PROOFS OF THE PARISI AND COPPERSMITH-SORKIN CONJECTURES IN THE RANDOM ASSIGNMENT PROBLEM

A DISSERTATION
SUBMITTED TO THE DEPARTMENT
OF ELECTRICAL ENGINEERING
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Chandrasekhar Madhavan Nair
May 2005

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Balaji Prabhakar Principal Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Abbas El Gamal

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Tsachy Weissman

Approved for the University Committee on Graduate Studies.

This thesis is dedicated to my Mother and Father.

Preface

This thesis concerns the resolution of the Parisi's and Coppersmith-Sorkin Conjectures in the Random Assignment Problem. The assignment problem is an optimization problem of interest in a variety of scenarios. It comes up in assigning jobs to machines to minimize costs, in assigning packets from input line cards to output line cards to maximize throughput in internet routers, in assigning flights and crews to routes to maximize airline revenue; to name a few.

Of interest are both the optimal assignment and the optimal value. In a deterministic setting, a famous algorithm called the Hungarian method provides an efficient way of computing the optimal assignment. In a random environment, i.e. where the costs are modeled as random variables, one is often interested in the properties of the minimum cost assignment.

Based on some numerical evidence Parisi conjectured an expression for the expected value of the minimum cost under the case when the cost variables were independent and exponentially distributed. This conjecture was later extended by Coppersmith and Sorkin to a more general setting.

This thesis makes the following main contribution: It resolves both the Parisi and Coppersmith-Sorkin conjectures regarding the expected cost of the minimum assignment. Further, it completes an argument put forth by Dotsenko that aimed to solve Parisi's conjecture. This thesis also contains some results and conjectures towards the entire distribution of the minimum cost.

Acknowledgements

I would like to begin by acknowledging the contributions of my adviser, Professor Balaji Prabhakar. He has been a constant source of encouragement and motivation throughout my graduate life, with his enthusiasm for my improvement as an academic bordering on the edge of obsessiveness. I would also like to acknowledge his contributions on the technical content in this thesis and for patiently listening through detailed case analysis as I was trying to get a better understanding of the problem. If not for his continued encouragement, I would have given up the attempt to solve these conjectures in any of the many barriers that I encountered during the process. On the personal front, he has been extremely kind to point out my misgivings that has clearly given me a better perspective of myself as seen from the outside.

I would like to acknowledge Professor Abbas El Gamal who has been, in the past few years that I came to know him, a mentor and a research collaborator. The clarity of his thoughts and his methodical approach to problems have influenced me to a great extent. Professor Tsachy Weissman has been an inspiration for me during the last couple of years, both as a research collaborator as well as a mentor. I would like to thank Professor Abbas El Gamal and Professor Tsachy Weissman for kindly agreeing to be in my Orals committee as well as my Reading committee.

There have been lots of people who have contributed both to my intellectual growth as well as personal growth over the years in which this thesis work was done. I would, in particular, like to mention Professors Cover, Diaconis, Kailath, Gill and Boyd who have influenced me in many ways. Denise Murphy, the administrative associate par excellence, has been a shining example of demonstrating how, being methodical is necessary and sufficient condition for multitasking. I would also like to thank the members of Microsoft research theory group, in particular Jennifer Chayes and Christian Borgs, with whom I had a wonderful experience as a research intern during the summer of 2004.

On the technical front, I would also like to thank Mayank Sharma. Without his fearless foray into this problem, it is highly unlikely that I would have ended up with this thesis and the results. In addition to providing a great start to this problem, he has also been kind in listening to my various approaches to solve this problem. Moreover, he has been a good friend throughout my graduate life.

I have been incredibly lucky to have really great friends throughout my graduate life. Shivnath Babu has been a truly remarkable friend and confidante and it is quite difficult to get words that can express my gratitude to him. Devavrat Shah (and later Salvia, his wife) has been a great friend, a great research collaborator and a wonderful person to have known for the last nine years. Rong Pan has been like an elder sibling to me and has given great counsel during difficult times. I would also like to thank my officemates who have made my graduate life very pleasant and enjoyable.

I am sure this work would not have been possible without the distractions outside the academic life. Over the last few years I have developed a passion, bordering on obsession, for riding my bike. The pain that these rides cause was probably the best cleanser of stuck approaches and my mind would be revitalized after a grueling ride. I would like to thank members of the Stanford cycling club and team for giving me company during these rides and in general, for being the way they are; relaxed and motivated, kind and brutal, happy and suffering.

Finally, once again I would like to mention that all this would not have been possible without my true mentors and support crew; my parents, my brother and my grandparents.

Contents

List of Figures

CHAPTER 1

Introduction

This thesis concerns the resolution of the Parisi's and Coppersmith-Sorkin's conjectures that are related to the expected cost of a minimum assignment in the *random assignment problem*. The assignment problem is a fundamental and well-studied problem in the area of combinatorial optimization. Assignment problems arise in a variety of scenarios of practical interest. For example, they arise in allocating jobs to machines to minimize cost; scheduling of packets within crossbar switches in the Internet to maximize throughput; in assigning flights and crews to routes in the airline industry to maximize profits, etc.

In its most general form, the problem can be stated as follows: There are a number of machines and a number of jobs. Any machine can be assigned to perform any job, incurring a cost that may vary depending on both the machine and the job. One is required to perform all the jobs by assigning exactly one machine to each job in such a way that the total cost of the assignment is minimized.

When the total cost of the assignment for all jobs is equal to the sum of the costs for each machine the problem is called the linear assignment problem. Usually when one speaks of the assignment problem without any additional qualifications one implies the linear assignment problem. Other kinds are the quadratic assignment problem, bottleneck assignment problem, etc.

The assignment problem is a special case of another optimization problem known as the transportation problem. The transportation problem is a special case of the maximal flow problem, which in turn is a special case of a linear program. While it is possible to solve any of these problems using the simplex algorithm, a general method for solving any linear program, each problem has more efficient algorithms designed to take advantage of its special structure.

There are algorithms that solve the linear assignment problem within a time bounded by a polynomial expression in the number of machines. The celebrated Hungarian method by Kuhn [K 55] is the earliest known algorithm that showed that the computation of the minimum cost assignment has a polynomial complexity.

This thesis deals with a random version of the assignment problem; i.e., when the costs associated with assigning tasks to machines are modeled by random variables. Here the quantity of interest is the random variable corresponding to the minimum cost among all possible assignments. This thesis resolves the conjectures made by Parisi [Pa 98] and Coppersmith-Sorkin [CS 99] concerning the expected value of the minimum cost.

1.1 Formal description of the problem

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 $\pi : \{1, \dots, n\} \to \{1, \dots, n\}$, 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 [Pa 98] made the following conjecture:

$$\mathbb{E}(C_n) = \sum_{i=1}^{n} \frac{1}{i^2}. \tag{1.1.1}$$

A k-assignment in an $m \times n$ matrix is defined as a set of k elements no two of which belong to the same row or column. Coppersmith and Sorkin [CS 99] 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:

$$C(k, m, n) = \sum_{i,j \geq 0, i+j<k} \frac{1}{(m-i)(n-j)}. \tag{1.1.2}$$

By definition, $C(n, n, n) = \mathbb{E}(C_n)$ and the right hand sides of (1.1.2) and (1.1.1) are equal when $k = m = n$. The proof of this equality can be found in the original paper of Coppersmith and Sorkin [CS 99].

In this thesis, we prove Parisi's conjecture by two different but related strategies. The first builds on the work of Sharma and Prabhakar [SP 02] 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 the work of Nair [Na 02] 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 [SP 02] first. We then show how those arguments also resolve the conjectures in [Na 02].

Before surveying the literature on this problem, it is important to mention that simultaneously and independently of this proof (originally published in [NPS 03]), Linusson and Wästlund [LW 04] have also announced a proof of the Parisi and Coppersmith-Sorkin conjectures based on a quite different approach.

1.2 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 [Al 92] 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 [St 97, CS 99]. 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 [La 93], 1.441 by Goemans and Kodialam [GK 93], and 1.51 by Olin [Ol 92]. The first upper bound of 3 was given by Walkup [Wa 79], who thus demonstrated that $\limsup_n E(C_n)$ is finite. Walkup's argument was later made constructive by Karp *et al* [KKV 94]. Karp [Ka 87] made a subtle use of LP duality to obtain a better upper bound of 2. Coppersmith and Sorkin [CS 99] further improved the bound to 1.94.

Meanwhile, it had been observed through simulations that for large n, $E(C_n) \approx 1.642$ [BKMP 86]. Mézard and Parisi [MP 87] used the non-rigorous *replica method* [MPV 87] of statistical physics to argue that $C^* = \frac{\pi^2}{6}$. (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: $\frac{\pi^2}{6} = \sum_{i=1}^{\infty} \frac{1}{i^2}$.) More interestingly, their method allowed them to determine the density of the edge-cost distribution of the limiting optimal matching. These sharp (but non-rigorous) asymptotic results, and others of a similar flavor that they obtained in several combinatorial optimization problems, sparked interest in the replica method and in the random assignment problem.

Aldous [Al 92] proved that 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 [Al 01] established that $C^* = \pi^2/6$, and obtained the same limiting optimal edge-cost distribution as [MP 87]. 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 [LW 00] 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 [LW 04] that they obtain proofs of the Parisi and Coppersmith-Sorkin conjectures. Buck, Chan and Robbins [BCR 02] 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 [W04] by Wästlund using a modification of the argument in [LW 04].

Alm and Sorkin [AS 02], and Linusson and Wästlund [LW 00] verified the conjectures of Parisi and Coppersmith-Sorkin for small values of k, m and n. Coppersmith and Sorkin [CS 02] 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. In [Do 00] Dotsenko made an incomplete claim regarding the proof of Parisi's conjecture. In this thesis we complete this proof thus showing that his claims were right even though his arguments were incomplete.

An outline of the thesis is as follows: in Section 1.4 we recall some previous work from [SP 02] and state Theorem 1.4.4, whose proof implies a proof of Parisi's conjecture. In Section 1.5 we describe an induction procedure for proving Theorem 1.4.4. We then state and prove some combinatorial properties of matchings in Chapter 2 that will be useful for the rest of the paper. Chapter 3 contains a proof of Theorem 1.4.4. Section 3.2 builds on the work of [Na 02] and contains a proof of Theorem 3.2.3. This implies a proof of the Coppersmith-Sorkin conjecture. In Chapter 4 we present some work towards determining the distribution of the smallest matching and some connections to the non-rigorous assumptions that are present in the methods of the physicists. We now present some conventions that are observed in the rest of the thesis.

1.3 Conventions

(1) The words 'cost' and 'weight' are used interchangeably and mean the same thing; the cost (or weight) of a collection of 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 $\frac{1}{\lambda}$; 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 thesis, 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 1.3.1. Note that all of our claims in Section 2.1 will go through even if we do not assume uniqueness. However, when there is a tie, the claims must be re-worded 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 2.1.13.

1.4 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 i^{th} 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 1.4.1 (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 1.4.2 (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 1.4.3. 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 2.1.2 that they are all distinct.

