A PROOF OF PARISI'S CONJECTURE ON THE RANDOM ASSIGNMENT PROBLEM

SVANTE LINUSSON AND JOHAN WÄSTLUND

ABSTRACT. An assignment problem is the optimization problem of finding, in an m by n matrix of nonnegative real numbers, k entries, no two in the same row or column, such that their sum is minimal. Such an optimization problem is called a random assignment problem if the matrix entries are random variables. We give a formula for the expected value of the optimal k-assignment in a matrix where some of the entries are zero, and all other entries are independent exponentially distributed random variables with mean 1. Thereby we prove the formula $1+1/4+1/9+\cdots+1/k^2$ conjectured by G. Parisi for the case $k=m=n$, and the generalized conjecture of D. Coppersmith and G. B. Sorkin for arbitrary k, m and n.

1. INTRODUCTION

The problem of minimizing the sum of k entries in a matrix of nonnegative real numbers under the condition that no two of them may be in the same row or column is called an *assignment problem.* A set of matrix positions no two in the same row or column is called an *independent* set. An independent set of k matrix positions will also be called a k-assignment.

A *random assignment problem* is given by a number k, and an m by n matrix ($\min(m,n) \geq k$) of random variables. If P is a random matrix, we denote by $F_k(P)$ the expected value of the minimal sum of an independent set of k matrix entries. We use this notation even if P is a deterministic matrix.

In this article we prove the following.

Theorem 1.1 (Parisi's Conjecture [P98]). *Let P be a k by k matrix with independent* $\exp(1)$ *entries. Then*

$$F_k(P) = 1 + \frac{1}{4} + \frac{1}{9} + \cdots + \frac{1}{k^2}.$$

We also prove the following two generalizations.

Theorem 1.2 (Conjectured by D. Coppersmith and G. B. Sorkin [CS98]). *Let P be an m by n matrix with independent* $\exp(1)$ *entries. Then*

$$F_k(P) = \sum_{\substack{i,j\geq 0 \\ i+j<k}} \frac{1}{(m-i)(n-j)}. \tag{1}$$

Date: 26th September 2003.

Theorem 1.3 (Conjectured in [LW00]). *Let P be an $m \times n$ matrix where some entries are zero and the other entries are independent* $\exp(1)$*-variables. Then*

$$F_k(P) = \frac{1}{mn} \sum_{i,j} \frac{d_{i,j,k}(P)}{\binom{m-1}{i}\binom{n-1}{j}}. \tag{2}$$

Here $d_{i,j,k}(P)$ is an integer coefficient defined in terms of the combinatorics of the set of zeros in P, see Section 2.2. The identity (2) will be referred to as the *cover formula.* When P has no zeros,

$$d_{i,j,k}(P) = \binom{m}{i}\binom{n}{j}.$$

Hence Theorem 1.3 $\Rightarrow$ Theorem 1.2. To see the implication Theorem 1.2 $\Rightarrow$ Theorem 1.1 (proved in [CS98]), note that if we put $m = n = k$ in (1), then the terms for which $\gcd(k-i, k-j) = d$ sum to $1/d^2$ for $d = 1, \dots, k$.

As a consequence of Theorem 1.1 we also obtain a new and completely different proof of the following theorem, conjectured by M. Mézard and G. Parisi [MP85].

Theorem 1.4 (D. Aldous [A92, A01]). *Let P_k be a k by k matrix with independent* $\exp(1)$ *entries. Then*

$$\lim_{k\to\infty} F_k(P_k) = \frac{\pi^2}{6}.$$

In Section 7 we mention some other corollaries of Theorem 1.3.

Note: Within a few hours of the announcement of our proof of Theorem 1.1, we were contacted by C. Nair, B. Prabhakar and M. Sharma, who told us that they too had a proof of this theorem [NPS03a, NPS03b], and that they were about to finish a paper on it! Interestingly, these authors use a completely different approach, building on results from [N02].

1.1. Outline of the proof.

A key result, Theorem 4.1, is a formula for the probability that a certain row in a random matrix is used in the optimal k-assignment. From [LW00] we know that the probability that an exponentially distributed entry in a matrix P is used in the optimal k-assignment can be written $F_k(P) - F_k(P')$, where P' is obtained from P by setting the matrix entry in question to zero, see Theorem 2.2. Therefore, the formula for the probability that a row (or column) is used gives certain linear equations for the values of the random assignment problems. Provided m or n is sufficiently large compared to k, this system of linear equations has a unique solution given by (2), see Section 5. Finally in Section 6 we prove that for fixed k, fixed m, and a fixed set of zeros, $F_k(P)$ is given by a rational function in n, which must then agree with (2).

1.2. Background.

Random assignment problems have attracted the attention of researchers from physics, optimization, and probability. There are experimental results in [O92, PR93]. Constructive upper and lower bounds on $F_k(P)$ have been given in [W79, O92, CS98, L93, K87, GK93]. M. Mézard and G. Parisi [MP85, MP87] used the non-rigorous *replica method* and arrived at the conjectured limit $\pi^2/6$. This limit was subsequently established rigorously by D. Aldous [A92, A01] using the *weak convergence* method on a weighted infinite tree model. In this paper we continue the *exact formulas*-approach inspired by [P98] and developed further in [AS02, BCR02, CS98, CS02, LW00, EES01].

There are also interesting results on similar problems such as finding a minimal spanning tree in a graph with random edge weights [BFM98, FM89, F85, EES01]. An intriguing question is why the Riemann ζ-function appears in the limit of both the spanning tree and bipartite matching problems.

A tool that we believe can be useful to a wider range of problems is Theorem 2.2 below and its generalization Theorem 7.3 of [LW00].

Acknowledgement: We thank Mireille Bousquet-Mélou and Gilles Schaeffer for introducing the problem to us. We also thank an anonymous referee for suggesting some improvements of the exposition.

2. Preliminaries

2.1. **Probabilistic preliminaries.** We say that a random variable X is exponentially distributed with rate a if $Pr(X > t) = e^{-at}$ for $t \geq 0$. This is sometimes written $X \sim \exp(a)$. The rate of X is denoted $I(X)$. We have $E(X) = 1/I(X)$.

A random matrix P is called *standard* if the matrix entries are either zero or independent $\exp(1)$ random variables. A standard random assignment problem is thus determined by the numbers k, m, n, and the set Z of zero entries.

The following is a well-known lemma.

Lemma 2.1. *Let $a_1, \dots, a_n$ be positive real numbers, and let $X_1, \dots X_n$ be independent random variables with $X_i \sim \exp(a_i)$ for $i = 1, \dots, n$. Then the probability that X_i is minimal among $X_1, \dots X_n$ is*

$$\frac{a_i}{a_1 + \cdots + a_n}.$$

If we let Y be the minimum, then $Y \sim \exp(a_1 + \cdots + a_n)$, and Y is independent of which variable is minimal. Under the condition that X_i is minimal, $X_i = Y$, and for $j \neq i$, we can write $X_j = Y + X'_j$, where $X'_j \sim \exp(a_j)$, and the variables Y and X'_j for $j \neq i$ are all independent.

We say that a (deterministic) matrix is *generic* if no two sums of nonzero matrix entries are equal. In a standard matrix, the nonzero matrix entries are independent and have continuous distributions. Such a matrix is generic with probability 1. In the generic case, a nonzero entry belongs either to every optimal k-assignment, or to none. Hence without ambiguity we can speak of the probability that a certain nonzero entry is used in *the* optimal k-assignment, without specifying whether we take this to mean some optimal k-assignment, or every optimal k-assignment.

Notice that this is not the case for zero entries. Even in a generic matrix, there may be several different optimal k-assignments, that differ in the choice of zero entries.

The following theorem, Theorem 2.10 of [LW00], is essential for the recursion equations in Section 5.

