

# At most $3.55^n$ stable matchings

Cory Palmer

Department of Mathematical Sciences  
University of Montana  
Missoula, MT USA  
cory.palmer@umontana.edu

Dömötör Pálvölgyi

MTA-ELTE Lendület Combinatorial Geometry Research Group  
Institute of Mathematics, Eötvös Loránd University  
Budapest, Hungary  
domotorp@gmail.com

**Abstract**—We improve the upper bound for the maximum possible number of stable matchings among  $n$  jobs and  $n$  applicants from  $131072^n + O(1)$  to  $3.55^n + O(1)$ . To establish this bound, we state a novel formulation of a certain entropy bound that is easy to apply and may be of independent interest in counting other combinatorial objects.

**Index Terms**—stable matchings; entropy

## I. INTRODUCTION

A *stable matching instance* is a complete bipartite graph  $G$  with an  $n$ -vertex class of *jobs*<sup>1</sup> and an  $n$ -vertex class of *applicants*<sup>1</sup> where each vertex has a strict order of *preferences* on the vertices in the opposite class. Given a perfect matching in  $G$ , a pair  $u, v$  is *unstable* if  $u$  and  $v$  both prefer each other to their currently matched pair. A perfect matching that contains no unstable pairs is called a *stable matching*. Gale and Shapley [12] showed that a stable matching exists in every stable matching instance, and gave an efficient algorithm to find one; see also [22, Chapter 1]. Since their introduction in 1962 by Gale and Shapley, stable matchings have found a wide variety of practical applications including college admissions [12], hospital residences [25], [27], [28], kidney exchanges [5], [6], and even pure math proofs [13]. Shapley and Roth were awarded the 2012 Nobel Memorial Prize in Economic Sciences for their work on stable matchings.

A long-standing problem, first posed by Knuth in 1976 [23] (see also Manlove [24, Section 2.2.2]) is to determine the maximum possible number  $SM(n)$  of stable matchings in a stable matching instance with  $n$  jobs and  $n$  applicants. This problem is interesting in its own right and also has implications for the run time of certain algorithms [21].

A very easy lower bound for  $SM(n)$  is  $2^{n/2}$ , while the best-known lower bound  $\Omega(2.28^n)$  is due to Irving and Leather [18]. This bound holds when  $n$  is a power of 2, but for all other  $n$ , Thurber [31] gave a slightly weaker bound of  $2.28^n / c^{\log n}$ .

The trivial upper bound on the number of stable matchings is  $n!$ . This was improved to  $O(n! / c^n)$  by Strathoudopolos [30] and  $\frac{3}{4}n!$  by Drgas-Burchardt and Świtalski [11]. Despite considerable attention [4], [11], [17], [30], [31], the first simply exponential upper bound was obtained only recently by Karlin, Oveis Gharan and Weber [19]. In that paper, the

authors remark that their argument gives a bound of  $c^n = 2^{17n}$  for  $n$  large enough, but that their method will not give a  $c$  close to the best-known lower bound. Their approach is a probabilistic argument, similar to one of the early proofs of the Crossing Lemma [1]. A recent survey of Cechlárová, Cseh and Manlove [9] highlights the importance of further improvements to the upper bound.

Our main result is to improve the upper bound for the maximum possible number of stable matchings; the base of our exponent gets close to 2.28, the base in the best lower bound.

**Theorem 1.** *The number of stable matchings  $SM(n)$  in an instance with  $n$  jobs and  $n$  applicants satisfies*

$$SM(n) \leq 3.55^n + O(1).$$

Apart from its theoretical importance, this also improves the known upper bound for the running time of certain algorithms. For example, the upper bound of the running time of the best-known algorithm for the popular roommates problem with strict preferences by Kavitha [21] is improved from roughly  $400000^n$  to  $10.65^n$ .

As mentioned in [19], Theorem 1 also gives an upper bound for the so-called *stable roommates problem*, defined similarly to the stable matching problem, but without requiring the graph  $G$  to be bipartite. This follows by a simple reduction to the stable matching problem where we take two copies of each vertex [10].

**Corollary 2.** *The number of stable matchings among  $n$  roommates is at most  $3.55^n + O(1)$ .*

Our starting point to prove Theorem 1 is similar to the starting point of [19]. We will observe some properties of the rotation poset (to be defined in Section III), and then use them to give an upper bound for  $SM(n)$  with a variant of the entropy method. Our Lemma 4 may be of independent interest, as it provides a simple-to-apply method for counting combinatorial objects compared to other techniques, like those that employ entropy or the container method [3].

**Notation.** We write the expected value of a function  $X$  of a random variable  $r$  as  $E_r[X(r)]$  when we want to emphasize that  $r$  is the random variable—this is to avoid confusion when  $X$  also depends on other (fixed) parameters. All logarithms

<sup>1</sup>In this paper we have decided to use the terms of *jobs* and *applicants* in place of the more traditional *men* and *women*; a similar approach was used in [10].

in this paper are natural and contain no artificial ingredients or added color.

**Outline of the paper.** In Section II we present our main result for counting objects, Lemma 4, and give an example of a simple application, a short proof of Bregman's theorem. In Section III we first outline why earlier results [15], [18], [19] imply that the number of stable matchings is bounded by the number of downsets of a so-called *tangled grid* poset. We then use our main lemma to prove that the number of downsets in a tangled grid is at most  $11.11^n$ . This gives a bound on  $SM(n)$  and is a combinatorial problem of independent interest. In Section IV we refine these methods to prove Theorem 1.

## II. MAIN LEMMA

In this section we present a method for bounding the cardinality of a family  $S$ . To this end, we will encode  $S$  as a family of  $n$ -tuples from  $A_1 \times A_2 \times \dots \times A_n$  for some carefully selected sets  $A_i$ . We use the term *component* both to refer to the sets  $A_i$  and to the elements  $s_i$  for  $s = (s_1, s_2, \dots, s_n) \in S$ . It will be convenient to think of each of the components  $s_i$  of  $s$  as carrying some information about  $s$  such that  $s$  is completely determined by the values  $s_1, s_2, \dots, s_n$ .

If  $S \subset A_1 \times \dots \times A_n$ , then a trivial upper bound on  $|S|$  is  $|A_1| \dots |A_n|$ . However, if the structure of  $S$  has sufficient symmetry, then we can get better bounds by revealing the components one-by-one as follows.

Fix a permutation  $\pi$  of  $\{1, \dots, n\}$  and an  $n$ -tuple  $s = (s_1, s_2, \dots, s_n) \in S$ . Imagine that the components of  $s$  are revealed according to  $\pi$  in the order  $s_{\pi(1)}, s_{\pi(2)}, \dots, s_{\pi(n)}$ . Let  $X_i(s, \pi)$  count the number of possible  $i$ th components of members of  $S$  whose first  $\pi^{-1}(i) - 1$  components under  $\pi$  (i.e., those components that are revealed before  $s_i$  is revealed) are identical to those of  $s$  under  $\pi$ . That is,

$$X_i(s, \pi) = |\{x_i : (x_1, \dots, x_n) \in S, x_j = s_j \text{ for } j \text{ satisfying } \pi^{-1}(j) < \pi^{-1}(i)\}|.$$

This immediately improves the earlier mentioned trivial bound  $|S| \leq |A_1| \dots |A_n|$  to

$$|S| \leq \prod_{i=1}^n \max_{s \in S} X_i(s, \pi) \quad (1)$$

for any permutation  $\pi$ . We give an example that illustrates how to compute  $X_i(s, \pi)$ .

**Example 3.** Put

$$S = \{(i, i, 0) : 1 \leq i \leq N\} \cup \{(i, 0, i) : 1 \leq i \leq N\} \cup \{(0, i, i) : 1 \leq i \leq N\}.$$

Thus,  $S \subset \{0, 1, \dots, N\}^3$  and  $|S| = 3N$ . Suppose that  $s = (s_1, s_2, s_3) = (i, i, 0)$  for some  $i \in \{1, \dots, N\}$  and consider all six possible permutations  $\pi$  of  $\{1, 2, 3\}$ . If  $\pi(1) = 1$ , i.e., coordinate 1 is revealed first, then  $X_1(s, \pi) = N + 1$ , as  $s_1$  can take any value from 0 to  $N$ . If  $\pi(3) = 1$ , i.e., coordinate 1 is revealed last, then  $X_1(s, \pi) = 1$ , as  $s_2 = i$  and  $s_3 = 0$

determine that  $s_1 = i$ . Otherwise,  $\pi(2) = 1$ , so  $\pi = (2, 1, 3)$  or  $\pi = (3, 1, 2)$ . If  $\pi = (2, 1, 3)$ , from  $s_2 = i$  we learn that  $s_1$  is either 0 or  $i$ , so there are only two options left for  $s_1$ . Finally, if  $\pi = (3, 1, 2)$ , from  $s_3 = 0$  we only learn that  $s_1 \neq 0$ , so there are still  $N$  options for  $s_1$ . We can repeat a similar analysis for  $s = (i, 0, i)$  and  $s = (0, i, i)$  to calculate all values of  $X_1(s, \pi)$ ; see Table I. Since  $S$  is symmetric under a permutation of the coordinates, the possible values of  $X_2(s, \pi)$  and  $X_3(s, \pi)$  would be given by a table that can be obtained from Table I by a permutation of rows and columns.

| $X_1(s, \pi)$ | 123     | 132     | 213 | 231 | 312 | 321 |
|---------------|---------|---------|-----|-----|-----|-----|
| $(i, i, 0)$   | $N + 1$ | $N + 1$ | 2   | 1   | $N$ | 1   |
| $(i, 0, i)$   | $N + 1$ | $N + 1$ | $N$ | 1   | 2   | 1   |
| $(0, i, i)$   | $N + 1$ | $N + 1$ | 2   | 1   | 2   | 1   |

TABLE I: Summary of values of  $X_1(s, \pi)$ . Rows represent the elements  $s \in S$ , while columns represent the permutations  $\pi$  of  $\{1, 2, 3\}$ .