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

C:

3	6	11
9	2	20

6	**11**
2	20

$\Rightarrow S_1 = 13;$

3	**11**
9	20

$\Rightarrow S_2 = 20;$

3	6
9	**2**

$\Rightarrow S_3 = 5 = T_0.$

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

We now state the conjectures of Sharma and Probhakar in [SP 02] that will establish Parisi's Conjecture. Since we prove the conjectures in this thesis we state it as a theorem rather than as a conjecture.

Theorem 1.4.4. *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 [SP 02] which show how Theorem 1.4.4 implies Parisi's conjecture.

Corollary 1.4.5. *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:

$$\begin{aligned}
\mathbb{E}(C_n) &= \int_0^\infty P(C_n > x)\, dx \\
&= \int_0^\infty P(c_{1j} > x - S_j,\ j = 1, \ldots, n)\, dx \\
&= \int_0^\infty P(c_{1\sigma(j)} > x - T_j,\ j = 0, \ldots, n-1)\, dx \qquad (1.4.1)
\end{aligned}$$

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 (1.4.1) that

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

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

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

$$= t_0 + 1 - \frac{1}{2}e^{-(t_1-t_0)} - \frac{1}{6}e^{-(2t_2-t_0-t_1)} - \cdots - \frac{1}{n(n-1)}e^{-((n-1)t_{n-1}-t_0-\cdots-t_{n-2})}.$$

Therefore,

$$\begin{aligned} \mathbb{E}(C_n) &= \mathbb{E}(T_0) + 1 - \sum_{i=1}^{n-1} \frac{1}{i(i+1)} \mathbb{E}\left(e^{-(iT_i-T_0-\cdots-T_{i-1})}\right) \\ &= \mathbb{E}(T_0) + 1 - \sum_{i=1}^{n-1} \frac{1}{i(i+1)} \mathbb{E}\left(e^{\sum_{j=1}^{i} -j(T_j-T_{j-1})}\right). \end{aligned} \quad (1.4.2)$$

However, from Theorem 1.4.4 (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 (1.4.2) gives

$$\mathbb{E}(C_n) = \mathbb{E}(T_0) + \frac{1}{n^2} + \frac{1}{n}\sum_{i=1}^{n-1} \frac{1}{i}. \quad (1.4.3)$$

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

$$\mathbb{E}(T_j) = \mathbb{E}(T_0) + \sum_{k=1}^{j} \mathbb{E}(T_k - T_{k-1}) = \mathbb{E}(T_0) + \sum_{k=1}^{j} \frac{1}{k(n-k)} \quad \text{(by Theorem 1.4.4).} \quad (1.4.4)$$

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

$$\begin{aligned} \mathbb{E}(S_1) &= \frac{1}{n}\sum_{j=0}^{n-1} \mathbb{E}(T_j) = \frac{1}{n}\mathbb{E}(T_0) + \frac{1}{n}\sum_{j=1}^{n-1}\left(\mathbb{E}(T_0) + \sum_{k=1}^{j} \frac{1}{k(n-k)}\right) \\ &= \mathbb{E}(T_0) + \frac{1}{n}\sum_{k=1}^{n-1} \frac{1}{k}. \end{aligned} \tag{1.4.5}$$

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 (1.4.5) we obtain

$$\mathbb{E}(T_0) = \sum_{k=1}^{n-1} \left(\frac{1}{k^2} - \frac{1}{nk}\right). \tag{1.4.6}$$

Using this at (1.4.3) we get

$$\mathbb{E}(C_n) = \sum_{i=1}^{n-1} \left(\frac{1}{i^2} - \frac{1}{ni}\right) + \frac{1}{n^2} + \frac{1}{n}\sum_{i=1}^{n-1} \frac{1}{i} = \sum_{i=1}^{n} \frac{1}{i^2}. \tag{1.4.7}$$

□

1.5 A Sketch of the proof of Theorem 1.4.4

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

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

Establishing this combinatorial property is one major thrust of the thesis.

3. We will then show that **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 **B** as the starting point in the next step of the induction and proceed.

Remark 1.5.1. We have seen above that $T_1 - T_0$ is independent of **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 **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 1.4.4 involves a combinatorial and a probabilistic part. We develop a number of combinatorial lemmas in the next Chapter. 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 1.4.4 are presented in Chapter 3.

CHAPTER 2

Combinatorial Properties of Assignments

2.1 Some combinatorial properties of matchings

To execute some of the proofs in this chapter, 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 2.1.1. *Consider an $m \times n$ matrix $\mathbf{C}$. For every $j \in Col(\mathcal{T}_0)$, we have $|Col(\mathcal{S}_j) \cap Col(\mathcal{T}_0)| = m-1$.*

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 Figure 2.1 where the bold edges belong to $\mathcal{T}_0$ and the dashed edges belong to $\mathcal{S}_1$.

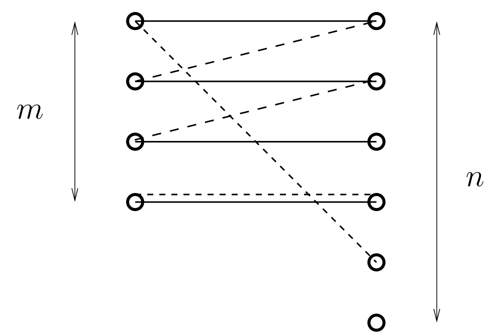

Figure 2.1: Subgraph formed by two matchings depicting an even-length path and a 2-cycle

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 sub-matchings 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 2.1.2. *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 2.1.1, $\mathcal{S}_i$ 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 2.1.3. *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 2.1.1 and the definition of S-matchings. □

We can use Lemma 2.1.1 and Corollary 2.1.3 to give an alternate characterization of the T-matchings that does not explicitly consider the S-matchings.

Lemma 2.1.4 (Alternate Characterization of the T-matchings). *Consider an $m \times n$ rectangular matrix, $\mathbf{C}$. Let $\mathcal{T}_0$ be the smallest matching of size m in this matrix. The rest of the T-matchings $\mathcal{T}_1, ..., \mathcal{T}_m$, 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 $\mathcal{R}_m = \{\mathcal{M} : Col(\mathcal{M}) \supsetneq (Col(\mathcal{T}_0) \cap Col(\mathcal{T}_1) \cdots \cap Col(\mathcal{T}_{m-1}))\}$. Then $\{\mathcal{T}_0, \ldots, \mathcal{T}_m\}$ are the* T-matchings *of* $\mathbf{C}$.

Proof. The proof is straightforward and is omitted. (Note that the alternate characterization was used in the definition of the T-matchings in [Na 02].)

Remark 2.1.5. 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 2.1.6. *Consider an $m \times n$ rectangular matrix,* $\mathbf{C}$*. 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, $\mathbf{C}$, 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 $j \in \{m+1, ..., m+k\}$ 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 2.1.6 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 2.1.1, proving the lemma. □

Lemma 2.1.7. *Consider a $m \times n$ cost matrix* $\mathbf{C}$*. Let* $\mathbf{D}$ *be an extension of* $\mathbf{C}$ *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 $\mathbf{D}$ as a complete bipartite graph $K_{m+r,n}$ and focus on the subgraph (see Figure 2.2) 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 2.1.1, the minimality of the two matchings under consideration prevents

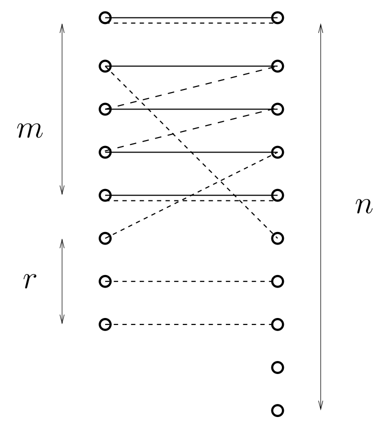

Figure 2.2: Subgraph depicting odd-length paths and a 2-cycle

cycles of 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, $Col(\mathcal{T}_0(\mathbf{C})) \subset Col(\mathcal{T}_0(\mathbf{D}))$. □

Remark 2.1.8. The following lemma was originally stated and proved for the case of an $n-1 \times n$ matrix by Sharma and Prabhakar. The proof of the claim for that case can be found in the thesis of Mayank Sharma.

Lemma 2.1.9. *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 shall first establish this claim for $m = n - 1$. Let us color the edges of $\mathcal{S}_{k_1}$ red (bold), the edges of $\mathcal{S}_{k_2}$ blue (dash) and the edges of $\mathcal{S}_{k_3}$ green (dash-dot). Consider the subgraph formed by the edges present in $\mathcal{S}_{k_1}$ and $\mathcal{S}_{k_2}$, i.e. the red and blue edges (see Figure 2.3 (a)). Clearly 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 two.

- Even length paths with degree-1 vertices on the left, since every vertex on left has degree two.

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. This implies that the single even length alternating path must have vertices 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 red and green edges can only consist of common edges and one even length alternating path, P_{13}, connecting vertices k_1 and k_3. Likewise, in the subgraph formed by green and blue edges we have, other than common edges, exactly one even length alternating path, P_{23}, connecting vertices k_2 and k_3.

We now proceed to 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 matching in $\mathbf{C} \setminus k_3$, say $\tilde{\mathcal{S}}_{k_3}$, using only edges belonging to $\mathcal{S}_{k_1}$, $\mathcal{S}_{k_2}$ and possibly some from $\mathcal{S}_{k_3}$ such that in the subgraph formed by the edges of $\mathcal{S}_{k_1}, \mathcal{S}_{k_2}$ and $\tilde{\mathcal{S}}_{k_3}$, the vertices on the left will have at most degree two. We will show that this new matching $\tilde{\mathcal{S}}_{k_3}$ 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 shall construct $\tilde{\mathcal{S}}_{k_3}$ in each of the following two cases.

- **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 red and green edges. Start traversing the path from k_3 along the red edge. Observe that one takes the red edge when going from a right vertex to a left vertex and a green edge when going from a left vertex to a right vertex. Let v be the first vertex along this path that also belongs to the alternating path, P_{12}, of red and blue edges.

  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 red edges (belonging to each of P_{13} and P_{12}) incident on v. But this is impossible, since the red edges belong to the same matching. Therefore, v must be on the right.

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

  – green edges in the path P_{13} starting from k_3 until vertex v

  – red edges in the path P_{12} starting from v to k_2

  – blue edges in the path P_{12} starting from v to k_1

  – the red edges from all the uncovered vertices on left.

  Note that, by construction of $\tilde{\mathcal{S}}_{k_3}$, on the subgraph formed by the edges of $\mathcal{S}_{k_1}, \mathcal{S}_{k_2}$ and $\tilde{\mathcal{S}}_{k_3}$ the vertices on the left have degree at most two (see Figure 2.3 b).

