Random partitions under the Plancherel–Hurwitz measure and high genus Hurwitz maps

Guillaume Chapuy
Université de Paris, IRIF, CNRS, F-75013 Paris, France

Baptiste Louf
Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden

Harriet Walsh
Univ Lyon, ENS de Lyon, CNRS, Laboratoire de Physique, F-69342 Lyon, France
Université de Paris, IRIF, CNRS, F-75013 Paris, France

Abstract

We study the asymptotic behaviour of random integer partitions under a new probability law that we introduce, the Plancherel–Hurwitz measure. This distribution, which has a natural definition in terms of Young tableaux, is a deformation of the classical Plancherel measure. It appears naturally in the enumeration of Hurwitz maps, or equivalently transposition factorisations in symmetric groups.

We study a regime in which the number of factors in the underlying factorisations grows linearly with the order of the group, and the corresponding maps are of high genus. We prove that the limiting behaviour exhibits a new, twofold, phenomenon: the first part becomes very large, while the rest of the partition has the standard Vershik–Kerov–Logan–Shepp limit shape. As a consequence, we obtain asymptotic estimates for unconnected Hurwitz numbers with linear Euler characteristic, which we use to study random Hurwitz maps in this regime. This result can also be interpreted as the return probability of the transposition random walk on the symmetric group after linearly many steps.

2012 ACM Subject Classification Mathematics of computing → Enumeration; Mathematics of computing → Distribution Functions; Mathematics of computing → Random graphs

Keywords and phrases Random partitions, limit shapes, transposition factorisations, map enumeration, Hurwitz numbers, RSK algorithm, giant components

Digital Object Identifier 10.4230/LIPIcs.CVIT.2016.23

Funding This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. ERC-2016-STG 716083 "CombiTop").
Baptiste Louf: supported by the Knut and Alice Wallenberg foundation.

Acknowledgements We thank Philippe Biane, Jérémie Bouttier and Andrea Sportiello for insightful conversations.

1 Random partitions, Plancherel and Plancherel–Hurwitz measures

Let $n \geq 1$ be an integer and let $\mathfrak{S}_n$ denote the group of permutations on $[n] := \{1, 2, \dots, n\}$. The famous *RSK algorithm* (Robinson, Schensted, Knuth) associates each permutation $\sigma \in \mathfrak{S}_n$ bijectively to a pair (P, Q) of standard Young tableaux of the same shape. It is impossible to overstate the importance of this construction in enumerative and algebraic combinatorics. At the enumerative level, the RSK algorithm gives a bijective proof of the following identity:

$$\sum_{\lambda \vdash n} (f_\lambda)^2 = n!, \tag{1}$$

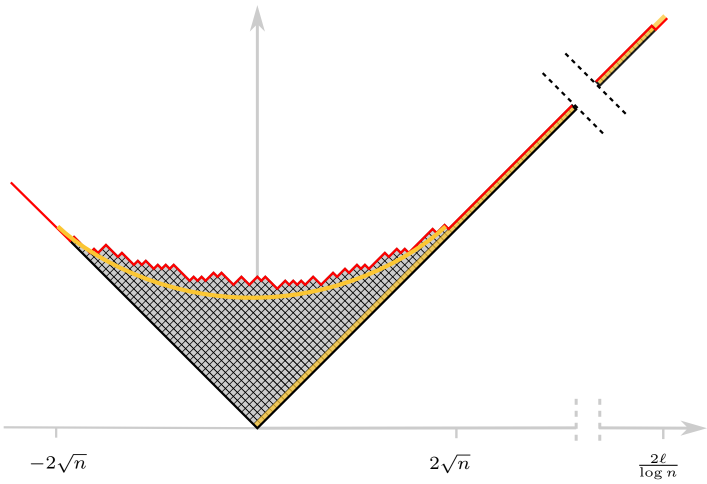

Figure 1 A random partition of $n = 2500$ under the Plancherel–Hurwitz measure $\mathbb{P}_{n,\ell}$ in the high genus regime $\ell = 2\lfloor 1.5n \rfloor$ (sampled via a Metropolis–Hastings algorithm). The twofold asymptotic behaviour is shown in yellow: the first part λ_1 is asymptotic to $\frac{2\ell}{\log(n)}$ and escapes the picture, while the rest of the partition scales in $\sqrt{n}$ with a VKLS limit shape. See Theorem 2. The profile of the partition is in red, while the VKLS limit shape scaled up to $\sqrt{n} \cdot \Omega(x/\sqrt{n})$ is the yellow curve.

where the sum is taken over integer partitions of n, and where f_λ is the number of SYT (standard Young tableaux) of shape λ (see Figure 2 or Section 3). If the permutation σ is chosen uniformly at random, the shape λ of the associated tableaux is a random partition of n distributed according to the probability measure

$$\lambda \mapsto \frac{1}{n!}(f_\lambda)^2, \tag{2}$$

which is the *Plancherel measure* of the symmetric group $\mathfrak{S}_n$.

The study of random partitions under the Plancherel measure is an immense subject in itself with many ramifications. One of the classical and most famous results is the fact, due independently to Logan and Shepp [15] and Vershik and Kerov [19], that when n goes to infinity, the diagram of a Plancherel distributed partition converges in some precise sense to a deterministic limit shape (Theorem 7 below) that we call the *VKLS limit shape* following initials of these authors. Other deep results deal with the behaviour of the largest part λ_1, which coincides with the *longest increasing subsequence* of the random permutation σ, which scales as $2\sqrt{n}$ with fluctuations of order $n^{1/6}$ driven by a Tracy–Widom distribution [14]. The book [17] is a delightful introduction to the subject.

