Annals of Mathematics **175** (2012), 1061–1091
http://dx.doi.org/10.4007/annals.2012.175.3.2

Replica symmetry of the minimum matching

By Johan Wästlund

Abstract

We establish the soundness of the replica symmetric ansatz introduced by M. Mézard and G. Parisi for the minimum matching problem in the pseudo-dimension d mean field model for $d \geq 1$. The case $d = 1$ corresponds to the $\pi^2/6$-limit for the assignment problem proved by D. Aldous in 2001.

We introduce a game-theoretical framework by which we establish the analogous limit also for $d > 1$.

1. Introduction

1.1. *Background.* It has been known for some time that methods of the statistical mechanics of disordered systems apply to certain problems of combinatorial optimization. Much of the work in this direction stems directly or indirectly from G. Parisi's solution [30] of the Sherrington-Kirkpatrick model [38] of spin glasses, established rigorously by M. Talagrand [40]. In [18], S. Kirkpatrick and G. Toulouse suggested the mean field traveling salesman problem (TSP) as an archetypal optimization problem sharing important features with spin glasses. M. Mézard and G. Parisi [22], [24], [23], [25], [31] and Mézard and W. Krauth [19] obtained several remarkably detailed predictions about minimum matching and the TSP with the replica and cavity methods. These predictions were based on the assumption of *replica symmetry* which is known to fail at low temperatures for models of spin glasses. It became clear that minimum matching and the TSP are different in this respect from models like the Sherrington-Kirkpatrick model and random k-SAT. Several authors have verified the consistency of the replica symmetric ansatz by testing its various predictions numerically and theoretically [5], [9], [10], [11], [16], [27], [29], [32], [34], [35], [39].

I gratefully acknowledge the hospitality of the Mittag-Leffler institute, Djursholm, Sweden, and the Theory Group of Microsoft Research, Redmond, USA.

Replica symmetry is interesting from an algorithmic point of view since it is linked to the efficiency of Belief Propagation heuristics; see for instance [26]. Recently J. Salez and D. Shah [37], [36] have obtained rigorous results in this direction for the assignment problem, and in view of our results, their conclusions should be valid in a broader range of random models.

In this paper we introduce a type of two-person game that we call *Exploration games*, which are played on graphs with lengths assigned to the edges. They are related to the family of *Geography* games that have been studied from the perspective of combinatorial game theory [12].

Exploration in its simplest form corresponds in a certain way to the minimum matching problem, while modifications of the rules will produce games corresponding to other optimization problems. In particular, a so-called *comply-constrain* version of Exploration corresponds to the TSP, but the analysis of such variations is beyond the scope of the present paper. We show that on certain infinite graphs with random edge lengths, Exploration has an almost surely well-defined game theoretical value, and we argue that this property is the essence of replica symmetry for the minimum matching problem.

The paper is organized as follows. In the rest of Section 1 we give further background, and in Section 1.3 we state our main result, Theorem 1.1. In Section 2 we develop some tools necessary for the proof of this theorem. The most important new concepts are the Exploration game and the analysis of this game on certain random trees in Sections 2.5–2.6. In Section 3 we use the results of Section 2 to conclude the proof of Theorem 1.1.

Acknowledgment. I thank the referees for suggestions that greatly improved the presentation.

1.2. *Random distance models.* By a *distance model* we mean a graph whose edges are assigned nonnegative real numbers that we think of as lengths. A *random distance model* is simply a distance model chosen according to some probability measure. In general, the underlying graph as well as the edge-lengths may be random. If the underlying graph is a fixed graph G and the edge-lengths are independent identically distributed according to a given measure, we say that the model is *mean field*. If the distribution of the edge-lengths $l_{i,j}$ is given by a density function ρ so that

$$P(a \leq l_{ij} \leq b) = \int_a^b \rho(l)\ \mathrm{d}l$$

for $0 \leq a \leq b$, then we denote the model by $G[\rho]$.

A graph of particular interest is the complete graph K_n on n vertices. A number of results have been established for $K_n[\rho]$ when ρ satisfies

$$\lim_{l \to 0^+} \rho(l) = 1. \tag{1}$$

It was shown by A. Frieze [13] that the total length of the minimum spanning tree in $K_n[\rho]$ converges in probability to $\zeta(3)$.

A *perfect matching* in a graph is a set of edges of which each vertex is incident to exactly one. In a distance model G, the minimum matching problem asks for the perfect matching of minimum total length. Assuming that a perfect matching exists, we denote the total length of the minimum matching by $M(G)$. For the minimum matching problem on $K_n[\rho]$ (assuming n is even), it was conjectured by M. Mézard and G. Parisi [22] that the limit of the total length $M(K_n[\rho])$ of the solution is $\pi^2/12$. This was proved by D. Aldous [2] in 2001. For the TSP, a similar limit was conjectured by Parisi, Mézard and Krauth [19], [22], and proved in [42].

Starting with the famous paper [8] by J. Beardwood, J. H. Halton and J. M. Hammersley, much work in the same direction has been done for *euclidean* models, where the vertices are distributed randomly in the unit cube in $\mathbb{R}^d$ and edge lengths are given by euclidean distance. It is interesting to compare euclidean and mean field models [11], [16]. What is difficult in one model can be easy in the other. The independence of the edge-lengths makes mean field models tractable, while in euclidean models the geometry can be exploited.

This has motivated the study of *mean field approximations* of euclidean models, where the idea is to choose the density function ρ in order to mimic distances in d-dimensional space. In this context, (1) corresponds to a 1-dimensional space, while in d dimensions, we expect the distribution of $l_{i,j}$ to satisfy

$$P(l_{i,j} < r) \sim \text{constant} \cdot r^d \tag{2}$$

for small r, since the probability of a point being within distance r of another should be proportional to the volume of a ball of radius r. We therefore say that a mean field model has *pseudo-dimension* d if

$$\frac{\rho(l)}{l^{d-1}} \to c \tag{3}$$

for some nonzero constant c, as $l \to 0$.

In the context of mean field models, there is no reason to require d to be an integer. In order to obtain a valid probability measure it suffices to assume that $d > 0$.

1.3. *Minimum matching and statement of the main theorem.* The method employed in [22] is applicable also for $d \neq 1$, and minimum matching seems to be the simplest problem that allows those ideas to be displayed in a nontrivial way. We start by discussing the normalization. We fix a density function ρ satisfying (3) and study the asymptotics of $M(K_n[\rho])$. This obviously requires n to be even unless we allow one vertex to be left out of the pairing, but this is a minor issue since we are mainly interested in the large n asymptotics.

It is not difficult to guess roughly how $M(K_n[\rho])$ scales with n. From an arbitrary vertex, the order of the distance to the nearest neighbors is obtained by setting $P(l_{i,j} < r) \approx 1/n$. In view of (3),

$$P(l_{i,j} < r) = \int_0^r \rho(l)\,\mathrm{d}l \approx c \cdot \int_0^r l^{d-1}\,\mathrm{d}l = \frac{c}{d} r^d,$$

which leads to

$$r \approx \text{constant} \cdot n^{-1/d}. \tag{4}$$

To simplify, we may assume that the constant c in (3) is equal to d, so that the constant in (4) is equal to 1 and $P(l_{i,j} < r) \approx r^d$. This leads to

$$\rho(l) \sim dl^{d-1}$$

for small l, or to be precise,

$$\lim_{l \to 0^+} \frac{\rho(l)}{dl^{d-1}} = 1. \tag{5}$$

If we believe that edge-lengths of order $n^{-1/d}$ will dominate the minimum matching, then since a matching contains $n/2$ edges, we expect $M(K_n[\rho])$ to scale with n like $n^{1-1/d}$. It is natural to conjecture that $M(K_n[\rho])/n^{1-1/d}$ converges in probability to a constant depending on d. Our main result is that this is true for $d \geq 1$.

THEOREM 1.1. *For every $d \geq 1$, there is a number $\beta(d)$ such that if ρ is fixed and satisfies* (5), *then as $n \to \infty$,*

$$\frac{M(K_n[\rho])}{n^{1-1/d}} \xrightarrow{\mathrm{p}} \beta(d). \tag{6}$$

Moreover, the proof leads to a method for computing $\beta(d)$ numerically. We believe that in principle the method applies also when $0 < d < 1$, but we have run into some difficulties that have so far prevented us from establishing (6) in that case.

1.4. *Examples.* It is worth pointing out how the normalization works in a couple of simple examples. Suppose we take the distribution of edge lengths as $l = \max(U, V)$, where U and V are independent and uniform in $[0, 1]$. Then

$$P(l < r) = P(U < r)^2 = r^2$$

if $r \leq 1$, and therefore the distribution belongs to the case $d = 2$.

We can think of $\beta(d)$ as the average cost per vertex of obtaining a perfect matching, measured in a unit r_0 of length such that the expected number of neighbors within distance r_0 of a given vertex is 1.

The unit r_0 will depend on n, and we can determine r_0 asymptotically by putting $P(l_{i,j} < r_0) \sim 1/n$, which gives

$$r_0 \sim \frac{1}{\sqrt{n}}.$$

Since there are n vertices, the total length of the minimum matching is asymptotically

$$\beta(2)\sqrt{n}.$$

If on the other hand we take the edge-lengths to be distributed like $U+V$, then $P(l < r) \sim r^2/2$, and the unit of length is given by $r_0^2/2 = 1/n$ or equivalently

$$r_0 = \frac{\sqrt{2}}{\sqrt{n}}.$$

In that case the asymptotical total length of the minimum matching is

$$\beta(2)\sqrt{2n}.$$

Although it does not follow from Theorem 1.1, our results apply, through a modification of Proposition 2.2, also to the *assignment problem*, in other words minimum matching on the complete bipartite graph $K_{n,n}$. In this model, each vertex is connected to n other vertices and therefore the nearest neighbor distances scale with n in the same way as in K_n, but the difference is that there are $2n$ vertices instead of n. With the two distributions $\max(U,V)$ and $U+V$, the minimum assignments will therefore have lengths asymptotically $2\beta(2)\sqrt{n}$ and $2\beta(2)\sqrt{2n}$ respectively.

