Acta Math., 204 (2010), 91–150
DOI: 10.1007/s11511-010-0046-7

The mean field traveling salesman and related problems

by

Johan Wästlund

Chalmers University of Technology
Gothenburg, Sweden

1. Introduction

In a complete graph on n vertices, the edges are assigned independent random costs from a fixed distribution μ on the non-negative real numbers. This is the *mean field model of distance*. Several well-known optimization problems consist in finding a set of edges of minimum total cost under certain constraints. Examples are minimum matching, spanning tree, and the traveling salesman problem (TSP). The distribution of the cost of the solution to these problems has been studied extensively, in particular when μ is either uniform on the interval $[0,1]$ or exponential of mean 1. These distributions both represent the so-called *pseudo-dimension* 1 case, in which a variable X of distribution μ satisfies

$$\frac{P(X<t)}{t}\to 1, \quad \text{as } t\to 0^+. \tag{1}$$

When the number of edges in a solution scales like n, a common phenomenon is that as $n\to\infty$, the cost of the solution converges in probability to some constant which is characteristic of the problem.

We let L_n denote the cost of the TSP. More precisely, L_n is the minimum sum of the edge costs of a cycle that visits each vertex precisely once. It can be proved by elementary methods [18], [21], [46] that there are positive constants c and C such that as $n\to\infty$, $P(c<L_n<C)\to 1$, and the same holds for minimum matching and spanning tree, but establishing convergence and determining the limit is a more difficult problem.

In this paper we study several optimization problems including variations of minimum matching as well as the TSP. We show that they are related to a 2-dimensional version of the urn process of M. Buck, C. Chan and D. Robbins [14], giving rise to a random plane region R_n. In several cases it turns out that as $n\to\infty$, R_n converges to a problem specific limit region $R^\star$ whose shape is described by a simple equation in two

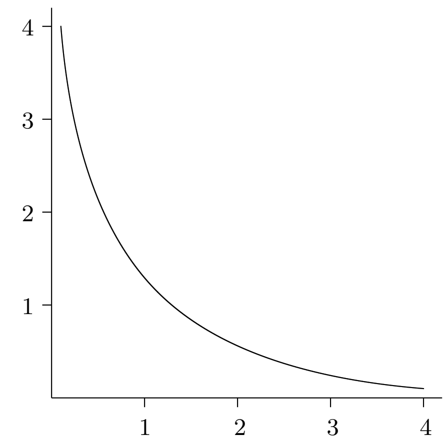

Figure 1. The curve $(1+x/2)e^{-x}+(1+y/2)e^{-y}=1$.

variables. Moreover, the area of $R^\star$ is equal to the limit cost of the optimization problem. Our prime example is the TSP, whose asymptotic cost is given by Theorem 1.1.

THEOREM 1.1. *If* μ *satisfies* (1), *then as* $n\to\infty$,

$$L_n \xrightarrow{\text{p}} \frac{1}{2}\int_0^\infty y\,dx, \tag{2}$$

where y, *as a function of* x, *is the positive solution to the equation*

$$\left(1+\frac{x}{2}\right)e^{-x}+\left(1+\frac{y}{2}\right)e^{-y}=1. \tag{3}$$

The limit, which we denote by $L^\star$, is therefore half of the area under the curve shown in Figure 1. There seems to be no simple expression for $L^\star$ in terms of known mathematical constants, but it can be evaluated numerically to

$$\lim_{n\to\infty} L_n = L^\star \approx 2.0415481864.$$

Our method applies to a variety of problems. For instance, the minimum matching problem leads to the conclusion of Theorem 1.1 with equation (3) replaced by the simpler equation $e^{-x}+e^{-y}=1$. In this case the equation has the explicit solution

$$y=-\log(1-e^{-x}),$$

and, as was conjectured in [13], [26], [27], [29] and [35], and established in [3] and [4], the limit cost is equal to

$$\frac{1}{2}\int_0^\infty -\log(1-e^{-x})\,dx = \frac{\pi^2}{12}.$$

Acknowledgments. This work has been influenced by discussions with several people, of which in particular I thank David Aldous, Svante Janson and Giorgio Parisi.

2. Background

The replica and cavity methods of statistical mechanics, originally developed in the study of spin glasses and other disordered systems, have led to a number of purely mathematical predictions [30], [43]. Several of these results are mathematically non-rigorous, and establishing them has become a stimulating challenge to probability theory. Some important achievements in this direction are David Aldous' proof of the $\zeta(2)$ limit in the assignment problem [4], Michel Talagrand's proof of the correctness of the Parisi solution of the Sherrington–Kirkpatrick model [19], [34], [40], [44], and the algorithmic and theoretical results on phase transitions in constraint satisfaction problems [1], [2], [31] and counting [11], [49]. See [25] for a recent introduction to this field.

One of the areas where the statistical mechanics approach has produced a series of remarkable conjectures is optimization in mean field models of distance. The simplest of these models is the one described in the introduction, but the analogous bipartite model has also been studied. In general one is interested in the large n limit cost of the problem, which corresponds physically to the *thermodynamical limit*.

Explicit limit theorems have been obtained only for distributions satisfying (1). A. Frieze showed [17] that the limit cost of the minimum spanning tree is $\zeta(3)$, and D. Aldous [3], [4] established the limit $\frac{1}{6}\pi^2$ for bipartite minimum matching. See also [6] and [8] for other results based on weak convergence. The theorem of Aldous resolved a conjecture by M. Mézard and G. Parisi from the mid-1980's [26], [27], [29], [30], [35]. In articles of Mézard, Parisi and W. Krauth [23], [27] and [28], a similar limit was conjectured for the TSP. The limit was predicted on theoretical grounds to be a certain constant, approximately 2.0415. This value was consistent with earlier less precise estimates obtained theoretically as well as by computer simulation [10], [22], [30], [41], [45], see also [4], [5], [7], [12], [15], [37], [38] and [39].

Theorem 1.1 confirms the Krauth–Mézard–Parisi conjecture, except that for a while it was not clear that $L^\star$ is the same as the number obtained with the cavity method. The numerical agreement was certainly convincing, and soon it was shown by Parisi (in personal communication) that the analytic characterization of $\lim_{n\to\infty} L_n$ obtained in equations (6) and (9) of [23] is indeed equivalent to the one given by (3). This leads to the satisfactory conclusion that the cavity result for the limit cost of the TSP is correct. Parisi's argument sheds some light on the relation between the cavity solution and the limit shape $R^\star$, but this is something that deserves further study.

3. Organization of the paper

In §4 we introduce a modified random graph model which is designed in order to simplify the analysis. In this model each vertex has a non-negative real *weight*, which governs the distribution of the edge costs, an idea originating in the paper of Buck, Chan and Robbins [14]. We study a certain type of optimization problem that we call *flow problem*, which includes natural relaxations of matching and the TSP. Our method relies on letting the weight of one of the vertices tend to zero. This idea was suggested already in [14], but not fully exploited. With this method we show in §5 that knowledge of the expected number of edges from a vertex in a solution to the flow problem allows us to find recursively the expected cost of the solution.

This strategy is carried out in §6 and §7, where we also introduce the 2-dimensional urn process. The expected cost of the flow problem is expressed as the area of a region R_n defined by the urn process. In §8 we study the behavior of R_n for large n. Using the Talagrand concentration inequality [42], we prove in §9 that under certain conditions, the cost of a flow problem converges in probability to the area of the limit shape of R_n. In §10 and §11, we show that the mean field TSP is well approximated by its relaxation to a flow problem, thereby completing the proof of Theorem 1.1. An important step is provided by a theorem of Frieze [18] stating that the mean field TSP is asymptotically equivalent to the mean field 2-factor problem.

In §12 we show that similar results can be obtained in the technically simpler bipartite graph model, for which our method produces exact results not only for the expectation, but also for the higher moments of the cost of a flow problem. For several problems, the limit cost on the bipartite graph $K_{n,n}$ is twice that of the same problem on the complete graph. Indeed, this is the reason for the factor $\frac{1}{2}$ in equation (2). The region shown in Figure 1 actually corresponds to the bipartite TSP, but the two limit regions are related through a simple change of variables.

For the minimum matching, which is the simplest flow problem, the urn process can be analyzed exactly, and this leads to several explicit formulas that are discussed in §13.

4. Flow problems and the friendly model

In this section we define a certain type of random graph model that we call the *friendly model*, and a class of optimization problems that we call *flow problems*.

4.1. The friendly model

There are n vertices $v_1, ..., v_n$. Each vertex v_i has a positive real *weight* γ_i. For each i and j there is a potentially infinite set of edges connecting v_i and v_j. If $i \neq j$, these edges have costs that are determined by the times of the events in a Poisson process of rate $\gamma_i \gamma_j$. If $i = j$, there is a set of loops at this vertex, and their costs are given by a Poisson process of rate $\frac{1}{2}\gamma_i^2$. All these Poisson processes are independent.

The friendly model can also be characterized in the following way: The total sequence of edge costs in the graph is generated by a Poisson process of rate

$$\tfrac{1}{2}(\gamma_1 + ... + \gamma_n)^2.$$

For every event in this process, an edge of cost equal to the time of the event is added to the graph by choosing its two endpoints independently according to the fixed probability measure on the vertices given by

$$P(v_i) = \frac{\gamma_i}{\gamma_1 + ... + \gamma_n}.$$

4.2. Flow problems

We let each vertex v_i have a non-negative integer *capacity* c_i. Let E denote the set of edges. For each $e \in E$, let $X_e \geqslant 0$ be the cost of e. For $e \in E$ and $v \in V$ we use the notation $\langle e, v \rangle$ for the number of times that e is connected to v. In other words,

$$\langle e, v \rangle = \begin{cases} 2, & \text{if } e \text{ is a loop at } v, \\ 1, & \text{if } e \text{ connects } v \text{ to another vertex}, \\ 0, & \text{otherwise.} \end{cases}$$

For a non-negative real number k, the flow problem asks for the function $\sigma: E \to [0, 1]$ that minimizes

$$\sum_{e \in E} \sigma_e X_e,$$

subject to

- the capacity constraints: for every i,

$$\sum_{e \in E} \langle e, v_i \rangle \sigma_e \leqslant c_i; \tag{4}$$

- the norm of σ is (at least) k:

$$\sum_{e \in E} \sigma_e \geqslant k. \tag{5}$$

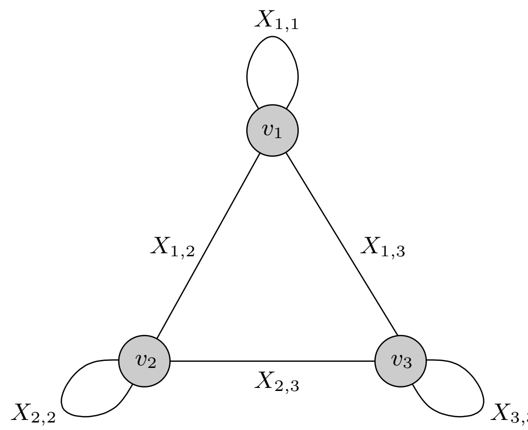

Figure 2. The friendly model, n=3.

Since the edge costs X_e are non-negative, we can obviously replace the inequality in (5) by equality.

A function σ that satisfies the constraints (that is, a feasible solution) is called a *flow*. If equality holds in (5), σ is called a k-flow. The left-hand side of (4) is the *degree* of v_i with respect to σ, and if equality holds in (4) we say that v_i has *full degree* in σ.

In the *generic* case, that is, when there are no linear relations between the edge costs, there is a unique minimum k-flow for every k that allows a feasible solution. In the friendly model, genericity holds with probability 1. We shall sometimes assume genericity without explicitly stating this assumption. The minimum k-flow is then denoted by $\sigma^{(k)}$ (without assuming genericity, this notation would be ambiguous). The cost of $\sigma^{(k)}$ is denoted by C_k, and the degree of v_i in $\sigma^{(k)}$ is denoted by $\delta_i^{(k)}$. We denote the integer vector of capacities by boldface $\mathbf{c}=(c_1, ..., c_n)$, and when dependence on the capacities is crucial, we write $\sigma^{(k)}(\mathbf{c})$, $\delta_i^{(k)}(\mathbf{c})$, etc.

In principle k denotes a non-negative real number, but it turns out that we only have to consider values of k for which $2k$ is an integer. The reason is that for given edge costs, C_k as a function of k is piecewise linear, and the points of non-differentiability can only be located at these particular values of k.

4.3. Example

As a simple example, suppose that n=3 and that the weights and capacities are all equal to 1. Then only the cheapest edge between each pair of vertices is relevant. The graph essentially looks like Figure 2, where the loops have costs $X_{1,1}$, $X_{2,2}$ and $X_{3,3}$ that are exponential of rate $\frac{1}{2}$, that is, mean 2, and the edges connecting distinct vertices have costs $X_{1,2}, X_{1,3}$ and $X_{2,3}$ that are exponential of mean 1.

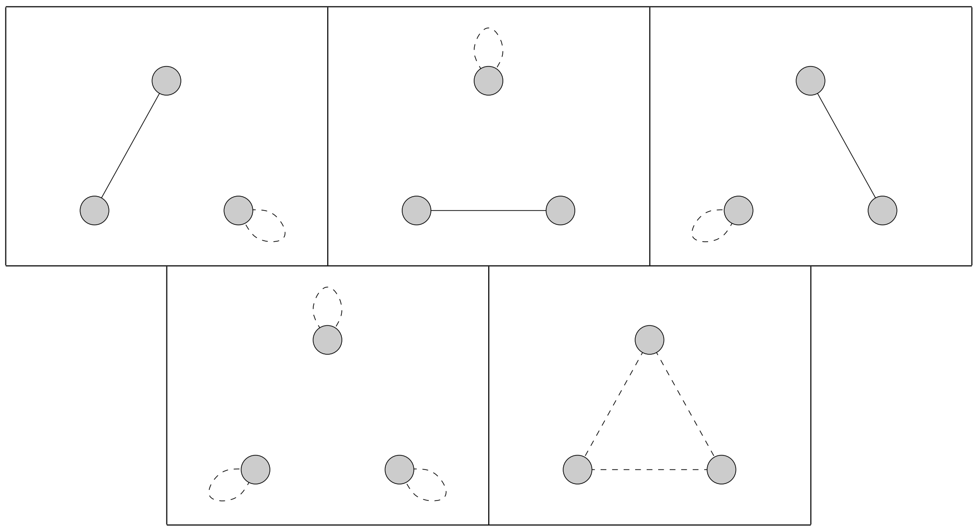

Figure 3. The five basic solutions.

Consider the k-flow problem for $k=\frac{3}{2}$, which is the maximum value of k since the sum of the vertex capacities is 3. Once the edge costs are given, the optimization problem consists in minimizing

$$\sigma_{1,1}X_{1,1}+\sigma_{2,2}X_{2,2}+\sigma_{3,3}X_{3,3}+\sigma_{1,2}X_{1,2}+\sigma_{1,3}X_{1,3}+\sigma_{2,3}X_{2,3},$$

subject to $\sigma_{i,j}\in[0,1]$,

$$\sigma_{1,1}+\sigma_{2,2}+\sigma_{3,3}+\sigma_{1,2}+\sigma_{1,3}+\sigma_{2,3}=\tfrac{3}{2}$$

and

$$\begin{aligned}2\sigma_{1,1}+\sigma_{1,2}+\sigma_{1,3}&\leqslant 1,\\ 2\sigma_{2,2}+\sigma_{1,2}+\sigma_{2,3}&\leqslant 1,\\ 2\sigma_{3,3}+\sigma_{1,3}+\sigma_{2,3}&\leqslant 1.\end{aligned}$$

It is shown in §6 that for the optimum solution in the generic case, $\sigma_{i,j}$ only takes the values 0, $\frac{1}{2}$ and 1. Therefore the optimum solution must be one of the five solutions shown in Figure 3. The dashed lines indicate edges of coefficient $\frac{1}{2}$.

The loops can only occur with coefficient $\frac{1}{2}$. Since they are exponential of mean 2, it seems that the model can be simplified by letting the loops be exponential of mean 1, and counting them only once in the degree of their vertices. Indeed, if we simulate this example by computer, we most easily generate exponential mean-1 variables $Y_{i,j}$ for $1\leqslant i\leqslant j\leqslant 3$, and then compute

$$\min\bigl\{Y_{1,1}+Y_{2,2}+Y_{3,3}, Y_{1,1}+Y_{2,3}, Y_{2,2}+Y_{1,3}, Y_{3,3}+Y_{1,2}, \tfrac{1}{2}(Y_{1,2}+Y_{1,3}+Y_{2,3})\bigr\}.$$

However, when the capacites are at least 2, loops can occur with coefficient 1, and it then turns out that the definitions of §4.1 and §4.2 are the correct ones.

A computer simulation in order to estimate the expectation of the cost of the minimum solution reveals that $EC_{3/2}\approx 0.86$. In fact, it follows from equation (54) that $EC_{3/2}=\frac{31}{36}$.

4.4. Integer flow problems

The optimization problem described in §4.2 is the *linear* flow problem. If we require that $\sigma_e\in\{0,1\}$, we get an *integer* flow problem. For such a problem, a solution can be regarded as a set of edges (the set of edges of coefficient 1).

The integer flow problem, specialized to the case that all capacities are 1, is well known as the *minimum matching problem*. A feasible solution, called a k-matching, is a set of k edges of which no two have a vertex in common.

A 2-*factor* is a set of edges for which every vertex has degree 2. The minimum 2-factor problem is an integer flow problem for which every vertex has capacity 2, and $k=n$. A 2-factor can also be described as a set of vertex-disjoint cycles. A solution to the TSP is a 2-factor with only one cycle, also known as a *tour*. Therefore the 2-factor problem is a relaxation of the TSP.

5. Description of the method

Our method is based on letting the weight of a vertex tend to zero. The results of this section are valid for the integer as well as the linear flow problem.

Suppose that we are given a random flow problem in the friendly model as described in §4. This means that we have specified the number n, the weights $\gamma_1, ..., \gamma_n$ and the capacities $c_1, ..., c_n$. We also assume that $k\geqslant 1$ and that k is small enough so that feasible solutions to the flow problem exist.

We now introduce an extra vertex v_{n+1} of weight γ_{n+1} and capacity 1. We compare the flow problem for the capacity vector $\mathbf{c}=(c_1, ..., c_n, 0)$, that is, the original problem, with the flow problem for the capacity vector $\mathbf{c}+\mathbf{1}_{n+1}=(c_1, ..., c_n, 1)$. The latter is called the *extended problem*. Here $\mathbf{1}_i$ denotes the vector that has a 1 in position i and zeros elsewhere.

For the moment, we fix a vertex v_i, $1\leqslant i\leqslant n$, and assume that its capacity c_i is non-zero. Since, in the extended flow problem, the capacity of v_{n+1} is still only 1, at most one edge (the cheapest) between v_i and v_{n+1} is relevant to the flow problem. We denote this edge by e and note that the cost X_e of e has exponential distribution of rate $\gamma_i\gamma_{n+1}$, in other words, the density is

$$\gamma_i\gamma_{n+1}e^{-\gamma_i\gamma_{n+1}t}$$

for $t\geqslant 0$. We are interested in the expected value of the coefficient $\sigma_e^{(k)}(\mathbf{c}+\mathbf{1}_{n+1})$ of e in the minimum k-flow in the extended problem, as a function of γ_{n+1}.