Observe that the trivial upper bound gives  $|S| \leq (N + 1)^3$ , while the slightly better bound in (1) gives  $|S| \leq (N + 1) \cdot N \cdot 1 = O(N^2)$  for any permutation  $\pi$ .

Lemma 4 (to be stated below) further improves this bound on  $|S|$ . The first part of Lemma 4 gives  $\log|S| \leq \log(N + 1) + \frac{1}{3} \log N + \frac{2}{3} \log 2 + \log 1$  (for any choice of  $\pi$ ), which implies  $|S| \leq 2^{\frac{2}{3}}(N + 1)N^{\frac{1}{3}} = O(N^{\frac{4}{3}})$ . The second part of Lemma 4, when we average over only  $\pi$ , chosen uniformly at random, gives the weaker  $\log|S| \leq 3 \cdot \frac{1}{6}(2 \log(N + 1) + \log N + \log 2 + 2 \log 1)$ , which implies  $|S| \leq (2(N + 1)^2 N)^{\frac{1}{2}} = O(N^{\frac{3}{2}})$ . The loss compared to the previous bound is due to taking the maximum over all  $s \in S$  instead of the expectation, but we have stated this result in this form because in some applications (like in the proof of Theorem 1) we have no control over  $s$ . Finally, Corollary 5 (to be stated below), which is easier to apply than Lemma 4, gives the much weaker bound  $|S| \leq (\frac{3N+6}{6})^3$ .

Now we state our main lemma.

**Lemma 4.** *Let  $S \subset A_1 \times \dots \times A_n$  and  $\pi$  be any fixed permutation of  $\{1, \dots, n\}$ . Choose  $s \in S$  uniformly at random. Then*

$$\log|S| \leq \mathbb{E}_s \left[ \sum_{i=1}^n \log X_i(s, \pi) \right].$$

*By averaging, if  $\pi$  is chosen at random from the permutations of  $\{1, \dots, n\}$  according to an arbitrary probability distribution,*

$$\begin{aligned} \log|S| &\leq \mathbb{E}_{(s, \pi)} \left[ \sum_{i=1}^n \log X_i(s, \pi) \right] \\ &\leq \sum_{i=1}^n \max_{s \in S} \mathbb{E}_{\pi} [\log X_i(s, \pi)]. \end{aligned}$$

Applying Jensen's inequality gives the following weaker but simpler form of Lemma 4.

**Corollary 5.** Let  $S \subset A_1 \times \dots \times A_n$ . Then for any probability distribution of the permutations

$$|S| \leq \prod_{i=1}^n \max_{s \in S} \mathbb{E}_\pi [X_i(s, \pi)].$$

Although we could not find Lemma 4 explicitly stated anywhere, very similar bounds are frequently used in entropy proofs. Indeed, the first line of Lemma 4 follows from Shannon's famous noiseless coding theorem [29]. In that language Lemma 4 says that if the letters of the alphabet  $S$  are evenly distributed in a random text, then any encoding uses at least  $\log|S|$  bits per character on average. Lemma 4 can be also easily proved using the basic properties of entropy; we sketch such a proof here.

*Proof of Lemma 4.* Define the entropy of a random variable  $Z$  taking finitely many values as

$$H(Z) = - \sum_z \Pr[Z = z] \cdot \log \Pr[Z = z].$$

Fix an arbitrary permutation  $\pi$  of  $\{1, 2, \dots, n\}$  and select  $s \in S$  uniformly at random. From the basic properties of entropy (see [26]) we have

$$\begin{aligned} \log|S| &= H(s) \\ &= \sum_{i=1}^n H(s_i : s_j \text{ for } j \text{ satisfying } \pi^{-1}(j) < \pi^{-1}(i)) \\ &\leq \sum_{i=1}^n H(s_i : X_i(s, \pi)) \\ &= \sum_{i=1}^n \sum_k \Pr_s[X_i(s, \pi) = k] \cdot H(s_i : X_i(s, \pi) = k) \\ &\leq \sum_{i=1}^n \sum_k \Pr_s[X_i(s, \pi) = k] \cdot \log k \\ &= \sum_{i=1}^n \mathbb{E}_s[\log X_i(s, \pi)]. \end{aligned} \quad \square$$

Taking averages over all permutations is also a widely used tool in proofs. Arguably, the core idea of this goes back to Katona's circle method [20]; see also the permutation method described in [14, Chapter 1]. Therefore, most of the arguments presented in this paper could be repeated using entropy and averaging, but we feel that Lemma 4 can be a useful tool for proving upper bounds because it can be applied without introducing entropy.

Let us mention one form of Lemma 4, not explicitly stated in the book by Alon-Spencer [2, Chapter 2, The Probabilistic Lens]. They defined for a non-negative random variable  $X$  taking finitely many possible values, all with non-zero probability, its *geometric mean* as  $G[X] = \prod_x x^{\Pr[X=x]}$ . As noted there,  $G[X] = e^{\mathbb{E}[\log X]} \leq \mathbb{E}[X]$ . In this language, Lemma 4 states  $|S| \leq \prod_{i=1}^n G[X_i(s, \pi)]$  where  $s$  is a uniform random variable and  $\pi$  is an arbitrary permutation.

Lemma 4 shows that in order to get an upper bound on  $|S|$ , it is enough to find components  $A_i$  so that  $\mathbb{E}_\pi[\log X_i(s, \pi)]$

is small for every  $s \in S$ . As a simple application, we prove Bregman's theorem [8]. The argument follows Alon-Spencer [2] and Radhakrishnan [26], the only difference is that entropy is replaced by a simple application of Lemma 4.

**Theorem 6** (Bregman [8]). *The number of perfect matchings in a bipartite graph  $G = (U, V; E)$  is at most  $\prod_i (d_i!)^{1/d_i}$  where  $d_i$  denotes the degree of vertex  $v_i \in V$ .*

*Proof.* Let  $S$  be the set of all perfect matchings in  $G$  and let  $A_i$  be the set of edges incident to  $v_i \in V$ . Then  $S \subset \prod_i A_i$ . Fix a perfect matching  $s \in S$  and order the vertices  $v_i \in V$  according to a random permutation  $\pi$ . We first reveal the edge of  $s$  incident to  $v_{\pi(1)}$ , then the edge of  $s$  incident to  $v_{\pi(2)}$  and so on. Recall that  $X_i(s, \pi)$  is the number of options remaining for the edge of  $s$  incident to  $v_i$  right before it is revealed. Let  $N_i(s, \pi)$  be the number the neighbors of  $v_i$  (necessarily in  $U$ ) that have not yet had their edge from  $s$  revealed. Then we have  $X_i(s, \pi) \leq N_i(s, \pi)$ , where the inequality can be strict, as some of these edges may not be extendable to a perfect matching.

Since the edges of  $s$  were ordered randomly,  $\Pr[N_i(s, \pi) = k] = 1/d_i$  for every  $1 \leq k \leq d_i$ . Therefore, by Lemma 4 we have that the number of perfect matchings  $|S|$  satisfies

$$\begin{aligned} \log|S| &\leq \sum_i \max_{s \in S} \mathbb{E}_\pi [\log X_i(s, \pi)] \\ &\leq \sum_i \max_{s \in S} \mathbb{E}_\pi [\log N_i(s, \pi)] \\ &= \sum_i \sum_{k=1}^{d_i} \frac{1}{d_i} \log k = \sum_i \log(d_i!)^{1/d_i}. \quad \square \end{aligned}$$

### III. GENERAL METHOD AND A SIMPLE BOUND FOR $SM(n)$

In this section we give our first bounds on  $SM(n)$  by counting downsets in the so-called tangled grid poset.

#### A. From stable matchings to downsets in the tangled grid poset

In this subsection we detail a connection between stable matchings and downsets of a tangled grid. We begin by discussing how stable matchings can be mapped to the downsets in the rotation poset, a result established by Irving and Leather [18]. (It is also possible to show a correspondence in the other direction, see [7], [16].) Following [19], we define a rotation in a stable matching.

