Notes on random optimization problems

Johan Wästlund

Department of Mathematics
Chalmers University of Technology, S-412 96 Göteborg, Sweden
wastlund@chalmers.se

2nd December 2008

These notes are under construction. They constitute a combination of what I have said in the lectures, what I will say in future lectures, and what I will not say due to time constraints. Some sections are very brief, and this is generally because they are not yet written. Some of the "problems and exercises" describe things that I am actually going to write down in detail in the text. This is because I have used the problems & exercises section in this way to take short notes of things I should not forget to mention.

1 Introduction (23/9)

I started by mentioning the Beardwood-Halton-Hammersley theorem [13] on euclidean TSP. Let L_n be the length of the minimum traveling salesman tour through n uniform independent random points in the unit square. Then it is known that for some β,

$$\frac{L_n}{\sqrt{n}} \xrightarrow{p} \beta.$$

Reminder: Convergence in probability to a constant β means that for every $\epsilon > 0$, $P(|L_n/\sqrt{n} - \beta| < \epsilon) \to 1$ as $n \to \infty$. Through simulations it is believed that $\beta \approx 0.71$, but the best rigourous bounds are embarassingly far apart, something like

$$0.625 < \beta < 0.922.$$

The situation is similar for the minimum spanning tree problem [39]. One of the reasons I mention this is to contrast with the situation in the *mean* *field model*, where the distances between points (in an abstract geometry) are taken as independent random variables. We will return to the euclidean model later, when we discuss methods for proving concentration, in particular the Talagrand isoperimetric inequality.

I listed some classical optimization problems on graphs:

- Minimum Spanning Tree (MST)
- Simple matching (simple = non-bipartite)
- Bipartite matching (=assignment problem)
- Traveling Salesman (TSP)

These problems have in common that we are required to find an edge set of minimum total cost satisfying certain conditions (for instance in the case of spanning tree, to contain a path between any pair of points). In the mean field model (where the triangle inequality need not hold) we can add one more problem to the list:

- Shortest Path

This does not make sense in the euclidean model because the shortest path is always a straight line.

1.1 Elementary bounds on the cost in the mean field model

For the moment, consider the bipartite matching problem (a. k. a. the assignment problem) with edge costs taken from uniform [0,1] distribution. What can we say about the cost of the optimal solution? First there is a rather elementary lower bound on the expectation. The minimum cost edge from a given node has cost

$$\min(X_1, \dots, X_n).$$

The expectation of this can be found by calculus:

$$\begin{aligned} E\min(X_1, \dots, X_n) &= \int_0^1 P(\min(X_1, \dots, X_n) > t)\, dt \\ &= \int_0^1 P(X_i > t)^n\, dt = \int_0^1 (1-t)^n\, dt = \frac{1}{n+1}. \end{aligned} \quad (1)$$

There is also a cute trick that gives this result without integration. It is not necessary here, but a similar one will be needed (and therefore promoted to *method*) in order to solve harder problems like finding the limit in the mean field TSP. It goes like this, let X_{n+1} be another random number generated in the same way, that is, uniform in [0,1] and independent of the others. What is

$$P(X_{n+1} < \min(X_1, \ldots, X_n))?$$

On one hand it is $1/(n+1)$ by symmetry, since X_{n+1} has the same chance as any of $X_1, \ldots, X_n$ of being smallest. On the other hand it is (by definition of uniform distribution) $E \min(X_1, \ldots, X_n)$.

It follows that (letting A_n denote the cost of the minimum assignment)

$$EA_n \geq \frac{n}{n+1},$$

and in particular it is bounded away from zero. We now turn to the less trivial upper bound.

Theorem 1.1 (Walkup 1979). *There is a constant C independent of n such that*

$$EA_n < C.$$

A couple of modifications will simplify the computations.

Modification 1: We consider exp(1)-distribution (exponential of rate 1). This distribution stochastically dominates uniform[0,1], so it is sufficient to prove Walkup's statement in the exp(1)-setting.

Modification 2: We introduce multiple edges between each pair of vertices. The potentially infinite sequence of edges have costs given by the times of a Po(1)-process. Since the first event of such a process comes after an exp(1)-distributed time, this modification obviously doesn't change the optimal solution.

We now give each edge a random orientation (by coin flipping) and consider the set of edges obtained by choosing the five cheapest edges from each vertex (on both sides of the graph). We want to show that with high probability, this set contains a perfect matching. Let V_1 and V_2 be the two vertex-sets.

Theorem 1.2 (Hall's criterion). *A set of edges contains a perfect matching iff for every subset S of V_1,*

$$|\Gamma(S)| \geq |S| \,. \tag{2}$$

Here $\Gamma(S)$ denotes the set of neighbors of S, that is, the set of vertices in V_2 that have an edge to some vertex in S. I gave a handwaving argument for Hall's theorem. For a proof and more discussion of Hall's theorem we refer to the Wikipedia article on "Marriage theorem".

If (2) fails for some $S \subseteq V_1$, then for the minimal S for which it fails, we must have $|\Gamma(S)| = |S| - 1$. Then let $T = V_2 - \Gamma(S)$. Then $|S| + |T| = n + 1$, and since $\Gamma(T) \subseteq V_1 - S$, T is also a counterexample to (2). Therefore if (2) fails, then it must fail for some $S \subseteq V_1$ or $T \subseteq V_2$ of size at most

$$\left\lceil \frac{n}{2} \right\rceil .$$

In order to estimate the probability of (2) failing, we use the following standard estimate:

Lemma 1.3.

$$\binom{n}{s} \leq \left(\frac{ne}{s}\right)^s .$$

This is a quite rough but still useful inequality.

Proof. We have

$$\binom{n}{s} \leq \frac{n^s}{s!},$$

and therefore it suffices to show that

$$s! \geq \left(\frac{s}{e}\right)^s ,$$

or equivalently that

$$\log s! = \log 2 + \log 3 + \cdots + \log s \geq s \log s - s.$$

Here

$$s \log s - s = \int_0^s \log x \, dx.$$

By drawing some boxes, we see that this is smaller than the right hand side. $\square$

The merit of this lemma is just its simplicity. It is easy to get better estimates from the Stirling formula (of which we will say more later).

$$P(\text{Hall's criterion fails}) \leq 2\sum_{s=2}^{\lceil n/2\rceil}\binom{n}{s}\binom{n}{s-1}\left(\frac{s-1}{n}\right)^{5s}$$
$$\leq 2\sum_{s=2}^{\lceil n/2\rceil}\binom{n}{s}^2\left(\frac{s}{n}\right)^{5s} \leq 2\sum_{s=2}^{\lceil n/2\rceil}\left(\frac{ne}{s}\right)^{2s}\left(\frac{s}{n}\right)^{5s} = 2\sum_{s=2}^{\lceil n/2\rceil}e^{2s}\left(\frac{s}{n}\right)^{3s}. \quad (3)$$

Here the summand is log-convex:

$$\frac{d^2}{ds^2}\left(\log\left(e^{2s}\left(\frac{s}{n}\right)^{3s}\right)\right) = \frac{3}{s} \geq 0.$$

Therefore the largest term is either the first one or the last (which one it is may depend on n). If we plug in $s = 2$, we get

$$e^4\left(\frac{2}{n}\right)^6 = O\left(\frac{1}{n^6}\right).$$

If on the other hand $s = n/2$ or $s = n/2 + 1/2$, the term is at most

$$e^{n+1}\left(\frac{1}{2}+\frac{1}{n}\right)^{3n/2} = e\cdot\left(\frac{e}{2^{3/2}}\right)^n\cdot\left[\left(1+\frac{1}{n/2}\right)^{n/2}\right]^3 \leq e^4\cdot\left(\frac{e}{2^{3/2}}\right)^n.$$

Here the “3” is the number of edges from each node minus 2. It is chosen very carefully so that $e/2^{3/2} < 1$. Therefore the size of the last term decreases exponentially, so that for large n, the first term ($s = 2$) will be largest. There are $O(n)$ terms in the sum, and therefore

$$P(\text{Hall's criterion fails}) = O\left(\frac{1}{n^5}\right).$$

This is of course not the same thing as Walkup's theorem. To finish the proof, we use another trick, randomly coloring the edges. We let the edges be red with probability p and blue with probability $1-p$, independently. The blue edges between each pair of vertices form a Po$(1-p)$-process, and with the method above, we can, with probability $1 - O(n^{-5})$, find a solution involving the five cheapest blue edges directed from each vertex. In the cases of failure, we take the cheapest red edges matching the first vertex in V_1 to the first in V_2 etc. This gives a method of obtaining an assignment. The blue edges directed from a particular vertex form a $Po((1-p)n/2)$-process, and the five cheapest edges therefore have expected costs

$$\frac{2}{(1-p)n}, \frac{4}{(1-p)n}, \dots, \frac{10}{(1-p)n}.$$

The expected cost of an edge in this set is therefore $6/((1-p)n)$.

In the cases of failure we use edges of expected cost $1/p$. Since there are n edges in an assignment, we get

$$EA_n \leq \frac{6}{1-p} + O\left(\frac{1}{n^5}\right) \cdot \frac{n}{p}.$$

To establish Theorem 1.1 it suffices to take $p = 1/2$, but if we choose p as a function of n so that $p \to 0$ but at the same time, $pn^4 \to \infty$, we see that we can take any $C > 6$.

Walkup's theorem will later be superseded by the exact formula (valid in the exp(1) setting):

$$EA_n = 1 + \frac{1}{4} + \frac{1}{9} + \cdots + \frac{1}{n^2}, \tag{4}$$

of which we will have more to say. There is, however, one sense in which the elementary argument outlined above gives more information than (4).

Theorem 1.4. *There are positive constants c and C such that when $n \to \infty$,*

$$P(c < A_n < C) \to 1.$$

Proof. For the lower bound we took the sum of the cheapest edges from each vertex in V_1. Since the terms are independent it follows from the Law of Large Numbers that their sum converges in probability to 1 (and therefore any $c < 1$ will do). Actually, letting $X_1, \dots, X_n$ be independent uniform in [0,1] and $Y = \min(X_1, \dots, X_n)$, we can use the trick again to compute $\mathrm{var}(Y)$. This is because EY^2 is the probability that X_{n+1} and X_{n+2} are both smaller than Y, and by symmetry, this probability is

$$\frac{1}{\binom{n+2}{2}} = \frac{2}{(n+2)(n+1)}.$$

Hence

$$\mathrm{var} Y = EY^2 - (EY)^2 = \frac{2}{(n+2)(n+1)} - \frac{1}{(n+1)^2} \le \frac{1}{n^2}.$$

The variance of the sum of n independent Y's is therefore smaller than $1/n$. Of course this says nothing about $\mathrm{var} A_n$, that is another story.

For the high probability upper bound (C), we again take the five cheapest edges from each vertex, and suppose that Hall's criterion (2) holds. In that case, A_n is upper bounded by the sum of the fifth cheapest edges from each vertex (both in V_1 and in V_2 so there are $2n$ of them). The fifth cheapest edge from a given vertex is of the form $X_1 + \cdots + X_5$, where X_i are independent exponentials of rate $n/2$. Summing over all vertices, we get a sum of $10n$ exponentials of rate $n/2$. Each term has mean $2/n$ and variance $4/n^2$ which means the sum has mean 20 and variance $40/n \to 0$. This rather sloppy estimate thus shows that we can take any $C > 20$. $\square$

One way of combining Walkup's theorem with the high probability upper bound in one statement is to say that

$$E \max(A_n, 20) \to 20.$$

This is a strong form of what is sometimes called *uniform integrability*. It shows that the limit distribution of A_n, if it exists, has compact support. It is therefore hard to imagine any type of large n behavior other than convergence in probability to some constant, but in the 1980's, no methods were available that would allow a proof of this.

2 Bounds on the cost of the TSP

Using some more tricks, we prove the analogous results for the bipartite TSP. This implies the corresponding bounds for spanning tree, and for the complete graph. The setting is again the same, n vertices on each side, and Po(1) processes for the edge costs between each pair of vertices.

We color the edges red and blue with probabilities $1/2 - p$ each, and save a fraction $2p$ of the edges (which we give random orientation) for other purposes. Step 1 is to find the optimal assignments of red and blue edges respectively. These matchings are independent, which means that relative to the red solution, the blue solution will describe a random permutation of n elements, drawn uniformly among the $n!$ possibilities.

If this permutation is cyclic, we have a tour. The bounds on the costs of the red and blue assignments will then carry over and give a high probability bound on the cost L_n of the tour. In case there are several cycles, we will use the extra edges to "patch" the cycles into a tour. In order for this to work, we have to show that with high probability there aren't that many cycles. To this end, we use the following quite clean result:

Lemma 2.1. *In a uniform random permutation of n elements, let Z_n be the number of cycles. Then*

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

Proof. The statement is obviously true if $n = 1$. We proceed by induction. A permutation of n elements is constructed from one of $n-1$ elements by either mapping the element n to itself (probability $1/n$) or inserting it in an already existing cycle. Hence

$$E(2^{Z_n}) = \frac{1}{n} \cdot E(2^{Z_{n-1}+1}) + \frac{n-1}{n} E(2^{Z_{n-1}}) = \frac{n+1}{n} E(2^{Z_{n-1}}).$$

□

It follows that

$$P(Z_n > 4\log n) = P(2^{Z_n} > 2^{4\log n}) = P(2^{Z_n} > n^{4\log 2}) \leq \frac{n+1}{n^{4\log 2}} = O\left(\frac{1}{n^{4\log 2 - 1}}\right). \quad (5)$$

And $4\log 2 - 1 \approx 1.77$.

If there are at most $4\log n$ cycles, then one of them must have size at least $n/(4\log n)$. We now take the remaining cycles, one by one, and patch them to the main cycle in the obvious way by using the extra edges, and to preserve independence we use edges directed from the small cycle to the main cycle. In one step of patching, we choose arbitrarily two adjacent vertices v_1 and v_2 of the small cycle, and consider the extra edges from v_1 and v_2 to adjacent pairs of vertices in the main cycle. We are therefore considering the minimum of at least $n/(4\log n)$ variables, each of which is a sum of two exponentials of rate p. It follows from Lemma 2.2 below that the expected cost of such a patching is

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

There are at most $4\log n$ patchings to perform, so the expected total cost of the patchings (given that there are at most $4\log n$ cycles) is

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

If there are more than $4\log n$ cycles, we use the extra edges to construct a tour of expected cost n/p. In all, we therefore have

$$EL_n \leq \frac{2}{1/2-p}\cdot EA_n + O\left(\frac{(\log n)^{3/2}}{pn^{1/2}}\right) + O\left(\frac{1}{n^{4\log 2-1}}\right)\cdot\frac{n}{p}.$$

If for instance we choose $p = n^{-1/4}$, it follows that $EL_n \leq 24 + o(1)$. Again an estimate of the variance of the upper bound (given at most $4\log n$ cycles) will show that there is a C such that $P(L_n < C) \to 1$ as $n \to \infty$.

The lemma that we promised to prove is the following:

Lemma 2.2. *For $1 \leq i \leq m$, let $Z_i = X_i + Y_i$, where X_i and Y_i are* exp(1)*-variables, all independent. Then*

$$E\min(Z_1,\dots,Z_m) = O\left(\frac{1}{\sqrt{m}}\right).$$

Proof. Let $1 \leq k \leq m$, we will optimize the choice later. Now suppose without loss of generality that $X_1,\dots,X_k$ are the k smallest of $X_1,\dots,X_m$. The expected value of $\max(X_1,\dots,X_k)$ is

$$\frac{1}{m}+\frac{1}{m-1}+\cdots+\frac{1}{m-k+1} \leq \frac{k}{m-k}.$$

Moreover, the expected value of $\min(Y_1,\dots,Y_k)$ is $1/k$. Hence

$$E\min(Z_1,\dots,Z_m) \leq \frac{k}{m-k}+\frac{1}{k}.$$

The statement follows by taking $k = \sqrt{m}+O(1)$. □

Another proof is sketched in Exercise 7.

3 Matching and TSP with the statistical mechanics approach (30/9)

The following is a combination of ideas originating in the papers by Mézard and Parisi, with the simplification by Aldous, and some refinements. It can be described as the physics approach without physics. The replica and cavity methods are often labeled non-rigorous, but to be fair, it should be pointed out that a large part of the physics literature on combinatorial optimization is devoted to the motivation of why the calculations should give the right answer. This part, which is what is hard to understand, is here left out. What remains is only a calculation leading from a certain equation of "self consistency" directly to what we now know is the right answer. What is presented here is therefore much less rigorous than the physics papers where the ideas are originally described.