Let us for the moment condition on the costs of all other edges. We let $f(x)$ denote the cost of the minimum k-flow with respect to $\mathbf{c}+\mathbf{1}_{n+1}$ given that $X_e=x$. We have

$$EC_k=Ef(X_e)=\gamma_i\gamma_{n+1}\int_0^\infty e^{-\gamma_i\gamma_{n+1}x}f(x)\,dx.$$

Moreover, the coefficient $\sigma_e^{(k)}(\mathbf{c}+\mathbf{1}_{n+1})$ is the derivative of $f(x)$ at $x=X_e$. Here we disregard the fact that there may be a finite number of points where f is non-differentiable. The following calculation by partial integration only requires f to be continuous. We have

$$E[\sigma_e^{(k)}(\mathbf{c}+\mathbf{1}_{n+1})]=\int_0^\infty \gamma_i\gamma_{n+1}e^{-\gamma_i\gamma_{n+1}x}f'(x)\,dx. \tag{6}$$

By partial integration, (6) is equal to

$$\begin{aligned}\gamma_i\gamma_{n+1}&\left[\int_0^\infty \gamma_i\gamma_{n+1}e^{-\gamma_i\gamma_{n+1}x}f(x)\,dx-f(0)\right]\\ &=\gamma_i\gamma_{n+1}[EC_k(\mathbf{c}+\mathbf{1}_{n+1})-(C_k(\mathbf{c}+\mathbf{1}_{n+1})\mid X_e=0)].\end{aligned} \tag{7}$$

We are still conditioning on all edge costs except X_e. In (7), we therefore regard

$$(C_k(\mathbf{c}+\mathbf{1}_{n+1})\mid X_e=0)$$

as a non-random quantity.

Since one way of obtaining a k-flow with respect to $\mathbf{c}+\mathbf{1}_{n+1}$ is to use the edge e together with the minimum $(k-1)$-flow with respect to $\mathbf{c}-\mathbf{1}_i$, we have

$$(C_k(\mathbf{c}+\mathbf{1}_{n+1})\mid X_e=0)\leqslant C_{k-1}(\mathbf{c}-\mathbf{1}_i).$$

If X is an exponential random variable of rate λ, then X can be defined as $X=Y/\lambda$, where Y is a rate-1 variable. This way, the edge costs can be generated from an underlying set of rate-1 variables. If these underlying variables constitute the probability space, then we can let $\gamma_{n+1}\to 0$ for a fixed point in this space. Then the costs of all edges of non-zero cost from v_{n+1} will tend to infinity. It follows that, pointwise,

$$C_k(\mathbf{c}+\mathbf{1}_{n+1})\to C_k \quad \text{and} \quad (C_k(\mathbf{c}+\mathbf{1}_{n+1})\mid X_e=0)\to C_{k-1}(\mathbf{c}-\mathbf{1}_i).$$

By the principle of dominated convergence, we conclude from (7) that, as $\gamma_{n+1}\to 0$,

$$\frac{E[\sigma_e^{(k)}(\mathbf{c}+\mathbf{1}_{n+1})]}{\gamma_{n+1}}\to \gamma_i(EC_k(\mathbf{c})-EC_{k-1}(\mathbf{c}-\mathbf{1}_i)). \tag{8}$$

So far, we have conditioned on all edge costs except the cost X_e of e. Now it is clear that (8) must hold also if we interpret the expectations as averages over all edge costs.

In §7.2 we show how to compute the expected degree of a vertex in the minimum flow. In particular this allows us to compute

$$\lim_{\gamma_{n+1}\to 0}\frac{1}{\gamma_{n+1}}\sum_e E[\sigma_e^{(k)}(\mathbf{c}+\mathbf{1}_{n+1})], \tag{9}$$

where the sum is taken over all edges e from v_{n+1} (only the cheapest edge to each other vertex is relevant). If we assume that $c_i>0$ for every i, then it follows from (8) that (9) is equal to

$$\biggl(\sum_{i=1}^n \gamma_i\biggr)EC_k(\mathbf{c})-\sum_{i=1}^n \gamma_i EC_{k-1}(\mathbf{c}-\mathbf{1}_i).$$

A convenient way to state this equation is

$$EC_k(\mathbf{c})=\frac{1}{\gamma_1+\ldots+\gamma_n}\lim_{\gamma_{n+1}\to 0}\frac{E[\delta_{n+1}^{(k)}(\mathbf{c}+\mathbf{1}_{n+1})]}{\gamma_{n+1}}+EC_{k-1}(\mathbf{c}-\mathbf{1}_i), \tag{10}$$

where the last term $EC_{k-1}(\mathbf{c}-\mathbf{1}_i)$ is interpreted as an average not only over the edge costs but also over a random choice of i, taken with probabilities proportional to the weights. In other words, the probability of choosing a particular i is

$$\frac{\gamma_i}{\gamma_1+\ldots+\gamma_n}.$$

6. Combinatorics of the linear flow problem

In this section we establish some combinatorial results on the structure of the minimum linear flows. These results are valid for arbitrary non-negative edge costs, and are therefore independent of the random model introduced in §4.1.

In the linear flow problem we allow a coefficient σ_e to be any number in the interval $[0,1]$. The following proposition shows that we can still restrict our attention to a finite number of potentially optimal solutions.

PROPOSITION 6.1. *Suppose that for a linear flow problem the capacities, the edge costs and the number k are given. Suppose moreover that the sum $2k$ of the degrees of the vertices in a solution is an integer. Then there is a minimum k-flow whose coefficients all belong to the set* $\left\{0, \frac{1}{2}, 1\right\}$.

In the proof of Proposition 6.1, we let σ be a k-flow, and assume that σ is not a convex combination of other k-flows. We let H be the subgraph consisting of all edges of non-integer coefficient in σ, and all vertices incident to such an edge. The main part of the proof consists of the following lemma.

LEMMA 6.2. *All but at most one of the components of H are cycles of odd length where all edges have coefficient $\frac{1}{2}$ and all vertices have full degree. There may or may not be an exceptional component, which is a path or cycle of odd length, or a cycle with an attached path.*

Proof. We first show that there cannot be any cycle of even length in H (and in particular no multiple edges). Let us say that a cycle is *non-trivial* if some edge occurs an odd number of times. We claim that H cannot contain a non-trivial cycle of even length. Suppose for a contradiction that there is such a cycle in H. Then we modify σ by an operation that we call *switching*. We construct a new flow σ' by letting $\sigma'_e = \sigma_e$ whenever e is not in the cycle, while if e is in the cycle, $\sigma'_e = \sigma_e \pm \varepsilon$, where the signs alternate around the cycle, and $\varepsilon > 0$ is chosen small enough that $\sigma'_e \in [0, 1]$. Each vertex has the same degree in σ' as in σ, so σ' is a flow. If we now construct σ'' by choosing the signs in the opposite way, we have

$$\sigma = \tfrac{1}{2}(\sigma' + \sigma'').$$

The assumption that there is some edge that occurs an odd number of times in the cycle implies that σ' and σ'' are distinct from σ, a contradiction.

For the same reason, there cannot be a path of even length in H connecting two vertices which do not have full degree in σ. This means that the components of H are trees or unicyclic (if a component has two distinct cycles, then it must be possible to find a non-trivial cycle of even length). Since the leafs in a tree component of H obviously do not have full degree (they do not have integer degree), the tree components must be paths of odd length (if there were three leafs, then two of them would be at even distance from each other).

In an odd cycle, there cannot be more than one vertex of less than full degree, since two such vertices would be at even distance from each other along some path. Actually there cannot even be two vertices of less than full degree in two different odd cycles, because then we can switch σ along these cycles, and start by increasing in one of them and start by decreasing in the other. Again this would make σ a convex combination of two other k-flows.

This argument works for cycles where some edges are used twice. Hence in all the unicyclic components of H, there cannot be more than a total of one vertex that does not have full degree in σ. In particular, the unicyclic components can have at most one leaf altogether.

In the same way, we can exclude the possibilities that there are two distinct path components in H, or that there is a path and a unicyclic component with a leaf. Therefore the structure of H must be as claimed.

It is clear that at most one component of H can have a vertex of less than full degree. Moreover, in a component of H where all vertices have full degree, the non-integer coefficients must add to integers at every vertex, and the only solution is a cycle of odd length where every edge has coefficient $\frac{1}{2}$. □

Proof of Proposition 6.1. We use Lemma 6.2 together with the assumption that $2k$ is an integer. If H has no exceptional component, the statement is clear. Suppose therefore that there is an exceptional component H_1, and notice that the edge coefficients in H_1 must add to half an integer.

If H_1 is a path or cycle, then the coefficients must alternate between x and $1-x$ for some $x \in (0,1)$, and we must have $x=\frac{1}{2}$.

It remains to consider the case that H_1 is a cycle with an attached path. The leaf (endpoint of the path) obviously does not have full degree in σ. Hence, by the previous argument, there cannot be another vertex that does not have full degree.

This implies that for some x and y, the coefficients in the path alternate between x and $1-x$, and the coefficients in the cycle alternate between y and $1-y$, with two y's meeting at the point where the path is attached. It follows that $x+2y$ is an integer.

Also, the sum of the coefficients in H_1 is either an integer$+y$ or an integer$+x+y$, depending on whether the path is of even or odd length. Since $y+(x+y)=x+2y$ which is an integer, it follows that if one of y and $x+y$ is half-integral, then so is the other, so in fact both are half-integral. It follows that $y=\frac{1}{2}$ and from this that $x=\frac{1}{2}$. This already proves Proposition 6.1, but in fact we reach a contradiction since it shows that $x+2y$ was not an integer after all. Hence when k is half-integral, the exceptional component must be a path or a cycle. □

Definition 6.3. We say that a flow is *stable* if all edges of non-integer coefficient go between vertices of full degree.

PROPOSITION 6.4. *Suppose that the edge costs are generic. If $2k$ is an integer and the minimum k-flow $\sigma^{(k)}$ is not stable, then*

$$\sigma^{(k+1/2)} = 2\sigma^{(k)} - \sigma^{(k-1/2)}. \tag{11}$$

Proof. Suppose that $\sigma^{(k)}$ has an edge e of coefficient $\frac{1}{2}$ incident to a vertex v which is not of full degree. Then e and v are in the exceptional component H_1, which is either a path or a cycle of odd length. If it is a path, then v has to be an endpoint, and the other endpoint u does not have full degree either (since it has non-integer degree). In any case, by alternating between increasing and decreasing the coefficients of the edges in H_1 by $\frac{1}{2}$ in the two possible ways, we obtain a $\left(k-\frac{1}{2}\right)$-flow σ' and a $\left(k+\frac{1}{2}\right)$-flow σ'' such that

$$\sigma = \tfrac{1}{2}(\sigma' + \sigma'').$$

Since the mean value of a $\left(k-\frac{1}{2}\right)$-flow and a $\left(k+\frac{1}{2}\right)$-flow is always a k-flow, this actually shows that σ' is the minimum $\left(k-\frac{1}{2}\right)$-flow and σ'' is the minimum $\left(k+\frac{1}{2}\right)$-flow. Consequently,

$$\sigma^{(k)} = \tfrac{1}{2}(\sigma^{(k-1/2)} + \sigma^{(k+1/2)}),$$

which is equivalent to (11). □

6.1. The nesting lemma

Finally we establish an analogue of the *nesting lemma* of [14].

LEMMA 6.5. *If the capacities and the edge costs are fixed, then the degree $\delta_i^{(k)}$ of a given vertex v_i is a non-decreasing function of k. Moreover, for half-integral k, if $\delta_i^{(k)}$ is an integer for every i, then $\delta^{(k+1/2)}$ is obtained from $\delta^{(k)}$ by either increasing the value by $\frac{1}{2}$ at two vertices, or increasing by 1 at one vertex. If $\delta_i^{(k)}$ is not an integer for every i, then there are precisely two vertices for which it is half-integral. In this case $\delta^{(k+1/2)}$ is obtained from $\delta^{(k)}$ by rounding up to the nearest integer at these two vertices.*

Proof. Suppose that $2k$ is not an integer and that σ is the minimum k-flow. Then there must be an exceptional component H_1 of H. Let k' and k'' be obtained by rounding k down and up, respectively, to the nearest half-integers. If H_1 is a path or a cycle, then plainly σ is a convex combination of a k'-flow σ' and a k''-flow σ'' obtained by switching in the two ways.

The corresponding holds also when H_1 is a cycle with an attached path. In the notation of the proof of Proposition 6.1, σ' and σ'' are then obtained by rounding y to the two nearest half-integer values, and setting x accordingly either to 0 or 1, so that the only vertex that changes its degree is the leaf.

Again, since every convex combination of flows is a flow, it follows that σ' and σ'' are minimal, and the statement of the lemma follows. □

7. Expected value of the minimum flow

In this section we establish the connection between the random flow problem and the 2-dimensional urn process. Several of the ideas go back to the paper [14] by Buck, Chan and Robbins. In particular, Proposition 7.1 is a generalization of their Lemma 5.

Here we consider the linear flow problem. Bounds on the expected cost of integer flow problems are obtained in §10.

7.1. The oracle process

It follows from the results of §6 that it suffices to consider minimum k-flows for half-integral k. From now on we therefore let k denote a number such that $2k$ is an integer. We think of an "oracle" who knows all the edge costs. At the beginning, these costs are unknown to us, but we get knowledge about them by asking questions to the oracle. The trick is to choose the questions so that we get knowledge about the location of the minimum flow while at the same time retaining control of the joint distribution of all the edge costs conditioning on the information we get.

We describe a generic step of the process, and we assume that for a certain k the minimum k-flow $\sigma^{(k)}$ is stable. In particular this implies that every vertex has integer degree. Moreover, we assume that the following information is known to us:

O1. The costs of all edges for which both endpoints have full degree in $\sigma^{(k)}$.

O2. The edges of non-zero coefficient in $\sigma^{(k)}$, and their costs (by the stability assumption, these edges have coefficient 1 unless they were included under O1).

O3. The minimum cost of the remaining edges between vertices that do not have full degree (but not the location of the edge that has this minimum cost).

O4. For each vertex v of full degree, the minimum cost of the remaining edges that connect v to a vertex that does not have full degree (but again not the location of this edge).

Using this information only, we can essentially compute the minimum $\left(k+\frac{1}{2}\right)$-flow $\sigma^{(k+1/2)}$. We know from §6.1 that $\sigma^{(k+1/2)}$ is obtained from $\sigma^{(k)}$ by an operation that we refer to as *switching* of an alternating path that connects two vertices not of full degree, that is, the coefficients of the edges in the path are alternatingly increased and decreased by $\frac{1}{2}$. The path can be degenerate in a number of ways, and in particular the two endpoints need not be distinct. The information in O1–O4 allows us to compute everything except the location of the endpoints of this path.

By the memorylessness property of the Poisson process, the unknown endpoints (whether one or two) are chosen independently among the vertices not of full degree in $\sigma^{(k)}$, with probabilities proportional to the weights γ_i. Notice that this holds also if the path consists of only one edge (and that this edge can turn out to be a loop).

After "essentially" computing $\sigma^{(k+1/2)}$, we ask the oracle for the information that will be needed according to O1–O4 in the next round of the process. We begin by asking for the locations of the endpoints of the alternating path. There are essentially three possibilities.

(1) If there are two distinct endpoints v_i and v_j, then their degrees increase by $\frac{1}{2}$, that is, $\delta_i^{(k+1/2)}=\delta_i^{(k)}+\frac{1}{2}$ and $\delta_j^{(k+1/2)}=\delta_j^{(k)}+\frac{1}{2}$. In this case $\sigma^{(k+1/2)}$ is not stable, and by Proposition 6.4, $\sigma^{(k+1)}$ is determined by (11) once $\sigma^{(k+1/2)}$ is known. The same two vertices will again increase their degrees by $\frac{1}{2}$, so that $\delta_i^{(k+1)}=\delta_i^{(k)}+1$ and $\delta_j^{(k+1)}=\delta_j^{(k)}+1$. Since $\sigma_e^{(k+1)}-\sigma_e^{(k)}$ is an integer for every e, the flow $\sigma^{(k+1)}$ is stable.

(2) Even if the alternating path starts and ends in two distinct edges, these edges can turn out to go to the same vertex v_i. Then $\delta_i^{(k+1/2)}=\delta_i^{(k)}+1$. If v_i gets full degree in $\sigma^{(k+1/2)}$, then $\sigma^{(k+1/2)}$ is stable. Otherwise Proposition 6.4 applies again, and $\delta_i^{(k+1)}=\delta_i^{(k)}+2$. In this case, $\sigma^{(k+1)}$ is stable.

(3) The third possibility is that the alternating path starts and ends with the same edge. It is then clear that the endpoints of the path will coincide. This endpoint v_i is again chosen among the vertices not of full degree, with probabilities proportional to the weights. In this case $\delta_i^{(k+1/2)}=\delta_i^{(k)}+1$, and $\sigma^{(k+1/2)}$ is stable.

7.2. The average degree of a vertex in the minimum *k*-flow

In this section we show how to compute the expected degree of a given vertex in the linear flow problem in the friendly model. Suppose that the weights and capacities are given, and recall that for

$$0\leqslant k\leqslant \frac{1}{2}\sum_{i=1}^{n} c_i,$$

$\delta^{(k)}$ is the degree vector for the minimum flow $\sigma^{(k)}$. Here we define another random process which is also based on the weights and capacities of the vertices. This process was introduced (for the bipartite assignment problem) by Buck, Chan and Robbins [14], and we refer to it as the Buck–Chan–Robbins urn process. It turns out that several quantities associated with the random flow problem have counterparts in the urn process. Here we introduce the 1-dimensional version of the urn process and show how $E\delta_i^{(k)}$ can be interpreted in terms of this process.

An urn contains the vertices $v_1, ..., v_n$. Vertices are drawn one at a time from the urn, and each time, the vertex to be drawn is chosen with probabilities proportional to the weights $\gamma_1, ..., \gamma_n$. The capacities $c_1, ..., c_n$ serve as a *replacement protocol*: The vertices that are drawn from the urn are put back into the urn as long as they have not yet been drawn a number of times equal to their capacity. When v_i has been drawn c_i times, it is removed.

The results in this section can be stated for a discrete time version of the process, where vertices are just drawn one at a time, but it is convenient to introduce an equivalent continuous time version, in which the vertices are drawn at random times independently of each other. This is achieved by letting the vertex v_i be drawn from the urn at times determined by a rate-γ_i Poisson process (and independently of all other vertices). After c_i events of the Poisson process, v_i is removed from the urn.

We record the outcome of the urn process by letting $D_i^{(h)}$ be the number of times that v_i occurs among the first h times that a vertex is drawn from the urn. Notice that the distribution of $D_i^{(h)}$ depends on the weights as well as on the capacities, but that it is independent of whether we regard time as discrete or continuous. The crucial result is the following.