1.5. *Replica predictions.* Within the framework of the replica method, Mézard and Parisi obtained an analytical characterization of $\beta(d)$ which is conjectured to be correct for all $d > 0$. They arrived at an integral equation equivalent to

$$F(x) = \exp\left(-d\int_0^\infty l^{d-1}F(l-x)\,\mathrm{d}l\right), \tag{7}$$

from which $\beta(d)$ is obtained as

$$\beta(d) = \frac{d^2}{2}\iint\limits_{\substack{-\infty<x,y<\infty\\ x+y\geq 0}} (x+y)^{d-1}F(x)F(y)\,\mathrm{d}x\,\mathrm{d}y.$$

The method is inherently nonrigorous, and it has not been established that (7) has a unique solution except in the case $d = 1$, where the solution

$$F(x) = \frac{1}{1+e^x}$$

leads to $\beta(1) = \pi^2/12$.

Numerical estimates of $\beta(d)$ for small integral d have been obtained in [16], [22]. Through an approximate solution of (7), we have found that

$$\beta(2) \approx 0.57175904959888.$$

We have no theoretical estimate of the error, but these decimals seem to be stable. Although we still do not know whether (7) has a solution, the numerical result can be regarded as an approximation of the fixed point of V_λ (see §2.7) for an appropriately chosen λ, and therefore apart from the numerical error the result is backed up rigorously.

1.6. *Rigorous results.* On the mathematical side there has been considerable progress on the case $d = 1$. In particular, the $\pi^2/6$-limit in the assignment problem has received several different proofs [2], [20], [28], [41]. From our point of view the result corresponds to the statement that for $d = 1$, the limit in (6) exists and $\beta(1) = \pi^2/12$, but the asymptotic equivalence between assignment and matching on the complete graph is by no means trivial and does not follow from [20], [28], [41]. The proofs together provide a quite detailed picture of the distribution of the total length as well as various local statistics of the optimum solution, and the analogous result for the TSP was established in [42]. However, the proofs in [20], [28], [41], [42] are very different from the approach in the physics literature and do not seem to generalize to $d \neq 1$. The original proof by Aldous [2] is the one that comes closest to justifying the replica symmetric ansatz (particularly in view of additional results in [6], [37]), but it seems to rely on finding a solution to (7).

In the present paper we show that the calculations in [22] are sound for quite general reasons. We prove that for $d \geq 1$, (6) holds, and we characterize $\beta(d)$ analytically in terms of certain integral equations similar to (7). Although we cannot find explicit solutions to these equations when $d \neq 1$, our results show that the numerical computation of $\beta(2)$ in [22] is correct in principle.

For $d = 1$, much more detailed results can be obtained. A more precise analysis of the $d = 1$ case with the present method and a clarification of its relation to the results of [42] is given in joint work with G. Parisi [33].

Our approach is “zero temperature,” but similar to the statistical physics method in that we reach the optimum solution through a limiting process. We introduce a parameter λ and study “diluted” problems where partial matchings are allowed but receive a penalty of λ for each missing edge. The original problem is recovered in the limit $\lambda \to \infty$. The parameter λ plays a role similar to the inverse temperature in statistical physics. Finite λ allows for a certain local freedom that destroys all long-range interactions. In particular, adding or deleting a vertex has only a local effect on the optimum solution. In [22] a similar assumption seems to be crucial for the renormalization that leads to (7).

2. Local convergence and the Exploration game

2.1. *Rooted distance models and total variation distance.* The first step towards proving Theorem 1.1 is to establish a certain local convergence result for $K_n[\rho]$. The idea is the same as in the papers [1], [2] by Aldous, but we use a stronger convergence concept which is more convenient in our approach.

We first introduce *rooted* distance models and the total variation distance, which defines our concept of convergence. To obtain local convergence, we then have to rescale the model $K_n[\rho]$ essentially by setting the "unit of distance" r_0 equal to 1. It seems that we could have done this immediately and formulated Theorem 1.1 accordingly. However, this would involve letting ρ depend on n. The problem is that in order for Theorem 1.1 to hold, it is not enough to assume that (3) holds for every n. Some uniformity in n is required, and therefore the present setting seems the most natural after all.

A rooted distance model is a distance model where one of the vertices is distinguished as the root. More generally, if m is a nonnegative integer, an m-rooted distance model is a distance model with m labeled roots. Technically, the "rooting" is a function from $\{1, \dots, m\}$ to the vertex set, and thus we do not *a priori* require the roots to be distinct. When we speak of rooted *random* distance models, we always assume that the number m of roots is fixed (nonrandom).

Two m-rooted distance models are *isomorphic* if there is a bijection between their vertex sets and a compatible bijection between their edge sets such that the i:th roots are mapped to each other for $1 \le i \le m$ and the graph structure and edge-lengths are preserved. We say that a property E of m-rooted distance models is *invariant* if it depends only on the isomorphism type of the model.

Suppose G and H are random m-rooted distance models. If E is an invariant property, then provided E is measurable in both G and H, the two random processes G and H will assign probabilities $P_G(E)$ and $P_H(E)$ to E. We will not discuss issues of measurability here, since all events that we actually consider will be measurable in all our models. We define a *total variation metric* on rooted random distance models as follows.

Definition 2.1. Let G and H be rooted random distance models (with the same number of roots). The *total variation distance* between G and H is

$$d_{\mathrm{TV}}(G, H) = \sup_{E \text{ invariant}} \left| P_G(E) - P_H(E) \right|.$$

Here taking absolute values in the right-hand side is strictly not necessary, since replacing E by its complementary event will change the sign of $P_G(E) - P_H(E)$. The total variation metric gives us a concept of convergence. We say that $G_n \to G$ in total variation if G_n is a sequence of (rooted) random distance models whose total variation distance to G tends to zero.

There is another way to look at the total variation distance, which is convenient although formally not necessary. The statement that $d_{\mathrm{TV}}(G, H) \leq \varepsilon$ is equivalent to the existence of a coupling of the two processes G and H with the property that with probability at least $1-\varepsilon$, G and H are isomorphic. This equivalence holds generally and can be established for finite probability spaces by a simple inductive argument. For general probability spaces, the proof requires some measure theory that we will not discuss here.

2.2. *Rescaling and local convergence.* To obtain local convergence, we will work with a rescaled version of the pseudo-dimension d mean field model. On the graph K_n, we take the edge-lengths to be $n^{1/d} l'_{i,j}$, where the variables $l'_{i,j}$ are given by a density function ρ satisfying (5). To avoid excessive indexing, we will suppress the dependence on d and ρ, and just write K_n for this model, and K_n^m when it is rooted in m distinct arbitrarily chosen vertices.

If k is a nonnegative integer and $\lambda > 0$, we define the (k, λ)-*truncation* of an m-rooted distance model to be the union of all paths of at most k edges, each of length at most λ, starting from some root. We let $K_n^m(k, \lambda)$ denote the (k, λ)-truncation of K_n^m.

For fixed d, m, k and λ, the sequence $K_n^m(k, \lambda)$ converges in total variation to a certain limit model which we now describe. Readers familiar with the papers [1], [2] by Aldous and [4] by Aldous and J. M. Steele will recognize that for $m = d = 1$, the limit is the (k, λ)-truncation of a *Poisson Weighted Infinite Tree* (PWIT). Here we do not need the full PWIT, and therefore describe the limit object directly.

The *d-dimensional m-rooted* (k, λ)*-forest* $T^m(k, \lambda)$ is constructed as follows. First we let each of the m roots give rise to a Galton-Watson process with Poisson(λ^d) offspring distribution, which is truncated after k generations, meaning that the vertices k steps away from the roots, if there are any, do not give rise to further offspring. Finally all edges are assigned independent lengths distributed like $X^{1/d}$ for X uniformly distributed on $[0, \lambda^d]$.

There are a couple of other equivalent ways of defining $T^m(k, \lambda)$. It can be defined as the (k, λ)-truncation of m independent d-dimensional PWITs [4]. In particular, instead of first constructing the graph and then assigning edge-lengths, we can generate the children of each new vertex through an inhomogeneous Poisson point process of rate dl^{d-1} on the interval $0 \leq l \leq \lambda$ that simultaneously produces the children and the lengths l of the corresponding edges. The fact that the various ways of constructing $T^m(k, \lambda)$ are equivalent follows from elementary properties of the Poisson process.

PROPOSITION 2.2. *As* $n \to \infty$,

$$K_n^m(k, \lambda) \to T^m(k, \lambda)$$

in total variation.

We do not prove Proposition 2.2 in detail. In principle it follows from the results of [1], and a similar argument in the context of Erdös-Rényi graphs is carried out in detail in [17, §5.2]. A sketch of the idea of the proof is as follows. First consider the structure of the underlying graph. Here it suffices to show that each finite m-rooted forest occurs as the underlying graph of $K_n^m(k, \lambda)$ with asymptotically the same probability as in $T^m(k, \lambda)$ when $n \to \infty$. There is a countable infinity of such forests, but since their probabilities in $T^m(k, \lambda)$ add to 1, convergence is automatically uniform. Then consider the edge-lengths. First specify a finite m-rooted forest, and for each edge, a subinterval of $[0, \lambda]$ where the length of the edge is supposed to be. Then one can show that the event of $K_n^m(k, \lambda)$ being this particular forest with edge-lengths in the given intervals has asymptotically the correct probability.

2.3. *The Exploration game.* The following two-person zero-sum game that we call *Exploration* was invented in an attempt to find a mathematically sound interpretation of (7). We are given a distance model with a particular vertex v chosen as a starting point, and a nonnegative parameter λ. Alice and Bob take turns choosing the next edge of a self-avoiding walk, with Alice starting the game from v. The player who makes a move pays an amount equal to the length of the chosen edge to the opponent. At each turn, the player who is about to move also has the option of, instead of moving, terminating the game and paying a penalty of $\lambda/2$ to the opponent. If the game terminates, each player's *payoff* is the total amount they received from the opponent minus the total amount they paid to the opponent. Each player tries to maximize their own payoff.

