A Proof of a Conjecture of Buck, Chan, and Robbins on the Expected Value of the Minimum Assignment

Johan Wästlund
Department of Mathematics, Linköping University, S-581 83 Linköping, Sweden; e-mail: jowas@mai.liu.se

Received 9 February 2004; accepted 28 September 2004
DOI 10.1002/rsa.20066
Published online 18 January 2005 in Wiley InterScience (www.interscience.wiley.com).

ABSTRACT: We prove the main conjecture of the paper "On the expected value of the minimum assignment" by Marshall W. Buck, Clara S. Chan, and David P. Robbins [Random Structures Algorithms 21 (2002), 33–58]. This is an exact formula for the expected value of a certain type of random assignment problem. It generalizes the formula $1 + 1/4 + \cdots + 1/n^2$ for the n by n exp(1) random assignment problem. © 2005 Wiley Periodicals, Inc. Random Struct. Alg., 26, 237–251, 2005

1. INTRODUCTION

This work is motivated by a conjecture made by Marshall W. Buck, Clara S. Chan, and David P. Robbins [4]. This conjecture in turn is a generalization of a conjecture made in 1998 by the physicist Giorgio Parisi [16]. Consider an n by n matrix of independent exp(1) random variables. Parisi conjectured that the expected value of the minimum sum of n matrix entries, of which no two belong to the same row or column, is given by the formula

$$1 + \frac{1}{4} + \frac{1}{9} + \cdots + \frac{1}{n^2}. \tag{1}$$

At the time, the main open question was the value of the limit of the expected minimum sum as n tends to infinity. Upper and lower bounds had been established in [17,3,15,7–9]. A nonrigorous argument due to Marc Mézard and Parisi [13] showed that the limit ought to be $\zeta(2) = \pi^2/6$. David Aldous [1] had proved, using an infinite model, that the limit exists. Aldous would later confirm the $\zeta(2)$ limit conjecture of Mézard and Parisi [2]. See also [6] for further background.

The striking conjecture of Parisi [obviously consistent with the conjectured $\zeta(2)$-limit] paved the way for an entirely new approach. It seemed likely that (1) would yield to an inductive argument, and that therefore the $\zeta(2)$-limit could be established by exact analysis of the "finite n" case. The Parisi formula was almost immediately generalized by Don Coppersmith and Gregory Sorkin [5] to

$$\sum_{\substack{i,j\geq 0 \\ i+j<k}} \frac{1}{(m-i)(n-j)} \tag{2}$$

for the minimum sum of k entries, no two in the same row or column, in an m by n matrix. It is not hard to verify that (2) specializes to (1) when $k = m = n$.

The Coppersmith-Sorkin conjecture was then generalized in two different directions by Buck, Chan, and Robbins [4] and by Svante Linusson and the author [10]. Meanwhile, the $\zeta(2)$ limit was established by Aldous. In 2003, the Parisi and Coppersmith-Sorkin formulas were proved simultaneously and independently by Linusson and the author [11] and by Chandra Nair, Balaji Prabhakar, and Mayank Sharma [14].

In this paper we show that by combining the methods and ideas of [11] and [4], we can prove a simultaneous generalization of the Buck-Chan-Robbins conjecture and the main theorem of [11]. At the same time, this provides a considerable simplification of the proof of (1) and (2).

2. THE BUCK-CHAN-ROBBINS FORMULA

Let X be an m by n matrix of nonnegative real numbers. The rows and columns of the matrix will be indexed by *weighted sets* R and C respectively. We may take $R = \{1, \dots, m\}$ and $C = \{1, \dots, n\}$. The sets R and C are endowed with weight functions $\|\cdot\|_R$ and $\|\cdot\|_C$ respectively, that associate a positive weight to each element of the set. A *k-assignment* is a set $\mu \subseteq R \times C$ of k matrix positions, or *sites*, of which no two belong to the same row or column. An assignment will also be called an *independent* set. The *cost* of μ is the sum

$$\text{cost}_X(\mu) = \sum_{(i,j)\in\mu} X(i,j)$$

of the matrix entries in μ. We let $\min_k(X)$ denote the minimum cost of all k-assignments in X, and we say that μ is a *minimum* k-assignment if $\text{cost}_X(\mu) = \min_k(X)$.

In [4], the Parisi and Coppersmith-Sorkin conjectures are generalized to a certain type of matrix with entries which are exponential random variables, but not necessarily with parameter 1. We say that a random variable x is *exponential of rate a* if $Pr(x > t) = e^{-at}$ for every $t \geq 0$. In this case we write $x \sim \exp(a)$. Buck, Chan, and Robbins considered the following type of matrix: For every $(i,j) \in R \times C$, $X(i,j)$ is $\exp(\|i\|_R \, \|j\|_C)$-distributed, and independent of all other matrix entries. To state the formula, we use the notation

$$\|\alpha\|_R = \sum_{i\in\alpha} \|i\|_R$$

and

$$\|\overline{\alpha}\|_R = \sum_{\substack{i \in R \\ i \notin \alpha}} \|i\|_R,$$

for the weight of a set of rows and its complement, respectively. We use similar notation for sets of columns. Where no confusion can arise, we will drop the subscripts R and C.

The formula conjectured by Buck, Chan, and Robbins in [4] is given in two different versions, a "combinatorial" formula, involving a binomial coefficient, and an equivalent "probabilistic" version of the formula. The situation is similar in the paper [10], whose main conjecture was proved in [11]. In this paper we work in the probabilistic setting, which seems to be the natural one. We remark, however, that both in [10] and [4], the discovery of the probabilistic formulas were made through formal manipulation of the combinatorial formulas. Therefore, the latter have played an important role in obtaining the results of this paper.

