SORTING PROBABILITY FOR LARGE YOUNG DIAGRAMS

SWEE HONG CHAN, IGOR PAK, AND GRETA PANOVA

ABSTRACT. For a finite poset $P = (X, \prec)$, let $\mathcal{L}_P$ denote the set of linear extensions of P. The *sorting probability* $\delta(P)$ is defined as

$$\delta(P) := \min_{x,y \in X} \left| \mathbf{P}\left[L(x) \leq L(y)\right] - \mathbf{P}\left[L(y) \leq L(x)\right]\right|,$$

where $L \in \mathcal{L}_P$ is a uniform linear extension of P. We give asymptotic upper bounds on sorting probabilities for posets associated with large Young diagrams and large skew Young diagrams, with bounded number of rows.

1. INTRODUCTION

Random linear extensions of finite posets occupy an unusual place in combinatorial probability by being remarkably interesting with numerous applications, and at the same time by being unwieldy and lacking general structure. One reason for this lies in the broad nature of posets, when some special cases are highly structured, extremely elegant and well studied, while there is no universal notion of "large poset" or "random poset" in the opposite extreme. As a consequence, the results in the area tend to range widely across the *generality spectrum*: from weaker results for large classes of posets to stronger results for smaller classes of posets.

In this framework, the famous $\frac{1}{3} - \frac{2}{3}$ *Conjecture* 1.1 is very surprising in both the scope and precision, as it bounds the *sorting probability* $\delta(P) \leq \frac{1}{3}$ for *all* finite posets P. There are numerous partial results on the conjecture, as well as the Kahn–Saks general upper bound $\delta(P) \leq \frac{5}{11}$. At the same time, the asymptotic analysis of $\delta(P)$ remains out of reach even for the most classical examples. In this paper we obtain sharp asymptotic upper bounds on $\delta(P)$ for large Young diagrams and large skew Young diagrams. These are the first asymptotic results of this type, as we are moving down the generality spectrum.

1.1. **Sorting probability.** Let $P = (X, \prec)$ be a finite poset with $n = |X|$ elements. A *linear extension* L of P is an order preserving bijection $L : X \to [n] = \{1, \ldots, n\}$, so that $x \prec y$ implies $L(x) < L(y)$ for all $x, y \in X$. The set of linear extensions is denoted $\mathcal{L}(P)$, and $e(P) = |\mathcal{L}(P)|$ is the *number of linear extensions* of P.

The *sorting probability of two elements* $x, y \in X$, $x \neq y$, is defined as

$$(1.1) \qquad \delta(P; x, y) := \left| \mathbf{P}\left[L(x) < L(y)\right] - \mathbf{P}\left[L(y) < L(y)\right]\right|,$$

where the probability is over uniform linear extensions $L \in \mathcal{L}(P)$. This is a measure of how independent random linear extensions on elements x and y are. The *sorting probability*[1] of P is defined as:

$$(1.2) \qquad \delta(P) := \min_{x,y \in X, \, x \neq y} \delta(P; x, y).$$

Clearly, $\delta(P) = 1$ when P is a chain, since all pairs of elements are comparable, so $\delta(P; x, y) = 1$ for all $x, y \in X$. The idea of the sorting probability $\delta(P)$ is to measure how close to $1/2$ can one get the probabilities in (1.1).

Conjecture 1.1 (The $\frac{1}{3} - \frac{2}{3}$ Conjecture)**.** *For every finite poset $P = (X, \prec)$ that is not a chain, we have* $\delta(P) \leq \frac{1}{3}$.

---

[1]There seem to be multiple conflicting notations for variations of the sorting probability used in the literature. Notably, in [BFT95, Sah21] the notation $\delta(P)$ means what we denote by $\frac{1}{2}\left(1 - \delta(P)\right)$. We hope this will not lead to confusion.

This celebrated conjecture was initially motivated by applications to sorting under partial information, but quickly became a challenging problem of independent interest, and inspired a great deal of work, including our investigation. To quote [BFT95], this "remains one of the most intriguing problems in the combinatorial theory of posets." We discuss the history and previous results on the conjecture later in the section, after we present our main results (see also §13.1).

1.2. **Main results.** Let $\lambda = (\lambda_1, \ldots, \lambda_d) \vdash n$ be an integer partition with at most d parts. We use $\ell(\lambda)$ to denote the number of parts and $|\lambda|$ the size of the partition. Denote by P_λ the poset associated with λ, with elements squares of the Young diagram, and the order defined by $(i,j) \preccurlyeq (i',j')$ if and only if $i \leq i'$ and $j \leq j'$. The linear extensions $L \in \mathcal{L}(P_\lambda)$ are exactly the *standard Young tableaux* of shape λ, see Figure 1.1.
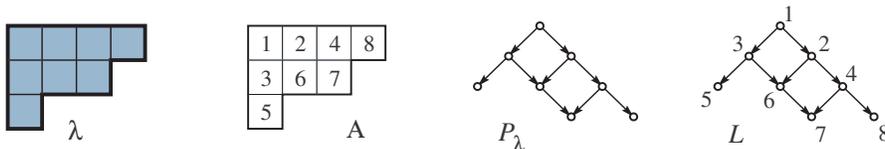

FIGURE 1.1. Young diagram $\lambda = (4,3,1)$, standard Young tableau $A \in \mathrm{SYT}(\lambda)$, poset P_λ, and the corresponding linear extension $L \in \mathcal{L}(P_\lambda)$.

We state our results, roughly, from less general to more general. Let $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}_+^d$, $\alpha_1 \geq \ldots \geq \alpha_d \geq 0$, and $|\alpha| = 1$, where $|\alpha| := \alpha_1 + \ldots + \alpha_d$. Such α are called *Thoma sequences*. Define a *Thoma–Vershik–Kerov* (TVK) α-shape $\lambda \simeq \alpha n$, to be partition $\lambda = (\lambda_1, \ldots, \lambda_d)$, with $\lambda_i = \lfloor \alpha_i n \rfloor$, for all $1 \leq i \leq d$. Note that $|\lambda| = n - O(1)$ in this case.

Theorem 1.2. *Fix $d \geq 2$. For every Thoma sequence $\alpha \in \mathbb{R}_{>0}^d$, there is universal constant C_α, s.t.*

$$\delta(P_\lambda) \leq \frac{C_\alpha}{\sqrt{n}},$$

where $\lambda \simeq \alpha n$ is a TVK α-shape.

We say that a partition $\lambda \vdash n$ is *ε-thick*, if the smallest part $\lambda_d \geq \varepsilon n$, where $d = \ell(\lambda)$.

Theorem 1.3. *Fix $d \geq 2$. For every $\varepsilon > 0$, there is a universal constant $C_{d,\varepsilon}$, such that for every ε-thick partition $\lambda \vdash n$ with $\ell(\lambda) = d$ parts, we have:*

$$\delta(P_\lambda) \leq \frac{C_{d,\varepsilon}}{\sqrt{n}}.$$

Clearly, every TVK α-shape is ε-thick when $0 < \varepsilon < \alpha_d$, and n is large enough. Thus, Theorem 1.3 can be viewed as an advanced generalization of Theorem 1.2.

Let $\lambda = (\lambda_1, \ldots, \lambda_d)$, $\mu = (\mu_1, \ldots, \mu_d)$ be two partitions with at most d parts, and such that $|\lambda/\mu| := |\lambda| - |\mu| = n$. We write $\mu \subset \lambda$, if $\mu_i \leq \lambda_i$ for all $1 \leq i \leq d$, and refer to λ/μ as *skew partition* (see Figure 1.2). Since poset P_μ is a subposet of P_λ, poset $P_{\lambda/\mu}$ is defined as their difference.

FIGURE 1.2. Skew Young diagram λ/μ and poset $P_{\lambda/\mu}$, where $\lambda = (5,5,4,2)$ and $\mu = (3,2,0,0)$.

Let $\alpha = (\alpha_1, \ldots, \alpha_d)$, $\beta = (\beta_1, \ldots, \beta_d) \in \mathbb{R}_+^d$, $\alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_d$, $\beta_1 \geq \ldots \geq \beta_d$, $\beta_i \leq \alpha_i$ for all $1 \leq i \leq d$, and $|\alpha| - |\beta| = 1$. Such (α, β) are called *Thoma pairs*. Define a *TVK (α, β)-shape* to be the skew partition λ/μ, where $\lambda \simeq \alpha n$ and $\mu \simeq \beta n$. Note that $|\lambda/\mu| = n + O(1)$ in this case.

Theorem 1.4. *Fix $d \geq 2$. For every Thoma pair (α, β), with $\alpha \in \mathbb{R}_{>0}^d$, $\beta \in \mathbb{R}_+^d$, there is a universal constant $C_{\alpha,\beta}$, s.t.*

$$\delta(P_{\lambda/\mu}) \leq \frac{C_{\alpha,\beta}}{\sqrt{n}},$$

where λ/μ is a TVK (α, β)-shape, i.e. $\lambda \simeq \alpha n$, $\mu \simeq \beta n$.

When $\beta = (0, \ldots, 0)$, we obtain Theorem 1.2 as a special case. We can now state our main result, the analogue of Theorem 1.3 for skew shapes.

We say that a partition λ is *ε-smooth*, if λ is ε-thick, and $\lambda_i - \lambda_{i+1} \geq \varepsilon n$, for all $1 \leq i < d$. For brevity, we say that a skew partition λ/μ is *ε-smooth* if λ is ε-smooth. Note that, despite the notation, this condition does not impose any restriction on μ.

Theorem 1.5 (Main theorem). *Fix $d \geq 2$. For every $\varepsilon > 0$, there is a universal constant $C_{d,\varepsilon}$, such that for every ε-smooth skew partition $\lambda/\mu \vdash n$, with $\ell(\lambda) = d$, we have:*

$$\delta(P_{\lambda/\mu}) \leq \frac{C_{d,\varepsilon}}{\sqrt{n}}.$$

In the TVK case, when $\alpha_1 > \ldots > \alpha_d > 0$, we obtain Theorem 1.4. However, when the inequalities are non-strict, there is no such implication. Similarly, Theorem 1.5 generalizes Theorem 1.3 for $\mu = \varnothing$, and λ is ε-smooth.

The results are proved by using random walks estimates and the technique Morales and the last two authors recently developed in a series of papers [MPP1]–[MPP4] on the *Naruse hook-length formula* (NHLF). Roughly, in order to estimate the sorting probabilities $\delta(P_\lambda; x, y)$, we need very careful bounds on the number of standard Young tableaux $f(\lambda/\nu) := |\mathrm{SYT}(\lambda/\nu)| = e(P_{\lambda/\nu})$ for the typical $\nu \subset \lambda$ obtained after removing x and/or y from λ. The NHLF gives a useful technical tool, which combined with various asymptotic estimates implies the result. We postpone further discussion of our results until after a brief literature review.

1.3. **Prior work on sorting probability.** The $\frac{1}{3} - \frac{2}{3}$ Conjecture 1.1 was proposed independently by Kislitsyn [Kis68] and Fredman [Fre75] in the context of sorting under partial information. The name is motivated by the following attractive equivalent formulation. In notation of (1.2), for every $P = (X, \prec)$ that is not a chain, there exist elements $x, y \in X$, such that

$$(1.3) \qquad\qquad \frac{1}{3} \leq \mathbf{P}[L(x) < L(y)] \leq \frac{2}{3}.$$

A major breakthrough was made by Kahn and Saks [KS84], who proved (1.3) with slightly weaker constants $\frac{3}{11} - \frac{8}{11}$. In our notation, they showed that $\delta(P) \leq \frac{5}{11} \approx 0.4545$ for all finite P. A much simplified geometric proof (with a slightly weaker bound) was given later in [KL91]. By utilizing technical combinatorial tools, the Kahn–Saks bound was slightly improved in [BFT95] to $\delta(P) \leq \frac{1}{\sqrt{5}} \approx 0.4472$, where it currently stands. For more on the history and various related results, we refer the reader to a dated but very useful survey [Bri99].

While the conjecture does not ask for an efficient algorithm for finding the desired elements $x, y \in X$, a nearly optimal sorting algorithm using $O(\log e(P))$ comparisons was found in [KK95]. See also [C+13] for a simpler version.

Note that the bound $\delta(P) \leq \frac{1}{3}$ in the conjecture is tight for a 3-element poset that is a union of a 2-chain and a single element. The effort to establish the conjecture and improve the constants remains very active. First, Linial [Lin84] proved that $\delta(P) \leq \frac{1}{3}$ for posets of width 2, where width(P) is the size of the maximal antichain in P. In this class, Aigner showed that the tight bound $\delta(P) = \frac{1}{3}$ can come only from decomposable posets, and Sah [Sah21] recently improved the bound to a slightly lower bound $\delta(P) < 0.3225$ in the indecomposable case (see also [Chen18]).

Conjecture 1.1 was further established for several other classes of posets, including semiorders [Bri89], N-free posets [Zag12], height 2 posets [TGF92], and posets whose cover graph is a forest [Zag19]. For posets with a nontrivial automorphism the conjecture was proved by Pouzet, see [GHP87], and a stronger bound $\delta(P) < 1 - \frac{2}{e} \approx 0.2642$ was shown by Saks [Saks85]. Closer to the subject of this paper, Olson and Sagan [OS18] recently applied Linial's approach to establish Conjecture 1.1 for all Young diagrams and skew Young diagrams.

There are very few results proving that $\delta(P_n) \to 0$ as $n \to \infty$ for a sequence $\{P_n\}$ of posets on n elements. Some of them are motivated by the following interesting conjecture of Kahn and Sacks [KS84].

Conjecture 1.6 (Kahn–Saks). *Let $\eta(d)$ denotes the supremum of $\delta(P)$ over all finite posets P of width d. Then $\eta(d) \to 0$ as $d \to \infty$.*

The most notable result in this direction is due to Komlós [Kom90], who proved it for height 2 posets, as well as posets with $n/f(n)$ minimal elements, for some undetermined, but possibly very slowly growing function $f(n) = \omega(1)$. Similarly, Korshunov [Kor94] proved that Conjecture 1.6 holds for *random posets*, which are known to have height 3 w.h.p. [KR75]. Note that these are the opposite extremes to our setting, as we consider posets $P_{\lambda/\mu}$ with width $d = O(1)$ and height $\Theta(n)$, see also §13.1.

Before we conclude, let us note that for general posets, counting the number $e(P)$ of linear extensions, as well as computing the sorting probability $\delta(P)$, is #P-complete [BW91]. Thus, there is little hope of getting good asymptotic bounds on $\delta(P)$, except possibly for one of several notions of "random poset" [Bri93] and "large poset" [Jan11]. In fact, the same complexity results hold for counting linear extensions of general 2-dimensional posets, as well as for posets of height 2; both results are recently proved in [DP18]. This makes (skew) Young diagrams refreshingly accessible in comparison.

1.4. Prior work on asymptotics for standard Young tableaux.
The combinatorics of standard Young tableaux is a classical subject, but until relatively recently, much of the work was on exact counting rather than on asymptotics and probabilistic aspects.

The *hook-length formula* (HLF) gives an explicit product formula for $e(P_\lambda) = |\operatorname{SYT}(\lambda)|$, see e.g. [Sta99]. In the *stable limit shape*, the Young diagram λ scaled by $\frac{1}{\sqrt{n}}$ in both directions $\to \pi$, a curve of area 1. Then the HLF gives a tight asymptotic bound for $e(P_\lambda)$ via *hook integral* [VK81] (see also [MPP4]). *Feit's determinant formula* is an exact formula for $f(\lambda/\mu)$, which can also be derived from the *Jacobi–Trudi identity* for skew shapes, see e.g. [Sta99]. Unfortunately, its determinantal nature makes finding exact asymptotics exceedingly difficult, see e.g. [BR10, MPP4].

For large skew shapes, Okounkov–Olshanski [OO98] and Stanley [Sta03] computed the asymptotics of $f(\lambda/\mu)$ for fixed μ, as $|\lambda| \to \infty$. Both papers rely on the *factorial Schur functions* introduced by Macdonald in [Mac92, §6]. The *Naruse hook-length formula* (NHLF) was introduced by Hiroshi Naruse in a talk in 2014, and given multiple proofs and generalizations in [MPP1, MPP2]. While the formula itself is algebro-geometric in nature, coming from the equivariant cohomology of the Grassmannian, some of the proofs are direct and combinatorial, using factorial Schur functions and explicit bijections [Kon, MPP1, MPP2] (see also [Pak21] for an overview).

In [MPP4], Morales–Pak–Panova used the NHLF and the hook integral approach to prove an exact asymptotic formula for $f(\lambda/\mu)$ when λ/μ have a TVK (α, β)-shape. In [MPT18], based on a bijection with lozenge tilings given in [MPP3] and the variational principle in [CKP01], Morales–Pak–Tassy proved an asymptotic formula for $f(\lambda/\mu)$ when both λ and μ have a stable limit shape. In a parallel investigation, Pittel–Romik [PR07] found limit curves for the shape of random Young tableaux of a rectangle. Most recently, Sun [Sun18] established existence of such limit curves for general skew stable limit shapes.

1.5. Some examples.
The main difficulty in estimating the sorting probability is finding the "right" *sorting elements* $x, y \in X$, such that, even when suboptimal, still give a good bound for $\delta(P; x, y)$. To better understand this issue, let us illustrate the sorting probability in some simple examples.

First, take $\lambda = (n, 1)$ and $\mu = (1)$. Then poset $P_{\lambda/\mu}$ consists of two chains, of length 1 and $(n-1)$. There is an easy optimal pair of elements $x = (1, \lfloor \frac{n+1}{2} \rfloor)$ and $y = (2, 1)$. Then $\delta(P_{\lambda/\mu}) = 0$ for even n, and $\delta(P_{\lambda/\mu}) = \frac{1}{n}$ for odd n. Similarly, let $\lambda = (n, 2)$ and $\mu = (2)$. The poset $P_{\lambda/\mu}$ again consists of two chains, of length 2 and $(n-2)$. In this case, the x as above give suboptimal $\delta(P_{\lambda/\mu}; x, y) \sim \frac{1}{2}$. Perhaps counterintuitively, the optimal sorting elements are $y = (2, 1)$ and $x = (1, m)$, where $m = n(1 - \frac{1}{\sqrt{2}}) + O(1)$. We have $\delta(P_{\lambda/\mu}; x, y) = \Theta(\frac{1}{n})$ bound in this case. We generalize this example in §3.2.

Now let $\lambda = (m, m)$, $\mu = \varnothing$, $n = 2m$. We have $f(\lambda/\mu) = |\operatorname{SYT}(m, m)| = \frac{1}{m+1}\binom{2m}{m}$, the *Catalan number*. One can check in this case that $\delta(P_{\lambda/\mu}; x, y) = \Omega(1)$ for $y = (2, 1)$ and every $x = (1, i)$. In fact, the bounds that work in this case are given by $x = (1, \frac{m}{2} + k)$ and $y = (2, \frac{m}{2} - k)$, for some $k = \Theta(\sqrt{m})$. We prove in [CPP21] that $\delta(P_{\lambda/\mu}) = O(n^{-5/4})$ by a direct asymptotic argument. A weaker $O(\frac{1}{\sqrt{n}})$ bound can be proved via standard bijection from standard Young tableaux $A \in \operatorname{SYT}(m, m)$ and Dyck paths

$(0,0) \to (m,m)$, which in the limit $m \to \infty$ converge to the Brownian excursion (see e.g. [Pit06]). This is the motivational example for this paper.

1.6. **Our work in context.** The differences between various approaches can now be explained in the way the authors look for the sorting elements. In [Lin84], Linial takes P of width two, breaks it into two chains, takes x to be the minimal element in one of them and looks for y in another chain. As the previous examples show, this approach can never give $\delta(P) = o(1)$ for general Young tableaux even with two rows. This approach has been influential, and was later refined and applied in a more general settings, see e.g. [Bri89, Zag12].

In [KS84] and followup papers [BFT95, KL91, Kom90, Zag12], a more complicated pigeonhole principle is used, at the end of which there is no clear picture of what sorting elements are chosen. In fact, the geometric approach in [KS84, KL91] can never give $\delta(P) = o(1)$, as they also point out, cf. [Saks85]. The paper most relevant to our paper is [OS18], where the authors look for elements x, y on the boundary $\partial\lambda$, and apply the pigeonhole principle, Linial-style. Already in the Catalan example this approach cannot be used to prove that $\delta(P; x, y) = o(1)$.

Now, following [PR07, Sun18], let $\lambda \vdash n$ be the stable limit shape. It is natural to take x and y from the same limit curve $C_\lambda(\alpha) := \partial\{(i,j) \in \lambda, A(i,j) \leq \alpha n\}$, where $0 < \alpha < 1$, and $A \in \mathrm{SYT}(\lambda)$ is a uniform standard Young tableau of shape λ. An example of these limit curves is given in Figure 1.3. Since the curves $C_\lambda(\alpha)$ have $\Theta(\sqrt{n})$ elements, and all $(i,j) \in C_\lambda(\alpha)$ can be permuted nearly independently, this could in principle give a small sorting probability. Making this precise would be both interesting and challenging, but this approach fails in our case, since we have $d = O(1)$ rows. It does have a few heuristic implications.
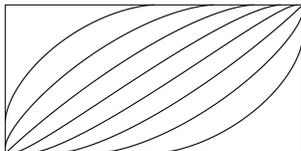

FIGURE 1.3. The limit curves in a $d \times 2d$ rectangle (created by Dan Romik, April 2020).

On the one hand, there are likely many good sorting pairs of elements $x = (i,j)$, $y = (i', j')$, for all $i < j$. On the other hand, in general, the limit curves do not have a closed-form formula of any kind, and arise as the solution of a variational problem [Sun18]. The same holds for the asymptotics of $f(\lambda/\mu)$ [MPT18]. As a consequence, we are essentially forced to make an indirect argument, which proves the result without explicitly specifying the exact location of x, y in λ.

Our approach is based on a combination of tools and ideas from algebraic combinatorics and discrete probability. The general philosophy is somewhat similar to the pigeonhole principle of Linial [Lin84], in a sense that we find a sorting pair $x = (1, a)$ and $y = (2, b)$ by searching over suitable choices of a, b. As in the Catalan case, we start with extreme cases $a = \lambda_1$, $b = \mu_2 + 1$, and decrease $(a - b)$ until the sorting probabilities of x and y becomes small. The main difficulty, of course, is estimating these sorting probabilities.

In fact, by analogy with the Catalan example, one can interpret random standard Young tableaux as random walks from $(0, \ldots, 0)$ to $(\lambda_1, \ldots, \lambda_d)$, which are confined to a certain simplex region in $\mathbb{N}^d$ defined by combinatorial constraints. The sorting probability $\delta(P_{\lambda/\mu}; x, y)$ can then be interpreted as the probability the walk passes below versus above of certain codimension-2 subspace. These probabilities are then bounded by comparing the simplex-confined lattice walk with the usual (unconstrained) lattice walk. This comparison is based on delicate estimates which largely rely on the Schur functions technology combined with the NHLF. This technical part occupies much of the paper.

1.7. **Structure of the paper.** We begin by reviewing standard definitions and notation in Section 2, where we also include a number of basic results in Algebraic Combinatorics and Discrete Probability. In the Warmup Section 3 we prove the $\frac{1}{3} - \frac{2}{3}$ Conjecture 1.1 for all Young diagrams. This is a known result, but the proof we give is new and the tools are a precursor of the proof of the Main Theorem 1.5. We also show how these tools easily give an upper bound on the sorting probability $\delta(P_\lambda)$, for $n - \lambda_1 = o(n)$, where $n = |\lambda|$. In fact, this short section has both the style and the flavor of the rest of the paper, cf. §4.8.