Notice that there is no randomness in the game. The players are assumed to have perfect information about the graph and the edge-lengths. We can immediately make the following observations:

(1) If the graph is finite, then there is a well-defined game-theoretical value (payoff under mutual optimal play). By convention, the value of a position in the game is the future payoff under optimal play for the player who moved to the position. Thus for consistency, the value of the starting position is Bob's payoff under optimal play.

(2) If the graph is infinite, there may or may not be such a value. For instance, if the graph is an infinite path where all edges have the same length $l < \lambda$, then each player will receive a better payoff if the other one terminates the game. It therefore appears that no player will ever terminate, but then there will not even be a well-defined payoff, let alone payoff under optimal play.

(3) Edges of length more than λ are irrelevant to the game. If Alice moves along such an edge, then Bob can terminate the game, and even though

this may not be Bob's best option, it would still have been better for Alice to terminate in the first place.

2.4. *The diluted matching problem.* There is a relaxation of the minimum matching problem that we refer to as the *diluted* matching problem. Instead of requiring each vertex to be covered by the matching, we allow for any partial matching, with a penalty of $\lambda/2$ for each unmatched vertex. This idea was considered in the context of multi-index matchings in [21].

For the moment we regard the parameter λ as fixed. If G is a finite distance model, we let $M_\lambda(G)$ be the cost of the diluted matching problem on G. More precisely, $M_\lambda(G)$ is the minimum, taken over all partial matchings, of the sum of the edge lengths in the matching plus $\lambda/2$ times the number of unmatched vertices. Moreover, we let $f(G, v)$ be Bob's payoff under mutual optimal play in Exploration on G with v as starting point.

Proposition 2.3. *Let G be a finite distance model, and let v be a vertex of G. Then*

$$f(G, v) = M_\lambda(G) - M_\lambda(G - v).$$

Proof. Suppose that the neighbors of v are $v_1, \dots, v_k$ and that the edges from v to these neighbors have lengths $l_1, \dots, l_k$. We first claim that f can be recursively characterized by

$$f(G, v) = \min(\lambda/2,\, l_i - f(G - v, v_i)). \tag{8}$$

Here the right-hand side is the minimum of $\lambda/2$, which is Bob's payoff if Alice chooses to terminate immediately, and $l_i - f(G - v, v_i)$, for $1 \leq i \leq k$, which we claim is Bob's payoff if Alice chooses to go from v to v_i in her first move and the rest of the game is played optimally. When Alice moves to v_i Bob receives a payment of l_i, and the rest of the game is played on the graph $G - v$ with v_i as starting point. In that game, Bob has the role of the first player, so his payoff in the rest of the game will be $-f(G - v, v_i)$, and for the whole game, $l_i - f(G - v, v_i)$. Since Alice wants to maximize her own payoff, she will make the choice that minimizes Bob's payoff, and this establishes (8).

On the other hand, the cost of the diluted matching problem satisfies the similar recursion

$$M_\lambda(G) = \min(\lambda/2 + M_\lambda(G - v), l_i + M_\lambda(G - v - v_i)). \tag{9}$$

Here we minimize over the options of leaving v unmatched, which will result in a total cost of $\lambda/2 + M_\lambda(G - v)$, or matching v to v_i, which leads to a total cost of $l_i + M_\lambda(G - v - v_i)$. Subtracting $M_\lambda(G - v)$ from both sides of (9) yields

$$M_\lambda(G) - M_\lambda(G - v) = \min(\lambda/2,\, l_i - (M_\lambda(G - v) - M_\lambda(G - v - v_i))).$$

Comparing this to (8), we see that $f(G, v)$ and $M_\lambda(G) - M_\lambda(G - v)$ satisfy the same recursion, and it follows by induction that they are equal. $\square$

It is clear from Proposition 2.3 and its proof that Alice achieves optimal payoff by starting along the edge of the optimum diluted matching if there is such an edge from v, and by terminating immediately if v is unmatched. By induction it follows that consistently playing along edges of the optimum diluted matching, and terminating when no such edge is available, is minimax optimal. Therefore under mutual optimal play, the path described by the game is the symmetric difference of the optimum diluted matchings on G and $G - v$.

Since the diluted matching problem can be solved efficiently by standard matching algorithms, it follows that Exploration can be played optimally with a polynomial time algorithm, but from our perspective this is beside the point. The advantage of introducing the game is that if the graph is infinite, there may still be a well-defined game-theoretical value. This value then replaces $M_\lambda(G) - M_\lambda(G - v)$ and allows for the equivalent of the renormalization argument of [22] in a mathematically consistent way.

2.5. *Exploration on* $T^1(k, \lambda)$. We now fix d and λ and assume that $d \geq 1$, although some of the results hold also for $0 < d < 1$. In view of the results of Section 2.2 it makes sense to study Exploration played on the d-dimensional (k, λ)-tree $T^1(k, \lambda)$. Our arguments will involve letting $k \to \infty$, and they are easiest to formulate on the potentially infinite tree obtained in the same way as $T^1(k, \lambda)$ but avoiding the truncation after k generations. We denote this tree by T, suppressing the dependence of λ. Hence T is a Galton-Watson tree based on Poisson(λ^d)-distribution, with independent identically distributed edge-lengths distributed as the d:th root of a variable of uniform distribution on $[0, \lambda^d]$.

If v is a vertex of T, we let $T(v)$ be the subgraph that can be reached from v by downward paths. (Here and in the following, we think of the trees as growing downwards, so that the root is at the top, and "down" means away from the root.)

We study Exploration played on T starting from the root, which we denote by ϕ. Since T is potentially infinite, it is not clear that concepts like optimal play are well defined. We would like to define a function f describing the value of having moved to a vertex v. More precisely, $f(v)$ should be the payoff under optimal play for the player who does not start, if the game is played on $T(v)$ starting from v. If such a value $f(v)$ can be defined consistently, then by an argument similar to the proof of Proposition 2.3, f must satisfy

$$f(v) = \min(\lambda/2, l_i - f(v_i)), \tag{10}$$

where l_i is the length of the edge to the i:th child v_i of v, and the minimum is taken over $\lambda/2$ and the sequence of $l_i - f(v_i)$ as v_i ranges over all children of v.

For a given realization of T, we say that a function f from its vertices to the real numbers is a *valuation* if it satisfies (10). A valuation can be regarded as a consistent way for a player to assess the positions of the game. We observe the following:

- A valuation must satisfy $-\lambda/2 \leq f(v) \leq \lambda/2$ at every vertex v.
- If v is a leaf of T, then $f(v) = \lambda/2$.
- If T is finite, then there is a unique valuation f, and $f(\phi)$ is Bob's payoff under optimal play.

A *partial valuation* is a function that takes arbitrary values in $[-\lambda/2, \lambda/2]$ on the vertices k steps away from the root, and which is defined recursively on all vertices closer to the root by (10). For each k, we define the two partial valuations f_A^k and f_B^k to be those that predict values most in favor of Alice and Bob respectively. More precisely, f_A^k is obtained by assigning values $\lambda/2$ at distance k from the root if k is odd and values $-\lambda/2$ if k is even, and f_B^k is obtained by assigning values the opposite way.

There is a simple interpretation of f_A^k and f_B^k. Given a realization of T, f_A^k and f_B^k are the bounds we get on any valuation if we examine T only to depth at most k. This is because if the rest of T is unknown to us, then from the perspective of a certain player, say Alice, the best possible scenario for the rest of T is that if Bob is to move from a vertex at distance k, then that vertex has no children, while if Alice is about to move, then it has an edge of zero length to a vertex that in turn has no children.

PROPOSITION 2.4. *For every realization of T, there exists a valuation.*

Proof. As k increases, the values $f_B^k(v)$ form a monotone sequence at each vertex v (decreasing at even levels, increasing at odd levels). Therefore there is a pointwise limit

$$f_B(v) = \lim_{k\to\infty} f_B^k(v),$$

and it is easily verified that f_B is a valuation. □

Informally, the valuation f_B can be thought of as the "optimistic" valuation from Bob's point of view, since it predicts the largest payoff for Bob which is consistent with the tree T. Similarly there is a limit valuation f_A of f_A^k which is "optimistic" from Alice's point of view. We can order the valuations from Bob's point of view by saying that $f_1 \leq f_2$ if whenever v is at even distance from the root, $f_1(v) \leq f_2(v)$, and whenever v is at odd distance from the root, $f_1(v) \geq f_2(v)$. In this ordering f_A is the unique minimal element and f_B is the unique maximal element. (In fact this ordering is a lattice.)

We are aiming to show that almost surely $f_A = f_B$, or equivalently, there is only one valuation. This holds trivially in the range $\lambda \leq 1$, since T is almost surely finite. For $\lambda > 1$, T is infinite with positive probability, and the scenario that we wish to exclude is that at some value of λ there occurs a "symmetry breaking" after which f_A is distinct from f_B with positive probability.

2.6. *The branching of near-optimal play.* For the moment we take f_B as our default valuation. This defines a strategy in an obvious way: From a vertex v, terminate if $f_B(v) = \lambda/2$, and otherwise move to the child v_i for which $f_B(v) = l_i - f_B(v_i)$. There seems to be the possibility of a tie in which several move options would be consistent with f_B, but f_B has the property that $f_B(v_i)$ depends only on $T(v_i)$. Therefore $l_i - f_B(v_i)$ has continuous distribution and is independent of $l_j - f_B(v_j)$ for $i \neq j$. It follows that the probability of a tie between move options is zero.

Let $\delta > 0$. We say that a move from v to v_i is *optimal* (with respect to f_B) if $l_i - f_B(v_i) = f_B(v)$ and δ*-reasonable* if $l_i - f_B(v_i) \leq f_B(v) + \delta$. Let R be the subtree of T formed by all paths from the root consisting of δ-reasonable moves by Alice and optimal moves by Bob. Let R_k be the set of vertices of R at distance k from the root.

PROPOSITION 2.5. *If δ is sufficiently small, then R is almost surely finite.*

In order to prove this proposition, we introduce the quantity

$$N_k = \#\{v \in R_k : f_B(v) < \lambda/2\} + \frac{1}{2} \cdot \#\{v \in R_k : f_B(v) = \lambda/2\}, \tag{11}$$

which essentially counts the vertices in R_k but gives weight $1/2$ to the vertices v where $f_B(v) = \lambda/2$. The proof of Proposition 2.5 consists in showing that for sufficiently small δ, $EN_k \to 0$ as $k \to \infty$.