In this paper we will be interested in a generalisation of this measure, motivated by the study of transposition factorisation, or Hurwitz maps, discussed in the next section. For an even integer $\ell \geq 0$, we let $H_{n,\ell}$ be the number of factorisations of the identity of $\mathfrak{S}_n$ into ℓ transpositions:

$$H_{n,\ell} = \#\{(\tau_1, \tau_2, \ldots, \tau_\ell) \in (\mathfrak{S}_n)^\ell, \tau_1\tau_2\cdots\tau_\ell = \mathrm{id},\ \text{each } \tau_i \text{ is a transposition}\}. \tag{3}$$

The *Frobenius formula* from representation theory of finite groups (see e.g. [12]) together with the combinatorial representation theory of $\mathfrak{S}_n$, gives an explicit expression for the number $H_{n,\ell}$ as a sum over partitions, vastly generalising (1) (which corresponds to $\ell = 0$). Indeed,

$$H_{n,\ell} = \frac{1}{n!}\sum_{\lambda \vdash n} f_\lambda^2 (C_\lambda)^\ell, \tag{4}$$

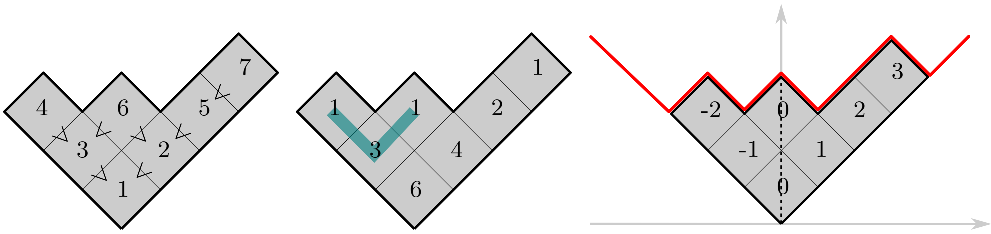

Figure 2 The Young diagram of the partition $(4,2,1) \vdash 7$ (in this paper we use the "Russian" representation where boxes are tilted by 45°). Left, its boxes are filled to produce a SYT of shape $(4,2,1)$; center, they are filled with their hook lengths, showing there are $f_{(4,2,1)} = 7!/(6\cdot4\cdot3\cdot2) = 35$ such tableaux; right, each box is filled with its content, which is the abscissa of its middle point in this representation. The content-sum is $C_{(4,2,1)} = -2-1+0+0+1+2+3 = 3$. The profile is the piecewise linear function represented here in red (coordinate axes are in grey).

where C_λ is a combinatorial quantity, namely the sum of contents of all boxes of the partition λ (see Figure 2). The RHS of this formula naturally gives rise to a certain measure on partitions, which is our main object of study:

▶ **Definition 1** (Main object). *For $n \in \mathbb{Z}_{>0}$, $\ell \in 2\mathbb{Z}_{\geq 0}$, the Plancherel–Hurwitz measure is the probability measure on partitions of n defined by*

$$\mathbb{P}_{n,\ell}(\lambda) := \frac{1}{n! H_{n,\ell}} f_\lambda^2 (C_\lambda)^\ell. \tag{5}$$

The measure $\mathbb{P}_{n,\ell}$ is left invariant by conjugation of a partition (vertical reflection of the diagrams), which sends the content-sum C_λ to its opposite. For $\ell > 0$ we will choose to work on the "positive half" of the measure, namely we let

$$\mathbb{P}^+_{n,\ell}(\lambda) := \mathbb{P}_{n,\ell}(\lambda | C_\lambda > 0) = 2 \cdot \mathbf{1}_{C_\lambda > 0} \cdot \mathbb{P}_{n,\ell}(\lambda). \tag{6}$$

A partition distributed under $\mathbb{P}_{n,\ell}$ for $\ell > 0$ can be thought of as a partition distributed under $\mathbb{P}^+_{n,\ell}$ which is reflected about a vertical axis with probability $\frac{1}{2}$.

When $\ell = 0$ the measure $\mathbb{P}_{n,\ell}$ is nothing but the Plancherel measure. Our main result deals instead with the case where ℓ grows linearly with n, corresponding to a high genus for the underlying map (see next section). The *rescaled profile* ψ_λ of a partition λ of n is the real function (piecewise linear with slope ±1) whose graph follows the contour of the diagram of λ in the coordinates of its tilted diagram representation, rescaled so that each box has area $1/n$ (see Figures 1 and 2).

▶ **Theorem 2** (Main result, see Figure 1). *Fix $\theta > 1$ and let $\lambda \vdash n$ be a random partition under the Plancherel–Hurwitz measure $\mathbb{P}^+_{n,\ell}$ in the "high-genus" regime given by $\ell = \ell(n) := 2\lfloor \theta n \rfloor$. Then, as $n \to \infty$:*

(i) *the first part λ_1 is equivalent to $\frac{2\ell}{\log n}$ (in probability)*

(ii) *the rest of the partition $\tilde\lambda = (\lambda_2, \lambda_3, \dots)$ has a VKLS limit shape. Namely, w.h.p.,*

$$\sup_x |\psi_{\tilde\lambda}(x) - \Omega(x)| \to 0, \quad \text{with} \quad \Omega(x) = \begin{cases} \frac{2}{\pi}\left(\arcsin\frac{x}{2} + \sqrt{4-x^2}\right), & |x| \leq 2 \\ |x|, & |x| > 2, \end{cases} \tag{7}$$

where $\psi_{\tilde\lambda}(x)$ is the rescaled profile of $\tilde\lambda$.