We begin with simple matching on the complete graph. A perfect matching requires even n, but we consider a certain relaxation. We allow for leaving a vertex out in a solution, but introduce a certain cost for doing so. In order for this to make sense, the punishment for leaving out vertices must scale like $1/n$, and to facilitate notation, we rescale the entire problem so that the edge costs are exponential of rate $1/n$ (mean n), and we therefore expect the cost per edge in the solution to converge in distribution.

Let the cost of leaving out a vertex be c. Then only edges of cost smaller than $2c$ are relevant, and this edge set will become very sparse for large n, so that the graph becomes locally tree-like. Let T_0 be the graph seen as a tree rooted at a particular vertex v_0, and let $v_1, v_2, \dots$ be the neighbors of v_0 in order of increasing edge costs, which we denote by $\xi_1, \xi_2, \dots$. Let T_i be the sub-tree rooted at v_i. We will regard these sub-trees as disjoint, although in reality they are of course connected somewhere. For a vertex set S, let $C(S)$ be the cost of the relaxed matching problem on the subgraph induced by S.

Now notice that

$$C(T_0) - C(T_0 - v_0) = \min(c, \xi_i - C(T_i - v_i) + C(T_i)). \tag{6}$$

To see this, think of the solution on $T_0 - v_0$ as given, and consider the problem on T_0, and where to connect v_0. If we put $X_i = C(T_i) - C(T_i - v_i)$ for $i \geq 0$, then the equation can be written

$$X_0 = \min(c, \xi_i - X_i). \tag{7}$$

Here X_i, for $i \geq 1$, are assumed to be independent variables of the same distribution as X_0. The ξ_i's are, in the pseudodimension 1 case, a Poisson process of rate 1. In this case it turns out that there is a simple explicit solution, but the method applies also for other distributions, as has been shown by Mézard, Parisi and others.

We first rewrite (7) in terms of its distribution. Let

$$F(u) = \frac{1}{2}P(X = u) + P(X > u),$$

where X has the distribution of X_i (clearly the only point where X has nonzero probability is c). It is obvious from (7) that $X_0 \leq c$, and since the ξ_i's are nonnegative and $X_i \leq c$, we also have $X \geq -c$. The distribution of X is a point mass at c plus a continuous distribution on $[-c, c]$. We focus on the nontrivial case $u \in [-c, c]$. The statement that $X_0 > u$ is then equivalent to saying that for every i,

$$\xi_i - X_i > u. \tag{8}$$

For a given ξ_i, the probability that (8) does not hold is

$$P(\xi_i - X_i \leq u) = P(X_i \geq \xi_i - u) = F(\xi_i - u).$$

$F(u)$ is the probability that there is no event at all in the Poisson process of such ξ_i's (for which (8) does not hold), which means that (7) is equivalent to

$$F(u) = \exp\left(-\int_0^\infty F(\xi - u)\, d\xi\right) = \exp\left(-\int_{-u}^\infty F(t)\, dt\right). \tag{9}$$

3.1 Solving for F

For this problem, there is a certain symmetry which can be used in order to obtain an explicit solution, but which is in general not necessary. Differentiating (9), we obtain

$$F'(u) = -F(u)F(-u). \tag{10}$$

This implies that $F'(u) = F'(-u)$, so that the continuous part of the distribution of X is symmetric under changing u to $-u$. Therefore if $-c < u < c$, then $F(-u) = A - F(u)$, where $A = 1 + P(X = c)$. Equation (10) can now be written

$$F'(u) = -F(u) \cdot (A - F(u)). \tag{11}$$

The advantage is that (11) is a so-called ordinary differential equation, meaning that it expresses $F'(u)$ in terms of $F(u)$ only, not in terms of $F(-u)$. It can be solved by the standard technique of "integrating factor":

$$-\frac{F'(u)}{F(u)(A-F(u))} = \frac{d}{du}\left(\frac{1}{A}\log\left(\frac{A-F(u)}{F(u)}\right)\right) = 1.$$

Hence for some constant B,

$$\log\left(\frac{A-F(u)}{F(u)}\right) = Au + B.$$

Since $F(0) = A/2$, we get $B = 0$. Solving for $F(u)$, we obtain

$$F(u) = \frac{A}{1+e^{Au}}.$$

Using the fact that $F(-c) = 1$, we get

$$1 = \frac{A}{1+e^{-Ac}},$$

so that

$$c = \frac{-\log(A-1)}{A}.$$

If instead we express this in terms of the proportion $p = 2 - A$ of vertices that are used in the solution (for which we don't pay the punishment), then

$$c = \frac{-\log(1-p)}{2-p}.$$

The fact that c can be found in terms of p is quite interesting and gives a very concrete conjecture about the behavior of incomplete matching problems. Suppose that for large n we find the minimum k-matching, that is, k vertex-disjoiont edges, where $k = pn/2 + O(1)$, so that a proportion p of the vertices participate in the solution. Then this is roughly equivalent of having a punishment of c. It is therefore reasonable to believe that the most expensive edge in the solution will have cost about $2c = -2\log(1-p)/(2-p)$. For instance, if $p = 1/2$ then $2c = 4\log(2)/3 \approx 0.924$.

3.2 The total cost of the solution

Having found the distribution of X, the question is if we can find the expected cost of the optimal solution. We therefore consider an arbitrarily chosen edge e, and compute the expected contribution of e to the total cost. We think of this edge as going between two vertices v_1 and v_2 that are roots of two infinite trees. If the edge e has cost z, then e will participate in the minimum solution if

$$z + C(T_1 - v_1) + C(T_2 - v_2) \leq C(T_1) + C(T_2).$$

This is equivalent to

$$z \leq X_1 + X_2,$$

where $X_1 = C(T_1) - C(T_1 - v_1)$ and $X_2 = C(T_2) - C(T_2 - v_2)$ are independent variables of the distribution of X. The contribution of e can therefore be computed as a scaling factor times

$$J = \int_0^\infty zP(X + Y \geq z)\, dz.$$

The cost is exponential of mean n, and therefore the density of the cost can be approximated by $1/n$ throughout the positive reals. The contribution from a single edge to the total cost is therefore

$$\frac{1}{n} \cdot J,$$

and since there are $\sim n^2/2$ edges, the total cost is

$$\sim n/2 \cdot J.$$

We have

$$\begin{aligned} J = \int_0^\infty zP(X_1 + X_2 \geq z)\, dz = E\left[\frac{((X_1 + X_2)^+)^2}{2}\right] \\ = \int_{-\infty}^\infty E\left((X_1 - u)I(-X_2 < u < X_1)\right)\, du \\ = \int_{-\infty}^\infty P(X_2 \geq -u) \cdot E((X_1 - u)^+)\, du \\ = \int_{-\infty}^\infty F(-u) \cdot \int_u^\infty F(x)\, dx\, du. \quad (12) \end{aligned}$$

In principle we can plug in the known expressions for F and compute this double integral, taking into account that the expression we have found for F is only valid in the interval $[-c, c]$. However, there is a simpler method that introduces the function

$$f(u) = \int_0^\infty F(\xi - u)\, d\xi = \int_{-u}^\infty F(t)\, dt. \tag{13}$$

In the physics literature, f is sometimes called the *order parameter function*. Differentiating (13) we get

$$f'(u) = F(-u). \tag{14}$$

The integral J is then simply

$$J = \int_{-\infty}^\infty f'(u) f(-u)\, du. \tag{15}$$

First, the handling of the boundary conditions becomes simpler since $f(u) = 0$ if $u < -c$: The integrand is zero outside the interval $[-c, c]$. Second, there is a simple interpretation of (15) as the area under the curve (in the positive quadrant) when $f(-u)$ is plotted against $f(u)$.

Since we know F explicitly, we can find f. If $-c \leq u \leq c$, then

$$f(u) = \int_{-u}^c F(t)\, dt = \int_{-u}^{-\log(1-p)/(2-p)} \frac{2-p}{1+e^{(2-p)u}}\, dt = \log\left(\frac{1+e^{(2-p)u}}{2-p}\right). \tag{16}$$

Hence

$$J = \int_{-c}^c f'(u) f(-u)\, du = \int_{\log(1-p)/(2-p)}^{-\log(1-p)/(2-p)} \frac{(2-p)e^{(2-p)u} \log\left(\frac{1+e^{-(2-p)u}}{2-p}\right)}{1+e^{(2-p)u}}\, du.$$

The only reasonable way to handle such an expression seems to be to differentiate it with respect to p. This is quite an exercise, but it can be verified that the derivative is

$$\frac{-2\log(1-p)}{2-p}.$$

Since $J = 0$ when $p = 0$, it follows that

$$J = \int_0^p \frac{-2\log(1-t)}{2-t}\, dt. \tag{17}$$

However, a simpler way to proceed is to find directly an equation that relates $f(u)$ to $f(-u)$. In fact this relation is

$$e^{-f(u)} + e^{-f(-u)} = 2 - p. \tag{18}$$

Once we have written down (18), it is easy to see that it follows from (16). We return in a moment to how we could have found the relation (18). If we plot $f(-u)$ against $f(u)$ in the x-y-plane, then the curve is given by

$$e^{-x} + e^{-y} = 2 - p, \tag{19}$$

or equivalently

$$y = -\log(2 - p - e^{-x}).$$

The integral J is simply the area under this curve. Hence for $p = 1$,

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

For smaller values of p we must take into account that y becomes zero at the point $x = -\log(1-p)$, so that

$$J = \int_0^{-\log(1-p)} -\log(2 - p - e^{-x})\,dx.$$

Again it can be verified that this is equal to (17) by differentiating with respect to p. This time it is actually doable.

3.3 A simpler solution

Starting from the equation (9), it is natural to first try to obtain an explicit description of the distribution of X by solving for F. Having in mind that the total cost of the solution is given by (15), we realize that it is only necessary to find the relation between $f(u)$ and $f(-u)$. This relation is given by (18), and since the rest is only a matter of routine calculus, we can regard (18) as essentially being the solution to the problem.

Let us go back to equation (9). It can be written

$$F(u) = e^{-f(u)}, \tag{20}$$

and at the same time, as we already noted in (14), $f'(u) = F(-u)$. Therefore an equivalent form of (9) in terms of f is

$$f'(u) = e^{-f(-u)}. \tag{21}$$

By replacing u with $-u$, we also have

$$f'(-u) = e^{-f(u)}. \tag{22}$$

With hindsight, we realize that (18) can be derived from (21) and (22), without first obtaining a closed expression for f. We simply differentiate the left hand side of (18), and verify that what we get is zero because of (21) and (22). Once we know that the left hand side of (18) is constant, the constant can be found by setting $u = c$ in (20).

3.4 The TSP

After the success in calculating the spectacular $\pi^2/12$ limit for the simple matching problem, Mézard and Parisi took the method one step further, attacking the famous traveling salesman problem. In a traveling salesman tour, each vertex has degree 2, and the relaxation we consider consists in requiring each vertex to have degree 2, with a punishment of c if the vertex has degree 1, and $2c$ if it has degree 0. Hence we allow for solutions that consist of several cycles, without punishment. We will not motivate here why allowing such solutions does not decrease the asymptotic cost. This is a fact that was proved by Alan Frieze in 2004 [36], but in view of the general lack of rigour in the approach, this is currently a minor problem.

The equation corresponding to (6) is

$$C(T_0) - C(T_0 - v_0) = \min[2](c, c, \xi_i - C(T_i - v_i) + C(T_i)).$$

Here $T_i - v_i$ does not mean that we remove v_i. It means that we decrease the required degree of v_i to 1. Just as for the matching problem, this can work either for or against a cheaper solution. On one hand we don't have to pay the punishment if v_i has degree 1, on the other hand, we do not have the right to use v_i more than once, which means that the cost of handling other vertices can increase.

min[2] means second-smallest. To see why the equation holds, again think of the solution for $T_0 - v_0$ as given, and consider what to do when the required degree of v_i increases. Either we pay the punishment, or we use a second edge from v_0.

The TSP analog of (9) is

$$F(u) = (1 + f(u))\, e^{-f(u)}, \tag{23}$$

since $F(u)$ is now the probability of having at most one point in the region $\xi_i - X_i \leq u$. Again f is defined by (13) and satisfies (14), so that the equation can be written in terms of f only as

$$f'(-u) = (1 + f(u))e^{-f(u)}. \tag{24}$$

Since the total cost of the solution is still given by (15), we can try to find a relation between $f(u)$ and $f(-u)$ similar to (18). Such a relation is supposed to come from integrating an expression which is equal to zero because of (24) and the accompanying equation

$$f'(u) = (1 + f(-u))e^{-f(-u)}. \tag{25}$$

By multiplying (24) and (25), we find that

$$\begin{aligned} f'(u)(1 + f(u))e^{-f(u)} &- f'(-u)(1 + f(-u))e^{-f(-u)} \\ &= f'(u)f'(-u) - f'(u)f'(-u) = 0. \end{aligned} \tag{26}$$

Integrating, and for convenience changing sign, we find that

$$(2 + f(u))e^{-f(u)} + (2 + f(-u))e^{-f(-u)}$$

is a constant, that we expect to depend on c.

For simplicity, we here only consider the case $c = \infty$. Then the boundary condition is just that when one of $f(u)$ and $f(-u)$ goes to infinity, the other will go to zero. Therefore the points $(f(u), f(-u))$ lie on the curve

$$(2 + x)e^{-x} + (2 + y)e^{-y} = 2. \tag{27}$$

This is as explicit as the solution gets. There seems to be no way of expressing y in terms of elementary functions of x. It can be verified by numerical integration that for the TSP, the integral (15), that is, the area under the curve given by (27), evaluates to approximately 4.083.

4 Euler's identity (not part of any lecture)

The evaluation of the integral

$$\int_0^\infty \log(1 - e^{-x})\, dx \tag{28}$$

(and several similar integrals) boils down to evaluating the sum

$$1 + \frac{1}{4} + \frac{1}{9} + \cdots = \sum_{k=1}^{\infty} \frac{1}{k^2} = \zeta(2). \tag{29}$$

The substitution $t = e^{-x}$ turns (28) into

$$\int_0^1 \frac{-\log(1-t)}{t}\, dt.$$

The last integrand has Taylor expansion

$$1 + \frac{t}{2} + \frac{t^2}{3} + \ldots,$$

and integrating termwise, we arrive at (29).

Theorem 4.1 (Euler).

$$\zeta(2) = \frac{\pi^2}{6}.$$

According to a widespread belief, Theorem 4.1 is a difficult (compared to single variable calculus) result whose proof requires Fourier series, residue calculus, or some equally sophisticated theory. There is a famous proof by Apostol that uses only real two-variable calculus. Apostol's proof is nice, but after having seen the following (much less well-known) proof by Yaglom and Yaglom, I no longer regard Apostol's proof as spectacularly simple.

