The Limit in the Mean Field Bipartite Travelling Salesman Problem

Johan Wästlund *

Department of Mathematics
Linköping University, S-581 83 Linköping, Sweden
jowas@mai.liu.se

April 17, 2006
Mathematics Subject Classification: Primary: 60C05, 90C27, 90C35

Abstract

The edges of the complete bipartite graph $K_{n,n}$ are assigned independent lengths from uniform distribution on the interval $[0, 1]$. Let L_n be the length of the minimum travelling salesman tour. We prove that as n tends to infinity, L_n converges in probability to a certain number, approximately 4.0831. This number is characterized as the area of the region

$$x, y \geq 0, \quad (1 + x/2) \cdot e^{-x} + (1 + y/2) \cdot e^{-y} \geq 1$$

in the x-y-plane.

1 Introduction

1.1 The travelling salesman problem

The travelling salesman problem, or TSP for short, is perhaps the most famous of all computational problems. For a set of n "cities", all pairwise

*This research was carried out in part while the author visited the Mittag-Leffler institute in spring 2005.

distances are known, and the task is to find the minimum length *tour*, that is, the shortest cycle that visits each city. The TSP is NP-hard [17], which indicates that there is little hope of finding an efficient algorithm that solves it completely. Work on search heuristics like branch-and-bound and local improvement has motivated the investigation of various stochastic models of the TSP.

1.2 The Euclidean model

A model that has been studied thoroughly is the so called *euclidean* or *random point* model. Here n points are chosen at random inside the unit square (or more generally, a bounded region of volume 1 in d-dimensional space). Since the distance between "neighbouring" points is of order $n^{-1/d}$, we expect the length L_n of the minimum tour to be proportinal to $n^{1-1/d}$. In fact it is known that for every $d \geq 2$,

$$\frac{L_n}{n^{1-1/d}} \xrightarrow{\text{p}} \beta(d),$$

as $n \to \infty$, where $\beta(d)$ is a constant depending only on d (and not on the shape of the region). This theorem goes back almost fifty years to the classic paper [6] by J. Beardwood, H. J. Halton and J. M. Hammersley.

Further results on algorithms and estimates of the rate of convergence have been obtained by Richard Karp [19], Michael Steele [39], and Wansoo Rhee and Michel Talagrand [37, 38]. However, there seems to be no hope of characterizing the limit constant $\beta(d)$ analytically, even in two dimensions.

1.3 Statistical mechanics and the mean field model

In the 1980's it was recognized that random models of optimization problems such as minimum matching, minimum spanning tree and the TSP have many features in common with the statistical mechanics of so called *disordered systems*.

S. Kirkpatrick and G. Toulouse [21] used methods from the theory of spin glasses to study the TSP in the so called *mean field* (or *random link*) model. Here the distances between pairs of points are taken as independent identically distributed random variables. This means that dependencies (such as the triangle inequality) arising from the geometry are eliminated. In [21], the distances are taken from uniform distribution on $[0, 1]$ for simplicity, although this would geometrically correspond to the trivial one-dimensional case. Kirkpatrick and Toulouse motivated their choice with the hope that this model would be analytically solvable: “This model is appealing for its simplicity and freedom from geometry. We hope that it may eventually prove, as has the S. K. model of spin glasses ... to be analytically tractable and provide a ’mean field’ limit of the statistical mechanics of a travelling salesman.”

Several other optimization problems had already been studied in the mean field model. In 1979, D. Walkup [42] showed that the expected cost of the minimum perfect matching in a complete bipartite graph with uniform $[0,1]$ edge costs is bounded as $n \to \infty$. In 1985 Alan Frieze [12] proved that the cost of the minimum spanning tree in the complete graph converges in probability to $\zeta(3)$.

Marc Mézard and Giorgio Parisi [27, 28] and later W. Krauth and Mézard [22] used the (non-rigorous) *replica* and *cavity* methods and arrived at an approximate value of the limit of L_n in the mean field model. The “theoretical” value (as opposed to values obtained by simulation) of 2.0415 obtained in [22] is consistent with the less precise results in [21, 28]. Allon Percus and Olivier Martin [35] give a relatively recent survey of these results, and of simulations supporting the cavity predictions.

For the minimum matching problem, Mézard and Parisi conjectured [26, 28, 29] that the limit should be $\pi^2/12$ for the complete graph and $\pi^2/6$ for the complete bipartite graph. These conjectures were proved by D. Aldous in 2001 [3, 4]. In these articles he introduced the so called *Poisson-weighted infinite tree* (PWIT), which can be regarded as a reformulation of the statistical physics viewpoint. In [4], he arrived at the recursive distributional equation (RDE):

$$X \stackrel{d}{=} \min_i (\xi_i - X_i), \tag{1}$$

where $\xi_1 \leq \xi_2 \leq \dots$ are the times of the events in a rate 1 Poisson process and $X_1, X_2, X_3, \dots$ are independent variables of the same distribution as X. The limit cost is obtained as

$$\int_0^\infty x \cdot Pr(X_1 + X_2 > x)\, dx, \tag{2}$$

where X_1 and X_2 are independent variables taken from the distribution given by (1). It was proved in [4] that the unique solution to (1) is the *logistic* distribution. In [4] Aldous conjectured that the limit cost of the random

TSP can be obtained in the same way. For the TSP, the corresponding equation is

$$X \stackrel{d}{=} \min_i[2](\xi_i - X_i), \tag{3}$$

where min[2] denotes second-smallest. The conjectured limit cost is again obtained by (2). Unfortunately this approach has so far not been made rigorous for the TSP, and it is not known whether there is a unique solution to (3) or whether the solution can be described explicitly.

1.4 The complete versus the bipartite graph

For the matching and spanning tree problems, it is known that the limit for the complete bipartite graph is twice the limit for the complete graph. This feature is believed to hold quite generally, but is not very well understood. A simple explanation is that in the complete graph, there are roughly twice as many edges to choose between. If we compare $K_{n,n}$ to K_{2n} (so that the solutions require the same or roughly the same number of edges), then the costs of the cheapest edges incident to a particular vertex are asymptotically a Poisson process of rate n in the former case, and of rate $2n$ in the latter.

It is not obvious why this should imply that the limit cost on the bipartite graph is exactly twice that of the complete graph, but strong evidence is provided from the local weak convergence to the PWIT [2, 3, 4].

If this holds also for the TSP, then the limit in the mean field bipartite TSP with uniform $[0, 1]$ costs is twice the number obtained by Krauth, Mézard and Parisi, that is, approximately 4.083. In the present article we obtain a rigorous proof of this limit.

1.5 Statement of our main result

We work in the bipartite mean field model. Suppose that the edges of the complete bipartite graph $K_{n,n}$ are assigned random independent lengths from uniform distribution on $[0, 1]$. For $n \geq 2$, let L_n denote the cost of the minimum travelling salesman tour, that is, the minimum sum of the edge lengths in a cycle that passes through each vertex exactly once. We prove that L_n converges in mean, and thereby in probability, to a certain number

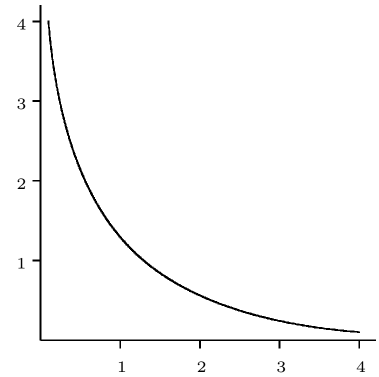

Figure 1: The curve $(1+\frac{x}{2})e^{-x} + (1+\frac{y}{2})e^{-y} = 1$.

that we denote τ. This number is defined by

$$\tau = \int_0^\infty y\,dx,$$

where $y(x)$ is the positive solution to the equation

$$\left(1+\frac{x}{2}\right)e^{-x} + \left(1+\frac{y}{2}\right)e^{-y} = 1,$$

see Figure 1. The number τ has been evaluated numerically to

$$\tau \approx 4.0830963728242648360909\,8032.$$

Our result, with an (admittedly quite weak) estimate of the rate of convergence, is

Theorem 1.1. *In the limit $n \to \infty$,*

$$E\,|L_n - \tau| = O\left(\frac{\log\log n}{(\log n)^{1/4}}\right).$$

We have not been able to extend our results to the complete graph, nor have we found a solution to the distributional equation (3). Hence there are at present three numbers which we conjecture to be equal, but for which we cannot yet prove equality of any two: The number $\tau/2$, the number given by the solution to (3), and the limit in the mean field TSP on the complete graph. Actually there does not seem to be any proof available that the two latter numbers are well-defined.

It has been recognized that it is more convenient to work with the exponential distribution rather than uniform distribution on $[0, 1]$. In this article too we work with exponential edge lengths. However, we want to show that the use of the exponential distribution is not just a way of simplifying the problem, but actually a method for solving it in the uniform setting as well. We have therefore chosen to formulate our main theorems for uniform distribution, although we prove them by first establishing the same results for exponential edge costs.

1.6 Outline of the method of proof

In 1998, G. Parisi conjectured that for the bipartite matching problem, if the edge costs are taken from exponential distribution, the expected minimum cost C_n of a perfect matching is given by the exact formula

$$E\left(C_n\right) = 1 + \frac{1}{4} + \frac{1}{9} + \cdots + \frac{1}{n^2}. \tag{4}$$

This formula was proved in 2003 independently by Chandra Nair, Mayank Sharma and Balaji Prabhakar [32] and by Svante Linusson and the author [24].

In their recent survey [1], D. Aldous and Antar Bandyopadhyay remarked on the Parisi formula (4) and its proofs in [24, 32] that "It seems unlikely that the applicability of exact methods extends far into the broad realm of problems amenable to asymptotic study".

We agree that it seems unlikely that further progress will be based on explicit formulas like (4) (and the related formula for the variance given in [47]). However, this does not mean that the scope of the so-called exact methods is limited to the assignment problem.

Although the method employed in this paper has little in common with the methods used in [24] and [32], they share a basic feature: They are based on "exact" (as opposed to asymptotic or approximate) statements, which are proved by induction. More precisely, these exact statements relate certain random optimization problems to another type of random process, the Buck-Chan-Robbins urn model. This process was introduced in [8] and further developed in [44, 45]. So far, it seems that the Buck-Chan-Robbins urn model can be defined only for a special class of optimization problems on bipartite graphs, although it is not yet clear what the scope is.

Our approach is based on finding a "friendly cousin" of the TSP. This friendly cousin is the Poisson(1) bipartite 2-factor problem (to be defined below). This problem is of the type that permits an "exact" solution in terms of the B-C-R urn process. At the same time, it is asymptotically close to the TSP. A 2-factor is a decomposition of the graph into vertex-disjoint cycles. Hence locally, a 2-factor looks like a tour.