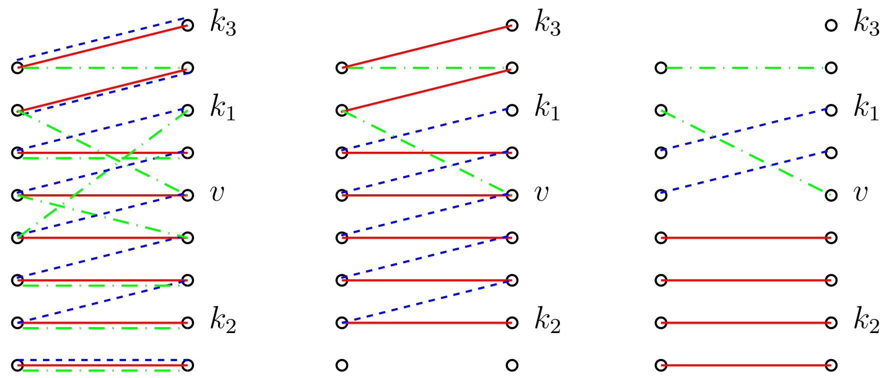

Figure 2.3: (a) Matchings $\mathcal{S}_{k_1}, \mathcal{S}_{k_2}, \mathcal{S}_{k_3}$ (b) P_{13} till vertex v and P_{12} (c) Matching $\tilde{\mathcal{S}}_{k_3}$

- **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:

  - red edges in the path P_{12} starting from k_3 to k_2
  - blue edges in the path P_{12} starting from k_3 to k_1
  - the red edges from all the uncovered vertices on left.

  Observe that, by construction, we again have that on the subgraph formed by the edges of $\mathcal{S}_{k_1}, \mathcal{S}_{k_2}$ and $\tilde{\mathcal{S}}_{k_3}$ the vertices on left have at most degree two.

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 $\mathbf{C} \setminus k_3$. 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

$$\text{sum of edges in } \mathcal{S}'_{k_3} > \text{ sum of edges in } \{E_1, E_2\} = \text{sum of edges in } \tilde{\mathcal{S}}'_{k_3}. \qquad (2.1.1)$$

The right hand side of the above inequality consists only of red and blue edges. Let E_1^c and E_2^c be the remaining red and blue edges, respectively. Adding the weights

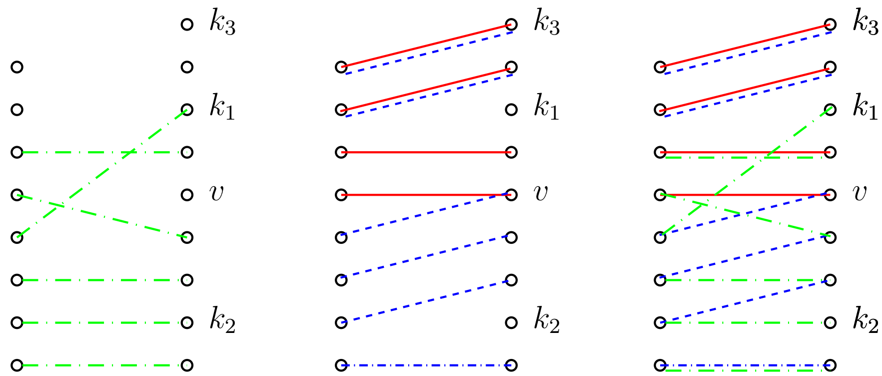

Figure 2.4: (a) Matching $\mathcal{S}'_{k_3}$ (b) Edges E_1^c and E_2^c (c) Matching $\mathcal{S}'_{k_3}$ and edges E_1^c and E_2^c

of these edges to both sides of (2.1.1), we are now required to show

$$\text{sum of edges in } \{\mathcal{S}'_{k_3}, E_1^c, E_2^c\} > S_{k_1} + S_{k_2}. \quad (2.1.2)$$

See Figure 2.4 for an illustration.

We establish (2.1.2) by showing that the left hand side splits into the weights of two matchings, one each in $\mathbf{C} \setminus k_1$ and $\mathbf{C} \setminus k_2$. The minimality of $\mathcal{S}_{k_1}$ and $\mathcal{S}_{k_2}$ will then complete the proof.

First observe that the edges in $\{\mathcal{S}'_{k_3}, E_1^c, E_2^c\}$ can be decomposed into the following:

- An alternating path of (red or blue) and green edges from v to k_1.
- An alternating path of (red or blue) and green edges from v to k_2.
- The common red/blue/green edges that are outside the vertices of P_{12}.

Form the first matching, say $\mathcal{M}$, in $\mathbf{C} \setminus k_2$ by taking the following edges:

- The green edges in the alternating path of red and green edges from v to k_1.
- The (red or blue) edges in the alternating path of blue and green edges from v to k_2.
- One of the red/blue/green edges that are outside the vertices of P_{12}.

Form the other matching, say $\mathcal{N}$, in $\mathbf{C} \setminus k_1$ by taking the following edges:

- The (red or blue) edges in the alternating path of red and green edges from v to k_1.
- The green edges in the alternating path of blue and green edges from v to k_2.
- The other set of red/blue/green edges that are outside the vertices of P_{12}.

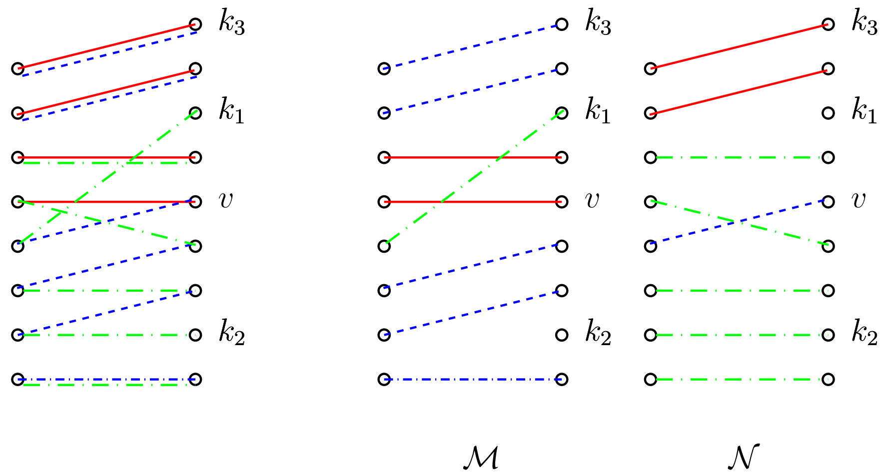

Figure 2.5: (a) $\mathcal{S}'_{k_3} \cup E_1^c \cup E_2^c$ (b) Splitting into matchings $\mathcal{M}$ and $\mathcal{N}$

This splitting into the two matchings establishes (2.1.1) 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 $\mathbf{D} \setminus \mathbf{C}$ 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. $\square$

Definition 2.1.10 (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 2.1.11. *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 2.1.9. □

Lemma 2.1.12. *Let ξ denote the locations row-wise minima of an $m \times n$ matrix $\mathbf{C}$. Let $Col(\xi)$ denote the columns occupied by the row-wise minima. Then we claim the following $Col(\mathcal{T}_0) \supset Col(\xi)$. .*

Proof. The proof is by contradiction.

Assume otherwise: That is, there is a row i_o such that the minimum lies outside $Col(\mathcal{T}_0)$. Form a new matching by replacing the entry of $\mathcal{T}_0$ in the row i_0 by the minimum entry in the row. (Since this entry lies outside $Col(\mathcal{T}_0)$ the new collection of entries is a matching.) This matching has lower weight and contradicts the minimality of $\mathcal{T}_0$.

□

2.1.1 General form of the lemmas

In this section, we state the lemmas for the case when the cost matrix $\mathbf{C}$ has non-unique elements or subset sums. Thus $\mathcal{T}_0$, the smallest matching (in weight), is potentially non-unique. 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 2.1.13. *There exists $\mathcal{S}_j \in \mathbb{S}_j$ for $j = 1, .., m$ such that Lemmas 2.1.1, 2.1.9 and 2.1.11 remain valid when the cost matrix $\mathbf{C}$ has non-unique elements or subset sums.*

Proof. Consider an $m \times n$ matrix $\mathbf{D}$ formed using independent random variables that are uniformly distributed in the interval $[0, 1]$. Define a matrix $\mathbf{C}^\epsilon = \mathbf{C} + \epsilon\mathbf{D}$. Note that for every $\epsilon > 0$, the matrix $\mathbf{C}^\epsilon$ has unique subset sums with probability one. Let $\mathcal{S}_j^\epsilon$ be the S-matchings of the matrix $\mathbf{C}^\epsilon$. Observe that as we let $\epsilon \to 0$, $\mathcal{S}_j^\epsilon \to \mathcal{S}_j \in \mathbb{S}_j$ such that Lemmas 2.1.1, 2.1.9 and 2.1.11 remain valid. □

For Lemma 2.1.4, 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 2.1.6 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 2.1.1; modification caused by the set of the S-matchings chosen according to Claim 2.1.13.

For Lemma 2.1.7 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.

For Lemma 2.1.12 to be valid, we need to state that there is no strict row-wise minima lying outside $Col(\mathcal{T}_0)$.

Remark 2.1.14. Note that though the lemmas are valid for general matrices, unless explicitly stated, we will assume in the rest of the thesis that all subset sums are unique.

CHAPTER 3

Proofs of the Conjectures

3.1 Proof of Theorem 1.4.4

We shall now execute the three steps mentioned in Section 1.5.

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 3.1.1. *$N_v > M$ for all $v \in \mathbf{A} \setminus Col(\mathcal{M})$ 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 2.1.6. □

Theorem 3.1.2. *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 $v \in \mathbf{A} \setminus Col(\mathcal{T}_0)$ and let $\mathcal{M}_v$ be the submatching of $\mathcal{N}_v$ (defined in Claim 3.1.1) such that $\mathcal{N}_v = v \cup \mathcal{M}_v$. Suppose $v > T_0 - M_v,\ \forall\ v \in \mathbf{A} \setminus Col(\mathcal{T}_0)$. Then Claim 3.1.1 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 3.1.3. 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 $\{v - (T_0 - M_v),\ v \in \mathbf{A} \setminus Col(\mathcal{T}_0)\}$ are i.i.d. exp(1).

From Lemma 2.1.1, $\mathcal{T}_1$ 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 3.1.4. A theorem in [Na 02] considers a slightly more general setting of matchings of size k in an $m \times n$ matrix. The argument used in Theorem 3.1.2 is an extension of the argument in [SP 02] for an $(n-1) \times n$ matrix. A similar argument was also used by Janson in [Ja 99] 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 3.1.2.

Remark 3.1.5. For any $v \notin Col(\mathcal{T}_0)$, $v - (T_1 - T_0) > 0$.

Proof. We know from the proof of Theorem 3.1.2 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 $\mathcal{T}_0 \setminus v_0$ 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 1.5. The matrix $\mathbf{B}$ is obtained from $\mathbf{A}$ 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 2.1.1 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.

$\mathbf{\Phi}$: Subtract $T_1 - T_0$ from each entry in $\mathbf{A} \setminus Col(\mathcal{T}_0)$ to get the $m \times n$ matrix $\mathbf{A}^*$. (Note that in the matrix $\mathbf{A}^*$ the entry f becomes $f^* = f - (T_1 - T_0)$).

$\mathbf{\Lambda}$: 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 $\mathbf{B}$.

Remark 3.1.6. 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 3.1.8). This randomized tie-breaking is essential for ensuring that $\mathbf{B}$ has i.i.d. exp(1) entries; indeed, if we were to choose e (or f^*) with probability 1, then the corresponding $\mathbf{B}$ would not have i.i.d. exp(1) entries.