Theorem 2.1 (The Buck-Chan-Robbins conjecture, combinatorial version). *Let X be a matrix as described above. Then*

$$\mathbf{E}[\min_k(X)] = \sum_{\substack{\alpha \subseteq R \\ \beta \subseteq C}} \binom{m+n-1-|\alpha|-|\beta|}{k-1-|\alpha|-|\beta|} \frac{(-1)^{k-1-|\alpha|-|\beta|}}{\|\overline{\alpha}\| \, \|\overline{\beta}\|}. \tag{3}$$

Notice that in order for the binomial coefficient to be nonzero, we must have $|\alpha|+|\beta| < k$, which resembles the condition $i+j < k$ in the Coppersmith-Sorkin formula (2). It is still not entirely obvious that (3) specializes to the Coppersmith-Sorkin formula when the row and column weights are set to 1. However, in [4], the formula (3) is shown to be equivalent to a formula given by an urn model.

3. THE BUCK-CHAN-ROBBINS URN MODEL

The following urn model is described in [4]: An urn contains a set of balls, each with a given positive weight. Balls are drawn one at a time without replacement, and each time the probability of drawing a particular ball is proportional to its weight. This simple model has perhaps been studied before, but the connection to random assignment problems is due to Buck, Chan, and Robbins.

To each weighted set we can associate such an urn process. In our applications, we consider the urn processes on the sets R and C of row and column indices. We consider a continuous time version of this process. Each ball i remains in the urn for an amount of time which is $\exp(\|i\|)$-distributed, after which it pops out of the urn. The times at which the balls leave the urn are all independent.

The urn process can be thought of as a continuous time random walk on the power set of the set of balls. If α is a set of balls, we denote by $Pr(\alpha)$ the probability that this random walk reaches α, in other words the probability that every ball in α is drawn before every ball not in α.

Example 3.1. *If there are three balls labeled* 1, 2, 3*, then*

$$Pr(\emptyset) = Pr(\{1,2,3\}) = 1,$$
$$Pr(\{1\}) = \frac{\|1\|}{\|\{1,2,3\}\|}$$

and

$$Pr(\{1,2\}) = \frac{\|1\| \cdot \|2\|}{\|\{1,2,3\}\| \cdot \|\{2,3\}\|} + \frac{\|1\| \cdot \|2\|}{\|\{1,2,3\}\| \cdot \|\{1,3\}\|},$$

since the set $\{1,2\}$ *can be obtained either by first drawing* 1 *and then* 2*, or the other way around.*

The following formula is stated in [4] and shown to be equivalent to (3):

Theorem 3.2 (The Buck-Chan-Robbins conjecture, probabilistic version).

$$\mathbf{E}[\min_k(X)] = \sum_{\substack{\alpha\subseteq R\\ \beta\subseteq C\\ |\alpha|+|\beta|<k}} \frac{Pr_R(\alpha)Pr_C(\beta)}{\|\overline{\alpha}\| \, \|\overline{\beta}\|}. \tag{4}$$

Notice that the Coppersmith-Sorkin formula (2) follows immediately from (4). If we set the row and column weights to 1, we can group together the terms for which $|\alpha| = i$ and $|\beta| = j$. The denominators are all $(m-i)(n-j)$, and the probabilities in the numerators sum to 1.

Example 3.3. *Let* X *be a* 2×2 *matrix, and let the row and column weights be* $\|i\|_R = a_i$*, and* $\|j\|_C = b_j$*. If* $k = 2$*, then according to* (4)*,*

$$\mathbf{E}[\min_2(X)] = \frac{1}{(a_1+a_2)(b_1+b_2)} + \frac{a_1}{a_1+a_2} \cdot \frac{1}{a_2(b_1+b_2)} + \frac{a_2}{a_1+a_2} \cdot \frac{1}{a_1(b_1+b_2)}$$
$$+ \frac{b_1}{b_1+b_2} \cdot \frac{1}{(a_1+a_2)b_2} + \frac{b_2}{b_1+b_2} \cdot \frac{1}{(a_1+a_2)b_1}.$$

By partial fraction decomposition, this expression simplifies to

$$\mathbf{E}[\min_2(X)] = \frac{-3}{(a_1+a_2)(b_1+b_2)} + \frac{1}{a_2(b_1+b_2)} + \frac{1}{a_1(b_1+b_2)}$$
$$+ \frac{1}{b_2(a_1+a_2)} + \frac{1}{b_1(a_1+a_2)}, \tag{5}$$

in accordance with (3). If we set the weights equal to 1, (5) evaluates to $5/4 = 1 + 1/4$, as predicted by (1).

4. MAIN THEOREM

The main theorem of [11] is a formula for the expected value of the minimum k-assignment in a matrix where a specified set of entries are set to zero, and the remaining entries are independent exp(1)-variables. In this article we prove a formula for the common generalization of the matrices considered in [4] and in [11]. We say that X is a *standard matrix* if the entries in a certain set of sites are zero, and the remaining entries are independent and distributed according to the row- and column-weights, that is, $X(i,j) \sim \exp(\|i\|\,\|j\|)$. This is an obvious generalization of the concept of standard matrix in [11].

Let $Z \subseteq R \times C$ be a set of sites. A *line* is a row or a column. Let λ be a set of lines. We say that λ is a *cover* of Z if every site in Z lies on a line that belongs to λ. By a cover of the matrix X we mean a cover of the set of zeros of X. By a $(k-1)$-*cover* we mean a cover consisting of $k-1$ lines. Finally, a *partial* $(k-1)$-*cover* is a subset of a $(k-1)$-cover.

Let $\mathcal{I}_k(X)$ be the set of partial $(k-1)$-covers of the zeros of X. Our main theorem states that

$$\mathbf{E}[\min_k(X)] = \sum_{\substack{\alpha\subseteq R\\ \beta\subseteq C\\ \alpha\cup\beta\in\mathcal{I}_k(X)}} \frac{Pr(\alpha)Pr(\beta)}{\|\overline{\alpha}\|\,\|\overline{\beta}\|}. \tag{6}$$

If X has no zeros, $\mathcal{I}_k(X)$ consists of all sets of at most $k-1$ lines. Hence (4) follows from (6).

