Proofs of the Parisi and Coppersmith-Sorkin Random Assignment Conjectures

Chandra Nair,* Balaji Prabhakar,† Mayank Sharma‡
350 Serra Mall, Packard 270, Stanford, California 94305;
e-mails: chandra.nair@gmail.com, balaji@stanford.edu, msharma@stanford.edu

Received 28 March 2003; revised 18 November 2003; accepted 16 May 2005
Published online 5 October 2005 in Wiley InterScience (www.interscience.wiley.com).
DOI 10.1002/rsa.20084

ABSTRACT: Suppose that there are n jobs and n machines and it costs c_{ij} to execute job i on machine j. The assignment problem concerns the determination of a one-to-one assignment of jobs onto machines so as to minimize the cost of executing all the jobs. When the c_{ij} are independent and identically distributed exponentials of mean 1, Parisi [Technical Report cond-mat/9801176, xxx LANL Archive, 1998] made the beautiful conjecture that the expected cost of the minimum assignment equals $\sum_{i=1}^{n}(1/i^2)$. Coppersmith and Sorkin [Random Structures Algorithms 15 (1999), 113–144] generalized Parisi's conjecture to the average value of the smallest k-assignment when there are n jobs and m machines. Building on the previous work of Sharma and Prabhakar [Proc 40th Annu Allerton Conf Communication Control and Computing, 2002, 657–666] and Nair [Proc 40th Annu Allerton Conf Communication Control and Computing, 2002, 667–673], we resolve the Parisi and Coppersmith-Sorkin conjectures. In the process we obtain a number of combinatorial results which may be of general interest. © 2005 Wiley Periodicals, Inc. Random Struct. Alg., 27, 413–444, 2005

Correspondence to: C. Nair
*Supported by the Stanford Graduate Fellowship and Stanford Networking Research Center Grant 1005544-1-WAAXI.
†Supported in part by the NSF Grant ANI-9985446.
‡Supported by Stanford Office of Technology Grant 2DTA112, Stanford Networking Research Center Grant 1005545-1-WABCJ, and NSF Grant ANI-9985446.

1. INTRODUCTION

Suppose there are n jobs and n machines and it costs c_{ij} to execute job i on machine j. An assignment (or a matching) is a one-to-one mapping of jobs onto machines. Representing an assignment as a permutation $\pi : \{1, \ldots, n\} \to \{1, \ldots, n\}$, the cost of the assignment π equals $\sum_{i=1}^{n} c_{i\pi(i)}$. The assignment problem consists of finding the assignment with the lowest cost. Let

$$C_n = \min_{\pi} \sum_{i=1}^{n} c_{i\pi(i)}$$

represent the cost of the minimizing assignment. In the *random* assignment problem the c_{ij} are i.i.d. random variables drawn from some distribution. A quantity of interest in the random assignment problem is the expected minimum cost, $\mathbb{E}(C_n)$.

When the costs c_{ij} are i.i.d. mean 1 exponentials, Parisi [19] made the following conjecture:

$$\mathbb{E}(C_n) = \sum_{i=1}^{n} \frac{1}{i^2}. \tag{1.1}$$

Coppersmith and Sorkin [6] 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) entries is:

$$C(k, m, n) = \sum_{i,j \geq 0, i+j<k} \frac{1}{(m-i)(n-j)}. \tag{1.2}$$

By definition, $C(n, n, n) = \mathbb{E}(C_n)$, and the right-hand sides of (1.2) and (1.1) are equal when $k = m = n$.

In this paper, we prove Parisi's conjecture by two different but related strategies. The first builds on the work of Sharma and Prabhakar [20] and establishes Parisi's conjecture by showing that certain increments of weights of matchings are exponentially distributed with a given rate and are independent. The second method builds on Nair [17] and establishes the Parisi and the Coppersmith-Sorkin conjectures. It does this by showing that certain other increments are exponentials with given rates; the increments are not required to be (and, in fact, aren't) independent.

The two methods mentioned above use a common set of combinatorial and probabilistic arguments. For ease of exposition, we choose to present the proof of the conjectures in [20] first. We then show how those arguments also resolve the conjectures in [17].

Before surveying the literature on this problem, it is important to mention that simultaneously and independently of our work, Linusson and Wästlund [14] have also announced a proof of the Parisi and Coppersmith-Sorkin conjectures based on a quite different approach.

1.1. Background and Related Work

There has been a lot of work on determining bounds for the expected minimum cost and on calculating its asymptotic value. Historically much of this work has been done for the case when the entries were uniformly distributed between 0 and 1. However, [1] shows that the asymptotic results carry over for exponential random variables as well.

Assuming that $\lim_n \mathbb{E}(C_n)$ exists, let us denote it by C^*. We survey some of the work; more details can be found in [21, 6]. Early work used feasible solutions to the dual linear programming (LP) formulation of the assignment problem for obtaining the following lower bounds for C^*: $(1+1/e)$ by Lazarus [12], 1.441 by Goemans and Kodialam [8], and 1.51 by Olin [18]. The first upper bound of 3 was given by Walkup [23], who thus demonstrated that $\limsup_n E(C_n)$ is finite. Walkup's argument was later made constructive by Karp, Rinnooy Kan, and Vohra, [11]. Karp [10] made a subtle use of LP duality to obtain a better upper bound of 2. Coppersmith and Sorkin [6] further improved the bound to 1.94.

Meanwhile, it had been observed through simulations that for large n, $E(C_n) \approx 1.642$ (Brunetti, Krauth, M. Mézard, and G. Parisi [4]). Mézard and Parisi [15] used the nonrigorous *replica method* [16] of statistical physics to argue that $C^* = \pi^2/6$. [Thus, Parisi's conjecture for the finite random assignment problem with i.i.d. exp(1) costs is an elegant restriction to the first n terms of the expansion: $\pi^2/6 = \sum_{i=1}^{\infty}(1/i^2)$.] More interestingly, their method allowed them to determine the density of the edge-cost distribution of the limiting optimal matching. These sharp (but nonrigorous) 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 [1] proved that C^* exists by identifying the limit as the average value of a minimum-cost matching on a certain random weighted infinite tree. In the same work he also established that the distribution of c_{ij} affects C^* only through the value of its density function at 0 (provided it exists and is strictly positive). Thus, as far as the value of C^* is concerned, the distributions $U[0, 1]$ and exp(1) are equivalent. More recently, Aldous [2] established that $C^* = \pi^2/6$, and obtained the same limiting optimal edge-cost distribution as [15]. He also obtained 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.

Generalizations of Parisi's conjecture have also been made in several ways. Linusson and Wästlund [13] conjectured an expression for the expected cost of the minimum k-assignment in an $m \times n$ matrix consisting of zeroes at some specified positions and exp(1) entries at all other places. Indeed, it is by proving this conjecture in their recent work [14] that they obtain proofs of the Parisi and Coppersmith-Sorkin conjectures. Buck, Chan, and Robbins [5] generalized the Coppersmith-Sorkin conjecture to the case where the c_{ij} are distributed according to $\exp(a_i b_j)$ for $a_i, b_j > 0$. In other words, if we let $\mathbf{a} = [a_i]$ and $\mathbf{b} = [b_j]$ be column vectors, then the rate matrix for the costs is of rank 1 and is of the form $\mathbf{ab^T}$. This conjecture has been subsequently established in [24] by Wästlund using a modification of the argument in [14].

Alm and Sorkin [3], and Linusson and Wästlund [13] verified the conjectures of Parisi and Coppersmith-Sorkin for small values of k, m and n. Coppersmith and Sorkin [7] studied the expected incremental cost, under certain hypotheses, of going from the smallest $(m-1)$-assignment in an $(m-1) \times n$ matrix to the smallest m-assignment in a row-augmented $m \times n$ matrix. However, as they note, their hypotheses are too restrictive and their approach fails to prove their conjecture.

An outline of the paper is as follows: In Section 2 we recall some previous work from [20] and state Theorem 2.4, whose proof implies a proof of Parisi's conjecture. In Section 3 we describe an induction procedure for proving Theorem 2.4. We then state and prove some combinatorial properties of matchings in Section 4 that will be useful for the rest of the paper. Section 5 contains a proof of Theorem 2.4. Section 6 builds on the work of [17] and contains a proof of Theorem 6.3. This implies a proof of the Coppersmith-Sorkin conjecture. We conclude in Section 7. We now present some conventions that are observed in the rest of the paper.

1.2. Conventions

1. The words "cost" and "weight" are used interchangeably and mean the same thing; the cost (or weight) of a collection entries is the sum of the values of the entries.
2. The symbol "$\sim$" stands for "is distributed as", and " $\perp\!\!\!\perp$ " stands for "is independent of".
3. By $X \sim \exp(\lambda)$ we mean that X is exponentially distributed with mean $1/\lambda$; i.e., $\mathbb{P}(X > x) = e^{-\lambda x}$ for $x, \lambda \geq 0$.
4. We use the term "rectangular matrices" to refer to $m \times n$ matrices with $m < n$.
5. We employ the following notation:
   — Boldface capital letters such as $\mathbf{A}, \mathbf{C}, \mathbf{M}$ represent matrices.
   — Calligraphic characters such as $\mathcal{R}, \mathcal{S}, \mathcal{T}, \mathcal{V}$ denote matchings.
   — The plain non-boldface version of a matching's name, e.g., R, S, T, V represent the weight of that matching.
6. $Col(\mathcal{S})$ to represent the set of columns used by the matching $\mathcal{S}$.
7. Throughout this paper, we shall assume that the costs are drawn from some continuous distribution. Hence, with probability 1, no two matchings will have the same weight. This makes the "smallest matching" in a collection unique; tie-breaking rules will not be needed.

Remark 1.1. Note that all of our claims in Section 4 will go through even if we do not assume uniqueness. However, when there is a tie, the claims must be reworded as "there exists a matching with the smallest weight satisfying," instead of "the smallest matching satisfies." The general statements when uniqueness is not assumed are stated in Section 4.11. However, we omit the proofs as it is similar to the case, but slightly more cumbersome than, when uniqueness is assumed.

2. PRELIMINARIES

Let $\mathbf{C} = [c_{ij}]$ be an $m \times n$ $(m < n)$ cost matrix with i.i.d. exp(1) entries. Let $\mathcal{T}_0$ denote the smallest matching of size m in this matrix. Without loss of generality, assume that $Col(\mathcal{T}_0) = \{1, 2, \ldots, m\}$. For $i = 1, \ldots, n$, let $\mathcal{S}_i$ denote the smallest matching of size m in the $m \times (n-1)$ submatrix of $\mathbf{C}$ obtained by removing its ith column. Note that $\mathcal{S}_i = \mathcal{T}_0$ for $i \geq m + 1$. Therefore, the $\mathcal{S}_i$'s are at most $m + 1$ distinct matchings.

Definition 2.1 (**S-matchings**). *The collection of matchings* $\{\mathcal{S}_1, \ldots, \mathcal{S}_m, \mathcal{S}_{m+1} (= \mathcal{T}_0)\}$ *is called the* S-matchings *of* $\mathbf{C}$ *and is denoted by* $\mathcal{S}(\mathbf{C})$.

Definition 2.2 (**T-matchings**). *Let* $\{\mathcal{T}_1, \ldots, \mathcal{T}_m\}$ *be a permutation of* $\{\mathcal{S}_1, \ldots, \mathcal{S}_m\}$ *such that* $T_1 < T_2 < \cdots < T_m$*; that is, the* $\mathcal{T}_i$*'s are a rearrangement of the* $\mathcal{S}_i$*'s in order of increasing weight. The collection of matchings* $\{\mathcal{T}_0, \mathcal{T}_1, \ldots, \mathcal{T}_m\}$ *is called the* T-matchings *of* $\mathbf{C}$ *and is denoted by* $\mathcal{T}(\mathbf{C})$.

Remark 2.3. Nothing in the definition of the S-matchings prevents any two of the $\mathcal{S}_i$'s from being identical; however, we will show in Corollary 4.2 that they are all distinct.

These quantities are illustrated below by taking $\mathbf{C}$ to be the following 2×3 matrix:

$$\mathbf{C}: \begin{array}{|c|c|c|}\hline 3 & 6 & 11 \\ \hline 9 & 2 & 20 \\ \hline \end{array}$$

$$\begin{array}{|c|c|}\hline 6 & \mathbf{11} \\ \hline \mathbf{2} & 20 \\ \hline \end{array} \Rightarrow S_1 = 13; \qquad \begin{array}{|c|c|}\hline 3 & \mathbf{11} \\ \hline \mathbf{9} & 20 \\ \hline \end{array} \Rightarrow S_2 = 20; \qquad \begin{array}{|c|c|}\hline \mathbf{3} & 6 \\ \hline 9 & \mathbf{2} \\ \hline \end{array} \Rightarrow S_3 = 5 = T_0.$$

In the above example, $T_0 = 5$, $T_1 = 13$, and $T_2 = 20$.

We now state the main result that will establish Parisi's Conjecture.

Theorem 2.4. *Consider an $m \times n$ $(m < n)$ matrix, $\mathbf{A}$, with i.i.d.* exp(1) *entries. Let $\{T_0, T_1, \ldots, T_m\}$ denote the weights of the* T-matchings *of* $\mathbf{A}$. *Then the following hold:*

- $T_j - T_{j-1} \sim \exp(m - j + 1)(n - m + j - 1)$, *for* $j = 1, \ldots, m$.
- $T_1 - T_0 \perp\!\!\!\perp T_2 - T_1 \perp\!\!\!\perp \cdots \perp\!\!\!\perp T_m - T_{m-1}$.

The proof of this theorem will be presented later. For completeness, we now reproduce the arguments from [20] which show how Theorem 2.4 implies Parisi's conjecture.

Corollary 2.5. *Let $\mathbf{C}$ be an $n \times n$ cost matrix with i.i.d.* exp(1) *entries. Let C_n denote the cost of the minimum assignment. Then*

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

Proof. The proof is by induction. The induction hypothesis is trivially true when $n = 1$ since $\mathbb{E}(C_1) = 1$. Let us assume that we have

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

Delete the top row of $\mathbf{C} \equiv [c_{ij}]$ to obtain the rectangular matrix $\mathbf{A}$ of dimensions $(n-1) \times n$. Let $\{S_1, \ldots, S_n\}$ and $\{T_0, \ldots, T_{n-1}\}$ be the weights of the matchings in $\mathcal{S}(\mathbf{A})$ and $\mathcal{T}(\mathbf{A})$ respectively.

