Towards the Resolution of Coppersmith-Sorkin Conjectures

Chandra Nair

Dept. of Electrical Engineering
Stanford University
Stanford, CA 94305.
mchandra@stanford.edu

Abstract

A *k-matching* is a set of k elements of a matrix, no two of which belong to the same row or column. The minimum weight *k-matching* of an $m \times n$ matrix C is the *k-matching* whose entries have the smallest sum. Coppersmith and Sorkin conjectured that if C is generated by choosing each entry independently from the exponential distribution of rate 1, then the expected value of the weight of the minimum weight *k-matching* is given by an explicit formula, whose proof is largely unknown. In this paper we describe our efforts to prove the Coppersmith-Sorkin conjecture by identifying the terms in the explicit formula to be the mean values of certain random variables which are functions of the matrix elements. We further conjecture that the distributions of these random variables are pure exponentials. We have partial theoretical backing and some simulation evidence for these conjectures. In the process we also prove a general combinatorial lemma about matchings in matrices.

1 Introduction

Suppose there are n jobs and n machines and the cost of executing a job i on machine j is a random variable c_{ij}, drawn from the same distribution for all i, j. The *random assignment problem* consists of determining a one-to-one assignment of the jobs to the machines so as to minimize the total cost of performing all the jobs. Mathematically, given a cost matrix $C = [c_{ij}]$, the problem is to determine the assignment π which solves

$$A_n = \min_{\pi} \sum_{i=1}^{n} c_{i,\pi(i)}.$$

The random assignment problem with costs distributed as i.i.d. exp(1) is of particular interest due to the following beautiful conjecture of Parisi [7]:

$$E(A_n) = \sum_{i=1}^{n} \frac{1}{i^2}.$$

Coppersmith and Sorkin [4] have proposed a larger class of conjectures which state that the expected cost of the minimum k-assignment in an $m \times n$ matrix of i.i.d. exp(1) is:

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

Observe $F(n,n,n) = E(A_n)$ and these conjectures subsume Parisi's conjecture.

Other attempts at resolving these conjectures have been in the direction of enlarging the set of conjectures to extended settings where either the rates are other than 1 [3]; or where the matchings are considered over a broader class of cost matrices [6]. There has also been another recent attempt at solving Parisi's conjecture, [5], and our paper has roughly the same flavor.

Our attack goes back to just considering the original setting of i.i.d exp 1 entries but trying to understand the Coppersmith-Sorkin conjecture in a different light. We try to prove the Coppersmith-Sorkin conjecture using an induction on the matching size k. In Section 2, a sequence of matchings indexed by their weight and size are defined. Certain differences of this sequence are conjectured to be exponentially distributed and their means are precisely the terms that appear in the Coppersmith-Sorkin conjecture. Thus summing up the expected values of the differences provides a proof for the Coppersmith-Sorkin conjecture. By proving a general result about matchings in matrices, we establish that the random variables conjectured to be exponentially distributed are all positive. Section 3 deals with these conjectures we make about the distributions of the differences, their positivity and the distributions of some of them. Section 4 concludes by stating what is known from simulations and other methods.

2 Matchings in Matrices

Consider a $m \times n$ matrix, C, with $1 \leq m \leq n$. A *k-matching* is defined as a set of k elements such that no two lie in the same row or column. C is assumed to have i.i.d. exp(1) entries. For the construction below, since the eventual interest is in the expected values of minimum weights, we will assume that any two *k-matchings* have different weight. The case where some two have the same weight is a probability zero event and can be ignored.

Consider $1 \leq k < m$. T_k^1 is defined as the weight of the *k-matching* in C that has the minimum weight. This matching has elements belonging to k distinct rows. Now, consider all *k-matchings* that contain elements from exactly $k-1$ of these k rows. Let T_k^2 be the weight of the matching that has the smallest weight among these matchings. There are precisely $k-1$ common rows among the first two matchings. Again consider all *k-matchings* that contain elements from exactly $k-2$ of these $k-1$ common rows. Define T_k^3 to be the weight of the matching that has the smallest weight among these matchings. Suppose we have defined T_k^1, T_k^2, …, T_k^i for $2 \leq i \leq k$ with the property that these matchings have $k-i+1$ rows in common among them. Then consider all *k-matchings* that contain elements from exactly $k-i$ of these $k-i+1$ common rows. Define T_k^{i+1} to be the weight of the matching that has the smallest weight among these matchings. Thus we inductively define the random variables T_k^1, T_k^2, …, T_k^{k+1}. With an abuse of notation, T_k^i may also be used to refer to the *k-matching* that has the weight T_k^i.