In the following, we will speak of an "arbitrary" vertex v of T and exploit the fact that $T(v)$ is equal in distribution to T itself. This seems clear from the way T is constructed, but the problem is that since v owes its very existence to the process that generates T, there is no way to choose v independently of T except if v is the root of T. Therefore, speaking of an arbitrary vertex v of T always technically involves conditioning on T. We simply think of T as constructed from the root and downwards so that conditioning on v being in T never involves the structure of $T(v)$. Now let v be an arbitrary vertex of T. Denote its children by v_i and the corresponding edge-lengths by l_i. By the *l-f-square* we mean the square $[0, \lambda] \times [-\lambda/2, \lambda/2]$.

LEMMA 2.6. *The points $(l_i, f_B(v_i))$ constitute a two-dimensional inhomogeneous Poisson point process on the l-f-square.*

Proof. The sequence of edge lengths l_i is a Poisson point process on $[0, \lambda]$ with rate dl^{d-1}. Since $f_B(v_i)$ depends only on $T(v_i)$, the $f_B(v_i)$'s are independent of each other and of the l_i's. This makes the joint process of pairs $(l_i, f_B(v_i))$ a Poisson process whose intensity measure is a product measure of the density dl^{d-1} of l_i on $[0, \lambda]$ and the probability measure of f_B on $[-\lambda/2, \lambda/2]$. ☐

We let μ_v be the intensity measure on the l-f-square associated with the Poisson process of pairs $(l_i, f_B(v_i))$. What we are trying to prove will imply that μ_v is the same for all v, but from what we have established so far it is conceivable that μ_v depends on whether v is at even or odd distance from the root.

We will need the following properties of the measure μ_v:

- The line $f = \lambda/2$ has positive measure; more precisely,

$$\mu_v(f = \lambda/2) = P(f_B(v_i) = \lambda/2) \geq P(v \text{ is a leaf in } T) = \exp(-\lambda^d). \tag{12}$$

- If $-\lambda/2 \leq a \leq b \leq \lambda/2$, then

$$\mu_v(l - f \in [a, b]) \leq \mu_v(a + \lambda/2 \leq l \leq b + \lambda/2). \tag{13}$$

Proof. Since $d \geq 1$, the density dl^{d-1} of l is increasing. Conditioning on f, we therefore have $P(a + f \leq l \leq b + f) \leq P(a + \lambda/2 \leq l \leq b + \lambda/2)$. ☐

To bound EN_{k+1} in terms of EN_k we bound the expected number of moves in R from a vertex $v \in R_k$ in four cases, depending on whether Alice or Bob is about to move and conditioning either on $f_B(v) < \lambda/2$ or on $f_B(v) = \lambda/2$. We first consider the case that Alice is about to move from a vertex $v \in R_k$, where thus k is even. Suppose first that $f_B(v) < \lambda/2$. This means that Alice's optimal move is given by a point $(l_i, f_B(v_i))$ above the diagonal $l - f = \lambda/2$ in the l-f-square. If we condition on $f_B(v_i) \in [a, b]$ for some a, b such that $-\lambda/2 \leq a \leq b < \lambda/2$, then

$$\begin{aligned} P(f_B(v_i) = \lambda/2) &= \frac{\mu_v(f = \lambda/2 \,\&\, a + \lambda/2 \leq l \leq b + \lambda/2)}{\mu_v(l - f \in [a, b])} \\ &\geq \frac{\mu_v(f = \lambda/2 \,\&\, a + \lambda/2 \leq l \leq b + \lambda/2)}{\mu_v(a + \lambda/2 \leq l \leq b + \lambda/2)} \qquad \text{(by (13))} \\ &= \mu_v(f = \lambda/2) \qquad \text{(since } \mu_v \text{ is a product measure)} \\ &\geq \exp(-\lambda^d) \qquad \text{(by (12))}. \end{aligned} \tag{14}$$

Conditioning on $v \in R_k$ and $f_B(v) < \lambda/2$, we therefore have $P(f_B(v_i) = \lambda/2) \geq \exp(-\lambda^d)$. Therefore with conditional probability at least $\exp(-\lambda^d)$, Alice's optimal move from v will go to a vertex v_i that contributes to EN_{k+1} by only $1/2$. Hence, still conditioning on $v \in R_k$ and $f_B(v) < \lambda/2$, Alice's optimal move from v contributes to EN_{k+1} by at most $1 - \exp(-\lambda^d)/2$.

The expected number of nonoptimal δ-reasonable moves is at most

$$\delta \cdot d\lambda^{d-1} = o(1).$$

By $o(1)$ we mean a term that can be made as small as we please by making δ small. Hence the expected contribution to N_{k+1} when Alice moves from a vertex v such that $f_B(v) < \lambda/2$ is at most

$$1 - \frac{1}{2}\exp(-\lambda^d) + o(1).$$

Consider now the case where Alice moves from a vertex $v \in R_k$ with $f_B(v) = \lambda/2$. Then there is no optimal move (the optimal decision is to terminate), and again the expected number of δ-reasonable moves is at most $\delta \cdot d\lambda^{d-1} = o(1)$. It follows that

$$\frac{EN_{k+1}}{EN_k} \leq \max\left(1 - \frac{1}{2}\exp(-\lambda^d) + o(1), \frac{o(1)}{1/2}\right) \leq 1 - \frac{1}{2}\exp(-\lambda^d) + o(1).$$

When Bob moves, there is no optimal move if $f_B(v) = \lambda/2$ and at most one if $f_B(v) < \lambda/2$. Therefore the "growth factor" for N_k over a pair of moves, one by Alice and one by Bob, satisfies

$$\frac{EN_{k+2}}{EN_k} \leq 1 - \frac{1}{2}\exp(-\lambda^d) + o(1) < 1$$

uniformly in k if δ is sufficiently small. It follows that $EN_k \to 0$ as $k \to \infty$, and this completes the proof of Proposition 2.5.

The upper bounds on the expected contributions to N_{k+1} when moving from a vertex $v \in R_k$ are summarized in the following table:

Player to move	Vertex	Contribution to EN_{k+1}
Alice	$f_B(v) < \lambda/2$	$1 - 1/2 \cdot \exp(-\lambda^d) + o(1)$
	$f_B(v) = \lambda/2$	$o(1)$
Bob	$f_B(v) < \lambda/2$	1
	$f_B(v) = \lambda/2$	0

LEMMA 2.7. *For sufficiently small δ, there is almost surely no infinite path starting anywhere in T and consisting of optimal moves by Bob and δ-reasonable moves by Alice.*

Proof. If such a path started from the root, it would be a subset of R, and R is almost surely finite. This event therefore has probability zero, and since T is countable, it follows that the probability of such a path anywhere in T is also zero. ☐

PROPOSITION 2.8. *There is almost surely only one valuation.*

Proof. It suffices to show that almost surely $f_A(\phi) = f_B(\phi)$. Suppose therefore that $f_A(\phi) < f_B(\phi)$. We want to demonstrate that some event of probability zero is a consequence of this assumption. We therefore let Alice play according to f_A and Bob according to f_B. Informally we can think of this as “optimistic” play, both players being convinced that a reward of $\lambda/2$ awaits them at infinity.

Alice’s strategy guarantees that if the game terminates, Bob’s payoff is at most $f_A(\phi)$. On the other hand Bob’s strategy guarantees that his payoff if the game terminates is at least $f_B(\phi)$. Therefore strict inequality $f_A(\phi) < f_B(\phi)$ implies that play has to continue forever.

Let $u_0, u_1, u_2, \dots$ be the path that the game follows (in particular $u_0 = \phi$), and let $l_{01}, l_{12}, \dots$ be the lengths of the edges of this path. We have

$$f_B(u_0) \leq l_{01} - f_B(u_1) = l_{01} - l_{12} + f_B(u_2) \leq l_{01} - l_{12} + l_{23} - f_B(u_3) = \dots,$$

and similarly

$$f_A(u_0) = l_{01} - f_A(u_1) \leq l_{01} - l_{12} + f_A(u_2) = l_{01} - l_{12} + l_{23} - f_A(u_3) \leq \dots.$$

It follows that the absolute difference between f_A and f_B along the path of the game is nondecreasing (with the sign of the difference alternating):

$$f_B(u_0) - f_A(u_0) \leq f_A(u_1) - f_B(u_1) \leq f_B(u_2) - f_A(u_2) \leq \dots.$$

On the other hand the sequence $|f_A(u_i) - f_B(u_i)|$ is obviously bounded since both f_A and f_B take values in $[-\lambda/2, \lambda/2]$. Fix $\delta > 0$ sufficiently small for the conclusion of Lemma 2.7 to hold. It follows that there is some k such that

$$|f_A(u_k) - f_B(u_k)| \geq \lim_{i\to\infty} |f_A(u_i) - f_B(u_i)| - \delta. \tag{15}$$

When Alice moves from some u_i with $i \geq k$, then i is even, and by (15),

$$f_B(u_i) - f_A(u_i) \geq f_A(u_{i+1}) - f_B(u_{i+1}) - \delta.$$

Rearranging this inequality and using the fact that by the definition of Alice’s strategy, $f_A(u_{i+1}) = l_{i,i+1} - f_A(u_i)$, we obtain

$$f_B(u_i) + \delta \geq l_{i,i+1} - f_B(u_{i+1}).$$

This means that Alice’s move from u_i to u_{i+1} is δ-reasonable with respect to f_B. Hence the path $u_k, u_{k+1}, \dots$ has the property that with respect to f_B, Bob’s moves are optimal and Alice’s moves are δ-reasonable. By Lemma 2.7, the existence of such a path is an event of probability zero. ☐

We need no longer distinguish between f_A and f_B, and we denote the almost surely unique valuation by f. Now recall the partial valuations f_A^k and f_B^k.