In Section 4, we give key new definitions which allow us to state the Main Lemma 4.3, and two bounds Lemmas 4.4 and 4.5 on the number $f(\lambda/\mu)$ of standard Young tableaux of shape λ/μ. The proofs of these lemmas occupy much of the paper. The technical outline of these proofs is the given in §4.7, so below we only give the structure of the paper in the broadest terms.

First, in Sections 5–7, we develop the technology of lattice path probabilities and their estimates, which culminates with the proof of Main Lemma 4.3 in Section 7. Then, in Section 8, we develop the technology of Young tableaux estimates, which allows us to prove Theorem 1.3 in Section 9. We then prove Lemma 4.4 and Main Theorem 1.5 in Section 10. Finally, Lemma 4.5 and Theorem 1.4 are proved in Section 11.

We conclude with Section 12, where we state several conjectures and open problems motivated by our results. We present final remarks in Section 13.

2. Definitions, notation and background results

2.1. **Standard conventions.** We fix the number of rows $d \geq 2$ throughout the paper. We consider only posets $P = (X, \prec)$ corresponding to partitions $\lambda \vdash n$, or skew partitions $\lambda/\mu \vdash n$. Unless stated otherwise, we have $|X| = n$.

We use $[n] = \{1, \ldots, n\}$, $\mathbb{N} = \{0, 1, 2, \ldots\}$, $\mathbb{Z}_+ = \{1, 2, \ldots\}$, $\mathbb{R}_+ = \{x \geq 0\}$, and $\mathbb{R}_{>0} = \{x > 0\}$. We denote by $\mathbb{P}_d \subset \mathbb{N}^d$ the set of partitions $(\lambda_1, \ldots, \lambda_d)$, where $\lambda_1 \geq \ldots \geq \lambda_d \geq 0$, and $\lambda_i \in \mathbb{N}$. We write $(a_1, \ldots, a_d) \trianglerighteq (b_1, \ldots, b_d)$, when $a_1 \geq b_1$, $a_1 + a_2 \geq b_1 + b_2$, $\ldots$, and $a_1 + \ldots + a_d = b_1 + \ldots + b_d$.

2.2. **Standard Young tableaux.** We adopt standard notation in the area. See e.g. [Mac95, Sag01, Sta99] for these results and further references.

Let $\lambda = (\lambda_1, \ldots, \lambda_d) \vdash n$, $\lambda_1 \geq \ldots \geq \lambda_d \geq 0$, be an *integer partition* of n. Here $n = |\lambda| := \lambda_1 + \ldots + \lambda_d$ denotes the size of λ, and $\ell(\lambda) \leq d$ is the number of parts of λ. We use λ' to denote a conjugate partition whose parts are the column lengths of the diagram λ.

A *skew partition* λ/μ is a pair of partitions $\lambda = (\lambda_1, \ldots, \lambda_d)$, $\mu = (\mu_1, \ldots, \mu_d)$, such that $\mu_i \leq \lambda_i$. In the vector notation above, $\lambda, \mu \in \mathbb{P}_d$, and $\lambda - \mu \in \mathbb{N}^d$. The *empty partition* is $\mu = (0, \ldots, 0)$, which we also denote $\varnothing$, e.g. $\lambda/\varnothing = \lambda$. The size $|\lambda/\mu| := |\lambda| - |\mu|$; we write $\lambda/\mu \vdash n$ for $|\lambda/\mu| = n$.

A *Young diagram* (shape), which we also denote by λ, is a set of squares $(i, j) \in \mathbb{N}^2$, such that $1 \leq i \leq d$, and $1 \leq j \leq \lambda_i$. Similarly, a *skew Young diagram*, which we also denote by λ/μ, is a set of squares $(i, j) \in \mathbb{N}^2$, such that $1 \leq i \leq d$ and $\mu_i < j \leq \lambda_i$. It can in principle have empty rows or be disconnected, although such cases are less interesting. We adopt the *English notation*, where i increases downwards, and j from left to right, as in Figure 1.1.

A *standard Young tableau* of shape λ/μ is a bijection $A : \lambda/\mu \to [n]$, which increases in rows and columns, see Figure 1.1. We use $\mathrm{SYT}(\lambda)$ to denote the set of standard Young tableaux of shape λ/μ. As in the introduction, we use $P_{\lambda/\mu} = (\lambda/\mu, \prec)$ to denote the poset on the set of squares of λ/μ, with the partial order defined by $(i, j) \preccurlyeq (i', j')$ if and only if $i \leq i'$ and $j \leq j'$. This is a standard definition of a *2-dimensional poset* associated with a set of points in the plane, see e.g. [Tro95].

Recall that the linear extensions $\mathcal{L}(P_{\lambda/\mu})$ are in natural bijection with the set $\mathrm{SYT}(\lambda)$ of standard Young tableaux. Whenever clear, we will use the latter from this point on. Denote by $\mathbf{P}_{\lambda/\mu}$ the uniform probability measure on $\mathrm{SYT}(\lambda/\mu)$. To simplify and unify the notation, from now on we use

$$f(\lambda/\mu) := \big|\mathrm{SYT}(\lambda/\mu)\big| = e\big(P_{\lambda/\mu}\big) = \big|\mathcal{L}\big(P_{\lambda/\mu}\big)\big|.$$

For straight shapes $\lambda \vdash n$, we have the *Frobenius formula*:

$$(2.1) \qquad f(\lambda) = \frac{n!}{\lambda_1! \cdots \lambda_d!} \prod_{1 \leqslant i < j \leqslant d} \frac{\lambda_i - \lambda_j + j - i}{\lambda_i + j - i},$$

see e.g. [FRT54] (cf. [Mac92, Ex. 1.1] and [Sta99, Lemma 7.21.1]).

2.3. **Schur polynomials.** A *semistandard Young tableau* of shape λ is an map $A : \lambda \to \mathbb{Z}_+$, such that A is weakly increasing in rows and strictly increasing in columns. We write $\mathrm{SSYT}(\lambda, d)$ for the set of such tableaux with all entries $\leq d$. The *Schur polynomial* is a symmetric polynomial defined as

$$(2.2) \qquad s_\mu(x_1, \ldots, x_d) := \det\big(x_j^{m_i}\big)_{i,j=1}^d \prod_{1 \leq i < j \leq d} (x_i - x_j)^{-1},$$

where $m_i = m_i(\mu) := \mu_i + d - i$. We call $(m_1, \ldots, m_d) = \mu + (d-1, \ldots, 1, 0)$ the *shifted partition* μ.

The combinatorics of Schur functions is given by

$$(2.3) \qquad s_\lambda(x_1, \ldots, x_d) := \sum_{A \in \text{SSYT}(\lambda, d)} \prod_{(i,j) \in \lambda} x_{A(i,j)} = \sum_{A \in \text{SSYT}(\lambda, d)} \prod_{i=1}^{d} (x_i)^{t_i(A)},$$

where

$$(2.4) \qquad t_i(A) := \big| \{ (j,k) \in \lambda/\mu \mid A(j,k) = i \} \big|, \qquad 1 \le i \le d.$$

The product formula below is classical and follows from (2.2) and (2.3):

$$(2.5) \qquad s_\mu(1, \ldots, 1) = \big| \text{SSYT}(\lambda, d) \big| = \prod_{1 \le i < j \le d} \frac{m_i - m_j}{j - i}.$$

2.4. **Hook-length formulas.** The *hook-length* of square $(i,j) \in \lambda$ is defined as

$$(2.6) \qquad h_\lambda(i,j) := \lambda_i - j + \lambda'_j - i + 1.$$

The *hook-length formula* (HLF) [FRT54] (see also [Sag01, Sta99]), is a product formula for the number of standard Young tableaux of straight shape:

$$(2.7) \qquad f(\lambda) = n! \prod_{(i,j) \in \lambda} \frac{1}{h_\lambda(i,j)}.$$

For skew Young diagrams, the number $f(\lambda/\mu)$ can be determined by the *Naruse hook-length formula* (NHLF), see [MPP1, MPP2]. Let $D \subset \lambda$ be a subset of squares with the same number of squares in each diagonal as μ. A subset D is called an *excited diagram* if and only if the relation $\preccurlyeq$ on squares of μ holds for the corresponding squares in D. Denote by $\text{ED}(\lambda/\mu)$ the set excited diagram of shape λ/μ. As shown in [MPP1], all $D \in \text{ED}(\lambda/\mu)$ can be obtained from μ by a sequence of *excited moves*: $(i,j) \to (i+1, j+1)$, for some $(i,j) \in D$, s.t. $(i+1, j), (i, j+1) \notin D$.

Theorem 2.1 (NHLF [MPP1]). *For all $\lambda/\mu \vdash n$, we have:*

$$(2.8) \qquad f(\lambda/\mu) = n! \sum_{D \in \text{ED}(\lambda/\mu)} \prod_{(i,j) \in \lambda \smallsetminus D} \frac{1}{h_\lambda(i,j)}.$$

When $\mu = \varnothing$, we obtain the HLF (2.7). The next result is a consequence of the NHLF. Define

$$(2.9) \qquad F(\lambda/\mu) := n! \prod_{(i,j) \in \lambda/\mu} \frac{1}{h_\lambda(i,j)}.$$

Theorem 2.2 ([MPP4, Thm 3.3]). *Let $\lambda/\mu \vdash n$, $\ell(\lambda) \le d$. Then*

$$F(\lambda/\mu) \le f(\lambda/\mu) \le \big| \text{ED}(\lambda/\mu) \big| \cdot F(\lambda/\mu).$$

In an effort to quantify excited diagrams, we follow an equivalent definition given in [MPP1, §3.3]. A *flagged tableau* of shape λ/μ is a tableaux $T \in \text{SSYT}(\mu)$, such that

$$(2.10) \qquad j + T(i,j) - i \le \lambda_{T(i,j)}, \quad \text{for all } (i,j) \in \mu.$$

The corresponding excited diagram is obtained by moving (i,j) for $T(i,j) - i$ steps down the southeast diagonal. The above inequality is a constraint that $D \subset \lambda$. We denote by $\text{FT}(\lambda/\mu)$ the set of flagged tableaux of shape λ/μ, so $|\text{FT}(\lambda/\mu)| = |\text{ED}(\lambda/\mu)|$.

Theorem 2.3 (Flagged NHLF [MPP1]). *For all $\lambda/\mu \vdash n$, we have:*

$$(2.11) \qquad f(\lambda/\mu) = n! \left[\prod_{(i,j) \in \lambda} \frac{1}{h_\lambda(i,j)} \right] \sum_{T \in \text{FT}(\lambda/\mu)} \prod_{(i,j) \in \mu} h_\lambda \big(T(i,j), j + T(i,j) - i \big).$$

2.5. **Bounds on binomial coefficients.** Recall an effective version of the *Stirling formula*:

$$(2.12) \qquad \sqrt{2\pi}\, n^{n+\frac{1}{2}}\, e^{-n} \ \leq\ n! \ \leq\ n^{n+\frac{1}{2}}\, e^{1-n}\,.$$

This implies the following standard result:

Proposition 2.4. *Let a, b be integers such that $a > b > 0$. Then*

$$\frac{\sqrt{2\pi}}{e^2}\,\sqrt{\frac{a}{b(a-b)}}\,\exp\!\big(a\,H(b/a)\big)\ \leq\ \binom{a}{b}\ \leq\ \frac{e}{2\pi}\,\sqrt{\frac{a}{b(a-b)}}\,\exp\!\big(a\,H(b/a)\big),$$

where $H(r) := -r\log r - (1-r)\log(1-r)$ is the binary entropy function.

2.6. **Concentration inequalities.** Consider a simple random walk $\boldsymbol{X} = \big(X_t\big)_{t\geq 0}$ on $\mathbb{R}^d$, with steps $V = \{v_1, \ldots, v_k\} \subset \mathbb{R}^d$ and probability distribution Q on $[k]$:

$$(2.13) \qquad X_0 \ =\ O,\ \ X_{t+1} \ =\ X_t + v_i\,, \quad \text{where}\ \ 1 \leq i \leq k\ \ \text{is chosen with probability}\ \ q_i := Q(i).$$

We will use the following concentration inequality that applies in much more general situation.

Theorem 2.5 (Hoeffding's inequality [Hoe63]). *Let $\boldsymbol{X} = (X_t)_{t\geq 0}$ be a random walk on $\mathbb{R}^d$ with steps V such that $\|v_i\| \leq 1$. Then, for every $t \geq 1$ and $c > 0$,*

$$\mathbf{P}\big[\,\|X_t - \mathbb{E}[X_t]\| \geq c\big] \ \leq\ 2\exp\big(-2c^2/t\big).$$

In §5.5, we will use Hoeffding's inequality for the set of steps $E = \{e_1, \ldots, e_d\}$ which forms the standard basis in $\mathbb{R}^d$, and a certain non-uniform distribution Q on $[d]$.

3. Warmup

In this short section we give a new proof and an extension of the $\frac{1}{3}$–$\frac{2}{3}$ Conjecture 1.1 for Young diagrams. We apply these to give an upper bound for the sorting probability for general Young diagrams.

3.1. **General Young diagrams.** The first part of the following theorem is the result by Olson and Sagan [OS18]. Below, we present a completely different proof of the result. In fact, our sorting pairs of elements are in a different location when compared to [OS18].

Theorem 3.1. *For every $\lambda \vdash n$, we have $\delta(P_\lambda) \leq \frac{1}{3}$. Moreover, $\delta(P_\lambda; x, y) \leq \frac{1}{3}$ for some $x = (1, k) \in \lambda$ and $y = (\ell, 1) \in \lambda$.*

As suggested by the second part of the Theorem, we need to estimate sorting probabilities for pairs of elements in the first row and the first column.

Lemma 3.2. *Let $\lambda \vdash n$, and $\mathbf{P}_\lambda$ denote the probability over uniform standard Young tableaux $A \in \mathrm{SYT}(\lambda)$. Denote*

$$q_i \ :=\ \mathbf{P}_\lambda\big[A(i,1) \,<\, A(1,2) \,<\, A(i+1,1)\big], \quad 1 \leq i \leq \ell - 1, \quad \text{and}$$

$$q_\ell \ :=\ \mathbf{P}_\lambda\big[A(\ell,1) \,<\, A(1,2)\big],$$

where $\ell = \ell(\lambda)$ is the length of the first column. Then $q_1 \geq \ldots \geq q_\ell$, and $q_1 + \ldots + q_\ell = 1$.

We present two proofs of the lemma: the traditional Young tableaux proof and the proof via the Naruse hook-length formula (Theorem 2.3). The former proof is simpler while the latter is amenable for generalizations and asymptotic analysis. We recommend the reader study both proofs.

First proof of Lemma 3.2. Since $A(1, 1) = 1 < A(1, 2)$, the number $A(1, 2)$ must fall in exactly one of the intervals in the lemma. Thus, we have $q_1 + \ldots + q_\ell = 1$.

Let $A \in \mathrm{SYT}(\lambda)$ be a standard Young tableau, such that $A(k, 1) < A(1, 2) < A(k + 1, 1)$, for some $1 \leq k < \ell$. Then $A(1, 1) = 1, \ldots, A(k, 1) = k$, and $A(1, 2) = k + 1$. The number of such tableaux A is then equal to $f(\lambda/\mu^k)$, where $\mu^k = (2, 1^{k-1}) \vdash k + 1$. In the notation of the lemma, we have:

$$(3.1) \qquad q_k \ =\ \frac{f(\lambda/\mu^k)}{f(\lambda)}\,.$$

Clearly $\mu^k \subset \mu^{k+1}$, and so $\lambda/\mu^{k+1} \subset \lambda/\mu^k$. Then $f(\lambda/\mu^{k+1})$ is equal to the number of tableaux $A \in \mathrm{SYT}(\lambda/\mu^k)$ with $A(k, 1) = 1$. Therefore, $f(\lambda/\mu^{k+1}) \leq f(\lambda/\mu^k)$ and $q_{k+1} \leq q_k$. $\qquad\square$

Second proof of Lemma 3.2. We follow the first proof until (3.1). At this point, recall the Naruse hook-length formula (2.8):

$$f(\lambda/\mu) \ = \ (n - |\mu|)! \prod_{(i,j)\in\lambda} \frac{1}{h_\lambda(i,j)} \sum_{D\in\mathrm{ED}(\lambda/\mu)} \prod_{(i,j)\in D} h_\lambda(i,j).$$

Combined with the hook-length formula (2.7), we have:

$$(3.2) \qquad q_k \ = \ \frac{f(\lambda/\mu)}{f(\lambda)} \ = \ \frac{(n-k-1)!}{n!} \sum_{D\in\mathrm{ED}(\lambda/\mu)} \prod_{(i,j)\in D} h_\lambda(i,j).$$

Now, let $\nu := (2, 1^k) \vdash k+2$. We similarly have:

$$(3.3) \qquad q_{k+1} \ = \ \frac{f(\lambda/\nu)}{f(\lambda)} \ = \ \frac{(n-k-2)!}{n!} \sum_{D'\in\mathrm{ED}(\lambda/\nu)} \prod_{(i,j)\in D'} h_\lambda(i,j).$$

Observe that excited diagrams $D' \in \mathrm{ED}(\lambda/\nu)$ are characterized by the locations of the squares $x_c \in D'$ in the diagonal $\{i - j = c\}$, where $-1 \le c \le k$ (see Figure 3.1).
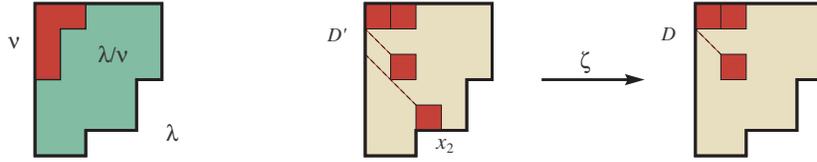

FIGURE 3.1. Skew Young diagram λ/ν, where $\lambda = (5,5,5,4,4,2)$ and $\nu = (2,1,1)$. Map $\zeta : D' \to D$, where $D' \in \mathrm{ED}(\lambda/\nu)$, $D \in \mathrm{ED}(\lambda/\mu)$, and $D' \smallsetminus D = x_2 = (5,3)$.

Consider a map $\zeta : \mathrm{ED}(\lambda/\nu) \to \mathrm{ED}(\lambda/\mu)$, $\zeta(D') = D$, where D is obtained from D' by removing the square x_k. From above and by definition of excited diagrams, map ζ is well defined. This gives:

$$(3.4) \qquad \sum_{D'\in\mathrm{ED}(\lambda/\nu)} \prod_{(i,j)\in D'} h_\lambda(i,j) \ = \ \sum_{D\in\mathrm{ED}(\lambda/\mu)} \prod_{(i,j)\in D} h_\lambda(i,j) \sum_{\substack{(i,j)\in\lambda,\, i-j=k \\ D\cup(i,j)\,\in\,\mathrm{ED}(\lambda/\nu)}} h_\lambda(i,j)$$

The sum on the right is at most

$$(3.5) \qquad h_\lambda(k+1,1) + h_\lambda(k+2,2) + \ldots = \lambda_{k+1} + \ldots + \lambda_\ell \le n - |\mu| = n - k - 1.$$

Combining these equations together, we obtain:

$$q_{k+1} \ =_{(3.2)} \ \frac{(n-k-2)!}{n!} \sum_{D'\in\mathrm{ED}(\lambda/\nu)} \prod_{(i,j)\in D'} h_\lambda(i,j)$$

$$\le_{(3.4)\,(3.5)} \ \frac{(n-k-2)!}{n!} \sum_{D\in\mathrm{ED}(\lambda/\mu)} \prod_{(i,j)\in D} h_\lambda(i,j)\,(n-k-1) \ =_{(3.3)} \ q_k\,,$$

as desired. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

Proof of Theorem 3.1. Without loss of generality, we can assume that

$$p_1 \ := \ \mathbf{P}_\lambda\big[A(1,2) < A(2,1)\big] \ \le \ \frac{1}{2}\,,$$

since we can conjugate diagram λ, otherwise. If $p_1 \ge \frac{1}{3}$, this implies $\delta(P_\lambda; x, y) \le \frac{1}{3}$ for $x = (1,2)$ and $y = (2,1)$, and proves the theorem.

Suppose now that $p \le \frac{1}{3}$. By the lemma, we have:

$$\frac{1}{3} \ \ge \ p_1 \ = \ q_1 \ \ge \ q_2 \ \ge \ \ldots \ \ge \ q_\ell.$$

Observe that

$$p_k \ := \ \mathbf{P}_\lambda\big[A(1,2) < A(k+1,1)\big] \ = \ q_1 + \ldots + q_k\,.$$

Since $q_k \leq \frac{1}{3}$ and $q_1 + \ldots + q_\ell = 1$ by the lemma, this implies that at least one of these probabilities $p_k \in \left[\frac{1}{3}, \frac{2}{3}\right]$. Therefore, the sorting probability $\delta(P_\lambda; x, y) \leq \frac{1}{3}$ for $x = (1, 2)$ and $y = (k + 1, 1)$, as desired. $\qquad\square$

3.2. **General upper bounds.** For a partition $\lambda \vdash n$ define the *imbalance* $\mathsf{q}(\lambda)$ as follows:

$$(3.6) \qquad \mathsf{q}(\lambda) := \frac{1}{n(n-1)} \sum_{i \leq j} h_\lambda(i,i)\, h_\lambda(j, j+1).$$

Note that

$$(3.7) \qquad \sum_{i \leq j} h_\lambda(i,i)\, h_\lambda(j, j+1) \leq \sum_i h_\lambda(i,i) \sum_j h_\lambda(j, j+1) \leq n\,(n-1),$$

so $0 \leq \mathsf{q}(\lambda) \leq 1$. The following result is a generalization of Theorem 3.1.

Theorem 3.3. *For every $\lambda \vdash n$, we have:*

$$\delta(P_\lambda) \leq \min\left\{ \mathsf{q}(\lambda),\, 1 - \mathsf{q}(\lambda),\, |1 - 2\,\mathsf{q}(\lambda)| \right\}.$$

Proof. In the notation of the proof above, let $k = 1$, $\mu = (2)$, and observe that excited diagrams $D \in \mathrm{ED}(\lambda/\mu)$ consist of two squares: (i, i) and $(j, j+1) \in \lambda$, s.t. $1 \leq i \leq j$. Therefore,

$$p_1 = q_1 = \frac{f(\lambda/\mu)}{f(\lambda)} = \frac{1}{n(n-1)} \sum_{D \in \mathrm{ED}(\lambda/\mu)} \prod_{(i,j) \in D} h_\lambda(i,j) =_{(3.6)} \mathsf{q}(\lambda).$$

There are three possibilities. First, if $q_1 \leq \frac{1}{3}$, then the sorting probability $\delta(P_\lambda) \leq q_k \leq q_1$. Similarly, if $q_1 \geq \frac{2}{3}$, by using $\mathsf{q}(\lambda') = 1 - q_1$, we have $\delta(P_\lambda) \leq 1 - \mathsf{q}(\lambda)$. Finally, if $\frac{1}{3} \leq q_1 \leq \frac{2}{3}$, we have $\delta(P_\lambda) \leq |1 - 2q_1|$ by definition of $p_1 = q_1$. This implies the result. $\qquad\square$

Lemma 3.4. *Let $\lambda \vdash n$, and $m = n - \lambda_1$. Then:*

$$\delta(P_\lambda) \leq \frac{mn + (m-1)(m-2)}{n(n-1)}.$$

Proof. We apply Theorem 3.3 to the conjugate partition λ'. We have $h_{\lambda'}(1,1) \leq n$, and

$$\sum_{j \geq 1} h_{\lambda'}(j, j+1) = \sum_{j \geq 1} h_\lambda(j+1, j) = m.$$

Thus, the first term $i = 1$ of the summation (3.6) for the imbalance $\mathsf{q}(\lambda')$, is at most mn. The remaining terms with $i \geq 2$ are equal to $\mathsf{q}(\tau')$, where $\tau = (\lambda_2 - 1, \lambda_3 - 1, \ldots)$ of size $\leq m - 1$. We conclude:

$$\mathsf{q}(\lambda') \leq \frac{1}{n(n-1)}\big(mn + \mathsf{q}(\tau')\big) \leq_{(3.7)} \frac{1}{n(n-1)}\big(mn + (m-1)(m-2)\big),$$

as desired. $\qquad\square$

Corollary 3.5. *Let $\lambda \vdash n$, $m = n - \lambda_1$, and suppose $m = o(n)$. Then $\delta(P_\lambda) = O\left(\frac{m}{n}\right)$.*

We refer to Section 12 for further discussion of general upper bounds.