Although the urn process is in principle much simpler than the original optimization problem, there does not seem to be a simple formula like (4) for the 2-factor problem. Instead we estimate the behaviour of the B-C-R urn process using simple probabilistic techniques. We prove that the expected cost of the minimum 2-factor is

$$\tau + O\left(\frac{(\log n)^{3/2}}{n^{1/2}}\right).$$

We conjecture that the true error term is $O(1/n)$, as it is for the assignment problem (this follows from (4)). We also obtain an upper bound on the variance of the cost of the minimum 2-factor:

$$\mathrm{var}(C_n) = O\left(\frac{(\log n)^{3/2}}{n^{1/2}}\right).$$

Here too we believe that a $O(1/n)$ bound can be achieved by a more detailed analysis of the urn process.

We then use methods from the paper [16] by Alan Frieze to turn the 2-factor into a tour by changing only a small (compared to n) number of edges. We show that with high probability, the number of cycles in the minimum 2-factor is small compared to n. We prove a $O(n \log\log n/\log n)$ bound, although the methods of [16] would give a bound of $O(n/\log n)$. We then use two different techniques to turn this collection of cycles into a tour at $o(1)$ expected cost. In a first phase, we use cheap edges to build a path connecting almost all vertices. In the second phase, we show that the remaining cycles can be absorbed into the path, and the path turned into a tour, using an operation called *rotation*.

Here all we really have to do is to show that the results of Frieze [16] hold also in the bipartite setting. In principle, it would suffice to go through Frieze's paper and just replace some of the constants in the calculations, since all the ideas work also in the bipartite setting. In practice, we have also been able to simplify some of the arguments by making use of the knowledge of the 2-factor problem gained through the connection to the urn process.

2 Random flow problems

In this section, we define and investigate a class of "friendly" random optimization problems that we call *flow problems*. We prove that the moments of the value of these problems can be expressed in terms of their corresponding B-C-R urn process.

2.1 Flows and flow problems

If V is a finite set, then by a *multiset* on V we mean a function $V \to \mathbb{N} \cup \{\infty\}$. If S is a multiset on V and $v \in V$, then $S(v)$ is the *multiplicity* of v in S. We allow for infinite multiplicities in order to make the algebra of multisets closed under arbitrary unions.

If S and T are multisets on V, then we say that S is a subset of T if $S(v) \leq T(v)$ for every $v \in V$. The *size* of S is the number

$$\sum_{v \in V} S(v),$$

in other words the number of elements of S counted with multiplicity. The union $S \cup T$ is given by $(S \cup T)(v) = \max(S(v), T(v))$, and the intersection $S \cap T$ by $(S \cap T)(v) = \min(S(v), T(v))$. The sum $S + T$ is the pointwise sum given by $(S + T)(v) = S(v) + T(v)$.

Let A be a multiset on V. We use language from matroid theory and say that a multiset S on V is *independent* with respect to A if $S \subseteq A$. This independence concept gives a special case of an *integral polymatroid* [11, 43]. For an introduction to matroid theory from the perspective of combinatorial optimization, we refer to [23].

The size of $S \cap A$, the maximal independent subset of S, is called the *rank* of S. Since we have allowed for infinite multiplicities, there is a unique largest superset of S that has the same rank as S. This set is called the *span* of S and is denoted $\sigma(S)$. The span can also be defined by

$$\sigma(S)(v) = \begin{cases} S(v), & \text{if } S(v) < A(v), \\ \infty, & \text{if } S(v) \geq A(v). \end{cases}$$

If $\sigma(S) = S$, then we say that S is a *subspace*. A minimal dependent set is called a *circuit*.

We let V and W be finite sets. We construct an infinite bipartite graph which we simply denote (V, W), with vertex sets V and W. For every pair $(v, w) \in V \times W$, there is a countably infinite sequence of edges

$$e_1(v, w), e_2(v, w), e_3(v, w), \dots$$

connecting v and w. We let $E = E(V, W)$ denote the set of all these edges.

If $F \subseteq E$ is a set of edges, then the *projections* F_V and F_W of F on V and W respectively are the multisets on V and W that count the number of edges of F incident to each vertex. Let A be a multiset on V and B a multiset on W. If F_V is independent with respect to A and F_W is independent with respect to B, then we say that F is a *flow* (with respect to (A, B)). A flow consisting of k edges is called a k-flow.

A *cost function* is a function $c : E \to \mathbb{R}$ satisfying, for every $v \in V$ and $w \in W$,

$$0 \leq c(e_1(v, w)) \leq c(e_2(v, w)) \leq c(e_3(v, w)) \leq \dots$$

If a cost function is given, and k is a nonnegative integer not larger than the size of either of A and B, then we can ask for the minimum cost k-flow, that is, the set $F \subseteq E$ that minimizes

$$c(F) = \sum_{e \in F} c(e)$$

under the constraint that F must be a k-flow. As we shall see next, this combinatorial optimization problem can be regarded as a special case of the *weighted matroid intersection problem* [23], for which polynomial time algorithms have been known since the 1970's.

2.2 Combinatorial properties of the flow problem

In this section we establish the necessary combinatorial properties of the flow problem. These properties are direct generalizations of the results of Section 3 of [47].

We assume that a cost function is given. We can regard the independence concept as defining two matroids on the set of edges. A set F of edges is independent with respect to A if $F_V \subseteq A$, and independent with respect to B if $F_W \subseteq B$. Let σ_A and σ_B be the closure operators with respect to these two matroids on the edge set.

Theorem 2.1. *Suppose that $F \subseteq E$ is an r-flow which is not of minimum cost. Then there is an r-flow F' of smaller cost than F which contains at most one edge outside $\sigma_A(F)$.*

Proof. We let G be a minimum cost r-flow, and let $H = F \triangle G$ (the symmetric difference of F and G). We orient the edges of H so that edges in F are directed from W to V, and edges in G are directed from V to W.

Now H can be partitioned into alternating paths and cycles in such a way that no path ends with an edge in G which belongs to $\sigma_B(F)$, that is, which goes to a vertex in W with degree in F equal to its multiplicity in B. Similarly, no path starts with an edge in G which belongs to $\sigma_A(F)$.

This partition is obtained as follows: At each vertex $v \in V$, if there are $A(v)$ edges in F incident to v, then the edges of $G - F$ incident to v are matched to the edges of $F - G$ incident to v (there are necessarily at least as many of the latter). Similarly, at each vertex $w \in W$, if there are $B(w)$ edges in F incident to w, then the edges of $G - F$ incident to w are matched to the edges of $F - G$ incident to w.

The components obtained in this way are either balanced in the sense that they have equally many elements from F as from G, or they have one more element from one of these sets than from the other. The components which are not balanced can be paired, so that H is partitioned into a number of balanced sets. Since G is a minimum flow and F is not, the elements of $G - F$ have smaller total cost than the elements of $F - G$. Hence one of these balanced sets P must be such that the cost of $F \triangle P$ is smaller than the cost of F. It is clear that $F \triangle P$ is a flow, and that it contains at most one element outside $\sigma_A(F)$. □

If $b \in W$ and $B(b) > 0$, then we define the *contraction* B/b of B by

$$B/b(w) = \begin{cases} B(w) - 1, & \text{if } w = b, \\ B(w) & \text{otherwise.} \end{cases}$$

In other words, a multiset S is independent with respect to B/b iff $S + b$ is independent with respect to B. The following is a corollary to Theorem

2.1. To simplify the statement, we assume that the edge costs are *generic* in the sense that no two different flows have the same cost. In the random model introduced in the next section, the edge costs are independent random variables of continuous distribution, which implies that genericity holds with probability 1.

Theorem 2.2. *Let $b \in W$ and suppose that $B(b) > 0$. Let F be the minimum r-flow with respect to $(A, B/b)$, and let G be the minimum $(r+1)$-flow with respect to (A, B). Then G contains exactly one edge outside $\sigma_A(F)$.*

Proof. In order to apply Theorem 2.1, we introduce an auxiliary element a. We let $V^\star = V \cup \{a\}$ and let $A^\star = A + a$.

We let the first edge (a, b) have nonnegative cost t, and let all other edges from a have infinite cost. If we put $t = 0$, then the minimum $(r+1)$-flow with respect to $(A^\star, B)$ consists of the edge (a, b) together with F, the minimum r-flow with respect to $(A, B/b)$. If we increase the value of t, then at some point the minimum $(r+1)$-flow with respect to $(A^\star, B)$ changes to G. If we let t have a value just above this threshold, so that the minimum $(r+1)$-flow in $(A^\star, B)$ is G, but no other $(r+1)$-flow with respect to $(A^\star, B)$ has smaller cost than $F + (a, b)$, then it follows from Theorem 2.1 that G contains exactly one edge outside $\sigma_A(F)$. $\square$

2.3 The random flow problem

We let λ and μ be discrete measures on V and W respectively. In other words, each element of V and W is given a nonnegative *weight*. A random cost function is chosen by letting the costs $c_i(v, w) = c(e_i(v, w))$ be the times at which the events occur in a Poisson process of rate $\lambda(v)\mu(w)$, and so that these Poisson processes for different pairs of vertices are all independent. We let

$$C_k(A, B)$$

denote the minimum cost of a k-flow with respect to (A, B).

Our aim is to obtain methods for computing the distribution of $C_k(A, B)$. The following is a key theorem that in principle summarizes what we need to know about the random flow problem. The characterization of the distribution of $C_k(A, B)$ in terms of the urn processes on V and W described below is then deduced by calculus.

As before, let F be the minimum r-flow with respect to $(A, B/b)$, where b is an element of nonzero multiplicity in B, and let G be the minimum $(r+1)$-flow with respect to (A, B). Moreover let v_G be the element of V incident to the unique edge of G which is not in $\sigma_A(F)$.

We make the following definition: If S is a multiset on V, then we let $S^\perp$ demote the set of $v \in V$ such that $S + v$ has greater rank than S.

Theorem 2.3 (The independence theorem). *If we condition on $\sigma(F_V)$ and the cost of F, then the cost of G is independent of v_G, and v_G is distributed on $F_V^\perp$ with probabilities proportional to the weights.*

Proof. We condition on (1) the costs of all edges in $\sigma_A(F)$, and (2) for each $w \in W$, the minimum cost of all edges to w which are not in $\sigma_A(F)$. By Theorem 2.1, we have thereby conditioned on the cost of G. The vertex $w_0 \in W$ to which the edge in $G - \sigma_A(F)$ is incident is also determined. It now suffices to prove that v_G is still distributed on the possible vertices in V according to the weights given by λ. The edges from $F_A^\perp$ to w_0 are a priori produced by Poisson processes of rates proportional to the weights of the vertices in $F_A^\perp$. We now disregard those of these edges that belong to F, and condition on the minimum of the remaining edges. It is a well-known property of the Poisson process that the first event occurs in a particular process with probability proportional to the rate, even conditioning on the time at which the first event occurs. □

2.4 The extension method and the normalized limit measure

We use a method that was introduced in [46], that we call the *extension method*. We extend the set V by introducing an auxiliary special element a in the same way as in the proof of Theorem 2.2. The multiset $A^\star$ on the ground set $V^\star = V \cup \{a\}$ is defined by letting $A^\star(a) = 1$, and $A^\star(v) = A(v)$ for $v \in V$.