Theorem 2.2. *Suppose P is a standard matrix. Suppose that the entry $P(i,j)$ is exponentially distributed. Let P' be as P except that $P'(i,j)$ is set to zero. Then the probability that (i,j) belongs to the optimal k-assignment in P is $F_k(P) - F_k(P')$.*

We include a proof for completeness.

Proof. We condition on the values of all matrix elements except $P(i,j)$. Let X be the value of the minimal k-assignment in P which does not use $P(i,j)$. Let Y be the value of the minimal $k-1$-assignment in P which does not use row i or column j. The optimal k-assignment in P either contains (i,j) and has value $Y + P(i,j)$, or does not contain (i,j) and has value X. Hence the probability that $P(i,j)$ belongs to the optimal k-assignment in P is equal to the probability that $P(i,j) < X - Y$.

We wish to show that this is equal to

$$F_k(P) - F_k(P') = \mathrm{E}(\max(0, \min(X - Y, P(i,j)))).$$

If $X \leq Y$, then both are zero. If $X > Y$, then we let $\delta = X - Y$. Then the probability that $P(i,j)$ is used in the optimal k-assignment in P is $1 - e^{-\delta}$. We compare this to

$$\mathrm{E}(\min(\delta, P(i,j))) = \delta e^{-\delta} + \int_0^\delta te^{-t}\,dt = \delta e^{-\delta} + 1 - (\delta+1)e^{-\delta} = 1 - e^{-\delta}.$$

This proves the theorem. □

2.2. **Covers.** We will consider sets of rows and columns in the matrices. A set λ of rows and columns is said to *cover* a set Z of matrix positions if every matrix position in Z is either in a row or in a column that belongs to λ. A cover with N rows and columns will be called an N-cover. The $k-1$-covers will be of particular importance. By a *partial $k-1$-cover* of Z, we mean a set of rows and columns which is a subset of a $k-1$-cover of Z.

When we speak of a cover of a matrix we mean a cover of its zeros. The *cover coefficient* $d_{i,j,k}(P)$ is the number of partial $k-1$-covers of P with i rows and j columns. For this to be nonzero, i and j have to be nonnegative integers with $i + j < k$. It is convenient to regard the cover coefficient as well-defined, but zero, for integers i, j outside this range.

Example Let P be an m by n standard matrix with zeros in positions $(1,1)$ and $(1,2)$.

$$P = \begin{pmatrix} 0 & 0 & P(1,3) & \dots \\ P(2,1) & P(2,2) & P(2,3) & \dots \\ \vdots & \vdots & \vdots & \end{pmatrix}$$

Then the cover coefficients $d_{i,j,3}(P)$ are given by

	$j=0$	1	2
$i=0$	1	n	1
1	m	n	
2	$m-1$		

According to Theorem 1.3, we have

$$F_3(P) = \frac{1}{mn}\left(\frac{1}{\binom{m-1}{0}\binom{n-1}{0}} + \frac{m}{\binom{m-1}{1}\binom{n-1}{0}} + \frac{m-1}{\binom{m-1}{2}\binom{n-1}{0}} + \frac{n}{\binom{m-1}{0}\binom{n-1}{1}} + \frac{n}{\binom{m-1}{1}\binom{n-1}{1}} + \frac{1}{\binom{m-1}{0}\binom{n-1}{2}}\right)$$

which simplifies to

$$\frac{1}{mn} - \frac{2}{m(n-1)} + \frac{1}{m(n-2)} + \frac{1}{(m-1)n} + \frac{1}{(m-1)(n-1)} + \frac{1}{(m-2)n}.$$

We say that a set λ of rows and columns is an *optimal cover* of P, if λ covers P, and λ has minimal cardinality among all covers of P. The following lemma is well-known. For a general introduction to matching theory we refer to [LP86].

Lemma 2.3 (Lattice structure of optimal covers). *The set of optimal covers of a matrix forms a lattice, where one of the lattice operations consists in taking union of row sets and intersection of column sets, and the opposite lattice operation is taking intersection of row sets and union of column sets.*

In particular, there is a row-maximal optimal cover containing every row that belongs to some optimal cover, and similarly a column-maximal optimal cover containing every column that belongs to some optimal cover.

Let us also recall a famous theorem of D. König and E. Egerváry, see e.g. [LP86].

Theorem 2.4. *If a matrix has no $k-1$-cover, then it has a zero-cost k-assignment.* □

We need the following lemmas.

Lemma 2.5. *Let $k \leq m, n$ be positive integers. Let P be an m by n matrix, with zeros in a certain set Z of positions, and (possibly random) positive values outside Z. Suppose that Z does not contain a $k+1$-assignment. Let λ be an optimal covering of Z. Then every row and every column of λ contains an element of every optimal k-assignment of P.*

Proof. Suppose that λ contains rows $1, \dots, r$ and no other rows. Since λ is optimal it follows by the König-Egerváry theorem that there is an r-assignment ν in Z containing no element from the columns in λ. Suppose (for a contradiction) that there is an optimal k-assignment μ which does not use row 1.

We now construct a sequence of matrix positions (all with zeros) as follows: Let ν_0 be the element of ν which is in the first row. Suppose that we have defined $\nu_0, \dots, \nu_h$ and $\mu_1, \dots, \mu_h$. Then if there is an element of $\mu \cap Z$ in the same column as ν_h, let this element be μ_{h+1}, and let ν_{h+1} be the element of ν which is in the same row as μ_{h+1}. Since μ does not contain any element from the first row, the sequences $\nu_0, \nu_1, \nu_2, \dots$ and $\mu_1, \mu_2, \dots$ cannot contain any repetitions o f the same element. Hence the sequence must end with an element ν_h such that there is no element of $\mu \cap Z$ in the same column.

We now consider two cases. Suppose first that no element in the column of ν_h belongs to μ. Since Z does not contain a $k+1$-assignment, the cardinality of λ is at most k. Each row and column of λ covers at most one element of μ, and row 1 does not cover any element of μ. Consequently there is an element (i,j) of μ which is not covered by λ, hence does not belong to Z. Then $\mu \setminus \{\mu_1, \dots, \mu_{h-1}, (i,j)\} \cup \{\nu_0, \dots, \nu_h\}$ is a k-assignment of smaller cost than μ, a contradiction. If on the other hand there is an element (i', j') of μ in the column of ν_h, then $\mu \setminus \{\mu_1, \dots, \mu_{h-1}, (i', j')\} \cup \{\nu_0, \dots, \nu_h\}$ is a k-assignment of smaller cost than μ. This contradiction proves the lemma. □

Lemma 2.6. *If λ is an optimal cover of a subset of Z, then there is an optimal k-assignment which intersects every row and every column of λ.*

Proof. This follows by a continuity argument. Let the values in the positions of Z not covered by λ be ϵ, and let ϵ tend to zero. Since there are only finitely many k-assignments, there has to be an assignment which is optimal for all sufficiently small ϵ, which by Lemma 2.5 has to intersect every row and column of λ. This assignment is also optimal for $\epsilon = 0$. □

Note that it is not true that every optimal k-assignment has to intersect the rows and columns of an optimal covering of a subset of Z.

2.3. **Recursions.** If P is an m by n matrix, and μ is an assignment, then we let $\mathrm{cost}_P(\mu)$ be the cost of μ, that is,

$$\mathrm{cost}_P(\mu) = \sum_{(i,j)\in\mu} P(i,j).$$

To recursively compute $F_k(P)$ we need Theorem 2.7 and Theorem 2.8 below from [LW00]. For completeness we include proofs here.