Define T_m^1 to be weight of the the smallest matching of size m. Consider the sequence $T_1^1, T_1^2, T_2^1, T_2^2, T_2^3, T_3^1, \ldots, T_{m-1}^m, T_m^1$. Observe that this sequence starts with weights of two matchings of size one, then has the three matchings of size two and so on till we reach the smallest matching of size m. Also, observe that within the matchings of the same size the sequence has the lower numbered matching appearing before the higher numbered one. For example, T_4^2 and T_4^3 are two matchings of size four and T_4^2 appears before T_4^3. We conjecture that the difference between two consecutive random variables in the sequence are exponentially distributed but with different rates, establish some of the claims and prove that these differences are in fact positive in the next section.

3 Conjectures

Conjecture 1

$$\begin{aligned}
T_1^1 &\stackrel{d}{=} \exp mn \\
T_1^2 - T_1^1 &\stackrel{d}{=} \exp (m-1)n \\
T_2^1 - T_1^2 &\stackrel{d}{=} \exp m(n-1) \\
T_2^2 - T_2^1 &\stackrel{d}{=} \exp (m-2)n \\
T_2^3 - T_2^2 &\stackrel{d}{=} \exp (m-1)(n-1) \\
T_3^1 - T_2^3 &\stackrel{d}{=} \exp m(n-2) \\
&\;\;\vdots \\
T_k^{i+1} - T_k^i &\stackrel{d}{=} \exp (m-k+i-1)(n-i+1) \quad 1 \le i \le k < m \\
&\;\;\vdots \\
T_{k+1}^1 - T_k^{k+1} &\stackrel{d}{=} \exp m(n-k) \quad 1 \le k < m \\
&\;\;\vdots \\
T_{m-1}^m - T_{m-1}^{m-1} &\stackrel{d}{=} \exp (m-1)(n-m+2) \\
T_m^1 - T_{m-1}^m &\stackrel{d}{=} \exp m(n-m+1)
\end{aligned}$$

By $\exp j$, we refer to an exponential random variable of rate j or mean $\frac{1}{j}$. Assuming the conjectures are true, the following corollary is immediate.

Corollary 1 Coppersmith-Sorkin Conjecture

Proof T_k^1 is the weight of the minimum weight *k-matching* of the entire matrix, by definition.

$$T_k^1 = (T_k^1 - T_{k-1}^k) + (T_{k-1}^k - T_{k-1}^{k-1}) + \cdots + (T_2^1 - T_1^2) + (T_1^2 - T_1^1) + T_1^1.$$

Hence, by linearity of expectation and from the above conjectures on the distribution of these random variables, we obtain

$$\begin{aligned}
E(T_k^1) &= \frac{1}{m(n-k+1)} + \frac{1}{(m-1)(n-k+2)} + \cdots + \frac{1}{m(n-1)} + \frac{1}{(m-1)n} + \frac{1}{mn} \\
&= \sum_{\substack{i,j \ge 0 \\ i+j<k}} \frac{1}{(m-i)(n-j)}
\end{aligned}$$

It is well known that Coppersmith-Sorkin conjecture evaluated at $k = m = n$ gives the Parisi's conjecture. ∎

Lemma 1 *The following statements hold.*

(a) $T_1^1 \stackrel{d}{=} \exp mn$

(b) $T_1^2 - T_1^1 \stackrel{d}{=} \exp{(m-1)n}$

(c) $T_2^1 - T_1^2 \stackrel{d}{=} \exp{m(n-1)}$

Proof

(a) T_1^1 is the smallest element in the matrix, the minimum of mn i.i.d $\exp 1$ random variables and hence is distributed as $\exp mn$.