**Definition 7** (Rotation). Let  $k \geq 2$ . A *rotation*  $\rho$  is an ordered list of edges

$$\rho = (u_0 v_0, u_1 v_1, \dots, u_{k-1} v_{k-1})$$

in a stable matching  $\mu$  with the property that for  $i = 0, 1, 2, \dots, k-1$  applicant  $v_{i+1}$  (where the indices are modulo  $k$ ) is the highest ranked applicant for  $u_i$ 's list of preferences satisfying:

- (1) job  $u_i$  prefers  $v_i$  to  $v_{i+1}$ , and
- (2) applicant  $v_{i+1}$  prefers  $u_i$  to  $u_{i+1}$ .

In this case we say that  $\rho$  is *exposed* in the stable matching  $\mu$ .

We say that a rotation  $\rho = (u_0v_0, u_1v_1, \dots, u_{k-1}v_{k-1})$  is *eliminated* if we replace the edges of the stable matching  $\mu$  appearing in  $\rho$  with  $u_0v_1, u_1v_2, \dots, u_{k-1}v_0$ . It is easy to see that the perfect matching resulting from the elimination of a rotation  $\rho$  from a stable matching is itself stable (see Lemma 4.5 in [18]). It was shown in [15] that any stable matching can be obtained from what we will call a *job-optimal* stable matching  $\mu_0$  through a sequence of eliminations of exposed rotations. This leads to a natural poset on the family  $\mathcal{R}$  of rotations exposed in any stable matching of a stable matching instance. Indeed, for rotations  $\rho, \rho' \in \mathcal{R}$  we say that  $\rho \prec \rho'$  if and only if for every stable matching  $\mu$  in which  $\rho'$  is exposed we must eliminate  $\rho$  to arrive at  $\mu$  from  $\mu_0$ . We call the poset  $\mathcal{R}$  the *rotation poset*.

Recall that a subset of a poset is a *downset* (or *downward closed set*) if for every element of the downset, all smaller elements of the poset are in the downset.

**Theorem 8** (Irving & Leather [18], Theorem 4.1). *Given a stable matching instance, there is a bijection between the family of stable matchings and the downsets of the rotation poset.*

As in [19], this correspondence is central to our argument—instead of investigating stable matchings directly, we may estimate the number of downsets in the rotation poset. We will need the following immediate consequence of an observation explicitly stated in [19], and implicit in [15]. We say that a rotation  $\rho$  *involves* vertex  $v$  if  $v$  appears among the vertices in the ordered list of edges in  $\rho$ .

Recall that a *chain* in a poset is a sequence of comparable elements  $a_1 \prec a_2 \prec \dots \prec a_k$ . We will use the term *top* and *bottom* of a chain to refer to elements  $a_1$  and  $a_k$ , respectively.

**Proposition 9** (Karlin, Oveis Gharan & Weber [19], Claim 4.6). *The rotations involving a fixed vertex  $v$  form a chain in the rotation poset.*

So each vertex of a stable matching instance defines a chain in the rotation poset. For a job  $u$ , let  $M(u)$  be the chain of rotations involving  $u$ . The family of chains  $\{M(u) : u \text{ is a job}\}$  are the *m-chains* in the rotation poset. Similarly, for an applicant  $v$ , let  $W(v)$  be the chain of rotations involving  $v$  and the family of chains  $\{W(v) : v \text{ is an applicant}\}$  are the *w-chains* in the rotation poset.

**Proposition 10** (Irving & Leather [18], Lemma 4.7). *A edge  $uv$  where  $u$  is a job and  $v$  is an applicant appears in at most one rotation.*

Recall that a *chain decomposition* of a poset is a partition of its elements into pairwise disjoint chains.

**Definition 11** (Tangled grid). A poset  $P$  is an  $n \times n$  *tangled grid* if  $P$  has two  $n$ -member chain decompositions called *m-chains* and *w-chains*, such that each m-chain and each w-chain intersect each other in exactly one element.

In an  $n \times n$  tangled grid  $P$ , the m-chains are pairwise-disjoint, and the w-chains are also pairwise-disjoint. Moreover, since each m-chain and w-chain intersect in a single element, each m-chain and w-chain is of length exactly  $n$ . These together imply that  $P$  has exactly  $n^2$  elements.

**Proposition 12.** *Every rotation poset of  $n$  jobs and  $n$  applicants is a subposet of an  $n \times n$  tangled grid.*

*Proof.* We will modify the collection of m-chains and w-chains in a rotation poset  $\mathcal{R}$  to arrive at m-chains and w-chains satisfying the definition of an  $n \times n$  tangled grid.

For each rotation  $\rho = (u_0v_0, u_1v_1, \dots, u_{k-1}v_{k-1}) \in \mathcal{R}$ , remove  $\rho$  from each m-chain and w-chain except for the m-chain  $M(u_0)$  and w-chain  $W(v_0)$ . By Proposition 10,  $u_0v_0$  never appears in another rotation, so  $M(u_0)$  and  $W(v_0)$  intersect only in  $\rho$ . Now every element of  $\mathcal{R}$  is contained in exactly one m-chain and in exactly one w-chain. Note that some of the chains might have become temporarily empty. At this point, the m-chains and w-chains both form chain decompositions of  $\mathcal{R}$ .

Now it remains to show that the resulting poset is a subposet of a tangled grid. It is possible that we have an m-chain and a w-chain that do not intersect. In this case, add a new element to the poset that is included in these two chains, and is larger than every existing element of the poset. Continuing this procedure until every m-chain w-chain pair has an intersection results in a tangled grid.  $\square$

Let  $TG(n)$  be the maximum number of downsets in an  $n \times n$  tangled grid and let  $SM(n)$  be the maximum number of stable matchings in a stable matching instance with  $n$  jobs and  $n$  applicants. Proposition 12 and Theorem 8 imply that  $SM(n) \leq TG(n)$ , so we may turn our attention to counting downsets in a tangled grid. This will be our approach in the next subsection.

### B. Simple bound for $TG(n)$

Let  $P$  be an  $n \times n$  tangled grid. For the sake of simpler notation, let us construct a new poset  $P^*$  by adding two new elements  $a_0$  and  $b_0$  to  $P$  that are less than all other elements in  $P$ .

Add  $a_0$  to the bottom of each m-chain in  $P$  and let  $M_1, M_2, \dots, M_n$  be the resulting chains in  $P^*$ . Let  $W_{n+1}, W_{n+2}, \dots, W_{2n}$  be the chains resulting from adding  $b_0$  to the bottom of the w-chains in  $P$ . With a slight abuse of notation, we will continue to use the term m-chain and w-chain to describe these chains in  $P^*$ . Observe that every pair of chains among the m-chains and w-chains intersects in exactly one element.

Every downset in  $P$  corresponds to a downset in  $P^*$  that contains  $a_0$  and  $b_0$ . Let  $\mathcal{D}$  be the family of downsets in  $P^*$  that contain both  $a_0$  and  $b_0$ . For a downset  $D \in \mathcal{D}$ , let  $\text{top}_D(M_i)$  and  $\text{top}_D(W_j)$  be the maximum element of  $D$  on chain  $M_i$  and  $W_j$ , respectively. Define

$$s_D = (\text{top}_D(M_1), \dots, \text{top}_D(M_n), \text{top}_D(W_{n+1}), \dots, \text{top}_D(W_{2n})).$$

As  $D$  is a downset, the intersection of  $D$  and a chain  $C$  is characterized by the maximum element  $\text{top}_D(C)$ . Therefore,  $s_D$  uniquely determines  $D$ . (In fact, the first  $n$  terms or last  $n$  terms of  $s_D$  are enough to determine  $D$ , but our argument relies on the full encoding above). Thus, it is enough to find the cardinality of  $S = \{s_D : D \in \mathcal{D}\}$ . Clearly,  $S \subset M_1 \times \dots \times M_n \times W_{n+1} \times \dots \times W_{2n}$ .

Reveal the components  $M_1, \dots, M_n, W_{n+1}, \dots, W_{2n}$  of  $S$  in order according to a random permutation  $\pi$  of  $\{1, \dots, 2n\}$ . Recall that  $X_i(s, \pi)$  is the number of options remaining for the  $i$ th component of  $s \in S$  right before it is revealed. Suppose that  $i \leq n$ , i.e., we are about to reveal an  $m$ -chain (the case when  $i > n$  is similar). So we are counting the number of options for  $t = \text{top}_D(M_i)$  in the  $m$ -chain  $M = M_i$ .

Denote the number of  $w$ -chains preceding  $M$  according to  $\pi$  by

$$L(i, \pi) = |\{j \in [n+1, 2n] : \pi^{-1}(j) < \pi^{-1}(i)\}|.$$

Define  $X_i(s, \pi : L(i, \pi) = \ell)$  to be the value of  $X_i(s, \pi)$  under the assumption that there are exactly  $\ell$   $w$ -chains revealed before  $M$ .

Now suppose that  $L(i, \pi) = \ell$ . Each of the  $\ell$   $w$ -chains intersects  $M$  in a unique element. For any of these  $\ell$   $w$ -chains  $W$  we know the value of  $\text{top}_D(W)$ . The intersection  $a$  of  $M$  and  $W$  is in the downset  $D$  if and only if  $\text{top}_D(W) \succeq a$  on  $W$ . Moreover,  $a$  is in  $D$  if and only if  $t$  is at least  $a$ . Therefore, revealing  $W$  determines whether  $a$  is in the downset  $D$  or not. This means that after revealing  $\ell$   $w$ -chains, there are (at least)  $\ell$  distinct elements on  $M$  that are known to be in  $D$  or not.