The relationship $C_n = \min_{j=1}^{n}\{c_{1j} + S_j\}$ allows us to evaluate $\mathbb{E}(C_n)$ as follows:

$$\begin{aligned} \mathbb{E}(C_n) &= \int_0^\infty P(C_n > x)\, dx \\ &= \int_0^\infty P(c_{1j} > x - S_j,\ j = 1, \ldots, n)\, dx \\ &= \int_0^\infty P(c_{1\sigma(j)} > x - T_j,\ j = 0, \ldots, n-1)\, dx, \end{aligned} \tag{2.1}$$

where $\sigma(\cdot)$ is a 1–1 map from $\{0, 1, \ldots, n-1\}$ to $\{1, 2, \ldots, n\}$ such that $c_{1\sigma(j)}$ is the entry in the first row of $\mathbf{C}$ that lies outside the columns occupied by the matching $\mathcal{T}_j$ in $\mathbf{A}$. Now, since the first row is independent of the matrix $\mathbf{A}$ and $\sigma(\cdot)$ is a bijection, the entries $c_{1\sigma(j)}$ are i.i.d. exp(1) random variables. We therefore have from (2.1) that

$$\mathbb{E}(C_n) = \mathbb{E}_{\mathbf{A}}\left(\int_0^\infty P(c_{1\sigma(j)} > x - t_j,\ j = 0, \ldots, n-1)\, dx \,\Big|\, \mathbf{A} \right).$$

We proceed by evaluating the expression inside the integral. Thus,

$$\begin{aligned}
&\int_0^\infty P(c_{1\sigma(j)} > x - t_j,\ j = 0, \ldots, n-1)\, dx \\
&= \int_0^\infty \prod_{j=0}^{n-1} P(c_{1\sigma(j)} > x - t_j)\, dx \qquad \text{(independence of } c_{1\sigma(j)}) \\
&= \int_0^{t_0} dx + \int_{t_0}^{t_1} e^{-(x-t_0)}\, dx + \cdots + \int_{t_{n-2}}^{t_{n-1}} e^{-((n-1)x - t_0 - \cdots - t_{n-2})}\, dx \\
&\quad + \int_{t_{n-1}}^\infty e^{-(nx - t_0 - \cdots - t_{n-1})}\, dx \text{ (since the } t_i\text{'s are increasing)} \\
&= t_0 + \left(1 - e^{-(t_1 - t_0)}\right) + \frac{1}{2}\left(e^{-(t_1-t_0)} - e^{-(2t_2 - t_0 - t_1)}\right) + \cdots \\
&\quad + \frac{1}{n-1}\left(e^{-((n-2)t_{n-2} - t_0 - \cdots - t_{n-3})} - e^{-((n-1)t_{n-1} - t_0 - \cdots - t_{n-2})}\right) + \frac{1}{n} e^{-((n-1)t_{n-1} - t_0 - \cdots - t_{n-2})} \\
&= t_0 + 1 - \frac{1}{2} e^{-(t_1 - t_0)} - \frac{1}{6} e^{-(2t_2 - t_0 - t_1)} - \cdots - \frac{1}{n(n-1)} e^{-((n-1)t_{n-1} - t_0 - \cdots - t_{n-2})}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathbb{E}(C_n) &= \mathbb{E}(T_0) + 1 - \sum_{i=1}^{n-1} \frac{1}{i(i+1)} \mathbb{E}\left(e^{-(iT_i - T_0 - \cdots - T_{i-1})}\right) \\
&= \mathbb{E}(T_0) + 1 - \sum_{i=1}^{n-1} \frac{1}{i(i+1)} \mathbb{E}\left(e^{\sum_{j=1}^{i} -j(T_j - T_{j-1})}\right).
\end{aligned} \tag{2.2}$$

However, from Theorem 2.4 (setting $m = n - 1$), we obtain

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

Substituting this in (2.2) gives

$$\mathbb{E}(C_n) = \mathbb{E}(T_0) + \frac{1}{n^2} + \frac{1}{n} \sum_{i=1}^{n-1} \frac{1}{i}. \tag{2.3}$$

We are left with having to evaluate $\mathbb{E}(T_0)$. First, for $j = 1, \ldots, n-1$,

$$\mathbb{E}(T_j) = \mathbb{E}(T_0) + \sum_{k=1}^{j} \mathbb{E}(T_k - T_{k-1}) = \mathbb{E}(T_0) + \sum_{k=1}^{j} \frac{1}{k(n-k)} \qquad \text{(by Theorem 2.4).} \tag{2.4}$$

Now, the random variable S_1 is the cost of the smallest matching of an $(n-1)\times(n-1)$ matrix of i.i.d. exp(1) random variables obtained by removing the first column of $\mathbf{A}$. Hence S_1 is distributed as C_{n-1}. However, by symmetry, S_1 is equally likely to be any of $\{T_0, \ldots, T_{n-1}\}$. Hence we get that

$$\begin{aligned} \mathbb{E}(S_1) &= \frac{1}{n}\sum_{j=0}^{n-1} \mathbb{E}(T_j) = \frac{1}{n}\mathbb{E}(T_0) + \frac{1}{n}\sum_{j=1}^{n-1}\left(\mathbb{E}(T_0) + \sum_{k=1}^{j}\frac{1}{k(n-k)}\right) \\ &= \mathbb{E}(T_0) + \frac{1}{n}\sum_{k=1}^{n-1}\frac{1}{k}. \end{aligned} \tag{2.5}$$

By the induction assumption, $\mathbb{E}(C_{n-1}) = \sum_{k=1}^{n-1}\frac{1}{k^2} = \mathbb{E}(S_1)$. Substituting this into (2.5), we obtain

$$\mathbb{E}(T_0) = \sum_{k=1}^{n-1}\left(\frac{1}{k^2} - \frac{1}{nk}\right). \tag{2.6}$$

Using this at (2.3), we get

$$\mathbb{E}(C_n) = \sum_{i=1}^{n-1}\left(\frac{1}{i^2} - \frac{1}{ni}\right) + \frac{1}{n^2} + \frac{1}{n}\sum_{i=1}^{n-1}\frac{1}{i} = \sum_{i=1}^{n}\frac{1}{i^2}. \tag{2.7}$$

∎

3. A SKETCH OF THE PROOF OF THEOREM 2.4

The proof uses induction and follows the steps below.

1. First, we prove that for any rectangular $m \times n$ matrix, $\mathbf{A}$, $T_1 - T_0 \sim \exp m(n-m)$.
2. The distribution of the higher increments is determined by an inductive procedure. We remove a suitably chosen row of $\mathbf{A}$ to obtain an $m-1 \times n$ matrix, $\mathbf{B}$, which has the following property: Let $\{T_0, \ldots, T_m\}$ and $\{U_0, \ldots, U_{m-1}\}$ be the weights of the T-matchings in $\mathcal{T}(\mathbf{A})$ and $\mathcal{T}(\mathbf{B})$ respectively. Then

   $$U_j - U_{j-1} = T_{j+1} - T_j \qquad \text{for } j = 1, 2, \ldots, m-1.$$

   Establishing this combinatorial property is one major thrust of the paper.
3. We will then show that $\mathbf{B}$ possesses a useful probabilistic property: Its entries are i.i.d. exp(1) random variables, independent of $T_1 - T_0$. This property, in conjunction with the results in 1 and 2 above, allows us to conclude (i) $T_2 - T_1 = U_1 - U_0 \sim \exp(m-1)(n-m+1)$ and (ii) $T_{j+1} - T_j \perp\!\!\!\perp T_1 - T_0$ for $j = 1, 2, \ldots, m-1$; in particular, $T_2 - T_1 \perp\!\!\!\perp T_1 - T_0$. We use the matrix $\mathbf{B}$ as the starting point in the next step of the induction and proceed.

Remark 3.1. We have seen above that $T_1 - T_0$ is independent of $\mathbf{B}$ and hence of all higher increments $T_{j+1} - T_j, j = 1, 2, \ldots, m-1$. This argument, when applied in the subsequent stages of the induction, establishes the independence of all the increments of $\mathbf{A}$.

The diagram below encapsulates our method of proof. We shall show that the first increments $T_1 - T_0, U_1 - U_0, \ldots, V_1 - V_0, \ldots,$ and $W_1 - W_0$ are mutually independent, that they are exponentially distributed with appropriate rates, and that they are each equal to a particular original increment $T_{j+1} - T_j$.

$$
\begin{array}{lcccccc}
\text{Matrix} & & & \text{T-matchings} & & & \\
\\
\mathbf{A}: & T_1 - T_0 & T_2 - T_1 & \ldots & T_{j+1} - T_j & \ldots & T_m - T_{m-1} \\
 & & \| & & \| & & \| \\
\mathbf{B}: & & U_1 - U_0 & \ldots & U_j - U_{j-1} & \ldots & U_{m-1} - U_{m-2} \\
 & & & & \| & & \| \\
\vdots & & & & \vdots & & \vdots \\
\mathbf{D}: & & & & V_1 - V_0 & \ldots & V_k - V_{k-1} \\
\vdots & & & & & & \vdots \\
 & & & & & & \| \\
\mathbf{F}: & & & & & & W_1 - W_0
\end{array}
$$

In summary, the proof of Theorem 2.4 involves a combinatorial and a probabilistic part. We develop a number of combinatorial lemmas in the next section. The lemmas and their proofs can be stated using conventional language; e.g., symmetric differences, alternating cycles and paths, or as linear optimizations over Birkhoff polytopes. However, given the straightforward nature of the statements, presenting the proofs in plain language as we have chosen to do seems natural. The probabilistic arguments and the proof of Theorem 2.4 are presented in Section 5.

4. SOME COMBINATORIAL PROPERTIES OF MATCHINGS

To execute some of the proofs in this section, we will use the alternate representation of an arbitrary $m \times n$ matrix $\mathbf{C}$ as a complete bipartite graph $\mathcal{K}_{m,n}$, with m vertices on the left and n vertices on the right corresponding to the rows and columns of $\mathbf{C}$, respectively. The edges are assigned weights c_{ij} with the obvious numbering.

In a number of these combinatorial lemmas we are interested in properties of "near optimal matchings." That is, suppose $\mathcal{M}$ is the smallest matching of size k in the matrix $\mathbf{C}$. Near optimal matchings of interest include (i) $\mathcal{M}'$: the smallest matching of size k which doesn't use all the columns of $\mathcal{M}$, or (ii) $\mathcal{M}''$: the smallest matching of size $k+1$. A generic conclusion of the combinatorial lemmas is that near-optimal matchings are "closely related" to the optimal matching $\mathcal{M}$. For example, we will find that $\mathcal{M}'$ uses all but one of the columns of $Col(\mathcal{M})$, and that the rows and columns used by $\mathcal{M}''$ are a superset of those used by $\mathcal{M}$.

Lemma 4.1. *Consider an $m \times n$ matrix $\mathbf{C}$. For every $j \in Col(\mathcal{T}_0)$, we have $|Col(\mathcal{S}_j) \cap Col(\mathcal{T}_0)| = m - 1$.*

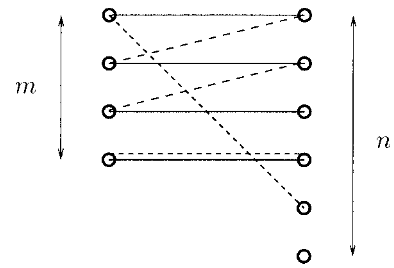

Fig. 1. Subgraph formed by two matchings depicting an even-length path and a 2-cycle.

Proof. We represent the matrix $\mathbf{C}$ as a complete bipartite graph $\mathcal{K}_{m,n}$. Without loss of generality, let $Col(\mathcal{T}_0)$ be the first m columns of $\mathbf{C}$, and let $j = 1$. Focus on the subgraph consisting of only those edges which are present in $\mathcal{T}_0$ and $\mathcal{S}_1$. For example, the subgraph is shown in Fig. 1 where the bold edges belong to $\mathcal{T}_0$ and the dashed edges belong to $\mathcal{S}_1$.

In general, a subgraph formed using two matchings in a bipartite graph can consist of the following components: cycles, and paths of even or odd lengths. We claim that it is impossible for the subgraph induced by the edges of $\mathcal{T}_0$ and $\mathcal{S}_1$ to have cycles of length greater than two, or paths of odd length. (Cycles of length two represent the entries common to $\mathcal{T}_0$ and $\mathcal{S}_1$.)

A cycle of length greater than two is impossible because it would correspond to two different submatchings being chosen by $\mathcal{T}_0$ and $\mathcal{S}_1$ on a common subset of rows and columns. This would contradict the minimality of either $\mathcal{T}_0$ or of $\mathcal{S}_1$.

An odd-length path is not possible because every vertex on the left has degree 2. Thus, any path will have to be of even length.

We now show that the only component (other than cycles of length 2) that can be present in the subgraph is a single path of even length whose degree-1 vertices are on the right. Every node on the left has degree 2 and hence even paths with two degree-1 nodes on the left are not possible. Now we rule out the possibility of more than one even length path. Suppose to the contrary that there are two or more paths of even length. Consider any two of them and note that at least one of them will not be incident on column 1. Now the edges of $\mathcal{T}_0$ along this path have smaller combined weight than the edges of $\mathcal{S}_1$ by the minimality of $\mathcal{T}_0$. Thus, we can append these bold edges to the dashed edges not on this path to obtain a new matching $\mathcal{S}_1'$ which would be smaller than $\mathcal{S}_1$. This contradicts the minimality of $\mathcal{S}_1$ amongst all matchings that do not use column 1.

Therefore, the subgraph formed by the edges of $\mathcal{T}_0$ and $\mathcal{S}_1$ can only consist of 2-cycles and one even length path. To complete the proof, observe that an even length path with two degree-1 vertices on the right implies that the edges of $\mathcal{S}_1$ in the path use exactly one column that is not used by the edges of $\mathcal{T}_0$ in the path (and vice-versa). This proves the lemma. ∎

Corollary 4.2. *The cardinality of $\mathcal{S}(\mathbf{C})$ is $m + 1$.*

Proof. From the definition of $\mathcal{S}_i$ it is clear that for $i \leq m$, $\mathcal{S}_i \neq \mathcal{T}_0$. We need to show that $\mathcal{S}_i \neq \mathcal{S}_j$ for $i \neq j$, $i, j \leq m$. From Lemma 4.1, $\mathcal{S}_i$ uses all the columns of $\mathcal{T}_0$ except column i. In particular, it uses column j and therefore is different from $\mathcal{S}_j$. ∎

Corollary 4.3. *For any $1 \leq k \leq m$, taking $i \in Col(\mathcal{T}_0) \cap Col(\mathcal{T}_1) \cdots Col(\mathcal{T}_k)$, an arrangement of $\mathcal{S}_i$ in increasing order gives the sequence $\mathcal{T}_{k+1}, T_{k+2}, \ldots, T_m$.*

Proof. The proof follows in a straightforward fashion from Lemma 4.1 and the definition of S-matchings. ∎

We can use Lemma 4.1 and Corollary 4.3 to give an alternate characterization of the T-matchings that does not explicitly consider the S-matchings.