4. Proof outline

We begin with a number of technical definitions which we present without any motivation. They allow us to state three key lemmas: Main Lemma 4.3, and two asymptotic upper bound Lemmas 4.4 and 4.5. These lemmas follow with a roadmap to the proofs of all theorems in the introduction.

4.1. **The balance function.** Define

$$(4.1) \qquad \Phi(\lambda/\mu) \ := \ \prod_{1 \leqslant i < j \leqslant d} \min\left\{ \mu_i - \mu_j + j - i, \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i} \right\}.$$

We refer to $\Phi(\lambda/\mu)$ as the *balance function* (not to be confused with the *balance constant*). We will need the following simple estimate:

Proposition 4.1. *For all $\lambda \vdash N$, we have:*

$$1 \ \leq \ \Phi(\lambda/\mu) \ \leq \ (dN)^{\frac{d(d-1)}{2}}.$$

Proof. The first inequality follows from

$$(4.2) \qquad\qquad\qquad \lambda_i + d - i \ \geq \ \lambda_i - \lambda_j + j - i, \quad \text{and}$$

$$(4.3) \qquad\qquad\qquad \mu_i - \mu_j + j - i \ \geq \ 1.$$

for all $1 \leqslant i < j \leqslant d$. The second inequality follows from:

$$(4.4) \qquad\qquad\qquad \lambda_i + d - i \ \leq \ dN, \quad \text{and}$$

$$(4.5) \qquad\qquad\qquad \mu_i - \mu_j + j - i \ \leq \ dN,$$

for all $1 \leqslant i < j \leqslant d$. □

4.2. **Definition of ε-admissible pairs.** Fix $\varepsilon > 0$. We say that (λ, μ) is an ε-*admissible pair of partition*, if $\mu \subset \lambda$, $\lambda, \mu \in \mathbb{P}_d$, and

$$(4.6) \qquad\qquad\qquad \lambda_i - \mu_i \ \geq \ \varepsilon|\lambda|, \quad \text{for all } \ 1 \leq i \leq d.$$

Denote by $\Lambda(n, d, \varepsilon)$ the set of ε-admissible pairs of partitions (λ, μ), such that $\lambda/\mu \vdash n$, and $\lambda, \mu \in \mathbb{P}_d$.

Proposition 4.2. *Let $(\lambda, \mu) \in \Lambda(n, d, \varepsilon)$ be an ε-admissible pair, $\lambda, \mu \in \mathbb{P}_d$. Then $\varepsilon \leq 1/d$.*

Proof. We have:

$$\varepsilon \ \leq \ \frac{1}{d} \sum_{i=1}^{d} \frac{\lambda_i - \mu_i}{|\lambda|} \ = \ \frac{|\lambda| - |\mu|}{d|\lambda|} \ \leq \ \frac{1}{d},$$

as desired. □

4.3. **Definition of ε-admissible triplets.** Fix $\varepsilon > 0$. Let $\lambda, \gamma, \mu \in \mathbb{P}_d$, such that $\mu \subseteq \gamma \subseteq \lambda$, and $\lambda/\mu \vdash n$. We say that a triplet (λ, γ, μ) is ε-*separated*, if

$$(4.7) \qquad\qquad \gamma_i - \mu_i \geq \frac{\varepsilon^3 |\lambda|}{2}, \quad \lambda_i - \gamma_i \geq \frac{\varepsilon^3 |\lambda|}{2}, \quad \text{for all } \ 1 \leq i \leq d.$$

In other words, condition (4.7) means that partition γ is bounded away from both μ and λ.

We say that (λ, γ, μ) is *progressive*, if

$$(4.8) \qquad\qquad\qquad \left\| \gamma - (1 - p)\mu - p\lambda \right\| \ \leq \ n^{\frac{3}{4}},$$

where $\| \cdot \|$ denote the ℓ_∞-distance in $\mathbb{R}^d$, and $p := p(\lambda, \gamma, \mu) \in [0, 1]$ is given by

$$(4.9) \qquad\qquad\qquad p \ := \ \frac{1}{n} \left(|\gamma| - |\mu| \right).$$

In other words, condition (4.8) means that γ is close to the weighted average of μ and λ.

Finally, we say that (λ, γ, μ) is an ε-*admissible triplet of partitions*, if $\mu \subseteq \gamma \subseteq \lambda$, the pair (λ, μ) is ε-admissible, and the triplet (λ, γ, μ) is both ε-separated and progressive. We use $\Omega(n, d, \varepsilon)$ to denote the set of ε-admissible triplets.

4.4. **Definition of solid triplets.** Let $(\lambda, \gamma, \mu) \in \Omega(n, d, \varepsilon)$ be an ε-admissible triplet defined above. We say that a triplet (λ, γ, μ) is *solid*, if the following inequalities hold:

$$(4.10) \qquad \frac{f(\gamma/\mu)}{F(\gamma/\mu)} \leq C \cdot \Phi(\gamma/\mu), \quad \frac{f(\lambda/\gamma)}{F(\lambda/\gamma)} \leq C \cdot \Phi(\lambda/\gamma) \quad \text{and} \quad \frac{f(\lambda/\mu)}{F(\lambda/\mu)} \geq \frac{1}{C} \cdot \Phi(\lambda/\mu),$$

where $\Phi(\cdot)$ is the balance function defined in (4.1). We refer to C as the *solid constant* of the triplet.

4.5. **Sorting probability of solid pairs.** Let $(\lambda, \mu) \in \Lambda(n, d, \varepsilon)$ be an ε-admissible pair. We say that a pair (λ, μ) is *solid*, if there is a constant $C_{\lambda,\mu} > 0$, such that every ε-admissible triplet $(\lambda, \gamma, \mu) \in \Omega(n, d, \varepsilon)$ is solid with the solid constant $C_{\lambda,\mu}$.

Lemma 4.3 (Main Lemma). *Fix $d \geq 2$ and $\varepsilon > 0$. Let $(\lambda, \mu) \in \Lambda(n, d, \varepsilon)$. Suppose further, that (λ, μ) is a solid pair, with a solid constant $C = C_{\lambda,\mu} > 0$. Then:*

$$(4.11) \qquad\qquad\qquad \delta(P_{\lambda/\mu}) \leq C_{d,\varepsilon} \, \frac{C^3 + 1}{\sqrt{n}} \,,$$

where $C_{d,\varepsilon} > 0$ is an absolute constant.

Main Lemma 4.3 is proved in Section 7.

4.6. **Asymptotics of $f(\lambda/\mu)$.** The key to proving the theorems in the introduction is proving that $f(\lambda/\mu)$ is equal, up to a multiplicative constant, to the product of $F(\lambda/\mu)$ and the balance function $\Phi(\lambda/\mu)$. Here is the precise statement of the reduction.

Lemma 4.4 (Smooth asymptotics). *Fix $d \geq 2$ and $\varepsilon > 0$. Let λ/μ be a skew partition, such that λ is ε-smooth, and $\lambda, \mu \in \mathbb{P}_d$. Then there exists an absolute constant $C_{d,\varepsilon} > 0$, such that*

$$\frac{1}{C_{d,\varepsilon}} \, \Phi(\lambda/\mu) \, \leq \, \frac{f(\lambda/\mu)}{F(\lambda/\mu)} \, \leq C_{d,\varepsilon} \, \Phi(\lambda/\mu).$$

This is the version we need for the proof of the Main Theorem 1.5. For Theorem 1.4, we need the following similar result.

Lemma 4.5 (TVK asymptotics). *Fix $d \geq 1$. Let (α, β), $\alpha, \beta \in \mathbb{R}_+^d$, be a Thoma pair. Then there is universal constant $C_{\alpha,\beta} > 0$, such that*

$$\frac{1}{C_{\alpha,\beta}} \, \Phi(\lambda/\mu) \, \leq \, \frac{f(\lambda/\mu)}{F(\lambda/\mu)} \, \leq C_{\alpha,\beta} \, \Phi(\lambda/\mu).$$

where λ/μ is a TVK (α, β)-shape, i.e. $\lambda \simeq \alpha n$, $\mu \simeq \beta n$.

4.7. **Roadmap for the rest of the paper.** The next three sections are dedicated to the proof of the Main Lemma 4.3. First, in Section 5, we relate sorting probabilities with the estimates on the number $f(\lambda/\mu)$ of standard Young tableaux, which we then compare with a certain lattice random walk in $\mathbb{R}^d$. The main result of this section is Lemma 5.4, which proves that the probability of having *any* non-ε-admissible triplets is exponentially small. In the following, completely independent Section 6, we obtain various Young tableaux estimates. Here the main result is Lemma 6.7 which gives an upper bound on the number of standard Young tableaux which contain a given ε-admissible triplet. This is the only result which will be used later on. Finally, in a short Section 7, we combine Lemma 5.4 and Lemma 6.7 to prove the Main Lemma 4.3.

We restart anew our analysis of the number $f(\lambda/\mu)$ in Section 8, this time with a different purpose of comparing it to the product $\Phi(\lambda/\mu) F(\lambda/\mu)$. The main results of this section are Lemma 8.3 and Corollary 8.5 which give upper and lower bounds. In Section 9, we prove conceptually simpler estimates required for Theorem 1.3. This section is both a culmination of earlier results, and a training bound for the next two sections.

In Section 9, we use results from Section 8 to prove Lemma 4.4. We then combine it with the Main Lemma 4.3 to prove Theorem 1.5 in a short Section 10. Similarly, in a much longer and more technical Section 11, we first prove Lemma 4.5, which is then combined with the Main Lemma 4.3 to prove Theorem 1.4.

4.8. **A tale of two styles.** The underlying logic of the paper is rather convoluted and somewhat buried in the avalanche of technical estimates, so let us clarify it a bit. There are really two things going on at the same time. On a higher level, we develop various probabilistic tools to obtain the desired estimates. While largely elementary from a technical point of view, these tools seem to be necessary. They are also unavoidably tedious largely because we are starting from scratch in the absence of such approach in the existing literature on the subject.

On a lower level, our probabilistic calculations employ a variety of highly technical estimates on a host of Young tableau parameters. Some of the tools involved, such as NHLF (2.8), are relatively recent and come from a long series of works in Algebraic Combinatorics, including some by the last two authors. While we make our presentation largely self-contained and clarify the NHLF in the Warmup Section 3, this technology remains difficult and yet to be fully understood on a conceptual level.

To make a musical comparison, we have a guitar duo with a new accessible melody played on a lead guitar, paired with a famously difficult theme on a rhythm guitar. The result may appear cacophonous at first, but we hope the reader can persevere, become oblivious to the noise, and learn to appreciate the tune.

5. Standard Young tableaux as lattice paths

We interpret the standard Young tableaux $A \in \mathrm{SYT}(\lambda/\mu)$ as lattice paths within a simplex in $\mathbb{N}^d$. We compare them to unconstrained lattice paths to estimate the sorting probabilities.

5.1. **Setup.** Let $\lambda/\mu \vdash n$, and let $L \in \mathrm{SYT}(\lambda/\mu)$ be a uniform random standard Young tableau. Denote by $\boldsymbol{Z} = (Z_0, Z_1, \ldots, Z_n)$ the sequence of $Z_0 = \mu$, $Z_n = \lambda$, and $Z_t = \{(i,j) \,|\, L(i,j) \le t\}$ is a partition. Denote by $\mathrm{Path}(\lambda/\mu)$ the set of all such lattice paths $\boldsymbol{Z} : \mu \to \lambda$. Note that $\mathrm{Path}(\lambda/\mu)$ is in bijection with $\mathrm{SYT}(\lambda/\mu)$.

We write $\boldsymbol{Z}$ as a sequence of vectors $\big(Z_t(1), \ldots, Z_t(d)\big)_{0 \le t \le n} \in \mathbb{P}_d$. From this point on, we think of $Z_t \in \mathbb{P}_d$ as a random vector, and the sequence $(Z_0, Z_1, \ldots, Z_n)$ as a random lattice path $\boldsymbol{Z} : \mu \to \lambda$ in $\mathbb{P}_d$. We refer to $\boldsymbol{Z}$ as *tableau random walk*. Recall that $\mathbf{P}_{\lambda/\mu}$ denotes the probability over uniform standard Young tableaux $A \in \mathrm{SYT}(\lambda/\mu)$. By a mild abuse of notation, we refer to tableau random walks $\boldsymbol{Z}$ as being sampled from $\mathbf{P}_{\lambda/\mu}$.

Below we give an upper bound for the sorting probability $\delta(P_{\lambda/\mu})$ in terms of the probability of the lattice path $(Z_t)_{t \ge 0}$ visiting a particular codimension 2 hyperplane in $\mathbb{R}^d$.

5.2. **Sorting probability via tableau random walks.** Let (a, b) be two integers, such that $\mu_1 < a \le \lambda_1$ and $\mu_2 < b \le \lambda_2$. Consider the event

$$\mathcal{A}(a,b) := \big\{ (Z_t)_{0 \le t \le n} \mid Z_t(1) = a,\, Z_t(2) = b, \text{ for some } t \ge 0 \big\}.$$

In other words, $\mathcal{A}(a,b)$ is the event that the tableau random walk $\boldsymbol{Z} = (Z_0, \ldots, Z_n)$ intersects the hyperplane in $\mathbb{R}^d$ given by $\big\{ (x_1, \ldots, x_d) \mid x_1 = a,\, x_2 = b \big\}$.

Lemma 5.1. *Let $\lambda, \mu \in \mathbb{P}_d$, $\lambda/\mu \vdash n$, and let $a \in \mathbb{N}$, s.t. $\mu_1 \le a \le \lambda_1$. Define*

$$\varphi(a) := \max_{\mu_2 < k \le \lambda_2} \mathbf{P}_{\lambda/\mu}\big[\,\mathcal{A}(a,k)\,\big]. \tag{5.1}$$

Then there exists $b \in \mathbb{N}$, such that $\mu_2 \le b \le a$,

$$\left| \mathbf{P}_{\lambda/\mu}\big[L(1,a) < L(2,b)\big] \,-\, \frac{1}{2} \right| \,\le\, \varphi(a).$$

In particular, we have

$$\delta(P_{\lambda/\mu}) \,\le\, 2\varphi(a).$$

Proof. Observe that $L(1,a) < L(2,b)$ in the language of paths means $Z_t(1) = a$ and $Z_t(2) < b$, for some $0 \le t \le n$. By taking the probabilities of both events, we then have

$$\mathbf{P}_{\lambda/\mu}\big[L(1,a) < L(2,b)\big] \,=\, \mathbf{P}_{\lambda/\mu}\big[Z_t(1) = a,\, Z_t(2) < b \text{ for some } t \ge 0\big] \,=\, \sum_{k=\mu_2}^{b-1} \mathbf{P}_{\lambda/\mu}\big[\mathcal{A}(a,k)\big].$$

Denote by $W(a, b)$ the sum on the right. It then suffices to show that $W(a, b) \in \left[\frac{1}{2} - \varphi, \frac{1}{2} + \varphi\right]$ for some $b \in [\mu_2, \lambda_2]$ and $\varphi > 0$.

Note that, when $b = \mu_2$, the sum has zero summands, so $W(a, b) = 0$. On the other hand, when $b = a$, we have $W(a, b) = 1$. As the sum is nondecreasing, there exists an integer $b' \in [\mu_2, a)$, such that $W(a, b') < \frac{1}{2}$, while $W(a, b' + 1) \geq \frac{1}{2}$. This completes our proof. $\qquad\square$

5.3. Conditioned lattice random walks are tableau random walks. Fix $\lambda/\mu \vdash n$, where $\lambda, \mu \in \mathbb{P}_d$ as above. Recall the notation in §2.6. Denote by $E = \{e_1, \ldots, e_d\}$ the standard basis in $\mathbb{R}^d$.

Define the *lattice random walk* $\boldsymbol{X} = (X_0, \ldots, X_n)$ on $\mathbb{N}^d$, as follows:

$$(5.2) \qquad X_0 = \mu, \quad X_{t+1} = X_t + e_i, \quad \text{where } i \in [d] \text{ is chosen with probability } q_i := \frac{1}{n}(\lambda_i - \mu_i).$$

Denote by

$$(5.3) \qquad\qquad\qquad\qquad \mathcal{C} := \{X_n = \lambda,\ \boldsymbol{X} \in \mathbb{P}_d\}$$

the event that $\boldsymbol{X} \in \mathrm{Path}(\lambda/\mu)$.

Proposition 5.2.

$$\mathbf{P}\big[\boldsymbol{X} \mid \mathcal{C}\big] = \frac{1}{f(\lambda/\mu)}\,.$$

The proposition is saying that the lattice random walk $\boldsymbol{X}$ conditioned to $\mathcal{C}$ coincides with the tableau random walk $\boldsymbol{Z}$ defined above.

Proof. Suppose $(X_0, \ldots, X_n) \in \mathrm{Path}(\lambda/\mu)$. Then $\boldsymbol{X}$ takes $(\lambda_i - \mu_i)$ steps e_i. Therefore,

$$\mathbf{P}\big[\boldsymbol{X} \mid \mathcal{C}\big] = \mathbf{P}\big[\boldsymbol{X} \mid X_n = \lambda,\ \boldsymbol{X} \in \mathbb{P}_d\big] \propto \prod_{i=1}^{d} (q_i)^{\lambda_i - \mu_i} = \prod_{i=1}^{d} \left(\frac{\lambda_i - \mu_i}{n}\right)^{\lambda_i - \mu_i}.$$

In other words, conditioned on $\mathcal{C}$, the random walk $\boldsymbol{X}$ is uniform in $\mathrm{Path}(\lambda/\mu)$. Since $f(\lambda/\mu) = \big|\mathrm{Path}(\lambda/\mu)\big|$ by definition, we obtain the result. $\qquad\square$

The reason for the non-uniform choice of distribution Q given above will become clear in the next subsection. For now, let us mention that this distribution is chosen so that $\mathbb{E}[X_n] = \lambda$. This is to ensure that the probability $\mathbf{P}\big[\mathcal{C}\big]$ decays polynomially rather than exponentially, i.e., so that the paths in $\mathrm{Path}(\lambda/\mu)$ are living in the typical regime and not the large deviation regime.

5.4. Polynomial decay. Let $\boldsymbol{X} = (X_t)_{1 \leq t \leq n}$ be the random walk on $\mathbb{Z}^d$ defined above. It follows from Proposition 5.2 that $\mathbf{P}[\boldsymbol{X} \mid \mathcal{C}]$ is uniform in $\mathrm{Path}(\lambda/\mu)$. The following lemma gives a lower bound on $\mathbf{P}[\mathcal{C}]$.

Lemma 5.3. *Fix $d \geq 2$. There exists an absolute constant $C_d > 0$ such that the following holds. Let $\lambda/\mu \vdash n$, and $\mu_i < \lambda_i$, for all $1 \leq i \leq d$. Then*

$$\mathbf{P}[\mathcal{C}] \geq C_d\, n^{-\frac{d^2 - 1}{2}}.$$

Proof. It follows from the proof of Proposition 5.2 that

$$\mathbf{P}\big[\mathcal{C}\big] = \frac{\mathbf{P}\big[\boldsymbol{X}\big]}{\mathbf{P}\big[\boldsymbol{X} \mid \mathcal{C}\big]} = f(\lambda/\mu) \prod_{i=1}^{d} \left(\frac{\lambda_i - \mu_i}{n}\right)^{\lambda_i - \mu_i}.$$

Recall the definition of $F(\lambda/\mu)$ in (2.9). Theorem 2.2 and definition (2.6) give:

$$f(\lambda/\mu) \geq F(\lambda/\mu) = n! \prod_{(i,j) \in \lambda\backslash\mu} \frac{1}{h_\lambda(i, j)} \geq n! \prod_{i=1}^{d} \frac{1}{(\lambda_i - \mu_i + d - i)!}.$$

Combining the two equations above, we then get that $\mathbf{P}\big[\mathcal{C}\big]$ is bounded from below by

$$
\mathbf{P}\big[\mathcal{C}\big] \;\geq\; \frac{n!}{n^n} \prod_{i=1}^{d} \frac{(\lambda_i - \mu_i)^{\lambda_i - \mu_i}}{(\lambda_i - \mu_i + d - i)!}
$$

$$
\geq \sqrt{2\pi n}\, e^{-n} \prod_{i=1}^{d} \frac{(\lambda_i - \mu_i)^{\lambda_i - \mu_i}}{(\lambda_i - \mu_i + d - i)^{\lambda_i - \mu_i + d - i + 1/2}\, e^{-\lambda_i + \mu_i - d + i + 1}}
$$

$$
\geq \sqrt{2\pi n}\, e^{-\frac{d(d-1)}{2}} \prod_{i=1}^{d} e^{-d+i+1} \frac{1}{(\lambda_i - \mu_i + d - i)^{d-i+1/2}}
$$

$$
\geq \sqrt{2\pi}\, e^{-d(d-2)} n^{-d^2/2 + 1/2}.
$$

Here we used Stirling's formula (2.12) to bound the factorials and

$$
\left(1 + \frac{d - i}{\lambda_i - \mu_i} \right)^{\lambda_i - \mu_i} \;\leq\; e^{d-i}.
$$

The assumption $\mu_i < \lambda_i$ for all $1 \leq i \leq d$, is used to conclude that $(\lambda_i - \mu_i + d - i) \leq n$. Taking $C_d = \sqrt{2\pi}e^{-d(d-2)}$ implies the result. $\qquad\square$

5.5. Most triplets are ε-admissible.
We can now prove the main result of this section, that the probability of (λ, Z_t, μ) not being ε-admissible is exponentially small.