Computing  $X_i(s, \pi : L(i, \pi) = \ell)$  directly is complicated, so we estimate it with a simpler random variable  $N_\ell$ .

**Definition 13.** Take  $n+1$  elements in a cyclic order, such that one of them is  $t$ . Pick  $\ell \leq n$  distinct elements uniformly at random, indexed in circular order as  $a_1, \dots, a_\ell$  (allowing  $a_j = t$  for some  $j$ ). Let  $N_\ell$  be the length of the circular interval  $[a_j, a_{j+1})$  (indexed cyclically) that contains  $t$ .

A random variable  $X$  is at most  $Y$  (in the stochastic order) if for all  $x$  we have  $\Pr[X \leq x] \leq \Pr[Y \leq x]$ .

**Claim 14.** For  $i \leq n$  and  $2 \leq \ell \leq n$ ,  $X_i(s, \pi : L(i, \pi) = \ell)$  is at most  $N_\ell$  (in the stochastic order).

*Proof.* Let  $a_1 \prec a_2 \prec \dots \prec a_\ell$  be the intersections of  $M$  with the  $\ell$  revealed  $w$ -chains. The bottom element of  $M$ ,  $a_0$ , satisfies  $a_0 \prec a_1$  as no  $w$ -chain contains  $a_0$ . If we imagine the elements of  $M$  arranged in the natural cyclic order around a cycle  $C$ , then  $a_0 \prec a_1 \prec \dots \prec a_\ell$  partition  $C$  into  $\ell+1$  intervals of the form  $[a_j, a_{j+1})$  (indexed modulo  $\ell+1$ ). As  $a_0$  is in  $D$ , there is a maximum  $j \geq 0$  such that  $a_j$  is in  $D$ . Thus,  $t$  must be on the interval  $[a_j, a_{j+1})$  of  $C$ . Let  $Y_\ell$  be the length of this interval. Observe that the number of options for  $t$ , i.e.,  $X_i(s, \pi : L(i, \pi) = \ell)$ , is at most  $Y_\ell$ . This allows us to ignore the structure of the poset  $P^*$  and estimate  $Y_\ell$ . For clarity, let us give an alternative definition of  $Y_\ell$ :

Take  $n+1$  elements in a cyclic order, such that one of them is  $a_0$ , and one of them is  $t$  (allowing  $a_0 = t$ ). Pick

$\ell$  distinct elements uniformly at random, indexed in circular order as  $a_1, \dots, a_\ell$  such that none of them is  $a_0$  (but allowing  $a_j = t$  for some  $j$ ). Let  $Y_\ell$  be the length of the circular interval  $[a_j, a_{j+1})$  (indexed modulo  $\ell+1$ ).

Note that  $Y_\ell$  depends on the relative position of  $a_0$  and  $t$ . Now it remains to show that  $Y_\ell$  is at most  $N_\ell$ . This is indeed the case because  $N_\ell$  is essentially the same as  $Y_\ell$ , except that  $a_0$  is not used when creating the intervals  $[a_j, a_{j+1})$ , so all these intervals are at least as long in case of  $N_\ell$ , as in case of  $Y_\ell$ . The only issue is that it is possible that for some  $j$  we have  $a_j = a_0$ , but in this case we get a distribution that is the same as  $Y_{\ell-1}$ , which is at least  $Y_\ell$  in the stochastic order.

Let  $Q$  be the event that  $a_0 \neq a_j$  for all  $1 \leq j \leq \ell$ . Then

$$\begin{aligned} \Pr[N_\ell \leq x] &= \Pr[N_\ell \leq x : Q] \Pr[Q] + \Pr[N_\ell \leq x : \bar{Q}] (1 - \Pr[Q]) \\ &\geq \Pr[Y_\ell \leq x] \Pr[Q] + \Pr[Y_{\ell-1} \leq x] (1 - \Pr[Q]) \\ &\geq \Pr[Y_\ell \leq x] \Pr[Q] + \Pr[Y_\ell \leq x] (1 - \Pr[Q]) \\ &= \Pr[Y_\ell \leq x]. \end{aligned} \quad \square$$

**Claim 15.** For  $2 \leq \ell \leq n$ ,

$$\mathbb{E}[N_\ell] \leq 2 \frac{n+1}{\ell+1}.$$

*Proof.* We will bound  $\mathbb{E}[N_\ell : t \neq a_j \forall 1 \leq j \leq \ell]$  and  $\mathbb{E}[N_\ell : \exists j : t = a_j]$  separately.

First, select  $\ell+1$  elements  $a'_0 \prec a'_1 \prec \dots \prec a'_\ell$  uniformly at random from all  $n+1$  elements of  $C$ . This partitions  $C$  into  $\ell+1$  intervals of the form  $[a'_j, a'_{j+1})$ . The expected length of each of these intervals is exactly  $\frac{n+1}{\ell+1}$ . Now, pick one of the  $a'_{j_0}$  uniformly at random, declare  $t = a'_{j_0}$  and reindex the rest of the  $a'_j$  into  $a_1, \dots, a_\ell$ . This generates the same distribution of  $t, a_1, \dots, a_\ell$  as in  $N_\ell$  given  $t \neq a_j \forall 1 \leq j \leq \ell$ , so the length of the interval  $[a_j, a_{j+1})$  containing  $t$  will be  $\mathbb{E}[N_\ell : t \neq a_j \forall 1 \leq j \leq \ell]$ . On the other hand, it is also the union of two intervals,  $[a'_{j_0-1}, a'_{j_0})$  and  $[a'_{j_0}, a'_{j_0+1})$ , thus its expected length is  $2 \frac{n+1}{\ell+1}$ . This proves  $\mathbb{E}[N_\ell : t \neq a_j \forall 1 \leq j \leq \ell] = 2 \frac{n+1}{\ell+1}$ .

The other case can be computed as follows.  $\mathbb{E}[N_\ell : \exists j_0 : t = a_{j_0}]$  is the same as the expected value of the smallest of  $\ell-1$  different numbers picked uniformly at random from  $[1, n]$ . These divide the  $n+1$  elements into  $\ell$  equal intervals, or we can compute the expectation as follows.

$$\begin{aligned} \mathbb{E}[N_\ell : \exists j_0 : t = a_{j_0}] &= \sum_{k=1}^{n-\ell+2} \Pr[\forall j \neq j_0 : a_j \geq k] \\ &= \sum_{k=1}^{n-\ell} \frac{\binom{n-k+1}{\ell-1}}{\binom{n}{\ell-1}} \leq \frac{\binom{n+1}{\ell}}{\binom{n}{\ell-1}} = \frac{n+1}{\ell}. \end{aligned}$$

If  $\ell \geq 1$ , then  $\frac{1}{\ell} \leq \frac{2}{\ell+1}$ , so both  $\mathbb{E}[N_\ell : t \neq a_j \forall 1 \leq j \leq \ell]$  and  $\mathbb{E}[N_\ell : \exists j : t = a_j]$  are less than  $2 \frac{n+1}{\ell+1}$ , thus so is  $\mathbb{E}[N_\ell]$ .  $\square$

We are ready to give an upper bound for  $E_\pi[X_i(s, \pi)]$  using  $\Pr[L(i, \pi) = \ell] = \frac{1}{n+1}$  for every  $0 \leq \ell \leq n$  by the uniformity of  $\pi$ . Observe that for  $i \leq n$

$$\begin{aligned} E_\pi[X_i(s, \pi)] &\leq \sum_{\ell=0}^n E_\pi[X_i(s, \pi) : L(i, \pi) = \ell] \cdot \Pr[L(i, \pi) = \ell] \\ &\leq \sum_{\ell=0}^1 \frac{n+1}{n+1} + \sum_{\ell=2}^n E[N_\ell] \cdot \frac{1}{n+1} \\ &\leq 2 + \sum_{\ell=2}^n \frac{2}{\ell+1} \leq 2 \log(n+1) + 3. \end{aligned} \quad (2)$$

An analogous argument for  $i > n$  considers a chain  $W = W_i$  and shows that (2) holds when  $i > n$  as well. Applying Corollary 5 with  $E[X_i(s, \pi)] \leq 2 \log(n+1) + 3$  gives

$$\begin{aligned} TG(n) &\leq \prod_{i=1}^{2n} \max_{s \in S} E_\pi[X_i(s, \pi)] \\ &\leq (2 \log(n+1) + 3)^{2n} \leq e^{2n \log \log n + O(n)}. \end{aligned}$$

This is worse than the bound given in [19], but better than the prior best bound of  $2^{n \log n - O(n)}$  from [30].

### C. Improved combinatorial bound for $TG(n)$

To get a better bound for  $|S|$ , we apply Lemma 4 using  $N_\ell$  as defined above.