Theorem 2.7. *Let P be a nonnegative real matrix with no set of k independent zeros. Let λ be an optimal cover, and let x be a positive real number smaller than or equal to the minimum of the entries not covered by λ. Let P' be obtained from P by subtracting x from the entries not covered by λ, and adding x to the doubly covered entries. Then*

$$F_k(P) = (k - |\lambda|)x + F_k(P').$$

Proof. Let μ be an optimal k-assignment of P' intersecting every row and column of λ. Such an optimal assignment exists by Lemma 2.6. Note that subtracting x from every entry not covered by λ, and adding x to every doubly covered entry is the same thing as first subtracting x from all entries not covered by the rows of λ, and then adding x to all the entries covered by the columns of λ. Let i and j be the number of rows and columns in λ, respectively. Then

$$\begin{aligned}(3)\quad \mathrm{cost}_P(\mu) &= \mathrm{cost}_{P'}(\mu) + x\cdot(k-i) - x\cdot j = \\ &= x\cdot(k-|\lambda|) + \mathrm{cost}_{P'}(\mu) = x\cdot(k-|\lambda|) + F_k(P').\end{aligned}$$

For every k-assignment ν we have

$$\mathrm{cost}_P(\nu) \geq x \cdot (k - |\lambda|) + \mathrm{cost}_{P'}(\nu) \geq x \cdot (k - |\lambda|) + F_k(P').$$

Hence μ is an optimal k-assignment also in P, and $F_k(P)$ is given by (3). □

In the typical use of the theorem, x is the minimum of all the non-covered entries, which gives a new zero when we subtract x. The theorem will be used for random assignment problems, by conditioning on the location of the minimal non-covered entry. The special case of this theorem where all doubly covered entries are known not to be in the optimal k-assignment was treated in [CS02] and [AS02], where it was used in the proof of Theorem 1.2 for $k \leq 4$.

The next theorem is a special case of Theorem 2.9 in [LW00].

Theorem 2.8. *Let P be a nonnegative random matrix. Suppose that a column c has at least k zero entries. Let $P \backslash c$ be the matrix obtained from P by deleting column c. Then*

$$F_k(P) = F_{k-1}(P \backslash c). \tag{4}$$

Similarly, if a row r conatins k zero entries, then $F_k(P) = F_{k-1}(P \backslash r)$.

Proof. Every k-assignment in P contains a $k-1$-assignment in $P \backslash c$. Hence $F_k(P) \geq F_{k-1}(P \backslash c)$. Conversely, since every $k-1$-assignment in $P \backslash c$ can be extended with a zero in column c to a k-assignment in P, $F_k(P) \leq F_{k-1}(P \backslash c)$. □

3. Combinatorics of two optimal k-assignments

In this section we prove two results, Lemmas 3.2 and 3.4, which are used in Section 4.

An assignment problem can be rephrased in a setting of bipartite graphs with weighted edges. In this setting, an assignment is a matching, and a cover of rows and columns is a vertex cover of the edges. When speaking of sets of matrix positions, we will borrow some terminology from graph theory. When we speak of a *component* of a set of matrix positions, we mean a maximal subset which is connected by rook moves, that is, a position is connected to positions in the same row and in the same col umn.

Consider a deterministic assignment problem. If μ and ν are two different k-assignments, their *symmetric difference* $(\mu \backslash \nu) \cup (\nu \backslash \mu)$ is denoted $\mu \triangle \nu$. By a (μ, ν)*-alternating path* we mean a sequence $(i_1, j_1), (i_2, j_2), \dots, (i_r, j_r)$ of matrix positions, where positions belonging to μ alternate with positions belonging to ν and $i_x = i_{x+1}$ or $j_x = j_{x+1}$, for all $x = 1, \dots, r-1$. The parameter r will be called the *length* of the path. Each component of $\mu \triangle \nu$ is a (μ, ν)-alternating path (possibly cyclic).

If all matrix entries in $\mu \triangle \nu$ are zero, we say that μ and ν are *equivalent*, $\mu \equiv \nu$.

Lemma 3.1. *Let A be a nonnegative matrix, and let μ be an optimal k-assignment. Let a be a position which does not belong to μ. Suppose that there is another optimal k-assignment ν that contains a. Then there is an optimal k-assignment ν' containing a, such that $\mu \triangle \nu'$ has at most two components. Moreover, if there are two components, then both are paths of* *odd length, and one of them starts and ends in μ, and the other starts and ends in ν'.*

Proof. Let T be the component of $\mu\triangle\nu$ that contains a. If T is of even length, then we let $\nu' = \mu\triangle T$. Since T is an alternating path, it contains equally many elements from μ and ν. Therefore ν' has size k. Moreover, since T is a component, ν' is an assignment. Similarly, $\mu' = \nu\triangle T$ is a k-assignment, and since $\text{cost}(\mu') + \text{cost}(\nu') = \text{cost}(\mu) + \text{cost}(\nu)$, both μ' and ν' are optimal.

If on the other hand T has odd length, then it contains one more element of one of μ and ν than of the other. Since μ and ν have the same size, there must be another component T' of $\mu\triangle\nu$ that balances, so that $T \cup T'$ has equally many elements from μ and ν. Then we let $\nu' = \mu\triangle(T \cup T')$. By the same argument, ν' is an optimal k-assignment. □

Lemma 3.2. *Let P be a nonnegative matrix. Let r_1 and r_2 be rows, and let c_1 and c_2 be columns. Suppose that $P(r_1, c_1) = 0$, $P(r_2, c_1) > 0$, and $P(r_2, c_2) > 0$. Suppose further that there is an optimal k-assignment that contains (r_2, c_2), but does not intersect row r_1. Then there is no optimal k-assignment that contains (r_2, c_1).*

Proof. Let μ be an optimal k-assignment that contains (r_2, c_2) but no entry in row r_1. Suppose for contradiction that there is an optimal k-assignment ν that contains (r_2, c_1). By Lemma 3.1, we may assume that $\mu\triangle\nu$ consists of at most two components, and that if there are two, both are of odd length.

There must be a zero element $(s_1, c_1) \in \mu$, otherwise (r_2, c_2) could be replaced by (r_1, c_1) in μ. Let S be the path of $(\mu\triangle\nu) \setminus \{(r_2, c_1)\}$ that contains (s_1, c_1). In other words, S is the set of matrix positions in the (μ, ν)-alternating path that starts at (s_1, c_1) and continues in the direction opposite to that of (r_2, c_1).

There must be an element (r_1, d) of ν since otherwise (r_2, c_1) could be replaced by (r_1, c_1) in ν. Since μ does not use row r_1, the element (r_1, d) is in a component of $\mu\triangle\nu$ which is a path with one end belonging to ν. If the other end belongs to μ, the path has an even number of elements, and must therefore be the same as the path containing (r_2, c_1). Otherwise the other end too belongs to ν. Then the other path has both ends in μ. In either case, S must end with an element of μ in a row which is not used by ν.

If $\text{cost}(\mu \cap S) > \text{cost}(\nu \cap S)$, then $\text{cost}\left((\mu\triangle S) \cup \{(r_1, c_1)\}\right) < \text{cost}(\mu)$. Since $(\mu\triangle S) \cup \{(r_1, c_1)\}$ is a k-assignment, this contradicts the optimality of μ. Otherwise $\text{cost}(\mu \cap S) \leq \text{cost}(\nu \cap S)$. Then $\text{cost}\left((\nu\triangle S) \setminus \{(r_2, c_1)\}\right) < \text{cost}(\nu)$, again a contradiction. □

Lemma 3.3. *Let P be a matrix with exactly two equivalence classes of optimal k-assignments, and suppose that $\mu \not\equiv \nu$ are optimal k-assignments. Then $\mu\triangle\nu$ has at most two components with some nonzero entry, and if there are two, then both are paths of odd length, and one of them starts and ends in μ, and the other starts and ends in ν.*

Proof. Suppose that T is a component of $\mu\triangle\nu$ which has even length and contains at least one nonzero entry. Then $\nu' = \mu\triangle T$ and $\mu' = \nu\triangle T$ are k-assignments. Since

$$\text{cost}(\mu') + \text{cost}(\nu') = \text{cost}(\mu) + \text{cost}(\nu),$$

both μ' and ν' must be optimal. We have assumed that P has exactly two equivalence classes of optimal k-assignments, and therefore $\nu \equiv \nu'$. Consequently T is the only component with nonzero entries in $\mu\triangle\nu$.