We could include λ_1 in the partition $\tilde\lambda$ of (7), since the supremum norm in this statement is insensitive to a small number of large parts. However, as Figure 1 indicates, λ_1 is the only part not scaling as $\sqrt{n}$ so we find this formulation more natural. Indeed, we can show that:

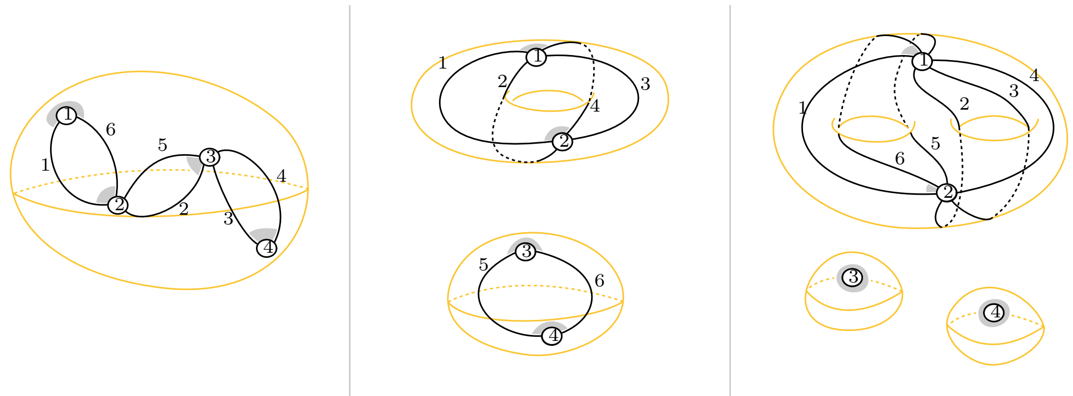

Figure 3 Three pure Hurwitz maps, each with 4 vertices, 6 edges and Euler characteristic $\chi = 0$. Left, the map corresponding to $(12)(23)(34)(34)(23)(12) = \text{id}$ is connected and has genus 0; center, the map corresponding to $(12)^4(34)^2 = \text{id}$ has two connected components, of genus 1 and 0; right, the map corresponding to $(12)^6 = \text{id}$ has three components, of genus 2, 0 and 0.

▶ **Proposition 3.** *Under the hypotheses of Theorem 2 we have*

(iii) *the second part satisfies* $\lambda_2 \leq (e + o(1))\sqrt{n}$ *w.h.p.*

We believe, as the figure strongly suggests, that the constant e can be replaced by 2 in the previous proposition. We have a strategy to prove this, but no full proof at the time of writing this extended abstract.

2 High genus maps, Hurwitz numbers, and random walks

Our original motivation to study the Plancherel–Hurwitz measure comes from the field of enumerative geometry and map enumeration. A *map* is a multigraph embedded on a compact oriented surface with simply connected faces, considered up to homeomorphisms. Equivalently it can be seen as a discrete surface, discretized by a finite number of polygons. Since the pioneering worsk of Tutte on planar maps (e.g. [18]) the enumeration of maps has proven to be particularly interesting, borrowing tools from physics, algebra and geometry and revealing their connections within combinatorics. These tools include matrix integral generating functions discovered by treating maps as Feynman diagrams [5], the topological recursion [11], and recurrence formulas based on integrable hierarchies [13]. Such exact methods have led to the asymptotic enumeration of many types of which notably exhibit a universal exponent of $-\frac{5}{2}$, and can extend to of surfaces with positive genus (e.g. [7]).

These methods do not, however, extend to maps whose genus grows with the number of polygons. This "high genus" regime is one of the most recent and exciting development in the field, due the inefficiency of existing generating-function or bijective methods which require to develop new tools. The first result in this direction was recently obtained by Budzinski and the second author [6], who showed the following estimate for the number of (connected) triangulations of size n on a surface of genus $g \sim \theta n$, by a combination of algebraic, combinatorial, and probabilistic methods:

$$T_{n,g} = n^{2g} \exp[c(\theta)n + o(n)], \quad g \sim \theta n, \tag{8}$$

where $c(\theta) > 0$ is a known continuous function. In this paper we will be interested in a different model of maps:

▶ **Definition 4** (Hurwitz map, see Figure 3). *A Hurwitz map with n vertices and ℓ edges is a map on a (non-necessarily connected) compact oriented surface, with vertices labelled from 1 to n and edges labelled from 1 to ℓ, such that the labels of edges around each vertex increase (cyclically) counterclockwise. In such a map each vertex is incident to precisely one corner which is an edge-label descent. If moreover each face of the map contains precisely one such corner, the Hurwitz map is called* pure.

It is classical, and easy to see, that Hurwitz maps of parameters n and ℓ are in bijection with tuples of transpositions $(\tau_1, \dots, \tau_\ell)$ in $\mathfrak{S}_n$, while *pure* Hurwitz maps are in bijection with tuples whose product is equal to the identity. The bijection only consists in identifying tranpositions with edges of the map, and their index with the edge-label, see Figure 3 (this construction is a special case and an adaptation of the classical construction of "constellations", see [4, 10]). The reader might find the definition of pure Hurwitz maps rather unnatural, however this model has a legitimate history in the field. In particular they are known [10], in the planar and fixed-genus cases, to belong to the same *universality class* as e.g. triangulations, quadrangulations, etc. We chose this model because among the natural models of maps, it is the one for which the connection to the Plancherel measure is the most combinatorial, and it is therefore a natural candidate to test the idea of using random partition techniques to study high genus random maps.

It is important to insist that our maps are not necessarily connected, which is an important difference with most of the literature. A pure Hurwitz map of parameters n and ℓ has necessarily n faces, and its Euler characteristic χ, its number of components κ, and its generalised genus G (sum of the genera, or number of handles, of each connected component) are related by Euler's formula:

$$\chi = \#\text{vertices} - \#\text{edges} + \#\text{faces} = 2n - \ell = 2\kappa - 2G. \tag{9}$$

This is why we call the regime $\ell \gg 2n$ the "high genus" regime.