The idea, introduced in [46], is to let the weight $\lambda(a)$ of the element a tend to zero. We let $\lambda(a) = h$. When $h \to 0$, the edges from a become very expensive and we essentially get the original problem back. Somewhat surprisingly, crucial information can be obtained by studying events involving the inclusion of an edge from a in the minimum flow. The probability of this is proportional to h, and hence tends to zero as $h \to 0$. What is interesting is the constant of proportionality. If ϕ is a random variable whose distribution depends on h, then we let

$$E^\star(\phi) = \lim_{h\to 0} \frac{1}{h} E(\phi).$$

We can informally regard $E^\star$ as a measure, the *normalized limit measure*. This measure is obtained by noting that the probability measure defined by the exponential distribution with rate h, scaled up by a factor $1/h$, converges to the Lebesgue measure on the positive real numbers as $h \to 0$.

Hence the normalized limit measure can be obtained as follows: We first define the measure space of cost functions on the edges from the auxiliary vertex a. This measure space is the set of all assignments of costs to these edges such that exactly one edge from a has nonnegative real cost, and the remaining edges from a have cost $+\infty$. For each $w \in W$, the cost functions for which the first edge $e_1(a, w)$ has finite cost are measured by $\mu(w)$ times Lebesgue measure on the positive reals. The measure space of all cost functions for the edges from a is the disjoint union of these spaces over all $w \in W$. The normalized limit measure, which we denote by $E^\star$, is the product of this space with the probability space of cost functions on the ordinary edges given by the independent Poisson processes as described earlier.

We use the notation $E^\star$ both for events (sets of cost functions) and random variables (functions defined in terms of the costs of the edges). However, we only use the normalized limit measure as an informal tool. $E^\star$ is defined as a limit, and when using the normalized limit measure, we have to make sure that we can interchange the limit $h \to 0$ with the integration with respect to the normalized limit measure.

2.5 A recursive formula

If T is a subspace with respect to A, then we let $I_k(T, A, B)$ be the indicator variable for the event that T is spanned by the projection on V of the minimum k-flow with respect to (A, B), in other words the event that there is a basis U of T such that for each $v \in V$, the minimum k-flow contains at least $U(v)$ edges from v.

In this section we prove the analogues of Theorems 4.1 and 4.2 of [47].

Theorem 2.4. *Let N be a positive integer, and let S be a rank $k-1$ subspace of A. For every $b \in W$, let I_b be the indicator variable for the event that the* *minimum k-flow with respect to $(A^\star, B)$ contains an edge (a, b) and that the projection on V of the remaining edges spans S. Then for every $b \in W$,*

$$\begin{aligned}&E\left(C_k(A,B)^N \cdot I_{k-1}(S,A,B/b)\right) = \\ &\quad E\left(C_{k-1}(A,B/b)^N \cdot I_{k-1}(S,A,B/b)\right) + \frac{N}{\mu(b)} E^\star\left(C_k(A^\star,B)^{N-1} \cdot I_b\right). \quad (5)\end{aligned}$$

Proof. We compute $N/\mu(b) \cdot E^\star\left(C_k(A^\star,B)^{N-1} \cdot I_b\right)$ by integrating over the cost, which we denote by t, of the first edge between a and b. We therefore condition on the costs of all other edges.

The density of t is he^{-ht}, and we therefore get the normalized limit by dividing by h and instead computing the integral with the density e^{-ht}. For every t this tends to 1 from below as $h \to 0$, and by the principle of dominated convergence, we can interchange the limits and compute the integral using the density 1. This is the same thing as using the normalized limit measure.

We have

$$\frac{d}{dt}\left(C_k(A^\star,B)^N \cdot I_{k-1}(S,A,B/b)\right) = N \cdot C_k(A^\star,B)^{N-1} \cdot I_b.$$

According to the normalized limit measure, since we are conditioning on the edge (a, b) being the one with finite cost, $E^\star$ is just $\mu(b)$ times Lebesgue measure. Hence the statement follows by the fundamental theorem of calculus, since putting $t = \infty$, we get

$$C_k(A^\star,B)^N \cdot I_{k-1}(S,A,B/b) = C_k(A,B)^N \cdot I_{k-1}(S,A,B/b),$$

while if we put $t = 0$, we get

$$C_k(A^\star,B)^N \cdot I_{k-1}(S,A,B/b) = C_{k-1}(A,B/b)^N \cdot I_{k-1}(S,A,B/b).$$

□

Theorem 2.5. *Let T be a rank k subspace of A. Then*

$$E\left(C_k(A,B)^N \cdot I_k(T,A,B)\right) = \sum_{\substack{S\subseteq T\\ \mathrm{rank}(S)=k-1}} \frac{\lambda(\{v\in V : \sigma_A(S+v)=T\})}{\lambda(S^\perp)}$$

$$\sum_{\substack{b\in W\\ B(b)>0}} \frac{\mu(b)}{\mu(\{w\in W : B(w)>0\})} E\left(C_{k-1}(A,B/b)^N \cdot I_{k-1}(S,A,B/b)\right)$$

$$+ \frac{N}{\mu(\{w\in W : B(w)>0\})} \sum_{\substack{S\subseteq T\\ \mathrm{rank}(S)=k-1}} \frac{\lambda(\{v\in V : \sigma_A(S+v)=T\})}{\lambda(S^\perp)}$$

$$\cdot E^\star\left(C_k(A^\star,B)^{N-1} \cdot I_k(S\cup a, A^\star, B)\right). \quad (6)$$

Proof. We multiply both sides of (5) by $\mu(b)$ and sum over all $b\in W$ such that $B(b)>0$. This way we obtain

$$\sum_{\substack{b\in W\\ \mathbf{B}(b)>0}} \mu(b)\cdot E\left(C_k(A,B)^N \cdot I_{k-1}(S,A,B/b)\right) =$$

$$\sum_{\substack{b\in W\\ B(b)>0}} \mu(b)\cdot E\left(C_{k-1}(A,B/b)^N \cdot I_{k-1}(S,A,B/b)\right)$$

$$+ N\cdot E^\star\left(C_k(A^\star,B)^{N-1} \cdot I_k(S+a,A^\star,B)\right). \quad (7)$$

We now use Theorem 2.3. Suppose that T is a rank k subspace of A, and that $b\in W$ and $B(b)>0$. If the projection on V of the minimum k-flow with respect to (A,B) spans T, then the projection of the minimum $(k-1)$-flow in $(A,B/b)$ must span a rank $k-1$ subspace of T. By summing over the possible subspaces, we obtain

$$E\left(C_k(A,B)^N \cdot I_k(T,A,B)\right)$$

$$= \sum_{\substack{S\subseteq T\\ \mathrm{rank}_A(S)=k-1}} \frac{\lambda(\{v\in V : \sigma_A(S+v)=T\})}{\lambda(S^\perp)}$$

$$\cdot E\left(C_k(A,B)^N \cdot I_{k-1}(S,A,B/b)\right). \quad (8)$$

Now we choose b randomly according to the weights, in other words, we multiply (8) by $\mu(b)/\mu(\{w' \in W : B(w) > 0\})$ and sum over all $b \in W$ such that $B(b) > 0$. This leaves the left hand side intact, and we get

$$E\left(C_k(A,B)^N \cdot I_k(T,A,B)\right) = \sum_{\substack{S \subseteq T \\ \mathrm{rank}_A(S) = k-1}} \frac{\lambda(\{v \in V : \sigma_A(S+v) = T\})}{\lambda(S^{\perp})}$$
$$\sum_{\substack{w \in W \\ B(w) > 0}} \frac{\mu(b)}{\mu(\{w \in W : B(w) > 0\})} \cdot E\left(C_k(A,B)^N \cdot I_{k-1}(S,A,B/b)\right). \quad (9)$$

Now we use equation (7) divided by $\mu(\{w \in W : B(w) > 0\})$ to rewrite the right hand side of (9). This establishes the theorem.

□

2.6 Interpretation in terms of the Buck-Chan-Robbins urn process

The Buck-Chan-Robbins urn process was introduced in [8] and further developed in [44, 45, 47]. Here we generalize this process to multisets. We still refer to this process as the B-C-R urn process, although the process we describe can no longer be interpreted as drawing balls from an urn, as in [8].

Let V be a set with a discrete measure λ as in the previous section. For each $v \in V$, let $0 < \xi(v,1) < \xi(v,2) < \dots$ be the times of the events in a Poisson process of rate $\lambda(v)$, and let all these processes be independent. We let $V(x)$ be the multiset of labels of the events that occur at time $\leq x$. We let W be another set with measure μ, and define a similar independent urn process on W, where $W(y)$ is the multiset of events occurring up to and including time y.

The "exact" theorems that we are going to prove all state that some quantity defined in terms of a random flow problem is equal to a corresponding quantity defined in terms of the B-C-R urn process. The most important theorem of this type is a straightforward generalization of the Buck-Chan-Robbins formula for the assignment problem [8, 44]: Let $R_k(A,B)$ be the region in the positive quadrant of the x-y-plane consisting of those points

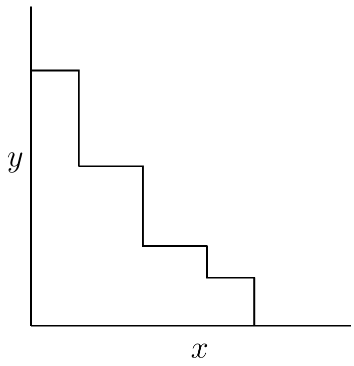

Figure 2: The typical shape of the region $R_k(A, B)$.

(x, y) for which

$$\text{rank}_A(V(x)) + \text{rank}_B(W(y)) < k.$$

Then the expected value of $C_k(A, B)$ is the same as the expected value of the area of $R_k(A, B)$, in analogy with the main theorem of [44]. We further show that the higher moments of $C_k(A, B)$ can also be characterized in terms of the distribution of $R_k(A, B)$, thereby generalizing the main results of [47].

We introduce an *extended urn process*. In the *N-th extension* of the urn process on V and W there are, in addition to the ordinary urn processes on V and W, N extra points $(x_1, y_1), \ldots, (x_N, y_N)$ in the positive quadrant of the x-y plane. These points are "chosen" according to Lebesgue measure on the positive real numbers, and therefore cannot be treated as random variables. Again we use $E^\star$ to denote the measure obtained by combining the probability measure on the urn process with Lebesgue measure on the extra points. $E^\star$ can be interpreted as the expected value (with respect to the ordinary urn process) of the Lebesgue measure in $2N$ dimensions of the set of points $x_1, \ldots, x_N, y_1, \ldots, y_N$ belonging to a particular event.