Claim 3.1.7. *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 3.1.5 establishes the positivity of the entries in the other columns. □

Lemma 3.1.8. *The following statements hold:*
(i) $\mathcal{T}_0^* = \mathcal{T}_0$ *and* $T_1^* = T_0^* = T_0$.
(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 2.1.6 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 2.1.1 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

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

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 3.1.9. $U_i - U_{i-1} = T_{i+1} - T_i, \; i = 1, 2, .., m-1.$

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 3.1.8. We shall prove (b) by showing that

$$T^*_i = U_{i-1} + v, \; i = 1, \ldots, m. \tag{3.1.1}$$

for some appropriately defined value v.

Remark 3.1.10. Since $T^*_1 = T^*_0$, equation (3.1.1) 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 3.1.8 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 2.1.6, 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 2.1.11. 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 3.1.11. The non-uniqueness amongst the weights of matchings introduced by forcing $T^*_1 = T^*_0$ does not affect the applicability of Lemma 2.1.11. Though we could resort to the generalized definition of S-matchings as defined by Claim 2.1.13; in this case, it is not necessary as with probability one, 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 2.1.1 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,

$$S^*_\ell = v + W_\ell. \tag{3.1.2}$$

But from Corollary 2.1.3 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,

$$T^*_i = U_{i-1} + v \text{ for } i = 1, ..., m. \tag{3.1.3}$$

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 Figure 3.1, where the entries of $\mathcal{T}_0$ and $\mathcal{T}_1$ are depicted by stars and

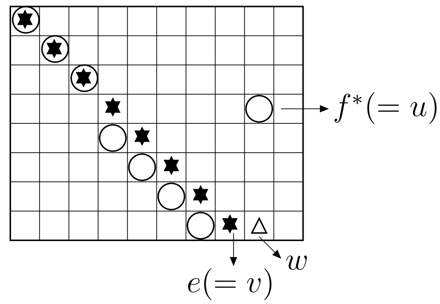

Figure 3.1: The entries e, f^*, w

circles, respectively. In the figure we assume that the row containing e was chosen to be removed by X (that is, $v = e$ and $u = f^*$).

As in Case 1, let $\mathcal{M}$ be the smallest matching of size $m-1$ in **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 **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 3.1.12. In the claim that follows, we will use a slightly unconventional method to prove a combinatorial fact implied by equation (3.1.1). 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 2.1.11 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 equation (3.1.1) for Case 2. A purely combinatorial argument also exists for this case which goes along the lines of Lemma 2.1.9. 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 3.1.13. *$P_i = T_i^*$ 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 2.1.11 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 $\mathbf{C}^\epsilon \setminus \mathbf{d}$. 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 $\mathbf{B} \setminus \mathbf{c}$. 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 $\mathbf{B} \setminus \mathbf{c}$ 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 $\mathbf{D}^\epsilon \setminus \mathbf{d}$. 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 $\mathbf{D}^\epsilon \setminus \mathbf{c}$. 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 $\mathbf{B} \setminus \mathbf{c}$ 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 2.1.11 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 (3.1.3) gives us the following:

$$Q_i = U_{i-1} + v - \epsilon, \text{ for } i = 1, 2, \ldots, m. \tag{3.1.4}$$

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 3.1.13 and equation (3.1.4) we have

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

This proves the lemma for Case 2 and hence completes the proof of Lemma 3.1.9. □

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 3.1.14. 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 3.1.15. 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 $Col(\mathcal{U}_0) = \{1, 2, ..., m-$

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 $\mathbf{M}$. Amongst 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 3.1.16. *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 $\mathcal{D}_\Lambda = \mathcal{F}_\Lambda \triangleq \bigcup_{\vec{r} \in \mathbb{R}_+^{m-1}} \mathcal{F}_\Lambda(\vec{r})$.

Proof. **($\boldsymbol{\alpha}$)** $\mathcal{D}_\Lambda \subset \mathcal{F}_\Lambda$: 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 3.1.14 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 3.1.9 that $v = u$ (the same as $e = f^*$ and this is Case 1 in the proof of Lemma 3.1.9). 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 3.1.16. 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 3.1.16 and hence $\mathbf{M} \in \mathcal{F}_\Lambda(\vec{r})$.

(β) $\mathcal{F}_\Lambda \subset \mathcal{D}_\Lambda$: Let $\mathbf{M} \in \mathcal{F}_\Lambda(\vec{r})$ for some $\vec{r}$. Then $\mathbf{M}$ satisfies condition (i) or (ii) of Lemma 3.1.16. Accordingly, this gives rise to two cases:

(a) $\mathbf{M}$ satisfies condition (i): We claim that $\Lambda(\mathbf{M}) = \mathbf{B}$. From Lemma 2.1.7 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, $\mathcal{D}^s_\Lambda$ and $\mathcal{D}^d_\Lambda$ correspond to the matrices in $\mathcal{D}_\Lambda$ which satisfy conditions (i) and (ii) of Lemma 3.1.16, respectively. Hence, when $\mathbf{M} \in \mathcal{D}^s_\Lambda$ we have $\Lambda(\mathbf{M}) = \mathbf{B}$ with probability one, and when $\mathbf{M} \in \mathcal{D}^d_\Lambda$ 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 3.1.16, 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

$$\mathcal{F}_1 = \bigcup_{\theta > 0, \mathbf{M} \in \mathcal{D}_\Lambda} F_1(\theta, \mathbf{M}). \tag{3.1.5}$$

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 3.1.16). 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

$$\mathcal{F}_2 = \bigcup_{\theta > 0, \mathbf{M} \in \mathcal{D}_\Lambda} F_2(\theta, \mathbf{M}). \tag{3.1.6}$$

Remark 3.1.17. 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 3.1.18. $\mathfrak{D}_m = \mathcal{F}_1 \cup \mathcal{F}_2$.

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 3.1.8 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 3.1.16 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 2.1.1, 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 3.1.19. 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 3.1.18.

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

$$\mathfrak{D}_m^s = F_1(\mathbb{R}_+, \mathcal{D}_\Lambda^s) \cup F_2(\mathbb{R}_+, \mathcal{D}_\Lambda^s) \quad \text{and} \quad \mathfrak{D}_m^d = F_1(\mathbb{R}_+, \mathcal{D}_\Lambda^d) \cup F_2(\mathbb{R}_+, \mathcal{D}_\Lambda^d). \quad (3.1.7)$$

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

$$f_v(\vec{v}) = \int_{\mathcal{R}_1} f_w(\vec{v}, \vec{x}) d\vec{x} + \frac{1}{2} \int_{\mathcal{R}_2} f_w(\vec{v}, \vec{x}) d\vec{x}. \tag{3.1.8}$$

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 3.1.16, 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 3.1.16. We set $H = J - U_0$, and $u_l = x_l - x_i$ for $l \neq i, k$. Finally, define

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

Thus, s_v denotes the sum of all of the entries of $\mathbf{A}$ except those in $\vec{x}$. As noted in the remark preceding Lemma 3.1.18, the value θ was added to precisely $m(n-m)$ entries. We have

$$\begin{aligned} &\int_{\mathcal{R}_1} f_w(\vec{v}, \vec{x}) d\vec{x} \\ &\quad \overset{(a)}{=} 2m \binom{n-m+1}{2} \int_0^H \iiint_0^\infty e^{-(s_v + (n-m+1)x_i + \sum_{l\neq i,k} u_l)} \prod_{l \neq i,k} du_l \, dx_i \\ &\quad = m(n-m) e^{-s_v} \left(1 - e^{-(n-m+1)H}\right). \end{aligned} \tag{3.1.9}$$

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 equations (3.1.5) and (3.1.6) respectively.

Similarly, on $\mathcal{R}_2$, we have that $x_i = J - U_0 \triangleq H$ and we shall set $u_l = x_l - x_i$ for

$l \neq i$ to obtain

$$\begin{aligned}
&\frac{1}{2}\int_{\mathcal{R}_2} f_w(\vec{v},\vec{x})d\vec{x} \\
&\overset{(b)}{=} \frac{1}{2}\left[2m(n-m)\iiint_0^\infty e^{-(s_v+(n-m+1)H+\sum_{l\neq i}u_l)}\prod_{l\neq i}du_l\right] \qquad (3.1.10)\\
&= m(n-m)e^{-s_v}e^{-(n-m+1)H}.
\end{aligned}$$

In equality (b) above, the factor $n-m$ accounts for the choice of i from $\{1,\dots,n-m+1\}$; 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 equations (3.1.5) and (3.1.6) respectively.

Substituting (3.1.9) and (3.1.10) into (3.1.8), we obtain

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

The above equation is summarized in the following lemma.

Lemma 3.1.20. *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 3.1.21. It is worth noting that part (ii) of Lemma 3.1.20 provides an alternate proof of Theorem 3.1.2.

From Lemma 3.1.9 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 3.1.22. $T_{k+1}(\mathbf{A}) - T_k(\mathbf{A})$ *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 1.5 required to prove Theorem 1.4.4. This completes the proof of Theorem 1.4.4 and hence establishes Parisi's conjecture.

3.2 The Coppersmith-Sorkin Conjecture

As mentioned in the introduction, Coppersmith and Sorkin [CS 99] 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:

$$F(k, m, n) = \sum_{i,j \geq 0, i+j<k} \frac{1}{(m-i)(n-j)}. \tag{3.2.1}$$

Nair [Na 02] has proposed a larger set of conjectures that identifies each term in equation (3.2.1) 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 3.1 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 3.2.1 (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 3.2.2 (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 3.2.3. *Let* $\mathbf{P}$ *be an* $m \times n$ *rectangular matrix,* $\mathbf{P}$*, of i.i.d.* exp(1) *entries. The* V-matchings *of* $\mathbf{P}$ *satisfy the following: for each* k*,* $1 \le k \le m-1$*:*

$$V_{i+1}^k - V_i^k \sim \exp(m-i)(n-k+i), \quad 0 \le i \le k-1 \tag{3.2.2}$$

and

$$V_0^{k+1} - V_k^k \sim \exp(m-k)n. \tag{3.2.3}$$

Further,

$$V_1^k - V_0^k \perp\!\!\!\perp V_2^k - V_1^k \perp\!\!\!\perp \cdots \perp\!\!\!\perp V_k^k - V_{k-1}^k \perp\!\!\!\perp V_0^{k+1} - V_k^k \tag{3.2.4}$$

Remark 3.2.4. We have grouped the increments according to the size of the matchings; so equations (3.2.2) and (3.2.3) both concern the k^{th} group. Equation (3.2.2) 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 equation (3.2.3) concerns the distribution of its difference with V_k^k.