Proposition 2.9. *$E\left(f_B^k(\phi) - f_A^k(\phi)\right) \to 0$ as $k \to \infty$.*

Proof. We have established that almost surely there is only one valuation. This means that almost surely, $f_B^k(\phi) - f_A^k(\phi) \to 0$ monotonely as $k \to \infty$. The statement now follows from the monotone convergence theorem. ☐

2.7. *Interpretation in terms of integral equations.* We want to obtain, to the extent possible, an analytical characterization of the distributions of $f_A^k(\phi)$, $f_B^k(\phi)$, and their common limit $f(\phi)$. We have

$$f_A^{k+1}(\phi) = \min(\lambda/2, l_i - f_A^{k+1}(v_i)), \tag{16}$$

where v_i ranges over the children of the root. Notice that

$$f_A^{k+1}(v_i) \overset{\mathrm{d}}{=} f_B^k(\phi). \tag{17}$$

Clearly the same holds with the roles of Alice and Bob interchanged.

Suppose now that we describe the distribution of $f_B^k(\phi)$ by the function

$$G_k(x) = P(f_B^k(\phi) \geq x),$$

and similarly

$$F_{k+1}(x) = P(f_A^{k+1}(\phi) \geq x).$$

Then for $-\lambda/2 \leq x \leq \lambda/2$, $F_{k+1}(x)$ is the probability that there is no event in the inhomogeneous Poisson process of v_i such that $l_i - f_A^{k+1}(v_i) < x$, or equivalently, that there is no l_i such that $f_A^{k+1}(v_i) > l_i - x$. Here it does not matter whether the inequality is strict or not, so for given x and l_i, in view of (17),

$$P\left(f_A^{k+1}(v_i) > l_i - x\right) = G_k(l_i - x).$$

The sequence of l_i such that $f_A^{k+1}(v_i) > l_i - x$ is therefore the set of points in a thinned Poisson point process of rate $dl^{d-1}G_k(l-x)$, and it follows that

$$F_{k+1}(x) = \exp\left(-d\int_0^{\lambda/2+x} l^{d-1}G_k(l-x)\,\mathrm{d}l\right).$$

Therefore we define an operator V_λ on functions on the interval $[-\lambda/2, \lambda/2]$ by

$$(V_\lambda F)(x) = \exp\left(-d\int_0^{\lambda/2+x} l^{d-1}F(l-x)\,\mathrm{d}l\right). \tag{18}$$

PROPOSITION 2.10. *The operator V_λ has a unique fixed point F, and*

$$F(x) = P(f(\phi) \geq x). \tag{19}$$

Proof. We have $F_{k+1} = V_\lambda(G_k)$, and by reversing the roles of Alice and Bob, $G_{k+1} = V_\lambda(F_k)$. The distributions of $f_A^k(\phi)$ and $f_B^k(\phi)$ are thus obtained by starting from $F_0 = 0$ and $G_0 = 1$ (on the interval $[-\lambda/2, \lambda/2]$) and iterating the operator V_λ. But since $G_1 = G_0$, it follows inductively that $F_2 = F_1$, $G_3 = G_2$ and so on. Therefore, in reality, there is only one sequence of functions, obtained by iterating V_λ starting from the zero function.

The operator V_λ is decreasing in the sense that if $F(x) \leq G(x)$ for every x, then $(V_\lambda F)(x) \geq (V_\lambda G)(x)$ for every x. It follows that if we start from the function that is identically zero (or identically 1) and iterate, the sequence of functions must either converge to a fixed point or approach an attractor of period 2. By Proposition 2.8 the sequence converges to a fixed point F that satisfies (19). Actually it follows that if we start from any real integrable function, then after two iterations we have a function which takes values in $[0, 1]$, in other words lies between F_0 and G_0. Therefore the subsequent iterates will be squeezed between F_k and G_k and thus converge to the same fixed point. In particular, V_λ has only one fixed point. □

The following observation of the fixed point functions F will be useful in Section 3.3.

Proposition 2.11. *For fixed d, as $\lambda \to \infty$, $F(\lambda/2) \to 0$.*

To avoid confusion, notice that here F depends on λ.

Proof. It follows from (18) and the fact that F is decreasing that

$$F(\lambda/2) = \exp\left(-d\int_0^\lambda l^{d-1}F(l-\lambda/2)\,\mathrm{d}l\right) \leq \exp\left(-d\int_0^\lambda l^{d-1}F(\lambda/2)\,\mathrm{d}l\right) = \exp(-\lambda^d F(\lambda/2)). \tag{20}$$

By (20) in turn,

$$\lambda \leq \left(\frac{-\log F(\lambda/2)}{F(\lambda/2)}\right)^{1/d},$$

from which the claim follows. □

The connection from optimization problems, via local convergence, to fixed point equations for integral operators is clear in the physics literature from the 1980's (see [27] and its references) despite the lack of explicit convergence results. These "recursive distributional equations" have been studied extensively in [3], [4], and a number of examples where symmetry breaking occurs for some values of a parameter corresponding to our λ were studied in [7], [15], [43].

The similarity of (18) to the Mézard-Parisi integral equation (7) is clearly visible. Naturally we may define an operator V_∞ by

$$(V_\infty F)(x) = \exp\left(-d\int_0^\infty l^{d-1}F(l-x)\,\mathrm{d}l\right).$$

It seems clear, both from numerical evidence and in view of the results we have established, that as $\lambda \to \infty$, the fixed points F of V_λ should converge uniformly to a limit function which is a unique fixed point to V_∞, in other words a unique solution to the Mézard-Parisi equation (7). We certainly believe that a more detailed analysis will show this to be the case (possibly the ideas of [37] can be extended to $d > 1$), but we leave it as an open conjecture since it is not necessary for our proof of Theorem 1.1. Moreover, the natural way to obtain numerical results from (7) is to approximate $F(x)$ by 1 for large negative x and by 0 for large positive x. Therefore, in practice, the numerical results based on (7) reduce to solving the equation $V_\lambda(F) = F$ on a bounded interval.

3. Applications to minimum matching problems

3.1. *The density of the minimum diluted matching.* We now return to the (rescaled) mean field model K_n on n vertices. Suppose that λ and $d \geq 1$ are fixed, and let the random variable q_n be the proportion of vertices that are not matched (for which we pay the penalty of $\lambda/2$) in the optimum diluted matching. Again we let F be the fixed point of V_λ.

PROPOSITION 3.1. *As* $n \to \infty$, $q_n \xrightarrow{\text{p}} F(\lambda/2)$.

Proof. We show that $Eq_n \to F(\lambda/2)$ and $\operatorname{var}(q_n) \to 0$. Let v be an arbitrary vertex of K_n, and notice that $Eq_n = P(v \text{ is not matched})$. Let k be a positive integer. We now compare K_n (rooted at v) to $T(k,\lambda)$. It follows from local convergence that we can couple K_n to $T(k,\lambda)$ in such a way that asymptotically almost surely as $n \to \infty$, the game theoretical value $f(K_n, v)$ of Exploration on K_n starting at v satisfies

$$f_A^k(\phi) \leq f(K_n, v) \leq f_B^k(\phi).$$

Therefore as $n \to \infty$,

$$\begin{aligned} P(f_A^k(\phi) = \lambda/2) - o(1) &\leq P(f(K_n, v) = \lambda/2) \\ &\leq P(f_B^k(\phi) = \lambda/2) + o(1). \end{aligned} \tag{21}$$

But $f(K_n, v) = \lambda/2$ is equivalent to v not being matched, and both sides of (21) converge to $F(\lambda/2)$ as $k \to \infty$. This shows that $Eq_n \to F(\lambda/2)$.

To bound the variance of q_n we simply take two vertices v_1 and v_2 of K_n and estimate the probability that neither is matched. To do this we apply Proposition 2.2 with $m = 2$. By local convergence we can approximate the (k,λ)-truncation of K_n rooted at v_1 and v_2 by $T^2(k,\lambda)$, which can be regarded as two independent copies of $T(k,\lambda)$. It follows that the probability that neither of v_1 and v_2 is matched converges to $F(\lambda/2)^2$, which means that $\operatorname{var}(q_n) \to 0$. □

3.2. *The length of the minimum diluted matching.* We wish to find the normalized limit cost of the minimum diluted matching. This cost splits naturally into the length of the participating edges and the cost of the penalties for the unmatched vertices. The penalties depend only on the number of unmatched vertices, which was discussed in the previous section, and therefore we concentrate on the participating edges. Still working in the rescaled model, we let

$M^\star_\lambda(K_n)$ be the total length of the edges participating in the optimum diluted matching (while M_λ, as defined earlier, is the total cost including penalties).

THEOREM 3.2. *For each λ and $d \geq 1$, there is a number $\beta_\lambda(d)$ such that*

$$\frac{M^\star_\lambda(K_n)}{n} \xrightarrow{\mathrm{p}} \beta_\lambda(d). \tag{22}$$

To prove Theorem 3.2, we use the same method as in [2], [22] to calculate the expected contribution of an edge to the optimum diluted matching. We consider the edge between two arbitrary vertices u and v. First we want to use Proposition 2.2 to conclude that the neighborhoods of u and v look like two independent (k, λ)-trees except for the edge between u and v.

First let k be an arbitrary integer that will eventually tend to infinity. We now apply Proposition 2.2 with $m = 2$ and u and v being the roots. The conclusion is that if we assume that the (k, λ)-neighborhoods of u and v behave like independent (k, λ)-trees, the error in any probability will tend to zero as $n \to \infty$.

Now we want to condition on the length of the edge between u and v without interfering with this conclusion. Technically what we do is to rerandomize the length of the edge between u and v; that is, we replace it by a new random variable of the same distribution. This obviously does not change the distance model in distribution; it is only an artificial trick to be able to apply Proposition 2.2. We now want to estimate the expected contribution of the new edge between u and v to the total length of the edges in the optimum diluted matching.

Recall that the rescaled edge-lengths are $n^{1/d} l'_{i,j}$, where the density function ρ of $l'_{i,j}$ satisfies $\rho(l) \sim dl^{d-1}$ as $l \to 0$. We let

$$\rho_n(l) = n^{-1/d} \rho(n^{-1/d} l) \tag{23}$$

be the density function of the rescaled length $n^{1/d} l'_{i,j}$. Since we are interested in the diluted matching problem, we are only concerned with the behavior of ρ_n on the interval $[0, \lambda]$.