Suppose instead that S is a component of $\mu\triangle\nu$ which has odd length and contains at least one nonzero entry. Then there must be another component S' of $\mu\triangle\nu$ that balances, so that $T = S \cup S'$ has equally many elements from μ and ν. With the same argument as above we can conclude that T contains all nonzero entries of $\mu\triangle\nu$. □

Lemma 3.4. *Let P be a nonnegative matrix. Suppose that there are exactly two equivalence classes of optimal k-assignments, and that $\mu \not\equiv \nu$ are two inequivalent optimal k-assignments. Suppose further that*

1. *μ has a nonzero entry in the last row and*
2. *ν does not use the last row.*

Then there is a unique row s such that

1. *ν has a nonzero entry in row s*
2. *There is a k-assignment $\mu' \equiv \mu$ that does not use row s.*

Proof. Existence: Choose a k-assignment $\mu' \equiv \mu$ which has as many matrix positions as possible in common with ν. Since the last row is used by μ' but not by ν, there has to be a row s which is used by ν but not by μ'. If ν has a zero entry in row s, then μ' must contain a zero entry in the same column. By replacing this zero by the zero in row s, we would obtain a k-assignment $\mu'' \equiv \mu'$ which has one more element in common with ν, a contradiction. Therefore, the entry in row s which belongs to ν must be nonzero.

Uniqueness: Assume on the contrary that there are two different rows s and t that contain nonzero entries of ν, say (s, c_1) and (t, c_2), and two equivalent k-assignments $\mu \equiv \mu'$ such that μ does not use row s and μ' does not use row t. If μ would not use row t either, then $\mu\triangle\nu$ would have two distinct paths (one path is impossible since it would have even length, starting and ending in ν) both with one end in a nonzero entry of ν, in the positions (s, c_1) and (t, c_2). This contradicts Lemma 3.3.

Hence there is a zero entry (s, c_3) of μ', and similarly a zero entry (t, c_4) of μ. Note that the symmetric difference $\mu\triangle\mu'$ contains only zeros and therefore it must consist of a number of (μ, μ')-alternating paths, all of even length. Let U be the one that contains (t, c_4). Unless U ends at row s, $\mu\triangle U$ will be a k-assignment equivalent to μ avoiding both rows s and t. We have already seen that this is impossible.

We may therefore assume that $\mu\triangle\mu'$ consists of a single path U from (t, c_4) to (s, c_3). Note that for every row r which is used in this path, there is a k-assignment $\mu'' \equiv \mu$ that avoids row r. This is obtained by choosing the zeros from μ in the part of U that goes towards (t, c_4), and choosing them from μ' in the part that goes towards (s, c_3).

Let (m, d) be the position in the last row used by μ and let L be the (μ, ν)-alternating path containing (m, d). First note that if L contains (s, c_1), it has to end there and thus be of even length, which means that it has passed through (t, c_2) first.

Case 1. L intersects U. In this case, the first element of L (starting from (m, d)) that belongs to U must be an element of μ. Suppose that this element is in row r. Then let $\mu'' \equiv \mu$ avoid row r. It follows that $\mu''\triangle\nu$ contains a path of even length starting at (m, d), and another path containing the nonzero element (s, c_1), contradicting Lemma 3.3.

Case 2. L does not intersect U. Let S be the path in $\mu\triangle\nu$ containing (s, c_1), (t, c_4) and (t, c_2) in this order.

We know that U and S intersect in (t, c_4). Of the positions in $U \cap S$, let (r, c) be the one which is closest to (s, c_3) in U, and let $\mu'' \equiv \mu$ avoid row r. Then $\mu''\triangle\nu$ will consist of a cycle containing the nonzero position (s, c_1), and two other paths, one containing (t, c_2) and one containing (m, d), contradicting Lemma 3.3. □

4. The probability that a row is used in the optimal assignment

Crucial for our proof is the following formula for the probability that a certain row without zeros intersects an optimal k-assignment. Let $\rho_k(P)$ denote this probability.

Theorem 4.1 (Row Inclusion Theorem). *Let P be a standard matrix, and let r be a row without zeros. The probability that some entry in r belongs to the optimal k-assignment is*

$$\rho_k(P) = \frac{1}{m}\sum_i \frac{\tilde{d}_{i,0,k}(P)}{\binom{m-1}{i}}, \tag{5}$$

where $\tilde{d}_{i,0,k}(P)$ is the number of partial $k-1$-covers of i rows not containing the row r.

Remarkably, $\rho_k(P)$ is independent of the number of columns in the matrix. Another consequence of (5) is that the probability of using a certain row without zeros in an optimal k-assignment does not change if further zeros are introduced in a row which belongs to an optimal cover. This observation turns out to be sufficient for the proof of the formula.

Lemma 4.2. *Let P be a standard matrix, and let r be a row that belongs to an optimal cover. Let P' be a matrix obtained from P by setting an entry in row r to zero. Suppose there is a row without zeros in P. Then*

$$\rho_k(P) = \rho_k(P').$$

Proof. Suppose that P is a standard matrix where the first ro w belongs to an optimal cover, and that there is an entry in the first row, say $(1, 1)$, which is not zero. Suppose further that the last row contains no zeros. We want to show that if we replace the entry in position $(1, 1)$ by zero, the probability that the last row is used in the optimal k-assignment does not change.

Let Ω_P be the probability space of all assignments of values to the random variables in P, that is, the space of all real nonnegative m by n matrices that have zeros in the positions where P has zeros. If $A \in \Omega_P$, we let $\varrho_k(A)$ be 1 if the last row is used by an optimal k-assignment, and 0 otherwise. We let A_x, for nonnegative real x, denote the matrix obtained from A by setting the entry in position $(1, 1)$ to x.

We construct a measure preserving involution φ on Ω_P with the property that (except possibly on a subset of probability zero) if $A \in \Omega_P$ is a matrix where the last row changes between being used and not being used in the optimal k-assignment when $A(1,1)$ is set to zero, then in $\varphi(A)$, the change goes the other way. In other words,

$$\varrho_k(A) - \varrho_k(A_0) = \varrho_k(\varphi(A)_0) - \varrho_k(\varphi(A)). \tag{6}$$

Let $A \in \Omega_P$. If $\varrho_k(A) = \varrho_k(A_0)$, we let $\varphi(A) = A$. Otherwise notice that if we decrease the entry in position $(1,1)$ continuously down to zero, there can be at most one point at which the location of the optimal k-assignment changes, and at this point, the entry in position $(1,1)$ goes from not being used to being used.

At the point A_f where the change occurs, the matrix has exactly two equivalence classes of optimal k-assignments. Let μ be one that contains an entry of the last row, and let ν be one that doesn't. By Lemma 3.4, there is a unique row s such that ν contains a nonzero entry in row s, and so that there is a k-assignment $\mu' \equiv \mu$ that contains no entry of row s. Notice that s cannot be the first row, since the first row is used by every optimal k-assignment.

We let $\varphi(A)$ be the matrix obtained from A by swapping the entries in row s with the corresponding entries (entries in the same column) in the last row, except in the columns where row s has zeros.

In the analysis of this mapping, we introduce some auxiliary matrices. Let A'_f be obtained from A_f by setting the entries in the last row that are in columns where row s has zeros, to zero. By Lemma 3.2, we can set these entries as small as we please without changing the location of the optimal assignments. Hence there will be no k-assignment of smaller cost than μ and ν in A'_f. Then since neither μ nor ν uses any zero entry in row s, these can of course be increased without changing the optimality of μ and ν. We let A''_f be the matrix obtained from A'_f by changing the zero positions in row s to the values of the corresponding entries in the last row of A. If A is generic, every optimal k-assignment in A''_f is equivalent either to μ or to ν.