Lemma 4.4 (Alternate Characterization of the T-Matchings). *Consider an $m \times n$ rectangular matrix,* **C**. *Let $\mathcal{T}_0$ be the smallest matching of size m in this matrix. The rest of the T-matchings $\mathcal{T}_1, \ldots, \mathcal{T}_m$, can be defined recursively as follows: $\mathcal{T}_1$ is the smallest matching in the set $\mathcal{R}_1 = \{\mathcal{M} : Col(\mathcal{M}) \supsetneq Col(\mathcal{T}_0)\}$, $\mathcal{T}_2$ is the smallest matching in the set $\mathcal{R}_2 = \{\mathcal{M} : Col(\mathcal{M}) \supsetneq (Col(\mathcal{T}_0) \cap Col(\mathcal{T}_1))\}$,..., and $\mathcal{T}_m$ is the smallest matching in the set $\mathcal{R}_m = \{\mathcal{M} : Col(\mathcal{M}) \supsetneq (Col(\mathcal{T}_0) \cap Col(\mathcal{T}_1) \cap \cdots \cap\ Col(\mathcal{T}_{m-1}))\}$. Then $\{\mathcal{T}_0, \ldots, \mathcal{T}_m\}$ are the T-matchings of* **C**.

Proof. The proof is straightforward and is omitted. (Note that the alternate characterization was used in the definition of the T-matchings in [17].) ∎

Remark 4.5. The next lemma captures the following statement. If a matching is locally minimal, then it is also globally minimal. The local neighborhood of a matchings is defined by the set of matchings whose columns differ from the matching under consideration by at most one column. That is, the lemma asserts that if a matching is the smallest among the matchings in its local neighborhood then it is also globally minimal.

Lemma 4.6. *Consider an $m \times n$ rectangular matrix,* **C**. *Suppose there is a size-m matching $\mathcal{M}$ with the following property: $M < M'$ for all size-m matchings $\mathcal{M}'(\neq \mathcal{M})$ such that $|Col(\mathcal{M}') \cap Col(\mathcal{M})| \geq m - 1$. Then $\mathcal{M} = \mathcal{T}_0$.*

Proof. Without loss of generality, assume $Col(\mathcal{M}) = \{1, 2, \ldots, m\}$. The lemma is trivially true for $n = m+1$. Let $k \geq 2$ be the first value such that there is a matrix, **C**, of size $m \times (m+k)$ which violates the lemma. We will show that this leads to a contradiction and hence prove the lemma.

Clearly, $Col(\mathcal{T}_0)$ must contain all the columns $\{m + 1, \ldots, m + k\}$. If not, there is a smaller value of k for which the lemma is violated. For any $j \in \{m+1, \ldots, m+k\}$ consider $Col(\mathcal{S}_j)$, where $\mathcal{S}_j$ is the smallest matching that does not contain column j.

The fact that k is the smallest number for which Lemma 4.6 is violated implies $\mathcal{S}_j = \mathcal{M}$. Hence $|Col(\mathcal{S}_j) \cap Col(\mathcal{T}_0)| \leq m - k \leq m - 2$. This contradicts Lemma 4.1, proving the lemma. ∎

Lemma 4.7. *Consider a $m \times n$ cost matrix* **C**. *Let* **D** *be an extension of* **C** *formed by adding r additional rows ($r < n - m$). Then $Col(\mathcal{T}_0(\mathbf{C})) \subset Col(\mathcal{T}_0(\mathbf{D}))$.*

Proof. As before, we represent the augmented matrix **D** as a complete bipartite graph $K_{m+r,n}$ and focus on the subgraph (see Fig. 2) consisting of only those edges that are part of $\mathcal{T}_0(\mathbf{C})$ (bold edges) and $\mathcal{T}_0(\mathbf{D})$ (dashed edges).

We proceed by eliminating the possibilities for components of this subgraph. As in Lemma 4.1, the minimality of the two matchings under consideration prevents cycles of

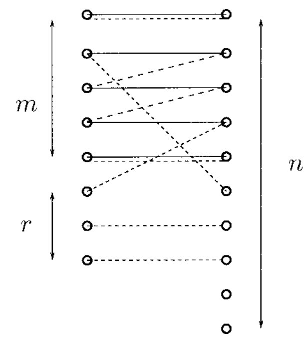

Fig. 2. Subgraph depicting odd-length paths and a 2-cycle.

length greater than 2 from being present. Note that 2-cycles (or common edges) are possible, and these do not violate the statement of the lemma.

Next we show that paths of even length cannot exist. Consider even-length paths with degree-1 vertices on the left. If such a path exists, then it implies that there is a vertex on the left on which a lone bold edge is incident. This is not possible since the edges of $\mathcal{T}_0(\mathbf{D})$ are incident on every vertex on the left.

Now consider even-length paths with degree-1 vertices on the right. These have the property that the solid and dashed edges use the same vertices on the left (i.e. same set of rows). Now, we have two matchings on the same set of rows and therefore by choosing the lighter one, we can contradict the minimality of either $\mathcal{T}_0(\mathbf{C})$ or $\mathcal{T}_0(\mathbf{D})$.

Consider odd-length paths. Since every vertex corresponding to rows in $\mathbf{C}$ must have degree 2, the only type of odd-length paths possible are those in which the number of edges from $\mathcal{T}_0(\mathbf{D})$ is one more than the number of edges from $\mathcal{T}_0(\mathbf{C})$. But in such an odd-length path, the vertices on the right (columns) used by $\mathcal{T}_0(\mathbf{C})$ are also used by $\mathcal{T}_0(\mathbf{D})$. Since the only components possible for the subgraph are odd length paths as above and common edges, $Col(\mathcal{T}_0(\mathbf{C})) \subset Col(\mathcal{T}_0(\mathbf{D}))$. ▪

Lemma 4.8. *Let $\mathbf{C}$ be an $m \times n$ rectangular matrix. Let $\mathcal{S}_k(i)$ denote the entry of $\mathcal{S}_k$ in row i. Consider three arbitrary columns k_1, k_2, k_3. For every row i, at least two of $\mathcal{S}_{k_1}(i)$, $\mathcal{S}_{k_2}(i)$, and $\mathcal{S}_{k_3}(i)$ must be the same.*

Proof. We first establish this claim for $m = n - 1$. Consider the subgraph formed by the edges in $\mathcal{S}_{k_1}$ and $\mathcal{S}_{k_2}$. This subgraph cannot have the following components:

- Cycles of length more than 2, since that would contradict the minimality of either $\mathcal{S}_{k_1}$ or $\mathcal{S}_{k_2}$.
- Odd length paths, since every vertex on the left has degree 2.
- Even length paths with degree-1 vertices on the left, since every vertex on left has degree 2.

Thus, the only possible components are even length paths with degree-1 vertices on the right, and common edges.

Now we use the fact that $m = n - 1$ to claim that there can only be one even length path. If there were two even length paths with degree-1 vertices on the right, then the edges in $\mathcal{S}_{k_1}$ will avoid at least two columns (one from each even length path). But $m = n - 1$ implies the edges in $\mathcal{S}_{k_1}$ can avoid only column k_1. Similarly the edges of $\mathcal{S}_{k_2}$ can avoid only column k_2. Thus, the single even length alternating path must have k_1 and k_2 as its degree-1 vertices. Let us call this path P_{12}. Arguing as above, we conclude that the subgraph formed by $\mathcal{S}_{k_1}$ and $\mathcal{S}_{k_3}$ consists only of common edges and one even length alternating path, P_{13}, connecting k_1 and k_3.

We now prove the lemma by contradiction. Suppose that $\mathcal{S}_{k_1}(i)$, $\mathcal{S}_{k_2}(i)$, and $\mathcal{S}_{k_3}(i)$ are all distinct for some row i. Our method of proof will be to construct a size $(n-1)$ matching in $\mathbf{C}\backslash k_3$, say $\tilde{\mathcal{S}}_{k_3}$, using only edges belonging to $\mathcal{S}_{k_1}$, $\mathcal{S}_{k_2}$, and possibly some from $\mathcal{S}_{k_3}$, which has a cost smaller than the cost of $\mathcal{S}_{k_3}$. This will contradict the minimality of $\mathcal{S}_{k_3}$ and hence prove the lemma. We need to consider two cases for the construction of $\tilde{\mathcal{S}}_{k_3}$.

Case 1: The vertex k_3 does not lie on the alternating path P_{12}: Consider the alternating path, P_{13}, from k_3 to k_1 consisting of edges from $\mathcal{S}_{k_1}$ and $\mathcal{S}_{k_3}$. Start traversing the path from k_3 along the edge belonging to $\mathcal{S}_{k_1}$. Observe that one takes the edge belonging to $\mathcal{S}_{k_1}$ when going from a right vertex to a left vertex and a $\mathcal{S}_{k_3}$-edge when going from a left vertex to a right vertex. Let v be the first vertex along this path that also belongs to P_{12}.

We claim that v must be on the right. Suppose that v is on the left. Since v is the first node common to P_{13} and P_{12}, it must be that there are two distinct edges belonging to $\mathcal{S}_{k_1}$ (belonging to each of P_{13} and P_{12}) incident on v. But this is impossible, since these edges belong to the same matching $\mathcal{S}_{k_1}$. Hence, v must be on the right.

Now form the matching $\tilde{\mathcal{S}}_{k_3}$ by taking the following edges:

- edges from $\mathcal{S}_{k_3}$ along P_{13} starting from k_3 until v,
- edges from $\mathcal{S}_{k_1}$ along P_{12} starting from v until k_2,
- edges from $\mathcal{S}_{k_2}$ along P_{12} starting from v until k_1,
- edges belonging to $\mathcal{S}_{k_1}$ from all the uncovered vertices on left.

Case 2: The vertex k_3 lies on P_{12}: We can construct $\tilde{\mathcal{S}}_{k_3}$ using the procedure stated in Case 1 if we take $v = k_3$. Then the matching $\tilde{\mathcal{S}}_{k_3}$ is formed by taking the following edges:

- edges from $\mathcal{S}_{k_1}$ along P_{12} starting from k_3 until k_2,
- edges from $\mathcal{S}_{k_2}$ along P_{12} starting from k_3 until k_1,
- edges belonging to $\mathcal{S}_{k_1}$ from all the uncovered vertices on left.

Observe that in both cases, by construction, the subgraph formed by the edges of $\mathcal{S}_{k_1}, \mathcal{S}_{k_2}$ and $\tilde{\mathcal{S}}_{k_3}$ is such that the vertices on left have at most degree 2.

To show that the cost of $\tilde{\mathcal{S}}_{k_3}$ is less than $\mathcal{S}_{k_3}$, we cancel edges that are common to the two matchings and thus obtain matchings $\tilde{\mathcal{S}}'_{k_3}$ and $\mathcal{S}'_{k_3}$ on $\mathbf{C}'$, a (possibly smaller) submatrix of $\mathbf{C}\backslash k_3$. Now $\tilde{\mathcal{S}}'_{k_3}$ consists of edges from either $\mathcal{S}_{k_1}$ or $\mathcal{S}_{k_2}$; denote these edges by E_1 and E_2 respectively. We have to show

$$\text{sum of edges in } \mathcal{S}'_{k_3} > \text{sum of edges in } \{E_1, E_2\} = \text{sum of edges in } \tilde{\mathcal{S}}'_{k_3}. \tag{4.1}$$

Let $E_1^c = \mathcal{S}_{k_1}\backslash E_1$ and $E_2^c = \mathcal{S}_{k_2}\backslash E_2$. Adding the weights of these edges to both sides of (4.1), we are now required to show

$$\text{sum of edges in } \{\mathcal{S}'_{k_3}, E_1^c, E_2^c\} > \mathcal{S}_{k_1} + \mathcal{S}_{k_2}. \tag{4.2}$$

We establish (4.2) by showing that the left hand side splits into the weights of two matchings, one each in $\mathbf{C}\backslash k_1$ and $\mathbf{C}\backslash k_2$. The minimality of $\mathcal{S}_{k_1}$ and $\mathcal{S}_{k_2}$ will then complete the proof.

But the decomposition into two appropriate matchings is immediate once we observe that in $\{\mathcal{S}'_{k_3}, E_1^c, E_2^c\}$, every vertex on the left has degree 2, and so does every vertex on the right, except k_1 and k_2. This splitting into the two matchings establishes (4.1) and thus shows that $S_{k_3} > \tilde{S}_{k_3}$. This contradiction proves the lemma when $m = n - 1$.

If $m < n-1$, append an $(n-m-1)\times n$ matrix to $\mathbf{C}$ to form an $(n-1)\times n$ matrix $\mathbf{D}$. The entries in $\mathbf{D}\backslash\mathbf{C}$ are i.i.d. random variables uniformly distributed on $[0, \epsilon/2(n-m)]$, where $\epsilon < \min\{|M - M'| : \mathcal{M}$ and $\mathcal{M}'$ are size-m matchings in $\mathbf{C}\}$. Then it is easy to see that for each i, $\mathcal{S}_i(\mathbf{D})$ contains $\mathcal{S}_i(\mathbf{C})$ since the combined weight of the additional edges from the appended part is too small to change the ordering between the matchings in $\mathbf{C}$.

Now apply the lemma to $\mathbf{D}$ to infer that at least two of $\mathcal{S}_{k_1}(i)$, $\mathcal{S}_{k_2}(i)$ and $\mathcal{S}_{k_3}(i)$ must be the same, where the $\mathcal{S}_{k_j}$ are size-m matchings of $\mathbf{C}$ and row i is in $\mathbf{C}$. This proves the lemma. ∎

Definition 4.9 (Marked elements). *An element of an $m \times n$ matrix $\mathbf{C}$ is said to be* marked *if it belongs to at least one of its* T-matchings.

Lemma 4.10. *An $m \times n$ matrix $\mathbf{C}$ has exactly two elements marked in each row.*

Proof. It is obvious that at least two such elements are present in each row. If there is any row that has three or more elements, by considering the S-matchings that give rise to any three of these elements we obtain a contradiction to Lemma 4.8. ∎

4.1. General Form of the Lemmas

In this section, we state the lemmas for the case when the cost matrix $\mathbf{C}$ has nonunique elements or subset sums. Thus $\mathcal{T}_0$, the smallest matching (in weight), is potentially nonunique. Choose any one of the equal weight matchings as $\mathcal{T}_0$. Since each $\mathcal{S}_j$ is defined as the smallest matching obtained by the removal of column j of $\mathcal{T}_0$, there may exist a set of matchings, $\mathbb{S}_j$, that have the same weight.

Claim 4.11. *There exists $\mathcal{S}_j \in \mathbb{S}_j$ for $j = 1, \ldots, m$ such that Lemmas 4.1, 4.8, and 4.10 remain valid when the cost matrix $\underline{C}$ has nonunique elements or subset sums.*

By following the arguments of the proof of the Lemmas 4.1, 4.8 and 4.10, one can show that it is always possible to define a set of matchings $\mathcal{S}_j \in \mathbb{S}_j$ such that Lemmas 4.1, 4.8, and 4.10 remain valid. The details are omitted.

For Lemma 4.4, similarly one can recursively choose a set of smallest-weight matchings $\mathcal{T}_j \in \mathcal{R}_j$ such that these are precisely the T-matchings alternately defined via the S-matchings.