We define a rank function r on the nonnegative real numbers (depending on the outcome of the extended urn process) by

$$r(x) = \text{rank}_A(V(x)) + \#\{i : x_i \leq x\}.$$

Similarly we let

$$s(y) = \text{rank}_B(W(y)) + \#\{i : y_i \leq y\}.$$

Our main theorem on random flow problems is the following:

Theorem 2.6.

$$E\left(C_k(A,B)^N\right) = E^\star\left(r(x_i) + s(y_i) \leq k + N \textit{ for } 1 \leq i \leq N\right). \tag{10}$$

When $N = 1$, the right hand side of (10) is the expected area of $R_k(A, B)$. For larger N, the condition $r(x_i) + s(y_i) \leq k + N$ for $1 \leq i \leq N$ implies that the points (x_i, y_i) all lie in $R_k(A, B)$. Therefore we have

Corollary 2.7.

$$E\left(C_k(A,B)^N\right) \leq E\left(\text{area}(R_k(A,B))^N\right),$$

with equality if $N = 1$.

This corollary will be sufficient for our applications to the 2-factor problem and TSP in Sections 3 and 4. We prove Theorem 2.6 by proving the following more precise form. We let V_k be the span of the first rank k multiset obtained in the urn process on V, and define W_k similarly.

Theorem 2.8. *Let T be a rank k subspace with respect to A. Then*

$$\begin{aligned} &E\left(C_k(A,B)^N \cdot I_k(T,A,B)\right) \\ &\qquad = E^\star\left(V_k = T \textit{ and } r(x_i) + s(y_i) \leq k + N \textit{ for } 1 \leq i \leq N\right). \end{aligned} \tag{11}$$

Proof. We use (6) (Theorem 2.5) together with induction on both k and N. Notice that (11) holds trivially when $k = 0$. Notice also that when $N = 0$, the second term of the right hand side of (6) vanishes, and that therefore only induction on k is needed.

Suppose therefore that (11) holds whenever k or N is replaced by a smaller number. Then the right hand side of (6) can be rewritten in terms of the urn process. We are going to show that the result is equal to the right hand side of (11).

We therefore split the "event" $V_k = T$ and $r(x_i) + s(y_i) \leq k + N$ for $i = 1, \dots, N$ into two cases. Let $S = V_{k-1}$. Let t be the time at which the extended urn process on W reaches rank 1, in other words, t is minimal such that $s(t) = 1$. We condition on the event occurring at time t.

Case 1. Let $b \in W$ and suppose that $t = \xi(b, 1)$. In other words, we assume that b is the first element such that $B(b) > 0$ which occurs in the urn process on W, and that the time at which b is first drawn is smaller than all the numbers $y_1, \dots, y_N$.

We couple to another extended urn process by letting

$$\xi'(b, i) = \xi(b, i + 1) - t,$$

and for all $w \neq b$ in W,

$$\xi'(w, i) = \xi(w, i) - t.$$

Moreover, let

$$y'_i = y_i - t.$$

We let $x'_i = x_i$ and for $v \in V$, $\xi'(v, i) = \xi(v, i)$. Let r' and s' be the rank functions with respect to A and B/b in the new extended urn process. Then $r' = r$, and $s'(y'_i) = s(y_i) - 1$. Hence for $1 \leq i \leq N$, $r'(x'_i) + s'(y'_i) \leq k - 1 + N$ if and only if $r(x_i) + s(y_i) \leq k + N$.

If we condition on $V_{k-1} = S$, then

$$Pr(V_k = T) = \frac{\lambda(\{v \in V : \sigma_A(S + v) = T\})}{\lambda(S^\perp)}.$$

By the induction hypothesis, $E^\star$(Case 1) is equal to the first term of the right hand side of (6).

Case 2. Suppose that $y_N = t$ (the other cases $y_i = t$ are of course identical) . We are going to couple Case 2 to the $(N - 1)$th extended urn process. The coupling is done in essentially the same way as in Case 1. Let

$$\xi'(b, i) = \xi(b, i) - t,$$

and for $1 \leq i \leq N - 1$ let

$$y'_i = y_i - t.$$

In the limit $h \to 0$, $\xi(a, 1)$ is measured by h times Lebesgue measure on the positive real numbers. Hence if we let

$$\xi'(a, 1) = x_N,$$

we obtain a coupling which is valid as $h \to 0$.

Again, condition on $V_{k-1} = S$,

$$Pr(V_k = T) = \frac{\lambda(\{v \in V : \sigma_A(S + v) = T\})}{\lambda(S^\perp)}.$$

Moreover,

$$\frac{1}{\mu(\{w \in W : B(w) > 0\})}$$

is the measure of the event that y_N is smaller than $\xi(v, 1)$ for every v such that $B(v) > 0$. By the induction hypothesis, $E^\star(\text{Case 2})$ is equal to the second term of the right hand side of (6).

Hence $E\left(C_k(A, B)^N \cdot I_k(T, A, B)\right) = E^\star(\text{Case 1}) + N \cdot E^\star(\text{Case 2})$. This completes the proof. □

Theorem 2.8 generalizes automatically to subspaces of rank $< k$.

Theorem 2.9. *Let T be a subspace with respect to A. Then*

$$E\left(C_k(A, B)^N \cdot I_k(T, A, B)\right) \\ = E^\star\left(T \subseteq V_k \text{ and } r(x_i) + s(y_i) \leq k + N \text{ for } 1 \leq i \leq N\right). \quad (12)$$

In particular if T is empty then this is Theorem 2.6. We remark that for the minimum matching problem (or *assignment problem*), the Parisi formula (4) can easily be derived from the case $N = 1$ of this theorem. This is done by taking V and W of size n, and letting A and B be the multisets where each element has multiplicity 1. By taking $N = 2$, one obtains, after some calculations, an exact formula for the variance. This was done in [47].

3 Applications to the random 2-factor problem

In this section, we apply the results of Section 2 to the 2-factor problem. We let $\tilde{L}_n$ denote the cost of the minimum 2-factor in $K_{n,n}$. We prove bounds on the expectation and variance of $\tilde{L}_n$.

First, however, it seems appropriate to show how the results of the previous section can be used to calculate $E\left(\tilde{L}_n\right)$ exactly for small n, although this is not necessary for the rest of the paper.

3.1 $E\left(\tilde{L}_n\right)$ for small n

We now apply Theorem 2.6 to the minimum 2-factor problem. We let V and W be sets of size n, and A and B are the multisets where each element has multiplicity 2. The measures λ and μ are taken as the counting measures, that is, each element has weight 1.

In this setting, a minimum $2n$-flow is the same thing as a minimum 2-factor. We compute $E\left(\tilde{L}_n\right)$ by summing over all the states of the two-dimensional urn process. For each state, we compute the expected amount of two-dimensional time spent in this state by integrating the probability of being in the particular state over x and y in the positive quadrant. Since the processes in V and W are independent, this integral is the product of two one-dimensional integrals. We obtain the following formula:

$$\sum_{i_1+2i_2+j_1+2j_2<2n} I(n, i_1, i_2) \cdot I(n, j_1, j_2),$$

where I is defined by

$$I(n, i_1, i_2) = \binom{n}{i_1, i_2, n - i_1 - i_2} \int_0^\infty (e^{-x})^{(n-i_1-i_2)} (xe^{-x})^{i_1} (1 - e^{-x} - xe^{-x})^{i_2}\, dx.$$

Here i_1 is the number of vertices for which exactly one event has occurred in the urn process, and i_2 is the number of vertices for which at least two events have occurred. For given values of n, i_1 and i_2, the integral is a rational number that can easily be calculated using a computer program like Maple. We obtain the following expected values for the 2-factor problem for the first few values of n.

$$E\left(\tilde{L}_1\right) = 3$$

$$E\left(\tilde{L}_2\right) = \frac{7}{2}$$

$$E\left(\tilde{L}_3\right) = \frac{3581}{972} \approx 3.684$$

$$E\left(\tilde{L}_4\right) = \frac{626981}{165888} \approx 3.779$$

$$E\left(\tilde{L}_5\right) = \frac{12953341271}{3375000000} \approx 3.838$$

$$E\left(\tilde{L}_6\right) = \frac{1526452234799}{393660000000} \approx 3.877$$

We can (non-rigorously) determine the limit shape as $n \to \infty$ of the region $R_{2n}(A, B)$ as follows: The expected value of $\text{rank}_A(V(x))$ and $\text{rank}_B(W(y))$ is n times the expected rank produced by a particular element. Consider an arbitrary element $v \in V$. With probability e^{-x}, no event has occurred in the urn process for v. With probability xe^{-x}, exactly one event has occurred, and in the remaining cases, that is, with probability $1 - e^{-x} - xe^{-x}$, at least two events have occurred. For large n, the point (x, y) will lie inside $R_{2n}(A, B)$ with high probability if the sum of the average ranks in V and W is smaller than 2, and will be outside $R_{2n}(A, B)$ with high probability if the sum of the average ranks is larger than 2. The limit shape of $R_{2n}(A, B)$ is therefore given by the equation

$$xe^{-x} + 2\left(1 - e^{-x} - xe^{-x}\right) + ye^{-y} + 2\left(1 - e^{-y} - ye^{-y}\right) = 2,$$

which simplifies to

$$\left(1 + \frac{x}{2}\right)e^{-x} + \left(1 + \frac{y}{2}\right)e^{-y} = 1. \tag{13}$$

It is therefore reasonable to expect that $E\left(\tilde{L}_n\right)$ converges to the number τ. In the next section we obtain a rigorous proof of this, with an explicit error term. For the assignment problem, the same reasoning would show that the limit shape is given by the curve $e^{-x} + e^{-y} = 1$. In this case, one can solve explicitly for y, obtaining $y = -\log(1 - e^{-x})$. Moreover, the area of the limit region can be calculated as

$$\int_0^\infty -\log(1 - e^{-x})\, dx = \frac{\pi^2}{6},$$

in agreement with the theorem of Aldous [4], although a rigorous proof would require an argument that the area of the limit shape is the same as the limit of the expected area.

3.2 A rough bound on the size of $R_{2n}(n)$

We let $R_k(n)$ denote the region $R_k(A, B)$ where V, W, A and B are as in the previous section. We know from the results of Section 2 that the expected value of the cost of the minimum $(2n)$-flow is the same as the expected area of the region $R_{2n}(n)$, and we therefore wish to obtain bounds on the area of this region. We are going to show that with high probability, the boundary of the region $R_{2n}(n)$ (except for the coordinate axes) lies entirely within a small distance of the curve $(1 + x/2) \cdot e^{-x} + (1 + y/2) \cdot e^{-y} = 1$. This allows us to bound the mean and variance of the length $\tilde{L}_n$ of the minimum 2-factor. The bounds we obtain are certainly not the best possible, but they are obtained with quite simple and general methods, and they are good enough to establish convergence in probability of $\tilde{L}_n$ to the number τ defined in Section 1.5.

