on the random Euclidean minimal spanning tree.

Problem 19 *The large deviation principle (Theorem 4.15) holds for a Poisson number of uniform random variables. Does this LDP hold for a fixed deterministic number of uniform random variables?*

Problem 20 *Generalize Theorem 4.15 by establishing a general large deviation principle for the random variables* $T(X_1, ..., X_n)$, *where* $X_i,\ i \geq 1$, *are i.i.d. with an arbitrary distribution on* $[0, 1]^d$. *Show that the rate function exhibits quadratic behavior.*

Problem 21 *Bi-partite matching. Given* $X_i,\ 1 \leq i < \infty$, *and* $Y_i,\ 1 \leq i < \infty$, *two independent sequences of random variables with the uniform distribution on* $[0, 1]^d$, *the bi-partite matching problem studies the behavior of*

$$T_n = \min_{\pi} \sum_{i=1}^{n} \|X_i - Y_{\pi(i)}\|,$$

where π *runs over all permutations of the integers* $1, 2, ..., n$. *When* $d \geq 3$ *subadditive methods ([28]) imply that* $\mathbb{E}T_n/n^{(d-1)/d} \to C$ *for some constant* C. *When* $d = 2$, *subadditive methods fail and it is known ([3], [90]) only that there are positive constants* C_1 *and* C_2 *such that*

$$C_1 \leq \mathbb{E}T_n/(n \log n)^{1/2} \leq C_2.$$

The logarithmic term in the denominator prevents us from applying the usual subadditive arguments to conclude that

$$\mathbb{E}T_n/(n \log n)^{1/2} \to C, \tag{4.28}$$

where C *is some positive constant. Proving or disproving (4.28) remains an intriguing unsolved problem.*

Problem 22 *Bi-partite TSP. Given* $X_i,\ 1 \leq i < \infty$, *and* $Y_i,\ 1 \leq i < \infty$, *two independent sequences of random variables with the uniform distribution on* $[0, 1]^2$, *the bi-partite TSP involves finding the length of the shortest tour through* $X_i,\ 1 \leq i \leq n$, *and* $Y_i, 1 \leq i \leq n$, *such that each* X *point is joined to two* Y *points and vice versa. If* T_n *denotes the length of the shortest such tour then clearly* T_n *is bounded below by the length of the bi-partite matching on the union of* $X_i,\ 1 \leq i \leq n$ *and* $Y_i,\ 1 \leq i \leq n$. *However, finding the asymptotics of* T_n *remains open.*

Problem 23 *Non-standard scaling. Theorem 4.10 tells us that whenever random variables* $X_1, X_2, ...$ *have an absolutely continuous part, then* $T(X_1, ..., X_n)$ *scales like* $n^{(d-1)/d}$. *This raises the following question: given an arbitrary increasing function* $f(n) = o(n^{(d-1)/d})$, *when do there exist i.i.d. random variables* $X_1, ..., X_n$ *such that*

$$\lim_{n\to\infty} T(X_1, ..., X_n)/f(n) = C,$$

where C *is a positive finite constant?*

References

[1] M. Ajtai, J. Komlós and E. Szemerédi *The first occurrence of Hamilton cycles in random graphs* Annals of Discrete Mathematics 27 (1985) 173-178.

[2] M. Akcoglu and U. Krengel, *Ergodic theorems for superadditive processes*, J. Reine Ang. Math. 323 (1981) 53-67.

[3] M. Ajtai, J. Komlós, and G. Tusnády, *On optimal matchings* Combinatorica 4 (1984) 259-264.

[4] D. Aldous and J. M. Steele, *Asymptotics for Euclidean minimal spanning trees on random points*, Prob. Theory and Related Fields 92 (1992) 247-258.

[5] K. Alexander, *Rates of convergence of means for distance-minimizing subadditive Euclidean functionals*, Ann. Appl. Prob. 4 (1994) 902-922.

[6] K. Alexander, *The RSW theorem for continuum percolation and the CLT for Euclidean minimal spanning trees*, Ann. Appl. Prob. 6 (1996) 466-494.

[7] N. Alon, J. H. Spencer, *The Probabilistic Method*, John Wiley and Sons, (1991).

[8] D. Angluin and L.G. Valiant, *Fast probabilistic algorithms for Hamilton circuits and matchings*, Journal of Computer and System Sciences 18 (1979) 155-193.

[9] J. Beardwood, J. H. Halton, and J. M. Hammersley, *The shortest path through many points*, Proc. Camb. Philos. Society 55 (1959) 299-327.