Lemma 4.6 and its proof carries over without any change to the general case that we are considering. Note that the contradiction now is based on the modified Lemma 4.1; modification caused by the set of the S-matchings chosen according to Claim 4.11.

For Lemma 4.7 to be valid, we need to state that one can choose one among the several smallest-weight matchings $\mathcal{T}_0(\mathbf{C})$ [and similarly $\mathcal{T}_0(\mathbf{D})$] such that the lemma remains valid.

Remark 4.12. Note that though the lemmas are valid for general matrices, unless explicitly stated, we will assume in the rest of the paper that all subset sums are unique.

5. PROOF OF THEOREM 2.4

We shall now execute the three steps mentioned in Section 3.

Step 1: $T_1 - T_0 \sim \exp m(n-m)$. We will show that if $\mathbf{A}$ is an $m \times n$ rectangular matrix with i.i.d. exp(1) entries, then $T_1 - T_0 \sim \exp m(n-m)$. We begin by the following characterization of $Col(\mathcal{T}_0)$. Let $\mathcal{M}$ be any matching that satisfies the property that it is the smallest size-m matching in the columns $Col(\mathcal{M})$ of $\mathbf{A}$. Consider any element, v, lying outside $Col(\mathcal{M})$. Let $N_v = \min\{N : v \in \mathcal{N}, |Col(\mathcal{N}) \cap Col(\mathcal{M})| = m - 1\}$. We make the following claim.

Claim 5.1. *$N_v > M$ for all $v \in \mathbf{A} \backslash Col(\mathcal{M})$ iff $Col(\mathcal{M}) = Col(\mathcal{T}_0)$.*

Proof. One of the directions of the implication is clear. If $Col(\mathcal{M}) = Col(\mathcal{T}_0)$, then $\mathcal{M} = \mathcal{T}_0$, and by the minimality of $\mathcal{T}_0$ we have $N_v > M$ for all v lying outside $Col(\mathcal{T}_0)$.

The reverse direction has already been established in Lemma 4.6. ∎

Theorem 5.2. *For an $m \times n$ matrix, $\mathbf{A}$, containing i.i.d.* exp(1) *entries,* $T_1 - T_0 \sim \exp(m(n-m))$.

Proof. Let $v \in \mathbf{A} \backslash Col(\mathcal{T}_0)$ and let $\mathcal{M}_v$ be the submatching of $\mathcal{N}_v$ (defined in Claim 5.1) such that $\mathcal{N}_v = v \cup \mathcal{M}_v$. Suppose $v > T_0 - M_v, \forall v \in \mathbf{A} \backslash Col(\mathcal{T}_0)$. Then Claim 5.1 implies that this is a necessary and sufficient condition to characterize the columns of $\mathcal{T}_0$.

We recall a well-known fact regarding exponentially distributed random variables.

Fact 5.3. Suppose $X_i,\ i = 1, \ldots, l$, are i.i.d. exp(1) random variables. Let $Y_i \geq 0,\ i = 1, \ldots, l$, be random variables such that $\sigma(Y_1, \ldots, Y_l) \subset \mathcal{F}$ for some σ-algebra $\mathcal{F}$. If $X_i \perp\!\!\!\perp \mathcal{F}\ \forall\ i$, then on the event $\{X_i > Y_i,\ i = 1, \ldots, l\}$, $X_i - Y_i$ are i.i.d. exp(1) random variables and independent of $\mathcal{F}$.

The above fact implies that the random variables $\{v - (T_0 - M_v),\ v \in \mathbf{A} \backslash Col(\mathcal{T}_0)\}$ are i.i.d. exp(1).

From Lemma 4.1, $\mathcal{T}_1$ has exactly one entry outside $Col(\mathcal{T}_0)$. Hence $T_1 - T_0 = \min_v N_v - T_0 = \min_v(v - (T_0 - M_v))$. Since the minimization is over $m(n-m)$ independent exp(1) random variables $v - (T_0 - M_v)$, we have that $T_1 - T_0 \sim \exp m(n-m)$. ∎

Remark 5.4. A theorem in [17] considers a slightly more general setting of matchings of size k in an $m \times n$ matrix. The argument used in Theorem 5.2 is an extension of the argument in [20] for an $(n-1) \times n$ matrix. A similar argument was also used by Janson in [9] for a problem regarding shortest paths in exponentially weighted complete graphs.

We note the following positivity condition that follows immediately from the proof of Theorem 5.2.

Remark 5.5. For any $v \notin Col(\mathcal{T}_0)$, $v - (T_1 - T_0) > 0$.

Proof. We know from the proof of Theorem 5.2 that, for any $v \notin Col(\mathcal{T}_0)$,

$$v - (T_0 - M_v) \geq \min_v(v - (T_0 - M_v)) = \min_v N_v - T_0 = T_1 - T_0.$$

This implies that $v - (T_1 - T_0) \geq (T_0 - M_v)$. Now, let v_0 be the entry of $\mathcal{T}_0$ in the same row as v. Consider the set of all matchings of size $m-1$ in $Col(\mathcal{T}_0)$ that do not contain an element in the same row as v. Then, both $\mathcal{T}_0 \backslash v_0$ and $\mathcal{M}_v$ are members of this set. But $\mathcal{M}_v$ has the smallest weight in this set. Hence $M_v \leq T_0 - v_0 < T_0$ which finally implies $v - (T_1 - T_0) \geq (T_0 - M_v) > 0$. ∎

Step 2: From $m \times n$ matrices to $(m-1) \times n$ matrices We will now demonstrate the existence of a matrix with one less row that preserves the higher increments as described in Section 3. The matrix **B** is obtained from **A** by applying the two operations Φ and Λ (which we will shortly define), as depicted below:

$$\mathbf{A} \xrightarrow{\Phi} \mathbf{A}^* \xrightarrow{\Lambda} \mathbf{B}.$$

To prevent an unnecessary clutter of symbols, we shall employ the following notation in this section:

- $\mathcal{T}(\mathbf{A}) = \{\mathcal{T}_0, \ldots, \mathcal{T}_m\}$,
- $\mathcal{T}(\mathbf{A}^*) = \{\mathcal{T}_0^*, \ldots, \mathcal{T}_m^*\}$,
- $\mathcal{T}(\mathbf{B}) = \{\mathcal{U}_0, \ldots, \mathcal{U}_{m-1}\}$.

From Lemma 4.1 we know that the matchings $\mathcal{T}_0$ and $\mathcal{T}_1$ have $m-1$ columns in common. Hence there are two well-defined entries, $e \in \mathcal{T}_0$ and $f \in \mathcal{T}_1$, that lie outside these common columns. We now specify the operations Φ and Λ.

Φ : Subtract $T_1 - T_0$ from each entry in $\mathbf{A}\backslash Col(\mathcal{T}_0)$ to get the $m \times n$ matrix $\mathbf{A}^*$. [Note that in the matrix $\mathbf{A}^*$ the entry f becomes $f^* = f - (T_1 - T_0)$.]
Λ : Generate a random variable X, independent of all other random variables, with $\mathbb{P}(X = 0) = \mathbb{P}(X = 1) = \frac{1}{2}$. If $X = 0$, then remove the row of $\mathbf{A}^*$ containing e, else remove the row containing f^*. Denote the resulting matrix of size $(m-1) \times n$ by **B**.

Remark 5.6. The random variable X is used to break the tie between the two matchings $\mathcal{T}_0^*$ and $\mathcal{T}_1^*$, both of which have the same weight (this shall be shown in Lemma 5.8). This randomized tie-breaking is essential for ensuring that **B** has i.i.d. exp(1) entries; indeed, if we were to choose e (or f^*) with probability 1, then the corresponding **B** would not have i.i.d. exp(1) entries.

Claim 5.7. *The entries of* $\mathbf{A}^*$ *are all positive.*

Proof. The entries in $Col(\mathcal{T}_0)$ are left unchanged by Φ; hence they are positive. Corollary 5.5 establishes the positivity of the entries in the other columns. ∎

Lemma 5.8. *The following statements hold:*

(i) $T_1^* = T_0^* = T_0$.
(ii) For $i \geq 1$, $T_{i+1}^* - T_i^* = T_{i+1} - T_i$.

Proof. Since $\mathcal{T}_0$ is entirely contained in the submatrix $Col(\mathcal{T}_0)$, its weight remains the same in A^*. Let $\mathcal{R}(\mathbf{A}^*)$ be the set of all matchings of size m in $\mathbf{A}^*$ that contain exactly one element outside $Col(\mathcal{T}_0)$. Then, every matching in $\mathcal{R}(\mathbf{A}^*)$ is lighter by exactly $T_1 - T_0$ compared to its weight in $\mathbf{A}$.

Thus, by the definition of $\mathcal{T}_1$, every matching in $\mathcal{R}(\mathbf{A}^*)$ has a weight larger than (or equal to) $T_1 - (T_1 - T_0) = T_0$. In other words, every size-m matching in $\mathbf{A}^*$ that has exactly one element outside $Col(\mathcal{T}_0)$ has a weight larger than (or equal to) T_0. Therefore, from Lemma 4.6 it follows that $\mathcal{T}_0$ is also the smallest matching in $\mathbf{A}^*$. Thus, we have $\mathcal{T}_0^* = \mathcal{T}_0$, and $T_0^* = T_0$.

From Lemma 4.1 we know that $\mathcal{T}_i^*,\ i \geq 1$, has exactly one element outside the columns of $Col(\mathcal{T}_0^*)$ $(= Col(\mathcal{T}_0))$. Hence, it follows that

$$T_i^* = T_i - (T_1 - T_0) \qquad \text{for } i \geq 1.$$

Substituting $i = 1$, we obtain $T_1^* = T_0$. This proves part (i). And considering the differences $T_{i+1}^* - T_i^*$ completes the proof of part (ii). ∎

To complete Step 2 of the induction we need to establish that $\mathbf{B}$ has the following properties.

Lemma 5.9. $U_i - U_{i-1} = T_{i+1} - T_i,\ i = 1, 2, \ldots, m-1.$

Proof. The proof of the lemma consists of establishing the following: For $i \geq 1$

$$\begin{aligned} T_{i+1} - T_i &\overset{(a)}{=} T_{i+1}^* - T_i^* \\ &\overset{(b)}{=} U_i - U_{i-1}. \end{aligned}$$

Observe that (a) follows from Lemma 5.8. We shall prove (b) by showing that

$$T_i^* = U_{i-1} + v,\ i = 1, \ldots, m. \tag{5.1}$$

for some appropriately defined value v.

Remark 5.10. Since $T_1^* = T_0^*$, Eq. (5.1) additionally shows that $T_0^* = U_0 + v$.

Two cases arise when applying the operation Λ: (1) e and f^* are in the same row, and (2) they lie in different rows. (Note that in Case 1, irrespective of the outcome of X, the common row will be removed.) As observed before, since f is in some column outside $Col(\mathcal{T}_0)$, its value is modified by the operation Φ to $f^* = f - (T_1 - T_0)$. The value of e, however, is left unchanged by the operation Φ. For simplicity, we will use the symbols e and f^* for both the names and the values of these entries.

Case 1: In this case, we claim that $e = f^*$ (as values). To see this, let $\mathcal{M}$ be the smallest matching of size $m - 1$ in the columns $Col(\mathcal{T}_0) \cap Col(\mathcal{T}_1)$ which does not have an entry in the same row as e and f^*. Then, clearly, $e \cup \mathcal{M} = \mathcal{T}_0$ and $f \cup \mathcal{M} = \mathcal{T}_1$. Hence, we obtain $e + M = T_0 = T_1 - (T_1 - T_0) = f + M - (T_1 - T_0) = f^* + M$. Therefore, in value, $e = f^*$; call this value v. From Lemma 5.8 we know that $T_0^* = T_0$ and this implies $e + M = T_0^* = f^* + M$.

Now consider any matching, $\mathcal{M}' \neq \mathcal{M}$, of size $m-1$ in $\mathbf{B}$ that has exactly one entry outside $Col(\mathcal{T}_0) \cap Col(\mathcal{T}_1)$. Clearly, one (or both) of the entries e and f^* could have chosen $\mathcal{M}'$ to form a candidate for $\mathcal{T}_0^*$. Since $v + M' > T_0^* = v + M$, we infer that $M' > M$ for all matchings $\mathcal{M}'$. Thus, from Lemma 4.6, we have that $\mathcal{M}$ equals $\mathcal{U}_0$. Therefore, $T_0 = T_0^* = T_1^* = U_0 + v$. This also implies that $Col(\mathcal{U}_0) = Col(\mathcal{T}_0) \cap Col(\mathcal{T}_1)$.

Next consider $\mathcal{S}_\ell^*$, the smallest matching in $\mathbf{A}^*$ obtained by deleting column $\ell \in Col(\mathcal{U}_0)$. Since this is T_k^* for some $k \geq 2$, $\mathcal{S}_\ell^*$ must use one of the entries e or f^* by Lemma 4.10. Hence $S_\ell^* = v + V_\ell$, where $\mathcal{V}_\ell$ is a matching of size $m-1$ in $\mathbf{B}$ that doesn't use the column ℓ. Therefore, $S_\ell^* \geq v + W_\ell$, where $\mathcal{W}_\ell$ is the smallest matching of size $m-1$ in $\mathbf{B}$ that doesn't use column ℓ.

Remark 5.11. The nonuniqueness amongst the weights of matchings introduced by forcing $T_1^* = T_0^*$ does not affect the applicability of Lemma 4.10. Though we could resort to the generalized definition of S-matchings as defined by Claim 4.11; in this case, it is not necessary as with probability 1, it is easy to see that there is a unique matching $\mathcal{S}_j$ in every $\mathbb{S}_j$.

We will now show that for $S_\ell^* \leq v + W_\ell$. Applying Lemma 4.1 to $\mathbf{B}$, we have that $\mathcal{W}_\ell$ has exactly one element outside $Col(\mathcal{U}_0)$. Therefore, $\mathcal{W}_\ell$ can pick either e or f^*, since both lie outside $Col(\mathcal{U}_0)$, to form a candidate for S_ℓ^*, with weight $v + W_\ell$. This implies $S_\ell^* \leq v + W_\ell$. Hence,

$$S_\ell^* = v + W_\ell. \tag{5.2}$$

But from Corollary 4.3 we know that arranging the matchings $\{\mathcal{S}_\ell^*,\ \ell \in Col(\mathcal{T}_0) \cap Col(\mathcal{T}_1)\}$, in increasing order gives us $T_2^*, \ldots, T_m^*$. And arranging the $\{W_\ell,\ \ell \in Col(\mathcal{U}_0) = Col(\mathcal{T}_0) \cap Col(\mathcal{T}_1)\}$ in increasing order gives us $U_1, \ldots, U_{m-1}$. Therefore,

$$T_i^* = U_{i-1} + v \qquad \text{for } i = 1, \ldots, m. \tag{5.3}$$

This proves the lemma under Case 1, i.e., when both the entries e and f are in the same row.