The Yaglom-Yaglom proof relies on the evaluation of the sum

$$S_N = \frac{1}{\sin^2\left(\frac{\pi}{4N}\right)} + \frac{1}{\sin^2\left(\frac{3\pi}{4N}\right)} + \frac{1}{\sin^2\left(\frac{5\pi}{4N}\right)} + \cdots + \frac{1}{\sin^2\left(\frac{(2N-1)\pi}{4N}\right)}.$$

Since $\sin(\pi/4) = 1/\sqrt{2}$, $S_1 = 2$. For $N = 2$ we have

$$S_2 = \frac{1}{\sin^2\left(\frac{\pi}{8}\right)} + \frac{1}{\sin^2\left(\frac{3\pi}{8}\right)},$$

and this can be evaluated to $4S_1 = 8$ by putting $x = \pi/4$ in the identity

$$\frac{1}{\sin^2(x/2)} + \frac{1}{\sin^2(\pi/2 - x/2)} = \frac{4}{\sin^2 x}. \tag{30}$$

Similarly, by putting $x = \pi/8$ and $x = 3\pi/8$ in (30), we find that $S_4 = 4S_2 = 32$, and in general, $S_{2N} = 4S_N$. Inductively, it follows that if N is a power of 2, then

$$S_N = 2N^2. \tag{31}$$

Yaglom-Yaglom prove that (31) holds for all N, but to prove Euler's identity, we only need it for some infinite sequence of values of N. Since $\sin x \leq x$ for nonnegative x, we have

$$\frac{1}{\sin^2 x} \geq \frac{1}{x^2}, \tag{32}$$

and

$$2N^2 = S_N \geq \frac{(4N)^2}{\pi^2} + \frac{(4N)^2}{9\pi^2} + \frac{(4N)^2}{25\pi^2} + \cdots + \frac{(4N)^2}{(2N-1)^2\pi^2}. \tag{33}$$

After dividing by N^2, (33) simplifies to

$$1 + \frac{1}{9} + \frac{1}{25} + \cdots + \frac{1}{(2N-1)^2} \leq \frac{\pi^2}{8}.$$

To prove that this inequality is sharp in the $N \to \infty$ limit, it suffices to show that the error in (32) is bounded, so that the error in (33) is $O(N)$. The simplest way to do this is probably to note that by another standard inequality, $\tan x \geq x$ for $0 \leq x < \pi/2$. Therefore

$$\frac{1}{\sin^2 x} \geq \frac{1}{x^2} \geq \frac{\cos^2 x}{\sin^2 x},$$

and here t he difference between the left hand side and the right hand side is 1. Therefore we have

$$1 + \frac{1}{9} + \frac{1}{25} + \cdots = \frac{\pi^2}{8}.$$

To finish the proof, notice that the remaining terms are

$$\frac{1}{4} + \frac{1}{16} + \frac{1}{36} + \cdots = \frac{1}{4} \cdot \left(1 + \frac{1}{4} + \frac{1}{9} + \dots\right),$$

so that

$$\frac{3}{4}\zeta(2) = \frac{\pi^2}{8}.$$

Note: There is also a very nice proof of Euler's identity by Mikael Passare [83] which is based on evaluation of the integral (28).

5 Frieze's $\zeta(3)$ limit (7/10)

There are two famous theorems concerning the number

$$\zeta(3) = 1 + \frac{1}{8} + \frac{1}{27} + \cdots \approx 1.202.$$

In 1977, Apéry proved that $\zeta(3)$ is irrational. And in 1985, Alan Frieze showed that the expected cost of the minimum spanning tree on K_n, with independent uniform $[0,1]$ edge costs, converges to $\zeta(3)$ when $n \to \infty$.

Apéry was first, and therefore $\zeta(3)$ is sometimes called "Apéry's constant". If it wasn't for Apéry's theorem, we could very well have called it "Frieze's constant".

5.1 The Galton-Watson process

Given a random distribution on the nonnegative integers, we start with a root, and generate a tree by letting the number of children of each vertex be taken independently from the given distribution. The most fundamental question is the probability that the process becomes extinct.

Let $f(x) = a_0 + a_1x + a_2x^2 + \dots$ be the generating function for the number of children. Then the probability that the process is extinct immediately is a_0. The probability that it becomes extinct after at most one generation is $f(a_0)$. After at most two generations it is $f(f(a_0))$ etc.

By drawing a graph of f and the iterations of f starting from $a_0 = f(0)$, we see that there are essentially two possibilities: If $f'(1) < 1$, then the process will become extinct with probability 1, and this holds also when $f'(1) = 1$ except in the degenerate case $f(x) = x$. If $f'(1) > 1$, then there is a nontrivial solution to $x = f(x)$, and this x is the probability of extinction.

5.2 The Poisson Galton-Watson process

We are first of all interested in the case when the number of children is Poisson distributed with a parameter s. In this case

$$a_k = \frac{s^k}{k!}e^{-s},$$

and the generating function f is given by

$$f(x) = e^{-s}\sum \frac{s^k}{k!}x^k = e^{-s}e^{sx} = e^{-s(1-x)}.$$

The probability $p(s)$ of extinction is zero for $0 \le s \le 1$, and for $s > 1$, we have

$$p(s) = e^{-s(1-p(s))}.$$

Inverting, we get

$$s = \frac{-\log(p(s))}{1-p(s)}.$$

5.3 The $\zeta(3)$ calculation

The following calculation that leads to the number $\zeta(3)$ is based on Aldous-Steele [5].

We will argue that with edge costs rescaled to $\exp(1/n)$, so that the edge costs from a given node are approximately a $Po(1)$ process, the expected contribution to the MST from a single edge is given by

$$\int_0^\infty s \cdot (2p(s) - p^2(s))\, ds. \tag{34}$$

Here $2p - p^2$ is the probability that of two Poisson G-W processes, at least one becomes extinct. The idea is that n is large and we consider an edge e of cost s. This edge will participate in the MST iff there is no path connecting its endpoints and consisting only of edges of cost $< s$.

If we start from the two endpoints and simultaneously search through such paths, we will approximately be observing two Po(s) G-W processes, as long as the number of vertices observed in the process is small compared to n (it has to be $<< n^{1/2}$, but we return to the details).

If one of the processes quickly becomes extinct, then clearly e will belong to the MST. If none of the processes becomes extinct within a reasonable horizon, then probably there is a connection between them, and e is not in the MST.

By partial integration, and noting that $p'(s) = 0$ for $0 \le s \le 1$, (34) can be written as

$$\int_0^\infty s^2(p(s)-1)p'(s)\,ds = \int_0^1 s^2(1-p)\,dp = \int_0^1 \frac{\log^2 p}{1-p}\,dp$$
$$= \int_0^1 \big(1+p+p^2+\dots\big)\log^2 p\,dp = \sum_{k=0}^\infty \int_0^1 p^k \log^2 p\,dp. \quad (35)$$

By applying partial integration twice, we find that

$$\int_0^1 p^k \log^2 p\,dp = -\int_0^1 \frac{p^{k+1}}{k+1}\cdot 2\log p \cdot \frac{1}{p}\,dp = -\frac{2}{k+1}\int_0^1 p^k \log p\,dp$$
$$= \frac{2}{(k+1)^2}\int_0^1 p^k\,dp = \frac{2}{(k+1)^3}. \quad (36)$$

Hence (35) is equal to $2\zeta(3)$. Since there are $\sim n^2/2$ edges in the graph, the total cost of the MST is $n\zeta(3)$, and with $\exp(1)$ edge costs, it is $\zeta(3)$.

5.4 Rigorous argument

The idea of the rigorous argument is that we pick a set of edges that we know belong to the MST because they have reasonably small cost, and one of the search processes dies out quickly. We then show that the expected total cost of this set (suitably rescaled) is at least $\zeta(3) - o(1)$. This provides an asymptotic lower bound. For the upper bound, we show that the set we have chosen has cost at most $\zeta(3) + o(1)$, and contains $(1 - o(1))n$ edges. Then we use the (by now standard) argument of coloring and orienting edges, and patch the $o(n)$ components into a single component at expected extra cost $o(1)$.

5.5 The Poisson G-W process again

Let $p_k(s)$ be the probability that the $\mathrm{Po}(s)$ G-W process becomes extinct after generating exactly k individuals (not counting the root). There is a miraculous exact formula, based on Cayley's formula for the number of trees.

It is

$$p_k(s) = \frac{(k+1)^{k-1}}{k!} s^k e^{-(k+1)s}.$$

By differentiating and solving $p_k'(s) = 0$, one finds that the maximum of $p_k(s)$ occurs for $s = k/(k+1)$, and the value is

$$\frac{k^k e^{-k}}{(k+1)!} = O\left(\frac{1}{k^{3/2}}\right),$$

by Stirling's formula. Summing from $k+1$ to infinity, we see that

$$p(s) - p_k(s) = O\left(\frac{1}{k^{1/2}}\right),$$

where the implicit constant is uniform in s.

5.6 Redoing the calculation

In K_n, with $\exp(1/n)$ edge costs, consider the search process from an endpoint of an edge e of cost s. We look for the cheapest edges to $n-2$ or $n-1$ vertices: all vertices except the vertex itself and its mother (if it has one). The costs of these edges stochastically dominate the times of the first $n-1$ or $n-2$ events of a Po(1)-process, so since we want to lower-bound the extinction probability, the fact that the edge costs are not really from a Poisson process does not bother us. A minor problem though is the possibility of collisions. If we examine a vertex from one search process, and it turns out to belong to the set of vertices from the other search process, then e does not belong to the MST even if the processes quickly become extinct. If we run the two search processes simultaneously, and examine vertices one at a time until at most k vertices have been found in each process, then there are at most $2k$ vertices to examine, and in each step the pro bability of a collision with earlier vertices is at most

$$\frac{2k+2}{n}.$$

Therefore the expected number of collisions is $O(k^2/n)$, and so is the probability of a collision.

We conclude that the probability of an edge of cost s participating in the MST is at least

$$2p_k(s) - p_k^2(s) - O\left(\frac{k^2}{n}\right) = 2p(s) - p^2(s) - O\left(\frac{1}{\sqrt{k}}\right) - O\left(\frac{k^2}{n}\right)$$
$$= 2p(s) - p^2(s) - O\left(\frac{1}{n^{1/5}}\right), \quad (37)$$

if we put $k = n^{2/5} + O(1)$.

Therefore the expected contribution to the MST from a single edge is at least

$$\frac{1}{n}\int_0^{\log n} e^{-s/n} \cdot s \cdot (2p(s) - p^2(s) - O(n^{-1/5}))\, ds$$
$$\geq \frac{1}{n} \cdot \left(1 - \frac{\log n}{n}\right) \cdot \int_0^{\log n} s \cdot (2p(s) - p^2(s) - O(n^{-1/5}))\, ds$$
$$= \frac{1}{n} \cdot \left(1 - \frac{\log n}{n}\right) \cdot \left[\int_0^{\log n} s \cdot (2p(s) - p^2(s))\, ds - O\left(\frac{\log^2 n}{n^{1/5}}\right)\right]. \quad (38)$$

Next we use the fact that $p(s) = O(p_0(s)) = O(e^{-s})$, and that consequently, $2p(s) - p^2(s) = O(e^{-s})$. For large s, this says that if the process dies (which is unlikely), then it is probably because the root has no children. And if this holds for large s, then it obviously holds uniformly as well.

If we change the upper limit of integration from $\log n$ to ∞, we introduce another error of order $\int_{\log n}^{\infty} se^{-s}\, ds = O(\log n/n)$, which is smaller than the present error term $O(\log^2 n/n^{1/5})$. Therefore (38) is

$$\frac{1}{n} \cdot \left(1 - \frac{\log n}{n}\right) \cdot \left[\int_0^{\infty} s \cdot (2p(s) - p^2(s))\, ds - O\left(\frac{\log^2 n}{n^{1/5}}\right)\right]$$
$$= \frac{1}{n} \cdot \left(1 - \frac{\log n}{n}\right) \cdot \left[2\zeta(3) - O\left(\frac{\log^2 n}{n^{1/5}}\right)\right] = \frac{1}{n} \cdot \left[2\zeta(3) - O\left(\frac{\log^2 n}{n^{1/5}}\right)\right]. \quad (39)$$

This leads to the desired lower bound.

5.7 The upper bound

For the upper bound, we start with the set of edges above. The edge costs in the search process are stochastically dominated by a $Po((n-2)/n)$-process, and all other estimates work both ways. Therefore we have an edge set of expected cost $\zeta(3) + O(\log^2 n/n^{1/5})$ which contains no cycles. We also want to show that the expected number of edges in this set is $(1 - o(1))n$, so that it connects the vertices into $o(n)$ components. In fact we get the expected number of edges by an integral similar to the one above, but with the factor s in the integrand deleted. The result is that the probability that a given edge belongs to our set is

$$\frac{1}{n} \cdot \left[2 - O\left(\frac{\log n}{n^{1/5}} \right) \right].$$

Here we should verify that

$$\int_0^\infty \left(2p(s) - p^2(s) \right) \, ds = 2.$$

The expected number of components we need to patch together is therefore $O(n^{4/5} \log n)$. We use the standard method of introducing an infinite sequence of edges between every pair of vertices, coloring edges red with probability p, and giving them a random direction. Then we do the construction above using the blue edges, and obtain an edge set of expected total cost (and now we scale back to the original $\exp(1)$ setting)

$$\frac{\zeta(3)}{1-p} + O(\log^2 n/n^{1/5}).$$

Then as long as there is more than one component, we take a component and patch it to some other component using the cheapest red edge, and requiring it to have a certain direction in order to preserve independence. The worst case is when we have to patch a component of size 1 or $n-1$, when we only have $n-1$ edges available. In any case, each round of patching costs (in expectation) at most $1/(np)$, and there are (independently of the costs) $O(n^{4/5} \log n)$ rounds of patching. This gives a spanning tree of expected cost at most

$$\zeta(3) + O(p) + O\left(\frac{\log^2 n}{n^{1/5}} \right) + O\left(\frac{\log n}{pn^{1/5}} \right).$$

If we put $p = n^{-1/10}$, then this shows that

$$E(T_n) = \zeta(3) + O\left(\frac{\log n}{n^{1/10}} \right).$$

6 A proof of the Parisi formula (14/10)

The following is taken from my article *An easy proof of the $\zeta(2)$ limit in the random assignment problem.*

6.1 Introduction

We consider the following random model of the assignment problem: The edges of an m by n complete bipartite graph are assigned independent exponentially distributed costs. A *k-assignment* is a set of k edges of which no two have a vertex in common. The cost of an assignment is the sum of the costs of its edges. Equivalently, if the costs are represented by an m by n matrix, a k-assignment is a set of k matrix entries, no two in the same row or column. We let $C_{k,m,n}$ denote the minimum cost of a k-assignment. We are primarily interested in the case $k = m = n$, where we write $C_n = C_{n,n,n}$.

The distribution of C_n has been investigated for several decades. In 1979, D. Walkup [104] showed that $E(C_n)$ is bounded as $n \to \infty$, a result which was anticipated already in [24]. Further experimental results and improved bounds were obtained in [18, 22, 40, 46, 52, 53, 57, 58, 77, 79]. In a series of papers [65, 66, 68] from 1985–1987, Marc Mézard and Giorgio Parisi gave strong evidence for the conjecture that as $n \to \infty$,

$$E(C_n) \to \frac{\pi^2}{6}. \tag{40}$$

The first proof of (40) was found by David Aldous in 2000 [6, 7].

In 1998, Parisi conjectured [82] that

$$E(C_n) = 1 + \frac{1}{4} + \frac{1}{9} + \cdots + \frac{1}{n^2}. \tag{41}$$

This suggested a proof by induction on n. The hope of finding such a proof increased further when Don Coppersmith and Gregory Sorkin [22] extended the conjecture (41) to general k, m and n. They suggested that

$$E(C_{k,m,n}) = \sum_{\substack{i,j\geq 0\\ i+j<k}} \frac{1}{(m-i)(n-j)}, \tag{42}$$

and showed that this reduces to (41) in the case $k = m = n$. In order to establish (42) inductively it would suffice to prove that

$$E(C_{k,m,n}) - E(C_{k-1,m,n-1}) = \frac{1}{mn} + \frac{1}{(m-1)n} + \cdots + \frac{1}{(m-k+1)n}. \tag{43}$$

Further generalizations and verifications of special cases were given in [10, 20, 23, 25, 59]. Of particular interest is the paper [20] by Marshall Buck, Clara Chan and David Robbins. They considered a model where each vertex is given a nonnegative weight, and the cost of an edge is exponential with rate equal to the product of the weights of its endpoints. In the next section we consider a special case of this model.

The formulas (41) and (42) were proved in 2003 independently by Chandra Nair, Balaji Prabhakar and Mayank Sharma [75] and by Svante Linusson and the author [60]. These proofs are quite complicated, relying on the verification of more detailed induction hypotheses. Here we give a short proof of (43) based on some of the ideas of Buck, Chan and Robbins. Finally in Section 6.4 we give a simple proof that $\mathrm{var}(C_n) \to 0$, thereby establishing that $C_n \to \pi^2/6$ in probability.

6.2 Some results of Buck, Chan and Robbins

In this section we describe some results of the paper [20] by Buck, Chan and Robbins. We include proofs for completeness. Lemma 6.1 follows from Lemma 2 of [20]. For convenience we assume that the edge costs are *generic*, meaning that no two distinct assignments have the same cost. In the random model, this holds with probability 1. We say that a vertex *participates* in an assignment if there is an edge incident to it in the assignment. For $0 \leq r \leq k$, we let σ_r be the minimum cost r-assignment.