5. MATRIX REDUCTION

Polynomial time algorithms for computing $\min_k(X)$ for a deterministic matrix X are well known. We do not focus on issues of computational efficiency, but we outline an algorithm whose special features will be of importance.

Let $Z \subseteq R \times C$ be a set of sites. We say that a cover of Z is *optimal* if it has the minimum number of lines among all covers of Z. The *rank* of a set of sites is the size of the largest independent subset. The following is a famous theorem due to D. König and E. Egerváry (see for instance [12]).

Theorem 5.1 (König-Egerváry theorem). *The number of lines in an optimal cover of Z is equal to* $\operatorname{rank}(Z)$.

The following was Lemma 2.5 of [11].

Lemma 5.2. *Let X be a nonnegative m by n matrix, and let λ be an optimal cover of X. Suppose that there is no zero cost $(k+1)$-assignment in X. Then every line in λ intersects every minimum k-assignment in X.*

The following matrix operation is fundamental for the algorithm. We refer to it as *matrix reduction*. Let X be a nonnegative m by n matrix, and let $\lambda = \alpha \cup \beta$ be an optimal cover of X, where α is the set of rows and β is the set of columns in λ. The *reduction X'* of X by λ is obtained from X as follows: Let t be the minimum matrix entry of X which is not covered by λ. If the site (i,j) is not covered by λ, we let $X'(i,j) = X(i,j) - t$. In particular, this means that X' will have a zero which is not covered by λ. If the site (i,j) is *doubly covered* by λ, that is, $i \in \alpha$ and $j \in \beta$, then we let $X'(i,j) = X(i,j) + t$. Finally if (i,j) is covered by exactly one line in λ, we let $X'(i,j) = X(i,j)$. Notice that the entries of X' are nonnegative.

Lemma 5.3. *Let X be a nonnegative matrix, and suppose that there is no zero cost $(k+1)$-assignment in X. Let X' be the reduction of X by an optimal cover λ. A k-assignment which is minimal in X is also minimal in X'.*

Proof. Let t be the minimum of the entries in X that are not covered by λ. For $s < t$, let X_s be the matrix obtained from X by subtracting s from the noncovered entries and adding s to the doubly covered entries. Since X_s has precisely the same zeros as X, λ is an optimal cover of X_s. By Lemma 5.2, every minimum k-assignment in X_s intersects every line of λ. By continuity, it follows that there is a minimum k-assignment in X' that intersects every line in λ. All k-assignments that intersect every line of λ are affected in the same way by the reduction from X to X', namely if μ is such a k-assignment, then $\mathrm{cost}_{X'}(\mu) = \mathrm{cost}_X(\mu) - (k - |\lambda|)t$ (as in Theorem 2.7 of [11]). Hence if μ is a minimum k-assignment in X, then μ is minimal also in X'. ∎

The following lemma is well-known. Again we refer to [12].

Lemma 5.4. *There is an optimal cover of Z containing every row that belongs to some optimal cover of Z, and similarly there is an optimal cover that contains every column that belongs to some optimal cover.*

These covers are called the *row-maximal* and the *column-maximal* optimal covers, respectively.

The proof of Theorem 6.3 is based on induction over matrix reduction. Therefore we need the following lemma:

Lemma 5.5. *Let $X = X_0$ be a nonnegative m by n matrix, and let $k \leq \min(m,n)$. For $i \geq 0$, let X_{i+1} be the reduction of X_i by the column-maximal optimal cover of X_i, and let Z_i be the set of zeros of X_i. Then $\mathrm{rank}(Z_{i+1}) \geq \mathrm{rank}(Z_i)$, and if equality holds, then the number of rows in the column-maximal optimal cover of Z_{i+1} is greater than the number of rows in the column-maximal optimal cover of Z_i. In particular, $X_{\binom{k+1}{2}}$ has a zero cost k-assignment.*

Proof. Let λ_i be the column-maximal optimal cover of Z_i. By the König-Egerváry theorem, Z_i has an independent subset μ_i containing exactly one site in each line of λ_i. Hence μ_i contains no site which is doubly covered by λ_i. It follows that $\mathrm{rank}(Z_{i+1}) \geq \mathrm{rank}(Z_i)$. Suppose that $\mathrm{rank}(Z_{i+1}) = \mathrm{rank}(Z_i)$. Notice that λ_i is the column-maximal optimal cover of the subset $Z_i \cap Z_{i+1}$ of Z_i consisting of the zeros of X_i, which are not doubly covered by λ_i. Since λ_{i+1} is also a cover of $Z_i \cap Z_{i+1}$, every row in λ_i belongs to λ_{i+1}. Since X_{i+1} has a zero which is not covered by λ_i, there has to be a row in λ_{i+1} which is not in λ_i.

Hence in each step of the reduction process, either the rank of the set of zeros increases, or the number of rows in the column-maximal optimal cover increases. Therefore, when the matrix has a zero cost $(k-1)$-assignment, it takes at most k more reductions until it has a zero cost k-assignment. The statement follows. ∎

A feature of matrix reduction that has been exploited in several papers [5,10,3] is that it keeps track of the cost of the minimum k-assignment. If t is as above and X reduces to X', then $\min_k(X) = (k - |\lambda|) \cdot t + \min_k(X')$. Therefore, we can compute $\min_k(X)$ recursively by iterating the reduction and keeping track of the values of t as well as the sizes of the optimal covers that are used. As long as the matrix entries are independent exponential variables, it is easy to compute the expected value of the minimum t, even for general m and n. However, since the doubly covered entries will eventually consist of sums of several dependent random variables, it is difficult to reach any conclusions valid for general k through this approach.

One of the insights that led to the proof of the Parisi formula in [11] was the fact that information about the probability that a certain matrix element belongs to a minimum assignment will give information about the expected minimum cost (Lemma 7.1 below). However, a problem with the reduction algorithm is that, in general, it loses track of the location of the minimum assignment.