Case 2: In this case, the entries e and f^* are in different rows and depending on the outcome of X, one of these two rows is removed. Let us denote by v the entry e or f^* (depending on X), that is in the row of $\mathbf{A}^*$ removed by Λ. Further, let $\mathbf{c}$ be the column in which v lies. Let $\mathcal{M}$ denote the matching of size $m-1$ in $Col(\mathcal{T}_0) \cap Col(\mathcal{T}_1)$ that v goes with to form $\mathcal{T}_0^*$ or $\mathcal{T}_1^*$ (depending on which of the two entries e or f^* is removed). Let us denote the entry, e or f^*, that was *not* removed by u. Let $\mathbf{d}$ be the column in which u lies. Let w denote the entry in the column of u and the row of v. These are represented in Fig. 3, where the entries of $\mathcal{T}_0$ and $\mathcal{T}_1$ are depicted by stars and circles, respectively. In the figure we assume that the row containing e was chosen to be removed by X (that is, $v = e$ and $u = f^*$).

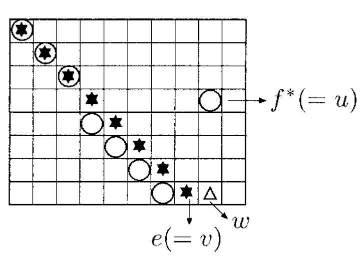

Fig. 3. The entries e, f^*, w.

As in Case 1, let $\mathcal{M}$ be the smallest matching of size $m-1$ in $\mathbf{B}$ that is contained in the columns $Col(\mathcal{T}_0) \cap Col(\mathcal{T}_1)$. Arguing as in the previous case yields $v+M = T_0 = T_0^* = T_1^*$.

This also implies that $w+M > T_0^* = T_0$. (In general, the definition of $\mathcal{T}_0^*$ only implies $w+M \geq T_0$. However, since the matchings in $\mathbf{A}$ have distinct weights, it is not hard to see that strict inequality holds when w is different from e and f.) Therefore, let $w = v + x$ for some $x > 0$.

Remark 5.12. In the claim that follows, we will use a slightly unconventional method to prove a combinatorial fact implied by Eq. (5.1). We believe it will be helpful to preface the proof by a brief description of the steps involved. Consider the elements v and w as defined above. First, we will reduce the value of w from $v+x$ to $v+\epsilon, x > \epsilon > 0$, and show that this does not alter the values of the matchings $\mathcal{T}_i^*, i \geq 0$. Next, we will perturb the value of both v and w slightly to $v-\epsilon$. By invoking Lemma 4.10 we will show that every matching $\mathcal{T}_i^*$ for the new matrix must use one of v or w. Moreover, we will also show that the matchings $\{\mathcal{T}_i^*\}$ are formed by combining v or w with the matchings $\{\mathcal{U}_i\}$. Since the values of the T-matchings are continuous in the entries of the matrix, we let ϵ tend to zero to conclude Eq. (5.1) for Case 2. A purely combinatorial argument also exists for this case which goes along the lines of Lemma 4.8. However, we feel that this approach is simpler.

Returning to the proof: Given any $0 < \epsilon < x$, let $\mathbf{C}^\epsilon$ be a matrix identical to $\mathbf{A}^*$ in every entry except w. The value of w is changed from $v+x$ to $v+\epsilon$. Let $\{\mathcal{P}_i\}$ denote the T-matchings of $\mathbf{C}^\epsilon$. Also recall that $\mathbf{c}$ is the column of v, and $\mathbf{d}$ is the column of both u and w.

Claim 5.13. *$P_i = T_i^*$ for every i.*

Proof. Since the only entry that was modified was w, it is clearly sufficient to show that w is not used by any of the matchings $\{\mathcal{T}_i^*\}$ or $\{\mathcal{P}_i\}$. From Lemma 4.10 we know that the matchings $\{T_i^*\}$ have only two marked elements in the row of w and one of them is v. The matching T_0^* or T_1^* (depending on the outcome of X) contains u and cannot use any entry from the column of v. Hence it must use another entry from the row of v (distinct also from w, as w lies in the column of u). Thus, since w is not one of the two marked elements in its row, it is not part of any T_i^*.

Now we have to show that w is not present in any of the $\{\mathcal{P}_i\}$. To establish this, we exhibit two distinct marked elements in the row of w that are different from w. Consider $\mathcal{S}_\mathbf{d}$: the smallest size m matching in $\mathbf{C}^\epsilon \backslash \mathbf{d}$. But the removal of column $\mathbf{d}$ in both $\mathbf{C}^\epsilon$ and $\mathbf{A}^*$ leads to the same $m \times n-1$ matrix. Hence, $\mathcal{S}_\mathbf{d}$ is formed by the entry v and $\mathcal{M}$, where $\mathcal{M}$ is the matching defined earlier. This implies v is a marked element.

Since $v+M = T_0^*$, it is clear that $\mathcal{M}$ is also the smallest matching of size $m-1$ in the matrix $\mathbf{B}\backslash\mathbf{c}$. Otherwise, v will pick a smaller matching and contradict the minimality of $\mathcal{T}_0^*$.

Consider next the matching $\mathcal{S}_\mathbf{c}$, the smallest matching in $\mathbf{C}^\epsilon$ obtained by deleting column $\mathbf{c}$. The only candidates we have to consider are the matchings involving w and the matching of weight T_0^* involving the element u. The smallest matching of size $m-1$ in the matrix $\mathbf{B}\backslash\mathbf{c}$ is $\mathcal{M}$, which implies that the best candidate for $\mathcal{S}_\mathbf{c}$ involving w is the matching formed by w and $\mathcal{M}$. However, this has weight $v+\epsilon+M > v+M = T_0^*$. Hence $\mathcal{S}_\mathbf{c}$ is the matching of weight T_0^* involving the element u. As before, this matching marks another element in the row of w which is different from either v or w. Since there are two marked elements in the row of w which are different from w, w cannot be in any of the matchings $\{\mathcal{P}_i\}$.

Thus the entry w is in neither of the set of matchings $\{\mathcal{T}_i^*\}$ or $\{\mathcal{P}_i\}$. Since w is the only entry that the two matrices $\mathbf{A}^*$ and $\mathbf{C}^\epsilon$ differ in, this proves the claim. ∎

Moving to the next step of the proof for Case 2, define a matrix $\mathbf{D}^\epsilon$ which is identical to the matrix $\mathbf{A}^*$ except for the entries v and w. We change the values of both v and w to $v - \epsilon$. Let the T-matchings of $\mathbf{D}^\epsilon$ be denoted by $\{\mathcal{Q}_i\}$.

Consider $\mathcal{S}_\mathbf{d}$, the smallest matching of size m in $\mathbf{D}^\epsilon \backslash \mathbf{d}$. It is easy to see that since v was the only entry that was modified in this submatrix, $\mathcal{S}_\mathbf{d}$ is formed by the entry v and the matching $\mathcal{M}$, and has weight $T_0 - \epsilon$. Hence v is a marked element.

Next, let $\mathcal{S}_\mathbf{c}$ be the smallest matching in $\mathbf{D}^\epsilon \backslash \mathbf{c}$. The only candidates we have to consider are the matchings involving w and the matching of weight T_0^* that includes the element u. As before, the smallest matching of size $m-1$ in the matrix $\mathbf{B}\backslash \mathbf{c}$ is $\mathcal{M}$ which implies that the best candidate for $\mathcal{S}_\mathbf{c}$ involving w is the matching formed by w and $\mathcal{M}$. This has weight $v - \epsilon + M < v + M = T_0^*$. Hence $\mathcal{S}_\mathbf{c}$ is the matching of weight $T_0 - \epsilon$ involving the element w. Hence w is a marked element.

Applying Lemma 4.10 to matrix $\mathbf{D}^\epsilon$, it is clear that the only two marked elements in the row of v are v and w. An argument similar to the one that proved (5.3) gives us the following:

$$Q_i = U_{i-1} + v - \epsilon, \text{ for } i = 1, 2, \ldots, m. \tag{5.4}$$

As $\epsilon \to 0$, the matrices $\mathbf{C}^\epsilon$ and $\mathbf{D}^\epsilon$ tend to each other. Since the weights of the T-matchings are continuous functions of the entries of the matrix, we have that in the limit $\epsilon = 0$, $P_i = Q_i$, and hence from Claim 5.13 and Eq. (5.4) we have

$$T_i^* = U_{i-1} + v \qquad \text{for } i = 1, 2, \ldots, m.$$

This proves the lemma for Case 2 and hence completes the proof of Lemma 5.9. ∎

We now note the following consequence of our previous arguments:

$$v + M = T_0 = T_0^* = T_1^* = U_0 + v.$$

This gives us the following:

Remark 5.14. Let $\mathcal{M}$ be the smallest matching of size $m-1$ in $\mathbf{A}^*$, contained in $Col(\mathcal{T}_0) \cap Col(\mathcal{T}_1)$. Then $M = U_0$.

In the next section we show that the matrix $\mathbf{B}$, obtained by deleting a row of $\mathbf{A}^*$ according to the action Λ, contains i.i.d. exp(1) entries.

Step 3: B has i.i.d. exp(1) **entries**. Let $\mathbf{B}$ be a fixed $(m-1) \times n$ matrix of positive entries. We compute the joint distribution of the entries of $\mathbf{B}$ and verify that they are i.i.d. exp(1) random variables. To do this, we identify the set, $\mathfrak{D}$, of all $m \times n$ matrices, $\mathbf{A}$, that have a positive probability of mapping to the particular realization of $\mathbf{B}$ under the operations Φ and Λ. We know that the entries of $\mathbf{A}$ are i.i.d. exp(1) random variables. So we integrate over $\mathfrak{D}$ to obtain the joint distribution of the entries of $\mathbf{B}$.

To simplify the exposition, we partition the set $\mathfrak{D}$ into sets $\{\mathfrak{D}_1, \ldots, \mathfrak{D}_m\}$ depending on the row removed by the operation Λ to obtain $\mathbf{B}$. We will characterize $\mathfrak{D}_m$, i.e., the set of all $m \times n$ matrices in which Λ removes the last row. All the other sets $\mathfrak{D}_i,\ i \neq m$, can be characterized similarly. The next few lemmas concern the complete characterization of the set $\mathfrak{D}_m$.

Let $\mathbf{B}$ be a fixed $(m-1) \times n$ matrix of positive entries. Let $\mathcal{D}_\Lambda = \Lambda_m^{-1}(\mathbf{B})$, i.e., the set of all matrices $m \times n$ matrices $\mathbf{A}^*$ such that when its last row is removed, one obtains the matrix $\mathbf{B}$. Now Λ is a random map, whose action depends on the value of X. This is related to e and f being on the same or different rows. Therefore, we may write $\mathcal{D}_\Lambda$ as the disjoint union of the sets $\mathcal{D}_\Lambda^s$ and $\mathcal{D}_\Lambda^d$, with the obvious mnemonics. Finally, $\mathfrak{D}_m = \Phi^{-1} \circ \Lambda_m^{-1}(\mathbf{B})$.

Remark 5.15. Since we are focusing just on $\mathfrak{D}_m$, the mapping $\Lambda^{-1}(\mathbf{B})$ from $\mathbb{R}_+^{m-1\times n}$ into $\mathbb{R}_+^{m\times n}$ will consist of the introduction of an additional row below $\mathbf{B}$ (hence the subscript Λ_m). When dealing with $\mathfrak{D}_i$, the additional row would be introduced after the $(i-1)^{th}$ row of $\mathbf{B}$.

Consider a matrix $\mathbf{M} \in \mathbb{R}_+^{m\times n}$, where the row vector $\vec{r} = (r_1, \ldots, r_{m-1}) \in \mathbb{R}_+^{m-1}$ denotes the elements in $Col(\mathcal{U}_0)$. W.l.o.g. let us assume that $Col(\mathcal{U}_0) = \{1, 2, \ldots, m-1\}$.

$$\mathbf{M} = \begin{bmatrix} & & & & \mathbf{B} & & & & \\ r_1 & r_2 & \cdot & \cdot & \cdot & r_{m-1} & x_1 & \cdot & \cdot & x_{n-m+1} \end{bmatrix}.$$

Let d be an element in $\mathbf{B} \setminus Col(\mathcal{U}_0)$. Let Δ_d be the cost of the smallest matching of size $m-1$, say $\mathcal{M}_d$, with entries in $Col(\mathcal{U}_0)$ but containing no entry from the row of d. Clearly $d \cup \mathcal{M}_d$ is a matching of size m in the matrix $\mathbf{M}$. Among all such choices of d, let $d_o \in \mathbf{B}\setminus Col(\mathcal{U}_0)$ be that entry which minimizes $d + \Delta_d$. Let $J = d_o + \Delta_{d_o}$, and denote the column of d_o by $\mathbf{j}$.

Given any $\vec{r} = (r_1, \ldots, r_{m-1}) \in \mathbb{R}_+^{m-1}$, the following lemma stipulates necessary and sufficient conditions that the vector $(x_1, \ldots, x_{n-m+1})$ must satisfy so that $\mathbf{M} \in \mathcal{D}_\Lambda$.

Lemma 5.16. *Given a $(m-1) \times n$ positive matrix $\mathbf{B}$ and a $\vec{r} \in \mathbb{R}_+^{m-1}$, let $\mathcal{F}_\Lambda(\vec{r})$ be the collection of all $m \times n$ matrices $\mathbf{M}$ such that one of the following two conditions hold:*

(*i*) *There exist i and k such that $x_i = x_k$, $x_i + U_0 < J$ and $x_l > x_i$ for all $l \neq i, k$.*
(*ii*) *There exists $x_i \notin \mathbf{j}$ such that $x_l > x_i$ for all $l \neq i$ and $x_i + U_0 = J$.*

Then $\mathcal{D}_\Lambda = \mathcal{F}_\Lambda \stackrel{\triangle}{=} \bigcup_{\vec{r}\in\mathbb{R}_+^{m-1}} \mathcal{F}_\Lambda(\vec{r})$.

Proof. **(α)** $\mathcal{D}_\Lambda \subset \mathcal{F}_\Lambda$: Let $\mathbf{M} \in \mathcal{D}_\Lambda$ be any matrix such that $\Lambda_m(\mathbf{M}) = \mathbf{B}$. Therefore, $\mathbf{B}$ consists of the first $m-1$ rows of $\mathbf{M}$. By the definition of Λ we know that the entry v occurs in the last row. From Corollary 5.14 we know that v chooses the matching $\mathcal{U}_0$ to form a matching of weight T_0^*, that is, $v + U_0 = T_0^*$. Hence v must be one of the x_i's. Again by definition, u lies outside $Col(\mathcal{U}_0) \cup \mathbf{c}$, where $\mathbf{c}$ is the column of v.