Lemma 5.4. *Fix $d \geq 2$ and $\varepsilon > 0$. Let $(\lambda, \mu) \in \Lambda(n, d, \varepsilon)$ be an ε-admissible pair. Suppose $t \in \mathbb{N}$ satisfies*

$$
(5.4) \qquad\qquad\qquad \varepsilon^2 \;\leq\; \frac{t}{n} \;\leq\; 1 - \varepsilon^2.
$$

Then there exists a constant $C_{d,\varepsilon} > 0$, such that

$$
\mathbf{P}_{\lambda/\mu}\big[(\lambda, Z_t, \mu) \notin \Omega(n, d, \varepsilon) \big] \;\leq\; C_{d,\varepsilon}\, n^{\frac{d^2 - 1}{2}}\, e^{-2\sqrt{n}}.
$$

Proof. Let

$$
\xi_t \;:=\; \frac{t}{n}(\lambda - \mu) + \mu.
$$

Note that $\xi_t \in \mathbb{R}_+^d$ is not necessarily in $\mathbb{P}_d$. Suppose that Z_t satisfies

$$
(5.5) \qquad\qquad\qquad \|Z_t - \xi_t\| \;\leq\; n^{3/4}.
$$

This assumption implies:

$$
|Z_t(i) - \mu_i| \;\geq\; |\xi_t(i) - \mu_i| - n^{3/4} \;=\; \frac{t}{n}(\lambda_i - \mu_i) - n^{3/4}
$$

$$
\geq_{(5.4)} \varepsilon^2(\lambda_i - \mu_i) - n^{3/4} \;\geq_{(4.6)}\; \varepsilon^3|\lambda| - n^{3/4} \;\geq\; \frac{\varepsilon^3}{2}|\lambda|,
$$

for all $1 \leq i \leq d$, and n large enough. By the same reasoning, the assumption (5.5) implies:

$$
|\lambda_i - Z_t(i)| \;\geq\; \frac{\varepsilon^3}{2}|\lambda|,
$$

for all $1 \leq i \leq d$, and n large enough. By the definitions (4.7) and (4.8) of ε-admissible triplets, the assumption (5.5) implies that $(\lambda, Z_t, \mu) \in \Omega(n, d, \varepsilon)$ for n large enough. We conclude:

$$
\mathbf{P}_{\lambda/\mu}\big[(\lambda, Z_t, \mu) \notin \Omega(n, d, \varepsilon) \big] \;\leq\; \mathbf{P}\big[\|Z_t - \xi_t\| \geq n^{3/4} \big],
$$

for n large enough.

By Proposition 5.2, the lattice random walk $\boldsymbol{X}$ conditioned to $\mathcal{C}$ coincides with $\boldsymbol{Z}$. Observe that $\xi_t = \mathbb{E}[X_t]$. Since $t \le n$, we have:

$$\mathbf{P}\big[\, \|Z_t - \xi_t\| \ge n^{3/4} \,\big] \le \mathbf{P}\big[\, \|X_t - \xi_t\| \ge n^{3/4} \mid \mathcal{C} \,\big] \le \frac{1}{\mathbf{P}[\mathcal{C}]} \cdot \mathbf{P}\big[\, \|X_t - \xi_t\| \ge n^{3/4} \,\big]$$

$$\le_{\text{(Thm 2.5)}} \frac{1}{\mathbf{P}[\mathcal{C}]} \cdot 2\,e^{-2\,n^{3/2}/t} \le \frac{1}{\mathbf{P}[\mathcal{C}]} \cdot 2\,e^{-2\sqrt{n}}$$

$$\le_{\text{(Lem 5.3)}} 2C_d\, n^{\frac{d^2-1}{2}}\, e^{-2\sqrt{n}}\,,$$

for n large enough, and where $C_d > 0$ is the constant from Lemma 5.3. This implies the result. $\qquad\square$

6. Asymptotics and bounds for lattice paths

This section contains bounds and estimates used to bound the sorting probability in the proof of Main Lemma 4.3.

6.1. **Asymptotics for hook-lengths.** In this subsection, we prove an asymptotic estimate for $F(\lambda/\mu)$ defined in (2.9), for all ε-admissible pairs (λ, μ). First, we need the following technical lemma.

Lemma 6.1. *Fix $d \ge 2$ and $\varepsilon > 0$. Let $(\lambda, \mu) \in \Lambda(n, d, \varepsilon)$ be an ε-admissible pair. Then:*

$$\frac{\lambda_i!}{(\lambda_i - \mu_i)!} \le \prod_{j=1}^{\mu_i} h_\lambda(i,j) \le \varepsilon^{-(d-i)} \frac{\lambda_i!}{(\lambda_i - \mu_i)!}\,,$$

for all $1 \le i \le d$.

Proof. The lower bound is clear since $h_\lambda(i,j) \ge \lambda_i - j + 1$. For the upper bound, let J be the largest nonnegative integer such that $\lambda_{J+i} \ge \mu_i$. It follows from the definition of hook-lengths that

$$(6.1) \qquad \prod_{j=1}^{\mu_i} h_\lambda(i,j) = \frac{\lambda_i!}{(\lambda_i - \mu_i)!} \cdot \prod_{k=1}^{d-i}(\lambda_i + k) \prod_{k=1}^{J} \frac{1}{\lambda_i - \mu_i + k} \prod_{k=J+1}^{d-i} \frac{1}{\lambda_i - \lambda_{i+k} + k}\,.$$

First, note that for all $k > J$, we have

$$\lambda_i - \lambda_{i+k} \ge \lambda_i - \mu_i \ge \varepsilon|\lambda|,$$

where the first inequality follows from the maximality of J, and the second inequality follows from (4.6). This implies

$$\prod_{k=1}^{d-i}(\lambda_i + k) \prod_{k=1}^{J} \frac{1}{\lambda_i - \mu_i + k} \prod_{k=J+1}^{d-i} \frac{1}{\lambda_i - \lambda_{i+k} + k}$$

$$\le \prod_{k=1}^{d-i}(|\lambda| + k) \prod_{k=1}^{J} \frac{1}{\varepsilon|\lambda| + k} \prod_{k=J+1}^{d-i} \frac{1}{\varepsilon|\lambda| + k} \le \prod_{k=1}^{d-i} \frac{|\lambda| + k}{\varepsilon|\lambda| + k} \le \varepsilon^{-(d-i)},$$

where the last inequality follows since $\varepsilon^{-1} \ge d \ge 1$, by Proposition 4.2. Together with (6.1), this completes the proof. $\qquad\square$

Lemma 6.2. *Fix $d \ge 2$ and $\varepsilon > 0$. Let $(\lambda, \mu) \in \Lambda(n, d, \varepsilon)$. Then:*

$$G(\lambda/\mu) \le F(\lambda/\mu) \le \varepsilon^{-\frac{d(d-1)}{2}}\, G(\lambda/\mu)\,,$$

where

$$(6.2) \qquad G(\lambda/\mu) := \frac{n!}{(\lambda_1 - \mu_1)! \cdots (\lambda_d - \mu_d)!} \prod_{1 \le i < j \le d} \frac{\lambda_i - \lambda_j + j - i}{\lambda_i + j - i}\,.$$

Proof. By definition (2.9), we have:

$$F(\lambda/\mu) \,=\, n! \prod_{(i,j)\in\lambda} \frac{1}{h_\lambda(i,j)} \prod_{(i,j)\in\mu} h_\lambda(i,j) \,=\, \frac{n!}{\lambda_1! \cdots \lambda_d!} \prod_{1\leqslant i<j\leqslant d} \frac{\lambda_i - \lambda_j + j - i}{\lambda_i + j - i} \prod_{(i,j)\in\mu} h_\lambda(i,j).$$

The result now follows by substituting the upper and lower bounds in Lemma 6.1 to the products of hooks on the RHS, over all $1 \leq i \leq d$. $\qquad\square$

6.2. **Asymptotics for binomial coefficients.** Consider a triplet of partitions (λ, γ, μ), such that $\lambda/\mu \vdash n$. Denote by $\mathbf{y} = \mathbf{y}(\lambda, \gamma, \mu)$ the vector $(y_1, \ldots, y_d) \in \mathbb{R}^d$, defined as

$$(6.3) \qquad y_i \,:=\, \frac{\gamma_i - (1-p)\mu_i - p\lambda_i}{\sqrt{n}}, \quad \text{where } p \in [0,1] \text{ is given by } (4.9).$$

Lemma 6.3. *Fix $d \geq 2$ and $\varepsilon > 0$. Let $(\lambda, \gamma, \mu) \in \Omega(n, d, \varepsilon)$ be an ε-admissible triplet. Then there exists an absolute constant $B_{d,\varepsilon} > 0$, such that*

$$\frac{\binom{\lambda_1-\mu_1}{\gamma_1-\mu_1} \cdots \binom{\lambda_d-\mu_d}{\gamma_d-\mu_d}}{\binom{n}{|\gamma|-|\mu|}} \,\leq\, B_{d,\varepsilon}\, n^{-\frac{(d-1)}{2}} \exp\left[-2\sum_{i=1}^d y_i^2\right],$$

where y_i are as in (6.3).

The lemma follows easily from Proposition 2.4 and the ε-separation property (4.7). We omit the details.

6.3. **Technical lemmas on the bounds.** Denote by ℓ_i, g_i, and m_i the shifted values of λ_i, γ_i, and μ_i, respectively:

$$(6.4) \qquad \ell_i \,:=\, \lambda_i + d - i, \quad g_i \,:=\, \gamma_i + d - i, \quad m_i \,:=\, \mu_i + d - i,$$

for all $1 \leq i \leq d$. Note that

$$(6.5) \qquad \ell_i - \ell_j \,\geq\, 1, \quad g_i - g_j \,\geq\, 1, \quad m_i - m_j \,\geq\, 1,$$

for all $1 \leq i < j \leq d$.

Lemma 6.4. *Let $d \geq 2$, $\varepsilon > 0$. Let $(\lambda, \gamma, \mu) \in \Omega(n, d, \varepsilon)$ be an ε-admissible triplet. Then:*

$$(6.6) \qquad \left(\frac{g_i - g_j}{\gamma_i + j - i}\right) \frac{\min\{m_i - m_j, \frac{g_i}{g_i - g_j}\} \min\{g_i - g_j, \frac{\ell_i}{\ell_i - \ell_j}\}}{\min\{m_i - m_j, \frac{\ell_i}{\ell_i - \ell_j}\}} \,\leq\, \frac{32}{d^2 \varepsilon^{12}} \left((y_i - y_j)^2 + 1\right),$$

for all $1 \leqslant i < j \leqslant d$.

We now build toward the proof of Lemma 6.4.

Sublemma 6.5. *Let $x, y, c \in \mathbb{R}_+$, we have:*

$$\min\{1, c\} \cdot \min\{x, y\} \,\leq\, \min\{x, cy\} \,\leq\, \max\{1, c\} \cdot \min\{x, y\}.$$

Sublemma 6.6. *For all $x, y, z \in \mathbb{R}_+$, we have:*

$$y \cdot \frac{\min\{x, \frac{1}{y}\} \cdot \min\{y, \frac{1}{z}\}}{\min\{x, \frac{1}{z}\}} \,\leq\, 4\left(y - \frac{x+z}{2}\right)^2 + 4.$$

Both sublemmas are elementary; we omit their proof for brevity.

Proof of Lemma 6.4. We start with estimating g_i, ℓ_i, and γ_i. Since (λ, γ, μ) is ε-admissible, we have the following upper bound for g_i and ℓ_i:

$$(6.7) \qquad g_i \leq \ell_i \,=\, \lambda_i + d - i \,\leq\, |\lambda| + d - i \,\leq\, d|\lambda| \,\leq_{(4.6)}\, \sum_{i=1}^d \frac{\lambda_i - \mu_i}{\varepsilon} \,=\, \frac{n}{\varepsilon}.$$

Similarly, we have the following lower bounds:

$$(6.8) \qquad \ell_i \,=\, \lambda_i + d - i \,\geq\, \lambda_i - \mu_i \,\geq_{(4.6)}\, \varepsilon|\lambda| \,\geq\, \varepsilon n,$$

$$(6.9) \qquad \gamma_i + j - i \geq \gamma_i - \mu_i \geq_{(4.7)} \frac{\varepsilon^3}{2} |\lambda| \geq \frac{\varepsilon^3}{2} n.$$

Finally, we have the following lower and upper bounds for p defined in (4.9):

$$(6.10) \qquad p = 1 - \frac{|\lambda| - |\gamma|}{n} \leq_{(4.7)} 1 - \frac{d\varepsilon^3 |\lambda|}{2n} \leq 1 - \frac{d\varepsilon^3}{2},$$

$$(6.11) \qquad p = \frac{|\gamma| - |\mu|}{n} \geq_{(4.7)} \frac{d\varepsilon^3 |\lambda|}{2n} \geq \frac{d\varepsilon^3}{2}.$$

Now note that

$$(6.12) \qquad \frac{g_i - g_j}{\gamma_i + j - 1} \leq_{(6.9)} \frac{2(g_i - g_j)}{\varepsilon^3 n}.$$

Using repeatedly Sublemma 6.5, Proposition 4.2 and the above inequalities, we obtain:

$$(6.13) \qquad \min\left\{m_i - m_j, \frac{g_i}{g_i - g_j}\right\} \leq_{(6.10),\,(6.7)} \frac{1}{d\varepsilon^4} \min\left\{2(1-p)(m_i - m_j), \frac{n}{g_i - g_j}\right\},$$

$$(6.14) \qquad \min\left\{g_i - g_j, \frac{\ell_i}{\ell_i - \ell_j}\right\} \leq_{(6.7)} \frac{2}{\varepsilon} \min\left\{g_i - g_j, \frac{n}{2p(\ell_i - \ell_j)}\right\},$$

$$(6.15) \qquad \min\left\{m_i - m_j, \frac{\ell_i}{\ell_i - \ell_j}\right\} \geq_{(6.8),\,(6.11)} \frac{d\varepsilon^4}{2} \min\left\{2(1-p)(m_i - m_j), \frac{n}{2p(\ell_i - \ell_j)}\right\}.$$

By dividing (6.13)–(6.15) by $\sqrt{n}$, and combining these upper bounds with (6.12), we conclude:

$$\text{LHS in (6.6)} \leq \frac{8}{d^2 \varepsilon^{12}} \left(\frac{g_i - g_j}{\sqrt{n}}\right) \frac{\min\left\{\frac{2(1-p)(m_i - m_j)}{\sqrt{n}}, \frac{\sqrt{n}}{g_i - g_j}\right\} \cdot \min\left\{\frac{g_i - g_j}{\sqrt{n}}, \frac{\sqrt{n}}{\ell_i - \ell_j}\right\}}{\min\left\{\frac{m_i - m_j}{\sqrt{n}}, \frac{\sqrt{n}}{\ell_i - \ell_j}\right\}}.$$

By Sublemma 6.6, the RHS of the equation above is bounded by

$$\frac{32}{d^2 \varepsilon^{12}} \left(B_{ij}^2 + 1\right), \quad \text{where} \quad B_{ij} = \frac{(g_i - g_j) - (1-p)(m_i - m_j) - p(\ell_i - \ell_j)}{\sqrt{n}} =$$

$$=_{(6.4)} \frac{(\gamma_i - \gamma_j) - (1-p)(\mu_i - \mu_j) - p(\lambda_i - \lambda_j)}{\sqrt{n}} =_{(6.3)} (y_i - y_j).$$

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

6.4. Upper bounds for solid triplets.
Recall the definition of solid triplets in §4.4. We can now give an upper bound for the probability that a tableau random walk $\boldsymbol{Z} : \mu \to \lambda$ goes through γ.

Lemma 6.7. *Fix $d \geq 2$ and $\varepsilon > 0$. Let $(\lambda, \gamma, \mu) \in \Omega(n, d, \varepsilon)$ be an ε-admissible solid triplet, with the solid constant C defined in (4.10). Let $k := |\gamma| - |\mu|$. Then*

$$(6.16) \qquad \mathbf{P}_{\lambda/\mu}\left[Z_k = \gamma\right] \leq C^3 C_{d,\varepsilon}\, n^{\frac{1-d}{2}} \prod_{1 \leq i < j \leq d} \left((y_i - y_j)^2 + 1\right) \cdot \exp\left[-2 \sum_{i=1}^d y_i^2\right],$$

for an absolute constant $C_{d,\varepsilon} > 0$.

Note that the RHS in the lemma does not depend on k. This is by design, as k will not be known, so we need a general upper bound.

Proof. By directly counting the number of lattice paths $\mu \to \gamma \to \lambda$, we obtain:

$$(6.17) \qquad \mathbf{P}_{\lambda/\mu}\left[Z_k = \gamma\right] = \frac{f(\gamma/\mu)\, f(\lambda/\gamma)}{f(\lambda/\mu)} \leq_{(4.10)} C^3 \left[\frac{F(\gamma/\mu)\, F(\lambda/\gamma)}{F(\lambda/\mu)}\right] \left[\frac{\Phi(\gamma/\mu)\, \Phi(\lambda/\gamma)}{\Phi(\lambda/\mu)}\right].$$

We now give an upper bound for the first product term:

$$
\frac{F(\gamma/\mu)\,F(\lambda/\gamma)}{F(\lambda/\mu)} \;\leq_{\text{Lem } 6.2}\; \varepsilon^{-d(d-1)}\,\frac{G(\gamma/\mu)\,G(\lambda/\gamma)}{G(\lambda/\mu)}
$$

$$
(6.18) \qquad \leq_{(6.2)}\; \varepsilon^{-d(d-1)}\left[\binom{\lambda_1-\mu_1}{\gamma_1-\mu_1}\cdots\binom{\lambda_d-\mu_d}{\gamma_d-\mu_d}\binom{n}{k}^{-1}\right]\prod_{1\leqslant i<j\leqslant d}\frac{g_i-g_j}{\gamma_i+j-i}
$$

$$
\leq_{\text{Lem } 6.3}\; \varepsilon^{-d(d-1)}\,B_{d,\varepsilon}\,n^{-\frac{(d-1)}{2}}\exp\left[-2\sum_{i=1}^{d}y_i^2\right]\prod_{1\leqslant i<j\leqslant d}\frac{g_i-g_j}{\gamma_i+j-i}.
$$

Combining the last products in RHS of (6.17) and (6.18), we have

$$
\prod_{1\leqslant i<j\leqslant d}\frac{g_i-g_j}{\gamma_i+j-i}\left[\frac{\Phi(\gamma/\mu)\,\Phi(\lambda/\gamma)}{\Phi(\lambda/\mu)}\right]
$$

$$
(6.19) \qquad \leq_{(4.1)}\; \prod_{1\leqslant i<j\leqslant d}\left(\frac{g_i-g_j}{\gamma_i+j-i}\right)\frac{\min\{m_i-m_j,\frac{g_i}{g_i-g_j}\}\,\min\{g_i-g_j,\frac{\ell_i}{\ell_i-\ell_j}\}}{\min\{m_i-m_j,\frac{\ell_i}{\ell_i-\ell_j}\}}
$$

$$
\leq_{\text{Lem } 6.4}\; \left(\frac{32}{d^2\,\varepsilon^{12}}\right)^{d(d-1)/2}\prod_{1\leqslant i<j\leqslant d}\big((y_i-y_j)^2+1\big).
$$

The lemma now follows by combining (6.17), (6.18) and (6.19). □

7. Sorting probability via lattice paths

We use an upper bound for the probability mass function of Z_t and the results of Section 5 which show that most triples are ε-admissible, see Lemma 7.1 below for a precise statement. The upper bounds are derived via some technical asymptotic bounds from Section 6.

7.1. Sorting probability of ε-admissible pairs. The following technical lemma is central to our proof.

Lemma 7.1. *Fix $d \geq 2$ and $\varepsilon > 0$. Let $(\lambda, \mu) \in \Lambda(n, d, \varepsilon)$ be an ε-admissible pair. Let a be an integer that satisfies*

$$
(7.1) \qquad\qquad \varepsilon \;\leq\; \frac{a-\mu_1}{\lambda_1-\mu_1} \;\leq\; 1-\varepsilon.
$$

Suppose there exists a constant $C = C_{\lambda,\mu} > 0$, such that for every γ for which $(\lambda, \gamma, \mu) \in \Omega(n, d, \varepsilon)$, this triplet is solid with solid constant C. Then, there exists an absolute constant $C_{d,\varepsilon} > 0$ such that

$$
\varphi(a) \;\leq\; C_{d,\varepsilon}\,\frac{C^3+1}{\sqrt{n}}\,.
$$

Proof. Let b be an arbitrary integer in $[\mu_2, \lambda_2]$. It follows from the definition of $\varphi(a)$ in (5.1) that it suffices to show that

$$
\mathbf{P}_{\lambda/\mu}\big[\,\mathcal{A}(a,b)\big] \;\leq\; C_{d,\varepsilon}\,\frac{C^3+1}{\sqrt{n}}, \quad \text{for all } \mu_2 \leq b \leq \lambda_2\,.
$$

We start with

$$
\begin{aligned}
\mathbf{P}_{\lambda/\mu}\big[\,\mathcal{A}(a,b)\big] \;&=\; \mathbf{P}_{\lambda/\mu}\big[Z_t(1)=a,\,Z_t(2)=b \text{ for some } t \geq 0\big] \\
(7.2) \qquad &=\; \mathbf{P}_{\lambda/\mu}\big[Z_t(1)=a,\,Z_t(2)=b,\,(\lambda, Z_t, \mu) \notin \Omega(n,d,\varepsilon) \text{ for some } t \geq 0\big] \\
&\quad+\; \mathbf{P}_{\lambda/\mu}\big[Z_t(1)=a,\,Z_t(2)=b,\,(\lambda, Z_t, \mu) \in \Omega(n,d,\varepsilon) \text{ for some } t \geq 0\big].
\end{aligned}
$$

We will bound each term in the RHS separately.