Lemma 6.1. *Suppose that $r < \min(m, n)$. Then every vertex that participates in σ_r also participates in σ_{r+1}.*

Proof. Let H be the symmetric difference $\sigma_r \triangle \sigma_{r+1}$ of σ_r and σ_{r+1}, in other words the set of edges that belong to one of them but not to the other. Since no vertex has degree more than 2, H consists of paths and cycles. We claim that H consists of a single path. If this would not be the case, then it would be possible to find a subset $H_1 \subseteq H$ consisting of one or two components of H (a cycle or two paths) such that H_1 contains equally many edges from σ_r and σ_{r+1}. By genericity, the edge sets $H_1 \cap \sigma_r$ and $H_1 \cap \sigma_{r+1}$ cannot have equal total cost. Therefore either $H_1 \triangle \sigma_r$ has smaller cost than σ_r, or $H_1 \triangle \sigma_{r+1}$ has smaller cost than σ_{r+1}, a contradiction. The fact that H is a path implies the statement of the lemma. $\square$

Here we consider a special case of the Buck-Chan-Robbins setting. We let the vertex sets be $A = \{a_1, \ldots, a_{m+1}\}$ and $B = \{b_1, \ldots, b_n\}$. The vertex

a_{m+1} is special: The edges from a_{m+1} are exponentially distributed of rate $\lambda > 0$, and all other edges are exponential of rate 1. This corresponds in the Buck-Chan-Robbins model to letting a_{m+1} have weight λ, and all other vertices have weight 1. The following lemma is a special case of Lemma 5 of [20], where the authors speculate that "This result may be the reason that simple formulas exist...". We believe that they were right.

Lemma 6.2. *Condition on the event that a_{m+1} does not participate in σ_r. Then the probability that it participates in σ_{r+1} is*

$$\frac{\lambda}{m-r+\lambda}. \tag{44}$$

Proof. Suppose without loss of generality that the vertices of A participating in σ_r are $a_1, \dots, a_r$. Now form a "contraction" K' of the original graph K by identifying the vertices $a_{r+1}, \dots, a_{m+1}$ to a vertex a'_{r+1} (so that in K' there are multiple edges from a'_{r+1}).

We condition on the cost of the minimum edge between each pair of vertices in K'. This can easily be visualized in the matrix setting. The matrix entries are divided into *blocks* consisting either of a single matrix entry $M_{i,j}$ for $i \leq r$, or of the set of matrix entries $M_{r+1,j}, \dots, M_{m+1,j}$, see Figure 1. We know the minimum cost of the edges within each block, but not the location of the edge having this minimum cost.

It follows from Lemma 6.1 that σ_{r+1} cannot contain two edges from $a_{r+1}, \dots, a_{m+1}$. Therefore σ_{r+1} is essentially determined by the minimum $(r+1)$-assignment σ'_{r+1} in K'. Once we know the edge from a'_{r+1} that belongs to σ'_{r+1}, we know that it corresponds to the unique edge from $\{a_{r+1}, \dots, a_{m+1}\}$ that belongs to σ_{r+1}. It follows from the "memorylessness" of the exponential distribution that the unique vertex of $a_{r+1}, \dots, a_{m+1}$ that participates in σ_{r+1} is distributed with probabilities proportional to the weights. This gives probability equal to (44) for the vertex a_{m+1}. $\square$

Corollary 6.3. The probability that a_{m+1} participates in σ_k is

$$1 - \frac{m}{m+\lambda} \cdot \frac{m-1}{m-1+\lambda} \cdots \frac{m-k+1}{m-k+1+\lambda} =$$
$$1 - \left(1 + \frac{\lambda}{m}\right)^{-1} \cdots \left(1 + \frac{\lambda}{m-k+1}\right)^{-1} =$$
$$\left(\frac{1}{m} + \frac{1}{m-1} + \cdots + \frac{1}{m-k+1}\right)\lambda + O(\lambda^2),$$

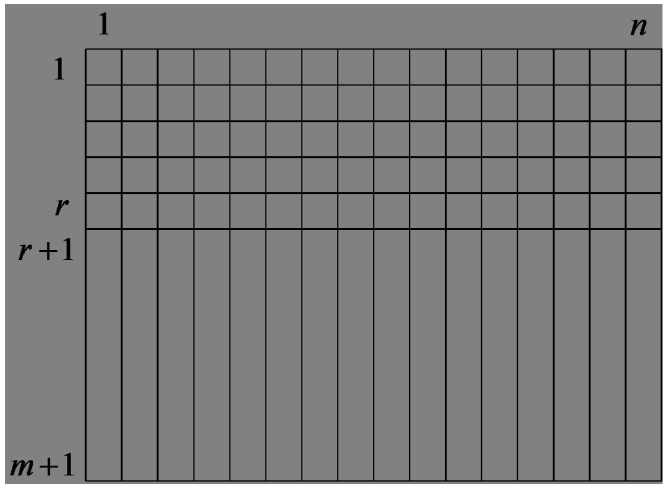

Figure 1: The matrix divided into blocks.

as $\lambda \to 0$.

Proof. This follows from Lemmas 6.1 and 6.2. $\square$

6.3 Proof of the Coppersmith-Sorkin formula

We show that the Coppersmith-Sorkin formula (42) can easily be deduced from Corollary 6.3. The reason that this was overlooked for several years is probably that it seems that by letting $\lambda \to 0$, we eliminate the extra vertex a_{m+1} and just get the original problem back.

We let X be the cost of the minimum k-assignment in the m by n graph $\{a_1, \dots, a_m\} \times \{b_1, \dots, b_n\}$ and let Y be the cost of the minimum $(k-1)$-assignment in the m by $n-1$ graph $\{a_1, \dots, a_m\} \times \{b_1, \dots, b_{n-1}\}$. Clearly X and Y are essentially the same as $C_{k,m,n}$ and $C_{k-1,m,n-1}$ respectively, but in this model, X and Y are also coupled in a specific way.

We let w denote the cost of the edge (a_{m+1}, b_n), and let I be the indicator variable for the event that the cost of the cheapest k-assignment that contains this edge is smaller than the cost of the cheapest k-assignment that does not use a_{m+1}. In other words, I is the indicator variable for the event that $Y + w < X$.

Lemma 6.4. *In the limit $\lambda \to 0$,*

$$E(I) = \left(\frac{1}{mn} + \frac{1}{(m-1)n} + \cdots + \frac{1}{(m-k+1)n}\right)\lambda + O(\lambda^2). \tag{45}$$

Proof. It follows from Corollary 6.3 that the probability that (a_{m+1}, b_n) participates in the minimum k-assignment is given by (45). If it does, then $w < X - Y$. Conversely, if $w < X - Y$ and no other edge from a_{m+1} has cost smaller than X, then (a_{m+1}, b_n) participates in the minimum k-assignment.

When $\lambda \to 0$, the probability that there are two distinct edges from a_{m+1} of cost smaller than X is of order $O(\lambda^2)$. $\square$

On the other hand, the fact that w is exponentially distributed of rate λ means that

$$E(I) = P(w < X - Y) = E\left(1 - e^{-\lambda(X-Y)}\right) = 1 - E\left(e^{-\lambda(X-Y)}\right).$$

Hence $E(I)$, regarded as a function of λ, is essentially the Laplace transform of $X - Y$. In particular $E(X - Y)$ is the derivative of $E(I)$ evaluated at $\lambda = 0$:

$$E(X-Y) = \frac{d}{d\lambda} E(I)\,|_{\lambda=0} = \frac{1}{mn} + \frac{1}{(m-1)n} + \cdots + \frac{1}{(m-k+1)n}.$$

This establishes (43) and thereby (42), (41) and (40).

6.4 A bound on the variance

The Parisi formula (41) shows that as $n \to \infty$, $E(C_n)$ converges to $\zeta(2) = \pi^2/6$. To establish $\zeta(2)$ as a "universal constant" for the assignment problem, it is also of interest to prove convergence in probability. This can be done by showing that $\mathrm{var}(C_n) \to 0$. The upper bound

$$\mathrm{var}(C_n) = O\left(\frac{(\log n)^4}{n(\log\log n)^2}\right)$$

was obtained by Michel Talagrand [96] in 1995 by an application of his isoperimetric inequality. In [110] it was shown that

$$\mathrm{var}(C_n) \sim \frac{4\zeta(2) - 4\zeta(3)}{n}. \tag{46}$$

These proofs are both quite complicated, and our purpose here is to present a relatively simple argument demonstrating that $\mathrm{var}(C_n) = O(1/n)$.

We first establish a simple correlation inequality which is closely related to the Harris inequality [43]. Let $X_1, \dots, X_N$ be random variables (not necessarily independent), and let f and g be two real valued functions of $X_1, \dots, X_N$. For $0 \le i \le N$, let $f_i = E(f|X_1, \dots, X_i)$, and similarly $g_i = E(g|X_1, \dots, X_i)$. In particular $f_0 = E(f)$, $f_N = f$, and similarly for g.

Lemma 6.5. *Suppose that for every i and every outcome of $X_1, \dots, X_N$,*

$$(f_{i+1} - f_i)(g_{i+1} - g_i) \ge 0. \tag{47}$$

Then f and g are positively correlated, in other words,

$$E(fg) \ge E(f)E(g). \tag{48}$$

Proof. Equation (47) can be written

$$f_{i+1}g_{i+1} \ge (f_{i+1} - f_i)g_i + (g_{i+1} - g_i)f_i + f_i g_i.$$

Notice that $f_{i+1} - f_i$, although not in general independent of g_i, has zero expectation conditioning on $X_1, \dots, X_i$ and thereby on g_i. It follows that $E\left((f_{i+1} - f_i)g_i\right) = 0$, and similarly for the second term. We conclude that $E(f_{i+1}g_{i+1}) \ge E(f_i g_i)$, and by induction that

$$E(fg) = E(f_N g_N) \ge E(f_0 g_0) = f_0 g_0 = E(f)E(g).$$

□

The random graph model that we use is the same as in the previous section, but we modify the concept of "assignment" by allowing an arbitrary number of edges from the special vertex a_{m+1} (but still at most one edge from each other vertex). This is not essential for the argument, but simplifies some details. Lemmas 6.1 and 6.2 as well as Corollary 6.3 are still valid in this setting.

We let C be the cost of the minimum k-assignment σ_k (with the modified definition), and we let J be the indicator variable for the event that a_{m+1} participates in σ_k.

Lemma 6.6.

$$E\left(C \cdot J\right) \le E\left(C\right) \cdot E\left(J\right). \tag{49}$$

Proof. Let $f = C$, and let $g = 1 - J$ be the indicator variable for the event that a_{m+1} does not participate in σ_k. As the notation indicates, we are going to design a random process $X_1, \ldots, X_N$ such that Lemma 6.5 applies. This process is governed by the edge costs, and $X_1, \ldots, X_N$ will give us successively more information about the edge costs, until σ_k and its cost are determined. A generic step of the process is similar to the situation in the proof of Lemma 6.2.

We let $M(r)$ be the matrix of "blocks" when σ_r is known, that is, the $r+1$ by n matrix of block minima as in Figure 1. Moreover we let $\theta_1, \ldots, \theta_k$ be vertices in A such that for $r \leq k$, σ_r uses the vertices $\theta_1, \ldots, \theta_r$.

When we apply Lemma 6.5, the sequence $X_1, \ldots, X_N$ is taken to be the sequence $M(0), \theta_1, M(1), \theta_2, M(2), \ldots, \theta_k$. Notice first that the cost f of the minimum k-assignment is determined by $M(k-1)$, and that of course $\theta_1, \ldots, \theta_k$ determines g, that is, whether or not a_{m+1} participates in σ_k.

In order to apply the lemma, we have to verify that each time we get a new piece of information in the sequence, the conditional expectations of f and g change in the same direction, if they change. By the argument in the proof of Lemma 6.2, we have

$$\begin{aligned} E\left(g|M(0), \theta_1, \ldots, \theta_r\right) &= E\left(g|M(0), \theta_1, \ldots, \theta_r, M(r)\right) \\ &= \begin{cases} 0, \quad \text{if } v_{n+1} \in \{\theta_1, \ldots, \theta_r\}, \\ \frac{m-r}{m-r+\lambda} \cdot \frac{m-r-1}{m-r-1+\lambda} \cdot \dots \cdot \frac{m-k+1}{m-k+1+\lambda}, \quad \text{if } v_{n+1} \notin \{\theta_1, \ldots, \theta_r\}. \end{cases} \end{aligned}$$

Therefore when we get to know another row in the matrix, the conditional expectation of g does not change, which means that for this case, the hypothesis of Lemma 6.5 holds. The other case to consider is when we already know $M(0), \ldots, M(r)$ and $\theta_1, \ldots, \theta_r$, and are being informed of θ_{r+1}. In this case the conditional expectations of f and g can obviously both change. For g, there are only two possibilities. Either $\theta_{r+1} = a_{m+1}$, which means that $g = 0$, or $\theta_{r+1} \neq a_{m+1}$, which implies that the conditional expectation of g increases.

To verify the hypothesis of Lemma 6.5, it clearly suffices to assume that $\{\theta_1, \ldots, \theta_r\} = \{1, \ldots, r\}$, and to show that if $\theta_{r+1} = a_{m+1}$, then the conditional expectation of f decreases. Since we know $M(r)$, we know to which "block" of the matrix the new edge belongs, that is, there is a j such that we know that exactly one of the edges in the set $E' = \{(a_i, b_j) : r+1 \leq i \leq m+1\}$ will belong to σ_{r+1}.

We now condition on the costs of all edges not in E'. Since we know $M(r)$, we also know the minimum edge cost, say α, in E'. We now make the very simple observation that if the minimum cost edge in E' is (a_{m+1}, b_j), then no other edge in E' can participate in σ_k, for in an assignment, any edge in E' can be replaced by (a_{m+1}, b_j). It follows that the value of f given that $M_{m+1,j} = \alpha$ is the same regardless of the costs of the other edges in E'. In particular it is the same as the value of f given that all edges in E' have cost α, which is certainly not greater than the conditional expecation of f given that some edge other than (a_{m+1}, b_j) has the minimum cost α in E'. It follows that Lemma 6.5 applies, and this completes the proof. □

The inequality (49) allows us to establish an upper bound on $\mathrm{var}(C_n)$ which is of the right order of magnitude (it is easy to see that $\mathrm{var}(C_n) \geq 1/n$, see [10]).

Theorem 6.7.

$$\mathrm{var}(C_n) < \frac{\pi^2}{3n}.$$

Proof. We let X, Y, I and w be as in Section 6.3, with I being the indicator variable of the event $Y + w < X$. Again C denotes the cost of σ_k, and J is the indicator variable for the event that a_{m+1} participates in σ_k.

Obviously

$$E\left(C\right) \leq E\left(X\right). \tag{50}$$

Again the probability that there are two distinct edges from a_{m+1} of cost smaller than X is of order $O(\lambda^2)$. Therefore

$$E\left(J\right) = nE\left(I\right) + O(\lambda^2) = n\lambda E\left(X - Y\right) + O(\lambda^2). \tag{51}$$

Similarly

$$E\left(C \cdot J\right) = nE\left(I \cdot (Y + w)\right) + O(\lambda^2). \tag{52}$$

If we condition on X and Y, then

$$\begin{aligned} E\left(I \cdot (Y + w)\right) &= \int_0^{X-Y} \lambda e^{-\lambda t}(Y + t)\, dt \\ &= Y(X - Y)\lambda + \frac{(X - Y)^2}{2}\lambda + O(\lambda^2) = \frac{1}{2}\left(X^2 - Y^2\right)\lambda + O(\lambda^2). \end{aligned} \tag{53}$$

If, in the inequality (49), we substitute the results of (50), (51), (52) and (53), then after dividing by $n\lambda$ we obtain

$$\frac{1}{2}E\left(X^2 - Y^2\right) \leq E\left(X\right)^2 - E\left(X\right)E\left(Y\right) + O(\lambda).$$

After deleting the error term, this can be rearranged as

$$\mathrm{var}(X) - \mathrm{var}(Y) \leq \left(E\left(X\right) - E\left(Y\right)\right)^2.$$

But we already know that $E\left(X\right) - E\left(Y\right)$ is given by (43). Therefore it follows inductively that

$$\mathrm{var}(C_n) \leq \sum_{i=1}^{n} \frac{1}{i^2}\left(\frac{1}{n} + \cdots + \frac{1}{n-i+1}\right)^2$$
$$\leq \sum_{i=1}^{n} \frac{1}{i^2}\left(\log(n+1/2) - \log(n+1/2-i)\right)^2.$$

If we replace the sum over i by an integral with respect to a continuous variable, then the integrand is convex. Hence

$$\mathrm{var}(C_n) \leq \int_0^{n+1/2} \frac{\left(\log(n+1/2) - \log(n+1/2-x)\right)^2}{x^2}\,dx$$
$$= \frac{1}{n+1/2}\int_0^1 \frac{\log(1-x)^2}{x^2}\,dx = \frac{2\zeta(2)}{n+1/2} < \frac{\pi^2}{3n}.$$

□

7 Concentration and universality (28/10)

In earlier lectures, there have been questions about whether the costs converge in probability to their limit expectations, and whether these results hold for other distributions of the edge cost. Several questions of this type, relating to the degree of universality of the limiting value, boil down to showing that with high probability no extremely expensive edge participates in the solution. In order to establish these universality results, we have to be a bit careful about exactly what we mean by "high probability" and "extremely expensive". In this section we show how some results can be derived by elementary methods, and in the next section, how a famous inequality of Talagrand can be applied to harder problems.