**Claim 16.** For  $\ell \geq 2$  and any  $1 \leq k \leq n$

$$\Pr[N_\ell = k] = k \frac{\binom{n-k}{\ell-2}}{\binom{n+1}{\ell}}.$$

*Proof.* There are  $k$  potential intervals  $[a_j, a_{j+1})$  of length  $k$  on  $C$  that contain  $t$ . Since the two ends  $a_j$  and  $a_{j+1}$  need to be two of the  $\ell$  known elements from  $a_1, \dots, a_\ell$ , there are  $\binom{n+1-(k+1)}{\ell-2}$  options for the remaining  $\ell-2$  known elements from  $a_1, \dots, a_\ell$ .  $\square$

**Claim 17.** For  $n \geq 1$ ,  $i \leq n$  and any  $s \in S$ ,

$$E_\pi[\log X_i(s, \pi)] \leq 1.2038.$$

The proof of Claim 17 follows from Claims 14 and 16 with a calculation that is similar to (2) but is more involved.

*Proof.*

$$\begin{aligned} E_\pi[\log X_i(s, \pi)] &= \sum_{\ell=0}^n E_\pi[\log X_i(s, \pi) : L(i, \pi) = \ell] \Pr[L(i, \pi) = \ell] \\ &\leq \sum_{\ell=0}^1 \frac{\log(n+1)}{n+1} + \sum_{\ell=2}^n E[\log N_\ell] \frac{1}{n+1} \\ &\leq \frac{2 \log(n+1)}{n+1} + \frac{1}{n+1} \sum_{\ell=2}^n \sum_k k \frac{\binom{n-k}{\ell-2}}{\binom{n+1}{\ell}} \log k \\ &= \frac{2 \log(n+1)}{n+1} + \frac{1}{n+1} \sum_{k=2}^n k \log k \sum_{\ell=2}^n \frac{\binom{n-k}{\ell-2}}{\binom{n+1}{\ell}}. \end{aligned} \quad (3)$$

To bound the inner sum above we use the following identity.

**Lemma 18** (Whitworth Identity [32]). If  $m \geq 0$  and  $n \geq m+a$ , then

$$\sum_{j=0}^m \frac{\binom{m}{j}}{\binom{n}{j+a}} = \frac{n+1}{(a+1) \binom{n-m+1}{a+1}}.$$

As the above identity appeared as an exercise in a book published over 100 years ago [32], we provide a proof below.

Using the Whitworth Identity with  $a = 2$  we get

$$\begin{aligned} \sum_{k=2}^n k \log k \sum_{\ell=2}^n \frac{\binom{n-k}{\ell-2}}{\binom{n+1}{\ell}} &= \sum_{k=2}^n k \log k \sum_{j=0}^{n-k} \frac{\binom{n-k}{j}}{\binom{n+1}{j+2}} \\ &= \sum_{k=2}^n k \log k \left( \frac{n+2}{3 \binom{k+2}{3}} \right) = 2(n+2) \sum_{k=2}^n \frac{\log k}{(k+1)(k+2)}. \end{aligned}$$

Putting this back to Inequality (3), we obtain

$$\begin{aligned} E_\pi[\log X_i(s, \pi)] &\leq \frac{2}{n+1} \log(n+1) + 2 \frac{n+2}{n+1} \sum_{k=2}^n \frac{\log k}{(k+1)(k+2)}. \end{aligned}$$

From here it is easy to confirm numerically that  $E_\pi[\log X_i(s, \pi)] \leq 1.2038$  for every  $n \geq 1$ .  $\square$

*Proof of the Whitworth Identity.* First count the number of ways to choose  $m$  objects with repetition from the union of sets of sizes  $a+1$  and  $b+1$ , respectively, to establish

$$\binom{a+b+m+1}{m} = \sum_{j=0}^m \binom{a+j}{j} \binom{b+m-j}{m-j}. \quad (4)$$

Now let us count the number of strings of length  $a+b+m$  using exactly  $a$  zeros,  $m-j$  ones,  $j$  twos and  $b$  threes in two ways:

$$\begin{aligned} \binom{a+b+m}{a+m} \binom{a+m}{m} \binom{m}{j} &= \binom{a+b+m}{a+j} \binom{a+j}{j} \binom{b+m-j}{m-j}. \end{aligned} \quad (5)$$

Rearranging terms in (5) and substituting into (4) gives

$$\begin{aligned} \binom{a+b+m+1}{m} &= \sum_{j=0}^m \frac{\binom{a+b+m}{a+m} \binom{a+m}{m} \binom{m}{j}}{\binom{a+b+m}{a+j}} \\ &= \binom{a+b+m}{a+m} \binom{a+m}{m} \sum_{j=0}^m \frac{\binom{m}{j}}{\binom{a+b+m}{a+j}}. \end{aligned}$$

Now making the substitution  $n = a+b+m$  and rearranging terms gives

$$\sum_{j=0}^m \frac{\binom{m}{j}}{\binom{n}{j+a}} = \frac{\binom{n+1}{m}}{\binom{n}{a+m} \binom{a+m}{m}} = \frac{n+1}{(a+1) \binom{n-m+1}{a+1}}$$

where the last equality comes from simplification.  $\square$

As in Section III-B, an analogous argument shows that Claim 17 holds for  $n+1 \leq i \leq 2n$  as well. Therefore, applying Lemma 4 to  $S \subset M_1 \times \dots \times M_n \times W_{n+1} \times \dots \times W_{2n}$  gives

$$\begin{aligned} \log|S| &\leq \sum_{i=1}^{2n} \max_{s \in S} \mathbb{E}_\pi[\log X_i(s, \pi)] \leq 1.2038 \cdot 2n \\ &= 2.4076n \end{aligned}$$

which implies  $TG(n) \leq |S| \leq e^{2.4076n} \leq 11.11^n$ .

#### D. Asymptotic upper bound for $TG(n)$

In the previous sections we have given exact formulas, but in fact, these are unnecessary, as when  $n \rightarrow \infty$ , instead of dealing with  $\binom{n+1}{\ell}$  options, we can pick each element with probability  $x = \frac{\ell}{n}$ . More precisely, let us define the random variable  $N_x$  as follows.

**Definition 19.** Let  $A$  be a random subset of the integers such that  $\Pr[j \in A] = x$  for every  $j \in \mathbb{Z}$  independently. Let  $N_x$  be the length of the shortest interval containing 0 formed by the elements of  $A$ , i.e.,

$$N_x = \min_{b \in A} b - \max_{a \in A, a \leq 0} a.$$

Then we have  $\Pr[N_x = k] = kx^2(1-x)^{k-1}$  because there are  $k$  intervals of length  $k$  that contain 0, its two endpoints need to be in  $A$ , and its  $k-1$  intermediate points should not be in  $A$ . We can now perform an analogous calculations to that of (3) as follows.

$$\begin{aligned} &\mathbb{E}_\pi[\log X_i(s, \pi)] \\ &= \sum_{\ell=0}^n \mathbb{E}_\pi[\log X_i(s, \pi) : L(i, \pi) = \ell] \cdot \Pr[L(i, \pi) = \ell] \\ &\leq o(1) + \int_0^1 \mathbb{E}[\log N_x] dx \\ &= o(1) + \int_0^1 \sum_k (\log k) \Pr[N_x = k] dx \\ &= o(1) + \int_0^1 \sum_k (\log k) kx^2(1-x)^{k-1} dx \\ &= o(1) + \sum_k \frac{2 \log k}{(k+1)(k+2)} \leq 1.2038. \end{aligned} \quad (6)$$

Because of the  $o(1)$ , this gives a slightly weaker bound  $TG(n) \leq 11.11^n + O(1)$ .

#### IV. IMPROVED BOUND FOR $SM(n)$

In this section we bound the number of downsets in the rotation poset which, by Theorem 8, will give a bound on  $SM(n)$ . Our approach is similar to that in Section III, but now instead of using the tangled grid poset, will focus on the rotation poset directly. Moreover, as in Section III we will estimate a random variable  $X_i(s, \pi)$  by investigating the number of w-chains revealed before a fixed m-chain  $M = M_i$ . However, we will also use intersections of some m-chains revealed before  $M$ .

##### A. Properties of the rotation poset

First we derive some further properties of the rotation poset.

**Claim 20.** The number of  $m$ -chains containing rotation  $\rho \in \mathcal{R}$  is equal to the number of  $w$ -chains containing  $\rho$ , and is at least 2.

*Proof.* By Proposition 9, a rotation  $\rho = (u_0v_0, u_1v_1, \dots, u_{k-1}v_{k-1})$  is contained in  $m$ -chains  $M(u_i)$  and  $w$ -chains  $W(v_i)$  for  $0 \leq i \leq k-1$  where  $k \geq 2$ .  $\square$

For a rotation  $\rho = (u_0v_0, u_1v_1, \dots, u_{k-1}v_{k-1})$ , let us pair up those  $m$ -chains and  $w$ -chains that contain  $\rho$ , in two ways:

- (1)  $M(u_i)$  and  $W(v_i)$  are paired for every  $i$ , and
- (2)  $M(u_i)$  and  $W(v_{i+1})$  are paired for every  $i$  (where the indices are modulo  $k$ ).