By the above correspondence, the number $H_{n,\ell}$ introduced in (3) is the number of pure Hurwitz maps with n vertices and ℓ edges. This number is called an *unconnected Hurwitz number* in the enumerative geometry literature. As a consequence of our analysis of the Plancherel–Hurwitz measure, we obtain the following estimate:

▶ **Theorem 5** (Asymptotics of high genus unconnected Hurwitz numbers). *As in* (3)*, let $H_{n,\ell}$ be the unconnected Hurwitz number counting not necessarily connected pure Hurwitz maps with n vertices and $\ell = \ell(n) = 2\lfloor \theta n \rfloor$ edges, for $\theta > 1$. Then, as $n \to \infty$,*

$$H_{n,\ell} = \left(\frac{n}{\log n}\right)^{2\ell} \exp\left[2(\log\theta - 2)\ell + o(n)\right]. \tag{10}$$

It is tempting to see this theorem as as strong (for our model) as the Budzinski–Louf estimate (8), but unfortunately this is not quite the case. The major difference is that our maps are not necessarily connected. Moreover, we can show that (the proof is omitted in this extended abstract but follows easily from our results)

▶ **Proposition 6.** *As $n \to \infty$, a uniform random unconnected Hurwitz map with n vertices and $\ell = 2\lfloor \theta n \rfloor$ edges contains a connected component with at least $\gamma(\theta)\ell$ edges, for some function $\gamma(\theta) > 0$, and $m = o(\ell)$ vertices, w.h.p.*

The fact that the "giant" edge-component in the previous proposition has a sublinear number of vertices seems to rule out the possibility of deducing asymptotics for the *connected* linear-genus regime from our results, at least not without new ideas.

At this point it is worth commenting that in map enumeration, the regime in which the genus is unconstrained, or superlinear, is often much easier to deal with than the linear case. In fact, the Plancherel–Hurwitz measure already appears (with no name) in the "super high genus" regime $\ell > \frac{1}{2} n \log n$, in work of Diaconis and Shahshahani on the transposition random walk on $\mathfrak{S}_n$. They famously showed [8] that when $\ell \geq \frac{1+\epsilon}{2} n \log n$, the walk is strongly mixed after ℓ steps, and the proof essentially consists in showing that the Plancherel–Hurwitz measure is dominated by the trivial partition (n) in this regime. In this context, our result (10) can also be interpreted as an estimate on the return probability of the random walk after linearly many steps – much before the cut-off time, at a time when the Plancherel–Hurwitz measure still has a more subtle behaviour than the trivial partition.

Finally, we note that related measures on partitions were studied by Biane in the context of the factorisation of characters of $\mathfrak{S}_n$ [1], related to the intermediate regime $\ell = 2\lfloor \theta \sqrt{n} \rfloor$ which we do not study in this work. The limit-shape phenomena observed in this reference are different from ours. We leave the study of a possible connection, and more generally the complete study of intermediate (sublinear) values of ℓ to further works.

3 Elements of the classical Plancherel case

Formally, a partition $\lambda = (\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{\ell(\lambda)})$ of n (we write "$\lambda \vdash n$") is a weakly decreasing sequence of $\ell(\lambda)$ positive integers (called parts) which sum to n. We represent it by its Young diagram in Russian convention (Figure 2). A *Standard Young Tableau (SYT)* of shape $\lambda \vdash n$ is a filling of the boxes of its diagram with all the numbers from 1 to n which is increasing along rows and columns. The number f_λ of such tableaux can be calculated by the "hook-length" formula

$$f_\lambda = \frac{n!}{\prod_{\square \in \lambda} h_\lambda(\square)} \tag{11}$$

where the hook length $h_\lambda(\square)$ is the number of boxes in a hook going down and right from the top edge of the diagram to $\square$ and up and right to the top edge (Figure 2 again).

The number f_λ is famously equal to the dimension of the irreducible representation V^λ of the symmetric group indexed by λ, a representation theoretic connection that we have no space to develop here. We will only point out, for interested readers, that the sum of all transpositions in $\mathfrak{S}_n$ acts on this module V^λ as a scalar C_λ, which is explicitly given by the sum of contents as defined in the introduction (Figure 2 again). These two facts, together with classical representation theory, are the main reasons behind the Frobenius formula (4) expressing the count of transposition factorisations in $\mathfrak{S}_n$ with tableaux-theoretic quantities.

For $\ell = 0$ the Plancherel–Hurwitz measure becomes the Plancherel measure $\mathbb{P}_{n,0}(\lambda) = \frac{1}{n!} f_\lambda^2$ which, as said in the introduction, is very well understood.

▶ **Theorem 7** (Vershik–Kerov–Logan–Shepp (VKLS) [15, 19]). *Let $\lambda \vdash n$ be a random partition under the Plancherel measure $\mathbb{P}_{n,0}$. Then, as $n \to \infty$ we have, w.h.p.,*

$$\sup_x |\psi_\lambda(x) - \Omega(x)| = 0 \qquad \textit{and} \qquad \lambda_1 \leq 2\sqrt{n} + o(\sqrt{n}),\ \ell(\lambda) \leq 2\sqrt{n} + o(\sqrt{n}) \tag{12}$$

where $\psi_\lambda(x)$ is the rescaled profile of λ and $\Omega(x)$ is the curve defined at (7).

Several proofs exists of this limit shape result. Perhaps the simplest and most conceptual ones use the formulation of the Plancherel measure in the language of fermions and the infinite wedge space, which provides direct connection with determinantal processes [3, 14].

Such approaches and their generalisations have grown into a vast field of research after the introduction of the theory of Schur processes (see e.g. [2] for an entry point).

In the case $\ell > 0$ that we study here, it is still possible (and natural) to formulate the Plancherel-Hurwitz measure in terms of the infinite wedge, see [16]. This leads to a deep connection with integrable hierachies (the KP and 2-Toda hierarchy in particular), and even to a simple looking recurrence formula to compute the number $H_{n,\ell}$ (more precisely, their connected counterpart, see [9]). However, we do not know how to use either of these tools to approach our problems (for readers familiar with the subject: the connection to determinantal processes in presence of a sandwiched content-sum scalar operator is unclear).