Before we prove Theorem 3.2.3, we show how it implies the Coppersmith-Sorkin conjecture:

Corollary 3.2.5.

$$F(k,m,n) = \sum_{i,j \ge 0, i+j<k} \frac{1}{(m-i)(n-j)}. \tag{3.2.5}$$

Proof. By definition $F(j+1,m,n) - F(j,m,n) = \mathbb{E}(V_0^{j+1} - V_0^j)$. Using equations (3.2.2) and (3.2.3) and by linearity of expectation we obtain

$$F(j+1,m,n) - F(j,m,n) = \sum_{0 \le i \le j} \frac{1}{(m-i)(n-j+i)} \tag{3.2.6}$$

Now using the fact that $\mathbb{E}(V_0^1) = \frac{1}{mn}$ and summing (3.2.6) over $j=0$ to $j=k-1$ we obtain

$$F(k,m,n) = \frac{1}{mn} + \sum_{j=1}^{k-1} \sum_{0 \le i \le j} \frac{1}{(m-i)(n-j+i)} = \sum_{i,j \ge 0, i+j<k} \frac{1}{(m-i)(n-j)}. \tag{3.2.7}$$

Thus Theorem 3.2.3 establishes the Coppersmith-Sorkin conjecture. ∎

We now proceed to the proof of Theorem 3.2.3.

Remark 3.2.6. We will establish the theorem for the k^{th} group inductively. The outline of the induction is similar to the one in Section 1.5 and the details of the proof are similar to those in Section 3.1. 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.

Proof of Theorem 3.2.3

In this section we will establish properties concerning the increments of the V-matchings in the k^{th} group of the cost matrix $\mathbf{P}$, i.e. the increments between the matchings $\{\mathcal{V}_0^k, ..., \mathcal{V}_k^k, \mathcal{V}_0^{k+1}\}$. 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

$$
\begin{aligned}
L_1^\gamma - L_0^\gamma &= V_{m-l+1}^k - V_{m-l}^k \\
L_2^\gamma - L_1^\gamma &= V_{m-l+2}^k - V_{m-l+1}^k \\
\cdots \quad & \cdots \quad \cdots \\
L_\gamma^\gamma - L_{\gamma-1}^\gamma &= V_{m-l+\gamma}^k - V_{m-l+\gamma-1}^k \\
L_0^{\gamma+1} - L_\gamma^\gamma &= V_0^{k+1} - V_k^k.
\end{aligned}
\tag{3.2.8}
$$

Induction Step:

Step 1: From $\mathbf{L}$, form a matrix $\mathbf{Q}$ of size $l - 1 \times n$. Let $\{\mathcal{Q}_0^{\gamma-1}, ..., \mathcal{Q}_{\gamma-1}^{\gamma-1}\}$ 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 equation (3.2.8) 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

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

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

Claim 3.2.7. *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

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

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 2.1.1 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 $\mathbf{L}' \setminus \mathbf{c}$ 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 3.1, 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 3.1 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 turned 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 $\{\mathcal{U}_0, \ldots, \mathcal{U}_{l-1}\}$ and $\{\mathcal{Q}_0^{\gamma-1}, \ldots, \mathcal{Q}_{\gamma-1}^{\gamma-1}, \mathcal{Q}_0^{\gamma}\}$ denote the T-matchings of the matrix $\mathbf{Q}'$ and the V-matchings of the matrix $\mathbf{Q}$, respectively. The following statements follow from Claim 3.2.7 applied to the zero-padded matrix $\mathbf{Q}'$.

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

and

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

Now from Lemma 3.1.9 in Section 3.1 we know that

$$T_{i+1} - T_i = U_i - U_{i-1} \text{ for } i = 1, \ldots, l-1. \tag{3.2.9}$$

Remark 3.2.8. 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 3.1. However since we adopted the same procedure to go from one matrix to the other, the proof continues to hold.

Finally, combining Equation (3.2.9) and Claim 3.2.7 we obtain:

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

This completes Step 1 of the induction.

Step 2: Again we reduce the problem to the one in Section 3.1 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 3.1. We borrow the notation from Section 3.1 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 3.1. 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 $\mathbf{A} \setminus Col(\mathcal{T}_0)$ equaled $n-m$; now the number of columns in $\mathbf{L}' \setminus Col(\mathcal{T}_0)$ 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 $\gamma = k-m+l-1$ non-zero 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 3.1, let

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

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 3.1. (The counterparts of x_a and x_b in Section 3.1 were x_i and x_k, and these were defined according to Lemma 3.1.16.) 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

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

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 3.1, 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

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

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 3.1.16). 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 (3.2.10) and (3.2.10), we obtain

$$\begin{aligned} f_v(\vec{v}) &= l(n-k+m-l)\; e^{-s_v} \\ &= e^{-(q_{1,1}+\ldots+q_{l-1,n})} l(n-k+m-l) e^{-l(n-k+m-l)\theta} e^{-(r_1+\ldots+r_{l+k-m-1})}. \end{aligned}$$

We summarize the above in the following lemma.

Lemma 3.2.9. *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 3.2.10. $V_{i+1}^k - V_i^k \sim \exp(m-i)(n-k+i)$ *for* $i = 0, 2, .., k-1$.

Note that we have shown that $\mathbf{Q}$ are $L_1^\gamma - L_0^\gamma$ independent. This implies that, in particular, $V_{i+1}^k - V_i^k$ is independent if all the higher increments. Hence this completes the proof of the equation 3.2.4 and shows that the increments under consideration are independent.

To complete the proof of Theorem 3.2.3 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 $\mathbf{K}$ 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 $\mathbf{P}$ to the matrix $\mathbf{K}$, 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 (stated and proved originally in [Na 02]) completes the proof of Theorem 3.2.3.

Lemma 3.2.11. *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.

Case 1: $\mathcal{K}_0^1$ and $\mathcal{K}_1^1$ occur in different rows. Then clearly $\mathcal{K}_0^2$, the smallest matching of size two is formed by the entries $\mathcal{K}_0^1$ and $\mathcal{K}_1^1$ and hence in this case $K_0^2 - K_1^1 = K_0^1 \sim \exp(m-k+1)n$.

Case 2: $\mathcal{K}_0^1$ and $\mathcal{K}_1^1$ occur in different rows. In this case, it is not hard to see that $K_0^2 = K_0^1 + K_1^1 + X$ where X is an independent exponential of rate $(m-k)n$. Hence in this case $K_0^2 - K_1^1$ is the sum of two independent exponentials of rates $(m-k+1)n$ and $(m-k)n$ respectively.

Thus $K_0^2 - K_1^1$ is a random variable having the following distribution: with probability $\frac{n-1}{n}$ it is an exponential of rate $(m-k)n$ and with probability $\frac{1}{n}$ it is the sum of independent exponentials of rates $(m-k+1)n$ and $(m-k)n$ respectively. It is not hard to see that such a random variable is distributed as an exponential of rate $(m-k)n$. □

Remark 3.2.12. There is a row and column interchange in the definitions of the V-matchings in [Na 02].

Thus, we have fully established Theorem 3.2.3 and hence the Coppersmith-Sorkin Conjecture.

This also gives an alternate proof to Parisi's conjecture since [CS 99] shows that $E_n = F(n,n,n) = \sum_{i=1}^{n} \frac{1}{i^2}$.

3.3 A generalization of Theorem 1.4.4

Let $\mathbf{Q}$ be an $m \times n$ 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 $\mathbf{Q}$; 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 $\mathbf{Q}$ has its row-wise minima placement according to ξ. For concreteness, let us assume that $\xi = \{\xi_1, ..., \xi_n\}$ where ξ_i denotes the number of row-wise minima present in column i of $\mathbf{Q}$. We claim that the following generalization of Theorem 1.4.4 holds:

Theorem 3.3.1. *The joint distribution of the vector* $\{\mathcal{T}_j^\xi - \mathcal{T}_{j-1}^\xi\}_{j=1}^{n-1}$ *is the same for all placements of the row-wise minima,* $\xi \in \Upsilon$. *That is,*

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

Remark 3.3.2. On the event $\tilde{\xi}$ where all the row-wise minima lie in different columns, it is quite easy to show that $T_i^{\tilde{\xi}} - T_{i-1}^{\tilde{\xi}} \sim \exp i(n-i)$ for $i = 1, \ldots, n-1$ and that these increments are independent.

Consider an $m \times n$ matrix $\mathbf{A}^\xi$ whose entries are independent and exponentially distributed with rate one and conditioned on the event that the locations of the row-wise minima are in agreement with ξ. Let $\mathcal{T}_i^\xi$ denote the T-matchings of the matrix $\mathbf{A}^\xi$. Further, let $\mathbf{B}^\xi$ be the matrix that is obtained after the operations Φ and Λ. Let ξ_B denote the placement of the row-wise minima in matrix $\mathbf{B}^\xi$.

We note the following property of the mapping Φ.

Lemma 3.3.3. *Let* $\mathbf{A}$ *be an* $m \times n$ *matrix of positive entries and let* $\mu(i), i = 1, .., m$ *denote the location of the row-wise minima, i.e.* $a_{i\mu(i)} \leq a_{ij}$. *Let* $\mathbf{A}^* = \Phi(\mathbf{A})$. *Then* $\{\mu(i)\}$ *continues to remain as the locations of the row-wise minima for* $\mathbf{A}^*$, *i.e.* $a^*_{i\mu(i)} \leq a_{ij}$.

Proof. The proof is by contradiction. Note that the mapping Φ only modifies the entries outside $Col(\mathcal{T}_0)$. Assume there exists an entry $a_{i_0j_0}$ in $\mathbf{A}^*$ lying outside $Col(\mathcal{T}_0)(= Col(\mathcal{T}_0^*))$ such that $a_{i_0j_0} < a_{i_0j} \forall j \in Col(\mathcal{T}_0)$, then we can form a matching $\tilde{\mathcal{T}}_0$ by replacing the entry of $\mathcal{T}_0$ in row i_0 by the smaller entry $a_{i_0j_0}$. However this will imply that $\mathcal{T}_0^* < \mathcal{T}_0$, a contradiction to Lemma 3.1.8 as from Lemma 3.1.8 we know that $\mathcal{T}_0^* = \mathcal{T}_0$. □

Remark 3.3.4. The above lemma can also be seen from the property in Lemma 2.1.12. From Lemma 3.1.8 we know that $\mathcal{T}_0^* = \mathcal{T}_0$. From Lemma 2.1.12 we know that $Col(\mathcal{T}_0^*)$

contains all the columns of the row-wise minima. Since Φ does not change the elements in these columns, the row-wise minima continue to be the same elements.

By a similar induction argument carried out before, it is easy to see that to prove Theorem 3.3.1 it suffices to prove the following:

Theorem 3.3.5. *The following holds:*

- $T_1^\xi - T_0^\xi \sim \exp m(n-m)$
- *The entries of the matrix* $\mathbf{B}^\xi$ *are independent and exponentially distributed with rate one and conditioned on the event that the row-wise minima are in agreement with* ξ_B.
- *The entries of matrix* $\mathbf{B}^\xi$ *are independent of* $T_1^\xi - T_0^\xi$.

Proof. For every $i \in \{1, 2, ..., m\}$ let $\mu(i)$ denote the column containing the minimum entry in row i. Further assume that the locations of the row-wise minima are according to ξ. Then it is easy to see that the entries of $\mathbf{A}^\xi \equiv [a_{ij}]$ are distributed as follows.

- $a_{i\mu(i)} \sim exp(n)$
- $a_{ij} \sim a_{i\mu(i)} + \exp(1)$

where all the exponentials are independent of each other.