PROPOSITION 7.1. *For every integer h such that* $0 \leqslant h \leqslant \sum_i c_i$,

$$E\delta_i^{(h/2)} = ED_i^{(h)}.$$

Proof. It follows from the results of §6 that the only way a vertex can have non-integer degree in the minimum $(h/2)$-flow is if for $i \neq j$, we have $\delta_i^{((h+1)/2)} = \delta_i^{((h-1)/2)} + 1$, $\delta_j^{((h+1)/2)} = \delta_j^{((h-1)/2)} + 1$ and $\delta^{(h/2)} = \frac{1}{2}(\delta^{((h-1)/2)} + \delta^{((h+1)/2)})$. In this case, $\sigma^{((h-1)/2)}$ and $\sigma^{((h+1)/2)}$ are both stable. We define a new random variable $\varepsilon^{(h/2)}$ which is equal to $\delta^{(h/2)}$ except that in the case of non-integer degrees above, we either choose $\varepsilon_i^{(h/2)} = \delta_i^{(h/2)} - \frac{1}{2}$ and $\varepsilon_j^{(h/2)} = \delta_j^{(h/2)} + \frac{1}{2}$ or vice versa, by tossing a coin. Then $\varepsilon^{(h/2)}$ takes only integer values. Whenever $\sigma^{(h/2)}$ is stable, $\varepsilon^{(h/2)} = \delta^{(h/2)}$, and for every h and i, $E\varepsilon_i^{(h/2)} = E\delta_i^{(h/2)}$. We claim that $\varepsilon^{(h/2)}$ has the same distribution as $D^{(h)}$. In the oracle process, suppose that we take into account only the answers given by the oracle about which vertices that are endpoints of the alternating path. Suppose further that when two endpoints are to be revealed, a coin is flipped to decide which one is revealed first. Then, if we condition on $\varepsilon_i^{((h-1)/2)}$ for every i, then $\varepsilon^{(h/2)}$ is obtained by choosing j among the vertices for which $\varepsilon_j^{((h-1)/2)} < c_j$, with probabilities proportional to the weights, and then letting $\varepsilon_j^{(h/2)} = \varepsilon_j^{((h-1)/2)} + 1$, and $\varepsilon_i^{(h/2)} = \varepsilon_i^{((h-1)/2)}$ for $i \neq j$. It is clear from the definition of the urn process that $D^{(h)}$ satisfies the same recursion. Hence $E\delta_i^{(h/2)} = E\varepsilon_i^{(h/2)} = ED_i^{(h)}$. □

Let $T^{(h)}$ be the hth time at which a vertex is drawn in the continuous time urn process.

LEMMA 7.2.
$$\lim_{\gamma_{n+1}\to 0} \frac{E[\delta_{n+1}^{(h/2)}(\mathbf{c}+\mathbf{1}_{n+1})]}{\gamma_{n+1}} = ET^{(h)}(\mathbf{c}).$$

Proof. By Proposition 7.1,

$$E[\delta_{n+1}^{(h/2)}(\mathbf{c}+\mathbf{1}_{n+1})] = ED_{n+1}^{(h)}(\mathbf{c}+\mathbf{1}_{n+1}).$$

We therefore have to prove the identity

$$\lim_{\gamma_{n+1}\to 0} \frac{ED_{n+1}^{(h)}(\mathbf{c}+\mathbf{1}_{n+1})}{\gamma_{n+1}} = ET^{(h)}(\mathbf{c}). \tag{12}$$

This identity concerns the urn process only. We have

$$ED_{n+1}^{(h)}(\mathbf{c}+\mathbf{1}_{n+1}) = P[t < T^{(h)}(\mathbf{c})],$$

where t is the time at which v_{n+1} is drawn for the first time. This is equal to

$$E[1-e^{-\gamma_{n+1}T^{(h)}(\mathbf{c})}]. \tag{13}$$

The left-hand side of (12) is the (right) derivative of (13) as $\gamma_{n+1}\to 0^+$. By differentiating, this is equal to

$$T^{(h)}(\mathbf{c})e^{-\gamma_{n+1}T^{(h)}(\mathbf{c})},$$

which tends to the limit $T^{(h)}(\mathbf{c})$ as $\gamma_{n+1}\to 0^+$. □

7.3. The 2-dimensional urn process

Suppose that the number n of vertices, the weights γ_i and the capacities c_i are given. The results of §5 and §7.2 lead to a "formula" for EC_k. This formula is expressed in terms of a 2-dimensional version of the urn process. In this version, there are two independent urn processes on the set of vertices. The processes take place in independent directions that we label the x- and z-axes, in a 2-dimensional time plane. For each vertex v_i, we let $P_i(x)$ be the number of times that the vertex v_i has been drawn in the first process up to time x. Similarly, $Q_i(z)$ is the number of times that v_i has been drawn in the second urn process up to time z. We define the rank of the process for the single vertex v_i at time (x,z) by

$$\mathrm{Rank}_i(x,z) = \min\{P_i(x), c_i\} + \min\{P_i(x)+Q_i(z), c_i\}.$$

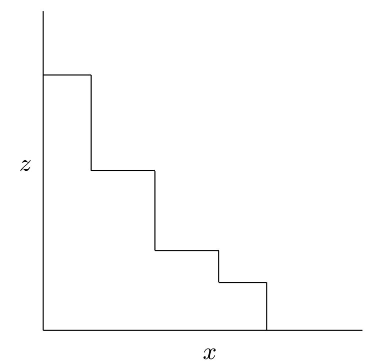

Figure 4. The typical shape of the region $R_h(\mathbf{c})$.

The total rank of the process is defined by

$$\mathrm{Rank}(x,z)=\sum_{i=1}^{n}\mathrm{Rank}_i(x,z).$$

We let $R_h=R_h(\mathbf{c})$ be the region in the positive quadrant of the xz-plane for which $\mathrm{Rank}(x,z)<h$, see Figure 4.

The following theorem gives an exact characterization of EC_k for half-integral $k=\frac{1}{2}h$.

Theorem 7.3.

$$EC_{h/2}=E\left[\mathrm{area}(R_h)\right].$$

Proof. This follows inductively from (10) and Proposition 7.1. Suppose by induction that $EC_{(h-1)/2}(\mathbf{c}-\mathbf{1}_i)=E[\mathrm{area}(R_{h-1}(\mathbf{c}-\mathbf{1}_i))]$ for $1\leqslant i\leqslant n$. Then by (10) and Lemma 7.2,

$$EC_{h/2}=\frac{1}{\gamma_1+\ldots+\gamma_n}ET^{(h)}+E[\mathrm{area}(R_{h-1}(\mathbf{c}-\mathbf{1}_i))]. \tag{14}$$

We have to show that the right-hand side of (14) is equal to $E[\mathrm{area}(R_h)]$. In the first of the two urn processes (the x-process), let x_0 be the first time at which a vertex is drawn. The expected area of the part of R_h that lies in the strip $0<x<x_0$ is

$$\frac{ET^{(h)}}{\gamma_1+\ldots+\gamma_n},$$

which is the first term in the right-hand side of (14).

The probability that v_i is the first vertex to be drawn in the x-process is

$$\frac{\gamma_i}{\gamma_1+\ldots+\gamma_n}.$$

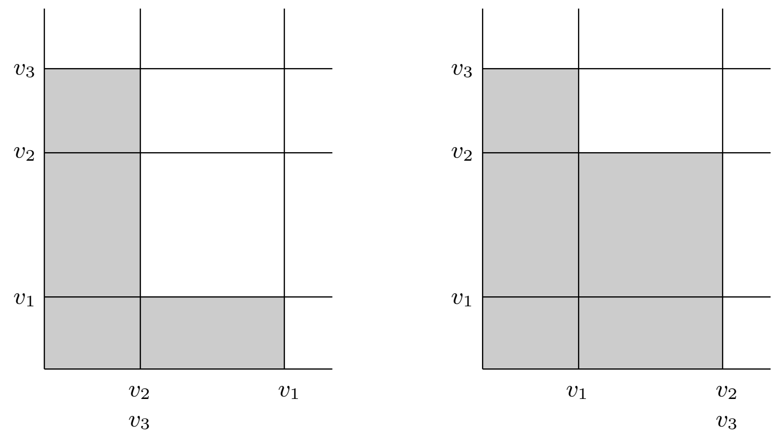

Figure 5. The urn process.

If this happens, then the expected area of the remaining part of R_h (for which $x>x_0$) is equal to

$$E[\text{area}(R_{h-1}(\mathbf{c}-\mathbf{1}_i))],$$

which is the second term of (14). □

7.4. Example

We continue the discussion of the example of §4.3. By symmetry, we may assume that in the z-process, the vertices are drawn in the order v_1, v_2, v_3. The region R_3 consists of four or five "boxes", depending on whether or not the first vertex to be drawn in the x-process is v_1, see Figure 5.

The expected areas of the boxes are indicated in Figure 6. In this example, the areas of the boxes are independent of whether or not they belong to R_3. Since the box of expected area $\frac{1}{4}$ belongs to R_3 with probability $\frac{1}{3}$, the expected cost of the minimum solution is equal to

$$\frac{1}{9}+\frac{1}{6}+\frac{1}{3}+\frac{1}{6}+\frac{1}{3}\cdot\frac{1}{4}=\frac{31}{36}.$$

8. Estimates of the expected cost of the minimum flow

In this section, we obtain estimates of the expected area of R_h, and thereby of EC_k. We are going to apply these results to the random TSP with edge costs satisfying (1). Therefore we assume that each vertex has weight 1 and capacity at most 2. In principle,

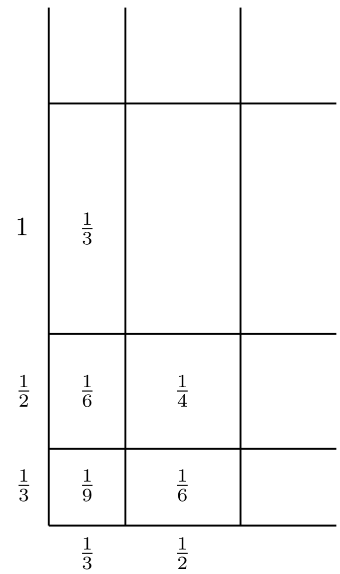

Figure 6. Expected areas of boxes.

Theorem 8.1 below is valid for an arbitrary upper bound on the capacities, but the implied constant will depend on this upper bound.

8.1. The area of the region R_h

Suppose that the positive integers n and $h=2k$ are given. For $i=1,2$, let n_i be the number of vertices of capacity at least i. We assume that $n=n_1$, in other words, there is no vertex of zero capacity. We wish to estimate the expected area of the random region $R=R_h(n_1,n_2)$ given by the points (x,z) for which

$$\operatorname{Rank}(x,z) < h.$$

Recall that the rank is defined by

$$\operatorname{Rank}(x,z) = \sum_{i=1}^{n} \operatorname{Rank}_i(x,z),$$

where

$$\operatorname{Rank}_i(x,z) = \min\{P_i(x), c_i\} + \min\{P_i(x)+Q_i(z), c_i\}.$$

We must assume that $h \leqslant n_1+n_2$, since otherwise the flow problem has no feasible solution, and the region R has infinite area.

We let $R^\star = R^\star_h(n_1, n_2)$ be the non-random region in the xz-plane given by

$$E(\mathrm{Rank}(x,z)) \leqslant h.$$

Our goal is to obtain the following upper bound on the difference between the area of $R^\star$ and the expected area of R.

THEOREM 8.1. *If $\gamma_i = 1$ and $c_i \leqslant 2$ for every i, then*

$$|E(\mathrm{area}(R)) - \mathrm{area}(R^\star)| = O\biggl(\frac{(\log n)^{3/2}}{n^{1/2}}\biggr).$$

This section is mainly devoted to the proof of Theorem 8.1. The following lemma shows that when estimating the expected area of R, it suffices to consider the area of $R \cap B$ for a sufficiently large rectangular box B.

LEMMA 8.2. *Let $x, y > 0$ and let B be the rectangular box with sides $[0, x]$ and $[0, y]$. Then*

$$E(\mathrm{area}(R\cap B)) \leqslant E(\mathrm{area}(R)) \leqslant \frac{E(\mathrm{area}(R\cap B))}{P(R\subseteq B)}.$$

Proof. The first inequality is trivial. To prove the second inequality, let B' be the strip $[0, x] \times [0, \infty]$. We have

$$\begin{aligned} E[\mathrm{area}(R)] &\leqslant E[\mathrm{area}(R\cap B')] + P((x,0)\in R)E[\mathrm{area}(R\setminus B') \mid (x,0)\in R] \\ &\leqslant E[\mathrm{area}(R\cap B')] + P((x,0)\in R)E[\mathrm{area}(R)], \end{aligned}$$

which can be rearranged as

$$E[\mathrm{area}(R)] \leqslant \frac{E[\mathrm{area}(R\cap B')]}{P((x,0)\notin R)}.$$

A similar argument shows that

$$E[\mathrm{area}(R\cap B')] \leqslant \frac{E[\mathrm{area}(R\cap B)]}{P((0,y)\notin R)}.$$

Consequently,

$$E[\mathrm{area}(R)] \leqslant \frac{E[\mathrm{area}(R\cap B)]}{P((x,0)\notin R)P((0,y)\notin R)} = \frac{E(\mathrm{area}(R\cap B))}{P(R\subseteq B)}. \qquad \square$$

In the following, we specifically choose B to be the box with sides $[0, 2]$ and $[0, 2\log n]$. To motivate this choice, we show that, with high probability, $R \subseteq B$. We shall use the following Chernoff-type bound, whose proof we omit.

LEMMA 8.3. *Suppose that $X = X_1 + \ldots + X_n$ is a sum of n independent variables that take the values* 0 *or* 1. *Let $\delta > 0$. Then*

$$P(X - EX \geqslant \delta) \leqslant e^{-\delta^2/2n}.$$

For non-negative x and z, and $l = 1, 2, 3, 4$, let $\theta_l(x, z)$ be the number of i for which $\mathrm{Rank}_i(x, z) \geqslant l$. Then

$$\mathrm{Rank}(x, z) = \theta_1(x, z) + \theta_2(x, z) + \theta_3(x, z) + \theta_4(x, z).$$

The point is that each θ_l is a sum of n independent 0, 1-variables.

LEMMA 8.4.

$$P((2, 0) \in R) \leqslant 2e^{-n/512}.$$

Proof. We have $h \leqslant \sum_{i=1}^n c_i$, and therefore

$$P((2, 0) \in R) = P(\mathrm{Rank}(2, 0) < h) \leqslant P\left(\mathrm{Rank}(2, 0) < \sum_{i=1}^n c_i\right).$$

Since $Q_i(0) = 0$,

$$\mathrm{Rank}(2, 0) = 2 \sum_{i=1}^n \min\{P_i(2), c_i\} = 2(\theta_2(2, 0) + \theta_4(2, 0)).$$

We claim that

$$E[\min\{P_i(2), c_i\}] \geqslant (1 - e^{-1})c_i > \tfrac{5}{8} c_i.$$

This is seen by dividing the interval $[0, 2]$ into c_i subintervals of length at least 1, and then counting the intervals that contain some event where v_i is drawn in the urn process. Consequently

$$E(\mathrm{Rank}(2, 0)) \geqslant \tfrac{5}{4} n \geqslant h + \tfrac{1}{4} n. \tag{15}$$

If $\theta_2(2, 0) + \theta_4(2, 0) < \frac{1}{2} h$, then either $\theta_2(2, 0)$ or $\theta_4(2, 0)$ must be smaller by at least $\frac{1}{16} n$ than their expected value. Since each $\theta_l(2, 0)$ is a sum of n independent 0, 1-variables, Lemma 8.3 shows that

$$P\left[\theta_2(2, 0) + \theta_4(2, 0) < \tfrac{1}{2} h\right] \leqslant 2e^{-n/512}. \qquad \square$$

LEMMA 8.5.

$$P((0, 2\log n) \in R) = O\left(\frac{\log n}{n}\right).$$

Proof. The probability that there is some vertex that is not drawn at least two times up to time z is at most

$$n(1+z)e^{-z},$$

and if $z=2\log n$ then this is equal to

$$\frac{1+2\log n}{n}. \qquad \square$$

Lemmas 8.2, 8.4 and 8.5 together imply the following result.

COROLLARY 8.6.

$$E(\text{area}(R)) = E(\text{area}(R\cap B))\left(1+O\left(\frac{\log n}{n}\right)\right) = E(\text{area}(R\cap B))+O\left(\frac{(\log n)^2}{n}\right).$$

The last equation comes from the fact that $\text{area}(R\cap B)\leqslant\text{area}(B)=O(\log n)$.

LEMMA 8.7. *For non-negative x and z, and $\varepsilon>0$, the probability that* $\text{Rank}(x,z)$ *deviates by at least εn from its expected value is at most*

$$4e^{-\varepsilon^2 n/32}.$$

Proof. If $\text{Rank}(x,z)$ deviates by at least εn from its expected value, then one of $\theta_l(x,z)$ for $l=1,...,4$ must deviate by at least $\frac{1}{4}\varepsilon n$ from its expected value. Since each $\theta_l(x,z)$ is a sum of n independent $0,1$-variables, it follows from Lemma 8.3 that the probability for such a deviation is at most $4e^{-\varepsilon^2 n/32}$. $\square$

We make the specific choice of ε so that

$$e^{-\varepsilon^2 n/32} = \frac{1}{n^2},$$

in other words we put

$$\varepsilon = \frac{8(\log n)^{1/2}}{n^{1/2}}.$$

We divide B into three regions B_1, B_2 and B_3 according to whether (for a point (x,z)) $E(\text{Rank}(x,z))$ is smaller than $h-\varepsilon n$, between $h-\varepsilon n$ and $h+\varepsilon n$, or greater than $h+\varepsilon n$, see Figure 7.

LEMMA 8.8.

$$\text{area}(B_2) = O\left(\frac{(\log n)^{3/2}}{n^{1/2}}\right).$$

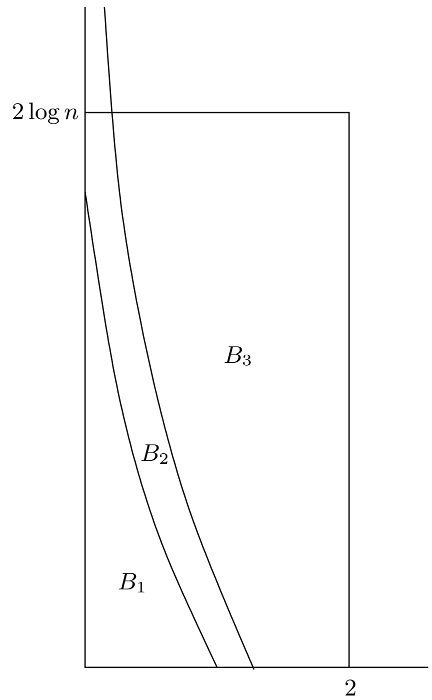

Figure 7. The box B and the regions B_1, B_2 and B_3.

