# On Parisi's Conjecture for the Finite Random Assignment Problem


Mayank Sharma
Dept. of EE
Stanford University
msharma@stanford.edu

Balaji Prabhakar
Depts. of EE and CS
Stanford University
balaji@stanford.edu



**Abstract**

The classical random assignment problem has received a lot of interest in the recent literature, mainly due to the following pleasing conjecture of Parisi: The expected value of the minimum-cost permutation in an $n \times n$ matrix with i.i.d. exp(1) entries equals $\sum_{i=1}^{n} \frac{1}{i^2}$. This crisp conjecture appears difficult to resolve despite the recent proof of Aldous [Al 01] that in the limit as $n \to \infty$, the expected minimum cost equals $\pi^2/6$. Nevertheless, the simplicity of the finite conjecture and of the asymptotic result suggest that there are interesting structural properties of matchings worth exploring. In this paper, we present some such properties and discuss their role in resolving Parisi's conjecture. For example, our analysis has led us to formulate a simpler conjecture which implies Parisi's conjecture, and highlights the role played by the exp(1) costs. We provide various verifications of our conjecture using analysis and simulation. Our approach also yields combinatorial results regarding the structure of minimum-cost permutations of size $\min\{m, n\}$ in rectangular matrices of dimensions $m \times n$.


## 1 Introduction

The *assignment problem* is described as follows: Suppose there are $n$ jobs and $n$ machines and it costs $c_{ij}$ to execute job $i$ on machine $j$. The 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. More formally, given a cost matrix $C = [c_{ij}]$, the problem is to determine the assignment $\pi$ which solves

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

In the *random* assignment problem the $c_{ij}$ are i.i.d. random variables drawn from some distribution. An equivalent formulation of the random assignment problem in graph-theoretic terms is this: Find the minimum-weight perfect matching in a complete $n \times n$ bipartite graph with i.i.d. edges weights $c_{ij}$. A quantity of interest in the random assignment problem is the expected minimum cost, $E(A_n)$, and its asymptotic value. Assuming, for now, that $\lim_n E(A_n)$ exists, let us denote it by $A^*$.

There has been a lot of work on determining bounds for the expected minimum cost. We survey some of the work; more details can be found in [St 97, CS 99]. Early work uses feasible solutions to the dual linear programming (LP) formulation of the assignment problem for obtaining the following lower bounds for $A^*$: $(1 + 1/e)$ by Lazarus [La 93], 1.441 by Goemans and Kodialam [GK 93], and 1.51 by Olin [Ol 92]. The first upper bound of 3 was given by Walkup [Wa 79], who thus demonstrated that $\limsup_n E(A_n)$ is finite. Walkup's argument was later made constructive by Karp *et al* [KKV 94]. Karp [Ka 84, Ka 87] made a subtle use of LP duality to obtain a better upper bound of 2. Coppersmith and Sorkin [CS 99] have further improved the bound to 1.94.

Meanwhile, it had been observed through simulations that for large $n$, $E(A_n) \approx 1.642$ [BKMP 86]. Mézard and Parisi [MP 87] used the *replica method* [MPV 87] of statistical physics to argue that $A^* = \frac{\pi^2}{6}$. More interestingly, their method allowed them to determine the density of the edge-weight distribution of the limiting optimal matching. These sharp (but non-rigorous) asymptotic results, and others of a similar flavor that they obtained in several combinatorial optimization problems, sparked interest in the replica method and in the random assignment problem.

Aldous [Al 92] proved that $A^*$ exists by identifying the limit as the value of a minimum-cost matching problem on a certain random weighted infinite tree. In the same work he also established that the distribution of $c_{ij}$ affects $A^*$ only through the value of its density function at 0 (provided it exists and is strictly positive). Thus, as far as the value of $A^*$ is concerned, the distributions $U[0,1]$ and $\exp(1)$ are equivalent. More recently, Aldous [Al 01] has established that $A^* = \pi^2/6$, and obtained the same limiting optimal edge-weight distribution as [MP 87]. He also obtains a number of other interesting results such as the asymptotic essential uniqueness (AEU) property - which roughly states that almost-optimal matchings have almost all their edges equal to those of the optimal matching.