Example 5.6. *Here $k = 2$, and after the final step, the matrix contains two zero-cost 2-assignments, of which only one was minimal in the original matrix. Hence the converse of Lemma 5.3 is not true.*

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 4 & 3 \end{pmatrix} \xrightarrow{\emptyset} \begin{pmatrix} 0 & 1 & 2 \\ 2 & 3 & 2 \end{pmatrix} \xrightarrow{\{\text{column}1\}} \begin{pmatrix} 0 & 0 & 1 \\ 2 & 2 & 1 \end{pmatrix} \xrightarrow{\{\text{row}1\}} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

The approach taken in this paper builds on an observation that has so far been overlooked, even in [11], namely that if the column-maximal optimal cover is used consistently, the matrix reduction algorithm keeps track of the set of rows that participate in the minimum k-assignment. Before we prove this, we cite a lemma from [4]. Although these authors make no claims of originality, we have not been able to trace it to any other source.

Lemma 5.7 (Nesting Lemma). *Let X be a real m by n matrix, and let $k_1 \leq k_2 \leq \min(m, n)$ be positive integers. If μ is an optimal k_1-assignment in X, then there is an optimal k_2-assignment μ' in X such that every line that intersects μ also intersects μ'. Moreover, if ν is an optimal k_2-assignment, then there is an optimal k_1-assignment ν' such that every line that intersects ν' intersects ν.*

We can now prove a partial converse of Lemma 5.3.

Lemma 5.8. *Let X be a nonnegative m by n matrix. Let λ be the column-maximal optimal cover of the zeros of X. Let $k \leq \min(m, n)$, and suppose that X has no zero cost k-assignment. Let X' be the reduction of X by λ. For every i, if there is a minimum k-assignment in X' that intersects row i, then there is a minimum k-assignment in X that intersects row i.*

Proof. By Lemma 5.5, the reduction process can be continued until there is a zero cost k-assignment. If at this point a row intersects a minimum (that is, zero cost) k-assignment, then the row must obviously contain a zero. Conversely, since a zero is a minimum 1-assignment, it follows from the Nesting Lemma that a row that contains a zero will intersect some minimum k-assignment. Hence to prove the lemma, it is sufficient to prove that whenever a new zero occurs in the reduction, the row containing this zero intersects some minimum k-assignment in the original matrix. Let l be the number of lines in λ. By the König-Egerváry theorem, $l < k$. By the Nesting Lemma, it suffices to prove that a row that contains a new zero of X' intersects a minimum $(l+1)$-assignment int X. If the new zero occurs in a row that belongs to some optimal cover of the zeros of X, then this is obvious, since that row must contain a zero of X. If the new zero occurs in a row that does not belong to any optuimal cover of the zeros of X, then since λ is column-maximal, it occurs in a position (i,j) which is not covered by any optimal cover of the zeros of X. By the König-Egerváry theorem, there is an $(l+1)$-assignment of cost $X(i,j)$ in X. Since every $(l+1)$-assignment must contain an entry which is not covered by λ, and $X(i,j)$ is minimal among these, it follows that (i,j) belongs to a minimum $l+1$ assignment in X. ▪

6. A FORMULA FOR THE PARTICIPATION PROBABILITY OF A ROW

In this section we establish a connection between the random assignment problem and the urn model by deriving a formula for the probability that a certain row intersects a minimum k-assignment. The special case of standard matrices without zeros was proved in [4]. Another special case, that of row- and column-weights equal to 1 (rate 1 exponential variables) was proved in [11] by a different method. A striking feature of this formula is that it is independent of the number of columns in the matrix, and of their weights.

Suppose that X is a matrix where some entries are zero, and the remaining entries are (possibly random) positive real numbers. We let $\mathcal{R}(X)$ be the set of rows in the row-maximal optimal cover of the zeros of X. Recall that $\mathcal{I}_k(X)$ is the set of partial $(k-1)$-covers of X. The following result was used in [11] although its proof is hidden in the analysis of case 1 in the proof of Theorem 4.1.

Lemma 6.1. *Suppose that $\alpha \subseteq R$ and $\alpha \in \mathcal{I}_k(X)$. Then $\alpha \cup \mathcal{R}(X) \in \mathcal{I}_k(X)$.*

Proof. It suffices to show that if a row $i \in \alpha$ is deleted from the matrix, the rows in $\mathcal{R}(X)$ belong to the row-maximal optimal cover of the remaining zeros. If the deletion of row i does not decrease the rank of the set of zeros, this is obvious. If on the other hand the deletion of row i decreases the rank of the set of zeros, then $i \in \mathcal{R}(X)$, and the deletion of this row from the row-maximal optimal cover obviously gives an optimal cover of the remaining zeros. ▪

Corollary 6.2. *Let X be a matrix as above, and let Z be the set of zeros of X. Let (i,j) be a site such that the set $Z' = Z \cup \{(i,j)\}$ has greater rank than Z. Let X' be a matrix whose set of zeros is Z'. Then if α is a set of rows, $\alpha \in \mathcal{I}_k(X')$ iff $\alpha \cup \{i\} \in \mathcal{I}_k(X)$.*

Proof. Since $i \in \mathcal{R}(X')$, this follows from Lemma 6.1. ▪

Theorem 6.3. *Let $|R| = m$ and $|C| = n$. Let $Z \subseteq R \times C$, and let $\mathcal{R}$ be the set of rows in the row-maximal optimal cover of Z. Let X be a random $R \times C$-matrix with the following properties: Z is the set of zeros of X. The remaining entries in the rows in $\mathcal{R}$ have arbitrary distribution on the nonnegative real numbers, allowing dependencies. For every $i \in R \backslash \mathcal{R}$ and $j \in C$, if $(i,j) \notin Z$, $X(i,j)$ is $\exp(\|i\|\,\|j\|)$-distributed, and independent of all other* *matrix entries. Let* $k \le \min(m, n)$. *Suppose that row* i_0 *has no zeros. Then the probability that row* i_0 *participates in a minimum k-assignment is*

$$\|i_0\| \sum_{\substack{\alpha \subseteq R \\ \alpha \in \mathcal{I}_k(X) \\ i_0 \notin \alpha}} \frac{Pr(\alpha)}{\|\overline{\alpha}\|}. \tag{7}$$

In our applications of this theorem, we are always dealing with standard matrices. However, to make the inductive proof go through, we must allow for arbitrary nonnegative values of the nonzero entries in the rows in $\mathcal{R}$. We first make a comment on how to interpret (7).