Proof. We show that at a given height z, the expected rank cannot stay between $h-\varepsilon n$ and $h+\varepsilon n$ for too long. We have

$$\frac{d}{dx}E(\mathrm{Rank}(x,z)) \geqslant \frac{d}{dx}\sum_{i=1}^{n} E\min\{P_i(x),c_i\}.$$

For each i,

$$\frac{d}{dx}E\min\{P_i(x),c_i\} = \begin{cases} e^{-x}, & \text{if } c_i=1,\\ (1+x)e^{-x}, & \text{if } c_i=2,\end{cases}$$

is decreasing in x, so that the minimum value occurs for $x=2$ and $c_i=1$. Therefore, inside B, we have

$$\frac{d}{dx}E\min\{P_i(x),c_i\} \geqslant e^{-2}, \tag{16}$$

and

$$\frac{d}{dx}E(\mathrm{Rank}(x,z)) \geqslant ne^{-2}.$$

It follows that, for a fixed z, the width of B_2 at height z is at most

$$\frac{2n\varepsilon}{ne^{-2}} = 2e^2\varepsilon = \frac{16e^2(\log n)^{1/2}}{n^{1/2}}.$$

Since the height of B is $2\log n$, the statement follows. □

LEMMA 8.9.

$$E[\operatorname{area}(R\cap B)]=\operatorname{area}(R^\star\cap B)+O\left(\frac{(\log n)^{3/2}}{n^{1/2}}\right).$$

Proof. In B_1, the probability that a point belongs to R is at least $1-4/n^2$, while in B_3 this probability is at most $4/n^2$. We conclude that the expected area of $R\cap B$ deviates from the area of $R^\star\cap B$ by at most

$$\operatorname{area}(B_2)+\frac{4}{n^2}\operatorname{area}(B)=O\left(\frac{(\log n)^{3/2}}{n^{1/2}}\right)+O\left(\frac{\log n}{n^2}\right)=O\left(\frac{(\log n)^{3/2}}{n^{1/2}}\right). \qquad \square$$

LEMMA 8.10.

$$\operatorname{area}(R^\star\setminus B)=O\left(\frac{\log n}{n^2}\right).$$

Proof. It follows from (15) that $R^\star$ lies entirely in the region $x<2$. For a single vertex v_i, the remaining capacity at time z, that is, $c_i-Q_i(z)$, is larger when the capacity is larger. To prove the lemma, we may therefore assume that $c_i=2$ for every i.

The average rank $E(\operatorname{Rank}_i(0,z))$ for a single vertex at time z is

$$\int_0^z(1+t)e^{-t}\,dt,$$

since the rank increases the first and the second time the vertex is drawn. Therefore the expected remaining capacity at time z is

$$E[c_i-Q_i(z)]=\int_z^\infty(1+t)e^{-t}\,dt=(2+z)e^{-z}.$$

By (16), the area of $R^\star\setminus B$ is at most

$$e^2\int_{2\log n}^\infty(2+z)e^{-z}\,dz=e^2\frac{3+2\log n}{n^2}. \qquad \square$$

Proof of Theorem 8.1. We have

$$\begin{aligned}|E(\operatorname{area}(R))-\operatorname{area}(R^\star)|\leqslant{}&|E(\operatorname{area}(R))-\operatorname{area}(R\cap B)|+|E(\operatorname{area}(R\cap B))-\operatorname{area}(R^\star\cap B)|\\&+|E(\operatorname{area}(R^\star\cap B))-\operatorname{area}(R^\star)|.\end{aligned}$$

From Corollary 8.6 and Lemmas 8.9 and 8.10 it follows that this is

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

8.2. Asymptotic estimates

For problems like relaxed matching and 2-factor that scale with n under fixed local constraints, the region $R^\star$ is independent of n. Therefore $R^\star$ is a natural limit of R_n, and by the results of §8.1, the expected cost of the optimization problem converges to the area of $R^\star$.

For the relaxed matching problem, we obtain $EC_{n/2}\to\frac{1}{12}\pi^2$, but this can also be established by an explicit calculation, as is shown by (54) in §13. Another consequence is that the expected cost of the relaxed 2-factor problem converges to the number $L^\star$ defined in the introduction.

THEOREM 8.11. *If $\gamma_i=1$ and $c_i=2$ for every i, then the area of the region $R^\star$ is the number $L^\star$ defined by* (2) *and* (3). *Hence, if C_n is the cost of the minimum relaxed 2-factor, then, as $n\to\infty$, $EC_n\to L^\star$.*

Proof. Let x and z be non-negative real numbers. The region $R^\star$ is defined by the inequality

$$E[\min\{P_i(x),2\}+\min\{P_i(x)+Q_i(z),2\}]\leqslant 2. \tag{17}$$

$P_i(x)$ and $Q_i(z)$ are the number of events in two independent rate-1 Poisson processes. This means that $P_i(x)$ is 0, 1 or at least 2, with probabilities e^{-x}, xe^{-x} and $1-e^{-x}-xe^{-x}$, respectively, and that the distribution of $Q_i(z)$ is given in the same way by z.

Since the two Poisson processes are independent, $P_i(x)+Q_i(z)$ has the same distribution as $Q_i(x+z)$. By the variable substitution $y=x+z$ and linearity of expectation, it follows that (17) is equivalent to

$$E\min\{P_i(x),2\}+E\min\{Q_i(y),2\}\leqslant 2.$$

Here

$$E\min\{P_i(x),2\}=0\cdot e^{-x}+1\cdot xe^{-x}+2\cdot(1-e^{-x}-xe^{-x})=2-2e^{-x}-xe^{-x},$$

and similarly for $Q_i(y)$. Therefore (17) can be written

$$4-2e^{-x}-xe^{-x}-2e^{-y}-ye^{-y}\leqslant 2.$$

Equality holds when

$$\left(1+\frac{x}{2}\right)e^{-x}+\left(1+\frac{y}{2}\right)e^{-y}=1, \tag{18}$$

and the region $R^\star$ is therefore defined by this equation together with the boundaries $x=0$ and $z=0$, the latter being equivalent to $x=y$. The change of variables from (x,z) to (x,y) is area-preserving. Since in the xy-plane $R^\star$ lies above the line $x=y$, its area is by symmetry half of the total area in the positive quadrant of the xy-plane under the curve given by (18). This means that the area of $R^\star$ is equal to the right-hand side of (2). $\square$

Since the relaxed 2-factor problem is a relaxation of the TSP, Theorem 8.11 already provides a lower bound on the expected cost of the TSP as a first step towards proving Theorem 1.1. Other problems for which the region $R^\star$ is essentially independent of n include incomplete matching and 2-factor problems, if k scales with n like pn for some constant p. A calculation similar to the one above shows that for matching, the expected cost converges to the area under the curve

$$e^{-x}+e^{-(x+z)}=2-2p, \quad 0\leqslant p\leqslant \tfrac{1}{2},$$

while for the incomplete 2-factor problem, the region is given by

$$\left(1+\frac{x}{2}\right)e^{-x}+\left(1+\frac{x+z}{2}\right)e^{-(x+z)}=2-p, \quad 0\leqslant p\leqslant 1. \tag{19}$$

Notice that for the incomplete problems, the limit region is bounded.

Proposition 8.12. *The area of the region given by* (19) *is continuous as a function of* p *on the interval* $0\leqslant p\leqslant 1$.

The area jumps to infinity for $p>1$, but the fact that it tends to $L^\star$ as p approaches 1 from the left has the following consequence for the cost of an incomplete 2-factor.

Corollary 8.13. *Consider the incomplete relaxed* 2*-factor problem. If* $k<n$ *but* k *scales with* n *in such a way that* $k/n\to 1$ *as* $n\to\infty$, *then*

$$EC_k\to L^\star.$$

9. Concentration

In this section we derive a concentration inequality for flow problems by a straightforward application of the Talagrand inequality [42]. As in §8, we assume that $\gamma_i=1$ and $c_i\leqslant 2$ for every i. We consider the linear flow problem. In principle, Talagrand's method applies also to integer flow problems, but they lead to some additional technical difficulties. A consequence of the results of §10 is that the concentration inequality for linear problems partly carries over to integer problems, and we thereby avoid these difficulties.

We are primarily interested in showing that the variance of the cost of certain flow problems tends to zero as $n\to\infty$. Throughout this section, $A_1, A_2, A_3, \ldots$ will denote positive constants for which numerical values could in principle be substituted, but whose actual values are not important.

Theorem 9.1. *If* $\gamma_i=1$ *and* $c_i\leqslant 2$ *for every* i, *then for the linear flow problem*, *and for every* k *that allows a solution*,

$$\operatorname{var}(C_k)=O\left(\frac{(\log n)^5}{n}\right). \tag{20}$$

By using the sharpest bounds available, one could probably obtain an $O((\log n)^2/n)$ bound, but we try to keep the argument simple without bothering about the exponent of the logarithm. We conjecture that the right-hand side of (20) can in reality be replaced by $O(1/n)$.

To apply Talagrand's method, we need a bound on the probability that the minimum flow uses some extremely expensive edge. Let $X_{\max}$ be the cost of the most expensive edge that has non-zero coefficient in $\sigma^{(k)}$. Our first objective is to prove the following.

Proposition 9.2. *If both n and z are sufficiently large, then*

$$P\left(X_{\max} > \frac{z(\log n)^2}{n}\right) \leqslant n^{-A_1 z}.$$

The expectation of $X_{\max}$ is at most

$$2E(C_k - C_{k-1/2}),$$

which is bounded by $O(\log n/n)$, but we need a stronger bound on large deviations of $X_{\max}$ than what follows from the expectation only.

To establish Proposition 9.2, we need some preliminary results. For $z>0$, let $D_z \subseteq E$ be the set of edges of cost at most $z \log n/n$. When necessary, we shall assume that z is sufficiently large. We begin by showing that with high probability, D_z has a certain expander property.

For $S \subseteq V$, let S' denote the set of D_z-neighbors of S, that is, the set of vertices that are connected to S by an edge in D_z.

Lemma 9.3. *Let S be a set of vertices chosen independently of the edge costs. If*

$$|S| \leqslant \frac{3n}{z \log n}, \tag{21}$$

then

$$|S'| \geqslant 0.3z(\log n)|S|, \tag{22}$$

with probability at least $1-n^{-A_2 z|S|}$.

Here the numbers 3 and 0.3 are chosen sufficiently large so that $1-e^{-3} > 3\cdot 0.3 > \frac{4}{5}$. It will become clear later why this is necessary. If the upper bound on c_i was larger, then 3 and 0.3 would have to be replaced by larger numbers.

Proof. Let $s=|S|$. For each vertex v we have

$$P(v \in S') = 1 - P(v \notin S') = 1 - e^{-sz(\log n)/n}.$$

From (21) it follows that $sz(\log n)/n \leqslant 3$, and therefore

$$P(v\in S') \geqslant \frac{1-e^{-3}}{3}\frac{sz\log n}{n}.$$

Let $p=P(v\in S')$. We have, for every $\lambda>0$,

$$E[e^{-\lambda|S'|}] = (pe^{-\lambda}+1-p)^n \leqslant e^{-np(1-e^{-\lambda})}.$$

By the Chebyshev inequality, and using the fact that $0.3\cdot 3/(1-e^{-3})<0.95$, we find that

$$\begin{aligned}
P(|S'|<0.3sz\log n) &\leqslant P\left(|S'|<\frac{0.3\cdot 3}{1-e^{-3}}pn\right)\\
&\leqslant P(|S'|<0.95pn)\\
&= P(e^{-\lambda|S'|}>e^{-0.95\lambda pn})\\
&\leqslant \frac{e^{-np(1-e^{-\lambda})}}{e^{-0.95\lambda pn}}\\
&= e^{-pn(1-e^{-\lambda}-0.95\lambda)}\\
&\leqslant \exp\left(-\frac{1-e^{-3}}{3}sz(\log n)(1-e^{-\lambda}-0.95\lambda)\right).
\end{aligned}$$

We now put $\lambda=1-0.95=0.05$. This gives

$$\frac{1-e^{-3}}{3}(1-e^{-\lambda}-0.95\lambda) > 0.0004.$$

Hence Lemma 9.3 holds with $A_2=0.0004$. □

LEMMA 9.4. *If z is sufficiently large, then with probability at least $1-n^{-A_3\cdot z}$, every set $S\subseteq V$ of vertices satisfies*

$$|S'| \geqslant \min\{0.9n, 0.3z(\log n)|S|\}. \tag{23}$$

Proof. It is sufficient to consider only sets of cardinality at most $|S|\leqslant 3n/z\log n$, and to show that with the desired probability, each of these sets satisfies (22). For larger sets, we can choose a subset of size $3n/z\log n$, rounded down. The number of sets S of size s is at most n^s. Therefore the probability that there is some set S of size s that does not satisfy (22) is at most

$$n^s e^{-A_2 sz\log n} = e^{-A_2 sz\log n + s\log n} \leqslant e^{-A_4 sz\log n},$$

if z is sufficiently large. Moreover,

$$\sum_{s=1}^{\infty} e^{-A_4 sz\log n} \leqslant e^{-A_3 z\log n} = n^{-A_3 z}.$$ □

Proof of Proposition 9.2. Fix n and z. Since we allow a failure probability of $n^{-A_1 z}$, we may assume that D_z has the desired expander property, that is, that every set S of vertices satisfies (23). Under this assumption we want to show that the maximum edge cost $X_{\max}$ in the minimum k-flow is at most $z(\log n)^2/n$.

Let σ be the minimum k-flow. All coefficients of σ are 0, $\frac{1}{2}$ or 1. Since $c_i \leqslant 2$, this implies that each vertex is incident to at most four edges of non-zero coefficient in σ (actually at most three by the results of §6, but any uniform bound will do).

Let e be an edge of non-zero coefficient in σ and let v_i and v_j be the endpoints of e (possibly, $i=j$). We want to construct a sequence of vertices $x_0, y_0, x_1, y_1, ..., x_m, y_m$ which constitutes an alternating path with respect to σ and D_z, and such that x_0 and y_m either belong to the set $\{v_i, v_j\}$ or do not have full degree in σ. We require that $(x_i, y_i) \in D_z$ for $0 \leqslant i \leqslant m$, and that (x_i, y_{i+1}) has non-zero coefficient in σ for $0 \leqslant i < m$.

We let $S_0 \subseteq V$ consist of v_i, v_j and all vertices that do not have full degree in σ. Then let S_i and T_i be the sets of vertices that can be reached by a path as the one above of length at most i. More precisely, supposing that S_i has been defined, let T_i be the set of $v \in V$ such that for some $u \in S_i$, there is an edge between u and v that belongs to D_z. Then supposing that T_i has been defined, let S_{i+1} be the set of $v \in V$ such that either $v \in S_0$ or there is a $u \in T_i$ and an edge between u and v that has non-zero coefficient in σ.

Suppose that T_i contains only vertices of full degree in σ. Then each $u \in T_i$ is incident to at least one edge in σ. Since on the other hand each vertex is incident to at most four edges in σ, we have

$$|S_{i+1}| \geqslant \tfrac{1}{4}|T_i|.$$

A consequence of this is that if

$$|T_i| > \tfrac{4}{5}n,$$

then T_i intersects S_{i+1}, and we can construct a path as above with $m=2i+1$. In view of (23), this will clearly happen for some $m=O(\log n)$.

Even if there are repetitions of the same vertex or edge in this sequence, there will be an $\varepsilon>0$ such that by "switching" (increasing and decreasing) the coefficients of the edges in the sequence by ε, we obtain a new flow. In particular, if the edge e has cost larger than the sum of the costs of the $m+1$ edges in D_z, then σ is not the minimum k-flow. If such a sequence can be found for every flow and every edge e, then

$$X_{\max} \leqslant \frac{(m+1)z\log n}{n} = O\left(\frac{z(\log n)^2}{n}\right).$$

This shows that there are constants A_5 and A_6 such that

$$P\left(X_{\max} > \frac{A_5 z(\log n)^2}{n}\right) \leqslant n^{-A_6 z}.$$

By rescaling, this is clearly equivalent to Proposition 9.2. □

We now turn to the proof of Theorem 9.1. We quote [42, Theorem 8.1.1], or rather the special case used in [42, §10] on the assignment problem. Let $v>0$ and let $Y_1, ..., Y_m$ be independent random variables with arbitrary distribution on the interval $[0, v]$. Let

$$Z = \min_{\lambda \in \mathcal{F}} \sum_{i \leqslant m} \lambda_i Y_i,$$

where $\mathcal{F}$ is a family of vectors in $\mathbb{R}^m$ (in our case the λ_i's are non-negative, but Talagrand's theorem holds without this assumption). Let

$$\varrho = \max_{\lambda \in \mathcal{F}} \sum_{i \leqslant m} \lambda_i^2.$$

Moreover, let M be a median for the random variable Z.

Theorem 9.5. (Talagrand [42]) *For every* $w>0$,

$$P(|Z-M| \geqslant w) \leqslant 4e^{-w^2/4\varrho v^2}.$$

When we apply this theorem to the random k-flow problem, the λ_i's will be at most 2, where the coefficients that are equal to 2 come from the possibility of using multiple edges or loops. Since the capacities are at most 2, a flow contains at most n edges. Although the coefficients λ_i cannot all be equal to 2, for simplicity we use the bound $\varrho \leqslant 4n$.

If we take $v=1$, the resulting bound will be too weak to be interesting. We therefore modify the problem by replacing the edge costs Y by $\min\{Y, \zeta(\log n)^2/n\}$. Let M_ζ be a median of the cost Z_ζ in the modified problem. It follows from Theorem 9.5 with $v=\zeta(\log n)^2/n$ that, for all $w>0$,

$$P(|Z_\zeta - M_\zeta| \geqslant w) \leqslant 4e^{-w^2 n/16\zeta^2 (\log n)^4},$$

and consequently, by Proposition 9.2, for all sufficiently large ζ,

$$P(|Z - M_\zeta| \geqslant w) \leqslant 4e^{-w^2 n/16\zeta^2 (\log n)^4} + n^{-A_7 \zeta}.$$

By taking

$$\zeta = \frac{w^{2/3} n^{1/3}}{(\log n)^{5/3}},$$

we obtain

$$\begin{aligned} P(|Z - M_\zeta| \geqslant w) &\leqslant 4 \exp\left(-\frac{w^{2/3} n^{1/3}}{16(\log n)^{2/3}}\right) + \exp\left(-\frac{A_7 w^{2/3} n^{1/3}}{(\log n)^{2/3}}\right) \\ &\leqslant 5 \exp\left(-\frac{A_8 w^{2/3} n^{1/3}}{(\log n)^{2/3}}\right). \end{aligned}$$

The requirement that ζ should be sufficiently large is equivalent to assuming that w is at least a certain constant times $n^{-1/2}(\log n)^{5/2}$. We therefore introduce yet another parameter t and put

$$w=\frac{t(\log n)^{5/2}}{n^{1/2}}.$$

We then conclude that, for all sufficiently large t,

$$P\left(|Z-M_\zeta|\geqslant\frac{t(\log n)^{5/2}}{n^{1/2}}\right)\leqslant 5e^{-A_8t^{2/3}\log n}. \tag{24}$$

A small remaining problem is that M_ζ depends on t. Equation (24) therefore only says that for every t, there is some interval of length

$$\frac{2t(\log n)^{5/2}}{n^{1/2}}$$

that contains Z with probability at least

$$1-5e^{-A_8t^{2/3}\log n}.$$

Since t has to be larger than some absolute constant, we may assume that

$$5e^{-A_8t^{2/3}\log n}<\tfrac{1}{2}.$$

Then the interval must also contain the median M of Z. We therefore conclude that, for sufficiently large t,

$$P\left(|Z-M|\geqslant\frac{2t(\log n)^{5/2}}{n^{1/2}}\right)\leqslant 5e^{-A_8t^{2/3}\log n}.$$

Here it is clear that the right-hand side tends to zero rapidly enough to give an $O((\log n)^{5/2}/n^{1/2})$ bound on the standard deviation of Z. This completes the proof of Theorem 9.1.