[10] E.A.Bender and E.R.Canfield, The asymptotic number of labelled graphs with given degree sequences, Journal of Combinatorial Theory, Series A 24 (1978) 296-307.

[11] B.Bollobás, *Random Graphs,* Academic Press, 1985.

[12] B. Bollobás. *The evolution of sparse graphs.* Graph Theory and Combinatorics. (Proc. Cambridge Combinatorics Conference in Honour of Paul Erdős (B. Bollobás; Ed)) Academic Press (1984) 35-57.

[13] B.Bollobás, *A probabilistic proof of an asymptotic formula for the number of labelled regular graphs*, European Journal on Combinatorics 1 (1980) 311-316.

[14] B.Bollobás, *Almost all regular graphs are Hamiltonian*, European Journal on Combinatorics 4, (1983) 97-106.

[15] B.Bollobás, *Complete matchings in random subgraphs of the cube*, Random Structures and Algorithms 1 (1990) 95-104.

[16] B.Bollobás, C.Cooper, T.I.Fenner and A.M.Frieze, *On Hamilton cycles in sparse random graphs with minimum degree at least k*, in A tribute to Paul Erdos, edited by A.Baker, B.B ollobas and A.Hajnal, (1991) 59-96.

[17] B.Bollobás, T.I.Fenner and A.M.Frieze, *An algorithm for finding Hamilton paths and cycles in random graphs*, Combinatorica 7 (1987) 327-341.

[18] B.Bollobás and A.M.Frieze, *On matchings and hamiltonian cycles in random graphs*, Annals of Discrete Mathematics 28 (1985) 23-46.

[19] A.Broder, A.M.Frieze and E.Shamir, *Finding hidden Hamilton cycles* Random Structures and Algorithms 5, (1994) 395-410.

[20] C. Cooper, *1-pancyclic Hamilton cycles in random graphs*, Random Structures and Algorithms 4 (1993) 469-472.

[21] C. Cooper and A.M.Frieze, *On the number of hamilton cycles in a random graph*, Journal of Graph Theory 13 (1989) 719-735.

[22] C. Cooper and A.M.Frieze, *Pancyclic random graphs*, in Proceedings of Random Graphs '87, Edited by M.Karonski, J.Jaworski and A.Rucinski, John Wiley and Sons, (1990) 29-39.

[23] C. Cooper and A.M.Frieze, *Multicoloured Hamilton cycles in random graphs: an anti-Ramsey threshold*, Electronic Journal of Combinatorics 2 (1995) R19.

[24] C. Cooper and A.M.Frieze, *Hamilton cycles in random graphs and directed graphs*, to appear.

[25] C.Cooper, A.M.Frieze and M.J.Molloy, *Hamilton cycles in random regular digraphs*, Combinatorics, Probability and Computing 3, (1994) 39-50.

[26] D. Coppersmith, P. Raghavan and M. Tompa, *Parallel Graph Algorithms That Are Efficient on Average*, Information and Computation 81 (1989) 318-333.

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

[28] V. Dobric and J. E. Yukich, *Exact asymptotics for transportation cost in high dimensions*, J. Theoretical Prob. (1995) 97-118.

[29] P. Erdős and A. Rényi, *On random graphs I*, Publ. Math. Debrecen 6 (1959) 290-297.

[30] P. Erdős and A. Rényi, *On the evolution of random graphs*, Publ. Math. Inst. Hungar. Acad. Sci. 5 (1960) 17-61.

[31] P. Erdős and A. Rényi, *On the strength of connectedness of a random graph*, Acta. Math. Acad. Sci. Hungar. 12 (1963) 261-267.

[32] P. Erdős and A. Rényi, *On random matrices*, Publ.Math. Hungar. Acad. Inst. Sci. 8 (1964) 455-461.

[33] P. Erdős and J.Spencer, *Evolution of the n-cube*, Comput. Math. Applics. 5 (1979) 33-39.

[34] T.I. Fenner and A.M.Frieze, *Hamiltonian cycles in random regular graphs*, Journal of Combinatorial Theory, Series B, (1984) 103-112.

[35] A.M. Frieze, *Partitioning random graphs into large cycles*, Discrete Mathematics 70 (1988) 149-158.

[36] A.M. Frieze, *On the exact solution of random travelling salesman problems with medium-sized integer costs*, SIAM Journal on Computing 16 (1987) 1052-1072.

[37] A.M. Frieze, *An algorithm for finding Hamilton cycles in random digraphs*, Journal of Algorithms 9 (1988) 181-204.