If $\mathcal{A}$ is a family of subsets of R, we let $T_R(\mathcal{A})$ denote the expected amount of time that the urn process spends in $\mathcal{A}$, that is, the expected amount of time for which the set of elements that have been drawn in the urn process is a set that belongs to $\mathcal{A}$. If $R \in \mathcal{A}$, then $T_R(\mathcal{A})$ is infinite. Otherwise it is given by the following formula.

Lemma 6.4.

$$T_R(\mathcal{A}) = \sum_{\alpha \in \mathcal{A}} \frac{Pr(\alpha)}{\|\overline{\alpha}\|}.$$

Proof. If the urn process on R reaches the set α, then the expected amount of time until another element is drawn from the urn is $1/\|\overline{\alpha}\|$. Hence the expected amount of time that the urn process spends in α is $Pr(\alpha)/\|\overline{\alpha}\|$. The formula follows by summing these times over all $\alpha \in \mathcal{A}$. ∎

It follows that

$$\sum_{\substack{\alpha \subseteq R \\ \alpha \in \mathcal{I}_k(X) \\ i_0 \notin \alpha}} \frac{Pr(\alpha)}{\|\overline{\alpha}\|} = T_R(\{\alpha \in \mathcal{I}_k(X) : i_0 \notin \alpha\}),$$

in other words, the second factor of (7) is the expected amount of time it takes until either i_0 is drawn, or the urn process reaches a set which is not in $\mathcal{I}_k(X)$.

At the same time, there is another natural interpretation of (7). If $\alpha \subseteq R$ and $i_0 \notin \alpha$, then

$$Pr(\alpha) \cdot \frac{\|i_0\|}{\|\overline{\alpha}\|}$$

is the probability that the urn process reaches α, and that the next element to be drawn is i_0. Since for different sets α, these events are mutually exclusive, (7) is the probability that i_0 belongs to the first set in the urn process which is not in $\mathcal{I}_k(X)$.

Proof of Theorem 6.3. The proof is by induction over matrix reduction, according to Lemma 5.5. The base of this induction is the case that X has a zero cost k-assignment. In this case, the probability that row i_0 participates in a minimum k-assignment is zero. So is (7), since $\mathcal{I}_k(X)$ is empty. For the induction step, suppose that the statement holds for all matrices with a larger set of independent zeros, or equally many independent zeros and more rows in the column-maximal optimal cover.

Let X' be the reduction of X by the column-maximal optimal cover. Let Z' be the set of zeros of X', and let $\mathcal{R}' = \mathcal{R}(X')$. We consider two cases, and we show that if we condition on being in one of these two cases, the right-hand side of (7) gives the participation probability in each case. By Lemma 5.8, we do not have to distinguish between the probability that row i_0 intersects a minimum k-assignment in X and X' respectively.

1. All new zeros are in rows that belong to $\mathcal{R}$. In this case we have $\operatorname{rank}(Z') = \operatorname{rank}(Z)$, and consequently $\mathcal{R}' = \mathcal{R}$. It follows from Lemma 6.1 that a set of rows belongs to $\mathcal{I}_k(X')$ if and only if it belongs to $\mathcal{I}_k(X)$. If we condition on Z', then X' satisfies the criteria of the theorem. Hence by induction, the probability that row i_0 participates in a minimum k-assignment in X', or equivalently in X, is given by (7).
2. There is a new zero $X'(i,j)$ such that $i \notin \mathcal{R}$. Since $X(i,j)$ has continuous distribution and is independent of all other matrix entries, we may assume that $X'(i,j)$ is the only new zero in X'. Since the site (i,j) is not covered by any optimal cover of Z, we have $\operatorname{rank}(Z') = 1 + \operatorname{rank}(Z)$. Hence $\mathcal{R}' \supseteq \mathcal{R} \cup \{i\}$.

If $i = i_0$, then every minimum k-assignment in X' must intersect row i_0. Suppose, on the other hand, that $i \neq i_0$. If we condition on Z', then X' satisfies the criteria of the theorem. Hence, by induction, the probability that i_0 participates in a minimum k-assignment in X' is equal to

$$\|i_0\| \sum_{\substack{\alpha \subseteq R \\ \alpha \in \mathcal{I}_k(X') \\ i_0 \notin \alpha}} \frac{Pr(\alpha)}{\|\overline{\alpha}\|}.$$

By Corollary 6.2, $\alpha \in \mathcal{I}_k(X')$ if and only if $\alpha \cup \{i\} \in \mathcal{I}_k(X)$. Therefore, if we condition on being in case 2, the participation probability of row i_0 is

$$\|i_0\| \cdot \left(\frac{1}{\|\overline{\mathcal{R}}\|} + \sum_{\substack{i \notin \mathcal{R} \\ i \neq i_0}} \frac{\|i\|}{\|\overline{\mathcal{R}}\|} \sum_{\substack{\alpha \subseteq R \\ \alpha \cup \{i\} \in \mathcal{I}_k(X) \\ i_0 \notin \alpha}} \frac{Pr(\alpha)}{\|\overline{\alpha}\|} \right). \tag{8}$$

We have to show that the second factor is equal to $T_R(\{\alpha \in \mathcal{I}_k(X) : i_0 \notin \alpha\})$. The sum

$$\sum_{\substack{\alpha \subseteq R \\ \alpha \cup \{i\} \in \mathcal{I}_k(X) \\ i_0 \notin \alpha}} \frac{Pr(\alpha)}{\|\overline{\alpha}\|} \tag{9}$$

is the expected amount of time that the urn process spends in sets α such that $\alpha \cup \{i\} \in \mathcal{I}_k(X)$ and $i_0 \notin \alpha$. Since this is independent of the time at which i is drawn, (9) is the expected amount of time that the urn process spends in $\{\alpha \in \mathcal{I}_k(X) : i_0 \notin \alpha\}$, given that the element i is drawn at time zero. By Lemma 6.1, this amount of time is also independent of the times at which the elements in $\mathcal{R}$ are drawn. Therefore, (9) is equal to the expected amount of time that the urn process remains in $\{\alpha \in \mathcal{I}_k(X) : i_0 \notin \alpha\}$ after the element i has been drawn, if we condition on i being the first element in $\overline{\mathcal{R}}$ to be drawn.