Since $Z_t(1) = a$, by definition of Z_t we have:

$$
a - \mu_1 \;\leq\; t \;\leq\; n - (\lambda_1 - a).
$$

From (7.1) and the assumption that $(\lambda, \mu) \in \Lambda(n, d, \varepsilon)$, it follows that

$$\frac{t}{n} \geq \frac{a - \mu_1}{n} \geq \varepsilon \frac{\lambda_1 - \mu_1}{n} \geq \varepsilon^2 \frac{|\lambda|}{n} \geq \varepsilon^2, \quad \text{and}$$

$$1 - \frac{t}{n} \geq \frac{a - \lambda_1}{n} \geq \varepsilon \frac{\lambda_1 - \mu_1}{n} \geq \varepsilon^2.$$

This implies that condition (5.4) holds. By Lemma 5.4, we then get:

$$\mathbf{P}_{\lambda/\mu}\big[Z_t(1) = a, \ (\lambda, Z_t, \mu) \notin \Omega(n, d, \varepsilon)\big] \leq C_{d,\varepsilon}\, n^{\frac{d^2-1}{2}}\, e^{-2\sqrt{n}},$$

for all $0 \leq t \leq n$, and for some absolute constant $C_{d,\varepsilon} > 0$. Thus, for the first term in the RHS of (7.2), we have:

$$\mathbf{P}_{\lambda/\mu}[Z_t(1) = a, Z_t(a) = b, \ (\lambda, Z_t, \mu) \notin \Omega(n, d, \varepsilon) \text{ for some } t \geq 0]$$

(7.3)
$$\leq \sum_{t=0}^{n-1} C_{d,\varepsilon}\, n^{\frac{d^2-1}{2}}\, e^{-2\sqrt{n}} = C_{d,\varepsilon}\, n^{\frac{d^2+1}{2}}\, e^{-2\sqrt{n}}.$$

For the second term in the RHS of (7.2), denote by $\mathcal{G}(a, b)$ the set of partitions given by

$$\mathcal{G}(a, b) := \big\{\gamma \mid \gamma_1 = a, \ \gamma_2 = b, \ \text{and} \ \mathcal{G}(a, b) \in \Lambda(n, d, \varepsilon)\big\}.$$

Then

$$\mathbf{P}_{\lambda/\mu}\big[Z_t(1) = a, \ Z_t(2) = b, \ (\lambda, Z_t, \mu) \in \Omega(n, d, \varepsilon) \text{ for some } t \geq 0\big] \leq \sum_{\gamma \in \mathcal{G}(a,b)} \mathbf{P}_{\lambda/\mu}\big[Z_t = \gamma\big]$$

(7.4)
$$\leq_{\text{Lem } 6.7} \sum_{\gamma \in \mathcal{G}(a,b)} C^3\, n^{-\frac{(d-1)}{2}} \exp\left[-2\sum_{i=1}^{d} y_i^2\right] \prod_{1 \leq i < j \leq d} \big((y_i - y_j)^2 + 1\big).$$

Using $(x - z)^2 + 1 \leq (x^2 + 1)(z^2 + 1)$, we obtain:

$$\exp\left[-2\sum_{i=1}^{d} y_i^2\right] \prod_{1 \leq i < j \leq d} \big((y_i - y_j)^2 + 1\big) \leq \prod_{i=1}^{d} \big(y_i^2 + 1\big)^{d-1} e^{-2y_i^2}.$$

Plugging this upper bound into (7.4), we obtain:

(7.5)
$$\text{RHS of (7.4)} \leq C^3\, n^{-\frac{(d-1)}{2}} \sum_{\gamma \in \mathcal{G}(a,b)} \prod_{i=1}^{d} (y_i^2 + 1)^{d-1} e^{-2y_i^2}.$$

Note that for all $\gamma \in \mathcal{G}(a, b)$, the value y_1 and y_2 is fixed by the assumption that $\gamma_1 = a$ and $\gamma_2 = b$. For $i \in \{3, \dots, d\}$, it follows from (6.3) that as γ varies between μ and λ, an increment of γ_i to $\gamma_i' = \gamma_i + 1$ would lead to an increment in the y's of order $\big|y_i' - y_i\big| = n^{-1/2}(1 - \frac{\lambda_i - \mu_i}{n}) \geq n^{-1/2}(d-1)\varepsilon$ (by ε-admissibility). Thus, we can bound each term for $i \in \{3, \dots, d\}$, as

$$\sum_{z \in n^{-1/2}(d-1)\varepsilon\, \mathbb{Z}} (z^2 + 1)^{d-1} e^{-2z^2} \leq \frac{\sqrt{n}}{(d-1)\varepsilon} \int_{-\infty}^{+\infty} (z^2 + 1)^{d-1} e^{-2z^2}\, dz \leq \sqrt{n}\, C_{d,\varepsilon}',$$

since the integral converges. This allows us to bound (7.5) as

$$\sum_{\gamma \in \mathcal{G}(a,b)} \prod_{i=1}^{d} (y_i^2 + 1)^{d-1} e^{-2y_i^2} \leq \left[\prod_{i=1,2} (y_i^2 + 1)^{d-1} e^{-2y_i^2}\right] n^{\frac{d-2}{2}} \big(C_{d,\varepsilon}'\big)^{d-2} \leq C_{d,\varepsilon}''\, n^{\frac{d-2}{2}},$$

where $C_{d,\varepsilon}'' := \big(\frac{d-1}{2}\big)^{2(d-1)} e^{-2(d+3)} (C_{d,\varepsilon}')^{d-2}$. Thus we get the following upper bound for the second term in the RHS of (7.2):

(7.6)
$$\mathbf{P}_{\lambda/\mu}\big[Z_t(1) = a, \ Z_t(2) = b, \ (\lambda, Z_t, \mu) \in \Omega(n, d, \varepsilon) \text{ for some } t \geq 0\big] \leq C_{d,\varepsilon}''\, C^3\, \frac{1}{\sqrt{n}}.$$

Using the upper bounds from (7.3) and (7.6) in (7.2), gives us:

$$\mathbf{P}_{\lambda/\mu}\big[\mathcal{A}(a, b)\big] \leq C_{d,\varepsilon}\, n^{\frac{d^2+1}{2}}\, e^{-2\sqrt{n}} + C_{d,\varepsilon}''\, C^3\, \frac{1}{\sqrt{n}}.$$

Since the second term dominates for sufficiently large n, we obtain:

$$\mathbf{P}_{\lambda/\mu}\left[\,\mathcal{A}(a,b)\right] \leq C_{d,\varepsilon}\,\frac{C^3+1}{\sqrt{n}}\,,$$

as desired. $\qquad\square$

7.2. **Proof of Main Lemma 4.3.** Let $a := \lfloor\frac{\mu_1+\lambda_1}{2}\rfloor$, so the first condition in Lemma 7.1 is satisfied. The second condition in Lemma 7.1 is satisfied by (4.10) and the definition of solid triplets. Lemma 7.1 combined with Lemma 5.1, gives:

$$\delta(P_{\lambda/\mu}) \leq 2C_{d,\varepsilon}\,\frac{C^3+1}{\sqrt{n}}\,,$$

for some absolute constant $C_{d,\varepsilon} > 0$, as desired. $\quad\square$

8. Upper bounds for the number of standard Young tableaux

8.1. **Upper bound via Schur polynomials.** In this subsection we give an upper bound to $f(\lambda/\mu)$ in terms of $F(\lambda/\mu)$ (see (2.9)), and evaluations of Schur polynomial (see (2.3)). For $\lambda/\mu \vdash n$, recall the definition of shifted values ℓ_i and m_i (see (6.4)).

Lemma 8.1. *Let λ be a partition. Then, for every $(i,j) \in \lambda$, and every $k \geq 0$, we have:*

$$\frac{h_\lambda(i+k,j+k)}{h_\lambda(i,j)} \leq \frac{\ell_{i+k}}{\ell_i}\,.$$

Proof. We have:

$$h_\lambda(i+k,j+k) \leq h_\lambda(i,j) - 2k + \lambda_{i+k} - \lambda_i = h_\lambda(i,j) - k + \ell_{i+k} - \ell_i\,.$$

Note that $h_\lambda(i,j) = \lambda_i - i + \lambda'_j - j + 1 \leq \lambda_i + d - i$. Hence:

$$\frac{h_\lambda(i+k,j+k)}{h_\lambda(i,j)} \leq 1 - \frac{\ell_i - \ell_{i+k} + k}{h_\lambda(i,j)} \leq 1 - \frac{\ell_i - \ell_{i+k} + k}{\ell_i} \leq \frac{\ell_{i+k}}{\ell_i}\,,$$

as desired. $\qquad\square$

We now apply Lemma 8.1 to derive an upper bound for the product of hooks of a flagged tableau. Let $T \in \mathrm{FT}(\lambda/\mu)$. Recall the notation (2.4), for the number $t_i(T)$ of i's in T. Lemma 8.1 immediately gives:

Corollary 8.2. *Let $d \geq 2$, and let T be a flagged tableau of λ/μ. Then:*

$$\prod_{(i,j)\in\mu} \frac{h_\lambda\big(T(i,j),\, j+T(i,j)-i\big)}{h_\lambda(i,j)} \leq \prod_{i=1}^d \frac{(\ell_i)^{t_i(T)}}{(\ell_i)^{\mu_i}}\,.$$

We now arrive to the main result of this subsection.

Lemma 8.3. *Let $d \geq 2$, and let λ,μ be partitions such that $\mu \subseteq \lambda$. Then*

$$1 \leq \frac{f(\lambda/\mu)}{F(\lambda/\mu)} \leq \frac{s_\mu(\ell_1,\ldots,\ell_d)}{\ell_1^{\mu_1}\ldots\ell_d^{\mu_d}}\,.$$

Proof. The lower bound is given in Theorem 2.2. For the upper bound, we have:

$$\frac{f(\lambda/\mu)}{F(\lambda/\mu)} =_{\text{Thm } 2.3} \sum_{T\in\mathrm{FT}(\lambda/\mu)} \prod_{(i,j)\in\mu} \frac{h_\lambda\big(T(i,j),\, j+T(i,j)-i\big)}{h_\lambda(i,j)} \leq_{\text{Cor } 8.2} \sum_{T\in\mathrm{FT}(\lambda/\mu)} \prod_{i=1}^d \frac{\ell_i^{t_i(T)}}{\ell_i^{\mu_i}}$$

$$\leq_{(2.10)} \frac{1}{\ell_1^{\mu_1}\ldots\ell_d^{\mu_d}} \sum_{T\in\mathrm{SSYT}(\mu)} \ell_1^{t_1(T)}\ldots\ell_d^{t_d(T)} \leq_{(2.3)} \frac{s_\mu(\ell_1,\ldots,\ell_d)}{\ell_1^{\mu_1}\ldots\ell_d^{\mu_d}}\,,$$

as desired. $\qquad\square$

8.2. **Interval decomposition upper bound.** In this subsection we give a refinement to the upper bound in Lemma 8.3. An interval decomposition of $[d] = \{1, \ldots, d\}$ is defined as the following collection of subsets: $\mathcal{B} := (B_1, \ldots, B_r)$, where

$$(8.1) \qquad B_1 := \{1, \ldots, b_1\}, \quad B_2 := \{b_1 + 1, \ldots, b_2\}, \quad \ldots \quad , B_r := \{b_{r-1} + 1, \ldots, d\},$$

for some $0 = b_0 < b_1 < b_2 < \ldots < b_r = d$ and $r \geq 1$.

For all $i, j \in [d]$, we write $i \overset{\mathcal{B}}{\sim} j$ when i and j are contained in the same partition in $B_1, \ldots, B_r$, and $i \overset{\mathcal{B}}{\nsim} j$ otherwise. We drop $\mathcal{B}$ when the partition is clear. Let

$$(8.2) \qquad N(\ell, \mathcal{B}) := \max \left\{ \frac{\ell_i}{\ell_i - \ell_j} \ \Big| \ 1 \leqslant i < j \leqslant d \ \text{ and } \ i \overset{\mathcal{B}}{\sim} j \right\},$$

and let $N(\lambda, \mathcal{B}) := 0$ for $r = d$. The main result of this section is the following upper bound:

Theorem 8.4. *Fix $d \geq 2$. Let $\lambda/\mu \vdash n$, such that $\lambda, \mu \in \mathbb{P}_d$, and let $\mathcal{B}$ be an interval decomposition of $[d]$. Then:*

$$(8.3) \qquad \frac{s_\mu(\ell_1, \ldots, \ell_d)}{\ell_1^{\mu_1} \ldots \ell_d^{\mu_d}} \ \leq \ C_d \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \ \sim \ j}} (m_i - m_j + N(\ell, \mathcal{B})) \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \ \nsim \ j}} \frac{\ell_i}{\ell_i - \ell_j},$$

for some absolute constant $C_d > 0$.

Lemma 8.3 and Theorem 8.4 immediately imply:

Corollary 8.5 (Interval Upper Bound). *In notation of Theorem 8.4,*

$$\frac{f(\lambda/\mu)}{F(\lambda/\mu)} \ \leq \ C_d \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \ \sim \ j}} (m_i - m_j + N(\ell, \mathcal{B})) \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \ \nsim \ j}} \frac{\ell_i}{\ell_i - \ell_j}.$$

8.3. **Expanding the determinant.** We now build toward the proof of Theorem 8.4. Our strategy is to break down the Schur function evaluated at the sequence $\ell_1, \ldots, \ell_d$ into evaluations of separate parts, and use either (2.2) when the values of ℓ_i are sufficiently distinct, or (2.5) when they are close. Denote

$$(8.4) \qquad \mathrm{M} := \left(x_j^{m_i} \right)_{i,j=1}^d.$$

Lemma 8.6. *Fix $d \geq 2$. Let $\mu \in \mathbb{P}_d$, and let $x_1 \geq \ldots \geq x_d > 0$. Then:*

$$0 \leq \det \mathrm{M} \leq x_1^{m_1} \ldots x_d^{m_d} \prod_{1 \leqslant i < j \leqslant d} \frac{(m_i - m_j)(x_i - x_j)}{(j - i) x_i}.$$

Proof. The first inequality follows from (2.2):

$$(8.5) \qquad \det \mathrm{M} = s_\mu(x_1, \ldots, x_d) \prod_{1 \leqslant i < j \leqslant d} (x_i - x_j) = \sum_{A \in \mathrm{SSYT}(\mu)} x_1^{t_1(A)} \ldots x_d^{t_d(A)} \prod_{1 \leqslant i < j \leqslant d} (x_i - x_j) \geq 0.$$

For the second inequality, since $x_1 \geq \ldots \geq x_d$, and $\mu_1 \geq t_1(A)$, $\mu_1 + \mu_2 \geq t_1(A) + t_2(A)$, $\ldots$, we have:

$$(8.6) \qquad x_1^{t_1(A)} \ldots x_d^{t_d(A)} \leq x_1^{\mu_1} \ldots x_d^{\mu_d}.$$

We conclude:

$$\det \mathrm{M} \ \leq_{(8.5),\,(8.6)} \ \sum_{A \in \mathrm{SSYT}(\mu)} x_1^{\mu_1} \ldots x_d^{\mu_d} \prod_{1 \leqslant i < j \leqslant d} (x_i - x_j)$$

$$\leq_{(2.5)} \ x_1^{\mu_1} \ldots x_d^{\mu_d} \prod_{1 \leqslant i < j \leqslant d} \frac{(m_i - m_j)(x_i - x_j)}{(j - i)},$$

which implies the result by the definition (6.4). $\qquad \square$

To simplify presentation, we use notation $\mathrm{DET}(A) := |\det(A)|$.

Lemma 8.7. *Fix $d \geq 2$. Let $\mu \in \mathbb{P}_d$, $x_1 \geq \ldots \geq x_d > 0$, and let $\mathcal{B}$ be an interval decomposition of $[d]$. Then:*

$$(8.7) \qquad s_\mu(x_1, \ldots, x_d) \leq \sum_{\sigma \in S_d} \frac{x_1^{m_{\sigma(1)}} \ldots x_d^{m_{\sigma(d)}}}{x_1^{d-1} \ldots x_d^{d-d}} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \sim j}} \frac{|m_{\sigma(i)} - m_{\sigma(j)}|}{j - i} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \nsim j}} \frac{x_i}{x_i - x_j}.$$

Proof. Apply the Laplace expansion of M along the interval decomposition $\mathcal{B} = (B_1, \ldots, B_r)$, defined as in (8.1). We get:

$$\det \mathrm{M} = \sum_{\sigma \in S_d / \mathrm{Stab}(\mathcal{B})} \mathrm{sign}(\sigma) \prod_{k=1}^r \det \left[x_j^{m_{\sigma(i)}} \right]_{i,j \in B_k},$$

where $\mathrm{Stab}(\mathcal{B}) \subset S_d$ is the stabilizer subgroup of $\mathcal{B}$, so $\mathrm{Stab}(\mathcal{B}) \simeq S_{b_1} \times S_{b_2 - b_1} \times \ldots \times S_{d - b_{r-1}}$. We have:

$$(8.8) \qquad \det \mathrm{M} \leq \sum_{\sigma \in S_d} \prod_{k=1}^r \mathrm{DET} \left[x_j^{m_{\sigma(i)}} \right]_{i,j \in B_k}.$$

We now analyze each term in the right side of (8.8) separately. We have for every $\sigma \in S_d$ that

$$\prod_{k=1}^r \mathrm{DET} \left[x_j^{m_{\sigma(i)}} \right]_{i,j \in B_k} \leq_{\mathrm{Lem\ 8.6}} \prod_{k=1}^r \prod_{i \in B_k} x_i^{m_{\sigma(i)}} \prod_{\substack{j \in B_k \\ j > i}} \frac{|m_{\sigma(i)} - m_{\sigma(j)}| \cdot (x_i - x_j)}{x_i (j - i)}$$

$$\leq x_1^{m_{\sigma(1)}} \cdots x_d^{m_{\sigma(d)}} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \sim j}} \frac{|m_{\sigma(i)} - m_{\sigma(j)}| \cdot (x_i - x_j)}{x_i (j - i)} .$$

Using the inequality above for the RHS of (8.8), we obtain:

$$\det \mathrm{M} \leq \sum_{\sigma \in S_d} x_1^{m_{\sigma(1)}} \cdots x_d^{m_{\sigma(d)}} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \sim j}} \frac{|m_{\sigma(i)} - m_{\sigma(j)}| \cdot (x_i - x_j)}{x_i (j - i)}$$

We conclude:

$$s_\mu(x_1, \ldots, x_d) =_{(2.2)} \det \mathrm{M} \prod_{1 \leqslant i < j \leqslant d} \frac{1}{x_i - x_j}$$

$$\leq \prod_{1 \leqslant i < j \leqslant d} \frac{1}{x_i - x_j} \sum_{\sigma \in S_d} x_1^{m_{\sigma(1)}} \cdots x_d^{m_{\sigma(d)}} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \sim j}} \frac{|m_{\sigma(i)} - m_{\sigma(j)}| \cdot (x_i - x_j)}{x_i (j - i)}$$

$$\leq \prod_{1 \leqslant i < j \leqslant d} \frac{x_i}{x_i - x_j} \sum_{\sigma \in S_d} \frac{x_1^{m_{\sigma(1)}} \cdots x_d^{m_{\sigma(d)}}}{x_1^{d-1} \cdots x_d^{d-d}} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \sim j}} \frac{|m_{\sigma(i)} - m_{\sigma(j)}| \cdot (x_i - x_j)}{x_i (j - i)}$$

$$\leq \sum_{\sigma \in S_d} \frac{x_1^{m_{\sigma(1)}} \cdots x_d^{m_{\sigma(d)}}}{x_1^{d-1} \cdots x_d^{d-d}} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \sim j}} \frac{|m_{\sigma(i)} - m_{\sigma(j)}|}{j - i} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \nsim j}} \frac{x_i}{x_i - x_j},$$

as desired. $\qquad \square$

8.4. **Simplifying the products.** In this subsection we will simplify the upper bound in Lemma 8.7. Our goal is to remove the dependence to $\sigma \in S_d$ in the RHS of (8.7). We start with the following two technical lemmas.

Let $\sigma \in S_d$, and let $1 \leq a < b \leq d$. We say that (a, b) is an *inversion* in σ, if $\sigma(a) > \sigma(b)$. Denote by (ab) a transposition in S_d.

Lemma 8.8. *Fix $m_1 \geq \ldots \geq m_d > 0$ and $x_1 \geq \ldots \geq x_d > 0$. Let $\sigma \in S_d$, and let $\tau = (ab) \in S_d$. Then:*

$$(8.9) \qquad \prod_{i=1}^d \frac{x_i^{m_{\sigma(i)}}}{x_i^{m_{\sigma\tau(i)}}} = \left(\frac{x_b}{x_a} \right)^{m_{\sigma(b)} - m_{\sigma(a)}} .$$

Furthermore, when (a, b) is an inversion of σ, we have

$$(8.10) \qquad \left(\frac{x_b}{x_a}\right)^{m_{\sigma(b)} - m_{\sigma(a)}} \leq 1.$$

Both claims are straightforward; we omit the proof.

Lemma 8.9. *Fix $x_1 \geq x_2 > 0$, and let $m \geq 0$. Then:*

$$m \left(\frac{x_2}{x_1}\right)^m \leq \frac{x_1}{e\,(x_1 - x_2)}.$$

Proof. Substitute $y = \frac{x_2}{x_1}$, and note that the function $m(1 - y)y^m$ achieves maximum at $y = (m-1)/m$, which $\to 1/e$ from below as $m \to \infty$. $\qquad\square$

Let $\sigma \in S_d$. For all $a = 1, \ldots, d$, define permutations σ_a and $\tau_a \in S_d$ recursively:

$$(8.11) \qquad \sigma_a := \sigma\tau_1 \ldots \tau_{a-1} \qquad \text{and} \qquad \tau_a := \begin{cases} 1 & \text{if } \sigma_a(a) = a, \\ \left(a\,\sigma_a^{-1}(a)\right) & \text{if } \sigma_a(a) \neq a. \end{cases}$$

In other words, at each step a, we modify the permutation σ_a so that the resulting permutation σ_{a+1} has a as a fixed point, by switching a and $\sigma_a^{-1}(a)$ if necessary. It follows from the construction that, at each step, either $a = \sigma_a^{-1}(a)$ or $(a, \sigma_a^{-1}(a))$ is an inversion of σ_a. Observe that $\sigma_1 = \sigma$ and $\sigma_d = 1$.

Denote by R_a the number

$$(8.12) \qquad R_a := \left(\frac{x_b}{x_a}\right)^c, \quad \text{where } b = \sigma_a^{-1}(a) \quad \text{and} \quad c = \frac{2}{d(d-1)}\left(m_{\sigma_a(b)} - m_{\sigma_a(a)}\right).$$

It follows from Lemma 8.8, that

$$(8.13) \qquad\qquad\qquad\qquad R_a \leq 1.$$

Indeed, either we have $a = b$, or by construction (8.11) we have (a, b) is an inversion in σ_a.

Let $\mathcal{B}$ be an interval decomposition of $[d]$. Recall from definition (8.2) that

$$(8.14) \qquad N(x, \mathcal{B}) := \max\left\{\frac{x_i}{x_i - x_j} \ \Big| \ 1 \leqslant i < j \leqslant d \text{ and } i \nsim j\right\} > 0.$$

For all $\ 1 \leq i < j \leq d$, denote

$$(8.15) \qquad H_a(i,j) := \left|m_{\sigma_a(i)} - m_{\sigma_a(j)}\right| + (a-1)\frac{d(d-1)}{e}\,N(x, \mathcal{B}).$$

It follows from the definition (8.11), that H_a satisfies

$$(8.16) \qquad H_{a+1}(i,j) = H_a(\tau_a(i), \tau_a(j)) + \frac{d(d-1)}{e}\,N(x, \mathcal{B}).$$

Note also that

$$H_a(i,j) = H_a(j,i) \quad \text{for all } 1 \leq i, j \leq d.$$

The following two lemmas utilize and clarify the properties of numbers R_a and $H_a(i,j)$ defined above. The idea is that we can now rewrite the RHS of (8.7) as

$$x_1^{\mu_1} \ldots x_d^{\mu_d} \sum_{\sigma \in S_d} \prod_{a=1}^{d-1} R_a^{\frac{d(d-1)}{2}} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \sim j}} \frac{H_1(i,j)}{j-i} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \nsim j}} \frac{x_i}{x_i - x_j}.$$

Our goal is to iteratively replace all $H_1(i,j)$'s (which depend on σ) with $H_d(i,j)$'s (which do not depend on σ), and R_a's will be the cost that we are paying for each iteration.

Lemma 8.10 (The same block estimate)**.** *Fix $m_1 \geq \ldots \geq m_d > 0$ and $x_1 \geq \ldots \geq x_d > 0$. Let $\mathcal{B}$ be an interval decomposition of $[d]$, and let $\sigma \in S_d$. Then, for all $1 \leq a \leq d-1$, such that $a \overset{\mathcal{B}}{\sim} \sigma_a^{-1}(a)$, we have:*

$$\prod_{\substack{1 \leqslant i < j \leqslant d \\ i \sim j}} H_a(i,j) \leq \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \sim j}} H_{a+1}(i,j).$$

Proof. It follows from (8.16) that

$$(8.17) \qquad \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \sim \, j}} H_{a+1}(i,j) \; \geq \; \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \sim \, j}} H_a\big(\tau_a(i), \tau_a(j)\big).$$

Note that the RHS can be rewritten as

$$(8.18) \qquad \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \sim \, j}} H_a\big(\tau_a(i), \tau_a(j)\big) \; = \; \prod_{k=1}^{r} \prod_{\substack{1 \leqslant i < j \leqslant d, \\ i,j \in B_k}} H_a\big(\tau_a(i), \tau_a(j)\big) \; = \; \prod_{k=1}^{r} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i,j \in \tau_a^{-1}(B_k)}} H_a(i,j) \,.$$

Now note that a and b are contained in the same block in $\mathcal{B}$, since $a \overset{\mathcal{B}}{\sim} b$ by assumption. Since $\tau_a = (ab)$, this implies that $\tau_a(B_k) = B_k$ for all $1 \leq k \leq r$. Thus, we have:

$$(8.19) \qquad \prod_{k=1}^{r} \prod_{\substack{1 \leqslant i < j \leqslant d, \\ i,j \in \tau_a^{-1}(B_k)}} H_a(i,j) \; = \; \prod_{k=1}^{r} \prod_{\substack{1 \leqslant i < j \leqslant d, \\ i,j \in B_k}} H_a(i,j) \; = \; \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \sim \, j}} H_a(i,j).$$

The lemma now follows by combining (8.17), (8.18), and (8.19). $\qquad \square$

Lemma 8.11 (The distinct blocks estimate)**.** *Fix $m_1 \geq \ldots \geq m_d > 0$ and $x_1 \geq \ldots \geq x_d > 0$. Let $\mathcal{B}$ be an interval decomposition of $[d]$, and let $\sigma \in S_d$. Then, for all $1 \leq a \leq d-1$, such that $a \overset{\mathcal{B}}{\not\sim} \sigma_a^{-1}(a)$, and all $1 \leqslant i < j \leqslant d$, we have:*

$$(8.20) \qquad R_a \, H_a(i,j) \; \leq \; H_{a+1}(i,j).$$

Proof. We first prove the following bound:

$$(8.21) \qquad R_a \big| H_a(\tau_a(i), j) - H_a(i,j) \big| \; \leq \; \frac{d(d-1)}{2e} \, N(x, \mathcal{B}),$$

all $1 \leqslant i < j \leqslant d$.