Now $\varphi(A)_f$ is obtained from A''_f by swapping row s with the last row. This means that in $\varphi(A)_f$, there are exactly two equivalence classes of optimal k-assignments, one that includes position $(1,1)$ and one that doesn't. But since the last row has been swapped with another row, we have, for $x \neq f$,

$$\varrho_k(\varphi(A)_x) = 1 - \varrho_k(A_x),$$

which implies (6).

In $\varphi(A)_f$, row s has the property expressed in the conclusion of Lemma 3.4, but with the roles of μ and ν interchanged. It follows that $\varphi(\varphi(A)) = A$, and in particular that φ is invertible. The mapping φ is piecewise linear, and on each piece, it is a permutation of variables. Hence it is measure preserving. □

Proof of Theorem 4.1. We now establish the Row Inclusion Theorem by an inductive argument. Let P be a standard matrix, and suppose the last row contains no zero entry. Notice that by the König-Egerváry theorem, the formula holds whenever P has a k-assignment of only zeros. Suppose that the formula has been established for $k-1$-assignments and for every standard matrix with fewer nonzero entries than P.

Case 1: Some optimal cover of P contains a row r with at least one nonzero entry. Then every optimal k-assignment must use row r. Let P' be as P but with another zero in row r. By Lemma 4.2, the probability that the last row is used is the same in P as in P'. Hence we only have to show that $\tilde{d}_{i,0,k}(P) = \tilde{d}_{i,0,k}(P')$ for $i = 0, \dots, k-1$. We have to show that if a set of rows can be extended to a $k-1$-cover of P, then it is possible to use row r in this $k-1$-cover. This will follow if we can show that row r belongs to the optimal cover of the remaining zeros.

It suffices to show that if a row $s \neq r$ is deleted from the matrix, row r still belongs to the optimal cover of the remaining zeros. If the deletion of s does not decrease the maximal number of independent zeros, this is obvious. Suppose therefore that the deletion of s decreases the number of independent zeros. Then s belongs to an optimal cover of P. Hence the row-maximal optimal cover of P contains both r and s. When row s is deleted, the remaining rows and columns including r will constitute an optimal cover of the remaining zeros.

Case 2: There is a row in P with only zeros. Let P' be the $m-1$ by n matrix obtained from P by deleting this row. Then $\rho_k(P) = \rho_{k-1}(P')$. By induction, $\rho_{k-1}(P')$ is equal to

$$\frac{1}{m-1}\sum_i \frac{\tilde{d}_{i,0,k-1}(P')}{\binom{m-2}{i}}.$$

We have

$$\tilde{d}_{i,0,k}(P) = \tilde{d}_{i,0,k-1}(P') + \tilde{d}_{i-1,0,k-1}(P'),$$

for every i. Hence

$$\begin{aligned}
(7)\quad \frac{1}{m}\sum_i \frac{\tilde{d}_{i,0,k}(P)}{\binom{m-1}{i}} &= \frac{1}{m}\sum_i \frac{\tilde{d}_{i,0,k-1}(P') + \tilde{d}_{i-1,0,k-1}(P')}{\binom{m-1}{i}} \\
&= \frac{1}{m}\sum_i \frac{\tilde{d}_{i,0,k-1}(P')}{\binom{m-1}{i}} + \frac{1}{m}\sum_i \frac{\tilde{d}_{i,0,k-1}(P')}{\binom{m-1}{i+1}} \\
&= \sum_i \tilde{d}_{i,0,k-1}(P')\left(\frac{1}{m\binom{m-1}{i}} + \frac{1}{m\binom{m-1}{i+1}}\right) \\
&= \frac{1}{m-1}\sum_i \frac{\tilde{d}_{i,0,k-1}(P')}{\binom{m-2}{i}} = \rho_{k-1}(P') = \rho_k(P).
\end{aligned}$$

Case 3: There is a unique optimal cover λ consisting of only columns. Then these columns will be used by the optimal k-assignment. Therefore the number of entries not covered by λ in the optimal k-assignment is independent of the random variables in the matrix. We condition on the position of the minimal entry not covered by λ. If in each case we subtract this minimum from all entries not covered by λ, the same nonzero entries will be used in the optimal k-assignments. Here we are using Theorem 2.7 in the special case of no doubly covered entries. Let P^t be the matrix obtained by conditioning on the minimal entry in P not covered by λ being in row t, and subtracting this minimum from all entries not covered by λ. Then P^t is a standard matrix, and the new zero occurring in row t means that row t

belongs to an optimal cover of P^t. Hence in case the new zero is in the last row, that row must be used in the optimal k-assignment, while if it is not, we can find the probability that the last row is used by induction. We let $P\backslash t$ be the matrix obtained by deleting row t from P. Cases 1 and 2 imply that $\rho_k(P^t) = \rho_{k-1}(P\backslash t)$. It follows that $\rho_k(P)$ is given by

$$\frac{1}{m} + \frac{1}{m(m-1)} \sum_t \sum_i \frac{\tilde{d}_{i,0,k-1}(P\backslash t)}{\binom{m-2}{i}}. \tag{8}$$

Note that if α is a partial $k-2$-cover with $i-1$ rows of $P\backslash t$ then $\alpha \cup \{\text{row } t\}$ is a partial $k-1$-cover with i rows of P. Hence we have

$$\sum_t \tilde{d}_{i,0,k-1}(P\backslash t) = (i+1)\tilde{d}_{i+1,0,k}(P),$$

since both sides are equal to the sum, taken over all t, of the number of partial $k-1$-covers with i rows of P that use row t (and not the last row). It follows that (8) equals

$$\begin{aligned} &\frac{1}{m} + \frac{1}{m(m-1)} \sum_i \frac{(i+1)\tilde{d}_{i+1,0,k}(P)}{\binom{m-2}{i}} \\ &\qquad = \frac{1}{m} + \frac{1}{m} \sum_i \frac{\tilde{d}_{i+1,0,k}(P)}{\binom{m-1}{i+1}} = \frac{1}{m} \sum_i \frac{\tilde{d}_{i,0,k}(P)}{\binom{m-1}{i}}. \end{aligned} \tag{9}$$

This completes the proof of Theorem 4.1. □

5. Proof of the cover formula for large m or large n

The results of the previous section enable us to prove that the cover formula (2) holds for standard matrices whenever either m or n is sufficiently large compared to k. We first prove that the cover formula is consistent with the Row Inclusion Theorem.