Hence we note the following: The density function of any particular realization of $\mathbf{A}$ is given by:

$$f(\mathbf{A}) = \begin{cases} n^m e^{-\sum_{i,j} a_{ij}}, & \text{on} \quad a_{i,j} \geq a_{i,\mu(i)} \quad \forall\ i, j \\ 0 & \text{else} \end{cases} \tag{3.3.1}$$

To complete the proof we need to perform the integral in equation (3.1.8) using the density function defined by the above equation (3.3.1). Assume that the matrix $\mathbf{B}$ is obtained by the removal of the row i_0 of matrix $\mathbf{A}^*$. Further let ξ_B be the vector corresponding to the arrangement of the row-wise minima of the matrix $\mathbf{B} \equiv [b_{ij}]$.

For this proof it is better to integrate out the $\vec{r}$ as well as $\vec{x}$ for reasons that will become apparent. Thus we write

$$f(\mathbf{B},\theta) = \int_{\mathcal{R}_1} f_w(\vec{v},\vec{x})d\vec{x}d\vec{r} + \frac{1}{2}\int_{\mathcal{R}_2} f_w(\vec{v},\vec{x})d\vec{x}d\vec{r}. \tag{3.3.2}$$

We consider two cases.

Case 1: $\mu(i_0)$ lies outside $Col(\mathcal{T}_0^{\mathbf{B}})$. In this case observe that x_i (as defined by Lemma 3.1.16) is the row-wise minimum and hence $r_j > x_i$ for all elements of $\vec{r}$. Let $\tilde{r}_j = r_j - x_i$. Note that x_i is the minimum entry in row i_o implies that $x_i + U > J$ since $\mathcal{U} = \mathcal{T}_0^{\mathbf{B}}$. Also observe that no point in region $\mathcal{R}_2$ possible in this case. This is because replacing x_i with the entry in row i_o chosen by the matching $\mathcal{T}_1$ (note that by definition of region $\mathcal{R}_2$ this entry must be in columns $Col(\mathcal{T}_0^B)$) leads to a matching of smaller weight and hence contradicts the minimality of $\mathcal{T}_1$ being the smallest matching that avoids at least one column of $\mathcal{T}_0$.

Let

$$s_{v_1} = b_{1,1} + b_{1,2} + \cdots + b_{m-1,n-m} + m(n-m)\theta$$

Thus, putting the density function defined by (3.3.1) in (3.3.2) we obtain

$$\begin{aligned} f(\mathbf{B},\theta) &= \int_{\mathcal{R}_1} f_w(\vec{v},\vec{x})d\vec{x}d\vec{r} \\ &= m(n-m)n^m \int_0^\infty\iiint_0^\infty e^{-(s_{v_1}+nx_i+\sum_{l\neq i,k} u_l)\sum_j \tilde{r}_j} \prod_{l\neq i,k} du_l \prod_j d\tilde{r}_j \; dx_i \\ &= m(n-m)n^{m-1}e^{-s_{v_1}}. \end{aligned} \tag{3.3.3}$$

Note that the factor 2 and $\binom{n-m+1}{2}$, that were present in the earlier integral in (3.1.9) are replaced by the single factor $n-m$ as we are given the location of the row minimum in row i_0. Hence we just have $n-m$ possible locations for x_k. The factor m appears (as before) from the fact that we have m choices for the row i_0, since the fact that the row-wise minima are in agreement with ξ does not change on row-permutations.

Hence

$$f(\mathbf{B},\theta) = m(n-m)n^{m-1}e^{-(b_{1,1}+b_{1,2}+\cdots+b_{m-1,n-m}+m(n-m)\theta)}. \qquad (3.3.4)$$

Note that this completes the proof of Theorem 3.3.5 under Case 1.

Case 2: $\mu(i_0)$ lies in $Col(\mathcal{T}_0^{\mathbf{B}})$. Let r_{i_0} denote the smallest entry. Let $\tilde{r}_j = r_j - r_{i_0}$ when $j \neq i_0$. In this case, in the region $\mathcal{R}_1$ we have $r_{i_0} < x_i < H$.

We first integrate out $\vec{x}$. Proceeding as in (3.1.9) and (3.1.10) we obtain

$$\begin{aligned}
&\int_{\mathcal{R}_1} f_w(\vec{v},\vec{x})d\vec{x} \\
&\quad = 2m\binom{n-m+1}{2}n^m \int_{r_{i_0}}^{H}\iiint_0^\infty e^{-(s_v+(n-m+1)x_i+\sum_{l\neq i,k}u_l)} \prod_{l\neq i,k} du_l\, dx_i \qquad (3.3.5)\\
&\quad = m(n-m)n^m e^{-s_v}\left(e^{-(n-m+1)r_{i_0}} - e^{-(n-m+1)H}\right).
\end{aligned}$$

and

$$\begin{aligned}
&\frac{1}{2}\int_{\mathcal{R}_2} f_w(\vec{v},\vec{x})d\vec{x} \\
&\quad = \frac{1}{2}\left[2m(n-m)n^m \iiint_0^\infty e^{-(s_v+(n-m+1)H+\sum_{l\neq i}u_l)} \prod_{l\neq i} du_l\right] \qquad (3.3.6)\\
&\quad = m(n-m)e^{-s_v}n^m e^{-(n-m+1)H}.
\end{aligned}$$

Combining (3.3.5) and (3.3.6) we obtain

$$f_v(\vec{v}) = m(n-m)n^m e^{-s_v}e^{-(n-m+1)r_{i_0}}$$

Note that

$$s_v = s_{v_1} + \sum_{j\neq i_0}\tilde{r}_j + (m-1)r_{i_0}$$

Thus we obtain

$$\begin{aligned} f(\mathbf{B},\theta) &= m(n-m)n^m \int_0^\infty \iiint_0^\infty e^{-s_{v_1}} e^{-\sum_{j\neq i_0} \tilde{r}_j} e^{-nr_{i_0}} \prod_{j\neq i_0} d\tilde{r}_j \, dr_{i_0} \\ &= m(n-m)n^{m-1} e^{-s_{v_1}} \\ &= m(n-m)n^{m-1} e^{-(b_{1,1}+b_{1,2}+\cdots+b_{m-1,n-m}+m(n-m)\theta)}. \end{aligned} \tag{3.3.7}$$

This completes the proof of Case 2 and hence of Theorem 3.3.5 (therefore of Theorem 3.3.1 as well).

□

3.4 The proof of a claim of Dotsenko

We now use Theorem 3.3.1 to complete an incomplete proof of the following claim of Dotsenko in [Do 00].

Let $\mathbf{A} \equiv [a_{ij}]$ be an $n \times n$ matrix with the following distribution for the entries: $a_{i,i} = 0, 1 \leq i \leq k$ and the rest of the entries are independent exponentials of mean 1. We assume $1 \leq k < n$. Let $\mathcal{T}^{\mathbf{A}}$ denote the smallest matching in matrix $\mathbf{A}$ of size n.

Let $\mathbf{B} \equiv b_{ij}$ be an $n \times n$ matrix with the following distribution for the entries: $b_{i,i} = 0, 1 \leq i \leq k-1$, $b(k,k-1) = 0$ and the rest of the entries are independent exponentials. Let $\mathcal{T}^{\mathbf{B}}$ denote the smallest matching in matrix $\mathbf{B}$ of size n.

Let $\mathbf{P}, \mathbf{Q}$ denote the matrices formed using the first k rows of $\mathbf{A}$ and $\mathbf{B}$ respectively and let $\{\mathcal{T}_i^{\mathbf{P}}, \mathcal{T}_i^{\mathbf{Q}}\}$ denote their T-matchings.

Claim 3.4.1. $\mathbb{E}(T^{\mathbf{B}}) - \mathbb{E}(T^{\mathbf{A}}) = \mathbb{E}(T_0^{\mathbf{Q}}) - \mathbb{E}(T_0^{\mathbf{P}})$

We make the following observation. Let $a_1, ..., a_k$ denote independent exponentials of rate n (independent of the entries in $\mathbf{A}$ and $\mathbf{B}$ as well). For $1 \leq i \leq k$, add a_i to the entries of the matrix $\mathbf{A}$ and $\mathbf{B}$ to form matrices $\mathbf{A}'$ and $\mathbf{B}'$ respectively. Now it is easy to see that the entries of $\mathbf{A}'$ and $\mathbf{B}'$ are distributed as independent exponentials of rate 1 and conditioned on the event that the minimum entry in row i, $1 \leq i \leq k$, occurs at the location of the zeroes in the matrices $\mathbf{A}$ and $\mathbf{B}$.

Further, if $\tilde{T}^{\mathbf{A}}$ and $\tilde{T}^{\mathbf{B}}$ denote the smallest matchings of the matrices $\mathbf{A}'$ and $\mathbf{B}'$, then one can see that $\tilde{T}^{\mathbf{A}} = T^{\mathbf{A}} + \sum_i a_i$ and $\tilde{T}^{\mathbf{B}} = T^{\mathbf{B}} + \sum_i a_i$. Let $\tilde{T}_i^{\mathbf{P}}, \tilde{T}_i^{\mathbf{Q}}$ denotes the T-matchings of size k of the $k \times n$ sub-matrices formed using the first k rows of the matrices $\mathbf{A}'$ and $\mathbf{B}'$ respectively. Observe that $\tilde{T}_0^{\mathbf{P}} = T_0^{\mathbf{P}} + \sum_i a_i$ and $\tilde{T}_0^{\mathbf{Q}} = T_0^{\mathbf{Q}} + \sum_i a_i$. This implies that to show Claim 3.4.1 it suffices to show that

$$\mathbb{E}(\tilde{T}^{\mathbf{B}}) - \mathbb{E}(\tilde{T}^{\mathbf{A}}) = \mathbb{E}(\tilde{T}_0^{\mathbf{Q}}) - \mathbb{E}(\tilde{T}_0^{\mathbf{P}}) \tag{3.4.1}$$

We prove equation (3.4.1) by proving a more general claim for which this is a special case. Let ξ_A^k and ξ_B^k denote the arrangement of the row-wise minimum in the first k, $1 \le k < m$, rows of an $m \times n$ matrix. Assume $m \le n$. Formally $\xi^k = \{\xi_1, ..., \xi_n\}$ and ξ_i denotes the number of the row-wise minima in the first k rows that occur in column i.

Remark 3.4.2. In the previous case observe that the matrices $\mathbf{A}'$ had a corresponding ξ vector $\{1, 1, .., 1, 0, 0, .., 0\}$ formed using k ones and $n-k$ zeroes where as $\mathbf{B}'$ had a ξ vector $\{1, 1, .., 1, 2, 0, 0, ..., 0\}$ formed using $k-2$ ones, 1 two and $n-k+1$ zeroes.