(b) T_1^2 is the smallest element in the matrix outside of the row that contains T_1^1. If C_{ij} is any element outside the row that contains T_1^1, then by the memory-less property of the exponential distribution $C_{ij} - T_1^1 \stackrel{d}{=} \exp 1$ and these increments are independent from one entry of this sub-matrix to another. By definition, $T_1^2 - T_1^1$ is the minimum of the increments of all the elements in the $(m-1) \times n$ sub-matrix obtained by removing the row of T_1^1. Hence it is the minimum of $(m-1)n$ independent $\exp 1$ random variables and therefore is distributed as $\exp{(m-1)n}$.

(c) The distribution of $T_2^1 - T_1^2$ can be computed as follows. In the event T_1^2 is in a different column as T_1^1, the smallest matching of size two is formed by the two elements T_1^1, T_1^2 and hence $T_2^1 = T_1^1 + T_1^2$.

In the event T_1^2 is in the same column as T_1^1, it is not hard to see from the memory-less property of exponential, that $T_2^1 = T_1^1 + T_1^2 + A$, where A is distributed as $\exp m(n-1)$ and is independent of T_1^1. To see this, we can assume without loss of generality that, T_1^1 was the $(1,1)$ element and T_1^2 was the $(2,1)$ element of the matrix C.

Consider the elements of the matrix in columns 2 to n. Now any element in the first row can be expressed as $T_1^1 + e_{1,i}, 2 \leq i \leq n$ and any element in the other rows can be expressed as $T_1^2 + e_{j,i}, 2 \leq j \leq m$ and $2 \leq i \leq n$. From the memory-less property of the exponential it is easy to see that these $e_{1,i}$'s and the $e_{j,i}$'s are exponentials of rate one and independent.

To form a matching of size two, we must take at least one element from the columns 2 to n. If one picks any element from the first row and columns 2 to n, then it will go with T_1^2 to form the smallest two matching. However, if one picks any element from the other rows and columns 2 to n, then it will go with T_1^1 to form the smallest two matching. Let $A = \min_{i \geq 2} e_{j,i}$. Then we see that $T_2^1 = T_1^1 + T_1^2 + A$ and A is distributed as $\exp m(n-1)$.

Combining these two events, we see that $T_2^1 - T_1^2$ is an $\exp mn$ with probability $\frac{1}{n}$ or sum of two independent exponentials of rate mn and $m(n-1)$ with probability $\frac{n-1}{n}$. The distribution of such a random variable is in fact $\exp m(n-1)$. ■

Theorem 1 *$T_k^2 - T_k^1 \stackrel{d}{=} \exp{(m-k)n}$.*

Proof The development of the proof is very similar to the one in Sharma and Prabhakar [5]. Let T_k^1 occupy the first k rows of the matrix, C. Consider the sub-matrix of size $(m-k) \times n$ formed by considering the $m-k$ rows below the first k rows. For any element c in this sub-matrix, denote by Δ_c, the weight of the smallest *(k-1)-matching* from the first k rows in the matrix C that does not have an element in the same column as c. Thus $c + \Delta_c$, will be the weight of the smallest *k-matching* that contains c and precisely

$k-1$ elements from the first k rows. First, we observe that $\Delta_c < T_k^1$ since T_k^1 has at least $k-1$ elements that are not present in the column of c and Δ_c has the option of choosing them. Observe, that the random variable, $T_k^1 - \Delta_c > 0$ is just a function of the random variables in first k rows. Therefore, from the memory-less property of the exponential we have, $c - (T_k^1 - \Delta_c)$ is an $\exp 1$ random variable and these increments are independent as we vary c. Now $T_k^2 - T_k^1 = \min_c c + \Delta_c - T_k^1$ and hence it is the minimum of $(m-k) \times n$ independent exponential random variables of rate 1. Therefore, $T_k^2 - T_k^1$ is distributed as a rate $(m-k)n$ random variable. ■

Theorem 2 $T_k^{i+1} - T_k^i > 0$ $a.s.$ $\forall\ 1 \leq k \leq m-1, 1 \leq i \leq k$.