The assignment problem with costs distributed as i.i.d. $\exp(1)$ continues to be of particular interest due to the following beautiful conjecture of Parisi [Pa 98]:

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

Note that this is an elegant restriction (seemingly true only for i.i.d. $\exp(1)$ costs) of the asymptotic result, in that $E(A_n)$ coincides with Euler's expansion for $\pi^2/6$ up to $n$ terms. Coppersmith and Sorkin [CS 99] have proposed a larger class of conjectures which state that the expected cost of of the minimum $k$-assignment in an $m \times n$ matrix of i.i.d. $\exp(1)$ is:

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

By definition, $F(n,n,n) = E(A_n)$ and their expression coincides with Parisi's conjecture.

## 1.1 Outline of the paper

Parisi's conjecture motivates the work in this paper, and we shall assume hereafter that the $c_{ij}$ are i.i.d. $\exp(1)$. Our development allows us to arrive at a simpler conjecture whose verification implies Parisi's conjecture and highlights the role of the exponential-cost distribution. We provide various verifications (proofs and simulations) of our conjecture.

While our conjecture and the development leading up to it are probabilistic in nature, our approach also yields some interesting combinatorial results concerning minimum-weight matchings in rectangular matrices. For instance, we show that in an $n \times (n+m)$

matrix precisely $(m+1)n$ elements participate in minimum-cost matchings of size $n$, and they belong to a class of canonical templates which we identify and characterize. In the special case when $m = 1$, we obtain a procedure for enumerating all such placements of $2n$ elements up to row and column permutations.

Section 2 contains the probabilistic analysis, Section 3 the combinatorial parts and Section 4 concludes the paper. Throughout the writing of this paper, we have been conscious of the constraint on space, and have therefore provided proofs only for results that are essential to the exposition.

# 2 Probabilistic Analysis

## 2.1 Preliminaries

Denote by $\tilde{C}$ the sub-matrix obtained by removing the first row of $C$. For each $i$, $i = 1, \ldots, n$, let $S_i$ be the cost of the minimum-cost permutation in the sub-matrix of $\tilde{C}$ obtained by deleting it's $i^{th}$ column. These quantities are illustrated taking $\tilde{C}$ to be the following $2 \times 3$ matrix.

$$\tilde{C}: \begin{array}{|c|c|c|}\hline 3 & 6 & 11 \\ \hline 9 & 2 & 20 \\ \hline \end{array} \rightarrow \begin{array}{|c|c|}\hline 6 & \mathbf{11} \\ \hline \mathbf{2} & 20 \\ \hline \end{array} \Rightarrow S_1 = 13; \quad \begin{array}{|c|c|}\hline 3 & \mathbf{11} \\ \hline \mathbf{9} & 20 \\ \hline \end{array} \Rightarrow S_2 = 20; \quad \begin{array}{|c|c|}\hline \mathbf{3} & 6 \\ \hline 9 & \mathbf{2} \\ \hline \end{array} \Rightarrow S_3 = 5.$$

Denote the ordered sequence of $S_i$ by $T_i$, $i = 1, \ldots, n$. That is, let $\sigma$ be the random permutation of $\{1, \ldots, n\}$ such that $S_{\sigma(1)} \leq \ldots \leq S_{\sigma(n)}$ a.s. Then $T_i = S_{\sigma(i)}$. In the above example, $T_1 = 5$, $T_2 = 13$ and $T_3 = 20$.

Now consider $E(A_n)$, abbreviated to $E_n$.

$$\begin{aligned} E_n &= \int_0^\infty P(A_n > x)\, dx = \int_0^\infty P(A_n(\pi) > x, \forall \pi)\, dx \\ &= \int_0^\infty P(c_{1j} > x - S_j, \forall j)\, dx = E_{\tilde{C}} \left( \int_0^\infty P(c_{1j} > x - s_j, \forall j)\, dx \mid \tilde{C} \right) \end{aligned}$$

where $E_{\tilde{C}}(.|\tilde{C})$ denotes the conditional expectation with respect to the matrix $\tilde{C}$. Note that conditioned on $\tilde{C}$, the $S_i$'s are constant and are therefore denoted by $s_i$. Next, consider