10. Integer flow problems

In this section we treat the integer flow problem in the friendly model. Unfortunately we are unable to obtain exact formulas for the expected cost of the solution, but we establish bounds that are good enough for determining limit costs as $n\to\infty$. The expected cost of the linear flow problem has already been established in Theorem 7.3, and this clearly gives a lower bound on the cost of the corresponding integer flow problem. Letting $\widetilde{C}_k$

here denote the cost of the integer flow problem, we have, in terms of the 2-dimensional urn process, the lower bound

$$E\widetilde{C}_k \geqslant E[\text{area}(R_{2k})].$$

Therefore it only remains to obtain a reasonably good upper bound. We do this by bounding the probability that the extra vertex v_{n+1} participates in the minimum flow with respect to $\mathbf{c}+\mathbf{1}_{n+1}$. In order to apply the same methods as for the linear problem, we need to establish a *nesting lemma* valid for integer flow problems. This is easier, since a solution is just a set of edges and all vertex degrees are integers.

LEMMA 10.1. *Consider the integer flow problem. For every k for which there exists a $(k+1)$-flow, each vertex is incident to at least as many edges in $\sigma^{(k+1)}$ as in $\sigma^{(k)}$. In other words, $\delta_i^{(k)} \leqslant \delta_i^{(k+1)}$.*

Proof. Let $H = \sigma^{(k)} \triangle \sigma^{(k+1)}$ be the symmetric difference of the minimum k- and $(k+1)$-flows, that is, the set of edges that belong to one of them but not to the other. We decompose H into paths and cycles in the following way: At each vertex v, if v has full degree in one of the two flows $\sigma^{(k)}$ and $\sigma^{(k+1)}$, then we pair up the edges (incident to v) of the other flow with these edges. Thus H is decomposed into paths and cycles in such a way that the symmetric difference of any of $\sigma^{(k)}$ and $\sigma^{(k+1)}$ with the union of a number of such paths and cycles is a flow.

By minimality and genericity, no such component can be balanced in the sense of containing equally many edges from $\sigma^{(k)}$ as from $\sigma^{(k+1)}$, since this would imply that either $\sigma^{(k)}$ or $\sigma^{(k+1)}$ could be improved by switching. For the same reason, there cannot be two components of H that together are balanced. The only remaining possibility is that H is a single path whose ends both belong to $\sigma^{(k+1)}$, and this proves the statement. □

When we derive the upper bound, we modify the protocol for the oracle process.

10.1. The protocol

As in §7.1, we ask questions to the oracle in order to successively find the minimum r-flow for $r = 1, ..., k$. Recall that $\sigma^{(r)}$ denotes the minimum r-flow, and that we can now regard $\sigma^{(r)}$ as a set of edges. We let Γ_r be the set of vertices of full degree in $\sigma^{(r)}$. At each stage (for $r = 0, ..., k-1$), the following information is available to us:

O1′. The edges of $\sigma^{(r)}$ and their costs.

O2′. The costs of all edges between vertices in Γ_r. (Strictly speaking, this includes infinitely many edges, but only finitely many are relevant in each problem instance.)

O3′. The costs of a further set of edges arising from collisions (see below). These are edges either from a vertex in Γ_r to a vertex of remaining capacity 1 (whose degree in $\sigma^{(r)}$ is one less than its capacity), or a loop at a vertex of remaining capacity 1. Each such edge is the cheapest edge not in $\sigma^{(r)}$ from its endpoint in Γ_r to a vertex outside Γ_r. In particular, from each vertex in Γ_r there is at most one such edge. If it is a loop, it must be the cheapest edge connecting vertices outside Γ_r.

O4′. For each $v \in \Gamma_r$, the cost of the cheapest edge from v other than the edges already specified in O1′–O3′ (in particular, this edge goes to a vertex outside Γ_r). Also notice that, since $r<k$, by assumption there must be at least one vertex outside Γ_r, otherwise a k-flow would be impossible for the given capacities.

O5′. The cost of the cheapest edge not included under O1′ or O3′ between two vertices not in Γ_r. In other words, the cost of the cheapest edge that does not fall under any of O1′–O4′.

Notice that in O4′ and O5′, it is only the costs that are known, not the locations of the edges of these costs.

Just as in §7 we now compute, using this information, the minimum cost of an $(r+1)$-flow under the assumption that there is no collision between unknown edges. By Lemma 10.1, $\sigma^{(r+1)}$ is obtained from $\sigma^{(r)}$ by switching an alternating path starting and ending at vertices (possibly the same) outside Γ_r.

If one or two of the ends of the alternating path fall under O4′ or O5′, so that the endpoints are unknown, then we ask the oracle to reveal these endpoints to us. The only reason our proposed flow may not be the minimum $(r+1)$-flow is that these two endpoints may be the same, and this vertex may already have degree in $\sigma^{(r)}$ only one less than its capacity.

If the oracle tells us that this is the case, then the ends of the alternating path are "colliding" edges that now fall under O3′ above. (The colliding edges must then be the minimum edges not in $\sigma^{(r)}$ from their endpoints in Γ_r to vertices outside Γ_r.) We further ask the oracle about the minimum cost of the remaining edges from these two vertices (if the collision consisted in a loop, we ask for the minimum cost of the remaining edges not in $\sigma^{(r)}$ connecting vertices not in Γ_r). Then we compute a new proposed minimum $(r+1)$-flow, and repeat until the oracle tells us that no collision takes place.

10.2. Estimate of the probability that v_{n+1} participates

When the oracle tells us that the proposed $(r+1)$-flow is valid, there are essentially three possibilities:

(1) The alternating path ends in two unknown vertices. Then we are conditioning on the event that there is no collision, that is, the two endpoints are chosen independently according to weights and conditioning on not both of them being equal to 1 and the same vertex of remaining capacity 1. Notice that this includes the case that $\sigma^{(r+1)}$ is obtained from $\sigma^{(r)}$ by adding the cheapest edge not in $\sigma^{(r)}$ between two vertices not in Γ_r, and that if the cheapest such edge is a loop of remaining capacity 1, this is treated like an ordinary collision.

(2) The alternating path has one end consisting of an edge whose cost is known according to O3′. Then we are conditioning on the event that the other endpoint is another vertex.

(3) The alternating path ends in two known edges. These edges must then end in vertices having remaining capacity 1.

Suppose that there is an adversary who can choose between these three options. The adversary tries to maximize the probability of v_{n+1} participating in the minimum flow. This can again be modeled by an urn process. Vertices are drawn from an urn two at a time, but with replacement. When some vertices reach remaining capacity 1, that is, they have been drawn a number of times equal to their capacity minus 1, the adversary can choose between three options every time a pair of vertices is drawn:

(1) To condition on the event that the two vertices that are drawn are not one and the same vertex of remaining capacity 1.

(2) To choose a vertex (other than v_{n+1}) of remaining capacity 1 and condition on the other vertex in the pair being different from this one.

(3) To choose two distinct vertices (other than v_{n+1}) of remaining capacity 1.

The procedure of drawing two vertices is repeated k times. The extra vertex v_{n+1} is thought of as having infinitesimal weight $\gamma=\gamma_{n+1}$. We can therefore disregard the possibility that v_{n+1} is drawn twice.

We derive an upper bound on the normalized probability of drawing v_{n+1}, which is valid under the conditions we need in our study of the traveling salesman problem. With similar methods it is certainly possible to derive bounds valid under more general conditions.

Suppose therefore that all vertices except v_{n+1} have weight 1. Further, let n_i denote the number of vertices of capacity at least i, and suppose that there is no vertex of capacity 3 or more. In other words, there are n_1 vertices of capacity 1 or 2, of which n_2 have capacity 2.

An upper bound on the probability of drawing v_{n+1} can be computed as follows: inductively define $I(k, n_1, n_2)$ for integers $0\leqslant n_2\leqslant n_1$ and $0\leqslant k\leqslant \frac{1}{2}(n_1+n_2)$ by

$$I(0, n_1, n_2)=0$$

and, for $k>0$,

$$I(k,n_1,n_2)=\max\left\{\begin{aligned}&\frac{1}{n_1^2-n_1+n_2}(2n_1+(n_1-n_2)(n_1-n_2-1)I(k-1,n_1-2,n_2)\\&\qquad\qquad+(2n_1-2n_2+1)n_2I(k-1,n_1-1,n_2-1)\\&\qquad\qquad+n_2(n_2-1)I(k-1,n_1,n_2-2)),\\&\frac{1}{n_1-1}(1+(n_1-n_2-1)I(k-1,n_1-2,n_2)\\&\qquad+n_2I(k-1,n_1-1,n_2-1)),\\&I(k-1,n_1-2,n_2)\end{aligned}\right\}. \tag{25}$$

Here the first expression in the right-hand side corresponds to option (1). There are n_1^2 ways of choosing two vertices independently, and n_1-n_2 of them consist in choosing the same vertex of capacity 1 twice. The normalized probability of choosing v_{n+1} as one of the two vertices is $2n_1$, and if this does not happen, the remaining terms give (an upper bound on) the probability of choosing v_{n+1} at a later stage. Similarly, the second and third arguments correspond to options (2) and (3).

The second argument should be taken into account only if $n_1\geqslant n_2+1$, and the third argument only if $n_1\geqslant n_2+2$. Under these conditions the right-hand side of (25) makes sense, because terms that are not well defined have coefficient zero. For instance, the term $I(k-1,n_1,n_2-2)$ is well defined only if $n_2\geqslant 2$, but if $n_2<2$ then the coefficient $n_2(n_2-1)$ is zero.

Hence, the normalized probability that v_{n+1} participates in the minimum integer k-flow is at most $I(k,n_1,n_2)$. This is compared with the recursive equation for the expected time, here denoted $J(2k,n_1,n_2)$, until $2k$ vertices have been drawn in the original urn process. We have

$$J(2k,n_1,n_2)=\frac{1+(n_1-n_2)J(2k-1,n_1-1,n_2)+n_2J(2k-1,n_1,n_2-1)}{n_1}. \tag{26}$$

If we recursively expand equation (26) one more step, we get

$$\begin{aligned}J(2k,n_1,n_2)=\frac{1}{n_1}+\frac{n_1-n_2}{n_1(n_1-1)}+\frac{n_2}{n_1^2}+\frac{(n_1-n_2)(n_1-n_2-1)}{n_1(n_1-1)}J(2k-2,n_1-2,n_2)\\+\left(\frac{n_2(n_1-n_2)}{n_1(n_1-1)}+\frac{n_2(n_1-n_2+1)}{n_1^2}\right)J(2k-2,n_1-1,n_2-1)\\+\frac{n_2(n_2-1)}{n_1^2}J(2k-2,n_1,n_2-2).\end{aligned} \tag{27}$$

We wish to establish an upper bound on the quantity

$$\Delta(k,n_1,n_2)=I(k,n_1,n_2)-J(2k,n_1,n_2).$$

PROPOSITION 10.2.

$$\Delta(k, n_1, n_2) \leqslant \frac{2}{n_1+n_2-2k+1}.$$

To prove this, we show inductively that

$$\Delta(k, n_1, n_2) \leqslant \frac{2}{n_1+n_2-2k+1} - \frac{2}{n_1+n_2+1}. \tag{28}$$

If $k=0$, which is the base of the induction, then $\Delta(k, n_1, n_2)=0$, and so is the right-hand side of (28). For the inductive step, we need two lemmas.

LEMMA 10.3. *For all $k>0$ and $n_1>n_2$, we have*

$$J(k-1, n_1-1, n_2) \leqslant J(k, n_1, n_2).$$

Proof. The right-hand side can be interpreted as the expected time until k vertices have been drawn given the capacities that correspond to (n_1, n_2), conditioning on the event that a certain vertex of capacity 1 is drawn at time zero. The fact that a certain vertex is drawn at time zero cannot increase the time until k vertices have been drawn. □

LEMMA 10.4.

$$J(k, n_1-1, n_2) \geqslant J(k, n_1, n_2-1).$$

Proof. We prove this inductively using equation (26). $J(k-1, n_1-1, n_2)$ is equal to $1/(n_1-1)$ plus a convex combination of $J(k-1, n_1-2, n_2)$ and $J(k-1, n_1-1, n_2-1)$, while $J(k-1, n_1, n_2-1)$ is $1/n_1$ plus a convex combination of $J(k-1, n_1-1, n_2-1)$ and $J(k-1, n_1, n_2-2)$. If any of these terms is undefined because $n_1-2<n_2$ or $n_2-2<0$, then the corresponding coefficient will be zero and the argument still holds. By induction,

$$J(k-1, n_1-2, n_2) \geqslant J(k-1, n_1-1, n_2-1) \geqslant J(k-1, n_1, n_2-2).$$

Hence

$$J(k, n_1-1, n_2) \geqslant \frac{1}{n_1-1} + J(k-1, n_1-1, n_2) \geqslant \frac{1}{n_1} + J(k-1, n_1-1, n_2) \geqslant J(k, n_1, n_2-1).$$

□

The induction step of the proof of Proposition 10.2 has three cases depending on which of the three values in the right-hand side of (25) is the largest.

Case 1. Suppose first that

$$\begin{aligned} I(k, n_1, n_2) = {} & \frac{2n_1}{n_1^2-n_1+n_2} + \frac{(n_1-n_2)(n_1-n_2-1)}{n_1^2-n_1+n_2} I(k-1, n_1-2, n_2) \\ & + \frac{(2n_1-2n_2+1)n_2}{n_1^2-n_1+n_2} I(k-1, n_1-1, n_2-1) \\ & + \frac{n_2(n_2-1)}{n_1^2-n_1+n_2} I(k-1, n_1, n_2-2). \end{aligned} \tag{29}$$

Comparing the first terms of the right-hand sides of (27) and (29), we find that

$$\frac{2n_1}{n_1^2-n_1+n_2}-\left(\frac{1}{n_1}+\frac{n_1-n_2}{n_1(n_1-1)}+\frac{n_2}{n_1^2}\right)=\frac{n_1^2(n_1-1)-n_1^2n_2+n_2^2}{(n_1^2-n_1+n_2)n_1^2(n_1-1)}$$
$$\leqslant\frac{n_1^2(n_1-1)}{(n_1^2-n_1+n_2)n_1^2(n_1-1)}=\frac{1}{n_1^2-n_1+n_2}.$$

Since

$$\frac{(n_1-n_2)(n_1-n_2-1)}{n_1^2-n_1+n_2}\leqslant\frac{(n_1-n_2)(n_1-n_2-1)}{n_1(n_1-1)}$$

and

$$\frac{n_2(n_2-1)}{n_1^2-n_1+n_2}\geqslant\frac{n_2(n_2-1)}{n_1^2},$$

it follows from Lemma 10.4 that

$$\begin{aligned}J(2k,n_1,n_2)\geqslant\frac{1}{n_1}+\frac{n_1-n_2}{n_1(n_1-1)}+\frac{n_2}{n_1^2}+\frac{(n_1-n_2)(n_1-n_2-1)}{n_1^2-n_1+n_2}J(2k-2,n_1-2,n_2)\\+\frac{(2n_1-2n_2+1)n_2}{n_1^2-n_1+n_2}J(2k-2,n_1-1,n_2-1)\\+\frac{n_2(n_2-1)}{n_1^2-n_1+n_2}J(2k-2,n_1,n_2-2).\end{aligned}$$

Therefore

$$\begin{aligned}\Delta(k,n_1,n_2)\leqslant\frac{1}{n_1^2-n_1+n_2}+\frac{(n_1-n_2)(n_1-n_2-1)}{n_1^2-n_1+n_2}\Delta(k-1,n_1-2,n_2)\\+\frac{(2n_1-2n_2+1)n_2}{n_1^2-n_1+n_2}\Delta(k-1,n_1-1,n_2-1)\\+\frac{n_2(n_2-1)}{n_1^2-n_1+n_2}\Delta(k-1,n_1,n_2-2).\end{aligned}$$

By the induction hypothesis, it follows that

$$\begin{aligned}\Delta(k,n_1,n_2)&\leqslant\frac{1}{n_1^2-n_1+n_2}+\frac{(n_1-n_2)(n_1-n_2-1)}{n_1^2-n_1+n_2}\left(\frac{2}{n_1+n_2-2k+1}-\frac{2}{n_1+n_2-1}\right)\\&\quad+\frac{(2n_1-2n_2+1)n_2}{n_1^2-n_1+n_2}\left(\frac{2}{n_1+n_2-2k+1}-\frac{2}{n_1+n_2-1}\right)\\&\quad+\frac{n_2(n_2-1)}{n_1^2-n_1+n_2}\left(\frac{2}{n_1+n_2-2k+1}-\frac{2}{n_1+n_2-1}\right)\\&=\frac{1}{n_1^2-n_1+n_2}+\frac{2}{n_1+n_2-2k+1}-\frac{2}{n_1+n_2-1}.\end{aligned}$$

In order for the induction to work, we have to verify that

$$\frac{1}{n_1^2-n_1+n_2}+\frac{2}{n_1+n_2-2k+1}-\frac{2}{n_1+n_2-1}\leqslant\frac{2}{n_1+n_2-2k+1}-\frac{2}{n_1+n_2+1},$$

or equivalently that

$$\frac{4}{(n_1+n_2+1)(n_1+n_2-1)} \geqslant \frac{1}{n_1^2-n_1+n_2}.$$

Since both denominators are positive, this reduces to showing that

$$4(n_1^2-n_1+n_2) \geqslant (n_1+n_2)^2-1. \tag{30}$$

If we fix n_1, then plainly (30) holds in the two extreme cases $n_2=0$ and $n_2=n_1$. Since, as a function of n_2, the left-hand side of (30) is linear while the right-hand side is convex, (30) holds in the other cases as well.