LEMMA 3.3. *As $n \to \infty$, $n\rho_n(l) \to dl^{d-1}$ uniformly on the interval $[0, \lambda]$.*

Proof. It suffices to show that

$$\frac{n\rho_n(l)}{dl^{d-1}} \to 1$$

uniformly on $0 \leq l \leq \lambda$. By (23), we have

$$\frac{n\rho_n(l)}{dl^{d-1}} = \frac{n^{1-1/d} \rho(n^{-1/d} l)}{dl^{d-1}},$$

which in view of (5) converges to 1 as $n \to \infty$. The convergence is uniform since as $n \to \infty$, $n^{-1/d} l \to 0$ uniformly on $0 \le l \le \lambda$. □

The expectation of $M_\lambda^\star(K_n)$ is the total number of edges in the graph times the expected contribution to $M_\lambda^\star(K_n)$ from a single edge. Let $h(l)$ be the probability that an edge participates in the optimum diluted matching conditioning on its length being l. Then

$$EM_\lambda^\star(K_n) = \binom{n}{2} \cdot \int_0^\lambda l \rho_n(l) \cdot h(l)\,\mathrm{d}l. \tag{24}$$

By Lemma 3.3 we can replace $\rho_n(l)$ by dl^{d-1}/n. Normalizing to obtain a quantity of order 1, we get

$$\frac{EM_\lambda^\star(K_n)}{n} = \frac{d}{2} \cdot \int_0^\lambda l^d \cdot h(l)\,\mathrm{d}l + o(1). \tag{25}$$

Now we regard the (k, λ)-neighborhoods of u and v as two independent trees distributed like $T(k, \lambda)$ whose roots u and v are connected by an edge of length l. By Proposition 2.2 this will only introduce small errors in probabilities as $n \to \infty$.

We denote these trees by T_u and T_v, and define the partial valuations f_A^k and f_B^k as in Section 2.5 on each of these trees. Suppose that we play Exploration starting at u (or v) and with the edge between u and v removed. Then the true payoff to the second player will be between $f_A^k(u)$ and $f_B^k(u)$, since it can in principle be calculated by assigning values to each vertex w at distance k equal to the payoff under optimal subsequent play given that the first k moves are the unique path to w in T_u (T_v).

Suppose now that we play Exploration starting at u on the graph including the edge of length l between u and v. The first player's payoff if moving to v satisfies

$$-l + f_A^k(v) \le \text{payoff on moving } (u \to v) \le -l + f_B^k(v). \tag{26}$$

The reason that we are talking about the "first player" is that if moving to v, the first player gets the role of "Bob" in T_v. If on the other hand the first player moves to another vertex (or terminates), the game will not leave T_u until at least k moves have been played. Then the first player will have the role of "Alice" in T_u, and therefore

$$-f_B^k(u) \le \text{payoff on } \textit{not} \text{ moving } (u \to v) \le -f_A^k(u). \tag{27}$$

If the right-hand side of (27) is smaller than the left-hand side of (26), then moving to v in the first move is clearly optimal. Conversely if the right-hand side of (26) is smaller than the left-hand side of (27), then moving to v

is not optimal. The move to v being optimal is equivalent to the edge (u, v) of lenght l participating in the optimum diluted matching. Hence

$$P\left(f_A^k(u) + f_A^k(v) \geq l\right) \leq h(l) \leq P\left(f_B^k(u) + f_B^k(v) \geq l\right).$$

If $k \to \infty$ with n, then by Proposition 2.9 we can replace $f_A^k(u)$ and $f_B^k(u)$ by a limit $f(u)$ and similarly replace $f_A^k(v)$ and $f_B^k(v)$ by $f(v)$. Hence (25) is equal to

$$\frac{d}{2} \cdot \int_0^\lambda l^d \cdot P\left(l \leq f(u) + f(v)\right) \mathrm{d}l + o(1)$$

for $f(u)$ and $f(v)$ independent and satisfying $P(f \geq x) = F(x)$, where F is the fixed point of V_λ. By partial integration it follows that

$$\frac{EM_\lambda^\star(K_n)}{n} \to \frac{d^2}{2} \cdot \iint\limits_{\substack{-\lambda/2 < x, y < \lambda/2 \\ x+y \geq 0}} (x+y)^{d-1} F(x) F(y) \, \mathrm{d}x \, \mathrm{d}y. \tag{28}$$

Therefore $\beta_\lambda(d)$ is equal to the right-hand side of (28). To see that (28) can be strengthened to convergence in probability as stated in (22) we again apply Proposition 2.2, this time with $m = 4$. It follows that the expected contribution from an arbitrary *pair* of edges to $M_\lambda^\star(K_n)^2$ is asymptotically the same as the square of the expected contribution of one edge, and that therefore $\operatorname{var}(M_\lambda^\star(K_n)) = o(n^2)$. This concludes the proof of Theorem 3.2.

3.3. *Perfect matching.* Assuming that n is even, we study the length of the minimum perfect matching in K_n. Naturally we expect perfect matching to correspond to infinite λ, and the remaining step in the proof of Theorem 1.1 essentially amounts to showing that we can interchange the order in which n and λ go to infinity. For the bipartite graph and $d = 1$, this was proved in [1]. Without claims of originality we give a self-contained proof based on expander properties of random graphs along the lines of [14]. I thank David Aldous for pointing out that the method applies here.

Recall that $\beta_\lambda(d)$ is defined as the right-hand side of (28). In view of (22), $\beta_\lambda(d)$ is increasing in λ. We define

$$\beta(d) = \lim_{\lambda \to \infty} \beta_\lambda(d).$$

We have not yet established that $\beta(d)$ is finite, but this will follow from Proposition 3.4 below. Theorem 1.1 states that if ρ satisfies (5), then as $n \to \infty$ through even values,

$$\frac{M(K_n[\rho])}{n^{1-1/d}} \xrightarrow{\mathrm{p}} \beta(d). \tag{29}$$

Since for every λ the cost of the λ-diluted matching problem is a lower bound on the cost of a perfect matching, Theorem 3.2 already establishes the required lower bound on $M(K_n[\rho])$, and what remains in order to prove (29) is to show that for every $\varepsilon > 0$, as $n \to \infty$,

$$P\left(\frac{M(K_n[\rho])}{n^{1-1/d}} \le \beta(d) + \varepsilon\right) \to 1. \tag{30}$$

It turns out that certain properties of the exponential distribution are convenient in the proof. Therefore we start by establishing (30) for a special distribution of edge-lengths based on the exponential distribution. It is obtained by taking edge-lengths $l = X^{1/d}$, where X is a mean 1 exponential variable. Equivalently, the density function is given by

$$\xi(l) = dl^{d-1}\exp(-l^d). \tag{31}$$

We establish the following slightly stronger result stating that a matching yielding (30) can be constructed using only short edges. The generalization in Section 3.4 to an arbitrary density function ρ satisfying (5) will then be a relatively simple matter.

PROPOSITION 3.4. *For every $\varepsilon > 0$ and $x > 0$, there is a λ such that asymptotically almost surely as $n \to \infty$ through even values, there exists a perfect matching in $K_n[\xi]$ that has total length at most $n^{1-1/d}(\beta_\lambda(d) + \varepsilon)$ and contains no edge longer than x.*

The rest of Section 3.3 is devoted to the proof of Proposition 3.4. We will work with a random distance model involving an infinite sequence of edges between each pair of vertices. For each pair of vertices, consider a rate 1 Poisson point process X_i $(i \ge 1)$ on the positive real numbers (independent processes for each pair of vertices), and let the sequence of edges have lengths $X_i^{1/d}$.

Since X_1 has exponential distribution, the length of the shortest edge between any two vertices is distributed according to ξ. Therefore it suffices to establish Proposition 3.4 for this "Poisson model."

We randomly color every edge red or green, where the probability of red is $1-p$ and the probability of green is p. Notice that for each pair of vertices, the two color classes of edges have lengths given by independent Poisson processes. The red edges have lengths $Y_i^{1/d}$ where Y_i are the points of a rate $1-p$ Poisson point process, and similarly the lengths of the green edges are $Z_i^{1/d}$ where Z_i are given by a process of rate p.

Let M_{Red} be the minimum λ-diluted matching on the red edges, where λ will be chosen later depending on ε. The length $Y_1^{1/d}$ of the first red edge between two vertices is equal in distribution to $1/(1-p)^{1/d}$ times the first edge without regard to color. Therefore the λ-diluted matching problem on the red edges is equal in distribution to a λ'-diluted matching problem in the original model scaled up by a factor $1/(1-p)^{1/d}$, where $\lambda' = \lambda \cdot (1-p)^{1/d}$. Hence if p is chosen so that

$$\frac{\beta(d)}{(1-p)^{1/d}} < \beta(d) + \varepsilon/2, \tag{32}$$

then since $\beta_{\lambda'}(d) < \beta(d)$ regardless of λ, the total length of M_{Red} is at most $n^{1-1/d}(\beta(d) + \varepsilon/2)$ asymptotically almost surely as $n \to \infty$.

Our aim is to show that we can combine a subset of the edges of M_{Red} with a set of relatively few green edges to obtain a perfect matching where the total length of the green edges is at most $\varepsilon/2 \cdot n^{1-1/d}$. Before looking at the lengths of the green edges, we choose arbitrarily a partition of the vertices into two sets A and B of size $m = n/2$ such that every edge of M_{Red} connects a vertex of A to a vertex of B. Then we consider the green edges that connect A to B and give each of them a random orientation by independent coin flips. We let D be the set consisting of the 13 cheapest green edges directed from each vertex to the opposite side of the partition.

If S is a set of vertices, we let S' denote the set of vertices that are connected to some vertex in S by an edge in D.