$$\begin{aligned} I &= \int_0^\infty P(c_{1j} > x - s_j, \forall j)\, dx = \int_0^\infty \prod_{j=1}^n P(c_{1j} > x - s_j)\, dx \qquad \text{(independence of } c_{1j}) \\ &= \int_0^{t_1} dx + \int_{t_1}^{t_2} e^{-(x-t_1)}\, dx + \ldots + \int_{t_{n-1}}^{t_n} e^{-((n-1)x - t_1 - \cdots - t_{n-1})}\, dx + \int_{t_n}^\infty e^{-(nx - t_1 - \cdots - t_n)}\, dx \\ &\qquad \text{(where the } t_i \text{ are obtained by ordering the } s_i) \\ &= t_1 + 1 - \frac{1}{2} e^{-(t_2 - t_1)} - \frac{1}{6} e^{-(2t_3 - t_1 - t_2)} - \cdots - \frac{1}{n(n-1)} e^{-((n-1)t_n - t_1 - \cdots - t_{n-1})}. \end{aligned}$$

Therefore,

$$\begin{aligned} E_n &= E(T_1) + 1 - \sum_{i=2}^n \frac{1}{i(i-1)} E\left(e^{-((i-1)T_i - T_1 - \cdots - T_{i-1})}\right) \\ &= E(T_1) + 1 - \sum_{i=2}^n \frac{1}{i(i-1)} E\left(e^{\sum_{j=1}^{i-1} -j(T_{j+1} - T_j)}\right) \qquad (1) \end{aligned}$$

To proceed further we make the following

**Conjecture 1** *For $j = 1, \ldots, n-1$, $T_{j+1} - T_j \sim \exp(j(n-j))$ and these increments are independent of each other.*

We will comment on the validity of this conjecture later. For now, assuming that it is true, we obtain

$$E(e^{-j(T_{j+1}-T_j)}) = \frac{n-j}{n-j+1}.$$

Therefore,

$$E\left(e^{\sum_{j=1}^{i-1} -j(T_{j+1}-T_j)}\right) = \prod_{j=1}^{i-1} E\left(e^{-j(T_{j+1}-T_j)}\right) = \prod_{j=1}^{i-1} \frac{n-j}{n-j+1} = \frac{n-i+1}{n}.$$

Substituting this in (1) gives

$$E_n = E(T_1) + \frac{1}{n^2} + \frac{1}{n}\sum_{i=1}^{n-1} \frac{1}{i}. \quad (2)$$

We are left with having to evaluate $E(T_1)$. It turns out that Conjecture 1 proves useful in doing this as well. That is, assuming that Conjecture 1 holds and by inductively assuming that $E_{n-1} = \sum_{i=1}^{n-1} \frac{1}{i^2}$, it is possible to eliminate $E(T_1)$ at (2) as follows.

The random variables $S_1, \ldots, S_n$ are all distributed as $A_{n-1}$. Therefore, $E(S_i) = E_{n-1}$, $\forall\ i$. Further, since $S_1$ is one of the $T_i$'s chosen uniformly at random, it is straightforward to note that $E(S_1) = \frac{1}{n}\sum_{j=1}^{n} E(T_j)$. But for $j = 2, \ldots, n$,

$$\begin{aligned} E(T_j) &= E(T_1) + \sum_{k=2}^{j} (E(T_k) - E(T_{k-1})) \\ &= E(T_1) + \sum_{k=2}^{j} \frac{1}{(k-1)(n-k+1)} \quad \text{(by Conjecture 1).} \end{aligned}$$

Thus,

$$\begin{aligned} E(S_1) &= \frac{1}{n}\sum_{j=1}^{n} E(T_j) = \frac{1}{n}\sum_{j=1}^{n}\left(E(T_1) + \sum_{k=2}^{j} \frac{1}{(k-1)(n-k+1)}\right) \\ &= E(T_1) + \frac{1}{n}\sum_{j=1}^{n-1} \frac{1}{j}. \quad (3) \end{aligned}$$

Assuming, by induction, that $E_{n-1} = \sum_{j=1}^{n-1} \frac{1}{j^2} = E(S_1)$, we obtain from (3):

$$E(T_1) = \sum_{j=1}^{n-1} \left(\frac{1}{j^2} - \frac{1}{nj}\right). \quad (4)$$

To conclude the induction step and establish Parisi's conjecture, we substitute $E(T_1)$ obtained above at (4) in (2) to get

$$E_n = \sum_{i=1}^{n-1} \left(\frac{1}{i^2} - \frac{1}{ni}\right) + \frac{1}{n^2} + \frac{1}{n}\sum_{i=1}^{n-1} \frac{1}{i} = \sum_{i=1}^{n} \frac{1}{i^2}. \quad (5)$$

## 2.2 Discussion of Conjecture 1

In this section we provide various indications of the validity of Conjecture 1; some proofs and some simulations.