Other, maybe more elementary, proofs of the VKLS theorem are based on a direct scaling of the hook-length formula (11) and variational calculus. We recommend the first chapters of the book [17] as a useful reference for such approaches. A key outcome of such an approach is the following estimate for the Plancherel measure of a partition λ in terms of its rescaled profile ψ_λ, see e.g. [17, Section 1.14]:

$$\mathbb{P}_{n,0}(\lambda) = \frac{1}{n!} f_\lambda^2 = \exp\left[-n\left(1 + 2I_{\text{hook}}(\psi_\lambda) + O\left(\frac{\log n}{\sqrt{n}}\right)\right)\right] \tag{13}$$

where $I_{\text{hook}}(\cdot)$ is an "energy" functional defined by an explicit integral formula. The VKLS function $\Omega(x)$ previously introduced is the unique continuous function satisfying $\int(\Omega(x) - |x|)dx = 1$ which minimises $I_{\text{hook}}(\Omega)$ (see e.g. [17, Section 1.17]). This implies the limit shape part of the VKLS theorem, since any partition whose profile is "far" from $\Omega(x)$ will appear with an exponentially small probability.

The upper bound on the first part λ_1 in the VKLS theorem does not directly follow from this limit shape analysis. Classical proofs usually depend either on the RSK algorithm or on the random growth process. The fact that neither of these tools exist in the context of factorisations ($\ell > 0$) will make our proofs become harder, see comments in the next section.

4 Proofs of our results

We will now sketch the proofs of Theorems 5 and 2, and Proposition 3. Throughout this section, we consider λ to be a random partition of n distributed by the Plancherel–Hurwitz measure $\mathbb{P}^+_{n,\ell}$ with $\ell = \ell(n) = 2\lfloor \theta n \rfloor$. Heuristically, a random partition under $\mathbb{P}^+_{n,\ell}$ is driven by two different "forces":

1. on the one hand, the "Plancherel entropy": the estimate (13) shows that there is an exponential cost for the partition, in terms of the Plancherel factor f_λ^2, to deviate from the VKLS shape.
2. on the other hand, the "content-sum entropy": the factor $(C_\lambda)^\ell$ can itself become exponentially high, so the partition may prefer to deviate from VKLS if this leads to a much higher content-sum.

Our main theorem shows that the best way for the partition to adapt to this situation, is to "throw" all its contribution to a large content-sum in the first part λ_1, and that after this the rest of the partition maximises the entropy classically. We establish this fact by successive refinements, in several steps.

We now go through the proofs. We will use the notation $Z_n(\Lambda) = \frac{1}{n!}\sum_{\lambda\in\Lambda} f_\lambda^2 (C_\lambda)^\ell$ for any set Λ of partitions of n, such that the partition function of our model is $H_{n,\ell} = Z_n(\{\lambda \vdash n\}) = \frac{1}{n!}\sum_{\lambda\vdash n} f_\lambda^2 (C_\lambda)^\ell$. We also fix $\varepsilon = \frac{1}{100}$ and split any partition $\lambda \vdash n$ into $\lambda = \lambda^+ \sqcup \lambda^-$ where λ^+ denotes the parts that are greater than $n^{1-\varepsilon}$ and λ^- the parts that are less than $n^{1-\varepsilon}$, see Figure 4. We will use the following immediate and convenient bounds.

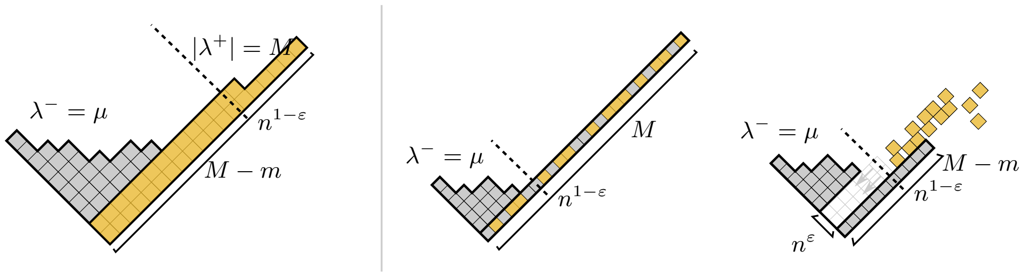

Figure 4 Left, a partition $\lambda \vdash n$ in $\Lambda(\mu, M, m)$, with λ^+ and λ^- indicated. Right, a SYT of shape $\lambda^0 \in \Lambda(\mu, M, 0)$ (the filling of the boxes is not shown) is transformed to a SYT of some shape $\lambda \in \Lambda(\mu, M, m)$ or to something else by the surjective operation used to prove Claim 13.

▶ **Lemma 8** (Useful bounds). *Let* $\lambda \vdash n$ *with* $\lambda^+ = (\lambda_1, \dots, \lambda_p)$*, then*

(i) $\frac{1}{n!} f_\lambda^2 \leq \frac{n^{|\lambda^+|}}{\prod_{i=1}^{p} (\lambda_i!)^2}$,

(ii) $C_\lambda \leq \frac{\lambda_1 n}{2}$,

(iii) $C_\lambda = \sum_{i=1}^{p} \left(\frac{\lambda_i(\lambda_i - 1)}{2} - (i-1)\lambda_i \right) - p|\lambda^-| + C_{\lambda^-}$.

We now proceed with the succession of lemmas that constitutes the core of our proof.

▶ **Lemma 9** (Bounding the partition function below). *We have*

$$H_{n,\ell} \geq \exp\left[2\ell(\log \ell - \log\log n) - \ell(2 - \log 2) + o(n)\right]. \tag{14}$$

Proof. Let $L := \frac{2\ell}{\log n}$ and $\lambda^* = L \sqcup \mu$ with μ maximising f_λ among all partitions of $n - L$ with non-negative content-sum. Using Lemma 8(iii) we have

$$C_{\lambda^*} = \frac{L(L-1)}{2} - |\mu| + C_\mu$$

from which it is not difficult to show that

$$Z_n(\{\lambda^*\}) \geq \exp\left[2\ell(\log \ell - \log\log n) - \ell(2 - \log 2) + o(n)\right]. \tag{15}$$

and this finishes the proof since $H_{n,\ell} = Z_n(\{\lambda \vdash n\}) \geq Z_n(\{\lambda^*\})$. ◀

The following lemma controls the contribution of "big parts" λ^+ in a Plancherel–Hurwitz random partition. The "truncation" threshold $n^{0.99}$ is somewhat arbitrary at this stage and will be improved to $O(\sqrt{n})$ at the very end of our analysis.