Let ξ_1^k and ξ_2^k be any two feasible ξ^k vectors corresponding to the arrangement of the row-wise minimum in the first k rows of an $m \times n$ matrix. Let $\mathbf{A}$ and $\mathbf{B}$ be two matrices formed using independent exponentials of rate 1 and conditioned on the event that their row-wise minima in the first k rows occur according to ξ_A^k and ξ_B^k respectively. As before, let $\mathbf{P}$ and $\mathbf{Q}$ denote the $k \times n$ sub-matrices formed using the first k rows of the matrices $\mathbf{A}$ and $\mathbf{B}$ respectively. Then we claim the following:

Theorem 3.4.3. $\mathbb{E}(T^{\mathbf{B}}) - \mathbb{E}(T^{\mathbf{A}}) = \mathbb{E}(T_0^{\mathbf{Q}}) - \mathbb{E}(T_0^{\mathbf{P}})$

Proof. Before going into the details of the proof, it is worth noting that from Remark 3.4.2 we see that Claim 3.4.1 is a special case of Theorem 3.4.3.

The proof is by an induction on the number, k, of the rows for whom the row-wise minima are according to a given arrangement. We start the induction at $k = m-1$ and then proceed down to $k = 1$.

k = m-1 : Here $\mathbf{P}$ is an $(m-1) \times n$ matrix of independent exponentials of rate one conditioned on the event that its row-wise minima are in agreement with ξ_1^{m-1}.

(Here, ξ_1^{m-1} is any arrangement of the row-wise minima in the first $m-1$ columns.) Let $\{a_1, ..., a_n\}$ denote the entries of the m^{th} row of $\mathbf{A}$. Note that $T^{\mathbf{A}} = \min_i(a_i + S_i^{\mathbf{P}})$.

Note that the matchings $\{\mathcal{T}_i^{\mathbf{P}}\}$ is just a re-ordering of the matchings $\{\mathcal{S}_i^{\mathbf{P}}\}$. Hence we can find a mapping, σ, of numbers $\{1, 2, ..., n\}$ to the numbers $\{0, 1, .., m-1\}$ such that $\mathcal{S}_i^{\mathbf{P}} = \mathcal{T}_{\sigma(i)}^{\mathbf{P}}$.

Note that, $T^{\mathbf{A}} = \min_i(a_i + T_{\sigma(i)}^{\mathbf{P}})$. Therefore,

$$T^{\mathbf{A}} - T_0^{\mathbf{P}} = \min_i(a_i + T_{\sigma(i)}^{\mathbf{P}} - T_0^{\mathbf{P}})$$

From Theorem 3.3.1 we know that the distribution of $\min_i(b_i + T_{\sigma(i)}^{\mathbf{P}} - T_0^{\mathbf{P}})$ does not depend on the arrangement of the row-wise minima in the matrix $\mathbf{P}$. Therefore the distributions of $\min_i(a_i + T_{\sigma(i)}^{\mathbf{P}} - T_0^{\mathbf{P}})$ and that of $\min_i(b_i + T_i^{\mathbf{Q}} - T_0^{\mathbf{Q}})$, defined similarly with the arrangement ξ_2^{n-1} instead of the arrangement ξ_1^{n-1}, are identical. This implies, in particular, that their expected values are same and therefore we obtain

$$\mathbb{E}(T^{\mathbf{A}} - T_0^{\mathbf{P}}) = \mathbb{E}(T^{\mathbf{B}} - T_0^{\mathbf{Q}}) \tag{3.4.2}$$

This completes the proof of the Theorem 3.4.3 for the case $k = m - 1$.

Now let us assume that Theorem 3.3.5 holds until say $k = j + 1$, i.e. it holds for $k = m - 1, m - 2, .., j + 1$. Let $\mathbf{P}$ and $\mathbf{Q}$ represent two $j \times n$ matrices of independent exponentials of rate one conditioned on the fact that their row-wise minima are arranged according to ξ_1^j and ξ_2^j respectively. Form matrices P_1 and Q_1 of size $(j + 1) \times n$ by adding a common row $\{r_1, .., r_n\}$ of independent exponential entries of rate one to the matrices $\mathbf{P}$ and $\mathbf{Q}$ respectively. We know from the first step of induction hypothesis that

$$\mathbb{E}(T^{\mathbf{P_1}} - T_0^{\mathbf{P}}) = \mathbb{E}(T^{\mathbf{Q_1}} - T_0^{\mathbf{Q}}) \tag{3.4.3}$$

Therefore it suffices to show that

$$\mathbb{E}(T^{\mathbf{A}} - T_0^{\mathbf{P_1}}) = \mathbb{E}(T^{\mathbf{B}} - T_0^{\mathbf{Q_1}}) \tag{3.4.4}$$

Let us further condition on the event, E_i, that the minimum entry in the row $j+1$ that was added occurs in column i. Let ξ_1^{j+1} and ξ_2^{j+1} denote the arrangement of the row-wise minima of the matrices P_1 and Q_1 conditioned on the event E_i.

From induction hypothesis we know that, for all $1 \leq i \leq n$,

$$\mathbb{E}(T^{\mathbf{A}} - T_0^{\mathbf{P_1}} | E_i) = \mathbb{E}(T^{\mathbf{B}} - T_0^{\mathbf{Q_1}} | E_i)$$

Averaging over the events E_i we obtain

$$\mathbb{E}(T^{\mathbf{A}} - T_0^{\mathbf{P_1}}) = \mathbb{E}(T^{\mathbf{B}} - T_0^{\mathbf{Q_1}}).$$

This completes the proof of the induction step and hence of Theorem 3.4.3. □

Thus, we see that that Claim 3.4.1 is indeed true. Using this claim one can obtain yet another completion of Parisi's conjecture following the line of arguments in [Do 00].

CHAPTER 4

Conclusion

This thesis provides a proof of the conjectures by Parisi [Pa 98] and Coppersmith-Sorkin [CS 99]. In the process of proving these conjectures, we have discovered some interesting combinatorial and probabilistic properties of matchings that could be of general interest. In this chapter we investigate a problem that our approach sheds new light on.

Of particular interest is the determination of the distribution of the C_n, the random variable that represents the cost of the minimum assignment. We use the methods we developed in the earlier chapters to obtain some conjectures regarding the complete distribution of C_n. In the last section of this chapter, we will touch upon some connections between the assumptions in the methods employed by the physicists and some of the results that we obtained rigorously in the earlier chapters.

4.1 Distribution of C_n

4.1.1 Early work

As mentioned in the introduction, using the non-rigorous Replica Method, a technique developed by Statistical Physicists to study interactions between particles, Mezard and Parisi, argued that the limit of $E(A_n)$ was $\frac{\pi^2}{6}$. They also computed the distribution of a randomly chosen entry that was part of the smallest matching. Further, they claimed that the assignment problem had the 'self averaging property', i.e. the distribution of A_n concentrates around the mean for large n.

In [Al 92], Aldous rigorously established that the limit of $E(A_n)$ exists. Later, in [Al 01], he established that the limit was $\frac{\pi^2}{6}$, as predicted by the physicists. He also recovered the distribution for a random entry in the smallest assignment. As further verification of the Physicists' approach, Talagrand showed that the variance decayed at a rate that was lower bounded by $\frac{1}{n}$ and upper bounded by $\frac{\log^4 n}{n}$.

For the case when the entries were independent exponentials of rate one, Parisi conjectured that for every finite n

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

From the proof of this conjecture as described in the previous chapter, we know that this expression can be broken down into sums of increments of matchings that are exponentially distributed. Though the smallest matching of size n is the sum of these increments, the increments (exponential random variables) are not independent and hence we are unable to compute easily the distribution of A_n.

In this chapter, we conjecture the exact nature of these correlations in the large n regime. If correct, these conjectures imply that

$$\sqrt{n}(A_n - E(A_n)) \stackrel{w}{\Rightarrow} N(0, 2)$$

This result is not surprising given the following two previous guesses. In [Al 01], Aldous commented that one would expect the limiting distribution to be Gaussian.

In [AS 02], Alm and Sorkin conjectured that the limiting variance of $\sqrt{n}(A_n - E(A_n))$ is 2. The basis of the conjecture regarding the variance, according to Alm and Sorkin, is based on a communication between Janson and them in which Janson guessed the exact distribution for every finite n. This guess turned out to be incorrect for $n \geq 3$ but seemed very close to the true distribution.

The conjecture in this chapter regarding the correlations in the large n regime, when applied to finite n will yield distributions that are largely similar to that of Janson's guess. However, the finer nature of our conjectures and the differences in some terms help us conclude that the limiting distribution is Gaussian rather easily.

4.2 Conjectures on correlations between increments

From Theorem 3.2.3 we have the following:

$$V_{i+1}^k - V_i^k \sim \exp(m-i)(n-k+i), \quad 0 \leq i \leq k-1 \tag{4.2.1}$$

and

$$V_0^{k+1} - V_k^k \sim \exp(m-k)n. \tag{4.2.2}$$

Further,

$$V_1^k - V_0^k \perp\!\!\!\perp V_2^k - V_1^k \perp\!\!\!\perp \cdots \perp\!\!\!\perp V_k^k - V_{k-1}^k \perp\!\!\!\perp V_0^{k+1} - V_k^k \tag{4.2.3}$$

Theorem 3.2.3 gives an explicit characterization of the distribution relating the difference between the smallest matching of size $k+1$ and the smallest matching of size k in terms of sums of independent exponentials.

Remark: Note that Theorem 3.2.3 does not give the entire distribution as it does not say anything regarding the dependence of the variables $V_0^{k+1} - V_0^k$ and $V_0^k - V_1^{k-1}$.

Consider a set of variables Δ_i^k defined by the following set of equations. All the random variables are assumed to be independent.

$$\begin{aligned}
\Delta_0^k &\sim \exp(n(n-k+1)), \text{ for } k = 1, ..., n \\
\Delta_i^k &\sim \begin{cases} 0 & \text{w.p. } \frac{n-i}{n-i+1} \\ \exp(n-i)(n-k+i+1) & \text{w.p. } \frac{1}{n-i+1} \end{cases} \quad {\scriptstyle 1 \le i \le k-1 \ , \ 2 \le k \le n}
\end{aligned}$$

Now define random variables R_i^k recursively according to the following relations:

$$\begin{aligned}
R_0^1 &= \Delta_0^1 \\
R_1^k - R_0^k &= \Delta_0^{k+1} && \text{for } k = 1, ..., n-1 \\
R_0^2 - R_1^1 &= R_0^1 + \Delta_1^2 \\
R_{i+1}^k - R_i^k &= R_i^{k-1} - R_{i-1}^{k-1} + \Delta_{i-1}^{k+1} && \text{for } i = 1, .., k-1 \\
R_0^{k+1} - R_k^k &= R_0^k - R_k^{k-1} + \Delta_{k+1}^{k+1} && \text{for } k = 2, .., n-1
\end{aligned}$$

It is easy to see that the R_i^k's satisfy the conditions of Theorem 3.2.3, i.e.

$$R_{i+1}^k - R_i^k \sim \exp(m-i)(n-k+i), \quad 0 \le i \le k-1 \tag{4.2.4}$$

and

$$R_0^{k+1} - R_k^k \sim \exp(m-k)n. \tag{4.2.5}$$

Further,

$$R_1^k - R_0^k \perp\!\!\!\perp R_2^k - R_1^k \perp\!\!\!\perp \cdots \perp\!\!\!\perp R_k^k - R_{k-1}^k \perp\!\!\!\perp R_0^{k+1} - R_k^k. \tag{4.2.6}$$

Observe that this equivalence of the marginals of the increments also implies $E(V_i^k) = E(R_i^k)$.