### 2.2.1 Exhaustive analysis for n=3

Consider the $2 \times 3$ matrix below and assume that the elements of $T_1, T_2$ and $T_3$ are chosen from the set $\{c_{11}, c_{12}, c_{22}, c_{23}\}$. (Theorem 2 shows that, up to row and column permutations, there are exactly 6 such statistically equivalent choices which admit the following common evaluation.)

| | | |
|---|---|---|
| $\mathbf{c_{11}}$ | $\mathbf{c_{12}}$ | $c_{13}$ |
| $c_{21}$ | $\mathbf{c_{22}}$ | $\mathbf{c_{23}}$ |

A tabulation of the cases, and probabilities conditioned on each case:

| Case | $T_1$ | $T_2$ | $T_3$ | $P(\mathcal{E}_i)$ | $P(T_2 - T_1 > t \mid \mathcal{E}_i)$ | $P(T_3 - T_2 > t \mid \mathcal{E}_i)$ |
|---|---|---|---|---|---|---|
| $\mathcal{E}_1$ | $c_{11} + c_{22}$ | $c_{12} + c_{23}$ | $c_{11} + c_{23}$ | $\frac{1}{72}$ | $e^{-2t}$ | $e^{-4t}$ |
| $\mathcal{E}_2$ | $c_{12} + c_{23}$ | $c_{11} + c_{22}$ | $c_{11} + c_{23}$ | $\frac{1}{72}$ | $e^{-2t}$ | $e^{-4t}$ |
| $\mathcal{E}_3$ | $c_{11} + c_{22}$ | $c_{11} + c_{23}$ | $c_{12} + c_{23}$ | $\frac{5}{144}$ | $e^{-2t}$ | $\frac{6}{5}(e^{-2t} - \frac{e^{-4t}}{6})$ |
| $\mathcal{E}_4$ | $c_{11} + c_{23}$ | $c_{11} + c_{22}$ | $c_{12} + c_{23}$ | $\frac{5}{144}$ | $e^{-2t}$ | $\frac{6}{5}(e^{-2t} - \frac{e^{-4t}}{6})$ |
| $\mathcal{E}_5$ | $c_{12} + c_{23}$ | $c_{11} + c_{23}$ | $c_{11} + c_{22}$ | $\frac{5}{144}$ | $e^{-2t}$ | $\frac{6}{5}(e^{-2t} - \frac{e^{-4t}}{6})$ |
| $\mathcal{E}_6$ | $c_{11} + c_{23}$ | $c_{12} + c_{23}$ | $c_{11} + c_{22}$ | $\frac{5}{144}$ | $e^{-2t}$ | $\frac{6}{5}(e^{-2t} - \frac{e^{-4t}}{6})$ |

Therefore,

$$P(T_2 - T_1 > t) = 6 \left( \left( \frac{2}{72} + 4.\frac{5}{144} \right) e^{-2t} \right) = e^{-2t}.$$

$$P(T_3 - T_2 > t) = 6 \left( \frac{2}{72} e^{-4t} + 4.\frac{5}{144}\, \frac{6}{5} \left( e^{-2t} - \frac{e^{-4t}}{6} \right) \right) = e^{-2t}.$$

Moreover, we have verified that $T_2 - T_1$ is independent of $T_3 - T_2$. We have also verified analytically that the conjecture holds for the case when $n = 4$.