We shall now show that $\mathbf{M} \in \mathcal{F}_\Lambda(\vec{r})$. Two cases occur:

(a) v and u are in the last row: In this case, we know from the proof of Lemma 5.9 that $v = u$ (the same as $e = f^*$ and this is Case 1 in the proof of Lemma 5.9). Since both

v and u are in the last row and outside $Col(\mathcal{U}_0)$, we know that $v = x_i$ and $u = x_k$ for some $i \neq k$. Therefore, $x_i = x_k$. We know that $v + U_0 = T_0^*$, hence from the minimality of T_o^* we have $x_i + U_0 < J$. Also, $x_i + U_0 < x_l + U_0$ for $l \neq i, k$ for the same reason. This implies $\mathbf{M}$ satisfies condition (i) of Lemma 5.16. Therefore, under (a) it follows that $\mathbf{M} \in \mathcal{F}_\Lambda(\vec{r})$.

(b) v is in the last row and u is not: Arguing as before, we conclude that $u = d_o$ and $v = x_i$. Thus, $T_0^* = v + U_0 = d_0 + \Delta_{d_0} = J$. We also know that v and u occur in different columns; hence $v = x_i$ for some $x_i \notin \mathbf{j}$. From the minimality of T_0^*, we also have that $x_i + U_0 < x_l + U_0$ for $l \neq i$. Thus, $\mathbf{M}$ satisfies condition (ii) of Lemma 5.16 and hence $\mathbf{M} \in \mathcal{F}_\Lambda(\vec{r})$.

($\boldsymbol{\beta}$) $\mathcal{F}_\Lambda \subset \mathcal{D}_\Lambda$: Let $\mathbf{M} \in \mathcal{F}_\Lambda(\vec{r})$ for some $\vec{r}$. Then $\mathbf{M}$ satisfies condition (i) or (ii) of Lemma 5.16. Accordingly, this gives rise to two cases:

(a) $\mathbf{M}$ satisfies condition (i): We claim that $\Lambda(\mathbf{M}) = \mathbf{B}$. From Lemma 4.7 we have that $\mathcal{T}_0(\mathbf{M})$ must use all the columns of $\mathcal{U}_0$. This implies that exactly one entry of $\mathcal{T}_0(\mathbf{M})$ lies outside $Col(\mathcal{U}_0)$. But, condition (i) implies that $x_i + U_0 \leq \min\{x_l + U_0, J\} = \min\{x_l + U_0, d + \Delta_d\}$. Since the last minimization is over all possible choices of the lone entry d that $\mathcal{T}_0(\mathbf{M})$ could choose outside $Col(\mathcal{U}_0)$, it follows that $T_0(\mathbf{M}) = x_i + U_0$. Condition (i) also implies that $x_k = x_i$. Hence $T_0(\mathbf{M}) = T_1(\mathbf{M}) = x_k + U_0$.

Since x_i and x_k are the entries of $\mathcal{T}_0(\mathbf{M})$ and $\mathcal{T}_1(\mathbf{M})$ outside $Col(\mathcal{U}_0)$, this implies u and v are x_i and x_k in some order. Observe that Λ removes the row in which v is present. Thus, we obtain $\Lambda(\mathbf{M}) = \mathbf{B}$ and therefore $\mathbf{M} \in \mathcal{D}_\Lambda$.

(b) $\mathbf{M}$ satisfies condition (ii): We claim that $\Lambda(\mathbf{M}) = \mathbf{B}$ with probability $\frac{1}{2}$. An argument similar to that in Case (a) yields $x_i + U_0 = T_0(\mathbf{M}) = T_1(\mathbf{M}) = J = d_o + \Delta_{d_o}$. Note that v and u are decided by the outcome of X. Hence $\mathbb{P}(v = x_i, u = d_o) = \frac{1}{2} = \mathbb{P}(u = x_i, v = d_o)$.

When $v = x_i$, by the definition of Λ we get that $\Lambda(\mathbf{M}) = \mathbf{B}$. When $v = d_o$ the row that is removed is the row containing d_o, hence $\Lambda(\mathbf{M}) \neq \mathbf{B}$ in this case. Therefore, with probability $\frac{1}{2}$ we will obtain $\mathbf{B}$ as the result of the operation $\Lambda(\mathbf{M})$. This implies $\mathbf{M} \in \mathcal{D}_\Lambda$.

Thus both cases in **($\boldsymbol{\beta}$)** imply that $\mathcal{F}_\Lambda \subset \mathcal{D}_\Lambda$, and this, along with **($\boldsymbol{\alpha}$)** implies $\mathcal{F}_\Lambda = \mathcal{D}_\Lambda$. ∎

Thus, $\mathcal{D}_\Lambda^s$ and $\mathcal{D}_\Lambda^d$ correspond to the matrices in $\mathcal{D}_\Lambda$ which satisfy conditions (i) and (ii) of Lemma 5.16, respectively. Hence, when $\mathbf{M} \in \mathcal{D}_\Lambda^s$, we have $\Lambda(\mathbf{M}) = \mathbf{B}$ with probability one, and when $\mathbf{M} \in \mathcal{D}_\Lambda^d$ we have $\Lambda(\mathbf{M}) = \mathbf{B}$ with probability $\frac{1}{2}$. We are now ready to characterize $\mathfrak{D}_m$.

Consider a matrix $\mathbf{M} \in \mathcal{D}_\Lambda$ and let $\theta \in \mathbb{R}_+$. Consider the column, say $\mathbf{k}$, in $\mathbf{M}$ which contains x_i. (Recall, from Lemma 5.16, that x_i is the smallest of the x_l's in the last row deleted by Λ.) Add θ to every entry in $\mathbf{M}$ outside $Col(\mathcal{U}_0) \cup \mathbf{k}$. Denote the resulting matrix by $F_1(\theta, \mathbf{M})$. Let

$$\mathcal{F}_1 = \bigcup_{\theta > 0, \mathbf{M} \in \mathcal{D}_\Lambda} F_1(\theta, \mathbf{M}). \tag{5.5}$$

Now consider the column, say ℓ, in $\mathbf{M}$ where the entry x_k or d_o is present (depending on whether $\mathbf{M}$ satisfies condition (i) or (ii) of Lemma 5.16). Add θ to every entry in $\mathbf{M}$ outside $Col(\mathcal{U}_0) \cup \ell$. Call the resulting matrix $F_2(\theta, \mathbf{M})$, and let

$$\mathcal{F}_2 = \bigcup_{\theta>0, \mathbf{M}\in\mathcal{D}_\Lambda} F_2(\theta, \mathbf{M}). \tag{5.6}$$

Remark 5.17. Note that $\mathcal{F}_1$ and $\mathcal{F}_2$ are disjoint since $\mathbf{k} \neq \ell$. Also, θ is added to precisely $m(n-m)$ entries in $\mathbf{M}$ in each of the two cases above.

Lemma 5.18. $\mathfrak{D}_m = \mathcal{F}_1 \cup \mathcal{F}_2$.

Proof. Consider $\mathbf{M}' \in \mathfrak{D}_m$. Subtracting $\theta = T_1(\mathbf{M}') - T_0(\mathbf{M}')$ from the entries of $\mathbf{M}'$ outside $Col(\mathcal{T}_0(\mathbf{M}'))$ leaves us with $\Phi(\mathbf{M}')$. From the proof of Lemma 5.8 we know that under Φ, the locations of the entries of T-matchings do not change; only the weights of $T_i(\mathbf{M}')$, $i \geq 1$, are reduced by $T_1(\mathbf{M}') - T_0(\mathbf{M}') = \theta$. It is clear that if e and f are in same row, then the last row of $\Phi(\mathbf{M}')$ satisfies condition (i) of Lemma 5.16 and hence $\mathbf{M}' = F_1(\theta, \Phi(\mathbf{M}'))$. If e and f are in different rows then the last row of $\Phi(\mathbf{M}')$ satisfies condition (ii) and therefore $\mathbf{M}' = F_2(\theta, \Phi(\mathbf{M}'))$. This implies $\mathbf{M}' \in \mathcal{F}_1 \cup \mathcal{F}_2$.

For the converse, consider the matrix $\mathbf{M}' = F_1(\theta, \mathbf{M})$ for some $\mathbf{M} \in \mathcal{D}_\Lambda$ and $\theta > 0$. Since $\mathcal{T}_0(\mathbf{M}) = x_i \cup \mathcal{U}_0$ and $\mathbf{M}'$ dominates $\mathbf{M}$ entry-by-entry, $\mathcal{T}_0(\mathbf{M}') = x_i \cup \mathcal{U}_0$ by construction. Consider every size-m matching in $\mathbf{M}'$ that contains exactly one element outside $Col(x_i \cup \mathcal{U}_0)$. By construction, the weight of these matchings exceeds the weight of the corresponding matchings in $\mathbf{M}$ by an amount precisely equal to θ. Using Lemma 4.1, we infer that $T_i(\mathbf{M}') - T_i(\mathbf{M}) = \theta$ for $i \geq 1$. Hence we have $T_1(\mathbf{M}') - T_0(\mathbf{M}') = T_1(\mathbf{M}) - T_0(\mathbf{M}) + \theta$. But, for any $\mathbf{M} \in \mathcal{D}_\Lambda$, $T_1(\mathbf{M}) = T_0(\mathbf{M}) = x_i + U_0$. Therefore, $T_1(\mathbf{M}') - T_0(\mathbf{M}') = \theta$.

Now, $\Phi(\mathbf{M}')$ is the matrix that results from subtracting θ from each entry outside the columns containing the matching $\mathcal{T}_0(\mathbf{M}') = x_i \cup \mathcal{U}_0$. But, by the definition of $F_1(\theta, \mathbf{M})$, $\Phi(\mathbf{M}')$ is none other than the matrix $\mathbf{M}$. Therefore, $\mathbf{M}' \in \mathfrak{D}_m$, and $\mathcal{F}_1 \subset \mathfrak{D}_m$.

Next, let $\mathbf{M}' = F_2(\theta, \mathbf{M})$. In this case too, $T_0(\mathbf{M}) = x_k + U_0$ (or $d_o + \Delta_{d_o}$) continues to be the smallest matching in $\mathbf{M}'$. An argument identical to the one above establishes that $\Phi(\mathbf{M}') = \mathbf{M}$. Hence, $\mathbf{M}' \in \mathfrak{D}_m$ and $\mathcal{F}_2 \subset \mathfrak{D}_m$, completing the proof of the lemma. ∎

Remark 5.19. Note that the variable θ used in the characterization of $\mathfrak{D}_m$ precisely equals the value of $T_1(\mathbf{M}') - T_0(\mathbf{M}')$, as shown in the proof of Lemma 5.18.

Continuing, we can partition $\mathfrak{D}_m$ into the two sets $\mathfrak{D}_m^s$ and $\mathfrak{D}_m^d$ as below:

$$\mathfrak{D}_m^s = F_1(\mathbb{R}_+, \mathcal{D}_\Lambda^s) \cup F_2(\mathbb{R}_+, \mathcal{D}_\Lambda^s) \quad \text{and} \quad \mathfrak{D}_m^d = F_1(\mathbb{R}_+, \mathcal{D}_\Lambda^d) \cup F_2(\mathbb{R}_+, \mathcal{D}_\Lambda^d). \tag{5.7}$$

Observe that whenever $\mathbf{M} \in \mathfrak{D}_m^s$, we have $\Phi(\mathbf{M}) \in \mathcal{D}_\Lambda^s$ and hence $\Lambda \circ \Phi(\mathbf{M}) = \mathbf{B}$ with probability 1. For $\mathbf{M} \in \mathfrak{D}_m^d$, $\Phi(\mathbf{M}) \in \mathcal{D}_\Lambda^d$ and $\Lambda \circ \Phi(\mathbf{M}) = \mathbf{B}$ with probability $\frac{1}{2}$. Recall also that $\mathfrak{D} = \cup_{i=1}^m \mathfrak{D}_i$.

Now that we have characterized $\mathfrak{D}$, we return to considering the matrix $\mathbf{A}$ (which has the same structure as $\mathbf{M}$), and "integrate out the marginals" $(r_1, \ldots, r_{m-1})$, $(x_1, \ldots, x_{n-m+1})$, and θ by setting

$$\vec{v} = (\mathbf{B}, \vec{r}, \theta) \quad \text{and} \quad \vec{w} = (\vec{v}, \vec{x}),$$

where $\mathbf{B} \equiv [b_{ij}] \in \mathbb{R}_+^{m-1\times n}$. Let $f_w(\vec{v},\vec{x})$ represent the density of an $\mathbf{M}$ matrix. Then the marginal density $f_v(\vec{v})$ is given by

$$f_v(\vec{v}) = \int_{\mathcal{R}_1} f_w(\vec{v},\vec{x})\, d\vec{x} + \frac{1}{2}\int_{\mathcal{R}_2} f_w(\vec{v},\vec{x})\, d\vec{x}. \tag{5.8}$$

The regions $\mathcal{R}_1$ and $\mathcal{R}_2$ are defined by the set of all $\vec{x}$'s that satisfy conditions (i) and (ii) of Lemma 5.16, respectively. The factor $\frac{1}{2}$ comes from the fact that on $\mathcal{R}_2$, e and f occur on different rows. Therefore, $\mathbf{A}$ is in $\mathfrak{D}^d = \cup_{i=1}^m \mathfrak{D}_i^d$ and will map to the desired $\mathbf{B}$ with probability $\frac{1}{2}$.

On $\mathcal{R}_1$, we have that $x_i = x_k < J - U_0$ for J as in Lemma 5.16. We set $H = J - U_0$, and $u_l = x_l - x_i$ for $l \neq i, k$. Finally, define

$$s_v = b_{1,1} + \cdots + b_{m-1,n} + r_1 + \cdots + r_{m-1} + m(n-m)\theta.$$

Thus, s_v denotes the sum of all of the entries of $\mathbf{A}$ except those in $\vec{x}$. As noted in Remark 5.17 preceding Lemma 5.18, the value θ was added to precisely $m(n-m)$ entries. We have

$$\begin{aligned}\int_{\mathcal{R}_1} f_w(\vec{v},\vec{x})\, d\vec{x} &\overset{(a)}{=} 2m\binom{n-m+1}{2}\int_0^H\iiint_0^\infty e^{-(s_v+(n-m+1)x_i+\sum_{l\neq i,k} u_l)}\prod_{l\neq i,k} du_l\, dx_i \\ &= m(n-m)e^{-s_v}\left(1-e^{-(n-m+1)H}\right).\end{aligned} \tag{5.9}$$

The factor $\binom{n-m+1}{2}$ in equality (a) accounts for the choices for i and k from $\{1,\ldots,n-m+1\}$; the factor m comes from the row choices available (i.e., the regions $\mathfrak{D}_1,\ldots,\mathfrak{D}_m$), and the factor 2 comes because $\mathbf{A}$ belongs to either $\mathcal{F}_1$ or $\mathcal{F}_2$ defined by Eqs. (5.5) and (5.6), respectively.