**Claim 21.** Suppose that an  $m$ -chain  $M(u)$  and a  $w$ -chain  $W(v)$  are paired for some rotation  $\rho$ . Then either  $M(u)$  and  $W(v)$  are paired at exactly one other rotation  $\rho'$  which is consecutive to  $\rho$  on both  $M(u)$  and  $W(v)$ , or  $\rho$  is at the top or at the bottom of both  $M(u)$  and  $W(v)$ . Moreover, neither  $M(u)$  or  $W(v)$  is paired with another chain at both  $\rho$  and  $\rho'$ .

*Proof.* By Proposition 10, there is at most one rotation where  $uv$  appears as an edge. On the other hand, there is at most one rotation  $\rho = (u_0v_0, u_1v_1, \dots, u_{k-1}v_{k-1})$  such that  $u = u_i$  and  $v = v_{i+1}$ . Indeed, after  $\rho$  is eliminated,  $u$  and  $v$  will be matched, so any rotation  $\rho'$  immediately above  $\rho$  on  $M(u)$  or  $W(v)$  has  $uv$  appearing as an edge.

The ‘moreover’ part of the statement follows from the fact that each applicant gets matched to a higher ranked job in its list of preferences after the elimination of a rotation.  $\square$

**Remark 22.** In order to avoid the case in Claim 21 where  $M(u)$  and  $W(v)$  are paired at only one rotation we could extend the poset  $\mathcal{R}$  as follows. We add two new elements  $a$  and  $b$  that are greater than and less than all other elements in  $\mathcal{R}$ , respectively. Then add  $a$  and  $b$  to the top and bottom, respectively, of each  $m$ -chain and  $w$ -chain in  $\mathcal{R}$ . If  $M(u)$  and  $W(v)$  are paired only at the top of both chains, then we also pair them at  $a$  and if the chains are paired only at the bottom of both chains, then we also pair them at  $b$ . Because of this, from now on we will assume that the  $m$ -chains and  $w$ -chains are paired twice to avoid technicalities in later arguments. Note that this is somewhat similar to how  $a_0$  and  $b_0$  were added in Section III-B to the tangled grid. This time, however, we will ignore the slight increase in the length of the chains, as our goal will only be to prove an asymptotic upper bound.

Let us begin with definitions similar to those in Section III-B. Let  $M_1, M_2, \dots, M_n$  be the  $m$ -chains in  $\mathcal{R}$  and let  $W_{n+1}, W_{n+2}, \dots, W_{2n}$  be the  $w$ -chains in  $\mathcal{R}$ . Let  $\mathcal{D}$  be the family of downsets in  $\mathcal{R}$ . For a downset  $D \in \mathcal{D}$ , let  $\text{top}_D(M_i)$

and  $\text{top}_D(W_j)$  be the maximum element of  $D$  on chain  $M_i$  and  $W_j$ , respectively. Define

$$s_D = (\text{top}_D(M_1), \dots, \text{top}_D(M_n), \text{top}_D(W_{n+1}), \dots, \text{top}_D(W_{2n})).$$

As in Section III-B,  $s_D$  uniquely determines  $D$ , so we will bound the cardinality of  $S = \{s_D : D \in \mathcal{D}\}$ . Clearly,  $S \subset M_1 \times \dots \times M_n \times W_{n+1} \times \dots \times W_{2n}$ .

Reveal the components  $M_1, \dots, M_n, W_{n+1}, \dots, W_{2n}$  of  $S$  in order according to a random permutation  $\pi$  of  $\{1, \dots, 2n\}$ . Recall that  $X_i(s, \pi)$  is the number of options remaining for the  $i$ th component of  $s \in S$  right before it is revealed. Suppose that  $i \leq n$ , i.e., we are about to reveal an m-chain (the case when  $i > n$  is similar). So we are counting the number of options for  $\text{top}_D(M_i)$  in the m-chain  $M = M_i$ .

**Remark 23.** Unlike in the case of tangled grids, we cannot assume that all the m-chains and w-chains have the same fixed length. However, Claim 21 does imply that each m-chain and w-chain has length at most  $n-1$ . Our proof can be adapted in a straightforward way to work for chains of length less than  $n-1$ , but to avoid technicalities we assume that all chains have length exactly  $n-1$ .

Moreover, for our m-chain  $M$ , whose elements we denote by  $\rho_1 \prec \dots \prec \rho_{n-1}$ , we will assume that  $\text{top}_D(M) = \rho_t$  with  $2 \leq t \leq n-3$ . It can be shown that the cases when  $t = 1, n-2$  or  $n-1$  give a better bound on  $X_i(s, \pi)$ , but we will not discuss these cases in detail.

#### B. Final calculations

We will follow the argument in Section III-D, as dealing with complicated sums of binomial coefficients, as in Section III-C, is very cumbersome and would not improve the base of the exponent in our upper bound.

Recall that we have fixed m-chain  $M = M_i$  whose elements are  $\rho_1 \prec \dots \prec \rho_{n-1}$  such that  $\text{top}_D(M) = \rho_t$  where we assume that  $2 \leq t \leq n-3$ . By Claim 21, there is a chain  $W$  paired to  $M$  both at  $\rho_t$  and at  $\rho_{t+1}$  and  $W$  is not paired to  $M$  at any other rotation. Moreover, the elements  $\rho_t$  and  $\rho_{t+1}$  are also consecutive on  $W$ .

Denote the number of w-chains different from  $W$  that are revealed before  $M$  according to  $\pi$  by

$$L(i, \pi) = |\{j \in [n+1, 2n] : \pi^{-1}(j) < \pi^{-1}(i), W_j \neq W\}|.$$

By Claim 21, each of these  $L(i, \pi)$  w-chains are paired to  $M$  at exactly two consecutive elements. As before, define  $X_i(s, \pi : L(i, \pi) = \ell)$  to be the value of  $X_i(s, \pi)$  under the assumption that there are exactly  $\ell$  w-chains different from  $W$  that are revealed before  $M$ . We will give an upper bound for  $X_i(s, \pi : L(i, \pi) = \ell)$  by bounding it by a simpler random variable.

**Definition 24.** Let  $A$  be a random subset of the integers such that  $\Pr[j \in A] = x$  for every  $j \in \mathbb{Z}$  independently. Let  $B$  be a random subset of  $\{-1, 0, 1, 2\}$  such that  $\Pr[j \in B] = x$  for every  $j \in \{-1, 0, 1, 2\}$  independently. Let  $N_x$  be a random variable defined as follows. With probability  $1-x$ ,  $N_x$  is

the length of the shortest interval containing 0 formed by the elements of  $A$  and  $B$ , i.e.,

$$N_x = \min_{\substack{b \in A \cup B \\ b > 0}} b - \max_{\substack{a \in A \cup B \\ a \leq 0}} a$$

and  $N_x = 1$  with probability  $x$ .

We also define a family of related random variables as follows. Let  $B'$  be a random subset of  $\{-1, 0, 1, 2\}$  such that the events  $j \in B'$  and  $j+1 \in B'$  are independent, and for all other pairs  $j, j' \in \{-1, 0, 1, 2\}$  we have either that the events  $j \in B'$  and  $j' \in B'$  are independent, or that  $j \in B'$  if and only if  $j' \in B'$ . This defines  $2^3 = 8$  distinct cases, but in fact because of the condition that  $j \in B'$  and  $j+1 \in B'$  are independent, there are only 5 cases, one of which is the above defined  $N_x$ . In each of the other cases we can define a random variable  $N = N(x)$  in the same way that we define  $N_x$  above for a fixed probability  $x$ . Let  $\mathcal{N}_x$  be the family of these 5 random variables  $N$  and note that  $N_x \in \mathcal{N}_x$ .

A random variable  $X = X(n)$  is *asymptotically at most*  $Y = Y(n)$  (in the stochastic order) if for all  $x$  we have  $\Pr[X \leq x] \leq \Pr[Y \leq x] + o(1)$ , where  $o(1)$  tends to 0 as  $n \rightarrow \infty$ . Denote this by  $X \leq (1 + o(1))Y$ .

**Claim 25.** Put  $x = \ell/n$ . If  $0 < x < 1$ , then there is an  $N \in \mathcal{N}_x$  such that

$$X_i(s, \pi : L(i, \pi) = \ell) \leq (1 + o(1))N$$

where  $o(1) \rightarrow 0$  as  $n \rightarrow \infty$ .

*Proof.* The chains  $M$  and  $W$  share elements  $\rho_t$  and  $\rho_{t+1}$ . Therefore, if  $\text{top}_D(M) = \rho_t$ , then  $\text{top}_D(W) = \rho_t$ . Thus, if  $W$  is revealed before  $M$ , we have  $X_i(s, \pi : L(i, \pi) = \ell) = 1$ . The probability that  $W$  is revealed before  $M$  is  $(\ell+1)/(n+1)$ , since we can generate the permutation  $\pi$  by first taking a random permutation of  $M$  and the w-chains that are different from  $W$ , and then inserting  $W$  (and the rest of the m-chains).