Theorem 5.1. *Let P be a standard matrix where the last row contains no zeros, and let P_c be obtained from P by setting the entry in column c of the last row to zero. If the cover formula (2) holds for every P_c, then it holds for P.*

Proof. By Theorem 2.2 the probability that the entry in column c in the last row belongs to an optimal k-assignment in P is equal to $F_k(P) - F_k(P_c)$. Hence

$$nF_k(P) - \sum_c F_k(P_c) = \frac{1}{m} \sum_i \frac{\tilde{d}_{i,0,k}(P)}{\binom{m-1}{i}}.$$

Suppose that the cover formula holds for each P_c. Then

$$nF_k(P) = \frac{1}{mn} \sum_c \sum_{i,j} \frac{d_{i,j,k}(P_c)}{\binom{m-1}{i}\binom{n-1}{j}} + \frac{1}{m} \sum_i \frac{\tilde{d}_{i,0,k}(P)}{\binom{m-1}{i}}.$$

In order to prove that the cover formula holds for P, it is sufficient to prove that

$$\sum_{i,j} \frac{d_{i,j,k}(P)}{\binom{m-1}{i}\binom{n-1}{j}} = \frac{1}{n}\sum_c \sum_{i,j} \frac{d_{i,j,k}(P_c)}{\binom{m-1}{i}\binom{n-1}{j}} + \sum_i \frac{\tilde{d}_{i,0,k}(P)}{\binom{m-1}{i}}. \tag{10}$$

Let $\tilde{d}_{i,j,k}(P)$ be the number of partial $k-1$-covers of P with i rows, not containing the last row, and j columns. If we write

$$\tilde{d}_{i,0,k}(P) = \sum_j \left(\frac{\tilde{d}_{i,j,k}(P)}{\binom{n}{j}} - \frac{\tilde{d}_{i,j+1,k}(P)}{\binom{n}{j+1}} \right),$$

and fix i and j, we see that (10) will follow from the identity

$$\frac{d_{i,j,k}(P)}{\binom{n-1}{j}} = \frac{\tilde{d}_{i,j,k}(P)}{\binom{n}{j}} - \frac{\tilde{d}_{i,j+1,k}(P)}{\binom{n}{j+1}} + \frac{1}{n\binom{n-1}{j}} \sum_c d_{i,j,k}(P_c).$$

Here, partial covers that contain the last row will contribute $1/\binom{n-1}{j}$ to both sides. It only remains to show that

$$\frac{\tilde{d}_{i,j,k}(P)}{\binom{n-1}{j}} = \frac{\tilde{d}_{i,j,k}(P)}{\binom{n}{j}} - \frac{\tilde{d}_{i,j+1,k}(P)}{\binom{n}{j+1}} + \frac{1}{n\binom{n-1}{j}} \sum_c \tilde{d}_{i,j,k}(P_c),$$

or equivalently

$$j\tilde{d}_{i,j,k}(P) + (j+1)\tilde{d}_{i,j+1,k}(P) = \sum_c \tilde{d}_{i,j,k}(P_c).$$

Here the first term of the left hand side counts the partial $k-1$-covers of P_c that contain column c, summing over all c. If a partial $k-1$-cover of P_c does not contain the column c or the last row it must be a subset of a $k-1$-cover that does. Since there is only one zero in the last row of P_c we may always add the column c and still have a partial $k-1$-cover. Hence the second term can be seen to count the covers of P_c that do not contain column c. Thus (10) holds. $\square$

Next we show that the cover formula is consistent with removing a column that has at least k zeros.

Theorem 5.2. *Let P be a standard $m \times n$-matrix, and suppose that the first column has at least k zeros. Let P' be the m by $n-1$ matrix obtained from P by deleting the first column. If the cover formula holds for P', then it holds for P.*

Proof. By Theorem 2.8 $F_k(P) = F_{k-1}(P')$. We have

$$\begin{aligned} \frac{1}{mn}\sum_{i,j} \frac{d_{i,j,k}(P)}{\binom{m-1}{i}\binom{n-1}{j}} &= \frac{1}{mn}\sum_{i,j} \frac{d_{i,j-1,k-1}(P') + d_{i,j,k-1}(P')}{\binom{m-1}{i}\binom{n-1}{j}} \\ &= \frac{1}{mn}\sum_{i,j} \left(\frac{d_{i,j,k-1}(P')}{\binom{m-1}{i}\binom{n-1}{j}} + \frac{d_{i,j,k-1}(P')}{\binom{m-1}{i}\binom{n-1}{j+1}} \right) \\ &= \frac{1}{m(n-1)}\sum_{i,j} \frac{d_{i,j,k-1}(P')}{\binom{m-1}{i}\binom{n-2}{j}} = F_{k-1}(P') = F_k(P). \end{aligned} \tag{11}$$

$\square$

Theorem 5.3. *If P is a standard m by n matrix and* $\max(m,n) > (k-1)^2$, *then*

$$F_k(P) = \frac{1}{mn}\sum_{i,j}\frac{d_{i,j,k}(P)}{\binom{m-1}{i}\binom{n-1}{j}}.$$

Proof. Suppose without loss of generality that $m > (k-1)^2$. We argue by induction on the number of nonzero entries of P. By Theorems 5.1 and 5.2, it suffices to consider the case that P has at least one zero in each row, and no column with k or more zeros.

However, these conditions together imply that P must have a zero cost k-assignment. In this case the cover formula clearly holds, since the cover coefficients are zero.

□

6. Rationality of $F_k(P)$ as a function of n

We know from Section 5 that for fixed k, whenever m or n is large, the cover formula (2) for standard matrices holds. In order to prove that the formula holds for smaller values of m and n, it is therefore sufficient to show that if k, m and the zero positions are fixed, and we let P_n be the standard matrix with n columns, then there is a rational function in the variable n that gives the value of $F_k(P_n)$ for every n which is at least as large as k and the number of columns with zeros. This rational function must then be equal to the one given by the cover formula. We prove this by induction over a class of matrices which includes the class of standard matrices as a special case.

In an *exponential* matrix, all entries are linear combinations with nonnegative rational coefficients of a set $X_1, \dots, X_p$ of independent exponentially distributed random variables (not necessarily of rate 1). If a variable X_i in an exponential matrix has rate 1, and occurs in one and only one matrix position, and this matrix entry is equal to X_i, then the variable is called a *standard variable*, and the matrix position where it occurs is called a *standard position*.

We introduce the concept of a *matrix sequence*. The idea is to treat a set of similar matrices with different number of columns in a uniform way, in order to prove that there is a rational expression in the number of columns that gives the value of $F_k(P)$ for each matrix P in the set.

We consider k and m to be fixed numbers throughout this section.

Definition 6.1. We say that a linear function $f(x) = ax + b$ in one variable is *k-positive*, iff $f(x) > 0$ whenever $x \geq k$, or equivalently, if $a \geq 0$ and $b > -ak$.

Obviously, the sum of two or more k-positive functions is k-positive.

Definition 6.2. A *matrix sequence* is a sequence P_n ($n \geq n_0$) of exponential matrices satisfying the following:

1. P_n is an exponential m by n matrix in a set of variables $X_{n,1}, \dots, X_{n,p(n)}$.
2. In the first n_0 columns, the matrix entries of the matrices P_n differ only in that the first index of the variables is changed. In other words, the coefficient of $X_{n,i}$ in a matrix entry in the first n_0 columns of P_n is equal to the coefficient of $X_{n_0,i}$ in the same position in P_{n_0}.

3. Beyond column n_0, P_n has only standard entries.

Note that it follows that the function $p(n)$ is linear of the form constant + $m(n - n_0)$. Moreover, we say that the matrix sequence is *well-behaved* if

1. For each i such that the variables $X_{n,i}$ occur in the first n_0 columns, there is a k-positive linear polynomial $f_i(n)$ such that $I(X_{n,i}) = f_i(n)$.
2. Nonzero nonstandard entries occur only in columns that belong to the column-maximal optimal cover of the zeros, or equivalently, columns that intersect every maximal set of independent zeros.

Theorem 6.3 (Rationality Theorem)**.** *Suppose P is a well-behaved matrix sequence. Then there is a rational function $f(x)$ in one variable such that*

1. *If x is a zero of the denominator of f, then $x < k$.*
2. *$F_k(P_n) = f(n)$ for every $n \geq n_0$.*

Definition 6.4. If u_1 and u_2 are linear combinations of a set $X_1, \dots, X_p$ of exponential variables, we say that $u_1 \leq u_2$ or that $u_1 < u_2$ etc whenever such a statement holds with probability 1. In other words, $u_1 \leq u_2$ iff the corresponding inequality holds for the coefficient of each variable.

Two such linear combinations u_1 and u_2 are *incomparable* if each of them is greater than the other with positive probability.

We say that u_i is *potentially minimal* among $u_1, \dots, u_p$ if for no j, $u_j < u_i$.