Similarly, on $\mathcal{R}_2$, we have that $x_i = J - U_0 \overset{\triangle}{=} H$ and we shall set $u_l = x_l - x_i$ for $l \neq i$ to obtain

$$\begin{aligned}\frac{1}{2}\int_{\mathcal{R}_2} f_w(\vec{v},\vec{x})\, d\vec{x} &\overset{(b)}{=} \frac{1}{2}\left[2m(n-m)\iiint_0^\infty e^{-(s_v+(n-m+1)H+\sum_{l\neq i} u_l)}\prod_{l\neq i} du_l\right] \\ &= m(n-m)e^{-s_v}e^{-(n-m+1)H}.\end{aligned} \tag{5.10}$$

In equality (b) above, the factor $n-m$ accounts for the choice of i from $\{1,\ldots,n-m+1\}$; the factor m comes from the row choices available and the factor 2 comes because $\mathbf{A}$ belongs to either $\mathcal{F}_1$ or $\mathcal{F}_2$ defined by Eqs. (5.5) and (5.6), respectively.

Substituting (5.9) and (5.10) into (5.8), we obtain

$$f_v(\vec{v}) = m(n-m)e^{-s_v} = e^{-(b_{1,1}+\cdots+b_{m-1,n})} \times m(n-m)e^{-m(n-m)\theta} \times e^{-(r_1+\cdots+r_{m-1})}.$$

The above equation is summarized in the following lemma.

Lemma 5.20. *For an i.i.d.* exp(1) *matrix* $\mathbf{A}$*, the following hold:*

- (i) $\mathbf{B}$ *consists of i.i.d.* exp(1) *variables.*
- (ii) $\theta = T_1(\mathbf{A}) - T_0(\mathbf{A})$ *is an* exp $m(n-m)$ *random variable.*
- (iii) $\vec{r}$ *consists of i.i.d.* exp(1) *variables.*
- (iv) $\mathbf{B}$, $T_1(\mathbf{A}) - T_0(\mathbf{A})$*, and* $\vec{r}$ *are independent.*

Remark 5.21. It is worth noting that part (*ii*) of Lemma 5.20 provides an alternate proof of Theorem 5.2.

From Lemma 5.9 we know that the increments $\{T_{k+1}(\mathbf{A}) - T_k(\mathbf{A}), k > 0\}$ are a function of the entries of $\mathbf{B}$. Given this and the independence of $\mathbf{B}$ and $T_1(\mathbf{A}) - T_0(\mathbf{A})$ from the above lemma, we get the following:

Corollary 5.22. *$T_{k+1}(\mathbf{A}) - T_k(\mathbf{A})$ is independent of $T_1(\mathbf{A}) - T_0(\mathbf{A})$ for $k > 0$.*

Thus we have established all the three steps mentioned in Section 3 required to prove Theorem 2.4. This completes the proof of Theorem 2.4 and hence establishes Parisi's conjecture.

6. THE COPPERSMITH-SORKIN CONJECTURE

As mentioned in the introduction, Coppersmith and Sorkin [5] conjectured that the expected cost of the minimum k-assignment in an $m \times n$ rectangular matrix $\mathbf{P}$ of i.i.d. exp(1) entries is

$$F(k, m, n) = \sum_{i,j \geq 0,\ i+j<k} \frac{1}{(m-i)(n-j)}. \tag{6.1}$$

Nair [17] has proposed a larger set of conjectures that identifies each term in Eq. (6.1) as the expected value of an exponentially distributed random variable corresponding to an increment of appropriately sized matchings in $\mathbf{P}$. We prove this larger set of conjectures using the machinery developed in Section 5 and therefore establish the Coppersmith-Sorkin conjecture.

We define two classes of matchings for $\mathbf{P}$, called W-matchings and V-matchings, along the lines of the S-matchings and T-matchings. But the W- and V-matchings will be defined for all sizes k, $1 \leq k < m$. Thus, the superscript associated with a matching will denote its size.

We now proceed to define these matchings for a fixed size $k < m$. Denote the smallest matching of size k by $\mathcal{V}_0^k$. Without loss of generality, we assume that $Col(\mathcal{V}_0^k) = \{1, 2, \ldots, k\}$. Let $\mathcal{W}_i^k$ denote the smallest matching in the matrix $\mathbf{P}$ when column i is removed. Note that for $i > k$, $\mathcal{W}_i^k = \mathcal{V}_0^k$.

Definition 6.1 (W-matchings). *Define the matchings $\{\mathcal{V}_0^k, \mathcal{W}_1^k, \ldots, \mathcal{W}_k^k\}$ to be the* W-matchings *of size k.*

Definition 6.2 (V-matchings). *Arrange the matchings $\{\mathcal{V}_0^k, \mathcal{W}_1^k, \ldots, \mathcal{W}_k^k\}$ in order of increasing weights. Then the resulting sequence $\{V_0^k, V_1^k, \ldots, V_k^k\}$ is called the* V-matchings *of size k.*

Finally, we refer to the smallest matching of size m as V_0^m.

We now state the following theorem regarding the distributions of the increments of the V-matchings.

Theorem 6.3. *Let* **P** *be an* $m \times n$ *rectangular matrix,* **P**, *of i.i.d.* exp(1) *entries. The* V-matchings *of* **P** *satisfy the following: for each* k, $1 \leq k \leq m-1$:

$$V_{i+1}^{k} - V_{i}^{k} \sim \exp(m-i)(n-k+i), \qquad 0 \leq i \leq k-1 \tag{6.2}$$

and

$$V_0^{k+1} - V_k^k \sim \exp(m-k)n. \tag{6.3}$$

Remark 6.4. We have grouped the increments according to the size of the matchings; so Eqs. (6.2) and (6.3) both concern the kth group. Equation (6.2) gives the distribution of the differences of matchings of size k. The matching V_0^{k+1} is the smallest one of size $k+1$, and Eq. (6.3) concerns the distribution of its difference with V_k^k.

Before we prove Theorem 6.3, we show how it implies the Coppersmith-Sorkin conjecture.

Corollary 6.5.

$$F(k,m,n) = \sum_{i,j \geq 0, i+j<k} \frac{1}{(m-i)(n-j)}. \tag{6.4}$$

Proof. By definition $F(j+1,m,n) - F(j,m,n) = \mathbb{E}(V_0^{j+1} - V_0^{j})$. Using Eqs. (6.2) and (6.3) and by linearity of expectation, we obtain

$$F(j+1,m,n) - F(j,m,n) = \sum_{0 \leq i \leq j} \frac{1}{(m-i)(n-j+i)}. \tag{6.5}$$

Now, using the fact that $\mathbb{E}(V_0^1) = 1/mn$ and summing (6.5) over $j=0$ to $j=k-1$, we obtain

$$F(k,m,n) = \frac{1}{mn} + \sum_{j=1}^{k-1} \sum_{0 \leq i \leq j} \frac{1}{(m-i)(n-j+i)} = \sum_{i,j \geq 0, i+j<k} \frac{1}{(m-i)(n-j)}. \tag{6.6}$$

Thus Theorem 6.3 establishes the Coppersmith-Sorkin conjecture. ∎

We now proceed to the proof of Theorem 6.3.

Remark 6.6. We will establish the theorem for the kth group inductively. The outline of the induction is similar to the one in Section 3, and the details of the proof are similar to those in Section 5. The key trick that will be used in this section is a zero-padding of the matrices under consideration in such a way that increments of the V-matchings of the zero padded matrix (the matrix $\mathbf{L}'$ defined below) and the actual matrix (the matrix $\mathbf{L}$ defined below) is identical.

6.1. Proof of Theorem 6.3

In this section we will establish properties concerning the increments of the V-matchings in the kth group of the cost matrix $\mathbf{P}$, i.e. the increments between the matchings $\{\mathcal{V}_0^k, \ldots, \mathcal{V}_k^k, \mathcal{V}_0^{k+1}\}$. Let $\mathbf{L}$ denote an $l \times n$ matrix with $l \leq m$. Consider its V-matchings of size $\gamma = k - m + l$ and denote them as $\{\mathcal{L}_0^\gamma, \ldots, \mathcal{L}_\gamma^\gamma\}$. Let $\mathcal{L}_0^{\gamma+1}$ denote the smallest matching of size $\gamma + 1$ in $\mathbf{L}$.

Inductive Hypothesis:

- The entries of $\mathbf{L}$ are i.i.d. exp(1) random variables.
- The increments satisfy the following combinatorial identities

$$\begin{aligned} L_1^\gamma - L_0^\gamma &= V_{m-l+1}^k - V_{m-l}^k, \\ L_2^\gamma - L_1^\gamma &= V_{m-l+2}^k - V_{m-l+1}^k, \\ &\cdots\cdots\cdots \\ L_\gamma^\gamma - L_{\gamma-1}^\gamma &= V_{m-l+\gamma}^k - V_{m-l+\gamma-1}^k, \\ L_0^{\gamma+1} - L_\gamma^\gamma &= V_0^{k+1} - V_k^k. \end{aligned} \tag{6.7}$$

Induction Step:

Step 1: From $\mathbf{L}$, form a matrix $\mathbf{Q}$ of size $l - 1 \times n$. Let $\{\mathcal{Q}_0^{\gamma-1}, \ldots, \mathcal{Q}_{\gamma-1}^{\gamma-1}\}$ denote its V-matchings of size $\gamma - 1$ and let $\mathcal{Q}_0^\gamma$ denote the smallest matching of size γ. We require that

$$\begin{aligned} Q_1^{\gamma-1} - Q_0^{\gamma-1} &= L_2^\gamma - L_1^\gamma, \\ Q_2^{\gamma-1} - Q_1^{\gamma-1} &= L_3^\gamma - L_2^\gamma, \\ &\cdots\cdots\cdots \\ Q_{\gamma-1}^{\gamma-1} - Q_{\gamma-2}^{\gamma-1} &= L_\gamma^\gamma - L_{\gamma-1}^\gamma, \\ Q_0^\gamma - Q_{\gamma-1}^{\gamma-1} &= L_0^{\gamma+1} - L_\gamma^\gamma. \end{aligned}$$

Step 2: Establish that the entries of $\mathbf{Q}$ are i.i.d. exp(1) random variables.

This completes the induction step since $\mathbf{Q}$ satisfies the induction hypothesis for the next iteration.

In Step 2 we also show that $L_1^\gamma - L_0^\gamma \sim \exp l(n - \gamma)$ and hence conclude from Eq. (6.7) that $V_{m-l+1}^k - V_{m-l}^k \sim \exp l(n - k + m - l)$.

The induction starts with matrix $\mathbf{L} = \mathbf{P}$ (the original $m \times n$ matrix of i.i.d. entries that we started with) at $l = m$ and terminates at $l = m - k + 1$. Observe that the matrix $\mathbf{P}$ satisfies the inductive hypothesis for $l = m$ by definition.

Proof of the Induction:

Step 1: Form the matrix $\mathbf{L}'$ of size $l \times n + m - k$ by adding $m - k$ columns of zeroes to the left of $\mathbf{L}$ as below

$$\mathbf{L}' = [\mathbf{0} \mid \mathbf{L}].$$

Let $\{\mathcal{T}_0, \ldots, \mathcal{T}_l\}$ denote the T-matchings of the matrix $\mathbf{L}'$. Then, we make the following claim:

Claim 6.7. *Let $\gamma = l - (m-k)$. Then the following hold*

$$\begin{aligned}\mathcal{T}_0 &= \mathcal{L}_0^\gamma,\\ \mathcal{T}_1 &= \mathcal{L}_1^\gamma,\\ &\cdots\\ \mathcal{T}_\gamma &= \mathcal{L}_\gamma^\gamma,\end{aligned}$$

and

$$T_{\gamma+1} = T_{\gamma+2} = \cdots = \mathcal{T}_l = \mathcal{L}_0^{\gamma+1}$$

Proof. Note that any matching of size l in $\mathbf{L}'$ can have at most $m-k$ zeroes. It is clear that the smallest matching of size l in $\mathbf{L}'$ is formed by picking $m-k$ zeroes along with the smallest matching of size γ in $\mathbf{L}$. Thus, $\mathcal{T}_0 = \mathcal{L}_0^\gamma$.

By Lemma 4.1 we know that the other T-matchings in $\mathbf{L}'$ drop exactly one column of $\mathcal{T}_0$. We analyze two cases: First, removing a column of zeroes and, next, removing a column containing an entry of $\mathcal{L}_0^\gamma$.

The removal of any column $\mathbf{c}$ containing zeroes leads to the smallest matching of size l in $\mathbf{L}'\backslash\mathbf{c}$ being a combination of $m-k-1$ zeroes with the smallest matching of size $\gamma+1$ in $\mathbf{L}$. Hence $m-k = l-\gamma$ of the T_i's, corresponding to each column of zeroes, have weight equal to $L_0^{\gamma+1}$.

If we remove any column containing $\mathcal{L}_0^\gamma$, then the smallest matching of size l in $\mathbf{L}$ is obtained by combining $m-k$ zeroes with the smallest matching of size γ in $\mathbf{L}$ that avoids this column. Hence, these matchings have weights L_i^γ for $i \in \{1, 2, \ldots, \gamma\}$.

We claim that $L_0^{\gamma+1}$ is larger than L_i^γ for $i \in \{0, 1, 2, \ldots, \gamma\}$. Clearly $L_0^{\gamma+1} > L_0^\gamma$. Further, for $i \geq 1$, we have a matching of size γ in $\mathcal{L}_0^{\gamma+1}$ that avoids the same column that $\mathcal{L}_i^\gamma$ avoids. But L_i^γ is the smallest matching of size γ that avoids this column. So we conclude that $L_0^{\gamma+1} > L_i^\gamma$.

Hence arranging the weights (in increasing order) of the smallest matchings of size l in $\mathbf{L}'$, obtained by removing one column of $\mathcal{T}_0$ at a time, establishes the claim. ∎

From the above it is clear that the matchings $\mathcal{T}_0$ and $\mathcal{T}_1$ are formed by $m-k$ zeroes and the matchings $\mathcal{L}_0^\gamma$ and $\mathcal{L}_1^\gamma$, respectively. Hence, as in Section 5, we have two elements, one each of $\mathcal{T}_0$ and $\mathcal{T}_1$ that lie outside $Col(\mathcal{T}_0) \cap Col(\mathcal{T}_1)$.

We now perform the procedure outlined in Section 5 for obtaining $\mathbf{Q}$ from $\mathbf{L}$ by working through the matrix $\mathbf{L}'$.

Accordingly, form the matrix $\mathbf{L}^*$ by subtracting the value $T_1 - T_0$ from all the entries in $\mathbf{L}'$ that lie outside $Col(\mathcal{T}_0)$. Generate a random variable X, independent of all other random variables, with $\mathbb{P}(X=0) = \mathbb{P}(X=1) = \frac{1}{2}$. As before, there are two well-defined entries, $e \in \mathcal{T}_0$ and $f \in \mathcal{T}_1$, that lie outside the common columns $Col(\mathcal{T}_0) \cap Col(\mathcal{T}_1)$. [Note that in the matrix, $\mathbf{L}^*$, the entry f has a value $f - (T_1 - T_0)$.] If X turns out to be 0, then remove the row of $\mathbf{L}^*$ containing the entry e; else remove the row containing the entry f. The resulting matrix of size $(l-1) \times n + m - k$ is called $\mathbf{Q}'$. In matrix $\mathbf{Q}'$ remove the $m-k$ columns of zeros to get the matrix $\mathbf{Q}$ of size $(l-1) \times n$.