Let $b := \sigma_a^{-1}(a)$ and suppose that $i \notin \{a,b\}$. Then $\tau_a(i) = i$ by the definition (8.11). It then follows that the LHS of (8.21) is equal to 0. Suppose now that $i \in \{a,b\}$. Equation (8.21) then becomes

$$R_a \big| H_a(b,j) - H_a(a,j) \big| \; \leq \; \frac{d(d-1)}{2e} \, N(x, \mathcal{B}).$$

Note that

$$\big| H_a(b,j) - H_a(a,j) \big| =_{(8.15)} \; \big| \, |m_{\sigma_a(b)} - m_{\sigma_a(j)}| - |m_{\sigma_a(j)} - m_{\sigma_a(a)}| \, \big| \; \leq \; \big| m_{\sigma_a(b)} - m_{\sigma_a(a)} \big|.$$

This implies that

$$R_a \big| H_a(b,j) - H_a(a,j) \big| \; \leq \; R_a \, \big| m_{\sigma_a(b)} - m_{\sigma_a(a)} \big| \; =_{(8.12)} \; \left(\frac{x_b}{x_a} \right)^c \big| m_{\sigma_a(b)} - m_{\sigma_a(a)} \big|,$$

where c is also defined in (8.12). Now note that (a,b) is an inversion of σ_a by construction (8.11). Apply Lemma 8.9 with $x_1 \leftarrow x_a$, $x_2 \leftarrow x_b$ and $m \leftarrow c$, to get

$$\left(\frac{x_b}{x_a} \right)^c \big| m_{\sigma_a(b)} - m_{\sigma_a(a)} \big| \; \leq \; \frac{d(d-1)}{2e} \, \frac{x_a}{x_a - x_b} \,.$$

Since $a \not\sim b$ and $a < b$ by the construction (8.11) of σ_a, we have $\frac{x_a}{x_a - x_b} \leq N(x, \mathcal{B})$ by (8.14), and the inequality (8.21) follows.

Therefore, we have:

$$H_{a+1}(i,j) - R_a \, H_a(i,j) \; =_{(8.16)} \; H_a\big(\tau_a(i), \tau_a(j)\big) + \frac{d(d-1)}{e} N(x, \mathcal{B}) - R_a H_a(i,j)$$

$$\geq_{(8.13)} \; R_a H_a\big(\tau_a(i), \tau_a(j)\big) + \frac{d(d-1)}{e} \, N(x, \mathcal{B}) - R_a H_a(i,j)$$

$$\geq \; R_a H_a\big(\tau_a(i), \tau_a(j)\big) - R_a H_a\big(\tau_a(i), j\big) + R_a H_a\big(\tau_a(i), j\big) - R_a H_a(i,j) + \frac{d(d-1)}{e} \, N(x, \mathcal{B})$$

$$\geq_{(8.21)} \; -\frac{d(d-1)}{2e} \, N(x, \mathcal{B}) - \frac{d(d-1)}{2e} \, N(x, \mathcal{B}) + \frac{d(d-1)}{e} \, N(x, \mathcal{B}) \; = \; 0.$$

This proves the lemma. $\qquad \square$

8.5. **Putting everything together.** We now combine Lemma 8.10 and Lemma 8.11 to get the following upper bound.

Lemma 8.12. *Fix $m_1 \geq \ldots \geq m_d > 0$ and $x_1 \geq \ldots \geq x_d > 0$. Let $\mathcal{B}$ be an interval decomposition of $[d]$, and let $\sigma \in S_d$. Then:*

$$(8.22) \qquad \frac{x_1^{m_{\sigma(1)}} \ldots x_d^{m_{\sigma(d)}}}{x_1^{m_1} \ldots x_d^{m_d}} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \sim \, j}} |m_{\sigma(i)} - m_{\sigma(j)}| \; \leq \; \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \sim \, j}} \left(m_i - m_j + \frac{d(d-1)^2}{e} N(x, \mathcal{B}) \right).$$

Proof. We have:

$$\frac{x_1^{m_{\sigma(1)}} \ldots x_d^{m_{\sigma(d)}}}{x_1^{m_1} \ldots x_d^{m_d}} \; =_{(8.9)} \; \prod_{a=1}^{d-1} \left(\frac{x_b}{x_a} \right)^{m_{\sigma_a(b)} - m_{\sigma_a(a)}} \; =_{(8.12)} \; \prod_{a=1}^{d-1} R_a^{\frac{d(d-1)}{2}}.$$

We can rewrite the inequality (8.22) in the lemma using the definition (8.15) as follows:

$$(8.23) \qquad \prod_{a=1}^{d-1} R_a^{\frac{d(d-1)}{2}} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \sim \, j}} H_1(i,j) \; \leq \; \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \sim \, j}} H_d(i,j).$$

First, note that the LHS of (8.23) is bounded from above by

$$\prod_{a=1}^{d-1} R_a^{\frac{d(d-1)}{2}} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \sim \, j}} H_1(i,j) \; \leq_{(8.13)} \; \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \sim \, j}} \left(H_1(i,j) \prod_{a=1}^{d-1} R_a \right).$$

Hence it suffices to show that

$$(8.24) \qquad \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \sim \, j}} \left(H_1(i,j) \prod_{a=1}^{d-1} R_a \right) \; \leq \; \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \sim \, j}} H_d(i,j).$$

First, for $a \sim \sigma_a^{-1}(a)$, we have:

$$\prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \sim \, j}} R_a H_a(i,j) \; \leq_{(8.13)} \; \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \sim \, j}} H_a(i,j) \; \leq_{\text{Lem 8.10}} \; \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \sim \, j}} H_{a+1}(i,j).$$

Otherwise, for $a \not\sim \sigma_a^{-1}(a)$, we have:

$$\prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \sim \, j}} R_a H_a(i,j) \; \leq_{(8.20)} \; \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \sim \, j}} H_{a+1}(i,j),$$

so we have the same inequality in both cases. Now (8.24) follows by induction on $a \in \{1, \ldots, d-1\}$. This completes the proof of the lemma. $\qquad\square$

Proof of Theorem 8.4. We have:

$$\frac{s_\mu(x_1, \ldots, x_d)}{x_1^{\mu_1} \ldots x_d^{\mu_d}} \; \leq_{\text{Lem 8.7}} \; \sum_{\sigma \in S_d} \frac{x_1^{m_{\sigma(1)}} \ldots x_d^{m_{\sigma(d)}}}{x_1^{m_1} \ldots x_d^{m_d}} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \sim \, j}} \frac{|m_{\sigma(i)} - m_{\sigma(j)}|}{j - i} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \not\sim \, j}} \frac{x_i}{x_i - x_j}$$

$$\leq \; \sum_{\sigma \in S_d} \frac{x_1^{m_{\sigma(1)}} \ldots x_d^{m_{\sigma(d)}}}{x_1^{m_1} \ldots x_d^{m_d}} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \sim \, j}} |m_{\sigma(i)} - m_{\sigma(j)}| \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \not\sim \, j}} \frac{x_i}{x_i - x_j}$$

$$\leq_{\text{Lem 8.12}} \; \sum_{\sigma \in S_d} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \sim \, j}} \left(m_i - m_j + \frac{d(d-1)^2}{e} N(\ell, \mathcal{B}) \right) \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \not\sim \, j}} \frac{x_i}{x_i - x_j}$$

$$\leq \; C_d \cdot \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \sim \, j}} \left(m_i - m_j + N(\ell, \mathcal{B}) \right) \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \, \not\sim \, j}} \frac{x_i}{x_i - x_j},$$

where $C_d := d! \max\{1, \frac{d(d-1)^2}{e}\}$. This completes the proof. $\qquad \square$

9. The case of thick Young diagrams

In this section we discuss the sorting probability for ε-thick Young diagrams and present the proof of Theorem 1.3.

9.1. **Using special interval decompositions.** Fix $\varepsilon > 0$ and $\mu = (0, \ldots, 0)$, so $\lambda/\mu = \lambda$. Throughout the section we assume that $\lambda \vdash n$ and λ is ε-thick. This assumption implies that (λ, μ) is ε-admissible, since $\lambda_i - \mu_i = \lambda_i \geq \lambda_d \geq \varepsilon n$.

For the rest of this section, let $\mathcal{B}$ be the interval decomposition of $[d] = \{1, \ldots, d\}$ that places $i, j \in [d]$, $i < j$, in the same block if and only if

$$(9.1) \qquad \lambda_i - \lambda_j + j - i \leq \sqrt{n}.$$

Lemma 9.1. *Fix $d \geq 2$. Let $\lambda \vdash n$, $\lambda, \gamma \in \mathbb{P}_d$, $\gamma \subseteq \lambda$, and let $\mathcal{B}$ as in* (9.1). *Then:*

$$\frac{f(\lambda/\gamma)}{F(\lambda/\gamma)} \ \leq \ C_d \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \overset{\mathcal{B}}{\sim} j}} (\gamma_i - \gamma_j + j - i + \sqrt{n}) \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \overset{\mathcal{B}}{\not\sim} j}} \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i} .$$

for some absolute constant $C_d > 0$.

Proof. It follows from Corollary 8.5, by substituting μ with γ, that

$$\frac{f(\lambda/\gamma)}{F(\lambda/\gamma)} \leq C_d \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \sim j}} (\gamma_i - \gamma_j + j - i + N(\ell, \mathcal{B})) \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \not\sim j}} \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i} ,$$

where $N(\ell, \mathcal{B})$ is as defined in (8.2), and $C_d > 0$ is an absolute constant. Note that

$$N(\ell, \mathcal{B}) = \max_{\substack{1 \leqslant i < j \leqslant d \\ i \not\sim j}} \left\{ \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i} \right\} \leq_{(9.1)} \frac{\lambda_1 + d - 1}{\sqrt{n}} \leq \frac{dn}{\sqrt{n}} = d\sqrt{n}.$$

We conclude:

$$\frac{f(\lambda/\gamma)}{F(\lambda/\gamma)} \leq C_d \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \sim j}} (\gamma_i - \gamma_j + j - i + d\sqrt{n}) \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \not\sim j}} \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i}$$

$$\leq d^{\frac{d(d-1)}{2}} C_d \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \sim j}} (\gamma_i - \gamma_j + j - i + \sqrt{n}) \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \not\sim j}} \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i},$$

which proves the lemma. $\qquad \square$

We now derive upper bounds for each term in the right side of Lemma 9.1. We collect these upper bounds in the following two lemmas.

Lemma 9.2 (Same blocks estimate). *Fix $d \geq 2$. Let $\lambda \vdash n$, $\lambda, \gamma \in \mathbb{P}_d$, $\gamma \subseteq \lambda$, and let $\mathcal{B}$ as in* (9.1). *Then, for all $1 \leq i < j \leq d$ satisfying $i \overset{\mathcal{B}}{\not\sim} j$, we have:*

$$\frac{\gamma_i - \gamma_j + j - i}{n} \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i} \leq d\left(|y_i - y_j| + 1\right),$$

where y_i are defined in (6.3).

Proof. We have

$$\frac{\gamma_i - \gamma_j + j - i}{n} \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i} \leq \frac{\gamma_i - \gamma_j + j - i}{n} \frac{dn}{\lambda_i - \lambda_j + j - i} = d\frac{\gamma_i - \gamma_j + j - i}{\lambda_i - \lambda_j + j - i}.$$

Note that

$$(9.2) \qquad \gamma_i - \gamma_j + j - i =_{(4.9)} \sqrt{n}(y_i - y_j) + p(\lambda_i - \lambda_j) + j - i \leq \sqrt{n}(y_i - y_j) + \lambda_i - \lambda_j + j - i.$$

Since $i \nsim j$, we have $\lambda_i - \lambda_j > \sqrt{n}$ by (9.1). Therefore:

$$\frac{\gamma_i - \gamma_j + j - i}{\lambda_i - \lambda_j + j - i} \leq 1 + \frac{\sqrt{n}\,(y_i - y_j)}{\sqrt{n}}\,.$$

Combining the inequalities implies the result. $\square$

Lemma 9.3 (Distinct blocks estimate). *Fix $d \geq 2$. Let $\lambda \vdash n$, $\lambda, \gamma \in \mathbb{P}_d$, $\gamma \subseteq \lambda$, and let $\mathcal{B}$ as in (9.1). Then, for all $1 \leq i < j \leq d$ satisfying $i \overset{\mathcal{B}}{\sim} j$, we have:*

$$\frac{\gamma_i - \gamma_j + j - i}{n}\left(\gamma_i - \gamma_j + j - i + \sqrt{n}\right) \leq 2\big(|y_i - y_j| + 1\big)^2.$$

Proof. By the same argument as in (9.2), we have

$$\gamma_i - \gamma_j + j - i \ \leq \ \sqrt{n}\,(y_i - y_j) + \lambda_i - \lambda_j + j - i.$$

Since $i \sim j$, it then follows that

$$\gamma_i - \gamma_j + j - i \leq \sqrt{n}\,(y_i - y_j) + \sqrt{n} \leq \sqrt{n}\,(y_i - y_j + 1).$$

This then implies that

$$\frac{\gamma_i - \gamma_j + j - i}{n}\left(\gamma_i - \gamma_j + j - i + \sqrt{n}\right) \leq \frac{\sqrt{n}\,(y_i - y_j + 1)}{n}\,\sqrt{n}\,(y_i - y_j + 2) \leq 2\,(|y_i - y_j| + 1)^2,$$

which completes the proof. $\square$

9.2. Lattice paths.

The main ingredient in the proof of Theorem 1.3 is the following lemma, a direct analogue for straight shapes of Lemma 6.7. Recall the definition of random integer paths $\boldsymbol{Z} = (Z_0, \ldots, Z_n)$ in §5.1, and the definition of y_i in (6.3).

Lemma 9.4. *Fix $d \geq 2$ and $\varepsilon > 0$. Let $\lambda \in \mathbb{P}_d$, such that λ is ε-thick. Then, for every $(\lambda, \gamma, \varnothing) \in \Lambda(n, d, \varepsilon)$, $\gamma \vdash k$, we have:*

$$\mathbf{P}_{\lambda/\mu}\big[Z_k = \gamma\big] \leq C_{d,\varepsilon}\, n^{-\frac{(d-1)}{2}} \exp\left[-2\sum_{i=1}^{d} y_i^2\right] \prod_{1 \leq i < j \leq d} \big((y_i - y_j)^2 + 1\big),$$

for some absolute constant $C_{d,\varepsilon} > 0$.

Proof. Following the proof of Lemma 6.7, we have:

$$(9.3) \qquad \mathbf{P}_{\lambda/\mu}\big[Z_k = \gamma\big] =_{(6.17)} \frac{f(\gamma)\,f(\lambda/\gamma)}{f(\lambda)} \ = \ f(\gamma)\,\frac{F(\lambda/\gamma)}{f(\lambda)}\,\frac{f(\lambda/\gamma)}{F(\lambda/\gamma)}.$$

We give a separate bound for each of the three terms in the RHS of (9.3).

First note that,

$$f(\gamma) =_{(2.1)} \frac{k!}{\gamma_1! \ldots \gamma_d!} \prod_{1 \leq i < j \leq d} \frac{\gamma_i - \gamma_j + j - i}{\gamma_i + j - i} \leq_{(4.7)} \frac{k!}{\gamma_1! \ldots \gamma_d!} \prod_{1 \leq i < j \leq d} \frac{\gamma_i - \gamma_j + j - i}{\frac{\varepsilon^3}{2}n}$$

$$\leq \left(\frac{\varepsilon^3}{2}\right)^{-\frac{d(d-1)}{2}} \frac{k!}{\gamma_1! \ldots \gamma_d!} \prod_{1 \leq i < j \leq d} \frac{\gamma_i - \gamma_j + j - i}{n}\,.$$

Note also that

$$\frac{F(\lambda/\gamma)}{f(\lambda)} =_{(2.9),\,(2.7)} \frac{(n-k)!}{n!} \prod_{(i,j) \in \gamma} h_\lambda(i,j) \leq_{(4.7),\,\text{Lem } 6.1} \left(\frac{\varepsilon^3}{2}\right)^{-\frac{d(d-1)}{2}} \frac{(n-k)!}{n!} \prod_{i=1}^{d} \frac{\lambda_i!}{(\lambda_i - \gamma_i)!}\,.$$

Finally, by Lemma 9.1 we have:

$$\frac{f(\lambda/\gamma)}{F(\lambda/\gamma)} \ \leq \ C_d \prod_{\substack{1 \leq i < j \leq d \\ i \overset{\mathcal{B}}{\sim} j}} \big(\gamma_i - \gamma_j + j - i + \sqrt{n}\big) \prod_{\substack{1 \leq i < j \leq d \\ i \overset{\mathcal{B}}{\nsim} j}} \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i}\,,$$

for some absolute constant $C_d > 0$, where $\mathcal{B}$ is defined in (9.1).

Substituting the above three estimates in (9.3), we obtain:

$$\mathbf{P}_{\lambda/\mu}\left[Z_k = \gamma\right] \le C_d \left(\frac{\varepsilon^6}{4}\right)^{-\frac{d(d-1)}{2}} \binom{\lambda_1}{\gamma_1} \cdots \binom{\lambda_d}{\gamma_d} \binom{n}{k}^{-1} \times$$

$$\times \prod_{\substack{1\leqslant i<j\leqslant d \\ i \,\sim\, j}} \frac{\gamma_i - \gamma_j + j - i}{n}\left(\gamma_i - \gamma_j + j - i + \sqrt{n}\right) \prod_{\substack{1\leqslant i<j\leqslant d \\ i \,\nsim\, j}} \frac{\gamma_i - \gamma_j + j - i}{n} \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i}\,.$$

By Lemmas 9.2 and 9.3, the last two products are bounded by

$$\prod_{\substack{1\leqslant i<j\leqslant d \\ i \,\sim\, j}} 2\left(|y_i - y_j| + 1\right)^2 \prod_{\substack{1\leqslant i<j\leqslant d \\ i \,\nsim\, j}} d\left(|y_i - y_j| + 1\right) \le \left[\left(2(d+2)\right]^{-\frac{d(d-1)}{2}} \prod_{1\leqslant i<j\leqslant d} \left((y_i - y_j)^2 + 1\right).$$

On the other hand, Lemma 6.3 gives

$$\binom{\lambda_1}{\gamma_1} \cdots \binom{\lambda_d}{\gamma_d} \binom{n}{k}^{-1} \le C_{d,\varepsilon}\, n^{-\frac{(d-1)}{2}}\, \exp\left[-2\sum_{i=1}^d y_i^2\right]$$

for some absolute constant $C_{d,\varepsilon} > 0$. Combining the last three inequalities, we conclude:

$$\mathbf{P}_{\lambda/\mu}\left[Z_k = \gamma\right] \le C'_{d,\varepsilon}\, n^{-\frac{(d-1)}{2}}\, \exp\left[-2\sum_{i=1}^d y_i^2\right] \prod_{1\leqslant i<j\leqslant d}\left((y_i - y_j)^2 + 1\right),$$

for some absolute constant $C'_{d,\varepsilon} > 0$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

9.3. **Proof of Theorem 1.3.** We follow the proof of Main Lemma 4.3 in §7.2. Let $a = \lfloor\frac{\lambda_1}{2}\rfloor$. Then the first condition in Lemma 7.1 is satisfied. Also note that the second condition in Lemma 7.1 is satisfied as a consequence of Lemma 9.4. We conclude:

$$\delta(P_{\lambda/\mu}) \le_{\text{Lem 5.1}} 2\varphi(a) \le_{\text{Lem 7.1, Lem 9.4}} 2C'_{d,\varepsilon}\frac{C^3_{d,\varepsilon} + 1}{\sqrt{n}},$$

for some absolute constants $C_{d,\varepsilon}, C'_{d,\varepsilon} > 0$. This completes the proof. $\square$

10. The case of smooth skew Young diagrams

In this section we prove Theorem 1.5.

10.1. **Proof of Lemma 4.4.** Let $\lambda/\mu \in \mathbb{P}_d$, $\lambda/\mu \vdash n$, and suppose λ is ε-smooth. Then we have:

$$(10.1) \qquad\qquad \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i} \le \frac{n + d - 1}{\varepsilon\, n} \le \frac{d}{\varepsilon}, \quad \text{for all} \quad 1 \le i < j \le d.$$

Therefore,

$$(10.2) \qquad\qquad 1 \le \min\left\{\mu_i - \mu_j + j - i, \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i}\right\} \le \frac{d}{\varepsilon}\,.$$

By the definition of $\Phi(\lambda/\mu)$, see (4.1), we get:

$$(10.3) \qquad\qquad\qquad 1 \le \Phi(\lambda/\mu) \le \left(\frac{d}{\varepsilon}\right)^{\frac{d(d-1)}{2}}.$$

i.e., function $\Phi(\lambda/\mu)$ is of a constant order. Therefore, the result follows from the following bounds:

$$(10.4) \qquad\qquad\qquad 1 \le \frac{f(\lambda/\mu)}{F(\lambda/\mu)} \le d!\left(\frac{d}{\varepsilon}\right)^{\frac{d(d-1)}{2}}.$$

The lower bound in (10.4) follows from Theorem 2.2. For the upper bound in 10.4, we use Corollary 8.5 applied to the interval decomposition $\mathcal{B} := \{B_1, \ldots, B_d\}$, where $B_i = \{i\}$. In this case Corollary 8.5 gives:

$$\frac{f(\lambda/\mu)}{F(\lambda/\mu)} \le d! \prod_{1\leqslant i<j\leqslant d} \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i} \le d!\left(\frac{d}{\varepsilon}\right)^{\frac{d(d-1)}{2}},$$

which proves the upper bound in (10.4). $\square$

10.2. **Proof of Theorem 1.5.** By Lemma 4.3, it suffices to check that for every $(\lambda, \gamma, \mu) \in \Omega(n, d, \varepsilon)$, we have

$$(10.5) \qquad \frac{f(\gamma/\mu)}{F(\gamma/\mu)} \leq C_{d,\varepsilon} \, \Phi(\gamma/\mu), \quad \frac{f(\lambda/\gamma)}{F(\lambda/\gamma)} \leq C_{d,\varepsilon} \, \Phi(\lambda/\gamma) \quad \text{and} \quad \frac{f(\lambda/\mu)}{F(\lambda/\mu)} \geq \frac{1}{C_{d,\varepsilon}} \, \Phi(\lambda/\mu),$$

for some absolute constant $C_{d,\varepsilon} > 0$. Note that the last two inequalities follow immediately from Lemma 4.4. We now prove that the first inequality holds.