The time until completion of the (one-dimensional) urn process (that is, until it reaches rank $2n$), is roughly $\log n$. Therefore with high probability, no point of $R_{2n}(n)$ will be outside the square $[0, 2\log n] \times [0, 2\log n]$. Our first aim is to state and prove this with a precise error term. The square $[0, 2\log n] \times [0, 2\log n]$ will be called the *basic square.*

Lemma 3.1. *The expected area of the part of $R_{2n}(n)$ which is outside the basic square is*

$$O\left(\frac{(\log n)^2}{n}\right).$$

Proof. The probability that at most one event has occurred up to time x in a Poisson(1) process is $(1+x)e^{-x}$. At time x, the expected number of vertices of V for which at most one event has occurred is $n(1 + x)e^{-x}$. Therefore

$$Pr((x, y) \in R_{2n}(n)) \leq Pr((x, 0) \in R_{2n}(n)) \leq n(1 + x)e^{-x}. \tag{14}$$

We also have

$$Pr((x, y) \in R_{2n}(n)) \leq Pr((x, 0) \in R_{2n}(n)) \cdot Pr((0, y) \in R_{2n}(n)) \\ \leq n^2(1 + x)(1 + y) \cdot e^{-x-y}. \tag{15}$$

Therefore we get an upper bound on the expected area of the part of $R_{2n}(n)$ which is outside a square of side T by

$$2T \cdot I + I^2,$$

where

$$I = \int_T^\infty n(1 + x)e^{-x}\, dx = n(2 + T)e^{-x}.$$

Hence

$$2T \cdot I + I^2 = 2nT(2+T)e^{-T} + n^2(2+T)^2 e^{-2T}.$$

If we put $T = 2\log n$, this is

$$O\left(\frac{(\log n)^2}{n}\right).$$

□

3.3 Bounds on the shape of $R_{2n}(n)$

We now establish more precise bounds on the shape of $R_{2n}(n)$ valid inside the basic square. We use the following Chernoff type bound, which is established using routine techniques:

Lemma 3.2. *Supppose that $X = X_1 + \cdots + X_n$ is a sum of n independent variables that take the values 0 or 1. Then*

$$Pr(|X - E(X)| \geq \delta) \leq 2e^{-\delta^2/2n}.$$

Proof. To simplify, we let $Y_i = X_i - E(X_i)$ and $Y = Y_1 + \cdots + Y_n = X - E(X)$. Each Y_i takes the value $-q$ with probability p, and the value p with probability q, for some nonnegative numbers p and q with $p + q = 1$. For nonnegative t, it is easily established that

$$E\left(e^{tY_i}\right) = pe^{-qt} + qe^{pt} \leq e^{t^2/2}.$$

We obtain the following Chernoff type bound for $\delta > 0$:

$$Pr(Y \geq \delta) = Pr(e^{tY} \geq e^{t\delta}) \leq \frac{E\left(e^{tY}\right)}{e^{t\delta}} \leq \exp(nt^2/2 - t\delta).$$

Putting $t = \delta/n$ to optimize the inequality, we obtain

$$Pr(Y \geq \delta) \leq \exp\left(-\frac{\delta^2}{2n}\right).$$

By symmetry, we obtain the same bound for the probability that $Y \leq -\delta$. □

We now turn to the urn process. For $x \geq 0$, we let $\theta(x) = \text{rank}_A(V(x))$. We can write

$$\theta(x) = \theta_1(x) + \theta_2(x),$$

where $\theta_i(x)$ is the number of vertices in V for which the urn process has produced at least i events up to time x. Each of $\theta_1(x)$ and $\theta_2(x)$ is a sum of n independent identically distributed 0-1 variables. By the Chernoff bound,

$$Pr\left(|\theta_i(x) - E(\theta_i(x))| \geq 2(n \log n)^{1/2}\right) \leq 2n^{-2}.$$

We are going to show that with high probability, $\theta(x)$ does not deviate from its expected value by more than $2(n \log n)^{1/2}$ for any x. The following lemma gives bounds on the derivative of $E(\theta_i(x))$ for $i = 1, 2$.

Lemma 3.3. *If x is a nonnegative real number, then*

$$0 \leq \frac{d}{dx} E(\theta_1(x)) \leq n,$$

and

$$0 \leq \frac{d}{dx} E(\theta_2(x)) \leq e^{-1} n.$$

Proof. We have

$$\frac{1}{n} \cdot \frac{d}{dx} E(\theta_1(x)) = 1 - e^{-x}$$

and

$$\frac{1}{n} \cdot \frac{d}{dx} E(\theta_2(x)) = 1 - e^{-x} - xe^{-x}.$$

The stated bounds are now obtained by elementary calculus. □

If $\theta(x)$ deviates more than $8(n \log n)^{1/2}$ from its mean, then either $\theta_1(x)$ or $\theta_2(x)$ must deviate at least $4(n \log n)^{1/2}$ from their mean. This in turn implies, by the upper bounds on the derivatives, that one of them must deviate by at least $2(n \log n)^{1/2}$ on an interval of length $2n^{-1/2}(\log n)^{1/2}$.

The probability that this happens somewhere in the interval $[0, 2 \log n]$ is at most

$$2n^{-2} \cdot \frac{2 \log n}{2n^{-1/2}(\log n)^{1/2}} = 2(\log n)^{1/2} n^{-3/2}.$$

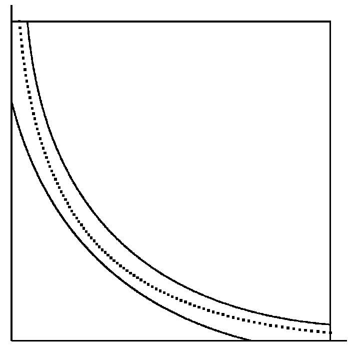

Figure 3: The basic square and the critical region. The dashed line is the curve $(1 + x/2)e^{-x} + (1 + y/2)e^{-y} = 1$.

The same thing holds for the urn process on W. We let $F(x, y)$ denote the normalized average rank of the two-dimensional urn process at the point (x, y) in the time plane. In other words,

$$F(x, y) = \frac{E(\theta(x)) + E(\theta(y))}{n} = 4 - 2e^{-x} - xe^{-x} - 2e^{-y} - ye^{-y}.$$

Notice that $F(x, y)$ does not depend on n. We also let $F(x) = 2-2e^{-x}-xe^{-x}$ so that $F(x, y) = F(x) + F(y)$.

We now obtain a high probability bound on the shape of $R_{2n}(n)$. In the following, we say that a statement holds with failure probability p if it holds with probability at least $1 - p$.

Theorem 3.4. *With failure probability*

$$O\left(\frac{(\log n)^{1/2}}{n^{3/2}}\right),$$

the part of the boundary of the region $R_{2n}(n)$ *which lies inside the basic square lies entirely within the region*

$$2 - \frac{8(\log n)^{1/2}}{n^{1/2}} \leq F(x, y) \leq 2 + \frac{8(\log n)^{1/2}}{n^{1/2}}.$$

The region $2 - 8n^{-1/2}(\log n)^{1/2} \leq F(x, y) \leq 2 + 8n^{-1/2}(\log n)^{1/2}$, $0 \leq x, y \leq 2\log n$ will be called the *critical region.* We now estimate the area of this region.

Lemma 3.5. *For $x \leq 1$,*

$$\frac{d}{dx}F(x, y) \geq 2e^{-1}.$$

Proof. This is established by elementary calculus. □

It now follows that the critical region has area bounded by

$$\frac{2}{2e^{-1}} \cdot 2\log n \cdot \frac{16(\log n)^{1/2}}{n^{1/2}} = O\left(\frac{(\log n)^{3/2}}{n^{1/2}}\right).$$

It remains to bound the area of the region $F(x, y) \leq 2$ which is outside the basic square. By a rough estimate, for large x we have $F(x) = 2 - 2e^{-x} - xe^{-x} \geq 2 - e^{-x/2}$, and for small y we have $F(y) \geq y/2$. The area of the limit region which lies outside the square of side $2\log n$ is therefore bounded by

$$\int_{2\log n}^{\infty} 2e^{-x/2}\, dx = \frac{4}{n}.$$

In the following estimate, this can be neglected, since it is much smaller than the error term that comes from the area of the critical region.

Theorem 3.6. *The expected value $E\left(\tilde{L}_n\right)$ of the minimum 2-factor in the Poisson(1) weighted complete bipartite graph $K_{n,n}$ satisfies*

$$E\left(\tilde{L}_n\right) = \tau + O\left(\frac{(\log n)^{3/2}}{n^{1/2}}\right).$$

Proof. The part of the region $R_{2n}(n)$ which lies outside the basic square has expected area

$$O\left(\frac{(\log n)^2}{n}\right),$$

which is negligible compared to the stated error term. The part of $R_{2n}(n)$ which lies inside the basic square has its boundary completely within the critical region with high probability. In the cases of failure, which have probability

$$O\left(\frac{(\log n)^{1/2}}{n^{3/2}}\right),$$

there is an error of at most the area of the basic square, that is, $4(\log n)^2$.

□

3.4 An upper bound on the variance of $\tilde{L}_n$

It follows from the results of Section 2 that $E\left(\tilde{L}_n^2\right) \leq E\left(\text{area}(R_{2n}(n))^2\right)$. We therefore want to bound the square of the area of $R_{2n}(n)$. With failure probability $O(\log n/n)$, the urn processes on V and W are both completed within time $2\log n$, which implies that $R_{2n}(n)$ lies entirely in the basic square. With even smaller failure probability, the boundary lies entirely in the critical region. If this holds, then the square of the area is

$$\tau^2 + O\left(\frac{(\log n)^{3/2}}{n^{1/2}}\right).$$

In the cases of failure, we bound the area by a rectangle each of whose sides is $2\log n$ plus the time it takes until (after time $2\log n$) at least one event has occurred in each of the n processes, plus the time until (after this) yet another event has occurred in each of the n processes.