[38] A.M. Frieze, *On matchings and Hamilton cycles in random graphs*, Surveys in Combinatorics (1989), Proceedings of British Combinatorial Conference, 84-114.

[39] A.M.Frieze, *Finding Hamilton cycles in sparse random graphs*, Journal of Combinatorial Theory B 44, (1988) 230-250.

[40] A.M. Frieze, *Parallel algorithms for finding Hamilton cycles in random graphs*, Information Processing Letters 25 (1987) 111-117.

[41] A.M. Frieze, *Random regular graphs with non-constant degree*, unpublished manuscript.

[42] A.M. Frieze, M.R. Jerrum, M. Molloy, R.W. Robinson and N.C. Wormald, *Generating and counting Hamilton cycles in random regular graphs*, Journal of Algorithms 21 (1996) 176-198.

[43] 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.

[44] A.M. Frieze and C.McDiarmid, *Algorithmic theory of random graphs*, Random Structures and Algorithms 10 (1997) 5-42.

[45] A.M. Frieze and B.D.McKay, *Multicoloured trees in random graphs*, Random Structures and Algorithms 5 (1994) 45-56.

[46] A.M. Frieze and G.Sorkin, *A note on random assignment and traveling salesman problems*, in preparation.

[47] A.M. Frieze and S.Suen, *Counting Hamilton cycles in random directed graphs*, Random Structures and Algorithms 3 (1992) 235-242.

[48] M. Goemans and D. Bertsimas, *Probabilistic analysis of the Held-Karp relaxation for the Euclidean traveling salesman problem, Math. Oper. Research 16 (1991) 72-89.*

[49] Y. Gurevich and S. Shelah, *Expected computation time for Hamiltonian path problem*, SIAM Journal on Computing 16 (1987) 486-502.

[50] M.Held and R.M.Karp, *The traveling salesman problem and minimal spanning trees*, Operations Research 18 (1970) 1138-1162.

[51] M.R. Jerrum and A.J. Sinclair, *Approximating the permanent*, SIAM Journal on Computing 18 (1989) 1149-1178.

[52] R. M. Karp, *The probabilistic analysis of some combinatorial search algorithms*, Algorithms and Complexity: New Directions and Recent Results (J. F. Traub, ed.), Academic Press, New York (1976) 1 - 19.

[53] R. M. Karp, *Probabilistic analysis of partitioning algorithms for the traveling salesman problem in the plane*, Math. Oper. Research 2 (1977) 209-224.

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

[55] 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.

[56] H. Kesten and S. Lee, *The central limit theorem for weighted minimal spanning trees on random points*, Ann. Appl. Prob. 6 (1996) 495-527.

[57] S. Lee, *The central limit theorem for Euclidean minimal spanning trees I*, Ann. Appl. Prob. 6 (1997) 996-1020.

[58] S. Lee, *Asymptotics of power-weighted Euclidean functionals* (1998), preprint.

[59] S. Lee, *Worst case asymptotics of power-weighted Euclidean functionals* (1999), preprint.

[60] S. Lee, *Rate of convergence of power-weighted Euclidean minimal spanning trees* (1999), Stoch. Proc. Applic., to appear.

[61] J. Komlós and E. Szemerédi, *Limit distributions for the existence of Hamilton circuits in a random graph*, Discrete Mathematics 43 (1983) 55-63.

[62] T. Łuczak, *Cycles in random graphs*, Discrete Mathematics 98 (1991) 231-236.

[63] P.D. MacKenzie and Q.F. Stout, *Optimal Parallel Construction of Hamiltonian Cycles and Spanning Trees in Random Graphs*, Proceedings of the 5th ACM Symposium on Parallel Algorithms and Architectures (1993), pp. 224-229.

[64] C. McDiarmid, *Clutter percolation and random graphs*, Combinatorial Optimization, (V.J.Rayward-Smith, Ed.) Mathematical Programming Study 13 (1980) 17-25.

[65] L. Pósa, *Hamiltonian circuits in random graphs*, Discrete Mathematics 14 (1976) 359-364.

[66] C. Redmond *Boundary rooted graphs and Euclidean matching algorithms* Ph.D. thesis (1993) Dept. of Mathematics, Lehigh University.

[67] C. Redmond and J. E. Yukich, *Limit theorems and rates of convergence for subadditive Euclidean functionals*, Ann. Appl. Prob. 4 (1994) 1057-1073.