Lemma 3.5. *Asymptotically almost surely as $n \to \infty$, D has the following expander property: If S is a set of vertices from one side of the partition and $1 \le |S| \le m/3$, then $|S'| > 2|S|$.*

Proof. If this condition is violated, then there is a positive integer $s \le m/3$ and a set of s vertices on one side of the partition such that all its $13s$ edges go into a certain set of $2s$ vertices on the other side. The probability that this happens is at most

$$2 \cdot \sum_{1 \le s \le m/3} \binom{m}{s}\binom{m}{2s}\left(\frac{2s}{m}\right)^{13s}. \tag{33}$$

Using the standard inequality

$$\binom{m}{k} \le \left(\frac{me}{k}\right)^k,$$

we find that (33) is at most

$$2 \cdot \sum_{1 \le s \le m/3} e^{3s} 2^{11s} \left(\frac{s}{m}\right)^{10s}.$$

By log-convexity of the summand, the maximum of the terms is attained either by the first term or by the last one. The first term ($s = 1$) is $O(1/m^{10})$ and the last one is at most

$$\left(\frac{e^3 2^{11}}{3^{10}}\right)^{m/3} \le 0.9^m.$$

For large m, the first term will dominate, and since there are $O(m) = O(n)$ terms, the failure probability of the expander property is $O(1/n^9)$. □

Having established the high probability expander property, we no longer need the orientations of the edges in D.

LEMMA 3.6. *Suppose D has the expander property of Lemma 3.5. Let M be a matching consisting of $m-k$ edges of arbitrary color connecting vertices in A to vertices in B. Then there is a matching of $m-k+1$ edges in $D \cup M$, of which all but at most $2\log_2(m/k)+O(1)$ belong to M.*

Proof. Let A_0 and B_0 be the sets of vertices in A and B respectively that are not matched by M. Consider paths starting from some vertex in A_0 and using alternatingly edges from D and M. Let A_i be the set of vertices in A that can be reached by such a path of length at most $2i$, in other words, in at most i steps, where each step consists of one edge in D and one in M. As before let S' denote the set of D-neighbors of S.

We have

$$|A_{i+1}| = k + |A_i' - B_0| \geq k + |A_i'| - k = |A_i'| \geq \min\left(\frac{2}{3}m, 2\,|A_i|\right).$$

It follows inductively that $|A_i| \geq \min(\frac{2}{3}m, k2^i)$, and similarly if B_i is the set of vertices that can be reached by an alternating path from B_0 in at most i steps, then $|B_i| \geq \min(\frac{2}{3}m, k2^i)$. Therefore there exists an $i \leq \log_2(m/k)+O(1)$ such that A_i and B_i are connected by an edge of D. This means that there is an alternating path connecting A_0 to B_0 in at most $2\log_2(m/k)+O(1)$ steps.

The required matching is now simply the symmetric difference of this path with M. □

We introduce yet another small positive parameter q to be chosen later as a function of ε.

LEMMA 3.7. *Suppose that the expander property of Lemma 3.5 holds for D. If M is a matching connecting vertices in A to vertices in B and containing at least $(1-q)m$ edges, then there is a perfect matching consisting of edges in $M \cup D$ where all but at most $O(qm\log(1/q))$ edges belong to M.*

Proof. We successively apply Lemma 3.6 for k from $\lfloor qm \rfloor$ down to 1, replacing the original matching M by matchings combining edges in M with edges in D. The total number of new edges needed to obtain a perfect matching is at most

$$2 \cdot \sum_{k=1}^{qm} \left(\log_2\left(\frac{m}{k}\right) + O(1)\right) = O\left(qm\log(1/q)\right). \qquad \square$$

Provided that M_{Red} matches all but at most qn vertices and the expander property of Lemma 3.5 holds for D, we can therefore complete M_{Red} to a perfect matching by combining some edges in M_{Red} with at most $O(qn\log(1/q))$ green edges. (Recall that $m = n/2$.)

The choice of edges in D that we use to complete the matching is done independently of their lengths so that the expected total length of the green edges in the matching is $O(nq\log(1/q))$ times the expected length of an edge in D, which is $O((1/np)^{1/d})$. Therefore assuming the expander property holds, with high probability as $n \to \infty$ the total length of the green edges in the matching is

$$O\left(\frac{nq\log(1/q)}{(np)^{1/d}}\right) \leq K \cdot n^{1-1/d} \cdot \frac{q\log(1/q)}{p^{1/d}}$$

for some constant K.

We now choose q so that

$$K \cdot \frac{q\log(1/q)}{p^{1/d}} \leq \frac{\varepsilon}{2}.$$

In view of (32), this shows that provided M_{Red} contains all but at most qn vertices, we can find the required perfect matching of total length at most $n^{1-1/d}(\beta(d)+\varepsilon)$ with high probability as $n \to \infty$.

Now, once p and q are fixed, it follows from Proposition 3.1 in combination with Proposition 2.11 that by taking λ large enough, with high probability M_{Red} has density at least $1-q$. The matching M_{Red} corresponds to a λ'-diluted matching, but λ' only differs from λ by a constant factor $(1-p)^{1/d}$.

This establishes that with high probability there is a perfect matching of total length at most $n^{1-1/d}(\beta(d)+\varepsilon)$, which is the first claim of Proposition 3.4. What remains is to show that with high probability, the longest edge in the matching we have constructed is $o(1)$ as $n \to \infty$. This property of the matching will be useful in completing the proof of Theorem 1.1.

The lengths of the edges in M_{Red} are bounded by $n^{-1/d}\lambda$, and λ was chosen independently of n, so we need only check the length of the longest green edge. Each green edge that is used is among the 13 cheapest green edges directed from some vertex into the opposite vertex-set in the partition. If we fix a vertex v, then for each vertex u on the other side of the partition, the green edges directed from v to u have lengths $Z_i^{1/d}$, where Z_i are the points of a Poisson process of rate $p/2$. Since there are $n/2$ vertices in the opposite part of v in the partition, the totality of green edges directed from v into the opposite part are given by the d:th roots of a Poisson process of rate $pn/4$. The maximum length of an edge in D is therefore the d:th root of the time it takes until each of n independent Poisson processes of rate $pn/4$ has had at least 13 events. This time is asymptotically almost surely

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

Therefore if we fix ε and thereby p, the perfect matching that we construct will asymptotically almost surely contain no edge longer than

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

where we allow the implied constant to depend on ε. This completes the proof of Proposition 3.4. □

3.4. *Proof of Theorem 1.1.* We have shown that (30) holds in the special case of the density function $\xi(l) = dl^{d-1}\exp(-l^d)$. To complete the proof of Theorem 1.1 we need to generalize this to an arbitrary density function ρ satsifying (5).

Since the probability distributions given by ρ and ξ are both continuous, we can find an order preserving coupling between them. Let us denote such a coupling by $l \mapsto l'$, where l has distribution given by ξ and l' is distributed according to ρ. This coupling has the property that the quotient l'/l will tend to 1 as $l \to 0$.

To obtain with high probability a perfect matching in $K_n[\rho]$ of length at most $n^{1-1/d}(\beta(d)+\varepsilon)$, we first find an x such that

$$\frac{l'}{l} \leq \frac{\beta(d)+\varepsilon}{\beta(d)+\varepsilon/2}$$

whenever $l \leq x$. Then we apply Proposition 3.4 with ε replaced by $\varepsilon/2$. This will give us a perfect matching in $K_n[\xi]$ of total length at most $n^{1-1/d}(\beta(d) + \varepsilon/2)$ and no edge longer than x, which in turn gives us the desired perfect matching in $K_n[\rho]$ through the coupling.

3.5. *The $\pi^2/12$-limit for $d = 1$.* The case $d = 1$ corresponds to the model studied by Aldous in [1], [2]. In our terminology his result (conjectured in [22]) is that $\beta(1) = \pi^2/12$. We briefly show how this follows from our present approach. In [2], [22] the calculations start from equation (7), which has the solution $F(x) = 1/(1+e^x)$. This corresponds to infinite λ, but in our approach we arrive at the equation $V_\lambda(F) = F$ for finite λ, which for $d = 1$ becomes

$$F(x) = \exp\left(-\int_0^{\lambda/2+x} F(l-x)\,\mathrm{d}l\right).$$

Since we know that the equation has a unique solution, it suffices to verify that it is satisfied by

$$F(x) = \frac{1+q}{1+e^{(1+q)x}},$$

where q is given by

$$\lambda = \frac{-2\log q}{1+q}.$$

Through (28) it can then be verified that the limit total length of the edges in the diluted matching for finite λ is given by

$$\beta_\lambda(1) = \int_q^1 \frac{-\log t}{1+t} dt, \tag{34}$$

from which the $\pi^2/12$-result is obtained by putting $q = 0$. The limit (34) for the minimum density $1-q$ matching also follows in a completely different way from the results of [42]. A more streamlined derivation along the present lines is given in [33].

3.6. *Concluding remarks.* It is easy to generalize Theorem 1.1 to length distributions satisfying (2) but not necessarily having a density function. This is simply because each distribution satisfying (2) is stochastically dominated by such a distribution that has a density function, and conversely dominates some other such distribution.

More interesting is to generalize the local convergence result to other types of graphs. The method of proof indicated for Proposition 2.2 works equally well for the complete bipartite graph $K_{n,n}$, provided the rescaling is done correctly. More generally it works for dense regular graphs, meaning graphs on n vertices where the degree of each vertex is $\alpha n + O(1)$ for some fixed $\alpha > 0$. Hence all our results on the diluted matching problem will generalize to such graphs. It seems that the completion to a perfect matching can be achieved by a modification of the arguments of Section 3.3 as long as $\alpha > 1/2$ but clearly cannot work if $\alpha < 1/2$.

There are several other optimization problems that yield to similar analysis. Analogs of Theorem 1.1 for the TSP and some other problems will be presented in forthcoming papers. There is a large family of problems, including matching and TSP, for which the case $d = 1$ can be analyzed with the method of [42]. It would be interesting to know if, for some general reason, replica symmetry must hold for all problems amenable to such methods.

References

[1] D. J. Aldous, Asymptotics in the random assignment problem, *Probab. Theory Related Fields* **93** (1992), 507–534. MR 1183889. Zbl 0767.60006. http://dx.doi.org/10.1007/BF01192719.

[2] ———, The $\zeta(2)$ limit in the random assignment problem, *Random Structures Algorithms* **18** (2001), 381–418. MR 1839499. Zbl 0993.60018. http://dx.doi.org/10.1002/rsa.1015.