We showed in in the previous section (though I didn't talk about this in the lecture) that the variance of the cost in the n by n exp(1) assignment problem tends to zero as $n \to \infty$. There are also other ways to prove this:

- The exact formula

$$\operatorname{var}(A_n) = 5 \cdot \sum_{k=1}^{n} \frac{1}{k^4} - 2\left(\sum_{k=1}^{n} \frac{1}{k^2}\right)^2 - \frac{4}{n+1}\sum_{k=1}^{n} \frac{1}{k^3}. \tag{54}$$

- The Talagrand inequality.

Let A_n' denote the cost of the minimum assignment under uniform [0,1] edge costs.

Proposition 7.1. *When $n \to \infty$,*

$$A_n' \xrightarrow{\mathrm{p}} \zeta(2).$$

Proof. The coupling $u \to -\log(1-u)$ immediately shows that A_n' is stochastically dominated by A_n. Since A_n converges in probability to $\zeta(2)$, all that remains to prove is that with high probability, A_n' is not very much smaller than A_n.

Since

$$\frac{-\log(1-u)}{u} \to 1,$$

when $u \to 0$, it suffices to show that with high probability, all edges in the minimum assignment have costs close to zero. Since the optimal assignments are not necessarily the same under exponential and uniform edge costs, we have to start from the uniform assignment, and show that whp all edges are cheap. If this is the case, then the coupling to exponential edge costs will give one assignment of cost not much greater than A_n', say $A_n' + \epsilon$. Now the inequality goes the right way, and A_n is squeezed:

$$A_n' \le A_n \le A_n' + \epsilon.$$

So how do we prove that whp all edges in the optimum solution are cheap? Again there are several possibilities. Working from scratch, the most natural approach is to use expander properties of random graphs, in the spirit of the proof of Walkup's theorem in Section 1. We come back to these methods in the section on Talagrand's inequality. But there is also a simple way based on the explicit formula (Coppersmith-Sorkin formula) (42) for incomplete matching.

The difference in cost of the minimum n-matching and the minimum $(n-1)$-matching is obviously an upper bound on the cost of the most expensive edge in the n-matching (since removing this edge gives an $(n-1)$-matching which is competing for the minimum).

It follows from (42) that

$$E(C_{n,n,n} - C_{n-1,n,n}) = \frac{2}{n+1} \cdot \left(1 + \frac{1}{2} + \cdots + \frac{1}{n}\right) = O\left(\frac{\log n}{n}\right).$$

To complete the proof, we need to carry this over to the uniform setting. Unfortunately, if we remove just the single most expensive edge from the optimal matching in the uniform case, then we cannot immediately conclude that the remaining matching transforms into a matching of cost considerably smaller than $\zeta(2)$, through the coupling, since the optimal uniform matching may transform into something which is not nearly optimal in the exponential setting.

The trick to overcome this difficulty is to decide on a threshold δ, and remove *all* edges of cost above δ from the optimal uniform matching. Trivially there cannot be essentially more than $\zeta(2)/\delta$ of them, so for large n what remains is an almost complete matching. Therefore the cost of this matching in the exponential setting is with high probability not much smaller than $\zeta(2)$. On the other hand, it is at most

$$A'_n \cdot \frac{-\log(1-\delta)}{\delta}.$$

This shows that A'_n is with high probability not much smaller than $\zeta(2)$, which completes the proof. $\square$

In fact the statement can be proved under the weaker assumption that the distribution of the edge costs satisfy

$$\frac{P(X < t)}{t} \to 1,$$

as $t \to 0+$. For the proof of this, we refer to Section 11 of [115].

8 Flow problems, the friendly model, and the Buck-Chan-Robbins urn process (4/11)

We will discuss the content of the first part of [115], emphasizing definitions and examples. The main theorem (expected cost of optimization problem = area of region defined by the urn process) is proved by introducing an extra vertex of degree tending to zero. I will not give the details of this proof in the lecture.

9 The random matching problem on K_n

(18/11). The following is taken from [112].

The edges of the complete graph on n vertices $v_1, \dots, v_n$ are assigned independent costs from exponential distribution with rate 1. A *k-matching* is a set of k edges of which no two have a vertex in common. We let $C_{k,n}$ denote the minimum total cost of the edges of a k-matching. In 1985 Marc Mézard and Giorgio Parisi [65, 68] gave convincing evidence that as $n \to \infty$,

$$E\left(C_{\lfloor n/2 \rfloor, n}\right) \to \frac{\pi^2}{12}. \tag{55}$$

This was proved in 2001 by David Aldous [6, 7]. He considered the related *assignment problem* on the complete bipartite graph, which is technically simpler. It is known that (55) follows from the results of [7] by a slight modification of Proposition 2 of [6].

We give a simple proof of (55) by establishing explicit upper and lower bounds on $E\left(C_{k,n}\right)$ valid for arbitrary k and n. For *perfect* matchings, that is, when n is even and $k = n/2$, we prove that

$$\frac{\pi^2}{12} < E\left(C_{n/2,n}\right) < \frac{\pi^2}{12} + \frac{\log n}{n}. \tag{56}$$

Notice that the difference between the upper and lower bounds in (56) is much smaller than the random fluctuations of $C_{n/2,n}$. It is not hard to show that the standard deviation of $C_{n/2,n}$ is at least of order $n^{-1/2}$.

9.1 The extended graph

In the *extended graph* there is an extra vertex v_{n+1}, and the costs of the edges from this vertex are exponentially distributed with rate $\lambda > 0$. In the end,

λ will tend to zero. We say that a vertex v *participates* in a matching if the matching contains an edge incident to v.

In the following, we shall assume that the edge costs are such that no two distinct matchings have the same cost (this holds with probability 1), and we let σ_k be the minimum cost k-matching in the extended graph. We let $P_k(n)$ denote the normalized probability that v_{n+1} participates in σ_k. More precisely,

$$P_k(n) = \lim_{\lambda \to 0} \frac{1}{\lambda} P(v_{n+1} \text{ participates in } \sigma_k).$$

Lemma 9.1.

$$E(C_{k,n}) - E(C_{k-1,n-1}) = \frac{1}{n} P_k(n), \tag{57}$$

and consequently

$$E(C_{k,n}) = \frac{1}{n} P_k(n) + \frac{1}{n-1} P_{k-1}(n-1) + \cdots + \frac{1}{n-k+1} P_1(n-k+1). \tag{58}$$

Proof. The right hand side of (57) is the normalized probability that a particular edge from v_{n+1}, say the edge $e = (v_n, v_{n+1})$, belongs to σ_k. Naturally, $C_{k,n}$ denotes the cost of the minimum k-matching on the vertices $v_1, \dots, v_n$. We couple $C_{k,n}$ and $C_{k-1,n-1}$ by letting $C_{k-1,n-1}$ be the cost of the minimum $(k-1)$-matching on the vertices $v_1, \dots, v_{n-1}$.

Let w be the cost of the edge e. If e participates in σ_k, then we must have $w \le C_{k,n} - C_{k-1,n-1}$. Conversely, if $w \le C_{k,n} - C_{k-1,n-1}$ then e will participate in σ_k unless there is some other edge from v_{n+1} that does. This can happen only if both e and some other edge from v_{n+1} have costs smaller than $C_{k,n}$. As $\lambda \to 0$, the probability for this is $O(\lambda^2)$. Hence

$$\begin{aligned}\frac{1}{n} P_k(n) &= \lim_{\lambda \to 0} \frac{1}{\lambda} P(w \le C_{k,n} - C_{k-1,n-1}) \\ &= \lim_{\lambda \to 0} \frac{1}{\lambda} E\left(1 - e^{-\lambda(C_{k,n} - C_{k-1,n-1})}\right) = E(C_{k,n}) - E(C_{k-1,n-1}).\end{aligned} \tag{59}$$

□

We therefore wish to estimate $P_k(n)$ for general k and n. For this purpose we design a random process driven by the edge costs. A convenient way to think about this process is to imagine that there is an oracle who knows all the edge costs. We ask questions to the oracle in such a way that we can control the conditional distribution of the edge costs while at the same time being able to determine whether v_{n+1} participates in σ_k or not. The following lemma is well-known in matching theory, but for completeness we include a proof.

Lemma 9.2. *Every vertex that participates in σ_r also participates in σ_{r+1}.*

Proof. Let H be the symmetric difference $\sigma_r \triangle \sigma_{r+1}$ of σ_r and σ_{r+1}, in other words the set of edges that belong to one of them but not to the other. Since no vertex has degree more than 2, H consists of paths and cycles. We claim that H consists of a single path. If this would not be the case, then it would be possible to find a subset $H_1 \subseteq H$ consisting of one or two components of H, such that H_1 contains equally many edges from σ_r and σ_{r+1}. By assumption, the edge sets $H_1 \cap \sigma_r$ and $H_1 \cap \sigma_{r+1}$ do not have equal total cost. Therefore either $H_1 \triangle \sigma_r$ has smaller cost than σ_r, or $H_1 \triangle \sigma_{r+1}$ has smaller cost than σ_{r+1}, a contradiction.

The fact that H is a path clearly implies the statement of the lemma. □

9.2 The lower bound

The lower bound on $E\left(C_{k,n}\right)$ is the simpler one and we establish it first. In Section 9.3, a modification of the method will yield a fairly good upper bound as well.

9.2.1 The process

The following protocol for asking questions to the oracle looks like an algorithm for finding the minimum k-matching, but what we are interested in is the probability that v_{n+1} participates in the minimum matching.

At each stage of the process, we say that a certain set of vertices are *exposed*, and the remaining vertices are *unexposed*. We have the following information:

1. We know the costs of all edges between exposed vertices.
2. For each exposed vertex v, we also know the minimum cost of the edges connecting v to the unexposed vertices.
3. Finally, we know the minimum cost of all edges connecting two unexposed vertices.

Another way to put this is to say that for every set A of at most two exposed vertices, we know the minimum cost of the edges whose set of exposed endpoints is precisely A. By well-known properties of independent exponential variables, the minimum is located with probabilities proportional to the rates of the corresponding exponential variables, and conditioning on a certain edge not being the one of minimum cost, its cost is distributed like the minimum plus another exponential variable of the same rate.

We also keep track of a nonnegative integer r which is such that σ_r contains only edges between exposed vertices. Moreover, we shall require that it can be verified from the information at hand that this matching is indeed the minimum r-matching. This is the reason why possibly some more vertices have to be exposed.

Initially, $r = 0$ and no vertex is exposed. At each stage of the process, the following happens:

- We compute a proposed minimum $(r+1)$-matching under the assumption that for all exposed vertices, their minimum cost edges to an unexposed vertex go to different unexposed vertices.

  By Lemma 9.2, σ_{r+1} will use at most two unexposed vertices. Hence either it contains the minimum cost edge connecting two unexposed vertices, or at most two of the minimum cost edges connecting an exposed vertex to an unexposed one.

- If the proposed minimum $(r+1)$-matching contains the minimum edge connecting two unexposed vertices, then it must indeed be the minimum $(r+1)$-matching. The two endpoints of the new edge are exposed (that is, we ask for the information required for them to be exposed). Finally, the value of r increased by 1.

- Otherwise the proposed matching includes up to two edges from exposed vertices to unexposed ones. Then the unexposed endpoints of these edges are revealed and exposed. Unless there are two such edges and they happen to have the same endpoint, the proposed matching is indeed the minimum $(r+1)$-matching, and the value of r is increased. If there are two edges to unexposed vertices and it turns out that they "collide", that is, they have the same unexposed endpoint, then the proposed matching is not valid. We have then exposed only one more vertex, and we complete the round of the process without updating the value of r.

9.2.2 A lower bound on $P_k(n)$

We wish to estimate the probability that v_{n+1} participates in σ_k. Suppose that at a given stage of the process there are m ordinary unexposed vertices (that is, not counting v_{n+1}). There are two cases to consider.

Suppose first that an edge between two unexposed vertices is going to be revealed. The total rate of the edges between unexposed vertices is

$$\binom{m}{2} + O(\lambda)$$

and the total rate of the edges from v_{n+1} to the other unexposed vertices is λm. Hence the probability that v_{n+1} is among the two new vertices to be exposed is

$$\frac{\lambda m}{\binom{m}{2} + O(\lambda)} = \frac{2\lambda}{m-1} + O(\lambda^2).$$

For convenience we suppose that the vertices are revealed one at a time, with a coin toss to decide which vertex to be revealed first. Then the probability that v_{n+1} is exposed is

$$\frac{\lambda}{m-1} + O(\lambda^2)$$

for both the first and the second vertex.

Secondly, suppose that the unexposed endpoint of an edge from an exposed vertex is going to be revealed. If at one stage of the process there are two such edges, then again we reveal the endpoints one at a time, flipping a coin to decide the order. In case there is a collision, this will be apparent when the first new vertex is exposed. If there are m ordinary remaining unexposed vertices, then the total rate of the edges from a particular exposed vertex v to them is $m + \lambda$, and consequently the probability that v_{n+1} is exposed is

$$\frac{\lambda}{m} + O(\lambda^2).$$

This will hold also for the second edge of two to be exposed at one stage of the process, provided m denotes the number of remaining unexposed vertices at that point.

If v_{n+1} is among the first $2k$ vertices to be exposed, then it will participate in σ_k. We have neglected the possibility that there is a collision at v_{n+1}, since this is an event of probability $O(\lambda^2)$. When m ordinary unexposed vertices remain, the probability that v_{n+1} is the next vertex to be exposed is at least λ/m. Hence

$$P_k(n) \geq \frac{1}{n}\left(\frac{1}{n} + \frac{1}{n-1} + \cdots + \frac{1}{n-2k+1}\right).$$

We can improve slightly on this inequality by noting that the normalized probability that v_{n+1} is one of the two first vertices to be exposed is exactly

$$\frac{n}{\binom{n}{2}} = \frac{2}{n-1}.$$

Taking this into account, we get

$$P_k(n) \geq \frac{1}{n}\left(\frac{2}{n-1} + \frac{1}{n-2} + \cdots + \frac{1}{n-2k+1}\right). \tag{60}$$

9.2.3 A lower bound on $E\left(C_{k,n}\right)$

From (58) and (60) it follows that

$$\begin{aligned} E\left(C_{k,n}\right) \geq\ & \frac{1}{n}\left(\frac{2}{n-1} + \frac{1}{n-2} + \cdots + \frac{1}{n-2k+1}\right) \\ & + \frac{1}{n-1}\left(\frac{2}{n-2} + \frac{1}{n-3} + \cdots + \frac{1}{n-2k+2}\right) \\ & \vdots \\ & + \frac{1}{n-k+1}\left(\frac{2}{n-k}\right). \end{aligned} \tag{61}$$

It is straightforward to prove by induction on n that when n is even and $k = n/2$, (61) becomes

$$E\left(C_{n/2,n}\right) \geq \frac{1}{2}\left(1 + \frac{1}{4} + \frac{1}{9} + \cdots + \frac{1}{(n/2)^2}\right) + \frac{1}{n}.$$

By a simple integral estimate,

$$E\left(C_{n/2,n}\right) \geq \frac{1}{2}\left(\frac{\pi^2}{6} - \int_{n/2}^{\infty} \frac{dx}{x^2}\right) + \frac{1}{n} = \frac{\pi^2}{12}. \tag{62}$$

9.3 The upper bound

By modifying the argument given in the previous section, we can also establish an upper bound on $E\left(C_{k,n}\right)$. For this purpose, we are going to design the process differently.