### 2.2.2 The distribution of $T_2 - T_1$

**Theorem 1** $T_2 - T_1 \sim \exp(n-1)$.

**Proof** For $i = 1, \ldots, n$ let $\mathbf{c}_i$ be the column such that $T_i$ is the cost of the minimum-cost matching on $\tilde{C} - \mathbf{c}_i \equiv M_i$. Then we make the following

**Claim 1** *The elements of* $\mathbf{c}_1$ *are completely specified by the conditions:*

$$c_{j1} + \Delta(c_{j1}) \geq T_1 \ , \ j = 1, \ldots, n-1, \tag{6}$$

*where for an element of column* $\mathbf{c}_1$ *say* $x$, $\Delta(x)$ *is defined as the smallest* $n-2 \times n-2$ *minor of* $x$ *drawn from* $\tilde{C} - \mathbf{c}_1$. *Further,* $T_1 - \Delta(c_{j1}) \geq 0 \ , \ j = 1, \ldots, n-1.$

Before establishing this claim, consider the following $4 \times 5$ matrix, and let $\mathbf{c_1}$ be the first column. Without loss of generality, let the dots represent the elements of $T_1$. Then the crosses represent a possible minor for $c_{11}$.

| | | | | |
|---|---|---|---|---|
| $c_{11}$ | $\bullet$ | | | |
| $c_{21}$ | | $\bullet$ | $\times$ | |
| $c_{31}$ | | $\times$ | $\bullet$ | |
| $c_{41}$ | $\times$ | | | $\bullet$ |

*Proof of Claim:* The conditions at (6) are necessary because $T_1$ is the smallest permutation in the whole matrix. They are sufficient because if any other column satisfies all the conditions, then the sub-matrix obtained by deleting it will contain $T_1$ as its minimum-cost permutation.

Now $T_1 - \Delta(c_{j1}) \geq 0$ , $j = 1, \ldots, n-1$ because $T_1$ contains an $n-2 \times n-2$ minor that could have been chosen by $c_{j1}$ and hence $\Delta(c_{j1})$ could only be smaller than this minor and hence smaller than $T_1$. ■

**Fact 1** *Suppose $X$ and $Y$ are independent, exp(1), and independent of $Z \geq 0$. Then on the event $\{X \geq Z, Y \geq Z\}$, $X - Z$ and $Y - Z$ are independent, exp(1), and independent of $Z$.*

Now (6) can be written as $c_{j1} \geq T_1 - \Delta(c_{j1}) \geq 0$ , $j = 1, \ldots, n-1$. This and Fact 1 imply that the variables

$$c_{j1} + \Delta(c_{j1}) - T_1 \sim \text{ exp}(1) \ , \ j = 1, \ldots, n-1$$

and are mutually independent (since the $c_{j1}$ are all independent). But notice that

$$T_2 - T_1 = \min_{j \leq n-1} \{c_{j1} + \Delta(c_{j1}) - T_1\}.$$

Thus $T_2 - T_1 \sim$ exp$(n-1)$, being the minimum of $n-1$ independent exp(1) variables. ■ Incidentally, we can also conclude that $T_2 - T_1$ independent of all the elements of $M_1$.

**Remark**: An analogous procedure can be used to calculate the other increments but, unfortunately, it does not give distributional statements . For example, define for an element from column $\mathbf{c}_i$, say $x$, the quantity $\Delta(x, \mathbf{c}_j)$ to be the smallest $n-2 \times n-2$ minor that is drawn from $\tilde{C} - \mathbf{c}_i$ but which contains an element from the column $\mathbf{c}_j$. Then with $i, j = 1, 2, \ i \neq j$

$$T_3 - T_2 = \min_{1 \leq k \leq n-2} \{c_{ki} + \Delta(c_{ki}, \mathbf{c}_j) - T_2\}.$$

But even though we are minimizing over $2(n-2)$ variables which might lead us to believe that the result might be an exp$(2(n-2))$ r.v., unlike in the case of $T_2 - T_1$, the individual r.v.s are neither exponential nor independent. So this approach does not extend for determining the distribution of $T_3 - T_2$ or for the other increments.