[68] C. Redmond and J. E. Yukich, *Asymptotics for Euclidean functionals with power weighted edges*, Stochastic Processes and Their Applications 61 (1996) 289-304.

[69] W. Rhee, *On the fluctuations of the stochastic traveling salesperson problem*, Math. Oper. Research 16 (1991) 482-489.

[70] W. Rhee, *On the traveling salesperson problem in many dimensions*, Random Structures and Algorithms 3 (1992) 227-233.

[71] W. Rhee, *On the stochastic Euclidean traveling salesperson problem for distributions with unbounded support*, Math. Oper. Research 18 (1993) 292-299.

[72] W. Rhee, *A matching problem and subadditive Euclidean functionals*, Ann. Appl. Prob. 3 (1993) 794-801.

[73] W. Rhee, *Boundary effects in the traveling salesperson problem,* Oper. Research Letters 16 (1994) 19-25.

[74] W. Rhee and M. Talagrand, *Martingale inequalities and NP-complete problems*, Math. Oper. Research 12 (1987) 177-181.

[75] W. Rhee and M. Talagrand, *A sharp deviation inequality for the stochastic traveling salesman problem*, Ann. Prob. 17 (1989) 1-8.

[76] W. Rhee and M. Talagrand, *On the long edges in the shortest tour through n random points*, Combinatorica 12 (1992) 323-330.

[77] R.W. Robinson and N.C. Wormald, *Almost all cubic graphs are Hamiltonian*, Random Structures and Algorithms 3 (1992) 117-126.

[78] R.W. Robinson and N.C. Wormald, *Almost all regular graphs are Hamiltonian*, Random Structures and Algorithms 5, (1994) 363-374.

[79] T. Seppäläinen and J. E. Yukich, *Large deviation principles for Euclidean functionals and other nearly additive processes*, (1999), preprint.

[80] E. Shamir, *How many random edges make a graph Hamiltonian?*, Combinatorica 3 (1983) 123-132.

[81] J. M. Steele, *Subadditive Euclidean functionals and non-linear growth in geometric probability*, Ann. Prob. 9 (1981) 365-376.

[82] J. M. Steele, *Probabilistic algorithm for the directed traveling salesman problem*, Math. Oper. Research 11 (1986) 343-350.

[83] J. M. Steele, *Growth rates of Euclidean minimal spanning trees with power weighted edges*, Ann. Prob. 16 (1988) 1767-1787.

[84] J. M. Steele, *Probabilistic and worst case analyses of classical problems of combinatorial optimization in Euclidean space*, Math. Oper. Research 15 (1990) 749-770.

[85] J. M. Steele, *Probability Theory and Combinatorial Optimization* (1997) SIAM.

[86] J. M. Steele and T. L. Snyder *Worst case growth rates of some classical problems of combinatorial optimization* Siam J. Computation 18 (1989) 278-287.

[87] M. Talagrand, *Complete convergence of the directed TSP*, Math. Oper. Research (1991) 881-887.

[88] M. Talagrand, *Concentration of measure and isoperimetric inequalities in product spaces*, Publications Mathématiques de l'I.H.E.S. 81 (1995) 73-205.

[89] M. Talagrand, *A new look at independence*, Ann. Prob. 24 (1996) 1-34.

[90] M. Talagrand and J. E. Yukich, *The integrability of the square exponential transportation cost*, Ann. Appl. Prob. 3 (1993) 1100-1111.

[91] A.Thomason, *A simple linear expected time algorithm for Hamilton cycles*, Discrete Mathematics 75 (1989) 373-379.

[92] D. Van Wieren and Q.F. Stout, *Random graph algorithms for the mesh with row and column subbuses*, Proceedings of the seconnd Workshop on Reconfigurable Architectures (1995), pp. 1-13.

[93] D.W. Walkup, *On the expected value of a random asignment problem*, SIAM J. Computing 8 (1979) 440-442.

[94] N.Wormald, *Models of random regular graphs*, Surveys in Combinatorics (1999) Proceedings of British Combinatorial Conference, 239-298.

[95] J. E. Yukich, *Asymptotics for the stochastic TSP with power weighted edges* Prob. Theory and Related Fields (1995) 203-220.

[96] J. E. Yukich, *Ergodic theorems for some classical optimization problems* Ann. Appl. Prob. (1996) 1006-1023.

[97] J. E. Yukich, *Probability Theory of Classical Euclidean Optimization Problems* Lecture Notes in Mathematics 1675 (1998) Springer Verlag.

This research was sponsored in part by National Science Foundation (NSF) grant no. CCR-0122581.