9.3.1 The process

We modify the process described in Section 9.2. This time, only the vertices that participate in σ_r will be exposed. At each stage, we have the following information:

1. The costs of all edges between exposed vertices. In particular, σ_r is known.
2. For each exposed vertex v, we know the minimum cost of the edges from v to unexposed vertices.
3. For some exposed vertices, we may also know to which vertex this minimum cost edge goes, and the cost of the second cheapest edge to an unexposed vertex.
4. We also know the minimum cost of the edges connecting two unexposed vertices.

As in the previous section, we assume that the information we have is sufficient to verify that the given r-matching is indeed of minimum cost. We also assume that for the exposed vertices for which the minimum cost edge to an unexposed vertex is known, this edge never goes to v_{n+1}. As will become clear below, this assumption is justified by the fact that such an edge is revealed only in case of a collision. The event of a collision at v_{n+1} has probability $O(\lambda^2)$, which is negligible since we are estimating a probability of order λ.

We of course assume that $2r + 2 \leq n$, so that there are at least two ordinary unexposed vertices. Given the information we have, we compute a proposed minimum cost $(r + 1)$-matching under the assumption that no collision takes place. Then we ask whether or not the proposed matching is valid. If it is invalid, that is, if there is a collision, then this must be between the minimum cost edges to unexposed vertices from two of the exposed vertices. This endpoint is revealed, and we can assume that it is not v_{n+1}, since the probability for this is negligible. We repeat this until we find that there is no collision, and that therefore the proposed matching is valid. It must then be the minimum $(r+1)$-matching.

9.3.2 An upper bound on $P_k(n)$

We analyze a particular stage of the process and we wish to obtain an upper bound on the probability that v_{n+1} is one of the two new vertices that are used in σ_{r+1}. We let $m = n - 2r$ be the number of ordinary unexposed vertices.

If σ_{r+1} has an edge between two unexposed vertices, then by the analysis of Section 9.2, the probability that it uses the vertex v_{n+1} is

$$\frac{\lambda m}{\binom{m}{2}} = \frac{2\lambda}{m-1},$$

neglecting a term of order λ^2.

Suppose instead that σ_{r+1} contains two edges from exposed vertices v_i and v_j to unexposed vertices. First assume that we do not know the minimum cost edge to an unexposed vertex for any of them (that is, none of them falls under (3) above). Then the probability that v_{n+1} participates in σ_{r+1} is

$$\frac{\lambda m + \lambda m}{m(m-1)} = \frac{2\lambda}{m-1},$$

since we are conditioning on no collision occurring.

If on the other hand for at least one of v_i and v_j the minimum cost edge to an unexposed vertex is known according to (3), then the probability that v_{n+1} participates in σ_{r+1} is even smaller, at most $\lambda/(m-1)$.

This gives the following upper bound on the probability that v_{n+1} participates in σ_k:

$$P_k(n) \leq \frac{2}{n-1} + \frac{2}{n-3} + \cdots + \frac{2}{n-2k+1}.$$

9.3.3 An upper bound on $E(C_{k,n})$

Inductively we obtain the following upper bound on the expected cost of σ_k:

$$
\begin{aligned}
E\left(C_{k,n}\right) \leq \frac{1}{n}\left(\frac{2}{n-1}+\frac{2}{n-3}+\cdots+\frac{2}{n-2k+1}\right) \\
+\frac{1}{n-1}\left(\frac{2}{n-2}+\cdots+\frac{2}{n-2k+2}\right) \\
\vdots \\
+\frac{1}{n-k+1}\left(\frac{2}{n-k}\right).
\end{aligned}
$$

We give a slightly weaker but simpler upper bound, valid for even n and $k = n/2$, in order to establish (56). If we replace all the terms of the form $2/(n-i)$ except the first one in each pair of parentheses by $1/(n-i)+1/(n-i-1)$, we obtain

$$
\begin{aligned}
E\left(C_{n/2,n}\right) &\leq \frac{1}{n}+\frac{1}{2(n-1)}+\cdots+\frac{1}{(n/2)(n/2+1)} \\
&\quad+\frac{1}{n}\left(1+\frac{1}{2}+\cdots+\frac{1}{n-1}\right)+\frac{1}{n-1}\left(\frac{1}{2}+\cdots+\frac{1}{n-2}\right)+ \\
&\quad\vdots \\
&\quad+\frac{1}{(n/2)((n/2)+1)} \\
&=\frac{1}{2}\left(1+\frac{1}{4}+\cdots+\frac{1}{(n/2)^2}\right)+\frac{1}{n+1}\left(1+\frac{1}{2}+\cdots+\frac{1}{n}\right) \\
&\leq \frac{1}{2}\left(\frac{\pi^2}{6}-\int_{n/2+1}^{\infty}\frac{dx}{x^2}\right)+\frac{1+\log n}{n+1}=\frac{\pi^2}{12}-\frac{1}{n+2}+\frac{1}{n+1}+\frac{\log n}{n+1} \\
&=\frac{\pi^2}{12}+\frac{\log n}{n}-\frac{\log n-\frac{n}{n+2}}{n(n+1)}\leq\frac{\pi^2}{12}+\frac{\log n}{n}, \qquad (63)
\end{aligned}
$$

since $\log n \geq n/(n+2)$ for $n \geq 2$. Together with (62), this establishes (56).

9.4 Concluding remarks

We have inductively established lower and upper bounds on $C_{k,n}$ from lower and upper bounds on the normalized probability $P_{k,n}$ that an extra vertex participates in the minimum matching. For the corresponding problem on the complete m by n bipartite graph (the so called *assignment problem*), the expected cost of the minimum k-matching is known [60, 75], and is given by the simple formula (42), conjectured in [22] as a generalization of the formula (4) suggested in [82] for the special case $k = m = n$. For the complete graph, no such formula has even been conjectured.

We can now see why finding the exact value is harder for the complete graph. The probability that the extra vertex is included when we pass from σ_r to σ_{r+1} depends on whether σ_{r+1} is obtained by adding an edge between two vertices that do not participate in σ_r, or by an alternating path that replaces edges in σ_r by other edges. For the bipartite graph, the probability that an extra vertex is included is the same in the two cases, and thereby known. A proof of (42) based on this method is given in [111].

10 The Talagrand inequality (11/11)

In the lecture, we discussed the proof of Talagrand's inequality given in Alon and Spencer [11].

This is a proof of the inequality

$$P(A)(1 - P(A_t)) \leq e^{-t^2/4}.$$

Here A_t is defined in terms of Talagrand's distance.

11 Application of the Talagrand inequality

We show how to apply the Talagrand inequality to the sort of minimization problems that we consider.

12 Stirling's formula (not part of any lecture)

The famous formula of Stirling is often useful for estimating factorials and binomial coefficients. We give a simple and self-contained proof.

We wish to estimate

$$\log n! = \sum_{k=1}^{n} \log k.$$

By a simple integral estimate we see that

$$\int_1^n \log x \, dx \le \log n! \le \int_1^{n+1} \log x \, dx.$$

This is often sufficient, but it is interesting to obtain a better estimate. Consider the graph $y = \log x$, and draw line segments between points corresponding to $x = k$ and $x = k + 1$ for every positive integer k. Compare this with the set of rectangles of base $[k, k + 1]$ and height $[0, \log(k + 1)]$. Since the log-function is concave, we have

$$\log n! \le \int_1^n \log x \, dx + \frac{1}{2} \log n.$$

For large n, the error in this estimate will converge to the total area between the line segments and the graph of $\log x$. Let δ_k be the area of the region between the line segment from $x = k$ to $x = k + 1$, and the graph of $\log x$. It is easy to see that

$$\delta = \sum_{k=1}^{\infty} \delta_k$$

is finite. Actually we can take the regions of areas δ_k for $k \ge n$, and translate them so that their leftmost corners coincide. Then they will not overlap, and since the derivative of $\log x$ at $x = n$ is $1/n$, they will all fit into a triangle of base 1 and height $1/n$. Hence

$$\sum_{k=n}^{\infty} \delta_k \le \frac{1}{2n}.$$

Therefore we get the following estimate of $\log n!$:

$$\int_1^n \log x \, dx + \frac{1}{2} \log n - \delta \le \log n! \le \int_1^n \log x \, dx + \frac{1}{2} \log n - \delta + \frac{1}{2n}.$$

Since

$$\int_1^n \log x \, dx = n \log n - n + 1,$$

this shows that

$$\log n! = \left(n + \frac{1}{2}\right) \log n - n + (1 - \delta) + \frac{\theta}{2n},$$

where $0 \le \theta \le 1$. By exponentiating, we can also write this as

$$n! = n^{n+1/2} \cdot e^{-n} \cdot C \cdot \left(1 + \frac{\theta'}{n}\right),$$

where θ' too is between 0 and 1.

It remains to determine the constant C, and this can be done with the Wallis product formula.

Let a_n be the nth coefficient of the Taylor expansion of

$$\frac{1}{\sqrt{1-x}} = \left(1 + x + x^2 + \dots\right)^{1/2} = 1 + \frac{1}{2}x + \frac{3}{8}x^2 + \frac{5}{16}x^3 + \frac{35}{128}x^4 + \dots.$$

It is easily verified that

$$a_n = (-1)^n \binom{-1/2}{n} = \frac{1 \cdot 3 \cdot 5 \cdot \dots \cdot (2n-1)}{2 \cdot 4 \cdot 6 \cdot \dots \cdot (2n)}.$$

Now let

$$W = \frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdots.$$

This is the Wallis product, and it is well-known that $W = \pi/2$. For a simple proof, see [114].

From the Wallis product, it follows that

$$\begin{aligned} a_n^2 &= \frac{1 \cdot 1 \cdot 3 \cdot 3 \cdot \dots \cdot (2n-1)(2n-1)}{2 \cdot 2 \cdot 4 \cdot 4 \cdot \dots \cdot (2n)(2n)} \\ &= \frac{1}{2n} \cdot \frac{1 \cdot 3 \cdot 3 \cdot \dots \cdot (2n-1)(2n-1)}{2 \cdot 2 \cdot 4 \cdot 4 \cdot \dots \cdot (2n)} \sim \frac{1}{2n} \cdot \frac{1}{W} = \frac{1}{n\pi}. \end{aligned} \tag{64}$$

At the same time we have

$$\begin{aligned} a_n &= \frac{1}{2^{2n}} \cdot \binom{2n}{n} = \frac{1}{2^{2n}} \cdot \frac{(2n)!}{(n!)^2} \\ &= \frac{1}{2^{2n}} \cdot \frac{(2n)^{2n} \cdot \sqrt{2n} \cdot e^{-2n} \cdot C \cdot (1 + \theta_1/2n)}{n^{2n} \cdot n \cdot e^{-2n} \cdot C^2 \cdot (1 + \theta_2/n)^2} \\ &= \frac{\sqrt{2}}{\sqrt{n} \cdot C} \cdot (1 + O(1/n)). \end{aligned} \tag{65}$$

From (64) and (65) it follows that

$$C = \sqrt{2\pi}.$$

We can therefore state Stirling's formula as

$$n! = \left(\frac{n}{e}\right)^n \cdot \sqrt{2\pi n} \cdot \left(1 + \frac{\theta}{n}\right),$$

where $0 \leq \theta \leq 1$.

13 Problems and exercises

1. If $x_1, \ldots, x_n$ are independent uniform$[0,1]$, then what is the expectation of $\min(x_1, \ldots, x_n)$? What if $x_1, \ldots, x_n$ are exponential? What about the higher moments?

2. How do you generate an exponentially distributed (pseudo-) random variable from a uniform$[0,1]$ variable? (Actually John von Neumann has published a paper on this, but his solution, to add uniform [0,1] variables until you get one which is larger than the previous one, is not needed with today's computers.)

3. Prove that von Neumann's method is correct, and prove that the expected number of uniform [0,1]'s that you need to generate is e.

4. Let X_n be a sequence of nonnegative real random variables. Consider the statements

   (i) EX_n is bounded.

   (ii) There is a constant C such that when $n \to \infty$, $P(X_n \leq C) \to 1$.

   Show by example that none of the statements implies the other.

5. (a) What is the probability that a uniformly chosen permutation of n elements is cyclic?

   (b) What is the probability that there is a cycle of length exactly, say, $n - 7$? (Hint: What is the probability that the first element belongs to such a cycle?).

(c) What is the limit probability that there is a cycle of length at least $n/2$?

6. If Z_n is the number of cycles of a uniform random permutation of n elements, then what is $E(3^{Z_n})$?

7. (a) Show that if Z is the sum of two independent exp(1) variables, then $P(Z > t) = (1+t)e^{-t}$. Notice that $P(Z > t)$ is the same thing as the probability that a Po(1) process has at most one event in the interval $[0, t]$.

   (b) Consider the minimum $\alpha_m = \min(Z_1, \dots, Z_m)$ of m independent variables of the same distribution as Z. Show that

$$P(\alpha_m > t) = ((1+t)e^{-t})^m.$$

   (c) Show that

$$E\alpha_m = \int_0^\infty ((1+t)e^{-t})^m \, dt.$$

   (d) Show that the inequality

$$(1+t)e^{-t} \geq e^{-t^2/2}$$

   holds for all $t \geq 0$.

   (e) Conclude that

$$E\alpha_m \geq \int_0^\infty e^{-mt^2/2} \, dt = \sqrt{\frac{\pi}{2}} \cdot \sqrt{m}.$$

   (f) Show that if $\epsilon > 0$, then there is some interval $0 \leq t \leq t_0$ in which $(1+t)e^{-t} \leq e^{(1/2-\epsilon)t^2}$, and some $c > 0$ such that the inequality

$$(1+t)e^{-t} \leq e^{-ct}$$

   holds whenever $t \geq t_0$.

   (g) Show that

$$\begin{aligned} E\alpha_m &\leq \int_0^{t_0} e^{-m\left(\frac{1}{2}-\epsilon\right)t^2} \, dt + \int_{t_0}^\infty e^{-mct} \, dt \\ &\leq \frac{\sqrt{\pi}}{2\sqrt{\frac{1}{2}-\epsilon}} \cdot \frac{1}{\sqrt{m}} + O(1/m), \end{aligned} \tag{66}$$

and conclude that

$$\sqrt{m} \cdot E\alpha_m \to \sqrt{\pi/2}.$$

(h) What is the asymptotics of the minimum of m sums of three independent exp(1)'s? What if exponentials are replaced by uniform [0,1]?

8. Find the derivative of the expression

$$-\int_{-u}^{\infty} F(t)\,dt$$

with respect to u. Obviously it is either $F(-u)$ or $-F(-u)$, but which one is it?

9. Evaluate the integral

$$\int_0^{\infty} -\log(1-e^{-x})\,dx$$

by making the substitution $t = e^{-x}$, taking the Taylor expansion of the new integrand, and integrating termwise.

10. Consider the (conjectured) cost

$$\mu(p) = \frac{-2\log(1-p)}{2-p}$$

of the most expensive edge in the solution to the incomplete matching problem where a proportion p of the vertices participate in the solution.

(a) Show that $\mu(p) = p + O(p^2)$ as $p \to 0$. Explain why this is what we should expect.

(b) Imagine that the matching problem could be solved by a greedy algorithm, so that a solution for larger p would consist in adding a set of edges to a solution for smaller p. Suppose $n/2$ edges were added in this way, and the cost of the most expensive one (the one that was last added) is all the time given by $\mu(p)$. Show that in this case the cost would asymptotically be a scaling factor times

$$\int_0^1 \mu(p)\,dp.$$

(c) Evaluate the integral.

11. Let C_k and C_{k-1} be the costs of the minimum k-matching and $(k-1)$-matching respectively. There is an obvious inequality between $C_k - C_{k-1}$ and the cost of the most expensive edge in the minimum k-matching. What does the result of Exercise 10 suggest about the degree of sharpness of this inequality? Actually the inequality is off by roughly a factor 2 for near-maximal k, but quite sharp for intermediate values of k.

12. Compute
$$\int_0^{1/2} \frac{-2\log(1-t)}{2-t}\,dt$$
numerically. Does it seem reasonable that this is the cost of an assignment that uses half of the vertices? Find a similar expression for the much easier limit cost of a set of $n/2$ edges in an n by n bipartite graph such that no two have a common vertex in V_1.