### 2.2.3 $T_3 - T_2$ is independent of $T_2 - T_1$

We will only give a sketch of the proof here: assume without loss of generality that $\arg\min_j \{c_{j1} + \Delta(c_{j1}) - T_1\} = 1$. Then, for $j \geq 2$, $\{c_{j1} + \Delta(c_{j1}) - T_2\} \sim$ exp(1); and is independent of $T_2 - T_1$. But we show that $T_3 - T_2$ is a function of only the elements of $M_1$ and the r.v's $c_{j1} + \Delta(c_{j1}) - T_2$ , $j \geq 2$; all of which are independent of $T_2 - T_1$. ■

### 2.2.4 An equivalence

We have obtained the expression for $E(T_1)$ at (4) using Conjecture 1:

$$E(T_1) = \sum_{j=1}^{n-1} \left( \frac{1}{j^2} - \frac{1}{nj} \right). \tag{7}$$

Coppersmith and Sorkin [CS 99] have separately conjectured that (see Section 1 for the definition of $F(m,n,k)$ )

$$E(T_1) = F(n-1,n,n-1) = \sum_{i,j\geq 0, i+j<n-1} \frac{1}{(n-1-i)(n-j)}. \tag{8}$$

A simple induction verifies the equivalence of (7) and (8).

### 2.2.5 Simulations

Figure 1 (a)-(d) display the distibutional fit of the increments $T_{j+1} - T_j \sim \exp(j(n-j))$, $j = 1,\ldots,n-1$ when $n = 5$. Subplots (e) and (f) show the agreement between the distributions of the sums of increments $T_5 - T_3 = (T_5 - T_4) + (T_4 - T_3) \sim \exp(4) + \exp(6)$, $T_4 - T_2 = (T_4 - T_3) + (T_3 - T_2) \sim \exp(6) + \exp(6)$ where the sums on the right hand side are those of independent exponentials of appropriate rates. In addition, we have evaluated moments of various orders and have found agreement with Conjecture 1.

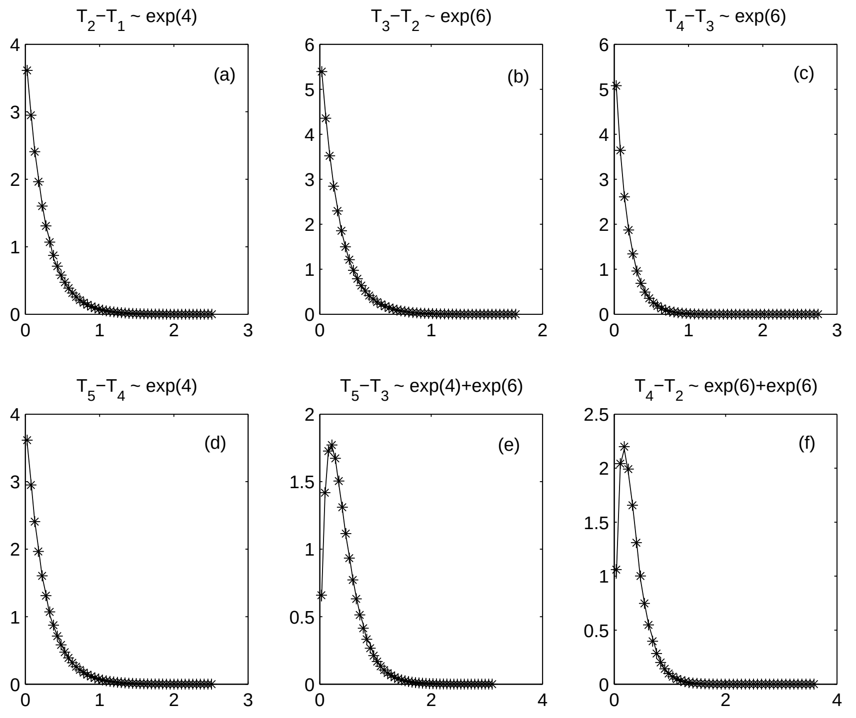


Figure 1: Distribution of increments and their sums for $n = 5$

# 3 Combinatorics

In the previous section we considered the $n-1 \times n$ matrix $\tilde{C}$ and studied probabilistic properties of the costs of the minimum-cost permutations in the square submatrices obtained by deleting a column at a time. Besides their cost, an interesting aspect of these minimum-cost permutations is the combinatorial structure they induce on $\tilde{C}$.

In this section we consider the following combinatorial problem: Consider an $n \times (n+m)$ matrix. Repeatedly delete $m$ columns at a time and mark the elements that constitute the minimum[1]-cost permutation on the remaining $n \times n$ matrix. Of interest is the number of entries marked in each row and in each column. The special case of matrices with the dimensions of $\tilde{C}$ corresponds to setting $m = 1$ and is considered first.