Otherwise, thus with probability  $1 - (\ell+1)/(n+1)$ , the chain  $M$  is revealed before  $W$  and we proceed by constructing sets  $A$  and  $B$  as follows. If the w-chain paired to  $M$  at both  $\rho_{t+j}$  and  $\rho_{t+j+1}$  is revealed before  $M$ , then we put  $j$  into  $A$  if  $j \geq 1$ , and we put  $j+1$  into  $A$  if  $j \leq -1$ . For any integer  $-t < j < n-t-1$  the probability that the w-chain paired to  $M$  at  $\rho_{t+j}$  and  $\rho_{t+j+1}$  is revealed before  $M$  is  $(1 + o(1))\frac{\ell}{n}$ .

Claim 20 implies that every  $\rho \in M$  is contained in at least one other m-chain. Moreover, Claim 21 implies that for each  $-1 \leq j \leq 1$  there are two distinct m-chains,  $M' \neq M$  and  $M'' \neq M$ , such that  $\rho_{t+1} \in M'$  and  $\rho_{t+j+1} \in M''$ . This is because the w-chain  $W'$ , paired to  $M$  at  $\rho_{t+j}$  and at  $\rho_{t+j+1}$ , is paired to another m-chain  $M'$  at  $\rho_{t+j}$  and a different m-chain  $M''$  at  $\rho_{t+j+1}$  by the ‘moreover’ part of Claim 21. Therefore, we can define four m-chains  $M^{(-1)} \neq M^{(0)} \neq M^{(1)} \neq M^{(2)}$ , each different from  $M$ , that contain  $\rho_{t-1}, \rho_t, \rho_{t+1}, \rho_{t+2}$ , respectively. Note that these chains are not necessarily all distinct, e.g.,  $M^{(0)} = M^{(2)}$  is possible.

Knowing  $\text{top}_D(M^{(j)})$  for  $-1 \leq j \leq 2$  will determine whether  $\rho_{t+j} \in D$  or not, which can decrease the number of

options remaining for  $\text{top}_D(M)$ . For each  $-1 \leq j \leq 2$  such that  $M^{(j)}$  is revealed before  $M$ , put  $j$  into  $B$ . The probability of that  $M^{(j)}$  is revealed before  $M$  is  $(1 + o(1))\frac{\ell}{n}$ .

Moreover, if  $n \rightarrow \infty$  and  $0 < x < 1$ , then any bounded collection of the above described events is almost independent, meaning that if we replace them by independent events, then the probability that we get a different outcome is  $o(1)$ . This is because a constant number of these events do not carry significant additional information compared to  $L(i, \pi) = xn$ .

The intersections of  $M$  and the previously-revealed w-chains as well as the previously-revealed m-chains  $M^{(-1)}$ ,  $M^{(0)}$ ,  $M^{(1)}$ ,  $M^{(2)}$  splits  $M$  into intervals. By definition,  $X_i(s, \pi : L(i, \pi) = \ell)$  is at most the length of the interval containing  $\rho_t$ . As in the proof of Claim 14 it is straightforward to see that  $X_i(s, \pi : L(i, \pi) = \ell) \leq (1 + o(1))N$  for the  $N \in \mathcal{N}_x$  corresponding to which of the four m-chains  $M^{(-1)}$ ,  $M^{(0)}$ ,  $M^{(1)}$ ,  $M^{(2)}$  coincide.  $\square$

**Claim 26.**  $E[\log N] \leq E[\log N_x]$  for any  $N \in \mathcal{N}_x$ .

Before proving Claim 26 we need the following simple lemma.

**Lemma 27.** Let  $a_0, a_1, a_2 > 0$ . Suppose that for two independent random variables,  $X_1$  and  $X_2$ , we have  $\Pr[X_1 = 0] = \Pr[X_2 = 0] = x$  and  $\Pr[X_1 = a_1] = \Pr[X_2 = a_2] = (1 - x)$ . Suppose that the random variables  $X'_1$  and  $X'_2$  have the same distribution as  $X_1$  and  $X_2$ , respectively, but they completely determine each other, i.e.,  $X'_1 = 0$  if and only if  $X'_2 = 0$ . Then

$$E[\log(a_0 + X_1 + X_2)] \geq E[\log(a_0 + X'_1 + X'_2)].$$

*Proof.* This follows from the following application of Jensen's inequality to the log function.

$$\begin{aligned} & E[\log(a_0 + X_1 + X_2)] \\ &= x^2 \log a_0 + x(1-x)(\log(a_0 + a_1) + \log(a_0 + a_2)) \\ &\quad + (1-x)^2 \log(a_0 + a_1 + a_2) \\ &\geq x^2 \log a_0 + x(1-x)(\log a_0 + \log(a_0 + a_1 + a_2)) \\ &\quad + (1-x)^2 \log(a_0 + a_1 + a_2) \\ &= x \log a_0 + (1-x) \log(a_0 + a_1 + a_2) \\ &= E[\log(a_0 + X'_1 + X'_2)]. \end{aligned} \quad \square$$

*Proof of Claim 26.* Now we are ready to show  $E[\log N] \leq E[\log N_x]$  for  $N \in \mathcal{N}_x$ . Suppose that, say,  $0 \in B'$  if and only if  $2 \in B'$ , but  $-1 \in B'$  and  $1 \in B'$  are independent from all other events. (The other cases go similarly.) First, randomly determine for each  $j \in \mathbb{Z}$  whether  $j \in A$  or not, and also whether  $-1, 1 \in B'$  or not, and let these sets be  $A_0$  and  $B_0$ . Notice that Proposition 27 with  $a_0 = 2$ ,  $a_1 = \max_{a \in A_0 \cup B_0, a \leq 0} a$ ,  $a_2 = \min_{b \in A_0 \cup B_0, b > 0} b$  gives

$$\begin{aligned} & E[\log N : A = A_0, B' \setminus \{-1, 1\} = B_0] \\ &\leq E[\log N_x : A = A_0, B \setminus \{-1, 1\} = B_0]. \end{aligned}$$

By averaging over all possible  $A_0, B_0$  with the appropriate weights, we get  $E[\log N] \leq E[\log N_x]$ , finishing the proof of Claim 26.  $\square$

To repeat the calculation similar to (6), we first need to calculate  $\Pr[N_x = k]$  for  $k \geq 2$ . (We do not need  $\Pr[N_x = 1]$  because this will be multiplied by  $\log N_x = 0$ .) Let  $I$  be the elements (excluding the largest) of the shortest interval containing 0 formed by the elements of  $A$  and  $B$ , i.e.,  $I = [\max_{a \in A \cup B, a \leq 0} a, \min_{b \in A \cup B, b > 0} b)$ . By the definition of  $N_x$ , for  $k \geq 2$ ,

$$\begin{aligned} \Pr[N_x = k] &= (1-x) \sum_{j \leq 0} \Pr[I = [j, j+k]] \\ &= (1-x) \sum_{j \leq 0} \Pr[j \in A \cup B] \cdot \Pr[j+k \in A \cup B] \\ &\quad \cdot \prod_{h=1}^{k-1} \Pr[j+h \notin A \cup B]. \end{aligned}$$

We compute  $\Pr[N_x = k]$  for  $k = 2$ ,  $k = 3$  and  $k \geq 4$  separately. We show all steps when  $k = 2$ . The remaining calculations are also straightforward, but are kept intentionally brief here.

$$\begin{aligned} \Pr[N_x = 2] &= (1-x)(\Pr[I = [-1, 1]] + \Pr[I = [0, 2]]) \\ &= (1-x)(\Pr[-1 \in A \cup B] \cdot \Pr[1 \in A \cup B] \cdot \Pr[0 \notin A, B] \\ &\quad + \Pr[0 \in A \cup B] \cdot \Pr[2 \in A \cup B] \cdot \Pr[1 \notin A, B]) \\ &= (1-x)(2(1-(1-x)^2)(1-(1-x)^2)(1-x)^2) \\ &= 2(1-x)^3(1-(1-x)^2)^2. \end{aligned}$$

$$\begin{aligned} \Pr[N_x = 3] &= (1-x)(\Pr[I = [-1, 2]] + \Pr[I = [-2, 1]] \\ &\quad + \Pr[I = [0, 3]]) \\ &= (1-x)^5(1-(1-x)^2)^2 + 2(1-x)^5(1-(1-x)^2)x. \end{aligned}$$

$$\begin{aligned} \Pr[N_x = k \geq 4] &= (1-x)(\Pr[I = [0, k]] + \Pr[I = [-k+1, 1]] \\ &\quad + \Pr[I = [-1, k-1]] + \Pr[I = [-k+2, 2]]) \\ &\quad + \sum_{-k+3 \leq h \leq -2} \Pr[I = [h, h+k]] \\ &= 2(1-x)^{k+2}(1-(1-x)^2)x \\ &\quad + 2(1-x)^{k+3}(1-(1-x)^2)x + (k-4)(1-x)^{k+4}x^2. \end{aligned}$$

Now we are ready to prove the upper bound for  $E_\pi[\log X_i(s, \pi)]$ .