Throughout the following, let λ be a random partition under the Plancherel–Hurwitz measure $\mathbb{P}^+_{n,\ell}$ at high genus, with $\ell = 2\lfloor \theta n \rfloor$.

▶ **Lemma 10** (Controlling big parts). *W.h.p., we have* $|\lambda^+| \in [\frac{1}{2}L, \frac{5}{2}L]$ *where* $L = \frac{2\ell}{\log n}$.

Proof. Let $R_\lambda = \frac{|\lambda^+| \log n}{\ell}$. For all $\lambda \vdash n$, by Lemma 8(i),

$$\frac{1}{n!} f_\lambda^2 \leq \exp(-(1 - 2\varepsilon) R_\lambda \ell + 2|\lambda^+| + o(n)). \tag{16}$$

On the other hand, by Lemma 8(ii)-(iii), if $C_\lambda \geq 0$, then

$$C_\lambda^\ell \leq \exp(2\ell(\log \ell - \log\log n) + \ell(\log(R_\lambda^2 + \frac{n^{2-\varepsilon} \log^2 n}{\ell^2}) - \log 2)). \tag{17}$$

Combining (16) and (17), and using (14), we obtain

$$\frac{Z_n(\{\lambda\})}{H_{n,\ell}} \leq \exp\left[\ell(2-(1-2\varepsilon)R_\lambda + \log\left(R_\lambda^2 + \frac{n^{2-\varepsilon}\log^2 n}{\ell^2}\right)\right] \tag{18}$$

hence for n large enough and $\lambda \vdash n$ with $R_\lambda \notin [1,5]$, $\mathbb{P}^+_{n,\ell}(\lambda) \leq \exp(-\ell/100)$, which entails the result since there are $e^{O(\sqrt{n})}$ partitions of n. ◀

▶ **Lemma 11** (Uniqueness of the big part). *W.h.p.,* $\lambda^+ = (\lambda_1)$.

The proof of Lemma 11 requires to compare the contribution of partitions having a single big part, to those having more than one (indeed, because we do not have exact formulas nor precise estimates on our partition functions, we can only rely on "comparison" of probabilities at this stage). We will perform this comparison among partitions having the same set of "small parts" (called μ below).

For non-negative integers M, m and partitions $\mu \vdash n - M$, we let $\Lambda(\mu, M) = \{\lambda||\lambda^+| = M, \lambda^- = \mu\}$ and $\Lambda(\mu, M, m) = \{\lambda \in \Lambda(\mu, M)|\lambda_1 = M - m\}$. We also use the notation $\lambda^0 = M \sqcup \mu$ so that $\Lambda(\mu, M, 0) = \{\lambda^0\}$. We will need the following two claims, whose proof is postponed to after that of the lemma.

▷ Claim 12. For all $\lambda \in \Lambda(\mu, M, m)$, we have $C_\lambda \leq C_{\lambda^0} - (m-1)\frac{M}{2}$.

▷ Claim 13. If $m > 0$ then, $\sum_{\lambda\in\Lambda(\mu,M,m)} f_\lambda \leq f_{\lambda^0} \exp[m(2\varepsilon \log n + 1)]$.

Proof of Lemma 11. By Lemma 10, we know that, w.h.p., $|\lambda^+| \in [\frac{\ell}{\log n}, 5\frac{\ell}{\log n}]$. We can thus assume this event for the rest of this proof.

We now condition on $|\lambda^+| = M$ and $\lambda^- = \mu$, with given $M \in [\frac{\ell}{\log n}, 5\frac{\ell}{\log n}]$ and $\mu \vdash n - M$. Combining Claims 12 and 13 for $m > 0$, one obtains

$$\frac{Z_n(\Lambda(\mu, M, m))}{Z_n(\{\lambda^0\})} \leq \exp\left[\ell \log\left(1 - \frac{(m-1)M}{2C_{\lambda^0}}\right) + 2m(2\varepsilon \log n + 1)\right]. \tag{19}$$

But we know that $C_{\lambda^0} \leq (1 + o(1))\frac{M^2}{2}$ and $M \leq 5\frac{\ell}{\log n}$. Hence

$$\frac{Z_n(\Lambda(\mu, M, m))}{Z_n(\{\lambda^0\})} \leq \exp\left[-\frac{m \log n}{100}\right] \tag{20}$$

Summing this over all $m > 0$ (recall that $m \geq n^{1-\varepsilon}$ if the set is non-empty), we have

$$\sum_{m>0} Z(\Lambda(\mu, M, m)) = o(Z(\{\lambda^0\})) \tag{21}$$

which is enough to conclude that $\lambda^+ = (\lambda_1)$ w.h.p. ◀

Proof of the claims. The first claim is direct. For the second one, we need to define a proper "redistribution" operation that enables us to compare the contribution of partitions with one big part to others. To do this, we will describe an operation taking as input a SYT of shape λ^0 plus some information, and outputting a SYT of some $\lambda \in \Lambda(\mu, M, m)$, or something else, such that this operation is surjective on $\Lambda(\mu, M, m)$.

Input: A SYT T of shape λ^0.

1. Create n^ε empty rows between the first row of T and the rest,
2. choose m numbers in the first row of T *($\binom{M}{m}$ choices),*
3. for each of these numbers, choose one of the newly created rows, and move it there *(n^ε choices each time),*

4. sort each row and delete the empty rows, output the result.