We can now interpret the second factor of (8). The term $1/\|\overline{\mathcal{R}}\|$ is the expected amount of time until the first element of $\overline{\mathcal{R}}$ is drawn. The probability that this element is i is $\|i\|/\|\overline{\mathcal{R}}\|$, and therefore the second factor of (8) is equal to the expected amount of time that the urn process spends in $\{\alpha \in \mathcal{I}_k(X) : i_0 \notin \alpha\}$. ∎

7. A FORMULA FOR $\mathbf{E}[\min_k(X)]$

In this section we prove the formula (6) for the expected cost of the minimum k-assignment. We first cite a lemma that appeared in [10]. The case $a = 1$ is Theorem 2.2 of [11]. This lemma is the reason why we are interested in the probabilities of certain matrix entries occurring in a minimum assignment.

Lemma 7.1 ([10]). *Let X be a random matrix where a particular entry $X(i,j) \sim \exp(a)$ is independent of the other matrix entries. Let X' be as X except that $X'(i,j) = 0$. Then the probability that (i,j) belongs to a minimum k-assignment in X is*

$$a \cdot (\mathbf{E}[\min_k(X)] - \mathbf{E}[\min_k(X')]).$$

Proof. We condition on all entries in X except $X(i,j)$. Let X_t be the deterministic matrix obtained by also conditioning on $X(i,j) = t$. Let $f(t) = \min_k(X_t)$. Then either f is constant, or f increases linearly up to a certain point after which it is constant. The key observation is that $f'(t) = 1$ if the site (i,j) belongs to a minimum k-assignment in X_t, and $f'(t) = 0$ otherwise (disregarding the possibility that t is equal to the point where f is not differentiable). Therefore, if $t \sim \exp(a)$, then the probability that (i,j) belongs to a minimum k-assignment in X_t is equal to $\mathbf{E}[f'(t)]$. By partial integration,

$$\begin{aligned} \mathbf{E}[f'(t)] &= a\int_0^\infty e^{-ax} f'(x)\,dx = a\int_0^\infty d(e^{-ax}f(x)) + a^2\int_0^\infty e^{-ax}f(x)\,dx \\ &= -af(0) + a\mathbf{E}[f(t)] = a\cdot(\mathbf{E}[\min_k(X)] - \mathbf{E}[\min_k(X')]). \end{aligned} \tag{10}$$

∎

Let X be a standard matrix with rows and columns indexed by the weighted sets R and C. We introduce the notation

$$F_k(X) = \sum_{\substack{\alpha\subseteq R \\ \beta\subseteq C \\ \alpha\cup\beta\in\mathcal{I}_k(X)}} \frac{Pr(\alpha)Pr(\beta)}{\|\overline{\alpha}\|\,\|\overline{\beta}\|}.$$

Then (6) states that

$$\mathbf{E}[\min_k(X)] = F_k(X). \tag{11}$$

The proof is inductive, and closely parallels the proof in [11]. We first prove that (11) is consistent with the row participation formula.

Lemma 7.2. *Let X be a standard matrix where row i_0 has no zeros. Let X_j be the matrix obtained from X by setting the entry in position (i_0,j) equal to zero. If $\mathbf{E}[\min_k(X_j)] = F_k(X_j)$ for every j, then $\mathbf{E}[\min_k(X)] = F_k(X)$.*

Proof. By Lemma 7.1, the probability that the site (i_0,j) belongs to a minimum k-assignment in X is

$$\|i_0\|\,\|j\|\,(\mathbf{E}[\min_k(X)] - \mathbf{E}[\min_k(X_j)]).$$

By summing over $j \in C$ we obtain the probability that row i_0 participates in a minimum k-assignment. Hence this sum is equal to the formula in Theorem 6.3:

$$\|i_0\| \sum_{j\in C} \|j\| \left(\mathbf{E}[\min_k(X)] - \mathbf{E}[\min_k(X_j)]\right) = \|i_0\| \sum_{\substack{\alpha\subseteq R\\ \alpha\in\mathcal{I}_k(X)\\ i_0\notin\alpha}} \frac{Pr(\alpha)}{\|\overline{\alpha}\|}.$$

We drop the factors $\|i_0\|$ and solve for $\mathbf{E}[\min_k(X)]$:

$$\mathbf{E}[\min_k(X)] = \sum_{j\in C} \frac{\|j\|}{\|C\|}\mathbf{E}[\min_k(X_j)] + \sum_{\substack{\alpha\subseteq R\\ \alpha\in\mathcal{I}_k(X)\\ i_0\notin\alpha}} \frac{Pr(\alpha)}{\|\overline{\alpha}\|\,\|C\|}. \tag{12}$$

We replace $\mathbf{E}[\min_k(X_j)]$ by $F_k(X_j)$:

$$\mathbf{E}[\min_k(X)] = \sum_{j\in C} \frac{\|j\|}{\|C\|} \sum_{\substack{\alpha\subseteq R\\ \beta\subseteq C\\ \alpha\cup\beta\in\mathcal{I}_k(X_j)}} \frac{Pr(\alpha)Pr(\beta)}{\|\overline{\alpha}\|\,\|\overline{\beta}\|} + \sum_{\substack{\alpha\subseteq R\\ \alpha\in\mathcal{I}_k(X)\\ i_0\notin\alpha}} \frac{Pr(\alpha)}{\|\overline{\alpha}\|\,\|C\|}.$$