**Claim 28.** For any  $i \leq n$  and  $s \in S$ , as  $n \rightarrow \infty$

$$E_\pi[\log X_i(s, \pi)] \leq 0.6331 + o(1).$$

*Proof.*

$$\begin{aligned}
& \mathbb{E}_\pi[\log X_i(s, \pi)] \\
&= \sum_{\ell=0}^n \mathbb{E}_\pi[\log X_i(s, \pi) : L(i, \pi) = \ell] \cdot \Pr[L(i, \pi) = \ell] \\
&\leq o(1) + \int_0^1 \mathbb{E}[\log N_x] dx \\
&= o(1) + \int_0^1 \sum_k (\log k) \Pr[N_x = k] dx \\
&= o(1) + \int_0^1 (\log 2) 2(1-x)^3(1-(1-x)^2)^2 \\
&\quad + (\log 3)((1-x)^5(1-(1-x)^2)^2 \\
&\quad + 2(1-x)^5(1-(1-x)^2)x) \\
&\quad + \sum_{k \geq 4} (\log k)(2(1-x)^{k+2}(1-(1-x)^2)x \\
&\quad + 2(1-x)^{k+3}(1-(1-x)^2)x \\
&\quad + (k-4)(1-x)^{k+4}x^2) dx \\
&= o(1) + (\log 2) \frac{1}{12} + (\log 3) \frac{23}{630} \\
&\quad + \sum_{k \geq 4} (\log k) \frac{2k(k+7)+72}{(k+3)(k+5)(k+6)(k+7)} \\
&\leq 0.6331 + o(1). \quad \square
\end{aligned}$$

As in the previous sections, an analogous argument holds for  $n+1 \leq i \leq 2n$  as well. Therefore, applying Lemma 4 to  $S \subset M_1 \times \cdots \times M_n \times W_{n+1} \times \cdots \times W_{2n}$  gives

$$\begin{aligned}
\log |S| &\leq \sum_{i=1}^{2n} \max_{s \in S} \mathbb{E}_\pi[\log X_i(s, \pi)] \leq (0.6331 + o(1)) \cdot 2n \\
&= 1.2662n + o(n).
\end{aligned}$$

which implies  $SM(n) \leq |S| \leq e^{1.2663n} + O(1) \leq 3.55^n + O(1)$ .

## V. CONCLUDING REMARKS

Note that in an (untangled)  $n \times n$  grid with its elements ordered as a ‘diamond’ (with a unique smallest and largest element), the number of downsets is exactly  $\binom{2n}{n}$ . This makes one think that possibly  $TG(n) \leq 4^n \cdot \text{poly}(n)$  holds, as was conjectured by the second author. This, however, turned out to be false; Clay Thomas (personal communication) gave an example showing  $TG(n) = \Omega(4.17^n)$  by modifying the construction from [18] showing  $SM(n) = \Omega(2.28^n)$ .

One approach to improve our bounds would be to somehow avoid dealing with the m-chains and w-chains separately and reduce the number of components from  $2n$  to  $n$ . This would imply that the base of the exponent could be reduced to its square root. Unfortunately, this is too good to be true for  $TG(n)$ , as the lower bound is larger than the square root of our upper bound,  $11.11^n$ . Because of this, it appears unlikely that there is a simple way to estimate  $SM(n)$  by investigating only one type of chain.

We do not believe our upper bounds to be tight; the truth is probably closer to the known lower bounds for both  $TG(n)$  and  $SM(n)$ . It may be possible to improve our bound on  $SM(n)$  by considering how the other chains intersect a given chain  $M$ . Unfortunately, this appears to involve a very tedious analysis that is unlikely to match the lower bound. Perhaps some step-by-step modification of the rotation poset could give some insight into its structure in the extremal case and ultimately give a sharp bound.

## ACKNOWLEDGMENTS

We would like to thank Padmini Mukkamala for pointing out Claim 15, Viktor Harangi and Gábor Kun for discussions about other proofs of Lemma 4, Ágnes Cseh for calling our attention to relevant literature about stable matchings, and Balázs Patkós for comments on the first version of this paper. This work has started while the first author visited the second at Eötvös Loránd University, supported by grant number LP2017-19/2017 of the Lendület program of the Hungarian Academy of Sciences (MTA).

## REFERENCES

- [1] M. Ajtai, V. Chvátal, M. M. Newborn, E. Szemerédi, Crossing-free subgraphs, in *Theory and Practice of Combinatorics*, North-Holland Mathematics Studies 60: 9–12, 1982.
- [2] N. Alon, J. H. Spencer, *The Probabilistic Method*, Fourth ed., Wiley, 2016.
- [3] J. Balogh, R. Morris, W. Samotij, The method of hypergraph containers, in *Proceedings of the International Congress of Mathematicians (ICM 2018)*, 3059–3092, 2019.
- [4] A. T. Benjamin, C. Converse and H. A. Krieger, How do I marry thee? Let me count the ways, *Discrete Applied Mathematics* 59: 285–292, 1995.
- [5] P. Biró, B. Haase et al., Building kidney exchange programmes in Europe—an overview of exchange practice and activities, *Transplantation* 103(7): 1514–1522, 2019.
- [6] P. Biró, J. van de Klundert, D. Manlove, W. Pettersson et al., Modelling and optimisation in European Kidney Exchange Programmes, *European Journal of Operational Research* 291(2): 447–456, 2021.
- [7] D. Gusfield, R. W. Irving, P. Leather, M. Saks, Every finite distributive lattice is a set of stable matchings for a small stable marriage, *J. Combin. Theory Ser. A* 44: 304–309, 1987.
- [8] L. Bregman, Some properties of nonnegative matrices and their permanents, *Soviet Math. Dokl.*, 14: 945–949, 1973.
- [9] K. Cechlárová, A. Cseh, D. Manlove, Selected open problems in matching under preferences, *Bulletin of the European Association for Theoretical Computer Science* 2.128, 2019.
- [10] B. C. Dean and S. Munshi, Faster algorithms for stable allocation problems, *Algorithmica* 58(1): 59–81, 2010.
- [11] E. Drgas-Burchardt and Z. Świąłski, A number of stable matchings in models of the Gale–Shapley type, *Discrete Applied Mathematics* 161(18): 2932–2936, 2013.
- [12] D. Gale, L. S. Shapley, College admissions and the stability of marriage, *American Mathematical Monthly* 69(1): 9–15, 1962.
- [13] F. Galvin, The list chromatic index of a bipartite multigraph, *J. Combin. Theory Ser. B* 63(1): 153–158, 1995.
- [14] D. Gerbner, B. Patkós, *Extremal Finite Set Theory*, CRC Press, 2018.
- [15] D. Gusfield and R. W. Irving, *The stable marriage problem: structure and algorithms*, MIT press, 1989.
- [16] D. Gusfield, R. W. Irving, P. Leather, M. Saks, Every finite distributive lattice is a set of stable matchings for a small stable marriage, *J. Combin. Theory Ser. A* 44: 304–309, 1987.
- [17] J. S. Hwang, The algebra of stable marriages, *International Journal of Computer Mathematics* 20: 227–243, 1986.
- [18] R. W. Irving and P. Leather, The complexity of counting stable marriages, *SIAM Journal on Computing* 15(3): 655–667, 1986.

- [19] A. R. Karlin, S. Oveis Gharan, R. Weber, A simply exponential upper bound on the maximum number of stable matchings, in Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing (STOC 2018), 920–925.
- [20] G. O. H. Katona, A simple proof of the Erdős–Chao Ko–Rado theorem, *J. Combin. Theory Ser. B* 13(2): 183–184, 1972.
- [21] T. Kavitha, Popular roommates in simply exponential time, in Proceedings of the 39th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2019), 20:1–20:15.
- [22] J. Kleinberg, and É. Tardos, *Algorithm Design*, Addison–Wesley, 2005.
- [23] D. E. Knuth, *Mariages stables et leurs relations avec d’autres problèmes combinatoires: introduction à l’analyse mathématique des algorithmes*, Presses de l’Université de Montréal, 1976.
- [24] D. Manlove, *Algorithmics of Matching Under Preferences*, World Scientific, 2013.
- [25] D. G. McVitie, L. B. Wilson, Stable marriage assignment for unequal sets, *BIT Numerical Mathematics* 10(3): 295–309, 1970.
- [26] J. Radhakrishnan, An entropy proof of Bregman’s theorem, *J. Combin. Theory Ser. A* 77(1): 161–164, 1997.
- [27] A. E. Roth, The evolution of the labor market for medical interns and residents: a case study in game theory, *Journal of Political Economy* 92(6): 991–1016, 1984.
- [28] A. E. Roth, On the Allocation of Residents to Rural Hospitals, *Econometrica* 54(2): 425–427, 1986.
- [29] C. E. Shannon, A mathematical theory of communication, *Bell System Technical Journal* 27(3): 379–423, (4): 623–656, 1948.
- [30] G. K. Stathopoulos, Variants of stable marriage algorithms, complexity and structural properties, MSc thesis, University of Athens, 2011.
- [31] E. G. Thurber, Concerning the maximum number of stable matchings in the stable marriage problem, *Discrete Mathematics* 248: 195–219, 2002.
- [32] W. A. Whitworth, *Choice and Chance*, Fifth ed., Cambridge, 1901.