Remark: The initial guess by the author was that the distribution of R_i^k was in fact the distribution of V_i^k. However this was observed not to be true for $n \ge 3$. Calculations for $n = 3$ and $n = 4$ demonstrated that the distribution of R_i^k and V_i^k are very close to each other though not exactly equal. Simulations for higher n

confirm this observation. This makes us conjecture that the under the correct scaling (i.e. multiplication by $\sqrt{n}$) the error terms are of lower order and they die down as n becomes large.

Conjecture 4.2.1. *Let* $F_n(x) = \mathbb{P}[\sqrt{n}(V_0^n - E(V_0^n)) \leq x]$ *and let* $G_n(x) = \mathbb{P}[\sqrt{n}(R_0^n - E(R_0^n)) \leq x]$. *Then* $|F_n(x) - G_n(x)| \to 0, \forall x$ *as* $n \to \infty$.

Assuming that Conjecture 1 is correct, then this would imply that if

$$\begin{aligned} \sqrt{n}(R_0^n - E(R_0^n)) &\overset{w}{\Rightarrow} N(0,2), \text{ then} \\ \sqrt{n}(A_n - E(A_n)) &\overset{w}{\Rightarrow} N(0,2), \text{ since } V_0^n = A_n \end{aligned}$$

We prove the first claim in the lemma below.

Lemma 4.2.2. $\sqrt{n}(R_0^n - E(R_0^n)) \overset{w}{\Rightarrow} N(0,2)$.

Proof. Writing R_0^n in terms of the random variables Δ_i^k we obtain the following relation.

$$R_0^n = \sum_{k=1}^{n} \sum_{i=0}^{k-1} (n-k+1)\Delta_i^k$$

Let $\mu_i^k = \mathrm{E}(\Delta_i^k)$ and let $\mu_n = E(R_0^n)$. Then we note the following:

$$\lim_n n(E(R_0^n - \mu_n)^2) = \lim_n n \sum_{k=1}^{n} \sum_{i=0}^{k-1} (n-k+1)^2 E(\Delta_i^k - \mu_i^k)^2 = 2 \quad (4.2.7)$$

$$\lim_n n^2 \sum_{k=1}^{n} \sum_{i=0}^{k-1} (n-k+1)^4 E(\Delta_i^k - \mu_i^k)^4 = 0 \quad (4.2.8)$$

The proofs of these two equations were obtained using *MATHEMATICA* and hence has been omitted from the paper.

Now we apply the Central Limit Theorem for arrays to finish the argument. Let $X_{n,k,i} = \sqrt{n}(n-k+1)(\Delta_i^k - \mu_i^k)$. Observe that $\sum_{k,i} X_{n,k,i} = \sqrt{n}(R_1^n - E(R_1^n))$.

Eqns (4.2.7) and (4.2.8) imply the following conditions for the zero-mean independent random variables $X_{n,k,i}$.

- $\lim_n \sum_{k,i} E(X_{n,k,i}^2) = 2$.
- $\lim_n \sum_{k,i} E(X_{n,k,i}^4) = 0$.

Hence they satisfy the Lyapunov conditions for CLT and thus we have

$$\sum_{k.i} X_{n,k,i} \overset{w}{\Rightarrow} N(0,2) \text{ as } n \to \infty$$

This completes the proof of the lemma and hence assuming Conjecture 1 is true, this establishes the limiting distribution of A_n. ■

Remark: Though the Lyapunov CLT is normally stated with the third moment rather than the fourth moment used here, it is easy to see that any $2+\delta$ moment is sufficient.)

4.2.1 Combinatorial evidence

Now consider the increment $R_0^{k+1} - R_0^k$. The distribution for this increment can be explicitly stated in terms of sums of independent exponentials as stated in Theorem 1. However, from the definition of the random variables R_i^k we get the following relation:

$$R_0^{k+1} - R_0^k = R_0^k - R_0^{k-1} + \sum_{i=0}^{k} \Delta_i^{k+1}$$

Hence $R_0^{k+1} - R_0^k > R_0^k - R_0^{k-1}$. The following lemma shows that this is also true for the V_i^k's.

Lemma 4.2.3. $V_0^{k+1} - V_0^k > V_0^k - V_0^{k-1}$

Proof. Re-arranging the terms it is sufficient to show that $V_0^{k+1} + V_0^{k-1} > 2V_0^k$.

Case 1: If the matching $\mathcal{V}_0^{k+1}$ contains one element that lies outside the rows and columns occupied by $\mathcal{V}_0^{k-1}$, then we can combine this element with the matching $\mathcal{V}_0^{k-1}$ and get a matching of size k. Note that the rest of the elements of $\mathcal{V}_0^{k+1}$ is a matching of size k. Therefore we can identify two matchings of size k from among the elements of $\mathcal{V}_0^{k-1}$ and $\mathcal{V}_0^{k+1}$. Therefore the combined weight of these two matchings of size k must be greater than twice the weight of the smallest matching of size k.

Case 2: When there is no element of $\mathcal{V}_0^{k+1}$ that lies outside the rows and columns of $\mathcal{V}_0^{k-1}$ we establish the lemma by using the following well-known property of matchings. The rows and columns used by the smallest matching of size k contains all the rows and columns used by the smallest matching of size $k-1$. Call this the *intersection* property of matchings.

Consider a bipartite graph formed by the elements of $\mathcal{V}_0^{k+1}$ and $\mathcal{V}_0^{k}$. From the intersection property this is a $k+1 \times k+1$ bipartite graph. Color the $k+1$ edges represented by the elements of $\mathcal{V}_0^{k+1}$ by red and the $k-1$ edges represented by the elements of $\mathcal{V}_0^{k-1}$ by green. Now from the minimality of these matchings there cannot be any cycles. The intersection property also implies that the alternating paths must be of odd length and must have one extra red edge. (If it is of even length or has one extra green edge then we see that property one is violated). Therefore, we can decompose the bipartite graph into common edges and two alternating paths each having one extra red edge.

Now form one matching of size k by picking the common edges, red edges from first alternating path and green edges from second alternating path. Form the second matching of size k by picking common edges, green edges from first alternating path and red edges from second alternating path. Observe that the total weight of these two matchings of size k is equal to $V_0^{k+1} + V_0^{k-1}$. But this should be greater than twice the weight of the smallest matching of size k. This completes the proof of the lemma for Case 2. ■

4.3 Final remarks

The random assignment problem has been of interest in various communities since the early sixties. In this thesis we have presented proofs of the Parisi and Coppersmith-Sorkin conjectures that are related to the expected cost of the smallest assignment. In addition, we have presented some conjectures towards resolving the entire distribution. The arguments that have been presented are both probabilistic and combinatorial in nature; moreover several combinatorial properties regarding matchings that may be of general interest have been discovered in this process.

A non-rigorous approach, called the *Replica Method*, had been employed by the Statistical Physicists to study the behavior of the minimum assignment in the asymptotic regime where the number of jobs and machines tend to infinity. This method yielded a limiting expected cost of $\frac{\pi^2}{6}$ for the smallest matching in the large number of jobs limit. This non-rigorous result was later proved to be accurate using a new probabilistic method, *Objective Method*, developed by Aldous.

The replica method and a similar non-rigorous method called the *Cavity Method* have recently been extensively employed by physicists to study combinatorial optimization problems and infer their limiting optimal values or optimal solutions. The success of these heuristic methods sparked considerable interest among mathematicians, computer scientists and electrical engineers, as the problems solved by these heuristics have considerable importance in these fields.

One of the crucial non-rigorous assumptions of the Replica method is in the assumption that a certain level of *Replica Symmetry Breaking* is optimal. The optimality of this assumption is backed by a local analysis; however a proof of global optimality remains a missing link in the rigorization of this method.

In this thesis we have seen instances where checking local optimality is sufficient to infer global optimality, c.f. Remark 2.1.5. There have been other instances as well in the properties of matchings where a local optimality guarantees global optimality. It seems an interesting avenue of further research to make these connections precise. One could make substantial advances in rigorizing the assumptions of the non-rigorous methods if one is able to make general mathematical verifications under which a local optimality proof implies a global optimality.

Bibliography

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

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

[AS 02] S. E. Alm and G. B. Sorkin. Exact expectations and distributions for the random assignment problem. *Combinatorics, Probability, and Computing*, vol 11 no.3 (2002) 217-248.

[BCR 02] M. W. Buck, C. S. Chan and D. P. Robbins. On the Expected Value of the Minimum Assignment. *Random Structures and Algorithms* 21 (2002) 33-58.

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

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

[CS 02] D. Coppersmith and G. Sorkin. On the Expected Incremental Cost of a Minimum Assignment. In *Contemporary Combinatorics* (B. Bollobas, ed.), *Bolyai Society Mathematical Studies*, vol 10, chapter 8, Springer, 2002

[Do 00] V. S. Dotsenko. Exact solution of the random bipartite matching model. *J. Phys. A*, vol 33, (2000) 2015-2030.

[GK 93] M. X. Goemans and M. S. Kodialam. A lower bound on the expected cost of an optimal assignment. *Math. Oper. Res.*, 18 (1993) 267-274.

[Ja 99] S. Janson. One, Two and Three times log n/n for paths in a complete graph with random weights. In *Comb. Prob. Comp*, 8 (1999) 347-361.

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

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

[K 55] H.W. Kuhn. The Hungarian method for the assignment problem. *Naval Research Logistic Quarterly*, 2 (1955) 83-97.

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

[LW 00] S. Linusson and J. Wästlund. A Generalization of the Random Assignment Problem. http://www.mai.liu.se/˜ svlin/

[LW 04] S. Linusson and J. Wästlund. A Proof of Parisi's Conjecture on the Random Assignment Problem. *Probability Theory and Related Fields*, 128(3) (2004) 419-440. http://www.mai.liu.se/˜ svlin/

[W04] J. Wästlund. A proof of a conjecture of Buck, Chan, and Robbins on the expected value of the minimum assignment. *Random Structures and Algorithms*, 26(1-2) (2004) 237-251.

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

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

[Na 02] C. Nair. Towards the Resolution of Coppersmith-Sorkin Conjectures. *Proceedings of the 40th Annual Allerton Conference on Communication, Control and Computing*, 2002.

[NPS 03] C. Nair, B. Prabhakar and M. Sharma. Proofs of the Parisi and Coppersmith-Sorkin Conjectures for the Finite Random Assignment Problem, *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, (2003) 168-178.

[NPS 05] C. Nair, B. Prabhakar and M. Sharma. Proofs of the Parisi and Coppersmith-Sorkin Random Assignment Conjectures, *Random Structures and Algorithms*, to be published, (2005).

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

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

[SP 02] M. Sharma and B. Prabhakar. On Parisi's Conjecture for the Finite Random Assignment Problem. *Proceedings of the 40th Annual Allerton Conference on Communication, Control and Computing* 2002.

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

[Ta 03] M. Talagrand. An assignment problem at high temperature. *Annals of Probability*, 31(2) (2003) 818 -848.

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