## 3.1 Notation and Results

Consider an $n \times n+1$ cost matrix $M \equiv [m_{ij}]$. Denote by $M_k, 1 \leq k \leq n+1$, the $n \times n$ matrix obtained by deleting the $k^{th}$ column of $M$, and let $\sigma_k(.) : \{1, \ldots, n\} \rightarrow \{1, \ldots, n+1\} - \{k\}$ be the minimum-cost permutation[2] in $M_k$. Further, let $N_i := \{\sigma_k(i) : 1 \leq k \leq n+1\}$ be the number of entries of row $i$ that participate in some minimum-cost permutation. A trivial observation is that $|N_i| \geq 2, \ \forall i$. But we state the following somewhat surprising theorem:

**Theorem 2** $|N_i| = 2, \ \forall \ i.$

**Proof** Provided in complete version of the paper.

An example $4 \times 5$ matrix of exp(1) r.v's is reproduced below in which the elements in bold are those that participate in the minimum-permutations in the five $4 \times 4$ submatrices. One can see that exactly two elements per row are marked. If with each marked element we identify a dot then we shall call the templates thus produced *two-dot patterns.*

| | | | | |
|---|---|---|---|---|
| 3.0691 | 0.3962 | **0.3240** | **0.1273** | 1.0787 |
| 0.1875 | 0.7545 | **0.3688** | 3.3028 | **0.3609** |
| **0.0433** | **0.0804** | 0.8464 | 0.2590 | 1.8943 |
| **0.3182** | 0.4228 | 2.5666 | 2.1152 | **0.5195** |

dot representation $\longrightarrow$

| | | | | |
|---|---|---|---|---|
| | | • | • | |
| | | • | | • |
| • | • | | | |
| • | | | | • |

This result generalizes to $n \times n+m, \ m = 0, 1, \ldots$ cost matrices $M$ in the following manner: Let $M_{k_1,\ldots,k_m}$ denote the matrix obtained by deleting the $m$ columns $k_1, \ldots, k_m$ from $M$, and let $\sigma_{k_1,\ldots,k_m}(.) : \{1, \ldots, n\} \rightarrow \{1, \ldots, n+m\} - \{k_1, \ldots, k_m\}$ be the minimum-cost permutation in $M_{k_1,\ldots,k_m}$. Write $N_i := \{\sigma_{k_1,\ldots,k_m}(i) : 1 \leq k_1 < \ldots < k_m \leq n+m\}$. Clearly, $|N_i| \geq m+1, \ \forall i$. We state without proof the following:

**Theorem 3** $|N_i| = m+1,$ *for* $m = 1, 2, \ldots; \ \forall \ i.$

**Proof** Provided in complete version of the paper.

## 3.2 Characterization of two-dot patterns

Consider an $n-1 \times n$ matrix of real numbers. Calculate the minimum weight permutations of the $n$ submatrices obtained by deleting one column at a time (call this procedure MinSubPerm). From Theorem 2 we know that if we mark all the elements involved in

[1]The same results apply if we consider the maximum-cost permutation

[2]Throughout we will break ties between equally-weighted permutations with common domain and range in favor of one of them consistently.

the above permutations, then exactly 2 elements per row will be marked and we will get a two-dot pattern (TDP).

The purpose of this section is to study certain properties of and thus characterize TDPs obtainable by the application of MinSubPerm to a matrix. Call the class of these patterns MinTDP$(n-1,n)$. Simply by using this definition it is not clear how to enumerate the patterns that belong to MinTDP$(n-1,n)$. We wish therefore to propose a constructive characterization of such patterns.

**Procedure**: Define SumTDP$(n-1,n)$ to be the class of patterns obtained by finding an assignment of non-negative integers to an $n-1\times n$ matrix such that:

- There are exactly two non-zero entries per row.
- The entries of each row and column sum to $n$ and $n-1$ respectively.