For α such that $i_0 \in \alpha$, the condition $\alpha \cup \beta \in \mathcal{I}_k(X_j)$ is equivalent to $\alpha \cup \beta \in \mathcal{I}_k(X)$. Hence

$$\begin{aligned}\mathbf{E}[\min_k(X)] = &\sum_{\substack{\alpha\subseteq R\\ \beta\subseteq C\\ i_0\in\alpha\\ \alpha\cup\beta\in\mathcal{I}_k(X)}} \frac{Pr(\alpha)Pr(\beta)}{\|\overline{\alpha}\|\,\|\overline{\beta}\|} + \sum_{\substack{\alpha\subseteq R\\ i_0\notin\alpha}}\sum_{\beta\subseteq C}\sum_{\substack{j\in C\\ \alpha\cup\beta\in\mathcal{I}_k(X_j)}} \frac{\|j\|}{\|C\|}\cdot\frac{Pr(\alpha)Pr(\beta)}{\|\overline{\alpha}\|\,\|\overline{\beta}\|}\\ &+ \sum_{\substack{\alpha\subseteq R\\ \alpha\in\mathcal{I}_k(X)\\ i_0\notin\alpha}} \frac{Pr(\alpha)}{\|\overline{\alpha}\|\,\|C\|}.\end{aligned} \tag{13}$$

We focus on the middle term. Since $i_0 \notin \alpha$ and the zero at position (i_0,j) in X_j is the only zero in its row, $\alpha \cup \beta$ can be extended to a $(k-1)$-cover of X_j if and only if this can be done by covering (i_0,j) with a column, that is, if and only if $\alpha \cup \beta \cup \{j\}$ can be extended to a $(k-1)$-cover of X. Notice that this statement is true whether or not $j \in \beta$. Therefore, the middle term can be written

$$\frac{1}{\|C\|}\sum_{\substack{\alpha\subseteq R\\ i_0\notin\alpha}} \frac{Pr(\alpha)}{\|\overline{\alpha}\|} \sum_{\substack{\beta\subseteq C\\ j\in C\\ \alpha\cup\beta\cup\{j\}\in\mathcal{I}_k(X)}} \frac{\|j\|\,Pr(\beta)}{\|\overline{\beta}\|}. \tag{14}$$

In the inner sum, the contribution from those β and j for which $j \in \beta$ is

$$\sum_{\substack{\beta\subseteq C\\ \alpha\cup\beta\in\mathcal{I}_k(X)}} \frac{\|\beta\|\,Pr(\beta)}{\|\overline{\beta}\|}. \tag{15}$$

The contribution from the β and j for which $j \notin \beta$ is

$$\sum_{\beta \subseteq C} \sum_{\substack{j \notin \beta \\ \alpha \cup \beta \cup \{j\} \in \mathcal{I}_k(X)}} \frac{\|j\| \, Pr(\beta)}{\|\overline{\beta}\|}. \tag{16}$$

If for every $\beta' \subseteq C$ we group together those terms for which $\beta \cup \{j\} = \beta'$, we see that [for nonempty β' such that $\alpha \cup \beta' \in \mathcal{I}_k(X)$] these terms sum to $Pr(\beta')$, since we are summing the probabilities of obtaining β' in the urn process by first drawing the elements of β and then drawing j. Hence (16) is equal to

$$\sum_{\substack{\beta \neq \emptyset \\ \alpha \cup \beta \in \mathcal{I}_k(X)}} Pr(\beta) = \sum_{\substack{\beta \neq \emptyset \\ \alpha \cup \beta \in \mathcal{I}_k(X)}} \frac{\|\overline{\beta}\| \, Pr(\beta)}{\|\overline{\beta}\|}. \tag{17}$$

When we add (15) to (17), we get a factor $\|\beta\| + \|\overline{\beta}\| = \|C\|$, which cancels the factor $1/\|C\|$ in (14). Hence (14) is equal to

$$\sum_{\substack{\alpha \subseteq R \\ i_0 \notin \alpha}} \sum_{\substack{\beta \subseteq C \\ \beta \neq \emptyset \\ \alpha \cup \beta \in \mathcal{I}_k(X)}} \frac{Pr(\alpha) Pr(\beta)}{\|\overline{\alpha}\| \, \|\overline{\beta}\|}. \tag{18}$$

We substitute this into (13), and notice that the third term in (13) corresponds to α and β such that $i_0 \notin \alpha$ and $\beta = \emptyset$. Hence the right-hand side of (13) equals $F_k(X)$, as was to be proved. ∎

Secondly, we show that (11) is consistent with removing a column that contains at least k zeros.

Lemma 7.3. *Suppose that* $\mathbf{E}[\min_{k-1}(Y)] = F_{k-1}(Y)$ *for every standard matrix* Y*. Let* X *be a standard matrix that has a column with at least* k *zeros. Then* $\mathbf{E}[\min_k(X)] = F_k(X)$.

Proof. Suppose that X has at least k zeros in column j_0. Let X' be the m by $n-1$ matrix obtained from X by deleting column j_0. Since every $(k-1)$-assignment in X' can be extended to a k-assignment in X by including a zero in column j_0, we have

$$\mathbf{E}[\min_k(X)] = \mathbf{E}[\min_{k-1}(X')] = \sum_{\alpha \subseteq R} \frac{Pr(\alpha)}{\|\overline{\alpha}\|} \sum_{\substack{\beta \subseteq C' \\ \alpha \cup \beta \in \mathcal{I}_{k-1}(X')}} \frac{Pr(\beta)}{\|\overline{\beta}\|}, \tag{19}$$

where $C' = C \backslash \{j_0\}$. The inner sum is equal to $T_{C'}(\{\beta : \alpha \cup \beta \in \mathcal{I}_{k-1}(X')\})$. Since there are k zeros in column j_0, every $(k-1)$-cover of X must include j_0. Therefore, if $\beta \subseteq C'$, then $\alpha \cup \beta \in \mathcal{I}_{k-1}(X')$ if and only if $\alpha \cup \beta \in \mathcal{I}_k(X)$, and this in turn holds if and only if $\alpha \cup \beta \cup \{j_0\} \in \mathcal{I}_k(X)$. Hence

$$T_{C'}(\{\beta : \alpha \cup \beta \in \mathcal{I}_{k-1}(X')\}) = T_C(\{\beta : \alpha \cup \beta \in \mathcal{I}_k(X)\}).$$