It is easily checked that this procedure can output any SYT of λ for any $\lambda \in \Lambda(\mu, M, m)$ (indeed, for such a λ, λ^+ must have at most $\frac{n}{n^{1-\varepsilon}} = n^\varepsilon$ rows). Hence we have

$$\sum_{\lambda \in \Lambda(\mu,M,g)} f_\lambda \leq \binom{M}{m} n^{\varepsilon g} f_{\lambda^0} \leq f_{\lambda^0} \exp(m(2\varepsilon \log n + 1) \tag{22}$$

where in the last inequality we used Stirling's approximation along with the facts that $\log M \leq \log n$ and $\log m \geq (1-\varepsilon)\log n$. ◁

We can now collect the fruits of the previous lemmas to obtain our main theorems.

Proof of Theorem 5. The previous lemmas imply that for a Plancherel–Hurwitz distributed partition λ, we have w.h.p. $C_\lambda = (1+o(1))\frac{\lambda_1^2}{2}$. On the other hand, we have $\frac{1}{n!} f_\lambda^2 \leq \frac{n!}{(\lambda_1!)^2 (n-\lambda_1)!}$, hence

$$Z_n(\{\lambda\}) \leq \exp\left[2\ell \log(\lambda_1) - \ell \log 2 - \lambda_1 \log n + o(n)\right]. \tag{23}$$

Now we substitute $\lambda_1 = \frac{R_\lambda \ell}{\log n}$ in the inequality above, and we obtain

$$Z_n(\{\lambda\}) \leq \left(\frac{n}{\log n}\right)^{2\ell} \exp\left[2(\log\theta - 2)\ell\right] \exp\left[\ell(2(\log R_\lambda - \log 2) + 2 - R_\lambda) + o(n)\right].$$

Now, since for $x > 0$ we always have $2(\log x - \log 2) + 2 - x \leq 0$, we get

$$Z(\{\lambda\}) \leq \left(\frac{n}{\log n}\right)^{2\ell} \exp\left[2(\log\theta - 2)\ell + o(n)\right].$$

This, together with the lower bound of Lemma 9, proves Theorem 5 since there are $e^{O(\sqrt{n})}$ partitions of n. ◀

Proof of Theorem 2, part (i). The last argument of the previous proof also implies that

$$\mathbb{P}_{n,\ell}(\lambda) \leq \exp\left[\ell(2(\log R_\lambda - \log 2) + 2 - R_\lambda) + o(n)\right]. \tag{24}$$

Now, the function on positive reals $x \longmapsto 2(\log x - \log 2) + 2 - x$ has a unique maximum at $x = 2$. Any non-negligible deviation of R_λ from this maximum thus entails an exponentially decreasing probability, which is enough to conclude that $R_\lambda = 2 + o(1)$ w.h.p. ◀

Proof of Theorem 2, part (ii). The previous discussions imply that, w.h.p., $C_\lambda = (1 + o(1))2\left(\frac{\ell}{\log n}\right)^2$ and $f_\lambda = \binom{n}{\lambda_1} f_{\tilde\lambda} e^{o(n)}$, which, by Theorem 5 and the Plancherel entropy estimate (13), lead to

$$\mathbb{P}_{n,\ell}(\lambda) \leq \exp\left(2n(I_{\text{hook}}(\psi_\lambda) - I_{\text{hook}}(\Omega)) + o(n)\right). \tag{25}$$

This implies, as in the classical Plancherel case (see [17, Section 1.17]), the almost sure convergence in supremum norm to the VKLS limit shape. ◀

It only remains to prove Proposition 3, i.e. to upper bound the size of λ_2. As we said already, the VKLS limit shape result in supremum norm does not imply such a bound, and even in the Plancherel case extra arguments are needed. We find convenient here to refer again to Romik's book where two bounds are given for the largest part in the Plancherel regime:

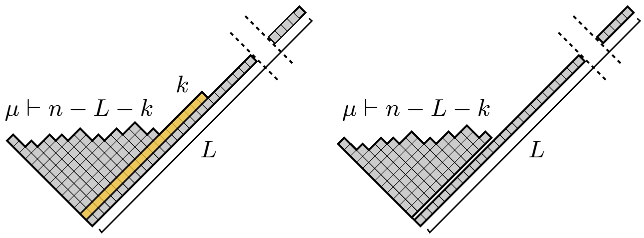

Figure 5 Partitions $L \sqcup k \sqcup \mu \vdash n$ and $L \sqcup \mu \vdash n-k$.

- an elementary bound, based on a first moment calculation and the RSK algorithm, which is enough to establish a bound of the form $(e + o(1))\sqrt{n}$ ([17, Lemma 1.4]).
- a more sophisticated bound based on the Cauchy–Schwartz inequality and on the existence of the corner-growth process for the Plancherel measure, which leads to the sharp bound $(2 + o(1))\sqrt{n}$ ([17, Section 1.19]).

In our context, we unfortunately do not have the analogue of the RSK algorithm (see next section), let alone of the corner growth process. The proof below mimics the first moment argument of the classical proof at the level of tableaux, and together with previous estimates on the partition functions enable us to reach the bound $(e + o(1))\sqrt{n}$. A subtler approach which tries to mimic the corner growth process as in the second proof should enable us to attain soon the (conjectured) bound $(2 + o(1))\sqrt{n}$ in which case it will appear in the journal version of this paper.