13. Evaluate
$$\int_0^1 \frac{(-\log t)^n}{t}\,dt$$
for general n.

14. What is the asymptotics of the cost of the MST with pseudodimension 2 edge costs?

15. Solve the following problem numerically with the non-rigorous method: Given an n by n bipartite graph. Consider the minimum total cost of a set of n edges under the constraint that each vertex can be incident to at most two of them. What is the limit as $n \to \infty$? What if the vertices in V_1 must have exactly one edge, and those in V_2 can have at most two?

16. Study the 2 by 2 random assignment problem with uniform $[0,1]$ edge costs. Find the expected cost, or give reasonably good bounds. Check by computer simulation that your answer seems correct. Another option is to argue that the answer has to be a rational number with a certain denominator, and then to find the numerator by simulation. Try the 3 by 3 case (here I don't know the answer). Don't try too hard though.

17. Prove that for every positive integer N,

$$\frac{1}{\sin^2\left(\frac{\pi}{4N}\right)} + \frac{1}{\sin^2\left(\frac{3\pi}{4N}\right)} + \frac{1}{\sin^2\left(\frac{5\pi}{4N}\right)} + \cdots + \frac{1}{\sin^2\left(\frac{(2N-1)\pi}{4N}\right)} = 2N^2.$$

18. (a) Show that the degree 4 analog is

$$\frac{1}{\sin^4\left(\frac{\pi}{4N}\right)} + \frac{1}{\sin^4\left(\frac{3\pi}{4N}\right)} + \frac{1}{\sin^4\left(\frac{5\pi}{4N}\right)} + \cdots + \frac{1}{\sin^4\left(\frac{(2N-1)\pi}{4N}\right)}$$
$$= \frac{8}{3}N^4 + \frac{4}{3}N^2.$$

   (b) Is there a simple explanation why this should be an even function of N (that is, why the odd degree coefficients are zero)?

   (c) Evaluate $\zeta(4)$.

19. Read Parisi's 1998 paper (a two-page note). Repeat his experiment: Study the random assignment problem with exponential edge costs for $n = 1, 2, 3, 4, 5$. For small n, we don't need a general polynomial time algorithm for the optimization problem, we can solve it by ad hoc methods. Verify that the expected cost seems to be 1, 5/4, 49/36 etc.

20. Do the same thing for the LP-relaxation of matching on the complete graph. Verify that the expected cost seems to be 1, 3/4, 31/36 etc.

21. Given the Parisi formula for exponential edge costs, what bounds can you establish for the expected cost of the assignment problem with uniform $[0, 1]$ edge lengths? In one of the directions there is an easy inequality. What are the obstacles when trying to establish a reasonable inequality in the other direction? More precisely, why isn't it easy to conclude that the expectation converges to $\pi^2/6$? Can you establish this anyway? If not, then make some explicit and reasonable assumptions, and show that convergence to $\pi^2/6$ follows from them.

22. Consider the n by n exponential assignment problem.

   (a) Using the Coppersmith-Sorkin formula, find the asymptotics of the expected difference $EC_{n,n,n} - EC_{n-1,n,n}$ between the cost of an n-assignment and an $(n-1)$-assignment.

(b) Show that whp, the cost of the most expensive edge in the minimum n-assignment is

$$(1 + o(1)) \cdot \frac{\log n}{n}.$$

Hint: Use Aldous 2^i-result for the probability of using the ith smallest edge, together with a simple analysis of the coupon collector's problem.

(c) Compare the two previous results, and try to find an intuitive explanation.

23. Let $n = 3$ and consider the minimum spanning tree problem with $\exp(1)$ edge costs.

(a) Show that an edge of cost s has probability $1 - (1 - e^{-s})^2$ of participating in the MST.

(b) Show that the expected contribution from a single edge to the cost of the MST is given by

$$\int_0^\infty e^{-s} \cdot s \cdot (1 - (1 - e^{-s}))\, ds$$

(c) Evaluate the integral.

(d) Let the edge costs be a, b and c. Show that $\min(a+b, a+c, b+c)$ can be expressed as a linear combination of $\min(a, b)$, $\min(a, c)$, $\min(b, c)$ and $\min(a, b, c)$.

(e) Use this to evaluate the expected cost of the MST.

24. Prove that for the complete graph with $\exp(1)$ edge costs, the expected cost of the minimum spanning tree is bounded as $n \to \infty$.

25. Choose three points A, B and C uniformly and independently in the unit square. What is the expected area of the triangle they determine? What is the expected area of the intersection of the unit square with the cone outside one of them (say A) determined by the lines to B and C?

26. Choose two points A and B as in the previous exercise.

(a) What is the expectation of the square of their distance? This is of course much simpler than the expectation of their distance.

(b) Hence what is the expectation of the area of the circle with center in A and with B on the perimeter?

(c) What is the expected area of the intersection of this circle (disk) with the unit square?

(d) Here is a harder one: What is the expected area of the intersection of the unit square with the disk on which AB is a diameter? You should be able to prove that it is smaller than $1/3$ (hint: If three points are chosen, then what is the probability that a given angle in the triangle is larger than $90°$?)

(e) By the way, what *is* the expected distance between A and B? We have all the moments of the difference in a given coordinate from exercise 1. Then we add two independent random variables, which means we still know all the moments . Is this enough? The square root of the answer to (a) is of course a bound, since the square-function is convex. Is this bound upper or lower?

27. If n points are chosen on a circle, what is the expectation of the smallest distance between them? What is the expectation of the largest gap? This can be solved in terms of barycentric subdivision of a simplex, but also by replacing the sizes of the gaps with exponential variables.

28. What is the probability that n points lie on one semicircle? Equivalently that the center of the circle is not in the convex hull of the n points? Two solutions, one geometrical based on splitting the choice, one in terms of gap sizes and exponential distribution.

29. In view of the previous exercises, suppose that the pairs of numbers $n, n+2$ that are both coprime to (product taken over the primes)

$$N = \prod_{p \le x} p$$

are distributed "randomly" on $\mathbf{Z}_N$ for large x. Show that if the largest gap between such pairs is not much greater than expected, then the twin primes conjecture holds, i. e. there are infinitely many primes p such that $p+2$ is also prime (this requires some basic knowledge of the distribution of primes). Similarly argue that the Goldbach conjecture is very likely true.

In fact evidence suggests that the pairs $n, n+2$ considered above are more evenly distributed than the obvious random model predicts, but the twin primes conjecture is still unsolved.

30. Consider the numbers n which are coprime to N, and assume that the maximum gap is not much larger than the random model predicts. Under this assumption, show that if $\theta > 1/2$, then for every sufficiently large x, there is a prime between x and $x + x^\theta$ (this is also predicted by the Riemann hypothesis).

31. Show that there is no interval of length $N/2$ in which the number of numbers coprime to N deviates by more than $2^{\pi(x)}$ from $\phi(N)/2$. Here $\pi(x)$ is the number of primes not exceeding x, and ϕ is the Euler ϕ-function.

    Conclude that the numbers coprime to N are, on a global scale, much more evenly distributed than would be predicted from the obvious random model.

    But this is not a course on sieve methods in number theory!

32. Let X be exponential of rate 1.

    (a) What is the median of X?

    (b) Prove that $E\,|X - \zeta|$ is minimized when ζ is the median of X.

    (c) Let ζ be the median of X, and calculate $E\,|X - \zeta|$. Is there a simple explanation for the answer?

    (d) Consider the "random optimization problem" $\min(X, \zeta)$, and use the participation probability-lemma.

    (e) Consider a sum $X_1 + X_2$ of two independent exponentials. Let ζ be the median of this sum. Prove that

    $$E\,|X_1 + X_2 - \zeta| = \frac{\zeta^2}{\zeta + 1}.$$

    (f) Generalize in the obvious way to several variables, for instance establish that for three variables,

    $$E\,|X_1 + X_2 + X_3 - \zeta| = \frac{\zeta^3}{\zeta^2 + 2\zeta + 2}.$$

33. Read Peter Winkler's paper *Games people don't play* (you find it on his web page). Solve the problem with the gladiators in the easy case by considering the expectation of $\min(x, y_1 + \cdots + y_n)$ for independent exponential variables.

34. Two variable optimization: How to turn a minimum of sums into a sum of minima. Example: Find the expectation of $\min(3x, x + y, 3y)$, where x and y are independent exp(1)-variables. Solution: $\min(3x, x + y, 3y) = \min(2x, y) + \min(x, 2y)$. If x and y are independent exp(1), this shows that the expectation is $4/3$. Generalize to minima of more than three sums (involving only two variables). What if $x \sim \exp(\alpha)$ and $y \sim \exp(\beta)$?

35. Study the derivations of the limits for matching and TSP using the cavity method. Study the following problem with the same method: Given a bipartite graph with n red and $2n$ blue vertices. Each edge has exponential(1) cost. The objective is to match each red vertex to two blue vertices so that each blue vertex is used exactly once.

    (a) Solve this exactly for $n = 1$ and $n = 2$ (or by simulation).

    (b) Find the limit non-rigorously with the cavity method. Express the answer as an integral of an elementary function, and evaluate this integral numerically. Think about the answer, does it seem reasonable?

    (c) Generalize.

36. Evaluate the TSP limit $L^\star$ numerically by using Maple (or some other program). How many digits can you get in a reasonable time?

37. Prove that the Coppersmith-Sorkin formula specializes to the Parisi formula in the case $k = m = n$. Show that if we collect terms for which $n - i$ and $n - j$ have gcd d, then the sum of these terms equals $1/d^2$.

38. Let n be a positive integer and let $0 \leq p \leq 1$. Choose m as the sum of n independent Bernoulli(p)-variables (0-1-variables with probability p of being 1). Show that the expected cost of the minimum m-assignment in an m by n matrix of independent exp(1)-variables is

$$p + \frac{p^2}{4} + \frac{p^3}{9} + \cdots + \frac{p^n}{n^2}.$$

39. If N points are chosen independently and uniformly in the unit square, what is the order of the distance between "nearby" points? What about in d dimensions?

40. The average distance to the nearest neighbor can be estimated by rescaling to a point set of density 1 in an infinite region ($\mathbf{R}^2$). This is simply a 2-dimensional Poisson process. What is the probability that a disk of radius r contains no point of the process?

41. Show that the average distance to the nearest neighbor is given by

$$\int_0^\infty e^{-\pi r^2}\, dr,$$

and evaluate the integral.

42. Show that in three dimensions, the average distance to nearest neighbor is

$$\frac{2^{1/3}\pi^{2/3}}{3^{7/6}\Gamma(2/3)}.$$

43. Can you give a rigorous lower bound on the β in the theorem of Beardwood-Halton-Hammersley?

44. Given that the volume of the d-dimensional unit ball is

$$\frac{\pi^{d/2}}{(d/2)!},$$

how does the distance to the nearest neighbor behave for large d? What is the highest dimension for which it is smaller than 1?

45. In dimension $d = 10^6$, estimate numerically the probability that the nearest neighbor is at distance smaller than 241 and smaller than 242 respectively.

46. Using the formula for the volume of the sphere in combination with Stirling's formula, make a guess at the asymptotics of $\beta(d)$ for large d.

47. In fact it was proved by Wansoo Rhee in 1992 that $\beta(d)/\sqrt{d}$ converges. Guess what the limit is.

48. For this problem, you need the correspondence between the assignment problem and the *aircraft passenger model*. Consider the n by n exponential assignment problem. Let a set S of edges have zero cost, while the remaining edges still have independent exponential costs.

    (a) What is the maximum participation probability that a nonzero edge can have?

    (b) Assume the very reasonable conjecture that there is always (unless S is the set of all edges) some nonzero edge whose participation in the minimum assignment is not positively correlated to the cost. Show that this gives a $O(\log n/n)$ bound on the variance of the cost.

49. For general minimization problems of sums of independent exp(1)-distributed variables, formulate a similar conjecture that there is always a variable whose participation is non-positively correlated with the value of the minimum.

    (a) Explain why this conjecture is "obviously true".

    (b) Explain why the argument in (a) is flawed.

    (c) Show that it would imply a general bound on the variance of the value of the minimization problem. There is more than one way of stating such an inequality, and it has to involve something like for instance the expectation of the largest value of a variable participating in the minimum.

    (d) Prove the conjecture for some infinite class of such problems (as general as you can).

50. The mean field model can be thought of as an approximation of a euclidean (geometric) model. The only difference is that in the mean field model the inter-point "distances" are independent. Will this work for or against shorter solutions to the classical optimization problems, spanning tree, matching, TSP?

51. Take the space to be a circle of perimeter 2. What is the distribution of the distance between two independent uniform points? Distance taken along the perimeter of the circle, that is, shortest of the two arcs. What are the limits of spanning tree, matching, TSP? Compare to the mean

field limits. Take n red and n blue points and consider the bipartite problems (this is not that trivial). Again compare.

52. Study matching of $2n$ points as follows: Suppose we choose two more points randomly on the circle, what is the probability that they divide the first $2n$ points in two sets of even size (a symmetry argument is available)? What is your conclusion?

53. Papers that compare mean field and euclidean models usually come to the conclusion that the mean field approximation is better in higher dimensions. Does this seem reasonable? Is there a simple explanation?

54. Given the n by n exponential random assignment problem. Given a number $c > 0$. What is the average number of assignments of cost at most c? What is your comment to the fact that this number is zero whp if $c < \zeta(2)$ and nonzero whp if $c > \zeta(2)$?

55. Read Viktor Dotsenko's paper [25] which claims to prove the Parisi conjecture. The paper is quite interesting, and was the first that gave a natural conjecture (that is, not containing a formula that came out of nowhere) that would have implied Parisi's conjecture. Unfortunately Dotsenko incorrectly claimed to have a proof of the conjecture, and the paper was therefore ignored by mathematicians. Find the mistake.

56. With the recursive equation for the higher moments of the cost in the assignment problem, recursively compute $E\sin(C_n)$ for $n = 1, 2, 3$.

57. For uniform $[0, 1]$ edge costs, consider the problem of maximizing the product of the edge costs in an assignment. What is the expected value of this for $n = 1, 2, 3$? Compute this using the Laplace transform (or equivalent) of the cost of the usual assignment problem.

58. For a complete graph on $3n$ vertices and exponential edge costs, consider the problem of covering the graph with n triangles. What is the limit cost?

59. Consider a 4 by 4 bipartite graph on vertex sets $U = \{u_1, \ldots, u_4\}$ and $V = \{v_1, \ldots, v_4\}$, exponential edge costs, and the problem of finding the minimum 2-assignment that uses one vertex of even label and one of odd label from each of U and V. What is the expected cost?

60. In an n by n matrix of exp(1) variables, pick, for $1 \le i \le n$, the element of row i which belongs to the minimum i-assignment in the submatrix consisting of the first i rows. What is the expected sum of the elements we have picked?

61. In an n by n matrix of exp(1) variables, let the edges be $e_1, \dots, e_{n^2}$, ordered by increasing cost. Pick the set of edges e_i such that the largest (in the sense of number of edges) matching in $\{e_1, \dots, e_i\}$ is larger than the largest matching in $\{e_1, \dots, e_{i-1}\}$. Find the asymptotic expected total cost of this edge set.'

62. Consider the following two-person game played on the complete graph K_n with exp(1) edge costs, and a randomly chosen starting point: The first player moves along an edge to a new vertex, and pays the cost of the edge to the other player. The second player now moves along an edge to another vertex, and pays the cost of this edge, and so on. A player cannot move to a vertex that has already been visited, and the game ends when all vertices have been visited.

    (a) Compute the expected payoff for $n = 2, 3, 4$.

    (b) Show by example that when $n = 4$, the best first move can be along the most expensive edge, and at the same time the one that belongs to the most expensive perfect matching.

    (c) Let $\phi(v)$ be the vertex to move to if the game starts at v. Show that when $n \to \infty$, $P(\phi^2(v) = v) \to 1$. I don't know how to solve this one, it is only a conjecture.

    (d) Let x be the cost of the optimal first move. Show that $n \cdot Ex \to \pi^2/6$. The same remark here, I don't know the answer.

    (e) Study the *twisted* version of this game: The player to move chooses two other vertex to go to, and the opponent then picks one of them. Equivalently, at each turn, the opponent can forbid one potential move.

63. Consider the *Exploration* game played on the square grid $\mathbf{Z}^2$, with edge cost which are 0 or 1 with equal probability. For what values of M does this game have a well-defined game theoretical value?