The time it takes until at least one event has occurred in each of n rate 1 Poisson processes is the sum of n independent exponential variables of rates $1, 2, \ldots, n$. Let Z be such a variable. Then

$$\text{var}(Z) = 1 + \frac{1}{4} + \frac{1}{9} + \cdots + \frac{1}{n^2} = O(1).$$

Hence $E(Z) = \log n + O(1)$ and $E(Z^2) = O((\log n)^2)$. The area of the rectangle that we use to bound $R_{2n}(n)$ is

$$U = (2\log n + Z_1 + Z_2)(2\log n + Z_3 + Z_4),$$

where $Z_1, \ldots, Z_4$ are independent variables of this distribution. By independence we have

$$\begin{aligned} E\left(U^2\right) &= E\left((2\log n + Z_1 + Z_2)^2\right) \cdot E\left((2\log n + Z_1 + Z_2)^2\right) \\ &= O((\log n)^2) \cdot O((\log n)^2) = O((\log n)^4). \quad (16) \end{aligned}$$

Hence the cases of failure contribute only

$$O\left(\frac{(\log n)^5}{n}\right)$$

to $E\left(\text{area}(R_{2n}(n))^2\right)$. We therefore conclude:

Theorem 3.7.

$$\text{var}(\tilde{L}_n) = O\left(\frac{(\log n)^{3/2}}{n^{1/2}}\right).$$

We expect that a more detailed analysis will show that the variance is actually of order $1/n$, as has been proved for the assignment problem [47].

Corollary 3.8. $\tilde{L}_n$ converges in probability to the number τ.

3.5 Bounds on $C_k(n)$ for $k < 2n$

We let $C_k = C_k(n)$ denote the cost of the minimum k-flow, that is, the minimum set of k edges of which no three share a vertex. Here we show that if k is only slightly smaller than $2n$, then $C_k(n)$ is rarely much smaller than $\tilde{L}_n$. The following lemma gives a bound on the expected increment in cost of the minimum k-flow as the value of k increases by 1.

Lemma 3.9. *For every $k \leq 2n$,*

$$E\left(C_k\right) - E\left(C_{k-1}\right) \leq \frac{4}{n+1}\left(1 + \frac{1}{2} + \cdots + \frac{1}{n}\right) = O\left(\frac{\log n}{n}\right).$$

Proof. $E\left(C_k\right) - E\left(C_{k-1}\right)$ is the expected area of the diagonal of rectangles in $R_k(n) - R_{k-1}(n)$. It is clear that $k = 2n$ is the worst case, so that it is sufficient to consider this case. We get an upper bound on the expected time from the point where the urn process reaches rank i to the point where it reaches rank $i+1$ by conditioning on the following event: The first two events in the urn process occur for the same vertex, then the third and fourth events occur for another vertex, and so on. It follows that the expected area of $R_{2n}(n) - R_{2n-1}(n)$ is upper bounded by

$$\frac{1}{1\cdot n} + \frac{1}{1\cdot n} + \frac{1}{2\cdot(n-1)} + \frac{1}{2\cdot(n-1)} + \cdots + \frac{1}{n\cdot 1} = \frac{4}{n+1}\left(1 + \frac{1}{2} + \cdots + \frac{1}{n}\right). \quad (17)$$

□

This means that if we take k in the interval

$$2n - (n \log n)^{1/2} \le k \le 2n,$$

then the bounds on the mean and variance of C_k given in Theorems 3.6 and 3.7 still hold (with a slightly worse constant). The upper bound on $E(C_k^2)$ is of course still valid.

Theorem 3.10. *If* $2n - (n \log n)^{1/2} \le k \le 2n$, *then*

$$E(C_k) = \tau + O\left(\frac{(\log n)^{3/2}}{n^{1/2}}\right),$$

and

$$\operatorname{var}(C_k) = O\left(\frac{(\log n)^{3/2}}{n^{1/2}}\right).$$

3.6 The lower bound on L_n

We have already obtained quite precise results on the cost of the minimum 2-factor in the Poisson(1) model. Here we wish to show that if the Poisson(1) costs are replaced by uniform $[0,1]$ costs, then the cost of the minimum 2-factor rarely decreases dramatically, hence it is rarely considerably smaller than τ.

Since in the Poisson model we allow for multiple edges, we consider the following model: For every pair of vertices, there are two edges, one with uniform $[0,1]$ cost, and one whose cost is the first one plus another independent uniform $[0,1]$ number. This corresponds exactly to the Poisson(1) model.

In the uniform model, take the minimum 2-factor and delete the $n^{1/2} + O(1)$ longest edges. Let $k = 2n - n^{1/2} + O(1)$ be the number of remaining edges, and let x be the length of the longest remaining edge. Now we apply the transformation

$$t \mapsto -\log(1-t)$$

to all edge costs, and to the increments in cost between the first and second edge of each pair. Let C be the cost of the k-flow in the Poisson(1) model that we have obtained.

We show that given C, we can obtain a lower bound on the cost of the original 2-factor in the uniform model. This cost must be at least

$$\frac{C \cdot x}{-\log(1-x)} + \sqrt{n} \cdot x \ge C \cdot (1-x) + \sqrt{n} \cdot x,$$

which is at least C, unless $C > \sqrt{n}$. If on the other hand $C \geq \sqrt{n}$, then

$$\frac{C \cdot x}{-\log(1-x)} + \sqrt{n} \cdot x \geq \frac{\sqrt{n} \cdot x}{-\log(1-x)} + \sqrt{n} \cdot x$$
$$= \sqrt{n}\left(x + \frac{x}{-\log(1-x)}\right) \geq \sqrt{n}. \quad (18)$$

Hence we have:

Lemma 3.11. *The cost of the minimum 2-factor in the uniform model is at least*

$$\min(C_k, \sqrt{n}),$$

where C_k is the cost of the minimum k-flow in the Poisson model, and $k = 2n - n^{1/2} + O(1)$ as above.

Proof. Since the flow of cost C that we have obtained is a k-flow, $C \geq C_k$. □

It remains to bound the contribution to $E(C_k)$ from the cases where $C_k > \sqrt{n}$. We have

$$E\left(C_k^2\right) \geq \sqrt{n} \cdot E\left(C_k \cdot I(C_k \geq \sqrt{n})\right).$$

Hence we can use the upper bound on the variance to lower bound the right hand side. We have

$$E\,(\text{cost of minimum tour in the uniform model})$$
$$\geq E\,(\text{cost of 2-factor in the uniform model})$$
$$\geq E\,(\text{cost of 2-factor in the uniform model with multiple edges permitted})$$
$$\geq E\left(\min(C_k, \sqrt{n})\right) \geq E\,(C_k) - E\left(C_k \cdot I(C_k \geq \sqrt{n})\right)$$
$$\geq E\,(C_k) - \frac{E\,(C_k^2)}{\sqrt{n}} = \tau + O\left(\frac{(\log n)^{3/2}}{n^{1/2}}\right) + O\left(\frac{1}{n^{1/2}}\right)$$
$$= \tau + O\left(\frac{(\log n)^{3/2}}{n^{1/2}}\right). \quad (19)$$

We let L_n denote the cost of the minimum tour in the uniform model, and if x is a real number, then we let

$$x^+ = \begin{cases} x, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0. \end{cases}$$

From the inequalities above, we conclude:

Theorem 3.12. *As $n \to \infty$,*

$$E(\tau - L_n)^+ = O\left(\frac{(\log n)^{3/2}}{n^{1/2}}\right).$$

4 Application to the random TSP

In this section, we finally apply the results we have obtained to the random TSP. We let L_n denote the length of the minimum tour on $K_{n,n}$. Since it only remains to establish an appropriate upper bound on L_n, we work with Poisson(1) edge costs. By stochastic dominance, the upper bound we obtain will be valid also with uniform $[0, 1]$ edge costs.

4.1 A high probability bound on L_n

We first establish a high probability upper bound on L_n. This bound permits the inference of convergence in mean from convergence in probability.

Theorem 4.1.

$$E\left(L_n - 4\zeta(2)\right)^+ = O\left(\frac{(\log n)^{3/4}}{n^{1/4}}\right), \quad \textit{as } n \to \infty.$$

Proof. We randomly colour the edges in three different colours, the first two colours (red and blue, say) have probability $1/2 - c$ each, and the third has probability $2c$. The edges of the third colour are also given a random orientation. The value of c will be chosen later. For each colour and direction, the edges appear as a Poisson process with rate given by these probabilities. Moreover, these four Poisson processes are independent. For the two first colours, red and blue, we find the minimum perfect matchings. From Theorem 7.1 of [47] we know that the cost of such a matching has expected value

$$\frac{1}{1/2 - c} \cdot (\zeta(2) - O(1/n)) = 2\zeta(2) + O(c) + O(1/n)$$

(for small c and large n), and variance

$$O\left(\frac{1}{n(1/2 - c)^2}\right) = O\left(\frac{1}{n}\right).$$

It follows from Jensen's inequality that the cost X of these two perfect matchings satisfies

$$E(X - 4\zeta(2))^+ = O\left(c + \frac{1}{\sqrt{n}}\right).$$

Since the red and blue matchings are independent, the number of cycles in the resulting 2-factor will be distributed like the number Z_n of cycles of a permutation taken uniformly from the set of all $n!$ permutations of n objects. It is easy to show by induction on n that this number satisfies

$$E\left(2^{Z_n}\right) = n + 1.$$

It follows from Markov's inequality that

$$Pr(Z_n > 4\log n) = Pr(2^{Z_n} > 2^{4\log n}) = Pr(2^{Z_n} > n^{4\log 2}) \le \frac{n+1}{n^{4\log 2}} < \frac{1}{n^{1.77}}.$$

We also want to bound the probability that there is no large cycle in the random permutation. Throughout, when we say that a cycle has *size* m, we mean that it contains m vertices in each of V and W, and consequently a total of $2m$ vertices and $2m$ edges. We have

$$\begin{aligned} &Pr(\text{no cycle of size } n/\log n \text{ or larger}) \\ &\qquad \le \left(\frac{2}{\log n}\right)^{(\log n)/2} = \frac{n^{\frac{1}{2}\log 2}}{n^{\frac{1}{2}\log\log n}} = O\left(\frac{1}{n^{1.77}}\right). \end{aligned} \tag{20}$$

Suppose now that there is a cycle of size at least $n/\log n$, and that there are at most $4\log n$ cycles. Then we use the directed edges of the third colour to patch the smaller cycles, one at a time, to the larger cycle. Such a patching is shown in Figure 4. To avoid dependencies among the edge costs under consideration, we patch each new cycle using edges directed from the new cycle to the main cycle.

We estimate the expected cost of the patching. For each patch, the expected cost is

$$O\left(\frac{(\log n)^{1/2}}{cn^{1/2}}\right),$$

and there are $4\log n$ patches to perform. Hence the expected cost of patching all the cycles to a tour is

$$O\left(\frac{(\log n)^{3/2}}{cn^{1/2}}\right). \tag{21}$$

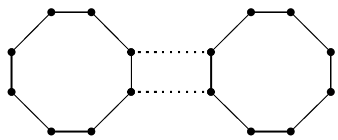

Figure 4: A simple patch turning two cycles into one.

If the number of cycles is larger than $4\log n$, or there is no cycle of size at least $n/\log n$, then we use the directed edges to create a tour of expected length $2n/c$. The contribution to $E(L_n - 4\zeta(2))^+$ from the cases of failure is

$$O\left(\frac{1}{cn^{0.77}}\right),$$

which is smaller than (21).