Proof $T_k^2 - T_k^1 > 0$ because T_k^1 is defined as the weight of the smallest matching of size k. Consider all *k-matchings* that contain at most *k-1* elements from the k rows occupied by T_k^1. Let the weight of this matching be $\tilde{T}_k^2$. We claim that $\tilde{T}_k^2 = T_k^2$. This will follow if we establish that the smallest *k-matching* that contains at most *k-1* elements from the k rows occupied by T_k^1, has exactly *k-1* elements from the k rows.

Suppose not and that the *k-matching* that gives $\tilde{T}_k^2$ has two or more elements outside the k rows occupied by T_k^1. Now consider the bipartite graph with m vertices on the left and n vertices to the right. Lay down a red colored edge between vertex i on the left to the vertex j on the right, if T_k^1 contains the element (i, j). Similarly place a blue colored edge between vertex i on the left to the vertex j on the right, if $\tilde{T}_k^2$ contains the element (i, j). If we remove the common edges from the graph, it is easy to see that the edges form disjoint combinations of one or more of the following structures.

(a) a cycle of alternating colored edges

(b) a path of alternating colored edges of even length

(c) a path of alternating colored edges of odd length that has one more red edge than a blue edge.

(d) a path of alternating colored edges of odd length that has one more blue edge than a red edge.

It is also easy to see that since the total number of blue edges and red edges are same the number of structures of type (c) must equal the number of structures of type (d). So, one can in fact pair them up so that you have one structure (e) that contains two structures, one of type (c) and one of type (d).

If any structure of type (a) was present, from the fact that T_k^1 is the smallest matching of size k, the red edges in the cycle have a total weight smaller than the blue edges in the cycle. Hence, we could have replaced the blue edges with the red edges to get a lower weight $\tilde{T}_k^2$. Since by this procedure we are not changing the rows occupied by $\tilde{T}_k^2$, this gives the required contradiction.

If any structure of type (b) or type (e) was present, by similar argument we could get a lower weight matching by replacing the blue edges with the red edges. However, we would be changing the rows occupied by $\tilde{T}_k^2$ by precisely one row. Therefore, if $\tilde{T}_k^2$ differed from T_k^1 by at least two rows, then we could do this to get a matching that has lower weight and has at most *k-1* rows in common to T_k^1. This will contradict minimality of $\tilde{T}_k^2$. Hence, $\tilde{T}_k^2$ can differ from T_k^1 from at most one row. Thus the smallest matching that contains at most *k-1* elements from the k rows occupied by T_k^1, has exactly *k-1* elements from the k rows. This shows $\tilde{T}_k^2 = T_k^2$.

Since, $T_k^3, T_k^4, \ldots, T_k^{k+1}$ had at most *k-1* rows in common with T_k^1, this establishes that T_k^2 is smaller than all of them. The rest of the proof follows along similar lines. Suppose that T_k^i is the smallest *k-matching* that has at most *k-i+1* elements in common to the *k-i+2* common rows of $T_k^1, T_k^2, \ldots, T_k^{i-1}$.

Define $\tilde{T}_k^{i+1}$ as the smallest *k-matching* that has at most *k-i* elements in common to the *k-i+1* common rows of $T_k^1, T_k^2, \ldots, T_k^i$. By a very similar reasoning as above we can show that in the bipartite graph matching defined by $\tilde{T}_k^{i+1}$ and T_k^i, we cannot have a cycle of alternating colored edges. Suppose the edges of T_k^i were colored red and the edges of $\tilde{T}_k^{i+1}$ were colored blue and in the cycle if the sum of red edges were smaller that that of blue edges, then we can replace the blue edges by red within the cycle and can get a smaller matching than $\tilde{T}_k^{i+1}$. On the other hand if the sum of blue edges were smaller that that of red edges, then we could get a smaller T_k^i by replacing the red edges with the blue edges of the cycle.

If $\tilde{T}_k^{i+1}$ has two or more rows different from T_k^i and if two or more structures of type (c) or (e) were present then it would contradict the minimality of T_k^i or $\tilde{T}_k^{i+1}$, depending on the case if the total weight of blue edges in the structure were lighter or heavier than the red edges in the structure. One has to be a bit careful here in the argument. T_k^i avoids exactly one of the $k-i+2$ common rows of $T_k^1, T_k^2, \ldots, T_k^{i-1}$. Call this the forbidden row. Since there are two or more structures of type (c) or (e), we can always pick one that does not contain the forbidden row and get the contradiction.