By the progressive assumption on $(\lambda, \gamma, \mu) \in \Lambda(n, d, \varepsilon)$, we have:

$$\gamma_i - \gamma_{i+1} \geq_{(4.8)} \; p \, (\lambda_i - \lambda_{i+1}) \, + \, (1-p)(\mu_i - \mu_{i+1}) \, - \, 2n^{\frac{3}{4}} \geq p \, (\lambda_i - \lambda_{i+1}) \, - \, 2n^{\frac{3}{4}} \geq p \, \varepsilon \, |\lambda| \, - \, 2n^{\frac{3}{4}},$$

for every $1 \leq i \leq d-1$. Similarly, by the ε-separation assumption on (λ, γ, μ), we have:

$$p \; = \; \frac{|\gamma| - |\mu|}{n} \; = \; \sum_{i=1}^{d} \frac{\gamma_i - \mu_i}{n} \; \geq_{(4.7)} \; \left(\frac{d\varepsilon^3}{2} \right) \frac{|\lambda|}{n} \; \geq \; \frac{d\varepsilon^3}{2}.$$

Thus, for sufficiently large n, we have:

$$(10.6) \qquad \gamma_i - \gamma_{i+1} \geq \frac{d\varepsilon^4}{2} |\lambda| \, - \, 2n^{\frac{3}{4}} \geq \frac{d\varepsilon^4}{4} \, |\lambda| \geq \frac{d\varepsilon^4}{4} \, |\gamma|.$$

By the same argument as above, for sufficiently large n, we have:

$$(10.7) \qquad \gamma_d \geq \frac{d\varepsilon^4}{4} \, |\gamma|.$$

Conditions (10.6) and (10.7) imply that γ/μ is $(d\varepsilon^4/4)$-smooth, for n large enough. Applying Lemma 4.4, we obtain the first inequality in (10.5). This completes the proof. $\square$

11. The case of TVK skew shapes

In this section we give upper and lower bounds for the number of standard Young tableaux corresponding to TVK pairs. We then prove Lemma 4.5 and Theorem 1.4.

11.1. **Conditions for interval decomposition.** We define three types of conditions for interval decomposition $\mathcal{B} = (B_1, \ldots, B_r)$ of $[d]$. These conditions will be used in combinations, to cover all possible TVK pairs. Formally, consider:

$$(11.1) \qquad \lambda_i - \lambda_j \geq \varepsilon |\lambda| \quad \text{for all} \quad i \overset{\mathcal{B}}{\not\sim} j, \quad 1 \leq i < j \leq d,$$

$$(11.2) \qquad \lambda_i - \lambda_j \leq 1 \quad \text{for all} \quad i \overset{\mathcal{B}}{\sim} j, \quad 1 \leq i < j \leq d,$$

$$(11.3) \qquad \mu_i - \mu_j \leq 1 \quad \text{for all} \quad i \overset{\mathcal{B}}{\sim} j, \quad 1 \leq i < j \leq d.$$

The motivation behind these conditions for TVK (α, β)-shapes will become apparent later in this section. For now, we treat them as abstract constraints on the interval decompositions,

11.2. **Upper bounds.** We start with estimating each term in the definition of $\Phi(\lambda/\mu)$, see (4.1), and we collect these estimates in the next three lemmas.

Lemma 11.1. *Fix $d \geq 2$ and $\varepsilon > 0$. Let $\lambda, \mu \in \mathbb{P}_d$, such that $\mu \subseteq \lambda$. Suppose (11.1) holds for the interval decomposition $\mathcal{B}$ of $[d]$. Then:*

$$(11.4) \qquad 1 \leq \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i} \leq \frac{d}{\varepsilon}, \quad \text{for all} \quad i \overset{\mathcal{B}}{\not\sim} j, \quad 1 \leq i < j \leq d.$$

In particular, we have:

$$(11.5) \qquad \frac{\varepsilon}{d} \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i} \leq \min \left\{ \mu_i - \mu_j + j - i, \; \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i} \right\} \leq \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i}.$$

Proof. The lower bound in (11.4) follows from (4.2). The upper bound in (11.4), follows verbatim (10.1). For the lower bound in (11.5), we have

$$\min\left\{\mu_i - \mu_j + j - i, \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i}\right\} \geq_{(4.2),\,(4.3)} 1 \geq_{(11.4)} \left(\frac{\varepsilon}{d}\right)\frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i},$$

as desired. $\qquad\qquad\square$

Lemma 11.2. *Fix $d \geq 2$ and $\varepsilon > 0$. Let $\lambda, \mu \in \mathbb{P}_d$, such that $\mu \subseteq \lambda$. Suppose* (11.2) *holds for the interval decomposition $\mathcal{B}$ of $[d]$. Then:*

$$(11.6) \qquad \frac{1}{d}\left(\mu_i - \mu_j + j - i\right) \leq \min\left\{\mu_i - \mu_j + j - i, \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i}\right\} \leq \mu_i - \mu_j + j - i,$$

for all $i \sim j$, $1 \leqslant i < j \leqslant d$.

Proof. The upper bound is straightforward. For the lower bound, it follows from (11.2), that

$$\frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i} \geq \frac{\lambda_i + d - i}{1 + j - i} \geq \frac{\lambda_i + d - i}{d} \geq \frac{\mu_i - \mu_j + j - i}{d}.$$

It then follows from the equation above that

$$\min\left\{\mu_i - \mu_j + j - i, \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i}\right\} \geq \min\left\{\mu_i - \mu_j + j - i, \frac{\mu_i - \mu_j + j - i}{d}\right\} = \frac{\mu_i - \mu_j + j - i}{d},$$

as desired. $\qquad\qquad\square$

Lemma 11.3. *Fix $d \geq 2$ and $\varepsilon > 0$. Let $\lambda, \mu \in \mathbb{P}_d$, such that $\mu \subseteq \lambda$. Suppose* (11.3) *holds for the interval decomposition $\mathcal{B}$ of $[d]$. Then* (11.6) *holds for all $i \sim j$, $1 \leqslant i < j \leqslant d$.*

Proof. The upper bound is straightforward. For the lower bound, it follows from (11.3) that

$$(11.7) \qquad\qquad\qquad \mu_i - \mu_j + j - i \leq 1 + j - i \leq d.$$

Therefore,

$$\min\left\{\mu_i - \mu_j + j - i, \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i}\right\} \geq_{(4.2),\,(4.3)} 1 \geq_{(11.7)} \frac{1}{d}\left(\mu_i - \mu_j + j - i\right),$$

as desired. $\qquad\qquad\square$

We now combine these three lemmas to give an estimate for the quantity $\Phi(\lambda/\mu)$ if (11.1) holds and either (11.2) or (11.3) holds. Denote

$$(11.8) \qquad K_{\mathcal{B}}(\lambda/\mu) := \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \sim j}} (\mu_i - \mu_j + j - i) \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \not\sim j}} \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i}.$$

Lemma 11.4. *Fix $d \geq 2$ and $\varepsilon > 0$. Let $\lambda, \mu \in \mathbb{P}_d$, such that $\mu \subseteq \lambda$. Suppose* (11.1) *and* (11.2) *hold for the interval decomposition $\mathcal{B}$ of $[d]$. Then:*

$$\left(\frac{\varepsilon}{d^2}\right)^{\frac{d(d-1)}{2}} K_{\mathcal{B}}(\lambda/\mu) \leq \Phi(\lambda/\mu) \leq K_{\mathcal{B}}(\lambda/\mu).$$

The same conclusion holds if condition (11.2) *is replaced with* (11.3).

Proof. By definition of $K_{\mathcal{B}}(\lambda/\mu)$, we have:

$$\frac{\Phi(\lambda/\mu)}{K_{\mathcal{B}}(\lambda/\mu)} = \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \not\sim j}} \frac{\min\left\{\mu_i - \mu_j + j - i, \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i}\right\}}{\frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i}} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \sim j}} \frac{\min\left\{\mu_i - \mu_j + j - i, \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i}\right\}}{\mu_i - \mu_j + j - i} \leq 1.$$

For the lower bound, note that each term in the first product is bounded from below by ε/d, by Lemma 11.1 and condition (11.1). Also note that each term in the second product is bounded from below by $1/d$, by Lemma 11.2 when (11.2) holds, or by Lemma 11.3 when (11.3) holds. This implies the result. $\qquad\square$

The main result of this subsection is the following upper bound.

Lemma 11.5. *Fix $d \geq 2$ and $\varepsilon > 0$. Let $\lambda, \mu \in \mathbb{P}_d$, such that $\mu \subseteq \lambda$. Suppose (11.1) and (11.2) hold for the interval decomposition $\mathcal{B}$ of $[d]$. Then:*

$$\frac{f(\lambda/\mu)}{F(\lambda/\mu)} \leq C_{d,\varepsilon}\Phi(\lambda, \mu).$$

where $C_{d,\varepsilon} > 0$ is an absolute constant. The same conclusion holds if condition (11.2) is replaced with (11.3).

Proof. For the first part, it follows from Lemma 11.4 that $\Phi(\lambda/\mu)$ is equal to $K_B(\lambda/\mu)$ up to a multiplicative constant. Therefore, it suffices to show that

$$\frac{f(\lambda/\mu)}{F(\lambda/\mu)} \leq C_{d,\varepsilon}\, K_{\mathcal{B}}(\lambda/\mu),$$

for some absolute constant $C_{d,\varepsilon} > 0$. Let $N(\ell, B)$ be as in (8.2). Then

$$N(\ell, \mathcal{B}) = \max_{\substack{1 \leqslant i < j \leqslant d \\ i \not\sim j}} \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i} \leq_{(11.1),\,(11.4)} \frac{d}{\varepsilon}.$$

Substituting this into Corollary 8.5, we get

$$\begin{aligned}
\frac{f(\lambda/\mu)}{F(\lambda/\mu)} &\leq C_d \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \sim j}} \left(\mu_i - \mu_j + j - i + \frac{d}{\varepsilon}\right) \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \not\sim j}} \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i} \\
&\leq C_d \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \sim j}} \left(1 + \frac{d}{\varepsilon}\right) (\mu_i - \mu_j + j - i) \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \not\sim j}} \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i} \\
&\leq C_d \left(1 + \frac{d}{\varepsilon}\right)^{\frac{d(d-1)}{2}} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \sim j}} (\mu_i - \mu_j + j - i) \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \not\sim j}} \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i} \\
&\leq_{(11.8)} C_{d,\varepsilon}\, K_{\mathcal{B}}(\lambda/\mu),
\end{aligned}$$

for some absolute constants $C_d, C_{d,\varepsilon} > 0$. This finishes the proof of the first part. The second part follows verbatim; we omit the details. $\qquad\square$

11.3. **Lower bounds.** Our first ingredient is the following estimate on the hook-lengths.

Lemma 11.6. *Fix $d \geq 2$ and $\varepsilon > 0$. Let $(\lambda, \mu) \in \Lambda(n, d, \varepsilon)$. Let $\mathcal{B}$ be an interval decomposition of $[n]$ such that (11.2) holds. Then, for all $(i, j) \in \mu$, and all $k \geq 0$ such that $i \overset{\mathcal{B}}{\sim} (i + k)$, we have:*

$$\frac{h_\lambda(i + k, j + k)}{h_\lambda(i, j)} \geq 1 - \frac{2d}{\varepsilon |\lambda|}.$$

Proof. By the definition (2.6) of the hook lengths, we have:

$$\begin{aligned}
\frac{h_\lambda(i + k, j + k)}{h_\lambda(i, j)} &\geq \frac{\lambda_{i+k} - j - k}{\lambda_i - i + d - j + 1} \geq_{[\text{since } i \sim (i+k)]} \frac{\lambda_i - 1 - j - k}{\lambda_i - i + d - j + 1} = 1 - \frac{d + k + 2 - i}{\lambda_i - i + d - j + 1} \\
&\geq 1 - \frac{2d}{\lambda_i - j + (d - i) + 1} \geq_{[\text{since } (i,j) \in \mu]} 1 - \frac{2d}{\lambda_i - \mu_i} \geq_{(4.7)} 1 - \frac{2d}{\varepsilon |\lambda|},
\end{aligned}$$

as desired. $\qquad\square$

We apply Lemma 11.6 to get a lower bound for the product of hooks of a flagged tableau, see (2.10). Let $\mathcal{B}$ be an interval decomposition of $[d]$. Denote by

$$\mathcal{D}_{\mathcal{B}} := \left\{ T \in \mathrm{FT}(\lambda/\mu) \mid i \overset{\mathcal{B}}{\sim} T(i, j) \text{ for all } (i, j) \in \mu \right\},$$

the set of flagged tableaux of λ/μ, for which the entries for each row i are drawn from the block of $\mathcal{B}$ that contains i.

Lemma 11.7. *Fix $d \geq 2$ and $\varepsilon > 0$. Let $(\lambda, \mu) \in \Lambda(n, d, \varepsilon)$, and let $\mathcal{B}$ be an interval decomposition of $[d]$, such that (11.2) holds. Then, for all $T \in \mathcal{D}_B$, we have:*

$$\prod_{(i,j) \in \mu} \frac{h_\lambda\big(T(i,j), j + T(i,j) - i\big)}{h_\lambda(i,j)} \geq C_{d,\varepsilon},$$

for some absolute constant $C_{d,\varepsilon} > 0$.

Proof. We have:

$$\prod_{(i,j) \in \mu} \frac{h_\lambda\big(T(i,j), j + T(i,j) - i\big)}{h_\lambda(i,j)} \geq_{\text{Lem 11.6}} \left(1 - \frac{2d}{\varepsilon |\lambda|}\right)^{|\mu|} \geq \left(1 - \frac{2d}{\varepsilon |\lambda|}\right)^{|\lambda|} \geq \left(\frac{1}{3}\right)^{2d/\varepsilon},$$

for sufficiently large $|\lambda|$. This implies the result. $\square$

Our second ingredient is the following lower bound on the cardinality of $\mathcal{D}_B$.

Lemma 11.8. *Fix $d \geq 2$ and $\varepsilon > 0$. Let $(\lambda, \mu) \in \Lambda(n, d, \varepsilon)$. Let $\mathcal{B}$ be an interval decomposition of $[n]$ such that (11.2) holds. Then there exists an absolute constant $C_{d,\varepsilon} > 0$ such that*

$$\big|\mathcal{D}_\mathcal{B}\big| \geq C_{d,\varepsilon} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \overset{\mathcal{B}}{\sim} j}} (\mu_i - \mu_j + j - i).$$

Proof. Let $\mathcal{D}'_\mathcal{B} := \mathcal{D}'_\mathcal{B}(\mu)$ be the set of semistandard Young tableau of shape μ given by

$$\mathcal{D}'_\mathcal{B} := \big\{ T \in \mathrm{SSYT}(\mu) \mid i \overset{\mathcal{B}}{\sim} T(i,j) \text{ for all } (i,j) \in \mu \big\}.$$

Note that $\mathcal{D}_\mathcal{B} = \mathcal{D}'_\mathcal{B} \cap \mathrm{FT}(\lambda/\mu)$. We will estimate $|\mathcal{D}_\mathcal{B}|$ via $|\mathcal{D}'_\mathcal{B}|$.

Recall the definition (8.1) of interval decompositions. For each $k \in \{1, \ldots, r\}$, denote by $\mu^{(k)}$ the partition obtained from μ by restricting to rows indexed by B_k:

(11.9) $\qquad \mu^{(k)} = \big(\mu_1^{(k)}, \mu_2^{(k)}, \ldots, \mu_{b_k - b_{k-1}}^{(k)}\big) := \big(\mu_{b_{k-1}+1}, \mu_{b_{k-1}+2}, \ldots, \mu_{b_k}\big).$

In this notation,

(11.10) $\qquad \mathcal{D}'_\mathcal{B} = \big\{ T \in \mathrm{SSYT}(\mu) \mid b_{k-1} < T(i,j) \leq b_k \text{ for all } (i,j) \in \mu, \, i \in B_k \big\}.$

Therefore, that the following map is a bijection:

(11.11)
$$\psi : \mathcal{D}'_\mathcal{B} \to \mathrm{SSYT}\big(\mu^{(1)}\big) \times \ldots \times \mathrm{SSYT}\big(\mu^{(r)}\big), \quad \psi(T) := \big(T^{(1)}, \ldots, T^{(r)}\big),$$
$$\text{where} \quad T^{(k)}(i,j) = T(i + b_{k-1}, j) - b_{k-1} \quad \text{for all} \quad (i,j) \in \mu^{(k)}.$$

In other words, the semistandard Young tableaux $T^{(k)}$ is obtained by restricting T to rows indexed by B_k and normalizing the smallest entries to start from 1. It now follows from (11.11) and (2.5), that

(11.12)
$$|\mathcal{D}'_\mathcal{B}| = \prod_{k=1}^{r} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i,j \in B_k}} \frac{\mu_i - \mu_j + j - i}{j - i} \geq \prod_{k=1}^{r} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i,j \in B_k}} \frac{\mu_i - \mu_j + j - i}{d - 1}$$
$$\geq (d-1)^{-\frac{d(d-1)}{2}} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \overset{\mathcal{B}}{\sim} j}} \big(\mu_i - \mu_j + j - i\big).$$

We claim that $\mathcal{D}_\mathcal{B} = \mathcal{D}'_\mathcal{B}$ for sufficiently large $|\lambda|$. It suffices to show that each $T \in \mathcal{D}'_\mathcal{B}$ is a flagged tableau of λ/μ, for sufficiently large $|\lambda|$. Let $(i,j) \in \mu$, and let k be the index such that B_k is the block of $\mathcal{B}$ that contains i. We have:

$$j + T(u) - 1 \leq \mu_i + d - 1 = \lambda_i - (\lambda_i - \mu_i) + d - 1 \leq_{(4.7)} \lambda_i - \varepsilon |\lambda| + (d - i)$$
$$\leq_{(11.2), \, i \overset{\mathcal{B}}{\sim} T(u)} (\lambda_{T(u)} + 1) - \varepsilon |\lambda| + (d - i) \leq \lambda_{T(u)},$$

for sufficiently large $\lambda_{T(u)}$. This proves the claim.

By (4.7), we have:

(11.13) $\qquad\qquad\qquad \lambda_1 \geq \ldots \geq \lambda_d \geq \lambda_d - \mu_d \geq \varepsilon |\lambda|,$

so the claim above assumes only that $|\lambda|$ is large enough. We conclude:

$$(11.14) \qquad |\mathcal{D}_\mathcal{B}| \;=_{(11.13)}\; |\mathcal{D}'_\mathcal{B}| \;\geq_{(11.12)}\; (d-1)^{-\frac{d(d-1)}{2}} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \overset{\mathcal{B}}{\sim} j}} \big(\mu_i - \mu_j + j - i\big),$$

for all $|\lambda|$ sufficiently large. This completes the proof. $\qquad\qquad\qquad\qquad\qquad\square$

The main result of this subsection is the following lower bound for $f(\lambda/\mu)$.

Lemma 11.9. *Fix $d \geq 2$ and $\varepsilon > 0$. Let $(\lambda, \mu) \in \Lambda(n, d, \varepsilon)$. Let $\mathcal{B}$ be an interval decomposition of $[n]$ such that (11.1) and (11.2) hold. Then there exists an absolute constant $C_{d,\varepsilon} > 0$ such that*

$$\frac{f(\lambda/\mu)}{F(\lambda/\mu)} \;\geq\; C_{d,\varepsilon}\, \Phi(\lambda/\mu).$$

Proof. We have:

$$\frac{f(\lambda/\mu)}{F(\lambda/\mu)} =_{\text{Thm } 2.3} \sum_{T \in \text{FT}(\lambda/\mu)} \prod_{(i,j) \in \mu} \frac{h_\lambda(T(i,j), j + T(i,j) - i)}{h_\lambda(i,j)} \geq \sum_{T \in \mathcal{D}_B} \prod_{(i,j) \in \mu} \frac{h_\lambda(T(i,j), j + T(i,j) - i)}{h_\lambda(i,j)}$$

$$\geq_{\text{Lem } 11.7} \sum_{T \in \mathcal{D}_B} C_{d,\varepsilon} = C_{d,\varepsilon} |\mathcal{D}_B| \geq_{\text{Lem } 11.8} C_{d,\varepsilon} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \,\sim\, j}} \big(\mu_i - \mu_j + j - i\big).$$

This implies that

$$\frac{f(\lambda/\mu)}{F(\lambda/\mu)} \geq_{(11.4)} C_{d,\varepsilon} \left(\frac{\varepsilon}{d}\right)^{\frac{d(d-1)}{2}} \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \,\sim\, j}} \big(\mu_i - \mu_j + j - i\big) \prod_{\substack{1 \leqslant i < j \leqslant d \\ i \,\not\sim\, j}} \frac{\lambda_i + d - i}{\lambda_i - \lambda_j + j - i}$$

$$=_{(11.8)} C_{d,\varepsilon} \left(\frac{\varepsilon}{d}\right)^{\frac{d(d-1)}{2}} K_\mathcal{B}(\lambda/\mu) \geq_{\text{Lem } 11.4} C_{d,\varepsilon} \left(\frac{\varepsilon}{d}\right)^{\frac{d(d-1)}{2}} \Phi(\lambda/\mu),$$

as desired. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

11.4. **Proof of Lemma 4.5.** Recall the definition of a Thoma pair (α, β) in §1.2. Let $\varepsilon := \varepsilon(\alpha, \beta)$ be given by

$$(11.15) \qquad \varepsilon := \frac{1}{2(\alpha_1 + \ldots + \alpha_d)} \min\left\{ \min_{1 \leqslant i < j \leqslant d} \{\alpha_i - \beta_i\},\; \min_{\substack{1 \leq i < j \leq d \\ \alpha_i \neq \alpha_j}} \{\alpha_i - \alpha_j\} \right\}.$$

We have $\varepsilon > 0$ since $\alpha_i > \beta_i$.

Let $\lambda \simeq \alpha n$, $\mu \simeq \beta n$ be a TVK (α, β)-shape. Note that (λ, μ) is ε-admissible for sufficiently large n. Indeed,

$$(11.16) \qquad \lambda_i - \mu_i \;=\; \lfloor \alpha_i n \rfloor - \lfloor \beta_i n \rfloor \;\geq\; (\alpha_i - \beta_i)n - 1 \;\geq\; \frac{\alpha_i - \beta_i}{2(\alpha_1 + \ldots + \alpha_d)} |\lambda| \;\geq\; \varepsilon |\lambda|,$$

for sufficiently large $|\lambda| = |\alpha| n + O(1)$, and for all $1 \leq i \leq d$.