Let λ be the row-maximal optimal cover of P. We prove the rationality theorem by induction on a number of parameters, in the following order:

1. The size of the largest independent set of zeros. A matrix sequence is considered simpler if it has a larger set of independent zeros.
2. The number of rows in λ. If the number of independent zeros are equal, the matrix sequence with fewer rows belonging to the row-maximal optimal cover is simpler.
3. The set of potentially minimal nonstandard entries not covered by λ. If two matrix sequences are equal by 1 and 2, then one of them is considered simpler than the other if its set of potentially minimal nonstandard entries not covered by λ is a subset of the corresponding set for the other one.
4. If 1–3 are equal, and there are two incomparable nonstandard entries not covered by λ, then a matrix sequence is simpler if there are fewer variables with different coefficients in the first two (in lexicographic order, say) incomparable potentially minimal nonstandard entries not covered by λ.
5. If 1–3 are equal, and there is a minimal non-covered nonstandard entry, then a matrix sequence is simpler if the number of variables occurring in this entry is smaller.

Let P be a well-behaved matrix sequence, and suppose that the rationality theorem holds for every simpler well-behaved matrix sequence (with the values of m and k under consideration). We may of course assume that $|\lambda| < k$. Since P is well-behaved, in each row not in λ, all but at most $k-1$ entries are standard and not covered by λ.

We show that $F_k(P)$ can be expressed in terms of rational functions in n, and values of simpler well-behaved matrix sequences.

Case 1: There are two or more non-covered incomparable nonstandard entries. Let u_1 and u_2 be the first two (in lexicographic order). We choose i and j such that the coefficient of $X_{n,i}$ is greater in u_1, and the coefficient of $X_{n,j}$ is greater in u_2. Let the coefficients be a_1, a_2, b_1 and b_2 so that $u_1 = a_1X_{n,i} + b_1X_{n,j} + \dots$ and $u_2 = a_2X_{n,i} + b_2X_{n,j} + \dots$.

Let Q and R be the matrix sequences obtained by conditioning on $(a_1 - a_2)X_{n,i}$ being smaller or greater than $(b_2 - b_1)X_{n,j}$, respectively.

The rates of $(a_1 - a_2)X_{n,i}$ and $(b_2 - b_1)X_{n,j}$ are $f_i(n)/(a_1 - a_2)$ and $f_j(n)/(b_2 - b_1)$ respectively. The probability of $(a_1 - a_2)X_{n,i}$ being smaller than $(b_2 - b_1)X_{n,j}$ is

$$\frac{f_i(n)/(a_1 - a_2)}{f_i(n)/(a_1 - a_2) + f_j(n)/(b_2 - b_1)},$$

and similarly, the probability of $(b_2-b_1)X_{n,j}$ being smaller than $(a_1-a_2)X_{n,i}$ is

$$\frac{f_j(n)/(b_2 - b_1)}{f_i(n)/(a_1 - a_2) + f_j(n)/(b_2 - b_1)}.$$

Therefore,

$$F_k(P_n) = \frac{f_i(n)F_k(Q_n)/(a_1 - a_2) + f_j(n)F_k(R_n)/(b_2 - b_1)}{f_i(n)/(a_1 - a_2) + f_j(n)/(b_2 - b_1)}. \tag{12}$$

We show that Q and R can be regarded as well-behaved matrix sequences. If we condition on $(a_1 - a_2)X_{n,i}$ being smaller than $(b_2 - b_1)X_{n,j}$, then we can write $(a_1 - a_2)X_{n,i} = Y_n$ and $(b_2 - b_1)X_{n,j} = Y_n + Z_n$, where

$$Y_n \sim \exp\left(\frac{f_i(n)}{a_1 - a_2} + \frac{f_j(n)}{b_2 - b_1}\right),$$

and

$$Z_n \sim \exp\left(\frac{f_j(n)}{b_2 - b_1}\right).$$

Y_n and Z_n are independent, and the rates are k-positive.

When replacing $X_{n,i}$ and $X_{n,j}$ by the new variables Y_n and Z_n, only the nonstandard entries are affected. Either u_1 is smaller than u_2 in Q, or at least the number of variables with distinct coefficients is smaller than in P, since u_1 and u_2 will get the same coefficient for Y_n.

Hence Q, and similarly R, are simpler than P. By induction, it follows that (12) gives a rational expression for $F_k(P)$, where the denominator is nonzero for $n \geq k$.

Case 2: There are no two non-covered incomparable nonstandard entries. Then either there is among the non-covered nonstandard entries a minimal one, or there are no non-covered nonstandard entries. We can treat these slightly different cases in the same way.

If there is a minimal non-covered nonstandard entry, let $aX_{n,i}$ be a term occurring in this matrix entry. We let S_n be a set of random variables consisting of $aX_{n,i}$ and all the non-covered standard variables (if there is no non-covered nonstandard entry, we let S_n consist only of the non-covered standard entries). There are at least $n - k + 1$ non-covered standard entries in each non-covered row. By grouping together the standard entries in each row, we can write the total rate of S_n in a uniform way as a sum of k-positive terms. Therefore the rate of the minimum Y_n of the terms in S_n is k-positive as a function of n.

We condition on the minimal entry in S_n. We can then replace the terms in S_n by new variables Y_n and $Z_{n,i}$, where $Y_n = \min(S_n)$ and $Z_{n,i}$ are the differences between the remaining terms in S_n and the minimum.

Since all non-covered nonstandard entries contain the variable $X_{n,i}$ with a coefficient of at least a, we can (by Theorem 2.7) subtract the minimum of S_n from every non-covered matrix entry, and add it to the doubly covered entries.

We get

$$F_k(P_n) = \frac{k - |\lambda|}{I(Y_n)} + \frac{1}{I(Y_n)} \sum_{t \in S_n} I(t) F_k(Q_n(t)), \tag{13}$$

where $Q_n(t)$ are the new matrices obtained by conditioning on the term t being smallest, and performing the change of variables and subtraction of the minimum, and $I(t)/I(Y_n)$ is the corresponding probability. In order to write this in a uniform way for the different values of n, we group together the cases where the minimum occurs in a particular row and beyond column n_0.

By permuting columns, we may assume that in those cases, the minimum always occurs in column $n_0 + 1$. The probability for each such case is $(n - n_0)/I(Y_n)$. In this way, (13) will contain the same terms for each $n > n_0$.

If we let S'_n be the subset of S_n consisting of variables occurring in the first n_0 columns, and we let S''_n be the set of standard variables in column $n_0 + 1$, we get

$$F_k(P_n) = \frac{k - |\lambda|}{I(Y_n)} + \frac{1}{I(Y_n)} \sum_{t \in S'_n} I(t) F_k(Q_n(t)) + \frac{n - n_0}{I(Y_n)} \sum_{t \in S''_n} I(t) F_k(Q_n(t)). \tag{14}$$

For $n = n_0$, the second sum will be empty. However, provided we can show that $Q(t)$ is a well-behaved matrix sequence simpler than P, it will follow by induction that the denominator of $F_k(Q(t))$ does not vanish for $n \geq k$. Hence the rational expression that occurs when multiplying the probability $(n - n_0)/I(S_n)$ with the expression for $F_k(Q(t))$ will vanish for $n = n_0$. Therefore the rational expression that results will give the correct value of $F_k(P)$ also for $n = n_0$.

It remains to show that $Q(t)$ is a well-behaved matrix sequence, and that it is simpler than P. We first show that $Q(t)$ is well-behaved. We have already seen that the rates of the variables occurring in $Q(t)$ are k-positive. We therefore turn to the distribution of nonstandard entries. Possibly there are some new nonstandard entries among the doubly covered entries. If there is no new zero among the non-covered entries, the columns of the doubly covered positions of course belong to the column-maximal optimal cover. Suppose that a new zero occurs among the non-covered entries. If the new zero is in a column that belongs to the column-maximal cover of P, then this is still the column-maximal cover of the new set of zeros. If the new zero is in a column that does not belong to the column-maximal cover of P, then since it is also in a row that does not belong to the row-maximal cover, there must be an independent set of zeros in $Q(t)$ which is larger than the largest independent set of zeros in P. Therefore, the column-maximal cover in P, extended with the column where the new zero has occurred, will be an optimal cover in $Q(t)$. Hence in any case, the columns of the column-maximal cover in P belong to the column-maximal cover of $Q(t)$.