It follows that (19) equals

$$\sum_{\alpha \subseteq R} \frac{Pr(\alpha)}{\|\overline{\alpha}\|} \sum_{\substack{\beta \subseteq C \\ \alpha \cup \beta \in \mathcal{I}_k(X)}} \frac{Pr(\beta)}{\|\overline{\beta}\|} = F_k(X).$$

∎

We are now in a position to prove that (11) holds whenever m is sufficiently large compared to k.

Lemma 7.4. *If X is a standard m by n matrix with $m > (k-1)^2$, then* $\mathbf{E}[\min_k(X)] = F_k(X)$.

Proof. By Lemmas 7.2 and 7.3, it is sufficient to prove that the statement holds when X has at least one zero in each row, and no column with k or more zeros. In this case each column can contain the leftmost zero of at most $k-1$ rows. Since there are more than $(k-1)^2$ rows, there must be at least k columns that contain the leftmost zero of some row. This implies that there is a zero cost k-assignment in X, that is, $\mathbf{E}[\min_k(X)] = 0$. Consequently, there is no $(k-1)$-cover, that is, $\mathcal{I}_k(X) = \emptyset$. It follows that $F_k(X) = 0$. ∎

Finally we prove that (11) holds also for smaller matrices by taking the limit as the weights of the exceeding rows tend to zero. Remarkably, this argument can be found (in a slightly different setting) in Theorems 5 and 6 of [4].

Theorem 7.5. *If X is a standard matrix, then* $\mathbf{E}[\min_k(X)] = F_k(X)$.

Proof. We prove this by downward induction on the number of rows. Suppose that X is a standard m by n matrix. Let X_ϵ be an augmented matrix of $m+1$ rows and n columns, so that the first m rows equal X, and the new row i_0 has no zeros and weight $\|i_0\| = \epsilon$. We can realize this by letting $x_1, \dots, x_n$ be exp(1) variables, each independent of all the others and of X, and letting $X_\epsilon(i_0, j) = x_j/\epsilon$. For every X and $x_1, \dots, x_n$, $\min_k(X_\epsilon)$ increases towards $\min_k(X)$ as $\epsilon \to 0$. By the principle of monotone convergence, it follows that

$$\mathbf{E}[\min_k(X)] = \lim_{\epsilon \to 0} \mathbf{E}[\min_k(X_\epsilon)].$$

We have to show that F has the same property, that is, that $F_k(X_\epsilon)$ converges to $F_k(X)$ as $\epsilon \to 0$. Let $R' = R \cup \{i_0\}$. Then

$$F_k(X_\epsilon) = \sum_{\substack{\alpha \subseteq R' \\ \beta \subseteq C \\ \alpha \cup \beta \in \mathcal{I}_k(X_\epsilon)}} \frac{Pr_\epsilon(\alpha) Pr(\beta)}{\|\overline{\alpha}\|_\epsilon \, \|\overline{\beta}\|}.$$

The subscripts indicate that $Pr(\alpha)$ and $\|\overline{\alpha}\|$ depend on ϵ. If $\alpha \subseteq R$, that is, α does not contain i_0, then $\alpha \cup \beta \in \mathcal{I}_k(X_\epsilon)$ if and only if $\alpha \cup \beta \in \mathcal{I}_k(X)$. As $\epsilon \to 0$, the probability that i_0 is the last row to be drawn in the urn process tends to 1. Therefore, $Pr_\epsilon(\alpha)$ converges to $Pr_R(\alpha)$ and $\|\overline{\alpha}\|_\epsilon$ converges to $\|\overline{\alpha}\|_R$. Hence, in the limit $\epsilon \to 0$, the contribution from the terms for which $i_0 \notin \alpha$ is $F_k(X)$. If, on the other hand, $i_0 \in \alpha$, then $Pr_\epsilon(\alpha) \to 0$ as $\epsilon \to 0$. Hence all the remaining terms tend to zero. This completes the proof. ∎

ACKNOWLEDGMENT

The author has had many interesting discussions with Svante Linusson. Svante Janson very carefully read and commented a draft of the paper.

References

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

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

[3] S. E. Alm and G. B. Sorkin, Exact expectations and distributions in the random assignment problem, Combin Probab Comput 11(3) (2002), 217–248.

[4] M. W. Buck, C. S. Chan, and D. P. Robbins, On the expected value of the minimum assignment, Random Structures Algorithms 21(1) (2002), 33–58.

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

[6] D. Coppersmith and G. B. Sorkin, On the expected incremental cost of a minimum assignment, Contemporary mathematics Ed. B. Bollobás, Vol. 10 of Bolyai Society Mathematical Studies. Springer, Berlin, 2002, pp. 277–288.

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

[8] R. M. Karp, An upper bound on the expected cost of an optimal assignment, Discrete Algorithms and Complexity: Proc Japan-U.S. Joint Sem, Academic Press, New York, 1987, pp. 1–4.

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

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

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

[12] L. Lovász and M. D. Plummer, Matching theory, North-Holland, Amsterdam, 1986.

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

[14] C. Nair, B. Prabhakar, and M. Sharma, Proofs of the Parisi and Coppersmith-Sorkin conjectures for the finite random assignment problem, Proc. IEEE FOCS, 2003, 168–178.

[15] B. Olin, Asymptotic properties of the random assignment problem, Ph.D. thesis, Kungl Tekniska Högskolan, Stockholm, Sweden, 1992.

[16] G. Parisi, A conjecture on random bipartite matching, arXiv:cond-mat/9801176, 1998.

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