Case 2. Now consider the case that

$$I(k,n_1,n_2) = \frac{1}{n_1-1} + \frac{n_1-n_2-1}{n_1-1} I(k-1,n_1-2,n_2) + \frac{n_2}{n_1-1} I(k-1,n_1-1,n_2-1).$$

Using the induction hypothesis, we have

$$\begin{aligned} I(k,n_1,n_2) &\leqslant \frac{1}{n_1-1} + \frac{n_1-n_2-1}{n_1-1}\left(J(2k-2,n_1-2,n_2) + \frac{2}{n_1+n_2-2k+1} - \frac{2}{n_1+n_2-1}\right) \\ &\quad + \frac{n_2}{n_1-1}\left(J(2k-2,n_1-1,n_2-1) + \frac{2}{n_1+n_2-2k+1} - \frac{2}{n_1+n_2-1}\right) \\ &= \frac{1}{n_1-1} + \frac{2}{n_1+n_2-2k+1} - \frac{2}{n_1+n_2-1} \\ &\quad + \frac{n_1-n_2-1}{n_1-1} J(2k-2,n_1-2,n_2) + \frac{n_2}{n_1-1} J(2k-2,n_1-1,n_2-1) \\ &= J(2k-1,n_1-1,n_2) + \frac{2}{n_1+n_2-2k+1} - \frac{2}{n_1+n_2-1}. \end{aligned}$$

To verify the induction step, we wish to show that this is at most

$$J(2k,n_1,n_2) + \frac{2}{n_1+n_2-2k+1} - \frac{2}{n_1+n_2+1}.$$

It suffices to note that, by Lemma 10.3,

$$J(2k-1,n_1-1,n_2) \leqslant J(2k,n_1,n_2).$$

Case 3. In the third case,

$$I(k,n_1,n_2) = I(k-1,n_1-2,n_2).$$

By induction,

$$I(k-1,n_1-2,n_2) \leqslant J(2k-2,n_1-2,n_2) + \frac{2}{n_1+n_2-2k+1} - \frac{2}{n_1+n_2-1}.$$

Therefore, in order to show that

$$\Delta(k,n_1,n_2)\leqslant\frac{2}{n_1+n_2-2k+1}-\frac{2}{n_1+n_2+1},$$

we only have to verify that

$$J(2k-2,n_1-2,n_2)-J(2k,n_1,n_2)\leqslant\frac{2}{n_1+n_2-1}-\frac{2}{n_1+n_2+1}.$$

Again, by Lemma 10.3, $J(2k-2,n_1-2,n_2)-J(2k,n_1,n_2)\leqslant 0$. This completes the proof of Proposition 10.2.

10.3. Bound on the difference in cost of the integer and linear flow problems

Equation (10), when specialized to the case of weights 1 and capacities at most 2, becomes

$$\begin{aligned}EC_k(n_1,n_2)=\frac{1}{n_1}\lim_{\gamma_{n+1}\to 0}&\frac{E[\delta^{(k)}_{n+1}(\mathbf{c}+\mathbf{1}_{n+1})]}{\gamma}\\&+\frac{n_1-n_2}{n_1}EC_{k-1}(n_1-1,n_2)+\frac{n_2}{n_1}EC_{k-1}(n_1,n_2-1).\end{aligned}\tag{31}$$

This equation holds for the integer as well as the linear flow problem, provided that δ stands for the degree in the optimal solution to the corresponding optimization problem. Let $C_k(n_1,n_2)$ and $\widetilde{C}_k(n_1,n_2)$ denote, respectively, the cost of the linear and integer flow problems. Similarly, let $\tilde{\delta}$ denote the degree with respect to the solution to the integer optimization problem. By Lemma 7.2 (with $h=2k$), we know that

$$\lim_{\gamma_{n+1}\to 0}\frac{E[\delta^{(k)}_{n+1}(\mathbf{c}+\mathbf{1}_{n+1})]}{\gamma_{n+1}}=ET^{(2k)}(\mathbf{c})=J(2k,n_1,n_2),$$

and, by Proposition 10.2, we therefore have

$$\lim_{\gamma_{n+1}\to 0}\frac{E[\tilde{\delta}^{(k)}_{n+1}(\mathbf{c}+\mathbf{1}_{n+1})]}{\gamma_{n+1}}-\lim_{\gamma_{n+1}\to 0}\frac{E[\delta^{(k)}_{n+1}(\mathbf{c}+\mathbf{1}_{n+1})]}{\gamma_{n+1}}\leqslant\frac{2}{n_1+n_2-2k+1}.$$

Hence, from (31) it follows that

$$\begin{aligned}E(\widetilde{C}_k(n_1,n_2)-C_k(n_1,n_2))\leqslant\frac{1}{n_1}&\frac{2}{n_1+n_2-2k+1}\\&+\max\{\widetilde{C}_{k-1}(n_1-1,n_2)-C_{k-1}(n_1-1,n_2),\\&\qquad\widetilde{C}_{k-1}(n_1,n_2-1)-C_{k-1}(n_1,n_2-1)\}.\end{aligned}\tag{32}$$

This allows us to establish the following bound.

PROPOSITION 10.5.

$$E(\widetilde{C}_k(n_1,n_2)-C_k(n_1,n_2)) \leqslant \frac{4}{n_1+n_2-k}\left(\frac{1}{n_1+n_2-2k+1}+\frac{1}{n_1+n_2-2k+2}+...+\frac{1}{n_1+n_2-k}\right). \tag{33}$$

Proof. By induction, we may assume that

$$\max\{\widetilde{C}_{k-1}(n_1-1,n_2)-C_{k-1}(n_1-1,n_2),\widetilde{C}_{k-1}(n_1,n_2-1)-C_{k-1}(n_1,n_2-1)\} \leqslant \frac{4}{n_1+n_2-k}\left(\frac{1}{n_1+n_2-2k+2}+...+\frac{1}{n_1+n_2-k}\right).$$

To verify that (33) follows from (32), we therefore only have to check that

$$\frac{2}{n_1(n_1+n_2-2k+1)} \leqslant \frac{4}{(n_1+n_2-k)(n_1+n_2-2k+1)}.$$

This inequality is equivalent to $n_2-k\leqslant n_1$, which clearly holds, since $n_2-k\leqslant n_2\leqslant n_1$. □

For simplicity, we replace (33) by a bound which depends only on $n_1=n$ and not on k and n_2.

THEOREM 10.6. *For weight-*1 *flow problems with capacities at most* 2,

$$E(\widetilde{C}_k(n_1,n_2)-C_k(n_1,n_2)) = O\left(\frac{\log n_1}{n_1}\right).$$

Proof. Since $k\leqslant\frac{1}{2}(n_1+n_2)$, we have

$$\frac{4}{n_1+n_2-k} \leqslant \frac{4}{n_1+n_2-(n_1+n_2)/2} = \frac{8}{n_1+n_2} \leqslant \frac{8}{n_1}.$$

Moreover,

$$\frac{1}{n_1+n_2-2k+1}+...+\frac{1}{n_1+n_2-k} \leqslant 1+\frac{1}{2}+...+\frac{1}{k} \leqslant 1+\frac{1}{2}+...+\frac{1}{n_1}.$$

Hence the right-hand side of (33) is at most

$$\frac{8}{n_1}\left(1+\frac{1}{2}+...+\frac{1}{n_1}\right) = O\left(\frac{\log n_1}{n_1}\right).$$ □

10.4. Concentration for the integer flow problem

Theorems 7.3, 8.1 and 10.6 imply that the expected cost of the (non-relaxed) matching problem on K_n for even n converges to $\frac{1}{12}\pi^2$ (see also equation (54)). Similarly, the expected cost of the 2-factor problem converges to $L^\star$.

Our next objective is to establish concentration inequalities strong enough to show that the costs of these problems converge in probability to their limits. Let k be a positive integer that allows a solution to the integer problem, and let C_k and $\widetilde{C}_k$ denote the costs of the linear and integer flow problems, respectively.

THEOREM 10.7. *Suppose that, for* $1\leqslant i\leqslant n$, $\gamma_i=1$ *and* $c_i\leqslant 2$. *Then*

$$E|\widetilde{C}_k-EC_k|=O\left(\frac{(\log n)^{5/2}}{n^{1/2}}\right). \qquad (34)$$

Proof. By a standard inequality,

$$E|C_k-EC_k|\leqslant\sqrt{\operatorname{var}(C_k)},$$

and by Theorem 9.1,

$$\operatorname{var}(C_k)=O\left(\frac{(\log n)^5}{n}\right).$$

Hence,

$$E|\widetilde{C}_k-EC_k|\leqslant E|C_k-EC_k|+E(\widetilde{C}_k-C_k)\leqslant\sqrt{\operatorname{var}(C_k)}+O\left(\frac{\log n}{n}\right)=O\left(\frac{(\log n)^{5/2}}{n^{1/2}}\right).$$

□

COROLLARY 10.8. *The cost of the* 2*-factor problem converges in probability to* $L^\star$, *and*, *for even* n, *the cost of the minimum matching converges in probability to* $\frac{1}{12}\pi^2$.

The fact that the cost of the minimum matching converges in probability to $\frac{1}{12}\pi^2$ is known to follow from the work of Aldous [3], [4] on the assignment (bipartite matching) problem, provided [3, Proposition 2] is modified so that it applies to the complete graph.

11. The mean field model

In this section we show that some of the asymptotic results carry over from the friendly model to the mean field model. In particular we complete the proof of Theorem 1.1 by proving that L_n converges in probability to $L^\star$ under the weaker assumption (1). We here consider only the traveling salesman problem, although with obvious minor changes the argument applies also to matching and related problems. The feasible solutions to the TSP contain no loops, multiple edges, or edges with coefficient $\frac{1}{2}$. Therefore, we can disregard all edges except the cheapest edge between each pair of distinct vertices. This makes the friendly model a special case of the mean field model described in the introduction, namely by taking μ to be the rate-1 exponential distribution.

We use the following notation: Let $Z_{k,n}$ denote the cost of the minimum integer k-flow in K_n given that each vertex has capacity 2, in other words an incomplete 2-factor. We let L_n denote the cost of a traveling salesman tour in K_n. Since we are considering the differences between various distributions, we shall here explicitly write out the dependence on the distribution under consideration. We let μ be a general pseudodimension-1 distribution. Further, we let exp denote rate-1 exponential distribution, and U denote uniform distribution on $[0,1]$.

It was shown by A. Frieze [18] that, as $n\to\infty$,

$$L_n(U)-Z_{n,n}(U)\xrightarrow{\mathrm{p}}0. \tag{35}$$

Every pseudodimension-1 distribution is stochastically dominated by a continuous pseudodimension-1 distribution, and conversely it dominates some other continuous pseudodimension-1 distribution. It is therefore sufficient to prove Theorem 1.1 under the assumption that μ is continuous. We first establish the following lower bound.

PROPOSITION 11.1. *If μ satisfies* (1), *then for every* $\varepsilon>0$,

$$P(L_n(\mu)\leqslant L^\star-\varepsilon)\to 0$$

as $n\to\infty$.

Proof. We establish the stronger result that

$$P(Z_{n-1,n}(\mu)\leqslant L^\star-\varepsilon)\to 0. \tag{36}$$

If X is a variable of distribution μ, then the function $F(t)=P(X<t)$ satisfies $F(0)=0$ and $F'(0)=1$ (only the right derivative is considered). Consider the minimum integer $(n-1)$-flow under distribution μ, and remove the k most expensive edges for an integer k chosen as a function of n. The only requirement on this function is that, as $n\to\infty$, we must have $k\to\infty$ and $k/n\to 0$. For instance, we can choose $k=\lfloor n^{1/2}\rfloor$. Let x be the cost of the most expensive remaining edge in the flow. Now we apply an order-preserving transformation of all edge costs, so that they become exponentially distributed with mean 1 (and still independent). This is done by the mapping

$$t\longmapsto\phi(t)$$

that satisfies

$$F(t)=P(X<t)=P(\widetilde{X}<\phi(t))=1-e^{-\phi(t)},$$

where $\widetilde{X}$ is rate-1 exponential. Solving for $\phi(t)$ gives

$$\phi(t)=-\log(1-F(t)).$$

We note that

$$\phi'(t)=\frac{F'(t)}{1-F(t)},$$

which in particular means that $\phi(t)\sim t$ for small t.

Let C be the cost of the $(n-1-k)$-flow of exponential cost edges that has been obtained this way. Obviously C stochastically dominates the cost $Z_{n-1-k,n}(\exp)$ of the *minimum* integer $(n-1-k)$-flow for capacities 2. By Corollary 8.13 and equation (34), $Z_{n-1-k,n}(\exp)$ converges in probability to $L^\star$ provided $k/n\to 0$. We want to show that given C and x, we can obtain a lower bound on the cost $Z_{n-1,n}(\mu)$ of the original minimum $(n-1)$-flow. The flow of cost C contains no edge of cost $>x$. Now define

$$f(y)=\inf_{0<t\leqslant y}\frac{t}{\phi(t)}.$$

Then $f(y)\to 1$ as $y\to 0^+$, and f is decreasing. We have

$$Z_{n-1,n}(\mu)\geqslant Cf(x)+kx\geqslant Z_{n-1-k,n}(\exp)f(x)+kx. \tag{37}$$

We want to show that, with high probability, the right-hand side of (37) is at least $L^\star-\varepsilon$. If μ (and thus f) is fixed, then for every ε_1 we can choose n large enough so that $f(x)>1-\varepsilon_1$ whenever $x<L^\star/k$. In that case, it follows that the right-hand side of (37) is at least $(1-\varepsilon_1)Z_{n-1-k,n}$. Since $Z_{n-1-k,n}\to L^\star$ in probability, the statement follows. □

We now turn to the upper bound on $L_n(\mu)$. In order to establish Theorem 1.1, it only remains to show the following.

PROPOSITION 11.2. *If μ satisfies* (1), *then for every $\varepsilon>0$,*

$$P(L_n(\mu)\geqslant L^\star+\varepsilon)\to 0$$

as $n\to\infty$.

Proof. We know that

$$Z_{n,n}(\exp)\xrightarrow{\text{p}}L^\star. \tag{38}$$

Since uniform distribution on $[0,1]$ is stochastically dominated by exponential distribution, it follows from (38) and Proposition 11.1 that

$$Z_{n,n}(U)\xrightarrow{\text{p}}L^\star.$$

Now it follows from (35) that

$$L_n(U)\xrightarrow{\text{p}}L^\star.$$

By (36), applied to the distribution U, it follows that

$$L_n(U)-Z_{n-1,n}(U)\xrightarrow{\text{p}}0.$$

Thus, the cost of the most expensive edge in the minimum tour under distribution U also converges in probability to zero, since by removing this edge we obtain an $(n-1)$-flow.

Hence, for fixed ε and letting $n \to \infty$, the order-preserving transformation that converts uniform $[0,1]$ variables to distribution μ will give a tour of cost at most $(1+\varepsilon)L^\star$ with probability tending to 1. This completes the proof of Proposition 11.2, and thereby of Theorem 1.1. □

12. The bipartite friendly model

We show that results similar to those established in §§4–8 hold for an analogous and technically simpler *bipartite friendly model*. Actually our results on the bipartite model are stronger, since we obtain exact results not only for the expectation but also for the higher moments of the cost of the flow problem.

12.1. Definitions and combinatorial results

In the bipartite friendly model, there are two vertex sets

$$U = \{u_1, ..., u_m\} \quad \text{and} \quad V = \{v_1, ..., v_n\},$$

and the vertices have weights $\alpha_1, ..., \alpha_m$ and $\beta_1, ..., \beta_n$, and capacities $\mathbf{a} = (a_1, ..., a_m)$ and $\mathbf{b} = (b_1, ..., b_n)$, respectively. For each pair u_i, v_j there is a sequence of edges whose costs are given by a Poisson process of rate $\alpha_i \beta_j$. There are no edges within U or within V.

The definition of the random flow problem in §4 carries over to the bipartite model with obvious minor changes in notation. An important difference that makes the bipartite model considerably simpler than the complete model is that the linear and integer problems are equivalent when k is an integer. This follows from the analysis of §6. The results of that section obviously hold for the bipartite graph too, but since there are no cycles of odd length, it follows that in the generic case, the minimum flow is stable if and only if k is an integer, and in that case the minimum linear flow is actually an integer flow.

Hence, in the bipartite case, we need only consider integer values of k, and integer flows. It is therefore natural to think of a flow as a set of edges. Following the method of §5, we introduce an extra vertex u_{m+1} of weight α_{m+1} which has edges only to the vertices $v_1, ..., v_n$. When we let $\alpha_{m+1} \to 0$ we obtain, paralleling the argument of §5, the equation

$$EC_k(\mathbf{a}, \mathbf{b}) = \frac{1}{\beta_1 + ... + \beta_n} \lim_{\alpha_{m+1} \to 0} \frac{E[\delta_{m+1}^{(k)}(\mathbf{a} + \mathbf{1}_{m+1}, \mathbf{b})]}{\alpha_{m+1}} + EC_{k-1}(\mathbf{a}, \mathbf{b} - \mathbf{1}_j), \qquad (39)$$

analogous to (10). Here $\delta^{(k)}_{m+1}(\mathbf{a}+\mathbf{1}_{m+1},\mathbf{b})$ is the degree of the extra vertex in the extended flow problem, while the last term $EC_{k-1}(\mathbf{a},\mathbf{b}-\mathbf{1}_j)$ refers to a random choice of $v_j\in V$ taken with probabilities proportional to the weights. But here we deviate from the argument of §7, and instead derive a result for all moments.

Let F be a flow (we now think of F as a subset of the set E of edges). We define the *span* of F (with respect to the capacities $\mathbf{a}$) to be the set of edges which either belong to F, or are incident to a vertex in U that has full degree in F.

LEMMA 12.1. *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 which is not in the span of F.*

Proof. Let G be a minimum cost r-flow, and let $H=F\triangle G$ be the symmetric difference of F and G. We split H into path and cycles that can be "switched", respecting the capacity constraints. Then the unbalanced components can be paired so that H is partitioned into a number of sets that contain equally many edges from F and G, and that can be switched. One of these sets must be such that by switching, the cost decreases. □

The following is a corollary to Lemma 12.1. To simplify the statement, we assume that the edge costs are generic.

LEMMA 12.2. *Let $v_j\in V$ and suppose that the capacity b_j is non-zero. Let*

$$F=\sigma^{(r)}(\mathbf{a},\mathbf{b}-\mathbf{1}_j) \quad \textit{and} \quad G=\sigma^{(r+1)}(\mathbf{a},\mathbf{b}).$$

Then G contains exactly one edge which is not in the span of F.

Proof. In order to apply Lemma 12.1, we introduce an auxiliary element u_{m+1} of capacity 1. We let the first edge between u_{m+1} and v_j have non-negative cost x, and let all other edges from u_{m+1} have infinite cost. If we put $x=0$, then the minimum $(r+1)$-flow with respect to $(\mathbf{a}+\mathbf{1}_{m+1},\mathbf{b})$ consists of the cheapest edge e between u_{m+1} and v_j together with F. If we increase the value of x, then at some point this minimum $(r+1)$-flow changes to G. If we let x have a value just above this point, so that the minimum $(r+1)$-flow is G, but no other $(r+1)$-flow has smaller cost than $F+e$, then it follows from Lemma 12.1 that G contains exactly one edge outside the span of F. □

The following theorem is the basis for our results on the higher moments in the bipartite model. We have been unable to find an analogous result for the complete model, and this is mainly why we have not established, or even conjectured, any exact results for the higher moments in the complete model.