We now show that $Q(t)$ is simpler than P. If $Q(t)$ contains a new zero, then either it gives a larger independent set of zeros, or it has to be covered by a column in the row-maximal cover in $Q(t)$. In either case, $Q(t)$ is simpler than P. If no new zero occurs, this must be because the minimal term in S_n was a term occurring in the minimal non-covered nonstandard entry. In this case, the number of variables in this entry will decrease, again making $Q(t)$ simpler than P.

Hence (13) gives a rational expression for $F_k(P_n)$ whose denominator is non-vanishing for $n \geq k$. This completes the proof of Theorem 6.3.

We are finally able to give a proof of Theorem 1.3, which we restate.

Theorem 1.3 *If P is a standard matrix, then*

$$F_k(P) = \frac{1}{mn} \sum_{i,j} \frac{d_{i,j,k}(P)}{\binom{m-1}{i}\binom{n-1}{j}}.$$

Proof. A standard matrix with n_0 columns can be extended to a well-behaved matrix sequence P_n, $(n \geq n_0)$ by inserting more columns without zeros. Hence Theorem 6.3 shows that there is a rational function f such that $F_k(P_n) = f(n)$ for $n \geq n_0$. Since the cover coefficient $d_{i,j,k}(P_n)$ can be expressed uniformly as a polynomial in n, the cover formula gives a rational function in n which takes the same values as f on the infinitely many integers $n > \max(n_0, (k-1)^2)$. Hence the cover formula must agree with f, and give the value of $F_k(P)$. $\square$

7. Some corollaries of the cover formula

Besides giving a new proof of Aldous' $\zeta(2)$-limit theorem, Theorem 1.3 also makes the conjectured limits of [LW00] rigorous. It follows already from the conjecture of Coppersmith and Sorkin, now Theorem 1.2, that if $\alpha, \beta \geq 1$, then as $k \to \infty$, the value of the optimal k-assignment in a $[\alpha k]$ by $[\beta k]$-matrix of exp(1)-variables converges to

$$\int_\Delta \frac{dx\,dy}{(\alpha - x)(\beta - y)},$$

where Δ is the triangle with vertices in $(0,0)$, $(1,0)$ and $(0,1)$. For instance, when $\alpha = 1$ and $\beta = 2$ the limit is equal to

$$\frac{\pi^2}{12} - \frac{(\log 2)^2}{2}.$$

It also follows that the value of a standard matrix with zeros in a region which is scaled up with k, m and n will converge to a similar integral. In particular it is shown in [LW00] that (assuming the cover formula) if P_k is a standard k by k matrix with zeros outside an inscribed circle, then

$$\lim_{k \to \infty} F_k(P_k) = \frac{\pi^2}{24}.$$

A result presented as a conjecture in [O92], and proved in [A01], states that in the case $k = m = n$ with no zeros, as $n \to \infty$, the probability that the smallest entry in a row belongs to the optimal assignment converges to $1/2$. We can now give an exact formula for this probability for finite k,

$$\frac{1}{2} + \frac{1}{2k}.$$

In the case $k = m$, the probability that the smallest entry in a particular row belongs to the optimal assignment is equal to the probability that the smallest entry in the entire matrix does. For arbitrary k, m, and n, the probability that the smallest entry in the matrix belongs to the optimal k-assignment is

$$1 - \frac{k(k-1)}{2mn}.$$

These theorems are obtained by combining the results of [LW00] and Theorem 1.3.

References

[A92] Aldous, D., *Asymptotics in the random assignment problem*, Pr. Th. Related Fields, **93** (1992) 507–534.

[A01] Aldous, D., *The $\zeta(2)$ limit in the rando m assignment problem*, Random Structures Algorithms **18** (2001), no 4. 381–418.

[AS02] 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.

[BFM98] Beveridge Andrew, Frieze, Alan M. and McDarmid Colin J.H., *Random minimum length spanning trees in regular graphs*, Combinatorica **18** (1998) 311–333.

[BCR02] Buck, M. W., Chan, C. S., Robbins, D. P., *On the expected value of the minimum assignment*, Random Structures Algorithms 21 (2002), no. 1, 33–58.

[CS98] Coppersmith, D., Sorkin, G. B., *Constructive Bounds and Exact Expectations For the Random Assignment Problem*, Random Structures Algorithms **15** (1999), 133–144.

[CS02] Coppersmith, D., Sorkin, G. B., *On the expected incremental cost of a minimum assignment.* In Contemporary Mathematics (B. Bollobás, ed.), Vol. 10 of *Bolyai Society Mathematical Studies*, Springer.

[EES01] Eriksson, H., Eriksson, K., and Sjöstrand, J. , *Exact expectation for random graphs and assignments*, Proceedings of FPSAC 2001, Arizona.

[FM89] Frieze, Alan M. and McDarmid Colin J.H., *On random minimum spanning trees*, Combinatorica **9** (1989).

[F85] Frieze, Alan M., *On the value of a random minimum spanning tree problem*, Disc Applied Math. **10** (1985) 47–56.

[GK93] Goemans, M. X., and Kodialam, M. S., *A lowe r bound on the expected cost of an optimal assignment*, Math. Oper. Res., **18** (1993), 267–274.

[K87] 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.

[L93] Lazarus, A., *Certain expected values in the random assignment problem*, Oper. Res. Lett., **14** (1993), 207–214.

[LW00] Linusson, S., and Wästlund, J., *A generalization of the random assignment problem*, arXiv:math.CO/0006146. Unpublished manuscript.

[LP86] Lovász, L., Plummer, M. D., *Matching Theory*, North-Holland 1986.

[MP85] Mézard, M. and Parisi, G., *Replicas and optimization*, J. Phys. Lett. **46**(1985), 771–778.

[MP87] Mézard, M. and Parisi, G., *On the solution of the random link matching problems*, J. Phys. Lett. **48** (1987), 1451–1459.

[N02] C. Nair, *Towards the resolution of Coppersmith-Sorkin conjectures.* Proceedings of the 40th annual Allerton conference on communication, control and computing, 2002.

[NPS03a] C. Nair, B. Prabhakar and M. Sharma, *A proof of the conjecture due to Parisi for the finite random assignment problem*, available at http://www.stanford.edu/~ balaji/rap.html.

[NPS03b] Chandra Nair, Balaji Prabhakar and Mayank Sharma, *Proofs of the Parisi and Coppersmith-Sorkin Conjectures for the Fininte Random Assignment Problem*, Proceedings of IEEE FOCS, 2003.

[O92] Olin, Birgitta, *Asymptotic properties of the random assignment problem*, Ph.D. thesis, Kungl Tekniska Högskolan, Stockholm, Sweden, (1992).

[PR93] 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.

[P98] Parisi, Giorgio, *A conjecture on random bipart ite matching*, Physics e-Print archive, http://xxx.lanl.gov/ps/cond-mat/9801176, January (1998).

[W79] Walkup, D. W., *On the expected value of a random assignment problem*, SIAM J. Comput., **8** (1979), 440–442.

Svante Linusson, Department of Mathematics, Linköpings universitet, SE-581 83 Linköping, Sweden

E-mail address: linusson@mai.liu.se

Johan Wästlund, Department of Mathematics, Linköpings universitet, SE-581 83 Linköping, Sweden

E-mail address: jowas@mai.liu.se