[3] D. J. Aldous and A. Bandyopadhyay, A survey of max-type recursive distributional equations, *Ann. Appl. Probab.* **15** (2005), 1047–1110. MR 2134098. Zbl 1105.60012. http://dx.doi.org/10.1214/105051605000000142.

[4] D. J. Aldous and J. M. Steele, The objective method: probabilistic combinatorial optimization and local weak convergence, in *Probability on Discrete Structures*, *Encyclopaedia Math. Sci.* **110**, Springer-Verlag, New York, 2004, pp. 1–72. MR 2023650. Zbl 1037.60008.

[5] S. Bacci and E. N. Miranda, The traveling salesman problem and its analogy with two-dimensional spin glasses, *J. Statist. Phys.* **56** (1989), 547–551. MR 1009516. http://dx.doi.org/10.1007/BF01044452.

[6] A. Bandyopadhyay, Endogeny for the logistic recursive distributional equation, *Z. Anal. Anwend.* **30** (2011), 237–251. MR 2793003. Zbl 1215.60009. http://dx.doi.org/10.4171/ZAA/1433.

[7] A. Bandyopadhyay and D. Gamarnik, Counting without sampling: asymptotics of the log-partition function for certain statistical physics models, *Random Structures Algorithms* **33** (2008), 452–479. MR 2462251. Zbl 1157.82006. http://dx.doi.org/10.1002/rsa.20236.

[8] J. Beardwood, J. H. Halton, and J. M. Hammersley, The shortest path through many points, *Proc. Cambridge Philos. Soc.* **55** (1959), 299–327. MR 0109316. Zbl 0118.35601. http://dx.doi.org/10.1017/S0305004100034095.

[9] S. Boettcher and A. Percus, Nature's way of optimizing, *Artificial Intelligence* **119** (2000), 275–286. Zbl 0949.90075. http://dx.doi.org/10.1016/S0004-3702(00)00007-2.

[10] R. Brunetti, W. Krauth, M. Mézard, and G. Parisi, Extensive numerical simulation of weighted matchings: Total length and distribution of links in the optimal solution, *Europhys. Lett.* **14** (1991), 295–301. http://dx.doi.org/10.1209/0295-5075/14/4/002.

[11] N. J. Cerf, J. Boutet de Monvel, O. Bohigas, O. C. Martin, and A. G. Percus, The random link approximation for the Euclidean traveling salesman problem, *J. de Phys.* **7** (1997), 117–136. http://dx.doi.org/10.1051/jp1:1997129.

[12] E. D. Demaine and R. A. Hearn, Constraint logic: a uniform framework for modeling computation as games, in *Twenty-Third Annual IEEE Conference on Computational Complexity*, IEEE Computer Soc., Los Alamitos, CA, 2008, pp. 149–162. MR 2513497. http://dx.doi.org/10.1109/CCC.2008.35.

[13] A. M. Frieze, On the value of a random minimum spanning tree problem, *Discrete Appl. Math.* **10** (1985), 47–56. MR 0770868. Zbl 0578.05015. http://dx.doi.org/10.1016/0166-218X(85)90058-7.

[14] ———, On random symmetric travelling salesman problems, *Math. Oper. Res.* **29** (2004), 878–890. MR 2104159. Zbl 1082.05517. http://dx.doi.org/10.1287/moor.1040.0105.

[15] D. Gamarnik, T. Nowicki, and G. Swirszcz, Maximum weight independent sets and matchings in sparse random graphs. Exact results using the local weak convergence method, *Random Structures Algorithms* **28** (2006), 76–106. MR 2187483. Zbl 1094.05050. http://dx.doi.org/10.1002/rsa.20072.

[16] J. Houdayer, J. H. Boutet de Monvel, and O. C. Martin, Comparing mean field and Euclidean matching problems, *Eur. Phys. J. B Condens. Matter Phys.* **6** (1998), 383–393. MR 1665617. http://dx.doi.org/10.1007/s100510050565.

[17] S. Janson, T. Łuczak, and A. Rucinski, *Random Graphs*, *Wiley-Intersci. Ser. Discrete Math. Optim.*, Wiley-Interscience, New York, 2000. MR 1782847. Zbl 0968.05003.

[18] S. Kirkpatrick and G. Toulouse, Configuration space analysis of travelling salesman problems, *J. Physique* **46** (1985), 1277–1292. MR 0807242. http://dx.doi.org/10.1051/jphys:019850046080127700.

[19] W. Krauth and M. Mézard, The cavity method and the travelling-salesman problem, *Europhys. Lett.* **8** (1989), 213–218. http://dx.doi.org/10.1209/0295-5075/8/3/002.

[20] S. Linusson and J. Wästlund, A proof of Parisi's conjecture on the random assignment problem, *Probab. Theory Related Fields* **128** (2004), 419–440. MR 2036492. http://dx.doi.org/10.1007/s00440-003-0308-9.

[21] O. C. Martin, M. Mézard, and O. Rivoire, Random multi-index matching problems, *J. Stat. Mech. Theory Exp.* (2005), P09006, 36 pp. MR 2174087. http://dx.doi.org/10.1088/1742-5468/2005/09/P09006.

[22] M. Mézard and G. Parisi, Replicas and optimization, *J. de Phys. Lett.* **46** (1985), 771–778. http://dx.doi.org/10.1051/jphyslet:019850046017077100.

[23] ______, Mean-field equations for the matching and the travelling salesman problems, *Europhys. Lett.* **2** (1986), 913–918. http://dx.doi.org/doi:10.1209/0295-5075/2/12/005.

[24] ______, A replica analysis of the travelling salesman problem, *J. de Phys.* **47** (1986), 1285–1296. http://dx.doi.org/10.1051/jphys:019860047080128500.

[25] ______, On the solution of the random link matching problems, *J. de Phys. Lett.* **48** (1987), 1451–1459. http://dx.doi.org/10.1051/jphys:019870048090145100.

[26] M. Mézard, G. Parisi, and R. Zecchina, Analytic and algorithmic solution of random satisfiability problems, *Science* **297** (2002), 812–815. http://dx.doi.org/10.1126/science.1073287.

[27] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond*, *World Sci. Lecture Notes Phys.* **9**, World Scientific Publishing Co., Teaneck, NJ, 1987. MR 1026102. Zbl 0992.82500.

[28] C. Nair, B. Prabhakar, and M. Sharma, Proofs of the Parisi and Coppersmith-Sorkin random assignment conjectures, *Random Structures Algorithms* **27** (2005), 413–444. MR 2178256. Zbl 1125.90026. http://dx.doi.org/10.1002/rsa.20084.

[29] A. Pagnani, G. Parisi, and M. Ratieville, Near-optimal configurations in mean-field disordered systems, *Phys. Rev. E* **68** (2003), 046706. http://dx.doi.org/10.1103/PhysRevE.68.046706.

[30] G. Parisi, A sequence of approximated solutions to the S-K model for spin glasses, *J. Phys. A* **13** (1980), L115. http://dx.doi.org/10.1088/0305-4470/13/4/009.

[31] G. Parisi, Spin glasses and optimization problems without replicas, in *Le Hasard et la Matière/Chance and Matter* (Les Houches, 1986) (J. Souletie, J. Vannimenus, and R. Stora, eds.), Elsevier Science Publishers B. V., Netherlands, 1987, pp. 525–552. MR 0985792.

[32] G. Parisi and M. Ratiéville, Neighborhood preferences in random matching problems, *Euro. Phys. J. B* **22** (2001), 229–237. http://dx.doi.org/10.1007/PL00011144.

[33] G. Parisi and J. Wästlund, Mean field matching and traveling salesman in pseudo-dimension 1, manuscript.

[34] A. G. Percus, Voyageur de commerce et problèmes stochastiques associés, 1997, Ph.D. thesis, Université Pierre et Marie Curie, Paris.

[35] A. G. Percus and O. C. Martin, The stochastic traveling salesman problem: finite size scaling and the cavity prediction, *J. Statist. Phys.* **94** (1999), 739–758. MR 1694076. Zbl 0963.82035. http://dx.doi.org/10.1023/A:1004570713967.

[36] J. Salez and D. Shah, Belief propagation: an asymptotically optimal algorithm for the random assignment problem, *Math. Oper. Res.* **34** (2009), 468–480. MR 2554069. Zbl 05881643. http://dx.doi.org/10.1287/moor.1090.0380.

[37] ———, *Optimality of Belief Propagation for Random Assignment Problem*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2009, 187–196. Available at http://www.informatik.uni-trier.de/~ley/db/conf/soda/soda2009.html#SalezS09.

[38] D. Sherrington and S. Kirkpatrick, Solvable model of a spin glass, *Phys. Rev. Lett.* **35** (1975), 1792–1796. http://dx.doi.org/10.1103/PhysRevLett.35.1792.

[39] N. Sourlas, Statistical mechanics and the travelling salesman problem, *Europhys. Lett.* **2** (1986), 919–923. http://dx.doi.org/10.1209/0295-5075/2/12/006.

[40] M. Talagrand, The Parisi formula, *Ann. of Math.* **163** (2006), 221–263. MR 2195134. Zbl 1137.82010. http://dx.doi.org/10.4007/annals.2006.163.221.

[41] J. Wästlund, An easy proof of the $\zeta(2)$ limit in the random assignment problem, *Electron. Commun. Probab.* **14** (2009), 261–269. MR 2516261. Zbl 1195.60018. http://dx.doi.org/10.1214/ECP.v14-1475.

[42] ———, The mean field traveling salesman and related problems, *Acta Math.* **204** (2010), 91–150. MR 2600434. Zbl 1231.90370. http://dx.doi.org/10.1007/s11511-010-0046-7.

[43] D. Weitz, Counting independent sets up to the tree threshold, in *STOC'06: Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, ACM, New York, 2006, pp. 140–149. MR 2277139. http://dx.doi.org/10.1145/1132516.1132538.

(Received: July 1, 2009)
(Revised: May 9, 2011)

Chalmers University of Technology and Gothenburg University,
Gothenburg, Sweden
E-mail: wastlund@chalmers.se