Let $\{\mathcal{U}_0, \ldots, \mathcal{U}_{l-1}\}$ denote the T-matchings of the matrix $\mathbf{Q}'$ and $\{\mathcal{Q}_0^{\gamma-1}, \ldots, \mathcal{Q}_{\gamma-1}^{\gamma-1}, \mathcal{Q}_0^{\gamma}\}$ denote the V-matchings of the matrix $\mathbf{Q}$. The following follows from Claim 6.7 applied to the zero-padded matrix $\mathbf{Q}'$.

$$\begin{aligned} \mathcal{U}_0 &= \mathcal{Q}_0^{\gamma-1}, \\ \mathcal{U}_1 &= \mathcal{Q}_1^{\gamma-1}, \\ &\cdots \\ \mathcal{U}_{\gamma-1} &= \mathcal{Q}_{\gamma-1}^{\gamma-1}, \end{aligned} \tag{6.7a}$$

and

$$\mathcal{U}_\gamma = \cdots = \mathcal{U}_{l-1} = \mathcal{Q}_0^{\gamma}.$$

Now from Lemma 5.9 in Section 5 we know that

$$T_{i+1} - T_i = U_i - U_{i-1} \qquad \text{for } i = 1, \ldots, l-1. \tag{6.8}$$

Remark 6.8. Though we have used the same notation, please bear in mind that we are referring to two different sets of matchings here and in Section 5. However, since we adopted the same procedure to go from one matrix to the other, the proof continues to hold.

Finally, combining Eq. (6.8), Eq. (6.7a) and Claim 6.7, we obtain

$$\begin{aligned} Q_1^{\gamma-1} - Q_0^{\gamma-1} &= L_2^{\gamma} - L_1^{\gamma}, \\ Q_2^{\gamma-1} - Q_1^{\gamma-1} &= L_3^{\gamma} - L_2^{\gamma}, \\ &\cdots\cdots\cdots \\ Q_{\gamma-1}^{\gamma-1} - Q_{\gamma-2}^{\gamma-1} &= L_{\gamma}^{\gamma} - L_{\gamma-1}^{\gamma}, \\ Q_0^{\gamma} - Q_{\gamma-1}^{\gamma-1} &= L_0^{\gamma+1} - L_{\gamma}^{\gamma}. \end{aligned}$$

This completes Step 1 of the induction.

Step 2: Again we reduce the problem to the one in Section 5 by working with the matrices $\mathbf{L}'$ and $\mathbf{Q}'$ instead of the matrices $\mathbf{L}$ and $\mathbf{Q}$. (Note that the necessary and sufficient conditions for $\mathbf{L}$ to be in the pre-image of a particular realization of $\mathbf{Q}$ is exactly same as the necessary and sufficient conditions for a $\mathbf{L}'$ to be in the pre-image of a particular realization of $\mathbf{Q}'$.)

Let $\mathcal{R}_1$ denote all matrices $\mathbf{L}$ that map to a particular realization of $\mathbf{Q}$ with e and f in the same row. Let $\mathcal{R}_2$ denote all matrices $\mathbf{L}$ that map to a particular realization of $\mathbf{Q}$ with e and f in different rows. Observe that in $\mathcal{R}_2$, $\mathbf{L}$ will map to the particular realization of $\mathbf{Q}$ with probability $\frac{1}{2}$ as in Section 5. We borrow the notation from Section 5 for the rest of the proof.

(Before proceeding, it helps to make some remarks relating the quantities in this section to their counterparts in Section 5. The matrix $\mathbf{A}$ had dimensions $m \times n$; its counterpart $\mathbf{L}'$ has dimensions $l \times (m-k+n)$. The number of columns in $\mathbf{A}\backslash Col(\mathcal{T}_0)$ equaled $n-m$; now the number of columns in $\mathbf{L}'\backslash Col(\mathcal{T}_0)$ equals $m-k+n-l$. This implies that the value $\theta = T_1 - T_0 = L_1^{\gamma} - L_0^{\gamma}$ will be subtracted from precisely $l(m-k+n-l)$ elements of $\mathbf{L}'$. Note also that the vector $\vec{r}$, of length $l-1$, has exactly $m-k$ zeroes and $\gamma = k-m+l-1$ nonzero elements. The vector x is of length $m-k+n-l+1$.)

To simplify notation, set $\eta = m - k + n - l$; the number of columns from which θ is subtracted. Thus, the vector x has length $\eta + 1$. As in Section 5, let

$$\vec{v} = (\mathbf{Q}, \vec{r}, \theta) \qquad \text{and} \qquad \vec{w} = (\vec{v}, \vec{x}).$$

We will evaluate $f_v(\vec{v}) = \int_{\mathcal{R}_1} f_w(\vec{v}, \vec{x})\, d\vec{x} + \frac{1}{2}\int_{\mathcal{R}_2} f_w(\vec{v}, \vec{x})\, d\vec{x}$, to obtain the marginal density of $\vec{v}$. As before the factor $\frac{1}{2}$ comes from the fact that on $\mathcal{R}_2$, e and f occur on different rows. Therefore, $\mathbf{L}$ will map to the desired $\mathbf{Q}$ with probability $\frac{1}{2}$.

On $\mathcal{R}_1$, we have that $x_i = x_j < H$ for H as in Section 5. (The counterparts of x_a and x_b in Section 5 were x_i and x_k, and these were defined according to Lemma 5.16.) We shall set $u_l = x_l - x_a$ for $l \neq a, b$. Finally, define

$$s_v = q_{1,1} + \cdots + q_{l-1,n} + r_1 + \cdots + r_{k-m+l-1} + l\eta\theta.$$

Thus, s_v denotes the sum of all of the entries of $\mathbf{L}$ except those in $\vec{x}$. We have

$$\int_{\mathcal{R}_1} f_w(\vec{v}, \vec{x})\, d\vec{x} \overset{(a)}{=} 2l\binom{\eta+1}{2}\int_0^H \iiint_0^\infty e^{-(s_v + (q+1)x_a + \sum_{l\neq a,b} u_l)} \prod_{l\neq a,b} du_l\, dx_a$$
$$= l\,\eta\, e^{-s_v}\left(1 - e^{-(q+1)H}\right).$$

The factor $\binom{\eta+1}{2}$ in equality (a) comes from the possible choices for a, b from the set $\{1, \ldots, \eta\}$, the factor l comes from the row choices available as in Section 5, and the factor 2 corresponds to the partition, $\mathcal{F}_1$ or $\mathcal{F}_2$ (defined likewise), that $\mathbf{L}$ belongs to.

Similarly, on $\mathcal{R}_2$, we have that $x_a = H$ and we shall set $u_l = x_l - x_a$ for $l \neq a$ to obtain

$$\frac{1}{2}\int_{\mathcal{R}_2} f_w(\vec{v}, \vec{x})\, d\vec{x} \overset{(b)}{=} \frac{1}{2}\left[2\, l\, \eta \iiint_0^\infty e^{-(s_v + (q+1)H + \sum_{l\neq a} u_l)} \prod_{l\neq a} du_l\right]$$
$$= l\,\eta\, e^{-s_v}\, e^{-(q+1)H}.$$

In equality (b) above, the factor η comes from the choice of positions available to x_a (note that x_a cannot occur in the same column as the entry d_o which was defined in Lemma 5.16). The factor l comes from the row choices available, and the factor 2 is due to the partition, $\mathcal{F}_1$ or $\mathcal{F}_2$, that $\mathbf{L}$ belongs to.

Substituting $\eta = n - k + m - l$ and adding (6.1) and (6.1), we obtain

$$f_v(\vec{v}) = l(n - k + m - l)\, e^{-s_v}$$
$$= e^{-(q_{1,1}+\cdots+q_{l-1,n})} l(n - k + m - l) e^{-l(n-k+m-l)\theta} e^{-(r_1+\cdots+r_{l+k-m-1})}.$$

We summarize the above in the following lemma.

Lemma 6.9. *The following hold:*

(i) $\mathbf{Q}$ *consists of i.i.d.* $\exp(1)$ *variables.*
(ii) $\theta = L_1^\gamma - L_0^\gamma$ *is an* $\exp l(n - k + m - l)$ *random variable.*
(iii) $\vec{r}$ *consists of i.i.d.* $\exp(1)$ *variables and* $m - k$ *zeroes.*
(iv) $\mathbf{Q}$, $L_1^\gamma - L_0^\gamma$, *and* $\vec{r}$ *are independent.*

This completes Step 2 of the induction. ∎

From the inductive hypothesis we have $L_1^\gamma - L_0^\gamma = V_{m-l+1}^k - V_{m-l}^k$. Further let us substitute $m - l = i$. Hence we have the following corollary.

Corollary 6.10. $V_{i+1}^k - V_i^k \sim \exp(m-i)(n-k+i)$ *for* $i = 0, 2, \ldots, k-1$.

To complete the proof of Theorem 6.3, we need to compute the distribution of the "level-change" increment $V_0^{k+1} - V_k^k$. At the last step of the induction, i.e., $l = m - k + 1$, we have a matrix **K** of size $m - k + 1 \times n$ consisting of i.i.d. exp(1) random variables. Let $\{\mathcal{K}_0^1, \mathcal{K}_1^1\}$ denote the V-matchings of size 1. Let $\mathcal{K}_0^2$ denote the smallest matching of size 2. From the induction carried out starting from the matrix **P** to the matrix **K**, we have random variables K_0^1, K_1^1, K_0^2 that satisfy the following: $K_1^1 - K_0^1 = V_k^k - V_{k-1}^k$ and $K_0^2 - K_1^1 = V_0^{k+1} - V_k^k$. The following lemma completes the proof of Theorem 6.3.

Lemma 6.11. *The following holds:* $K_0^2 - K_1^1 \sim \exp(m-k)n$.

Proof. This can be easily deduced from the memoryless property of the exponential distribution; equally, one can refer to Lemma 1 in [17] for the argument. ∎

Remark 6.12. There is a row and column interchange in the definitions of the V-matchings in [17].

Thus, we have fully established Theorem 6.3 and hence the Coppersmith-Sorkin conjecture.

This also gives an alternate proof to Parisi's conjecture since [6] shows that $E_n = F(n, n, n) = \sum_{i=1}^n (1/i^2)$.

7. CONCLUDING REMARKS

This paper provides a proof of the conjectures by Parisi [19] and Coppersmith-Sorkin [6]. In the process of proving these conjectures, we have discovered some interesting combinatorial and probabilistic properties of matchings that could be of general interest. Those related to the resolution of the conjectures have been presented in the paper. Others will appear in forthcoming publications. We mention one particularly interesting property below.

Let **Q** be an $(n-1) \times n$ matrix of i.i.d. exp(1) entries and let $\{\mathcal{T}_i\}$ denote its T-matchings. Let Υ denote the set of all possible configurations of the row-wise minimum entries of **Q**; for example, all the row-wise minima lie in the same column, all lie in distinct columns, etc. Consider any fixed configuration $\xi \in \Upsilon$ and let $\mathcal{T}_i^\xi$ denote the T-matchings conditioned on the event that **Q** has its row-wise minima placement according to ξ. Then the following statement holds:

> *Property 1: The joint distribution of the vector* $\{\mathcal{T}_i^\xi - \mathcal{T}_{i-1}^\xi\}_{i=1}^{n-1}$ *is the same for all placements of the row-wise minima,* $\xi \in \Upsilon$.

On the event, ξ_1, where all the row-wise minima lie in different columns, it is quite easy to show that $T_i^{\xi_1} - T_{i-1}^{\xi_1} \sim \exp i(n-i)$ for $i = 1, \ldots, n-1$ and that these increments are independent. Combining this with Property 1 above, one can obtain an alternate proof of Theorem 2.4 and hence of Parisi's conjecture.

However, the argument we currently have for proving Property 1 uses the machinery in this paper for proving Theorem 2.4. It would be nice if another, simpler, argument could be advanced for proving Property 1 since this would not only yield a simpler proof of Theorem 2.4 but would give some interesting new insight into the problem.

ACKNOWLEDGMENTS

The authors wish to thank the referee for making several valuable suggestions that greatly improved the presentation of the paper.

REFERENCES

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

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

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

[4] R. Brunetti, W. Krauth, M. Mezard, and G. Parisi, Extensive numerical simulations of weighted matchings: Total length and distribution of links in the optimal solution, Europhys Lett 14 (1991), 295.

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

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

[7] D. Coppersmith and G. Sorkin, "On the expected incremental cost of a minimum assignment," Contemporary combinatories, Ed., B. Bollobas, Bolyai Society Mathematical Studies, Volume 10, Chapter 8. Springer, New York, 2002.

[8] 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.

[9] S. Janson, One, two and three times $\log n/n$ for paths in a complete graph with random weights, Combination Probab Compu 8 (1999), 347–361.

[10] 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, New York, 1987, pp. 1–4.

[11] R. M. Karp, A. H. G. Rinnooy Kan, and R. Vohra, Average case analysis of a heuristic for the assignment problem, Math Oper Res, 19(3) (1994), 513–522.

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

[13] S. Linusson and J. Wästlund, A generalization of the random assignment problem, http://www.mai.liu.se/~svlin/

[14] S. Linusson and J. Wästlund, A proof of Parisi's Conjecture on the random assignment problem, Probabil Theory Related Fields 128(3) (2004), 419–440, http://www.mai.liu.se/~svlin/

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

[16] M. Mézard, G. Parisi, and M. A. Virasoro, Spin glass theory and beyond, World Scientific, Singapore, 1987.

[17] C. Nair, Towards the resolution of Coppersmith-Sorkin Conjectures, Proc 40th Annu Allerton Conf Communication, Control and Computing, Allerton, Illinois, 2002, 667–673.

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

[19] G. Parisi, A conjecture on random bipartite matching, Technical Report cond-mat/9801176, XXX LANL Archive, Allerton, Illinois, 1998, 657–666.

[20] M. Sharma and B. Prabhakar, On Parisi's Conjecture for the finite random assignment problem, Proc 40th Annu Allerton Conf Communication, Control and Computing, 2002.

[21] J. M. Steele, Probability theory and combinatorial optimization, CBMS-NSF Regional Conference Series in Applied Mathematics 69, SIAM, Philadelphia, PA, 1997.

[22] M. Talagrand, An assignment problem at high temperature, Ann Probab, 31(2) (2003), 818–848.

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

[24] J. Wästlund, A proof of a conjecture of Buck, Chan, and Robbins on the expected value of the minimum assignment, Random Structures Algorithms 26(1–2) (2004), 237–251.