It remains to choose the value of c in order to minimize

$$O\left(c + \frac{1}{\sqrt{n}}\right) + O\left(\frac{(\log n)^{3/2}}{cn^{1/2}}\right).$$

We must choose $c >> 1/\sqrt{n}$, otherwise the second term will not tend to zero as $n \to \infty$. Hence the first term is $O(c)$, and to minimize, we put

$$c = \frac{(\log n)^{3/2}}{cn^{1/2}},$$

that is,

$$c = \frac{(\log n)^{3/4}}{n^{1/4}}.$$

This gives the bound stated in the theorem. □

It then remains to estimate the bounded variable $\min(L_n, 4\zeta(2))$.

4.2 The longest edge in the minimum 2-factor

Let X_n be the length of the longest edge in the minimum 2-factor.

Theorem 4.2.

$$E(X_n) \leq \frac{4(\log n + 1)}{n+1} = O\left(\frac{\log n}{n}\right).$$

Proof. A $(2n-1)$-flow can be obtained by deleting the longest edge in the minimum $2n$-flow. Hence an upper bound for $E(X_n)$ is $E(C_{2n} - C_{2n-1}) = E(C_{2n}) - E(C_{2n-1})$. The statement now follows from Lemma 3.9 □

We remark that Theorem 4.2 actually gives the right order of magnitude of $E(X_n)$, since already the maximum over vertices v of the length of the shortest edge from v is of order $\log n/n$.

We also remark that Alan Frieze [16] has shown (for the complete graph, but his argument goes through for the bipartite graph as well) that there is an absolute constant A such that $X_n \leq A \log n/n$ with high probability as $n \to \infty$. We conjecture that

$$\frac{X_n \cdot n}{\log n}$$

converges in probability to some number between 1 and 4.

4.3 A bound on the number of small cycles

Our idea is to show that the minimum 2-factor can be changed into a tour by replacing a relatively small number of edges, and that this can be done at $o(1)$ extra cost. The arguments given in Section 4.1 show that this would be easy if it could be established that the cycle structure of the minimum 2-factor is roughly the same as that of a random 2-factor taken from uniform distribution on all 2-factors. Unfortunately we do not know of any method to establish such a theorem. We must, however, obtain a $o(n)$ bound on the number of cycles in order for our approach to work. If the average length of the cycles in the minimum 2-factor would be, say, about 100, then we would have to change at least $n/100$ edges to construct a tour, and this cannot be done at $o(1)$ expected cost.

Fortunately, for large n, it is unlikely that one can find a cycle starting at a given point and returning in not more than 100 steps without using some edge of length $>> \log n/n$. Hence we can use the bound on the expected cost of the longest edge in the minimum 2-factor to establish a bound on the number of short cycles. We know from Theorem 4.2 that the expected value of the longest edge in the minimum 2-factor is $O(\log n/n)$. From this and the Markov inequality it immediately follows that:

Lemma 4.3. *With failure probability $O(1/\log n)$, there is no edge longer than*

$$\frac{(\log n)^2}{n}$$

in the minimum 2-factor.

From this, we obtain a bound on the number of extremely short cycles. In [16], Alan Frieze obtains a similar bound. Using more sophisticated techniques, Frieze also obtains a stronger $O(n/\log n)$ bound on the total number of cycles, but it turns out that Lemma 4.4 below is sufficient for our purposes.

Lemma 4.4. *With a failure probability of $O(1/\log n)$, the minimum 2-factor contains fewer than $n^{0.81}$ cycles of size at most*

$$\frac{\log n}{5\log\log n}.$$

Proof. We estimate the number of cycles of size at most k with no edge longer than $(\log n)^2/n$ (in the whole graph, without regard to the minimum 2-factor). The expected number of such cycles is at most

$$\sum_{l=1}^{k} \frac{(n!)^2}{(n-l)!^2} \cdot \frac{(\log n)^{4l}}{n^{2l}} \leq k(\log n)^{4k}.$$

We now take

$$k = \frac{\log n}{5\log\log n}.$$

Then the expected number of cycles of length at most k is at most

$$\frac{\log n}{5\log\log n} \cdot (\log n)^{4\log n/(5\log\log n)} = \frac{\log n}{5\log\log n} \cdot \exp\left(\frac{4\log n}{5\log\log n} \cdot \log\log n\right) = O(n^{4/5}\log n). \quad (22)$$

□

4.4 Simple path extension

In this and the following sections we prove that with high probability, the minimum 2-factor can be changed to a tour at $o(1)$ extra cost. A technical obstacle is that if we condition on the minimum 2-factor, the costs of the remaining edges in the graph will no longer have the same distribution. To overcome this, we use the same method as in Section 4.1. We randomly "colour" some (relatively small fraction) of the edges with a different colour. We find the minimum 2-factor on the remaining edges, and then use the coloured edges to "patch" the 2-factor into a tour.

It turns out that we are going to need three extra "colours" of edges, and the coloured edges will also be assigned a random orientation. The extra edges will be called *path extension edges*, *rotation edges* and *connecting edges*. We let these three colours have rate c each, so that the "normal" edges occur with rate $1 - 3c$. The choice of c will be optimized later, but in any case c will tend to zero as $n \to \infty$. Throughout, we allow a failure probability of $O(1/\log n)$, which is actually much smaller than we need in order to establish Theorem 1.1.

In this section, we show that with high probability, the path extension edges can be used to turn the minimum 2-factor into a set of edges consisting of one long path, and $o(n/\log n)$ cycles, at $o(1)$ extra cost.

We consider the minimum 2-factor on the normal edges. We assume that this 2-factor satisfies the conclusion of Lemma 4.4, that is, there are fewer than $n^{0.81}$ cycles of size at most $\log n/(5\log\log n)$.

Now we use the path extension edges to connect most of these cycles by what we call *simple path extension*. We start with an arbitrary vertex u_0 and choose the shortest path extension edge directed from u_0 to a vertex v_1 in a different cycle. Then we let u_1 be a vertex adjacent to v_1 in this cycle, and connect u_1 to a vertex in yet another cycle by choosing the shortest path extension edge directed from u_1, and so on.

We continue this process until the number of vertices that are not connected to the path is at most

$$\frac{n}{(\log n)^{1/2}}.$$

We estimate the total cost of the simple extension phase. The total number of cycles in the minimum 2-factor is at most

$$n^{0.81} + O\left(\frac{n\log\log n}{\log n}\right) = O\left(\frac{n\log\log n}{\log n}\right).$$

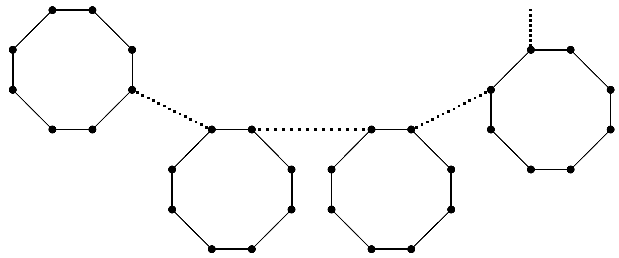

Figure 5: Simple path extension.

This is an upper bound on the number of steps in the simple extension phase. The expected cost of each step is at most

$$O\left(\frac{(\log n)^{1/2}}{cn}\right).$$

Hence the expected total cost of the simple extension phase is

$$O\left(\frac{\log\log n}{c(\log n)^{1/2}}\right).$$

4.5 An expander theorem

The final phase of turning the 2-factor into a tour uses the rotation and connecting edges. For simplicity we again take these edges to be oriented. We choose a particular set of rotation edges that we call the *expander set*. The expander set consists of the seven cheapest rotation edges directed away from each vertex. We show that with failure probability $O(1/\log n)$ (actually much less), the set of expander edges has a certain good expander property (which is why we call them expander edges).

Definition 4.5. If S is a subset of V, then we let S' be the set of vertices in W that are connected to S by an expander edge.

The expander property we want to obtain is the following: For every subset S of V with $|S| \leq n/8$, we have $|S'| \geq 4\,|S|$.

Theorem 4.6 (Expander theorem)**.** *With failure probability $O(1/\log n)$, the set of expander edges has the desired expander property.*

Proof. We estimate the failure probability for the expander property. This probability is bounded by

$$\sum_{s=1}^{n/8} \binom{n}{s}\binom{n}{4s}\left(\frac{4s}{n}\right)^{7s} \le \sum_{s=1}^{n/8} \binom{2n}{5s}\left(\frac{4s}{n}\right)^{7s} \le \sum_{s=1}^{n/8} \left(\frac{2n\cdot e}{5s}\right)^{5s}\left(\frac{4s}{n}\right)^{7s}$$
$$\le \sum_{s=1}^{n/8} \left(\frac{(2e)^5\cdot 4^7\cdot s^2}{5^5\cdot n^2}\right)^{s} \le \sum_{s=1}^{n/8} \left(29\frac{s^2}{n^2}\right)^{s}. \quad (23)$$

For $s \ge 4\log n$, the terms are smaller than

$$\left(\frac{29}{64}\right)^{4\log n} < \frac{1}{n^2},$$

while for $s < 4\log n$, the terms are smaller than

$$4^2\cdot 29\cdot \frac{(\log n)^2}{n^2}.$$

The sum is therefore bounded by

$$O\left(n\cdot \frac{(\log n)^2}{n^2}\right),$$

and in particular by

$$O\left(\frac{1}{\log n}\right).$$

□

4.6 Rotation phase

When we enter the rotation phase, we assume that there is a path that contains all but $O(n/(\log n)^{1/2})$ vertices, and that these remaining vertices make up at most

$$\frac{5n\log\log n}{(\log n)^{3/2}}$$

cycles. We also assume that the expander edges constitute a set with the good expander property.

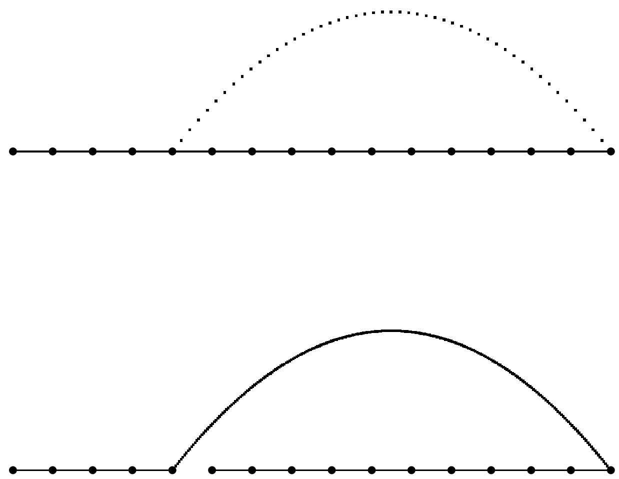

Figure 6: Rotation of the main path.

We will show that each of the remaining cycles can be absorbed into the path by using $O(\log n)$ expander edges possibly together with one connecting edge.