Proof of Proposition 3. Under the Plancherel-Hurwitz measure, if we condition on the first part being $\lambda_1 = \frac{2\ell}{\log n} = L$, the distribution of the second part is

$$\mathbb{P}(\lambda_2 = k | \lambda_1 = L) = \frac{1}{n! Z_n[L]} \sum_{\mu \vdash n-L-k} f^2_{L \sqcup k \sqcup \mu} C^{\ell}_{L \sqcup k \sqcup \mu}. \tag{26}$$

where $Z_n[L] \equiv Z_n(\{\lambda \vdash n | \lambda_1 = L\})$. Comparing SYT of shape $L \sqcup k \sqcup \mu \vdash n$ with ones of shape $L \sqcup \mu \vdash n-k$, obtained by removing the second part, and the contents of the partitions, we have

$$f_{L \sqcup k \sqcup \mu} \leq \binom{n}{k} f_{L \sqcup \mu}, \qquad C_{L \sqcup k \sqcup \mu} = C_{L \sqcup \mu} - |\mu| + \frac{k^2}{2} = C_{L \sqcup \mu}(1 + o(1)) \tag{27}$$

and from there we obtain

$$\mathbb{P}(\lambda_2 = k | \lambda_1 = L) = \binom{n}{k}^2 \frac{(n-k)!}{n!} \frac{Z_{n-k}[L]}{Z_n[L]} (1 + o(1)). \tag{28}$$

Now, following an application of the identity $n f_\mu = \sum_{\nu : \mu \nearrow \nu} f_\nu$ for $\mu \vdash n$, where "$\mu \nearrow \nu$" means that ν is obtained from μ by adding one box, and using elementary bounds on the variation of the content-sum when a single box is added, it is possible to show that

$$Z_n[L] = Z_{n-1}[L] e^{o(1)}. \tag{29}$$

It follows that

$$\mathbb{P}(\lambda_2 = k | \lambda_1 = L) = \frac{n!}{k!^2 (n-k)!} e^{o(k)} \leq \frac{n^k}{(k/e)^{2k}} e^{o(k)}, \tag{30}$$

and, to conclude the proof,

$$\forall \varepsilon > 0, \qquad \lim_{n \to \infty} \mathbb{P}(\lambda_2 = (1 + \varepsilon) e \sqrt{n} | \lambda_1 = L) = 0 \tag{31}$$

and we have $\lambda_2 \leq e(1 + o(1))\sqrt{n}$ w.h.p. as $n \to \infty$. ◀

5 Open questions and perspectives

Maybe the main open question that follows our work is the following: *does there exist an analogue of the RSK algorithm proving combinatorially the identity* (4)*?* If this is the case, then our results about λ_1 and λ_2 probably translate into distributional limit theorems for certain parameters of random factorisations (or random pure Hurwitz maps). To start with, can one identify the "meaning" of the statistic λ_1 on the Hurwitz side?

Another question is, of course, to know if one can use the Plancherel–Hurwitz approach to say anything about *connected* Hurwitz maps of high genus. This would be very interesting. It may also be interesting to combine this approach with the technology of integrable hierarchies, which have been so fruitful but have so far not directly led to precise asymptotic estimates nor limit theorems for connected random maps or Hurwitz numbers of high genus.

References

1. P. Biane. Approximate factorization and concentration for characters of symmetric groups. *International Mathematics Research Notices*, 2001(4):179–192, 01 2001.
2. A. Borodin and V. Gorin. Lectures on integrable probability. In *Probability and Statistical Physics in St. Petersburg*, volume 91 of *Proceedings of Symposia in Pure Mathematics*, pages 155–214. AMS, 2016. `arXiv:1212.3351`.
3. A. Borodin, A. Okounkov, and G. Olshanski. Asymptotics of Plancherel measures for symmetric groups. *J. Amer. Math. Soc.*, 13(3):481–515, 07 2000.
4. M. Bousquet-Mélou and G. Schaeffer. Enumeration of planar constellations. *Advances in Applied Mathematics*, 24(4):337–368, 2000.
5. E. Brézin, C. Itzykson, G. Parisi, and J. B. Zuber. Planar diagrams. *Communications in Mathematical Physics*, 59(1):35 – 51, 1978.
6. T. Budzinski and B. Louf. Local limits of uniform triangulations in high genus. *Inventiones mathematicae*, 223, 01 2021.
7. Guillaume Chapuy, Michel Marcus, and Gilles Schaeffer. A bijection for rooted maps on orientable surfaces. *SIAM J. Discrete Math.*, 23(3):1587–1611, 2009.
8. P. Diaconis and M. Shahshahani. Generating a random permutation with random transpositions. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 57:159–179, 1981.
9. B. Dubrovin, D. Yang, and D. Zagier. Classical hurwitz numbers and related combinatorics. *Moscow Mathematical Journal*, 17:601–633, 10 2017.
10. E. Duchi, D. Poulalhon, and G. Schaeffer. Bijections for simple and double hurwitz numbers, 2014. `arXiv:1410.6521`.
11. B. Eynard and N. Orantin. Invariants of algebraic curves and topological expansion. *Communications in Number Theory and Physics*, 1, 03 2007.
12. W. Fulton. *Young Tableaux: With Applications to Representation Theory and Geometry*. London Mathematical Society Student Texts. Cambridge University Press, 1996.
13. I.P. Goulden and D.M. Jackson. The KP hierarchy, branched covers, and triangulations. *Advances in Mathematics*, 219(3):932–951, 2008.
14. K. Johansson. The longest increasing subsequence in a random permutation and a unitary random matrix model. *Mathematical Research Letters*, 5:63–82, 01 1998.
15. B.F. Logan and L.A. Shepp. A variational problem for random Young tableaux. *Advances in Mathematics*, 26(2):206–222, 1977.
16. A. Okounkov. Toda equations for Hurwitz numbers. *Math. Res. Lett.*, 7(4):447–453, 2000.
17. Dan Romik. *The Surprising Mathematics of Longest Increasing Subsequences*. Institute of Mathematical Statistics Textbooks. Cambridge University Press, 2015.
18. W. T. Tutte. A census of planar maps. *Canadian Journal of Mathematics*, 15:249–271, 1963.
19. A.M. Vershik and S.V. Kerov. Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tableaux. *Doklady Akademii Nauk*, 233(6):1024–1027, 1977.