The only remaining case is when exactly one structure either of type (c) or type (e) is present. But then observe that the rows of $\tilde{T}_k^{i+1}$ can differ from the rows of T_k^i by exactly one. This shows that $\tilde{T}_k^{i+1}$ is in fact the same as T_k^{i+1} since it contains exactly *k-i* elements in common to the *k-i+1* common rows of $T_k^1, T_k^2, \ldots, T_k^i$. Since $T_k^{i+2}, \ldots, T_k^{k+1}$ had at most *k-i* elements in common to the *k-i+1* common rows of $T_k^1, T_k^2, \ldots, T_k^i$, this establishes that T_k^{i+1} is smaller than $T_k^{i+2}, \ldots, T_k^{k+1}$. This completes the proof of the Theorem. ■

Remark: We have assumed throughout the proof that the weight of any two different set of edges will always be different. Since these weights come from an exponential distrtibution, the cases where there is a tie has probability zero and hence can be ignored for computing the distributions. This remark also holds for the lemma below.

Lemma 2 $T_{k+1}^1 > T_k^{k+1}$ *a.s.* $\forall\ 1 \leq k \leq m-1$

Proof By definition, there is one common row for the matchings $T_k^1, \ldots, T_k^k$. T_k^{k+1} is defined as the weight of the smallest *k-matching* that avoids the common row. Since T_{k+1}^1 is a matching of size *k+1*, it has at least a matching of size *k* outside any row and hence T_k^{k+1} could have chosen a *k-matching* from the elements of T_{k+1}^1. This establishes the claim. (Note that the elements of the matrix are positive and hence any smaller subset always has smaller weight.) ■

It may be of independent interest to note that T_{k+1}^1 in fact contains all the rows occupied by T_k^1. The proof is omitted here.

We have thus shown using Theorem 2 and Lemma 2 that all the random variables that we have conjectured to be exponentially distributed are positive *a.s.*

4 Conclusion

Simulations have been performed for matrices up to size $m, n \leq 5$ and they "verify" the conjectures. A case by case analysis was theoretically performed for the 3×3 and the

3×4 matrices and the conjectures were verified. This was however done by an exhaustive case analysis and does not contribute to the general understanding or the solvability of the conjecture. Hence, this has been omitted from the paper. I would also like to bring to the attention of the interested reader that exponentiality of these increments seem to be a very general trend if we are to believe the simulations. For example, fix any k and choose the smallest weight matchings in all the $\binom{m}{k}$ ways of picking k distinct rows. Order them in the increasing order as $S_k^1, ..., S_k^{\binom{m}{k}}$. It appears that $S_k^2 - S_k^1, \ldots, S_k^{k+1} - S_k^k$ are also exponentially distributed. Note that, these random variables are different from $T_k^1, T_k^2, \ldots, T_k^{k+1}$, though $T_k^1 = S_k^1$ and $T_k^2 = S_k^2$.

There is also an interesting duality between the rows and columns. We could get a different (but similar) set of conjectures if we interchange the role of the columns and rows, and these will also imply the Coppersmith and Sorkin conjectures.

5 Acknowledgment

The author wishes to thank Mayank Sharma and Balaji Prabhakar for several useful discussions and suggestions.

References

[1] D. J. Aldous. Asymptotics in the random assignment problem. *Probab. Th. Rel. Fields*, 93 (1992) 507-534.

[2] D. J. Aldous. The zeta (2) Limit in the Random Assignment Problem. *Random Structures and Algorithms* 18 (2001) 381-418.

[3] Marshall W. Buck, Clara S. Chan, David P. Robbins. On the Expected Value of the Minimum Assignment. *Random Structures and Algorithms* 21 (2002) 33-58.

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

[5] M. Sharma and B. Prabhakar. On Parisi's Conjecture for the Finite Random Assignment Problem. *submitted to Allerton* 2002.

[6] S. Linusson and J. Wästlund. A generalization of the random assignment problem, Preprint 2000, http://arXiv.org/abs/math.CO/0006146.

[7] G. Parisi. A conjecture on random bipartite matching. Technical Report cond-mat/9801176, XXX LANL Archive, 1998