The rotation operation is carried out as follows (see Figure 6). Let P be a path of even length, that is, with one endpoint a in V and the other in W. If there is a rotation edge from a to another vertex b in P, then by replacing one of the edges from b by this rotation edge, we obtain a new path (on the same set of vertices) with one of the neighbours a' of b as the endpoint in V. The operation can then be iterated by using rotation edges from a' and so on.

Let $E_0 = \{a\}$, and let E_i be the set of vertices in V that can become enpoints of the path by performing at most i rotations. If for a particular value of i, fewer than $n/8$ vertices belong to E_i, then by the expander property, either there is a rotation edge from one of the vertices of E_i to a vertex outside P, or there are at least twice as many vertices in E_{i+1} as in E_i.

This shows that the size of E_i will grow exponentially until either there is a rotation edge to a vertex outside P, or at least one eighth of the vertices in $P \cap V$ can become enspoints of the main path by performing at most i rotations.

In the former case, we can extend the main path by using rotation edges only. In the latter case, we pick one of the cycles outside the main path, and we connect it to the main path by using the cheapest connecting edge from this cycle to one of the possible endpoints of the main path.

Finally, after absorbing all the remaining cycles into the main path, we turn this path into a tour by performing the same operation once more, this time treating the endpoint in W as the cycle to be absorbed.

4.7 Proof of Theorem 1.1

We can now estimate the cost of the TSP, thereby completing the proof of Theorem 1.1. We fist estimate the cost of the 2-factor on the normal edges, and the path extension phase. The normal edges have density $1 - 3c$, and we solve the minimum 2-factor problem. The expected cost of this is

$$\frac{\tau}{1-3c},$$

and the variance is of order $(\log n)^{3/2}/n^{1/2}$. The standard deviation is of order $(\log n)^{3/4}/n^{1/4}$, and therefore we can afford to let the algorithm fail if the cost of the minimum 2-factor is larger than $\tau/(1-3c) + (\log n)^{7/4}/n^{1/4}$.

We then turn to the cost of the rotation phase. There are

$$O\left(\frac{n \log\log n}{(\log n)^{3/2}}\right)$$

steps, and each step uses $O(\log n)$ expander edges. The expander edges each have expected cost $O(1/(cn))$. This gives a total expected cost of

$$O\left(\frac{\log\log n}{c(\log n)^{1/2}}\right)$$

for the expander edges, that is, the same as the cost of the simple extension phase. The connecting edges will cost only $O(1/(cn))$ each, so the cost of the connecting edges can be absorbed into this term.

Summing up, the expected cost of $|L_n - \tau|$, given that the algorithm succeeds, is bounded by

$$O(c) + O\left(\frac{(\log n)^{7/4}}{n^{1/4}}\right) + O\left(\frac{\log\log n}{c(\log n)^{1/2}}\right).$$

We now put

$$c = \frac{1}{(\log n)^{1/4}},$$

and obtain the error term

$$O\left(\frac{\log\log n}{(\log n)^{1/4}}\right).$$

The extra contribution from the cases of failure is, by Theorem 4.1, only of order $O(1/\log n)$. This completes the proof of Theorem 1.1.

References

[1] Aldous, David and Bandyopadhyay, Antar, *A Survey of Max-type Recursive Distributional Equations*, Annals of Applied Probability 15 (2005), 1047-1110.

[2] Aldous, David, Steele, Michael J., *The Objective Method: Probabilistic Combinatorial Optimization and Local Weak Convergence* (2003).

[3] Aldous, David, *Asymptotics in the random assignment problem*, Probab. Theory Relat. Fields, **93** (1992) 507–534.

[4] Aldous, David, *The $\zeta(2)$ limit in the random assignment problem*, Random Structures Algorithms **18** (2001), no 4. 381–418.

[5] Aldous, David, *Percolation-like Scaling exponents for Minimal Paths and Trees in the Stochastic Mean Field Model*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 461 (2005), 825–838.

[6] Beardwood, Jillian, Halton, H. J. and Hammersley, J. M., *The shortest path through many points*, Proc. Cambridge Phil. Soc., **55** (1959), 299–327.

[7] Beveridge, A., Frieze, A. and McDiarmid, C., *Random minimum length spanning trees in regular graphs*, Combinatorica **18** (1998), 311–333.

[8] Buck, Marshall W., Chan, Clara S. and Robbins, David P., *On the Expected Value of the Minimum Assignment*, Random Structures & Algorithms 21 (2002), no. 1, 33–58.

[9] Cerf, N. J., Boutet de Monvel, J., Bohigas, O,. Martin, O. C. and Percus, A. G., *The Random Link Approximation for the Euclidean Traveling Salesman Problem*, J. Physique **7** (1997), 117–136.

[10] Dyer, M. E., Frieze, A. and McDiarmid, C., *On linear programs with random costs*, Mathematical Programming **35** (1986), 3–16.

[11] Edmonds, J., *Submodular functions, matroids and certain polyhedra*, Proc. Int. Conf. on Combinatorics, Calgary 1970, Gordon and Breach, New York, 69–87.

[12] Frieze, A. M., *On the value of a random minimum spanning tree problem*, Discrete Appl. Math. **10** (1985), 47–56.

[13] Frieze, A. and McDiarmid, C., *On random minimum length spanning trees*, Combinatorica **9** (1989) 363–374.

[14] Frieze, A. M. and Sorkin, G. B., *The probabilistic relationship between the assignment and asymmetric traveling salesman problems*, Proceedings of the 15th Annual ACM-SIAM Symposium on Discrete Algorithms, Baltimore MD (2001), 652–660.

[15] Frieze, A. M. and Yukich, J., *Probabilistic analysis of the Traveling Salesman Problem*, in: The traveling salesman problem and its variations, G. Gutin and A.P. Punnen (Eds.), Kluwer Academic Publisher (2002), 257-308.

[16] Frieze, A. M., *On random symmetric travelling salesman problems*, Mathematics of Operations Research, Vol. 29, No. 4, November 2004, 878–890.

[17] Garey, M. R., Graham, R. L. and Johnson, D. S., *Some NP-complete geometric problems*, Proc. of the 8th ACM Symposium on Theory of Computing (1976), 10–29.

[18] Janson, Svante, *The minimal spanning tree in a complete graph and a functional limit theorem for trees in a random graph*, Random Structures and Algorithms 7 (1995), 337-355.

[19] Karp, Richard M., *Probabilistic analysis of partitioning algorithms for the traveling-salesman problem in the plane*, Mathematics of Operations Research Vol. 2, No 3, August 1977, 209–224.

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

[21] Kirkpatrick, S. and Toulouse, G., *Configuration space analysis of travelling salesman problems*, Journal de Physique **46** (1985), 1277–1292.

[22] Krauth, W., Mézard, M., *The Cavity Method and the Travelling Salesman Problem*, Europhys. Lett. **8** (3) (1989), 213–218.

[23] Lawler, Eugene, Combinatorial Optimization: Networks and Matroids, Holt, Rinehart and Winston 1976.

[24] Linusson, S. and Wästlund, J., *A proof of Parisi's conjecture on the random assignment problem*, Probab. Theory Relat. Fields **128** (2004), 419–440.

[25] Mézard, M., Parisi, G., Sourlas, N., Toulouse, G. and Virasoro, M., *Replica symmetry breaking and the nature of the spin glass phase*, Journal de Physique **45** (1984), 843–854.

[26] Mézard, Marc and Parisi, Giorgio *Replicas and optimization*, J. Phys. Lett. **46**(1985), 771–778.

[27] Mézard, M. and Parisi, G., *A replica analysis of the travelling salesman problem*, J. Phys. France **47**:1285–1296 (1986).

[28] Mézard, M. and Parisi, G., *Mean-field equations for the matching and the travelling salesman problems*, Europhys. Lett. **2** (1986) 913–918.

[29] Mézard, Marc and Parisi, Giorgio *On the solution of the random link matching problems*, J. Physique Lettres **48** (1987),1451–1459.

[30] Mézard, Marc and Parisi, Giorgio, *The cavity method at zero temperature*, J. Statist. Phys., **111** (2003), 1–34.

[31] Mézard, M., Parisi, G. and Virasoro, M. A., Spin Glass Theory and Beyond, World Scientific, Singapore 1987.

[32] Nair, Chandra, Prabhakar, Balaji and Sharma, Mayank, *Proofs of the Parisi and Coppersmith-Sorkin conjectures for the finite random assignment problem*, Proceedings of IEEE FOCS, 2003.

[33] Orland, H., *Mean-field theory for optimization problems*, Journal de Physique **46** (1985), Lettres 763–770.

[34] Parisi, Giorgio, *A conjecture on random bipartite matching*, arXiv:cond-mat/9801176, 1998.

[35] Percus, Allon G., and Martin, Olivier C., *The stochastic travelling salesman problem: Finite size scaling and the cavity prediction*, arXiv:cond-mat/9802295 (1998).

[36] Percus, A. G., *Voyageur de commerce et problèmes stochastiques associés*, PhD thesis, Université Pierre et Marie Curie, Paris 1997.

[37] Rhee, Wansoo T. and Talagrand, Michel, *Martingale Inequalities and NP-complete Problems*, Mathematics of Operations Research Vol 12, No 1, February 1987, 177–181.

[38] Rhee, Wansoo T. and Talagrand, Michel, *Martingale Inequalities, interpolation and NP-complete Problems*, Mathematics of Operations Research Vol 14, No 1, February 1987, 91–96.

[39] Steele, J. Michael, *Complete Convergence of Short Paths and Karp's Algorithm for the TSP*, Mathematics of Operations Research Vol. 6 No. 3, August 1981, 374–378.

[40] Talagrand, Michel, *Concentration of measure and isoperimetric inequalities in product spaces*, Inst. Hautes Études Sci. Publ. Math. **81** (1995), 73–205.

[41] Vannimenus, J. and Mézard, M., *On the statistical mechanics of optimization problems of the travelling salesman type*, Journal de Physique **45** (1984), Lettres 1145–1153.

[42] Walkup, D. W., *On the expected value of a random assignment problem*, SIAM J. Comput., **8** (1979), 440–442.

[43] Welsh, D. J. A., *Matroid Theory*, Academic Press, London 1976.

[44] Wästlund, J., *A Proof of a Conjecture of Buck, Chan and Robbins on the Expected Value of the Minimum Assignment*, Random Structures and Algorithms (2005) 237–251.

[45] Wästlund, J., *Exact formulas and limits for a class of random optimization problems*, Linköping Studies in Mathematics No. 5, 2005.

[46] Wästlund, J., *A simple proof of the Parisi and Coppersmith-Sorkin formulas for the random assignment problem*, Linköping Studies in Mathematics No. 6, 2005.

[47] Wästlund, J, *The variance and higher moments in the random assignment problem*, Linköping Studies in Mathematics No. 8, 2005.