**Theorem 4** *: MinTDP$(n-1,n)$ = SumTDP$(n-1,n)$.*

**Proof** We only prove MinTDP$(n-1,n) \subseteq$ SumTDP$(n-1,n)$ here. The proof of the other direction is more involved and is provided in the complete version of the paper. To show that every pattern in MinTDP$(n-1,n)$ also belongs to SumTDP$(n-1,n)$ assign to each dot a value equal to the number of permutations produced by MinSubPerm, that it is a part of. Then the pattern of numbers will satisfy the properties of an element of SumTDP$(n-1,n)$, because there will be exactly two non-zero entries per row, atleast one non-zero entry per column; further since each row is part of all $n$ permutations and each column a part of $n-1$ permutations, all rows and columns sum to $n$ and $n-1$ respectively. Hence proved. ■

As an example, the following figure depicts the three possible two-dot patterns for a $4\times 5$ matrix (upto row and column permutations).

| | | | | |
|---|---|---|---|---|
| 1 | 4 | | | |
| 1 | | 4 | | |
| 1 | | | 4 | |
| 1 | | | | 4 |

| | | | | |
|---|---|---|---|---|
| 4 | 1 | | | |
| | 3 | 2 | | |
| | | 2 | 3 | |
| | | | 1 | 4 |

| | | | | |
|---|---|---|---|---|
| 4 | | | 1 | |
| | 4 | | 1 | |
| | | 4 | | 1 |
| | | | 2 | 3 |

# 4 Conclusions

We have considered the average-case analysis of the random assignment problem, motivated by a conjecture of Parisi. The main contribution is a refined conjecture which, we believe, is more tractable since it reduces Parisi's conjecture to proving that certain quantities are distributed as independent exponentials. Our analysis has also led to the discovery that the elements which participate in minimum-weight matchings in rectangular matrices belong to a class of canonical templates. A procedure for enumerating these templates for $n\times(n+1)$ matrices is provided. When edge-weights are random, we believe that a small number of these canonical templates, which share a substantial fraction of elements, become highly probable. This suggests the following connection to Aldous' AEU property: As $n\to\infty$, the minimum-weight matching utilizes one of these highly probable templates; therefore, almost-optimal matchings share almost all of their edges with the optimal matching.

# References


[Al 92] D. J. Aldous. Asymptotics in the random assignment problem. *Probab. Th. Rel. Fields*, 93 (1992) 507-534.

[Al 01] D. J. Aldous. The zeta (2) Limit in the Random Assignment Problem. *Random Structures and Algorithms* 18 (2001) 381-418.

[BKMP 86] Brunetti, Krauth, Mézard and Parisi (Unpublished).

[CS 99] D. Coppersmith and G. Sorkin. Constructive bounds and exact expectations for the random assignment problem. *Random Structures and Algorithms*, 15 (1999) 113-144.

[GK 93] 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.

[Ka 84] R. M. Karp. An upper bound on the expected cost of an optimal assignment. Technical Report, Computer Science Division, University of California, Berkeley CA, 1984.

[Ka 87] R. M. Karp. An upper bound on the expected cost of an optimal assignment. In *Discrete Algorithms and Complexity: Proceedings of the Japan-U.S Joint Seminar*, pages 1-4, Academic Press, 1987.

[KKV 94] R. M. Karp, A. H. G Rinnooy Kan and R. Vohra. Average case analysis of a heuristic for the assignment problem. *Mathematics of Operations Research*, 19(3) (1994) 513-522.

[La 93] A. Lazarus. Certain expected values in the random assignment problem. *Oper. Res. Letters*, 14 (1993) 207-214.

[MP 87] M. Mézard and G. Parisi. On the solution of the random link matching problem. *J. Physique*, 48 (1987) 1451-1459.

[MPV 87] M. Mézard, G. Parisi and M. A. Virasoro. *Spin Glass Theory and Beyond.* World Scientific, Singapore, 1987.

[Ol 92] *Asymptotic Properties of Random Assignment Problems.* PhD thesis, Kungl Tekniska Högskolan, Stockholm, Sweden, 1992.

[Pa 98] G. Parisi. A conjecture on random bipartite matching. Technical Report cond-mat/9801176, XXX LANL Archive, 1998

[St 97] J. Michael Steele. *Probability Theory and Combinatorial Optimization.* Number 69 in CBMS-NSF Regional Conference Series in Applied Mathematics. SIAM, Philadelphia PA, 1997.

[Wa 79] D. W. Walkup. On the expected value of a random assignment problem. *SIAM J. Comput.*, 8 (1979) 440-442.