As above, let $F=\sigma^{(r)}(\mathbf{a},\mathbf{b}-\mathbf{1}_j)$ and $G=\sigma^{(r+1)}(\mathbf{a},\mathbf{b})$. Moreover, let u_G be the element of U incident to the unique edge of G which is not in the span of F.

THEOREM 12.3. *If we condition on the span of F and the cost of F, then the cost of G is independent of u_G, and u_G is distributed on the set of vertices in U that do not have full degree in F, with probabilities proportional to the weights.*

Proof. We condition on (1) the costs of all edges in the span of F, and (2) for each $v \in V$, the minimum cost of all edges to v which are not in the span of F. By Lemma 12.1, we have thereby conditioned on the cost of G, and by the memorylessness of the Poisson process, the unknown endpoint u_G of the edge that goes outside the span of F is still distributed, among the vertices that do not have full degree in F, with probabilities proportional to the weights. □

12.2. The normalized limit measure

As we shall see, the method of §5 together with Theorem 12.3 allow us to establish exact results for all moments of the cost of a bipartite flow problem. We extend U by an extra vertex u_{m+1} of capacity 1 and study the probability that u_{m+1} participates in the minimum solution.

The method, which relies on finally letting the weight α_{m+1} tend to zero, leads one to think of α_{m+1} as infinitesimal. This suggests an alternative way of interpreting equations (8), (10) and (39). The limit expectation as $\alpha_{m+1} \to 0$ of a random variable like $\sigma_e^{(k)}$, normalized by dividing by α_{n+1}, can be regarded as a measure in its own right. If ϕ is a function of the edge costs, we let

$$E^\star \phi = \lim_{\alpha_{m+1} \to 0} \frac{E\phi}{\alpha_{m+1}}.$$

This is the *normalized limit measure* of ϕ. We use the same notation for events, with the obvious interpretation, replacing expectation by probability. With this notation, we can write equation (39) as

$$EC_k(\mathbf{a}, \mathbf{b}) = \frac{1}{\beta_1 + \ldots + \beta_n} E^\star[\delta_{m+1}^{(k)}(\mathbf{a} + \mathbf{1}_{m+1}, \mathbf{b})] + EC_{k-1}(\mathbf{a}, \mathbf{b} - \mathbf{1}_j), \qquad (40)$$

but the interesting thing is that we can interpret the normalized limit measure directly, without regarding it as a limit. Notice that the exponential distribution with rate $\alpha_{m+1}\beta_j$, scaled up by a factor $1/\alpha_{m+1}$, converges to β_j times Lebesgue measure on the positive real numbers as $\alpha_{m+1} \to 0$. Therefore, we can construct a measure space whose points are assignments of costs to the edges, where exactly one edge from u_{m+1} has finite cost, and the others have cost $+\infty$. For each j, the cost assignments where the edge (u_{m+1}, v_j) has finite cost are measured by β_j times Lebesgue measure on the positive reals. The union of these spaces for $1 \leqslant j \leqslant n$ is combined with the original probability measure on the costs of the other edges.

From this starting point, equation (40) could have been derived without referring to the principle of dominated convergence, but we shall not discuss this alternative approach in detail here.

12.3. A recursive formula

If $\mathbf{g}=(g_1, ..., g_m)$ is a vector of non-negative integers, then we let $I_k(\mathbf{g}, \mathbf{a}, \mathbf{b})$ be the indicator variable for the event that for every $u_i \in U$, the minimum k-flow with respect to $(\mathbf{a}, \mathbf{b})$ contains at least g_i edges from u_i. Naturally we must have $g_i \leqslant a_i$ for every i, and $g_1+...+g_m \leqslant k$, in order for $I_k(\mathbf{g}, \mathbf{a}, \mathbf{b})$ to be non-zero.

Let N be a positive integer, and let $\mathbf{g}$ be as above. Assume moreover that

$$g_1+...+g_m = k-1.$$

For $v_j \in V$, let I_j be the indicator variable for the event that the minimum k-flow with respect to $(\mathbf{a}+\mathbf{1}_{m+1}, \mathbf{b})$ contains an edge between u_{m+1} and v_j and that for $1 \leqslant i \leqslant m$, it contains exactly g_i edges from u_i.

LEMMA 12.4.

$$E[C_k(\mathbf{a}, \mathbf{b})^N I_{k-1}(\mathbf{g}, \mathbf{a}, \mathbf{b}-\mathbf{1}_j)] = E[C_{k-1}(\mathbf{a}, \mathbf{b}-\mathbf{1}_j)^N I_{k-1}(\mathbf{g}, \mathbf{a}, \mathbf{b}-\mathbf{1}_j)] + \frac{N}{\beta_j} E^\star[C_k(\mathbf{a}+\mathbf{1}_{m+1}, \mathbf{b})^{N-1} I_j]. \tag{41}$$

Proof. We compute $(N/\beta_j)E^\star[C_k(\mathbf{a}+\mathbf{1}_{m+1}, \mathbf{b})^{N-1} I_j]$ by integrating over the cost, which we denote by t, of the first edge between u_{m+1} and v_j. We therefore condition on the costs of all other edges.

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

$$\frac{d}{dt}(C_k(\mathbf{a}+\mathbf{1}_{m+1}, \mathbf{b})^N I_{k-1}(\mathbf{g}, \mathbf{a}, \mathbf{b}-\mathbf{1}_j)) = N C_k(\mathbf{a}+\mathbf{1}_{m+1}, \mathbf{b})^{N-1} I_j.$$

According to the normalized limit measure, since we are conditioning on the first edge between u_{m+1} and v_j being the one with finite cost, $E^\star$ is just β_j times Lebesgue measure. Therefore (41) now follows from the fundamental theorem of calculus: putting $t=\infty$, we get

$$C_k(\mathbf{a}+\mathbf{1}_{m+1}, \mathbf{b})^N I_{k-1}(\mathbf{g}, \mathbf{a}, \mathbf{b}-\mathbf{1}_j) = C_k(\mathbf{a}, \mathbf{b})^N I_{k-1}(\mathbf{g}, \mathbf{a}, \mathbf{b}-\mathbf{1}_j),$$

while if $t=0$, we get

$$C_k(\mathbf{a}+\mathbf{1}_{m+1},\mathbf{b})^N I_{k-1}(\mathbf{g},\mathbf{a},\mathbf{b}-\mathbf{1}_j)=C_{k-1}(\mathbf{a},\mathbf{b}-\mathbf{1}_j)^N I_{k-1}(\mathbf{g},\mathbf{a},\mathbf{b}-\mathbf{1}_j). \qquad \square$$

PROPOSITION 12.5. *Let* $\mathbf{h}=(h_1,...,h_m)$ *be a non-negative integer vector such that* $h_i\leqslant a_i$ *for every* i, *and* $h_1+...+h_m=k$. *Moreover, let*

$$\beta=\sum_{\substack{1\leqslant j\leqslant n\\ b_j>0}}\beta_j.$$

Then

$$\begin{aligned}
&E[C_k(\mathbf{a},\mathbf{b})^N I_k(\mathbf{h},\mathbf{a},\mathbf{b})]\\
&=\sum_{\substack{1\leqslant i\leqslant m\\ h_i>0}}\frac{\alpha_i}{\alpha_i+\sum_{\substack{j\neq i\\ h_j<a_j}}\alpha_j}\sum_{\substack{1\leqslant j\leqslant n\\ b_j>0}}\frac{\beta_j}{\beta}E[C_{k-1}(\mathbf{a},\mathbf{b}-\mathbf{1}_j)^N I_{k-1}(\mathbf{h}-\mathbf{1}_i,\mathbf{a},\mathbf{b}-\mathbf{1}_j)]\\
&\quad+\frac{N}{\beta}\sum_{\substack{1\leqslant i\leqslant m\\ h_i>0}}\frac{\alpha_i}{\alpha_i+\sum_{\substack{j\neq i\\ h_j<a_j}}\alpha_j}E^\star[C_k(\mathbf{a}+\mathbf{1}_{m+1},\mathbf{b})^{N-1}I_k(\mathbf{h}-\mathbf{1}_i+\mathbf{1}_{m+1},\mathbf{a}+\mathbf{1}_{m+1},\mathbf{b})].
\end{aligned} \tag{42}$$

Proof. We multiply both sides of (41) by β_j and sum over all j for which $b_j>0$. This way we obtain

$$\begin{aligned}
\sum_{\substack{1\leqslant j\leqslant n\\ b_j>0}}&\beta_j E[C_k(\mathbf{a},\mathbf{b})^N I_{k-1}(\mathbf{g},\mathbf{a},\mathbf{b}-\mathbf{1}_j)]\\
&=\sum_{\substack{1\leqslant j\leqslant n\\ b_j>0}}\beta_j E[C_{k-1}(\mathbf{a},\mathbf{b}-\mathbf{1}_j)^N I_{k-1}(\mathbf{g},\mathbf{a},\mathbf{b}-\mathbf{1}_j)]\\
&\quad+NE^\star[C_k(\mathbf{a}+\mathbf{1}_{m+1},\mathbf{b})^{N-1}I_k(\mathbf{g}+\mathbf{1}_{m+1},\mathbf{a}+\mathbf{1}_{m+1},\mathbf{b})].
\end{aligned} \tag{43}$$

Next, we apply Theorem 12.3. Let $\mathbf{h}$ be as above, and suppose that, for a certain j, $b_j>0$. If the minimum k-flow with respect to $(\mathbf{a},\mathbf{b})$ contains exactly h_i edges from v_i for every i, then the minimum $(k-1)$-flow with respect to $(\mathbf{a},\mathbf{b}-\mathbf{1}_j)$ must have h_i-1 edges from a certain vertex v_i, and h_l edges from v_l for every $l\neq i$. By summing over the possible values of i, we obtain

$$E[C_k(\mathbf{a},\mathbf{b})^N I_k(\mathbf{h},\mathbf{a},\mathbf{b})]=\sum_{\substack{1\leqslant i\leqslant m\\ h_i>0}}\frac{\alpha_i}{\alpha_i+\sum_{\substack{l\neq i\\ h_l<a_l}}\alpha_l}E[C_k(\mathbf{a},\mathbf{b})^N I_{k-1}(\mathbf{h}-\mathbf{1}_i,\mathbf{a},\mathbf{b}-\mathbf{1}_j)]. \tag{44}$$

We average the right-hand side over the possible values of j (those for which $b_j>0$), chosen with probabilities according to the weights. In other words, we multiply (44) by β_j/β and sum over all j such that $b_j>0$. This leaves the left-hand side intact, and we get

$$E[C_k(\mathbf{a},\mathbf{b})^N I_k(\mathbf{h},\mathbf{a},\mathbf{b})] = \sum_{\substack{1\leqslant i\leqslant m \\ h_i>0}} \frac{\alpha_i}{\alpha_i+\sum_{\substack{l\neq i \\ h_l<a_l}} \alpha_l} \times \sum_{\substack{1\leqslant j\leqslant n \\ b_j>0}} \frac{\beta_j}{\beta} E[C_k(\mathbf{a},\mathbf{b})^N I_{k-1}(\mathbf{h}-\mathbf{1}_i,\mathbf{a},\mathbf{b}-\mathbf{1}_j)]. \tag{45}$$

We now rewrite the right-hand side of (45) using equation (43) with $\mathbf{g}=\mathbf{h}-\mathbf{1}_i$. This establishes (42). □

12.4. Interpretation in terms of the urn process

The urn process corresponding to the bipartite case is, as should be expected, symmetric with respect to the two sides of the graph. The coordinate axes are now labeled by x and y. An urn process with the vertices of U runs along the x-axis, and an independent urn process with the vertices of V runs along the y-axis. Again, each vertex is drawn at times of the events in a Poisson process of rate equal to the weight of the vertex. We let $P_i(x)$ be the number of times that u_i has been drawn up to time x, while $Q_j(y)$ similarly denotes the number of times that v_j has been drawn up to time y in the other process. The region R_k is now defined as the region in the xy-plane for which

$$\sum_{i=1}^{m} \min\{P_i(x), a_i\} + \sum_{j=1}^{n} \min\{Q_j(y), b_j\} < k.$$

Denoting, as before, the cost of the minimum k-flow by C_k, it turns out, in analogy with Theorem 7.3, that

$$EC_k = E[\operatorname{area}(R_k)]. \tag{46}$$

Equation (46) is a generalization of the formula conjectured by Buck, Chan and Robbins [14] and proved in [47]. It can be established by following the same route as the proof of Theorem 7.3 in §7, but we shall give a different proof, generalizing to higher moments.

To describe the higher moments, we introduce an *extended urn process*. In the Nth *extension* of the urn process on U and V there are, in addition to the ordinary urn processes, N extra points $(x_1, y_1), ..., (x_N, y_N)$ in the positive quadrant of the xy-plane.

These points are "chosen" according to Lebesgue measure on the positive real numbers, and therefore cannot be treated as random variables. In analogy with the notation for the normalized limit measure, we let $E^\star$ denote the measure obtained by combining the probability measure on the ordinary urn process with Lebesgue measure on the extra points. The measure $E^\star$ is the expected value (with respect to the ordinary urn process) of the Lebesgue measure in $2N$ dimensions of the set of points $x_1, ..., x_N, y_1, ..., y_N$ belonging to a particular event. It can also be interpreted as a normalized limit by letting $x_1, ..., x_N, y_1, ..., y_N$ be independent exponentially distributed of rate $\lambda \to 0$.

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

$$r(x) = \sum_{i=1}^{m} \min\{P_i(x), a_i\} + \#\{i : x_i \leqslant x\}.$$

Similarly we let

$$s(y) = \sum_{j=1}^{n} \min\{Q_j(y), b_j\} + \#\{j : y_j \leqslant y\}.$$

Our main exact theorem on bipartite random flow problems is the following.

THEOREM 12.6.

$$E[C_k(\mathbf{a}, \mathbf{b})^N] = E^\star\{(x, y) : r(x_i) + s(y_i) \leqslant k + N \textit{ for } 1 \leqslant i \leqslant N\}. \tag{47}$$

When $N=1$, the right-hand side of (47) is the expected area of $R_k(\mathbf{a}, \mathbf{b})$. For larger N, the condition $r(x_i)+s(y_i) \leqslant k+N$ for $1 \leqslant i \leqslant N$ implies that the points (x_i, y_i) all lie in $R_k(\mathbf{a}, \mathbf{b})$. Therefore we have the following result.

COROLLARY 12.7.

$$E(C_k(\mathbf{a}, \mathbf{b})^N) \leqslant E[\operatorname{area}(R_k(\mathbf{a}, \mathbf{b}))^N],$$

with equality if $N=1$.

This in turn implies that

$$\operatorname{var}(C_k(\mathbf{a}, \mathbf{b})) \leqslant \operatorname{var}(\operatorname{area}(R_k(\mathbf{a}, \mathbf{b}))),$$

so that the methods of §8 can be applied in order to bound the variance of $C_k(\mathbf{a}, \mathbf{b})$, but we shall not pursue this further here.

We prove Theorem 12.6 by proving the following more precise form. Let $\mathbf{u}_k$ be the multiset of the first k vertices to be drawn in the urn process (with replacement protocol $\mathbf{a}$) on U, and define $\mathbf{v}_k$ similarly.

THEOREM 12.8. *Let* $\mathbf{h}=(h_1,...,h_m)$ *be such that* $h_i\leqslant a_i$ *for every* i, *and*

$$h_1+...+h_m=k.$$

Then

$$E(C_k(\mathbf{a},\mathbf{b})^N I_k(\mathbf{h},\mathbf{a},\mathbf{b}))=E^\star[\mathbf{u}_k=\mathbf{h} \text{ and } r(x_i)+s(y_i)\leqslant k+N \text{ for } 1\leqslant i\leqslant N]. \tag{48}$$

Proof. We use (42) together with induction on both k and N. Notice that (48) holds trivially when $k=0$. Notice also that when $N=0$, the second term of the right-hand side of (42) vanishes.

Suppose therefore that (48) holds whenever k or N is replaced by a smaller number. Then the right-hand side of (42) can be rewritten in terms of the urn process. We will show that the result is equal to the right-hand side of (48).

We therefore split the "event" $\mathbf{u}_k=\mathbf{h}$ and $r(x_i)+s(y_i)\leqslant k+N$ for $i=1,...,N$ into two cases. Let $\mathbf{g}=\mathbf{u}_{k-1}$. Let t be the time at which the first event occurs in the extended urn process on V, in other words, we condition on t being minimal such that $s(t)=1$.

For $1\leqslant i\leqslant m$ and positive integers l, let $\xi(i,l)$ be the time at which u_i is drawn for the lth time, and similarly let $\psi(j,l)$ be the time at which v_j is drawn for the lth time. We think of these times as being defined for all i and j regardless of the capacities.

Case 1. Let $v_j\in V$ and suppose that the event that occurs at time t is that v_j is drawn from the urn. This means that v_j is the first element (of non-zero capacity) to be drawn from the urn in the y-process, and moreover that the time t at which this happens is smaller than all the numbers $y_1,...,y_N$.

We couple to another extended urn process by letting

$$\xi'(v_j,l)=\xi(v_j,l+1)-t$$

and, for all $i\neq j$,

$$\xi'(v_i,l)=\xi(v_i,l)-t.$$

Moreover, for $1\leqslant i\leqslant N$, let

$$y_i'=y_i-t.$$

Let $x_i'=x_i$ and, for $u_i\in U$, $\xi'(u_i,l)=\xi(u_i,l)$. Let r' and s' be the rank functions with respect to $\mathbf{a}$ and $\mathbf{b}-\mathbf{1}_j$, respectively, in the primed extended urn process. Then $r'=r$ and $s'(y_i')=s(y_i)-1$. Hence, for $1\leqslant i\leqslant N$,

$$r'(x_i')+s'(y_i')\leqslant k-1+N \quad \text{if and only if} \quad r(x_i)+s(y_i)\leqslant k+N.$$

If we condition on $\mathbf{u}_{k-1}=\mathbf{g}=\mathbf{h}-\mathbf{1}_i$, then

$$P(\mathbf{u}_k=\mathbf{h})=\frac{\alpha_i}{\alpha_i+\sum_{\substack{l\neq i\\ h_l<a_l}}\alpha_l}.$$

By the induction hypothesis, it follows that $E^\star$(case 1) is equal to the first term of the right-hand side of (42).

Case 2. Suppose that $y_N=t$ (the cases $y_i=t$ for $1\leqslant i\leqslant N-1$ are of course identical). We couple case 2 to the $(N-1)$-th extended urn process in essentially the same way as we did in case 1. Let

$$\xi'(v_j,l)=\xi(v_j,l)-t,$$

and, for $1\leqslant i\leqslant N-1$, let

$$y_i'=y_i-t.$$

In the limit $\alpha_{n+1}\to 0$, the point $\xi(u_{m+1},1)$ is measured by α_{n+1} times Lebesgue measure on the positive real numbers. Hence, if we let

$$\xi'(u_{m+1},1)=x_N,$$

we obtain a coupling which is valid as $\alpha_{m+1}\to 0$.