Let $\mathcal{B}$ be the interval decomposition of $[d]$ that puts two integers $i, j \in [d]$ in the same block if and only if $\alpha_i = \alpha_j$. Then (11.1) holds for sufficiently large $|\lambda|$, since for all $i \overset{\mathcal{B}}{\not\sim} j$, $1 \leqslant i < j \leqslant d$, we have:

$$(11.17) \qquad \lambda_i - \lambda_j \;=\; \lfloor \alpha_i n \rfloor - \lfloor \alpha_j n \rfloor \;\geq\; (\alpha_i - \alpha_j)n - 1 \;\geq\; \frac{\alpha_i - \alpha_j}{2(\alpha_1 + \ldots + \alpha_d)} |\lambda|, \;\geq\; \varepsilon |\lambda|,$$

Similarly, (11.2) holds, since for all $i \overset{\mathcal{B}}{\sim} j$, $1 \leqslant i < j \leqslant d$, we have:

$$(11.18) \qquad\qquad\qquad \lambda_i - \lambda_j \;=\; \lfloor \alpha_i n \rfloor - \lfloor \alpha_j n \rfloor \;=\; 0.$$

By Lemma 11.5 and Lemma 11.9, this implies that there exists an absolute constant $C_{\alpha,\beta} > 0$, such that

$$\frac{1}{C_{\alpha,\beta}} \Phi(\lambda/\mu) \;\leq\; \frac{f(\lambda/\mu)}{F(\lambda/\mu)} \;\leq\; C_{\alpha,\beta}\, \Phi(\lambda/\mu),$$

for sufficiently large n. This implies the result. $\square$

11.5. **Proof of Theorem 1.4.** Let $\varepsilon := \varepsilon(\alpha, \beta)$ be as in (11.15). By Lemma 4.3, it suffices to check that for every ε-admissible triplet (λ, γ, μ), we have:

$$(11.19) \qquad \frac{f(\gamma/\mu)}{F(\gamma/\mu)} \leq C_{\alpha,\beta}\,\Phi(\gamma/\mu)\,, \quad \frac{f(\lambda/\gamma)}{F(\lambda/\gamma)} \leq C_{\alpha,\beta}\,\Phi(\lambda/\gamma)\,, \quad \text{and} \quad \frac{f(\lambda/\mu)}{F(\lambda/\mu)} \geq \frac{1}{C_{\alpha,\beta}}\,\Phi(\lambda/\mu),$$

for some absolute constant $C_{\alpha,\beta} > 0$. Note that the third inequality in (11.19) is proved in Lemma 4.5.

For the second inequality in (11.19), let $\mathcal{B}$ be the interval decomposition of $[d]$ that puts two integers $i, j \in [d]$ in the same block if and only if $\alpha_i = \alpha_j$. By the same argument as in (11.17) and (11.18), we have (11.1) and (11.2) hold for the pair (λ, γ) and $\mathcal{B}$, and for sufficiently large n. By Lemma 11.5, we get the second inequality for sufficiently large n.

For the first inequality in (11.19), let $\mathcal{B}'$ be the interval decomposition of $[d]$ that puts $i, j \in [d]$ in the same block if and only if $\beta_i = \beta_j$. Let $\varepsilon' := \varepsilon'(\alpha, \beta)$ be the constant defined by

$$\varepsilon' := \frac{d\varepsilon^3}{8(\alpha_1 + \ldots + \alpha_d)} \min_{\substack{1 \leq i < j \leq d \\ \beta_i \neq \beta_j}} \{\beta_i - \beta_j\}.$$

For all $i \overset{\mathcal{B}'}{\sim} j$, $1 \leq i < j \leq d$, we have:

$$(11.20) \qquad \gamma_i - \gamma_j \geq_{(4.8)} p(\lambda_i - \lambda_j) + (1-p)(\mu_i - \mu_j) - 2(|\lambda| - |\mu|)^{\frac{3}{4}} \geq (1-p)(\mu_i - \mu_j) - 2n^{\frac{3}{4}}.$$

Note that

$$(11.21) \qquad \mu_i - \mu_j = \lfloor \beta_i n \rfloor - \lfloor \beta_j n \rfloor \geq (\beta_i - \beta_j)|\lambda| - 1 \geq \frac{\beta_i - \beta_j}{2(\alpha_1 + \ldots + \alpha_d)}\,|\lambda|,$$

for sufficiently large n. Note also that

$$(11.22) \qquad 1 - p =_{(4.9)} \frac{|\lambda| - |\gamma|}{|\lambda| - |\mu|} = \sum_{i=1}^{d} \frac{\lambda_i - \gamma_i}{|\lambda| - |\mu|} \geq_{(4.7)} \frac{d\varepsilon^3}{2} \frac{|\lambda|}{|\lambda| - |\mu|} \geq \frac{d\varepsilon^3}{2}.$$

Substituting (11.21) and (11.22) into (11.20), we get

$$(11.23) \qquad \gamma_i - \gamma_j \geq \frac{d\varepsilon^3}{4} \frac{\beta_i - \beta_j}{\alpha_1 + \ldots + \alpha_d}\,|\lambda| - 2n^{\frac{3}{4}} \geq \frac{d\varepsilon^3}{8} \frac{\beta_i - \beta_j}{\alpha_1 + \ldots + \alpha_d}\,|\lambda| \geq \varepsilon'|\lambda| \geq \varepsilon'|\gamma|,$$

for $|\lambda| = \Theta(n)$ large enough. On the other hand, for all $i \overset{B'}{\sim} j$, $1 \leq i < j \leq d$, we have:

$$(11.24) \qquad \mu_i - \mu_j = \lfloor \beta_i n \rfloor - \lfloor \beta_j n \rfloor = 0.$$

It follows from (11.23) and (11.24), that (11.1) and (11.3) hold for this case when n is sufficiently large. Thus, the first inequality in (11.19) follows by Lemma 11.5. This completes the proof of the theorem. $\square$

12. Conjectures and open problems

We believe our results can be further strengthened in several directions, and would like to mention a few possibilities.

12.1. **Sorting probability.** The bound $\delta(P_{\lambda/\mu}) = O\big(\frac{1}{\sqrt{n}}\big)$ that we obtain in Theorems 1.2–1.5 is likely not tight. In fact, $\Omega\big(\frac{1}{n}\big)$ is the only lower bound that we know in some cases (see §1.5). The results in Corollary 3.5 and [CPP21] also seem to suggest that $O\big(\frac{1}{n}\big)$ is perhaps the best one can aim for in full generality. We believe the TVK shapes are likely the easiest case to make progress as they are most similar to the Catalan poset case:

Conjecture 12.1. *There is a universal constant $C > 0$, such that for all $d \geq 2$, and for every Thoma sequence $\alpha \in \mathbb{R}^d_{>0}$, we have:*

$$\delta(P_\lambda) \leq \frac{C}{n^{5/4}}\,,$$

where $\lambda \simeq \alpha n$ is a TVK α-shape.

We believe the same bound holds for more general cases. To understand our reasoning, note that we take $a = \lfloor \lambda_1/2 \rfloor$ in this case to minimize the sorting probability. Even if the bound we obtain is tight, by varying a one is likely to obtain lower global minimum in the definition of the sorting probability. In fact, we believe the following general claim with a weaker bound:

Conjecture 12.2. *There is a universal constant $C > 0$, such that for every $\lambda \vdash n$, $\lambda \neq (n), (1^n)$, we have:*

$$\delta(P_\lambda) \leq \frac{C}{\sqrt{n}}.$$

This conjecture is suggesting that the constants $C_{d,\varepsilon}$ in Theorem 1.3 and Theorem 1.5 can be made independent of parameters d and ε, even though the proofs give dependence that is relatively wild. See, e.g. the last line of the proof of Theorem 8.4. At the moment, we cannot even prove that $\delta(P_\lambda) \to 0$ for general partitions λ, with $n = |\lambda| \to \infty$.

In a different direction, suppose λ is a 3-dimensional diagram defined as lower ideals in $\mathbb{N}^3$. The tools of this paper are heavily based on the HLF (2.7), NHLF (2.8), asymptotics of Schur functions and other Algebraic Combinatorics results. None of these are available for 3-dimensional diagrams, even for the boxes (products of three chains). Finding new tools to establish such bounds would be a major breakthrough.

Conjecture 12.3. *Fix $d, r \geq 2$. Denote by $P_{d,r,m}$ the 3-dimensional poset given by a $[d \times r \times m] \subset \mathbb{N}^3$ box (product of chains on size d, and m, respectively). Then:*

$$\delta(P_{d,r,m}) = O\!\left(\frac{1}{m}\right), \quad as \ m \to \infty.$$

A more general problem would be to find conditions on the poset $P = (X, \prec)$ of bounded width, which would guarantee that the sorting probability $\delta(P) \to 0$ as the size $|X| \to \infty$.

12.2. **Technical estimates.** The tools of this paper are based on bounds for $f(\lambda/\mu) = |\operatorname{SYT}(\lambda/\mu)|$, which are of independent interest. Recall the definition of $F(\lambda/\mu)$ in (2.9) and the bound in Theorem 2.2. Recall also the balance function $\Phi(\lambda/\mu)$ defined in (4.1) and the bounds in Lemmas 4.4 and 4.5. The following conjecture is a natural generalization.

Conjecture 12.4. *Fix $d \geq 2$. Let $\lambda/\mu \vdash n$, $\ell(\lambda) \leq d$. Then:*

$$(12.1) \qquad \frac{1}{C_d}\, \Phi(\lambda/\mu) \ \leq \ \frac{f(\lambda/\mu)}{F(\lambda/\mu)} \ \leq \ C_d\, \Phi(\lambda/\mu),$$

for an absolute constant $C_d > 0$.

One can generalize the definition of $\Phi(\lambda/\mu)$ to continuous setting:

$$\Phi(\boldsymbol{x}/\mu) := \prod_{1 \leqslant i < j \leqslant d} \min\left\{\mu_i - \mu_j + j - i,\ \frac{x_i}{x_i - x_j}\right\},$$

where $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$, and $\mu = (\mu_1, \ldots, \mu_d)$ is an integer partition.

Conjecture 12.5. *Fix $d \geq 2$ and $\varepsilon > 0$. Then, for every $\mu = (\mu_1, \ldots, \mu_d)$ and $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$, such that $x_1 > \ldots > x_d > \varepsilon x_1 > 0$, we have:*

$$(12.2) \qquad \frac{s_\mu(x_1, \ldots, x_d)}{x_1^{\mu_1} \cdots x_d^{\mu_d}} \ \leq \ C_{d,\varepsilon}\, \Phi(\boldsymbol{x}/\mu),$$

for an absolute constant $C_{d,\varepsilon} > 0$.

We obtain partial results in favor of this conjecture: a lower bound in Lemma 8.3 and an upper bound in Theorem 8.4. Let us present the former with simplified notation, as it also gives connection between Conjectures 12.4 and 12.5.

Theorem 12.6 (= Lemma 8.3**).** *Let λ/μ be a skew shape, where $\lambda = (\lambda_1, \ldots, \lambda_d)$, $\mu = (\mu_1, \ldots, \mu_d)$. Then:*

$$1 \ \leq \ \frac{f(\lambda/\mu)}{F(\lambda/\mu)} \ \leq \ \frac{s_\mu(\lambda_1 + d - 1, \lambda_2 + d - 2, \ldots, \lambda_d)}{(\lambda_1 + d - 1)^{\mu_1} (\lambda_2 + d - 2)^{\mu_2} \cdots \lambda_d^{\mu_d}}.$$

Remark 12.7. Conjecture 12.5 in the earlier version of the paper had a matching lower bound:

$$\frac{1}{C_{d,\varepsilon}} \; \Phi(\boldsymbol{x}/\mu) \;\leq\; \frac{s_\mu(x_1,\ldots,x_d)}{x_1^{\mu_1} \cdots x_d^{\mu_d}} \;.$$

Unfortunately, this bound fails for the substitution $x_i \leftarrow q^i$ for $q \in \left[\varepsilon^{1/(d-1)}, 1\right)$, by the *hook-content formula*, see [Sta99, Thm 7.21.2]. On the other hand, the upper bound (12.2) is easy to check in this case.

13. FINAL REMARKS

13.1. Although much of the paper is motivated by the work surrounding the $\frac{1}{3}$–$\frac{1}{3}$ Conjecture 1.1, we do not resolve the conjecture in any new cases. As mentioned in the introduction, for all skew Young diagrams the conjecture was already established in [OS18]. In fact, when compared with the Kahn–Saks Conjecture 1.6, our results are counterintuitive since we obtain the conclusion of the conjecture in a strong form, while the width of our posets remains bounded. Clearly, much of the subject remains misunderstood and open to further exploration.

13.2. The technical assumption in Theorem 1.3, that λ is ε-thick is likely unnecessary, but at the moment we do not know how to avoid it. The same applies for the ε-smooth assumption, and the Main Theorem 1.5 most likely holds under much weaker assumptions. Let us remark though, that in some formal sense these two assumptions are equivalent. Indeed, let $\lambda = (\lambda_1,\ldots,\lambda_d)$ and $\mu = (\lambda_{i+1},\ldots,\lambda_{i+1},\lambda_{i+2},\ldots,\lambda_d)$. The skew shape $\nu := \lambda/\mu = (\lambda_1 - \lambda_{i+1},\ldots,\lambda_i - \lambda_{i+1})$ is then the straight shape, so the ε-smooth condition $\lambda_i - \lambda_{i+1} \geq \varepsilon n$ becomes the ε-thick condition for ν.

13.3. For a fixed number of rows $d = \ell(\lambda)$, Corollary 3.5 shows that $\delta(P_\lambda) = O\left(\frac{1}{n}\right)$ for all $\lambda \vdash n$, such that $\lambda_2 = O(1)$. This is the opposite extreme of ε-thick diagrams λ, suggesting that the ε-thick assumption in Theorem 1.3 might be unnecessary indeed.

13.4. The upper bound in (11.2) and (11.3) can be replaced with an arbitrary constant K at the cost of changing the positive constant $C_{d,\varepsilon}$ in our results into the positive constant $C_{d,\varepsilon,K}$, which now also depends on K. The rest of the proof follows verbatim and gives a slight extension of Theorem 1.4 under weaker conditions $|\lambda_i - \alpha_i n| \leq K$, and the same for the μ. We omit the details.

13.5. The Naruse's hook-length formula (2.8) works well when $|\lambda/\mu|$ is relatively small compared to $|\lambda|$. On the other hand, when $|\mu|$ is very small, there is another positive formula due to Okounkov and Olshanski [OO98], which was observed in [OO98, Sta99] to give sharp estimates in that regime. In [MPP1, §9.4], the authors suggested that this rule is equivalent to the Knutson–Tao "equivariant puzzles" rule. This was proved in [MZ+], which reworked the Okounkov–Olshanski formula in the NHLF-style. It would be interesting to see if this formula can be used in place of NHLF to obtain sharper bounds on the sorting probability of skew Young diagrams, at least in some cases.

13.6. When $\lambda = (m^d)$ is a rectangle, one can estimate $\delta(P_\lambda)$ without the NHLF, since $f(\lambda/\mu)$ can be computed by the hook-length formula (2.7). This greatly simplifies the calculations, and is an approach take in [CPP21] for the Catalan numbers example $\lambda = \left(\frac{n}{2}, \frac{n}{2}\right)$, see §1.5.

13.7. As we mentioned in the previous section, there are several places where our bounds are likely not sharp. First, the argument in §5.2, is a quantitative version of Linial's pigeonhole principle argument, which we also employ in §3.1. But the real obstacle to improving the $O\left(\frac{1}{\sqrt{n}}\right)$ bound is not apparent until Section 8, where the interval decompositions are introduced and a different pigeonhole argument is used.

13.8. Most recently, Conjecture 1.1 was generalized to all Coxeter groups [GG20]. It would be interesting to see if our results extend to this setting.

References

[Aig85] M. Aigner, A note on merging, *Order* **2** (1985), 257–264.

[BR10]   Y. Baryshnikov and D. Romik, Enumeration formulas for Young tableaux in a diagonal strip, *Israel J. Math.* **178** (2010), 157–186.

[Bri89] G. R. Brightwell, Semiorders and the $\frac{1}{3}-\frac{2}{3}$ conjecture, *Order* **5** (1989), 369–380.

[Bri93] G. R. Brightwell, Models of random partial orders, in *Surveys in combinatorics*, Cambridge Univ. Press, Cambridge, UK, 1993, 53–83.

[Bri99] G. R. Brightwell, Balanced pairs in partial orders, *Discrete Math.* **201** (1999), 25–52.

[BFT95] G. R. Brightwell, S. Felsner and W. T. Trotter, Balancing pairs and the cross product conjecture, *Order* **12** (1995), 327–349.

[BW91]   G. R. Brightwell and P. Winkler, Counting linear extensions, *Order* **8** (1991), 225–242.

[BW92]   G. R. Brightwell and C. Wright, The $\frac{1}{3}-\frac{2}{3}$ conjecture for 5-thin posets, *SIAM J. Discrete Math.* **5** (1992), 467–474.

[C+13]   J. Cardinal, S. Fiorini, G. Joret, R. M. Jungers and J. I. Munro, Sorting under partial information (without the ellipsoid algorithm), *Combinatorica* **33** (2013), 655–697.

[CPP21] S. H. Chan, I. Pak and G. Panova, Sorting probability of Catalan posets, *Advances Applied Math.* **129** (2021), Paper No. 102221, 13 pp.

[Chen18] E. Chen, A family of partially ordered sets with small balance constant, *Electron. J. Combin.* **25** (2018), Paper 4.43, 13 pp.

[CKP01] H. Cohn, R. Kenyon and J. Propp, A variational principle for domino tilings, *Jour. AMS* **14** (2001), 297–346.

[DP18]   S. Dittmer and I. Pak, Counting linear extensions of restricted posets, preprint (2018), 33 pp.; `arXiv:1802.06312`.

[Feit53] W. Feit, The degree formula for the skew-representations of the symmetric group, *Proc. AMS* **4** (1953), 740–744.

[FRT54] J. S. Frame, G. de B. Robinson and R. M. Thrall, The hook graphs of the symmetric group, *Canad. J. Math.* **6** (1954), 316–324.

[Fre75] M. L. Fredman, How good is the information theory bound in sorting?, *Theoret. Comput. Sci.* **1** (1975), 355–361.

[GG20]   C. Gaetz and Y. Gao, Balance constants for Coxeter groups, preprint (2020), 27 pp.; `arXiv:2005.09719`.

[GHP87] B. Ganter, G. Häfner and W. Poguntke, On linear extensions of ordered sets with a symmetry, *Discrete Math.* **63** (1987), 153–156.

[Hoe63] W. Hoeffding, Probability inequalities for sums of bounded random variables, *J. Amer. Stat. Assoc.* **58** (1963), 13–30.

[Jan11] J. Janson, Poset limits and exchangeable random posets, *Combinatorica* **31** (2011), 529–563.

[KL91]   J. Kahn and N. Linial, Balancing extensions via Brunn–Minkowski, *Combinatorica* **11** (1991), 363–368.

[KK95]   J. Kahn and J. H. Kim, Entropy and sorting, *J. Comput. Syst. Sci.* **51** (1995), 390–399.

[KS84]   J. Kahn and M. Saks, Balancing poset extensions, *Order* **1** (1984), 113–126.

[Kis68] S. S. Kislitsyn, A finite partially ordered set and its corresponding set of permutations, *Math. Notes* **4** (1968), 798–801.

[KR75]   D. J. Kleitman and B. L. Rothschild, Asymptotic enumeration of partial orders on a finite set, *Trans. AMS* **205** (1975), 205–220.

[Kom90] J. Komlós, A strange pigeonhole principle, *Order* **7** (1990), 107–113.

[Kon]   M. Konvalinka, A bijective proof of the hook-length formula for skew shapes, *Europ. J. Combin.* **88** (2020), 103104, 14 pp.

[Kor94] A. D. Korshunov, On linear extensions of partially ordered sets (in Russian), *Trudy Inst. Mat.* (*Novosibirsk*) **27** (1994), 34–42, 179.

[Lin84] N. Linial, The information-theoretic bound is good for merging, *SIAM J. Comput.* **13** (1984), 795–801.

[Mac92] I. G. Macdonald, Schur functions: theme and variations, in *Sém. Lothar. Combin.*, Strasbourg, 1992, 5–39.

[Mac95] I. G. Macdonald, Symmetric functions and Hall polynomials (Second ed.), Oxford U. Press, New York, 1995, 475 pp.

[MPP1]   A. H. Morales, I. Pak and G. Panova, Hook formulas for skew shapes I. *q*-analogues and bijections, *J. Combin. Theory, Ser. A* **154** (2018), 350–405.

[MPP2]   A. H. Morales, I. Pak and G. Panova, Hook formulas for skew shapes II. Combinatorial proofs and enumerative applications, *SIAM J. Discrete Math.* **31** (2017), 1953–1989.

[MPP3]   A. H. Morales, I. Pak and G. Panova, Hook formulas for skew shapes III. Multivariate and product formulas, *Algebraic Combin.* **2** (2019), 815–861.

[MPP4]   A. H. Morales, I. Pak and G. Panova, Asymptotics of the number of standard Young tableaux of skew shape, *European J. Combin.* **70** (2018), 26–49.

[MPT18] A. H. Morales, I. Pak and M. Tassy, Asymptotics for the number of standard tableaux of skew shape and for weighted lozenge tilings, *Combinatorics, Probability and Computing*, to appear, 25 pp.; `arXiv:1805.00992`.

[MZ+]   A. H. Morales and D. G. Zhu, On the Okounkov–Olshanski formula for standard tableaux of skew shapes, preprint (2020), 37 pp.; `arXiv:2007.05006`.

[OO98]   A. Okounkov and G. Olshanski, Shifted Schur functions, *St. Petersburg Math. J.* **9** (1998), 239–300.

[OS18]   E. J. Olson and B. E. Sagan, On the $\frac{1}{3}-\frac{2}{3}$ conjecture, *Order* **35** (2018), 581–596.

[Pak21] I. Pak, Skew shape asymptotics, a case-based introduction, *Sém. Lothar. Combin.* **84**, Art. B84a (2021), 26 ipp.

[Pit06] J. Pitman, *Combinatorial stochastic processes*, Springer, Berlin, 2006, 256 pp.

[PR07]   B. Pittel and D. Romik, Limit shapes for random square Young tableaux, *Adv. Appl. Math.* **38** (2007), 164–209.

[Sag01] B. E. Sagan, *The symmetric group* (Second ed.), Springer, New York, 2001, 238 pp.

[Sah21] A. Sah, Improving the $\frac{1}{3}-\frac{2}{3}$ conjecture for width two posets, *Combinatorica* **41** (2021), 99–126.

[Saks85]   M. Saks, Balancing linear extensions of ordered sets, *Order* **2** (1985), 327–330.

[Sta99]   R. P. Stanley, *Enumerative Combinatorics*, vol. 1 (second ed.) and vol. 2, Cambridge Univ. Press, 2012 and 1999.

[Sta03]   R. P. Stanley, On the enumeration of skew Young tableaux, *Adv. Appl. Math.* **30** (2003), 283–294.

[Sun18]   W. Sun, Dimer model, bead and standard Young tableaux: finite cases and limit shapes, preprint (2018), 67 pp.; `arXiv:1804.03414`.

[Tro95]   W. T. Trotter, Partially ordered sets, in *Handbook of combinatorics*, Vol. 1, Elsevier, Amsterdam, 1995, 433–480.

[TGF92]   W. T. Trotter, W. G. Gehrlein and P. C. Fishburn, Balance theorems for height-2 posets, *Order* **9** (1992), 43–53.

[VK81]   A. M. Vershik and S. V. Kerov, The asymptotic character theory of the symmetric group, *Funct. Anal. Appl.* **15** (1981), 246–255.

[Zag12]   I. Zaguia, The $\frac{1}{3}-\frac{2}{3}$ conjecture for N-free ordered sets, *Electron. J. Combin.* **19** (2012), no. 2, Paper 29, 5 pp.

[Zag19]   I. Zaguia, The $\frac{1}{3}-\frac{2}{3}$ conjecture for ordered sets whose cover graph is a forest, *Order* **36** (2019), 335–347.

(Swee Hong Chan) Department of Mathematics, UCLA, Los Angeles, CA 90095.
Email address: `sweehong@math.ucla.edu`

(Igor Pak) Department of Mathematics, UCLA, Los Angeles, CA 90095.
Email address: `pak@math.ucla.edu`

(Greta Panova) Department of Mathematics, USC, Los Angeles, CA 90089.
Email address: `gpanova@usc.edu`