64. Let G be a finite graph with two distinguished vertices a and b. Suppose that the edges of G are given independent exp(1) costs. Show that the following operation computes the expected distance from a to b along the edges: Construct a matrix M whose rows are indexed by all vertex sets containing a but not b, and whose columns are indexed by all vertex sets containing b but not a. Let the entries of M be zero if the row and column respresent sets that intersect, and if they are disjoint, the number of edges connecting them. Then the expected distance from a to b is the sum of the entries of M^{-1}.

65. Consider the square grid $\mathbf{Z}^2$, with edge costs given as in the previous exercise. Find upper and lower bounds on the distance between $(0,0)$ and $(1,0)$.

66. If $n = 3$, then what are the possible values of Talagrand's distance $\rho(A,x)$?

67. Consider the following 2-person game: A counter is placed on square 0. There are squares numbered $1,\dots,N$, and to each square is associated a random variable X_i. The variables X_i are i.i.d, say uniform in $[0,1]$. The players take turns moving the counter one or two steps, and the player who moves to square i pays X_i to the opponent. Analyze the game. What if the player receives X_i (equivalently, if X_i is uniform in $[-1,0]$)?

References

[1] Achlioptas, Dimitris, Naor, A. and Peres, Yuval, *Rigorous location of phase transitions in hard optimization problems*, Nature **435** (205), 759–764.

[2] Achlioptas, Dimitris and Peres, Yuval, *The threshold for random k-SAT is* $2^k \log 2 - O(k)$, Journal of the AMS **17** (2004), 947–973.

[3] Aigner, M. and Dowling, T. A., *Matching theory for combinatorial geometries*, Trans. Amer. Math. Soc. **158** (1971), 231–245.

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

[5] Aldous, David, Steele, Michael J., *The Objective Method: Probabilistic Combinatorial Optimization and Local Weak Convergence* (2003). The paper can be found at

`http://www-stat.wharton.upenn.edu/~steele/Publications/PDF/OMSurvey.pdf`

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

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

[8] Aldous, David and Percus, Allon G., *Scaling and universality in continuous length combinatorial optimization*, Proceedings of the National Academy of Sciences of the United States of America, vol. 100 no. 20 (2003), 11211–11215.

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

[10] Alm, Sven Erick and Sorkin, Gregory B, *Exact expectations and distributions in the random assignment problem*, Combin. Probab. Comput. **11** (2002), no. 3, 217–248.

[11] Alon, N. and Spencer, J., Probabilistic Methods.

[12] Bacci, S. and Miranda, E. N., *The traveling salesman problem and its analogy with two-dimensional spin glasses*, J. Statist. Phys. **56** (1989), no. 3-4, 547–551.

[13] Beardwood, Jillian, Halton, John H. and Hammersley, John Michael, *The shortest path through many points*, Proc. Cambridge Phil. Soc., **55** (1959), 299–327. See also

`http://myweb.lsbu.ac.uk/~whittyr/MathSci/TheoremOfTheDay/OR/BHH/TotDBHH.pdf`

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

[15] Billera, Louis J., Brown, Kenneth S. and Diaconis, Persi, *Random Walks and Plane Arrangements in Three Dimensions*, The Mathematica Monthly **106** 6, June-July 1999.

[16] Boettcher, Stefan and Percus, Allon, *Nature's Way of Optimizing*, arXiv:cond-mat/9901351.

[17] Bollobás, Béla, Gamarnik, David, Riordan, Oliver and Sudakov, Benny, *On the value of a random minimum weight Steiner tree*, Combinatorica **24** (2) (2004), 187–207.

[18] Brunetti, R., Krauth, W., Mézard, M. and Parisi, G., *Extensive Numerical Simulation of Weighted Matchings: Total Length and Distribution of Links in the Optimal Solution*, Europhys. Lett. **14** 4 (1991), 295–301.

[19] Brunetti, R., Parisi, G. and Ritort, F., *On the Replica Symmetry for Random Weighted Matchings*, J. Phys. **A** (Math. Gen.) **24** (1991), 5077.

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

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

[22] Coppersmith, Don and Sorkin, Gregory B., *Constructive Bounds and Exact Expectations For the Random Assignment Problem*, Random Structures & Algorithms **15** (1999), 133–144.

[23] Coppersmith, Don, Sorkin, Gregory B., *On the expected incremental cost of a minimum assignment*. In Contemporary Mathematics (B. Bollobás, ed.), Vol. 10 of *Bolyai Society Mathematical Studies*, Springer.

[24] Donath, W. E., *Algorithm and average-value bounds for assignment problems*, IBM J. Res. Dev. **13** (1969), 380–386.

[25] Dotsenko, V., *Exact solution of the random bipartite matching model*, J. Phys. A, **33** (2000), 2015–2030, arXiv, cond-mat/9911477.

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

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

[28] Eriksson, Henrik, Eriksson, Kimmo and Sjöstrand, Jonas, *Exact Expectations for Random Graphs and Assignments*, Combinatorics, Probability and Computinf **12** (2003), 401–412.

[29] Fischer, K. H., and Hertz, J. A., *Spin Glasses*, Cambridge University Press 1991.

[30] Flaxman, Abraham D., Frieze, Alan M. and Krivelevich, Michael, *On the Random 2-Stage Minimum Spanning Tree*, Random Structures and Algorithms (2005?).

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

[32] Frieze, Alan M. and Grimmett, G. R., *The shortest path problem for graphs with random arc-lengths*, Discrete Appl. Math. **10** (1985). 57–77.

[33] Frieze, Alan M. and McDiarmid, Colin, *On random minimum length spanning trees*, Combinatorica **9** (1989) 363–374.

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

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

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

[37] Gamarnik, David, *The expected value of random minimal length spanning tree of a complete graph*, SODA05.

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

[39] Gilbert, E. N., *Random Minimal Trees*, Journal of the Society for Industrial and Applied Mathematics, Vol. 13, No. 2 (Jun., 1965), pp. 376–387.

[40] Goemans, M. X., and Kodialam, M. S., *A lower bound on the expected cost of an optimal assignment*, Math. Oper. Res., **18** (1993), 267–274.

[41] Guerra, F., *Replica Broken Bounds in the Mean Field Spin Glass Model*, Communications Math. Phys. **233** (2003), 1–12.

[42] Held, Michael and Karp, Richard M., *The traveling-salesman problem and minimum spanning trees*, Operations Research **18** Issue 6, Nov/Dec 1970, 1138–1162.

[43] Harris, T. E., *Lower bound for the critical probability in a certain percolation process*, Proc. Cambridge Phil. Soc. **56** (1960), 13–20.

[44] van der Hofstad, Remco, Hooghiemstra, Gerard and Van Mieghem, Piet, *Size and weight of shortest path trees with exponential link weights.* To appear in Combinatorics, Probability and Computing. Available at `http://www.win.tue.nl/~ rhofstad/`.

[45] van der Hofstad, Remco, Hooghiemstra, Gerard and Van Mieghem, Piet, *The Weight of the Shortest Path Tree.* Preprint 2005, available at `http://www.win.tue.nl/~ rhofstad/`.

[46] Houdayer, J., Boutet de Monvel, J. H. and Martin, O. C. *Comparing mean field and Euclidean matching problems*, Eur. Phys. J. B. **6** (1998), 383–393.

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

[48] Janson, Svante, *One, two and three times log n/n for paths in a complete graph with random weights*, Combin. Probab. Comput. **8** (1999), 347–361.

[49] Janson, Svante and Wästlund, Johan, *Addendum to "The minimal spanning tree in a complete graph and a functional limit theorem for trees in a random graph"*, Random Structures and Algorithms vol. 28 no. 4 (2006), 511–512.

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

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

[52] Karp, R. M., *An upper bound on the expected cost of an optimal assignment*, In Discrete Algorithms and Complexity: Proceedings of the Japan-U.S. Joint Seminar, Academic Press, 1987, 1–4.

[53] Karp, Richard M., Rinnooy Kan, A. H. G. and Vohra, R. V., *Average case analysis of a heuristic for the assignment problem*, Math. Oper. Res. **19** (1994), 513–522.

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

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

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

[57] Lazarus, A. J., *The assignment problem with uniform (0,1) cost matrix*, Master's thesis, Department of Mathematics, Princeton University, Princeton NJ.

[58] Lazarus, A. J., *Certain expected values in the random assignment problem*, Oper. Res. Lett., **14** (1993), 207–214.

[59] Linusson, Svante and Wästlund, Johan, *A generalization of the random assignment problem*, arXiv:math.CO/0006146.

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

[61] Lovász, L., Plummer, M. D., *Matching Theory*, North-Holland 1986.

[62] Martin, O. C., Mézard, M. and Rivoire, O., *A frozen glass phase in the multi-index matching problem*, cond-mat/0407623, Phys. Rev. Lett. **93** (2004), 217205.

[63] Mézard, M. and Montanari, A., Information, Physics and Computation. Book in preparation.

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

[65] Mézard, Marc and Parisi, Giorgio, *Replicas and optimization*, Journal de Physique Lettres **46** (1985), 771–778.

[66] Mézard, Marc and Parisi, Giorgio, *A replica analysis of the travelling salesman problem*, Journal de Physique **47**:1285–1296 (1986).

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

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

`http://hal.archives-ouvertes.fr/docs/00/21/05/75/PDF/ajp-jphys_1987_48_9_1451_0.pdf`

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

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

[71] Mézard, Marc and Parisi, Giorgio, *The Euclidean Matching Problem*, Journal de Physique **49** (1988), 2019.

[72] Mézard, Marc, Parisi, Giorgio and Zecchina, Riccardo, *Analytic and algorithmic solutions of random satisfiability problems*, Science 297 (2002), 812.

[73] Mézard, Marc, Parisi, Giorgio and Zecchina, Riccardo, *The random K-satisfiability problem: from an analytic solution to an efficient algorithm*, Phys. Rev. E **66** (2002), 056126.

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

[75] Nair, Chandra, Prabhakar, Balaji and Sharma, Mayank, *Proofs of the Parisi and Coppersmith-Sorkin random assignment conjectures*, Random Structures and Algorithms **27** No. 4 (2005), 413–444.

[76] Nair, Chandra, *Proofs of the Parisi and Coppersmith-Sorkin conjectures in the finite random assignment problem*, PhD thesis, Stanford University 2005.

[77] Olin, Birgitta, *Asymptotic properties of the random assignment problem*, Ph.D. thesis, Kungl Tekniska Högskolan, Stockholm, Sweden, (1992).

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

[79] Pardalos, P. M. and Ramakrishnan, K. G., *On the expected optimal value of random assignment problems: Experimental results and open questions*, Comput. Optim. Appl.**2** (1993), 261–271.

[80] Parisi, Giorgio, *A sequence of approximate solutions to the S-K model for spin glasses*, J. Phys. **A 13** (1980), L-115.

[81] Parisi, Giorgio, *Spin glasses and optimization problems without replicas*, in Les Houches, Session 46, 1986 —Le hasard et la matière/Chance and matter (editors J. Souletie, J. Vannimenus and R. Stora), Elsevier Science Publishers B. V., Netherlands 1987.

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

[83] Passare, Mikael, *How to Compute* $\sum 1/n^2$ *by Solving Triangles*, American Mathematical Monthly, October 2008.

[84] Percus, Allon G. and Martin, Olivier C., *Finite Size and Dimensional Dependence in the Euclidean Traveling Salesman Problem*, Phys. Rev. Lett. **76**, Issue 8 - 19 February 1996, 1188–1191.

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

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

[87] Ratiéville, M., , Ph.D. thesis 2002, Université Pierre et Marie Curie, Paris and Università degli Studi di Roma "La Sapienza", Rome.

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

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

[90] Roberts, David P., *Nash equilibria of Cauchy-random zero-sum and coordination matrix games.*

[91] Semerjian, Guilhem, *Modèles dilués en physique statistique: Dynamiques hors d'équilibre et algorithmes d'optimisation*, PhD thesis, Université paris 6 (2004-06-28). oai:arXiv.org:cond-mat/0406682.

[92] Sherrington, D. and Kirkpatrick, S., *Solvable model of a spin glass*, Phys. Rev. Lett. **35** (1975), 1792–1796.

[93] Sourlas, N., *Statistical Mechanics and the Travelling Salesman Problem*, Europhys. Lett., **2** No. 12 (1986), 919–923.

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

[95] Steele, J. Michael, *Minimal Spanning Trees for Graphs with Random Edge Lengths*, Mathematics and Computer Science II: Algorithms, Trees, Combinatorics and Probabilities, B. Chauvin, Ph. Flajolet, D. Gardy, and A. Mokkadem (eds.), Birkhäuser, Boston, 2002.

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

[97] Talagrand, Michel, *Huge Random Structures and Mean Field Models for Spin Glasses*, Proceedings of the Berlin international Congress of Mathematicians. Documenta Math., Extra Volume I, 1998.

[98] Talagrand, Michel, *Spin glasses: a new direction for probability theory?*, Atti Accad. Naz. Lincei Cl. Sc. Fis. Math. Nat. Rendi. Lincei Mat. Appl., special issue (2000), 127–146.

[99] Talagrand, M., *An assignment problem at high temperature*, preprint (2000).

[100] Talagrand, M., *Verres de spin et optimisation combinatoire*, Séminaire Bourbaki Mars 1999, Asterisque 266, 287–317 (2000).

[101] Talagrand, Michel, Spin Glasses: A Challenge for Mathematicians — Cavity and Mean Field Models, Springer 2003.

[102] Talagrand, Michel, *The Parisi Formula*, Annals of Mathematics **163** no. 1 (2006), 221–263.

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

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

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

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

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

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

[109] Wästlund, Johan, *Evaluation of Janson's constant for the variance in the random minimum spanning tree problem*, Linköping Studies in Mathematics No. 7 (2005).

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

[111] Wästlund, J,. *An easy proof of the zeta(2) limit in the random assignment problem*, submitted for publication.

[112] Wästlund, J,. *Random matching problems on the complete graph*, Electronic Communications in Probability 2008.

[113] Wästlund, J., *Random assignment and shortest path problems*, invited paper in Proceedings of the Fourth Colloquium on Mathematics and Computer Science, September 2006, Institut Élie Cartan, Nancy France.

[114] Wästlund, J., *An elementary proof of Wallis' product formula for pi*, The American Mathematical Monthly, December 2007,

[115] Wästlund, J., *The mean field traveling salesman and related problems*, submitted for publication.

[116] Yukich, J. E., *Asymptotics for the Euclidean TSP with power weighted edges*, Probability Theory and Related Fields **102** (1995), 203–220.

14 Laboration

1. Generera en exponentialfördelad pseudoslumpvariabel och testa dess medelvärde och varians. Om man startar med en likformigt fördelad variabel U i $[0,1]$, får man en exponentialfördelad variabel $X = -\log(1-U)$. Givetvis skulle $-\log U$ gå lika bra, men det finns två fördelar med $-\log(1-U)$. Dels ger detta en koppling där exp-variabeln alltid är större än den likformiga, dels undviker man en krasch om $U = 0$, vilket kan hända i vissa implementeringar.

2. Verifiera att $E \min(X_1, \ldots, X_n) = 1/n$. Vad är variansen?

3. Gör Parisis experiment [82], för $n = 2, 3, \ldots$. Testa också att variansen stämmer med formeln (54).

4. Simulera den LP-relaxering av matchning på K_3 som beskrivs i avsnitt 4.3 i [115]. Hur ofta utgör respektive konfiguration minimum?

5. Implementera en generell (effektiv) algoritm för bipartit matchning. Testa matchningsproblem med exp(1)-variabler. Hur stort n klarar algoritmen på rimlig tid? Stämmer det att gränsvärdet blir $\pi^2/6$, dvs ungefär 1.64?

   Ungefär hur ofta används det minsta elementet i en rad i matrisen, dvs den billigaste kanten från en given nod? Hur ofta används den näst minsta?

   Hur dyr är den dyraste kanten som ingår i den optimala lösningen?

   Koppla likformiga och exponentialfördelade variabler. Testa hur stor avvikelsen i allmänhet blir, dvs hur mycket mindre blir kostnaden för den minimala matchningen då variablerna är likformiga i stället för exp-fördelade?

   Hur mycket skiljer sig de två optimala matchningarna åt, dvs hur många kanter byts ut då man byter kostnaderna $U_{i,j}$ mot $-\log(1-U_{i,j})$?

6. Modifiera algoritmen så att den klarar LP-relaxerad matchning på kompletta grafen, inklusive loopar. Hur vanligt är kanter med koefficient 1/2 i den optimala lösningen? Hur verkar variansen bete sig som funktion av n?