Again, conditioning on $\mathbf{u}_{k-1}=\mathbf{g}=\mathbf{h}-\mathbf{1}_i$,

$$P(\mathbf{u}_k=\mathbf{h})=\frac{\alpha_i}{\alpha_i+\sum_{\substack{l\neq i\\ h_l<a_l}}\alpha_l}.$$

Moreover,

$$\frac{1}{\beta}=\frac{1}{\sum_{\substack{1\leqslant l\leqslant n\\ b_l>0}}\beta_l}$$

is the measure of the event that y_N is smaller than $\xi(v_j,1)$ for every j such that $b_j>0$. By the induction hypothesis, it now follows that $E^\star$(case 2) is equal to the second term of the right-hand side of (42).

Hence

$$E(C_k(\mathbf{a},\mathbf{b})^N I_k(\mathbf{h},\mathbf{a},\mathbf{b}))=E^\star(\text{case } 1)+NE^\star(\text{case } 2).$$

This completes the proof. □

Theorem 12.8 generalizes automatically to vectors $\mathbf{h}$ whose sum is smaller than k.

THEOREM 12.9. *Let* $\mathbf{h}\leqslant\mathbf{a}$. *Then*

$$E(C_k(\mathbf{a},\mathbf{b})^N I_k(\mathbf{h},\mathbf{a},\mathbf{b}))=E^\star[\mathbf{u}_k\geqslant\mathbf{h} \text{ and } r(x_i)+s(y_i)\leqslant k+N \text{ for } 1\leqslant i\leqslant N].$$

In particular, if $\mathbf{h}=\mathbf{0}$, then this is Theorem 12.6.

12.5. The area of the region R_k

With the method used in §8, we can establish a similar estimate of the area of R_k for the bipartite case. Suppose again that each vertex has weight 1 and capacity at most 2. Now $R=R_k(m_1,m_2,n_1,n_2)$ is the region for which

$$\sum_{i=1}^{m} P_i(x)+\sum_{j=1}^{n} Q_j(y)<k.$$

Naturally we take $R^\star=R^\star_k(m_1,m_2,n_1,n_2)$ to be the region for which

$$\sum_{i=1}^{m} E(P_i(x))+\sum_{j=1}^{n} E(Q_j(y))<k.$$

For simplicity, we assume that m and n are of the same order of magnitude. We briefly outline a proof that the statement of Theorem 8.1 (with the new definitions and a different implied constant) also holds in the bipartite case, that is,

$$|E(\operatorname{area}(R))-\operatorname{area}(R^\star)|=O\biggl(\frac{(\log n)^{3/2}}{n^{1/2}}\biggr). \tag{49}$$

Outline of proof. Lemma 8.2 still holds in the bipartite setting. We take B to be the box with sides $[0,2\log m]$ and $[0,2\log n]$. Lemma 8.5 still holds, and in Corollary 8.6 the last error term becomes $O((\log n)^3/n)$, but this is not important for the following.

In Lemma 8.7, the rank of the urn process at (x,y) can be written $\theta_1(x,y)+\theta_2(x,y)$, where each θ is a sum of $m+n$ independent 0,1-variables. The probability that the rank deviates by at least $\varepsilon(m+n)$ from its expectation is at most $2e^{-\varepsilon^2(m+n)/8}$. To make this equal to $1/(m+n)^2$, we put $\varepsilon=4(\log(m+n))^{1/2}/(m+n)^{1/2}$.

We divide B into regions B_1, B_2 and B_3 as before. The estimate of the area of B_2 in Lemma 8.8 still holds, although in the bipartite setting we also have to estimate the height of B_2 for a given value of x. Lemmas 8.9 and 8.10 hold with minor changes in the proofs, and as before, this makes the area of B_2 the main error term in (49). □

12.6. An example

In the bipartite case there is an optimization problem that provides an intermediate step between matching and 2-factor. Consider a bipartite graph with vertex sets U of size n, and V of size $2n$. Suppose that we wish to connect every vertex in U to two vertices in V in such a way that every vertex in V is used exactly once. This is the $2n$-flow problem with capacities 2 in U and 1 in V. Here too the region $R^\star$ is independent of n. The limit shape is given by the equation

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

and consequently the expected cost of the optimization problem converges to

$$\int_0^\infty y\,dx=\int_0^\infty -\log\left(1-e^{-x}-\frac{x}{2}e^{-x}\right)dx\approx 2.614522.$$

This is also predicted (non-rigorously) from the replica-cavity ansatz by a modification of the argument given in [29].

13. Exact formulas for matching and relaxed matching

In this section we show that the exact results on the flow problem established in §7 and §12 lead to some simple exact formulas for the linear flow problems when each vertex has weight 1 and capacity 1.

13.1. Expected values

The weight-1 capacity-1 case is simpler than the general case because conditioning on the time $T^{(i)}$ of the ith event in the urn process, the amount $T^{(i+1)}-T^{(i)}$ of time that we have to wait until the next vertex is drawn is independent of the set of vertices that were drawn up to time $T^{(i)}$. Since $n-i$ vertices remain, all of which have weight 1, the expected time until the next vertex is drawn is

$$\frac{1}{n-i}.$$

We can therefore compute the expected area of each of the rectangles that constitute the region R_h, both for the complete and the bipartite model.

In the bipartite case, the expected values of the increments $T^{(1)}, T^{(2)}-T^{(1)}, T^{(3)}-T^{(2)}, ..., T^{(k)}-T^{(k-1)}$ are $1/m, 1/(m-1), ..., 1/(m-k+1)$, respectively, for the process along the x-axis, and similarly $1/n, 1/(n-1), ..., 1/(n-k+1)$ along the y-axis. The rectangles that constitute R_k have expected areas

$$\frac{1}{(m-i)(n-j)}$$

for non-negative i and j such that $i+j<k$. This leads to the explicit result

$$EC_k=\sum_{\substack{i,j\geqslant 0\\ i+j<k}}\frac{1}{(m-i)(n-j)}. \tag{50}$$

Equation (50) was conjectured by D. Coppersmith and G. Sorkin [16], as a generalization of the formula

$$EC_n = 1+\frac{1}{4}+\frac{1}{9}+\ldots+\frac{1}{n^2} \tag{51}$$

for the complete n by n matching conjectured by Parisi in [36]. The formula (50) was established independently in [33] and [24], and it is easily verified that (50) specializes to (51) when $k=m=n$.

There is a direct analogue of (50) for the linear relaxation of matching in the complete graph. In this case, it is probably easier to derive the formula directly from (10) and Proposition 7.1, than to go via the urn process. When the capacities and weights are equal to 1, it follows by induction on k and n that

$$EC_{k/2} = \sum_{\substack{0\leqslant i\leqslant j \\ i+j<k}} \frac{1}{(n-i)(n-j)}. \tag{52}$$

If we put $m=n$ in (50), then (50) contains all terms of (52) twice, except those for which $i=j$. This means that if moreover $k=n$, so that (50) specializes to (51), then (52) becomes

$$\frac{1}{2}\left(1+\frac{1}{4}+\ldots+\frac{1}{n^2}+\sum_{n/2<i\leqslant n}\frac{1}{i^2}\right), \tag{53}$$

which in turn simplifies to

$$EC_{n/2} = 1-\frac{1}{4}+\frac{1}{9}-\ldots+(-1)^{n-1}\frac{1}{n^2}. \tag{54}$$

In addition to showing that $EC_{n/2}\to\frac{1}{12}\pi^2$, this alternating analogue of (51) shows that although the linear relaxation makes it possible to find feasible solutions also for odd n, the expected cost is always slightly higher than for even n.

13.2. The variance of bipartite matching

We establish an exact formula for the variance of the cost of the bipartite matching problem. This formula is described in [48]. Let C_n denote the cost of the minimum perfect matching in the n by n complete bipartite graph. The problem of establishing a good upper bound on $\operatorname{var}(C_n)$ for large n has been considered by several researchers [9], [32], [42]. The first proof that $\operatorname{var}(C_n)\to 0$ was obtained by M. Talagrand [42] with the method described in §9.

The cost of the minimum k-assignment is denoted by $C_{k,m,n}$, so that in particular $C_{n,n,n}=C_n$.

PROPOSITION 13.1. *The second moment of* $C_{k,m,n}$ *is given by*

$$E(C_{k,m,n}^2) = 2 \sum_{\substack{0\leqslant i_1\leqslant i_2\\ 0\leqslant j_2\leqslant j_1\\ i_1+j_1<k\\ i_2+j_2<k}} \frac{1}{(m-i_1)(m-i_2)(n-j_1)(n-j_2)} + 2 \sum_{\substack{0\leqslant i_1\leqslant i_2\\ 0\leqslant j_1\leqslant j_2\\ i_2+j_2<k-1}} \frac{1}{(m-i_1)(m-i_2)(n-j_1)(n-j_2)}. \tag{55}$$

Proof. Since all weights are equal to 1, the time between the ith and the $(i+1)$-th vertex to be drawn from U in the urn process is exponentially distributed with mean $1/(m-i)$ and independent of the set of vertices that have been drawn before. Similarly the time between the jth and the $(j+1)$-th vertex to be drawn from V is exponential with mean $1/(n-j)$. In the extended urn process there are two points (x_1, y_1) and (x_2, y_2). We first consider the event that $x_1<x_2$ and $y_1>y_2$. We fix i_1, i_2, j_1 and j_2, and consider the event that there are exactly i_1 elements of U that are drawn before x_1 and i_2 elements that are drawn before x_2, and similarly that exactly j_1 elements in V are drawn before y_1 and j_2 elements are drawn before y_2. The expected measure of the set of choices for x_1, x_2, y_1, y_2 is equal to

$$\frac{1}{(m-i_1)(m-i_2)(n-j_1)(n-j_2)}.$$

Notice that this holds also if $i_1=i_2$ or $j_1=j_2$. When we take the sum over all possible values of i_1, i_2, j_1, j_2, we obtain the first sum in (55). The case that $x_1>x_2$ and $y_1<y_2$ is similar and gives the factor 2.

In a similar way, the second sum comes from the case that $x_1<x_2$ and $y_1<y_2$, and in this case it is required that $(x_2, y_2)\in R_{k-1}$, which means that $i_2+j_2<k-1$. Again, the case that $x_1>x_2$ and $y_1>y_2$ is similar and gives a factor 2 on the second sum. □

It was shown in [48] that the case $k=m=n$ (perfect matching) leads to the simple formula

$$\operatorname{var}(C_n) = 5\sum_{i=1}^{n} \frac{1}{i^4} - 2\left(\sum_{i=1}^{n} \frac{1}{i^2}\right)^2 - \frac{4}{n+1}\sum_{i=1}^{n} \frac{1}{i^3}. \tag{56}$$

The derivation of (56) from (55) uses some identities from the remarkably similar calculation in [20] of the variance of the cost of the shortest path tree.

Notice that since $5\zeta(4)=2\zeta(2)^2=\frac{1}{18}\pi^4$, the first two terms of (56) essentially cancel as $n\to\infty$. By an elementary integral estimate,

$$\sum_{i=1}^{n} \frac{1}{i^s} = \zeta(s) - \frac{1}{(s-1)n^{s-1}} + O\left(\frac{1}{n^s}\right),$$

and therefore we have

$$\operatorname{var}(C_n) = 5\zeta(4) - 2\left(\zeta(2) - \frac{1}{n}\right)^2 - \frac{4}{n}\zeta(3) + O\left(\frac{1}{n^2}\right) = \frac{4\zeta(2) - 4\zeta(3)}{n} + O\left(\frac{1}{n^2}\right).$$

References

[1] Achlioptas, D., Naor, A. & Peres, Y., Rigorous location of phase transitions in hard optimization problems. *Nature*, 435 (2005), 759–764.

[2] Achlioptas, D. & Peres, Y., The threshold for random k-SAT is $2^k \log 2 - O(k)$. *J. Amer. Math. Soc.*, 17 (2004), 947–973.

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

[4] — The $\zeta(2)$ limit in the random assignment problem. *Random Structures Algorithms*, 18 (2001), 381–418.

[5] — 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] Aldous, D. & Bandyopadhyay, A., A survey of max-type recursive distributional equations. *Ann. Appl. Probab.*, 15 (2005), 1047–1110.

[7] Aldous, D. & Percus, A. G., Scaling and universality in continuous length combinatorial optimization. *Proc. Natl. Acad. Sci. USA*, 100:20 (2003), 11211–11215.

[8] Aldous, D. & Steele, J. M., The objective method: probabilistic combinatorial optimization and local weak convergence, in *Probability on Discrete Structures*, Encyclopaedia Math. Sci., 110, pp. 1–72. Springer, Berlin–Heidelberg, 2004.

[9] Alm, S. E. & Sorkin, G. B., Exact expectations and distributions for the random assignment problem. *Combin. Probab. Comput.*, 11 (2002), 217–248.

[10] Bacci, S. & Miranda, E. N., The traveling salesman problem and its analogy with two-dimensional spin glasses. *J. Stat. Phys.*, 56 (1989), 547–551.

[11] Bandyopadhyay, A. & Gamarnik, D., Counting without sampling: asymptotics of the log-partition function for certain statistical physics models. *Random Structures Algorithms*, 33 (2008), 452–479.

[12] Boettcher, S. & Percus, A., Nature's way of optimizing. *Artificial Intelligence*, 119 (2000), 275–286.

[13] Brunetti, R., Krauth, W., Mézard, M. & Parisi, G., Extensive numerical simulation of weighted matchings: total length and distribution of links in the optimal solution. *Europhys. Lett.*, 14 (1991), 295–301.

[14] Buck, M. W., Chan, C. S. & Robbins, D. P., On the expected value of the minimum assignment. *Random Structures Algorithms*, 21 (2002), 33–58.

[15] Cerf, N. J., Boutet de Monvel, J., Bohigas, O., Martin, O. C. & Percus, A. G., The random link approximation for the Euclidean traveling salesman problem. *J. Physique*, 7 (1997), 117–136.

[16] Coppersmith, D. & Sorkin, G. B., Constructive bounds and exact expectations for the random assignment problem. *Random Structures Algorithms*, 15 (1999), 113–144.

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

[18] — On random symmetric travelling salesman problems. *Math. Oper. Res.*, 29 (2004), 878–890.

[19] Guerra, F., Broken replica symmetry bounds in the mean field spin glass model. *Comm. Math. Phys.*, 233 (2003), 1–12.

[20] van der Hofstad, R., Hooghiemstra, G. & Van Mieghem, P., The weight of the shortest path tree. *Random Structures Algorithms*, 30 (2007), 359–379.
[21] Karp, R. M., An upper bound on the expected cost of an optimal assignment, in *Discrete Algorithms and Complexity* (Kyoto, 1986), Perspect. Comput., 15, pp. 1–4. Academic Press, Boston, MA, 1987.
[22] Kirkpatrick, S. & Toulouse, G., Configuration space analysis of travelling salesman problems. *J. Physique*, 46:8 (1985), 1277–1292.
[23] Krauth, W. & Mézard, M., The cavity method and the travelling salesman problem. *Europhys. Lett.*, 8 (1989), 213–218.
[24] Linusson, S. & Wästlund, J., A proof of Parisi's conjecture on the random assignment problem. *Probab. Theory Related Fields*, 128 (2004), 419–440.
[25] Mézard, M. & Montanari, A., *Information, Physics, and Computation*. Oxford Graduate Texts. Oxford University Press, Oxford, 2009.
[26] Mézard, M. & Parisi, G., Replicas and optimization. *J. Physique*, 46 (1985), 771–778.
[27] — Mean-field equations for the matching and the travelling salesman problems. *Europhys. Lett.*, 2 (1986), 913–918.
[28] — A replica analysis of the travelling salesman problem. *J. Physique*, 47 (1986), 1285–1296.
[29] — On the solution of the random link matching problems. *J. Physique*, 48 (1987), 1451–1459.
[30] Mézard, M., Parisi, G. & Virasoro, M. A., *Spin Glass Theory and Beyond*. World Scientific Lecture Notes in Physics, 9. World Scientific, Teaneck, NJ, 1987.
[31] Mézard, M., Parisi, G. & Zecchina, R., Analytic and algorithmic solutions of random satisfiability problems. *Science*, 297 (2002), 812.
[32] Nair, C., *Proofs of the Parisi and Coppersmith–Sorkin Conjectures in the Finite Random Assignment Problem*. Ph.D. Thesis, Stanford University, Stanford, CA, 2005.
[33] Nair, C., Prabhakar, B. & Sharma, M., Proofs of the Parisi and Coppersmith–Sorkin random assignment conjectures. *Random Structures Algorithms*, 27 (2005), 413–444.
[34] Parisi, G., A sequence of approximated solutions to the S–K model for spin glasses. *J. Physique*, 13 (1980), L–115.
[35] — Spin glasses and optimization problems without replicas, in *Le hasard et la matière* (Les Houches, 1986), pp. 525–552. North-Holland, Amsterdam, 1987.
[36] — A conjecture on random bipartite matching. Preprint, 1998. `arXiv:cond-mat/9801176`.
[37] Percus, A. G., *Voyageur de commerce et problèmes stochastiques associés*. Ph.D. Thesis, Université Pierre et Marie Curie, Paris, 1997.
[38] Percus, A. G. & Martin, O. C., Finite size and dimensional dependence in the Euclidean traveling salesman problem. *Phys. Rev. Lett.*, 76 (1996), 1188–1191.
[39] — The stochastic traveling salesman problem: Finite size scaling and the cavity prediction. *J. Stat. Phys.*, 94 (1999), 739–758.
[40] Sherrington, D. & Kirkpatrick, S., Solvable model of a spin glass. *Phys. Rev. Lett.*, 35 (1975), 1792–1796.
[41] Sourlas, N., Statistical mechanics and the travelling salesman problem. *Europhys. Lett.*, 2 (1986), 919–923.
[42] Talagrand, M., Concentration of measure and isoperimetric inequalities in product spaces. *Inst. Hautes Études Sci. Publ. Math.*, 81 (1995), 73–205.
[43] — *Spin Glasses: A Challenge for Mathematicians*. Modern Surveys in Mathematics, 46. Springer, Berlin–Heidelberg, 2003.
[44] — The Parisi formula. *Ann. of Math.*, 163 (2006), 221–263.
[45] Vannimenus, J. & Mézard, M., On the statistical mechanics of optimization problems of the travelling salesman type. *J. Physique*, 45 (1984), 1145–1153.

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

[47] Wästlund, J., A proof of a conjecture of Buck, Chan, and Robbins on the expected value of the minimum assignment. *Random Structures Algorithms*, 26 (2005), 237–251.

[48] — The variance and higher moments in the random assignment problem. Linköping Studies in Mathematics, 8. Preprint, 2005.

[49] Weitz, D., Counting independent sets up to the tree threshold, in *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, pp. 140–149. ACM, New York, 2006.

Johan Wästlund
Department of Mathematical Sciences
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
wastlund@chalmers.se

Received April 3, 2008
Received in revised form June 9, 2009