Limit shapes for Gibbs partitions of sets

Ibrahim Fatkullin
University of Arizona

Jianfei Xue
University of Missouri

This study extends a prior investigation of limit shapes for grand canonical Gibbs ensembles of partitions of integers, which was based on analysis of sums of geometric random variables. Here we compute limit shapes for partitions of sets, which lead to the sums of Poisson random variables. Under mild monotonicity assumptions on the energy function, we derive all possible limit shapes arising from different asymptotic behaviors of the energy, and also compute local limit shape profiles for cases in which the limit shape is a step function.

I. Introduction

In this work we extend the methods introduced in [9] to analyze limit shapes for Gibbs ensembles of partitions of integers to partitions of sets. The primary difference is that the former lead to sums of geometric random variables, while the latter lead to sums of Poisson random variables. Analysis carried out here applies also in a variety of contexts related to measures on permutations and in other combinatorial contexts [8, 3, 7, 6, 2, 15].

We use the same notation and definitions as in [9], which may be consulted for additional information. A sequence of non-negative integers, $\boldsymbol{p} = (p_k)$, $k \in \mathbb{N}$, corresponds to a partition of an integer

$$M = \mathscr{M}(\boldsymbol{p}) := \sum_{k=1}^{\infty} k p_k. \tag{1}$$

We say that $\boldsymbol{p}$ partitions M (provided the sum above is finite) and denote this by $\boldsymbol{p} \vdash M$. The function $\mathscr{M}(\boldsymbol{p})$ is called the *mass* of the partition $\boldsymbol{p}$. For a given partition, p_k-s represent the numbers of summands of size k. For example, the partition $21 = 1+2+2+2+4+4+6$ corresponds to $\boldsymbol{p} = (1, 3, 0, 2, 0, 1, 0\ldots)$; respectively, $\mathscr{M}(\boldsymbol{p}) = 21$. We casually employ the polymer physics language referring to summands in a partition as *polymers* and to the individual units as *monomers*. Thus the mass of a partition may also be thought of as the total number of monomers in the corresponding polymeric system.

I.i. Gibbs ensembles of set partitions

Partitions of sets and integers are related: a sequence $\boldsymbol{p} = (p_k)$ corresponds to a unique partition of the integer $M = \mathcal{M}(\boldsymbol{p})$ and also to

$$\frac{M!}{\prod_{k=1}^{\infty}(k!)^{p_k}p_k!} \tag{2}$$

different partitions of a set with cardinality M. Partitions of integers correspond to *indistinguishable* monomers, while partitions of set — to *distinguishable* monomers. The combinatorial factor in formula (2) is exactly the number of ways to distribute M labels among unlabeled units in a partition of M. It appears as a background weight when we introduce measures on partitions of sets. See, e.g., the treatise by J. Pitman [14] for a discussion on related topics.

Gibbs measures on partitions are characterized by the weights, $\mathrm{e}^{-\beta\mathcal{H}(\boldsymbol{p})}$, where $\mathcal{H}(\boldsymbol{p})$ is the energy (Hamiltonian) of a partition $\boldsymbol{p}$, and β is the inverse temperature of the system. We consider energies of the form,

$$\mathcal{H}(\boldsymbol{p}) = \sum_{k=1}^{\infty} E_k p_k, \tag{3}$$

where the numbers E_k represent the internal energies of polymers of size k, $k \in \mathbb{N}$. Thus the total energy of our system is equal to the sum of the individual energies of all polymers.

Canonical Gibbs measures are defined on partitions of sets with prescribed cardinalities, or alternatively, on sequences $\boldsymbol{p}$, such that $\mathcal{M}(\boldsymbol{p}) = M$. These measures are prescribed by the following probabilities:

$$\mathbb{P}_{M,\beta}\{\boldsymbol{p} = \boldsymbol{P}\} = \frac{1}{Q_{M,\beta}}\, \mathrm{e}^{-\beta\mathcal{H}(\boldsymbol{P})} \prod_{k=1}^{\infty} \frac{1}{(k!)^{P_k}P_k!}; \qquad Q_{M,\beta} = \sum_{\boldsymbol{p} \vdash M} \prod_{k=1}^{\infty} \frac{\mathrm{e}^{-\beta\mathcal{H}(\boldsymbol{P})}}{(k!)^{P_k}P_k!}. \tag{4}$$

The grand canonical Gibbs measures are defined on all partitions with finite mass ($\mathcal{M}(\boldsymbol{p}) < \infty$) by prescribing

$$\mathbb{P}_{\mu,\beta}\{\boldsymbol{p} = \boldsymbol{P}\} = \frac{1}{Q_{\beta}(\mu)}\, \mathrm{e}^{-\beta\mathcal{H}(\boldsymbol{P})-\mu\mathcal{M}(\boldsymbol{P})} \prod_{k=1}^{\infty} \frac{1}{(k!)^{P_k}P_k!}; \qquad Q_{\beta}(\mu) = \sum_{M=0}^{\infty} Q_{M,\beta}\, \mathrm{e}^{-\mu M}. \tag{5}$$

These measures are superpositions of the canonical measures $\mathbb{P}_{M,\beta}$ with weights proportional to $\mathrm{e}^{-\mu M}$. The parameter μ is called *chemical potential*. It regulates the expected total mass of the system, $\mathbb{E}_{\mu,\beta}\,\mathcal{M}(\boldsymbol{p})$.

As noted by A. Vershik [16], the grand canonical measures are multiplicative:

$$\mathbb{P}_{\mu,\beta}\{\boldsymbol{p} = \boldsymbol{P}\} = \prod_{k=1}^{\infty} \mathbb{P}_{\mu,\beta}^{(k)}\{p_k = P_k\}; \qquad \mathbb{P}_{\mu,\beta}^{(k)}\{p_k = N\} = \mathrm{e}^{-\alpha_k}\frac{\alpha_k^N}{N!}, \qquad \alpha_k = \frac{\mathrm{e}^{-\beta E_k - \mu k}}{k!}. \tag{6}$$

This implies that p_k-s are independent Poisson random variables with parameters α_k. This is the principal difference between the partitions of sets and integers: the latter induce measures for which p_k-s are geometric random variables, cf. [9]. Note that in some cases (see Section II.iv) these multiplicative measures may be supported on sequences for which $\mathcal{M}(\boldsymbol{p})$ is infinite, such sequences do not correspond to partitions of any finite integers or sets.

Equivalence of ensembles and the thermodynamic limit. One of the central problems in statistical mechanics of partitions is to understand various features of the canonical measures in the limit when the mass of the system, M, tends to infinity. The grand canonical measures were originally introduced as an aid in this task. Their utility is due to the fact that p_k-s (which are not independent in the canonical setting) become independent in the grand canonical setting. Equation $M = \mathbb{E}_{\mu,\beta}\mathcal{M}$ provides the correspondence between M (canonical ensembles) and μ (grand canonical ensembles). Consequently, instead of the limit as $M \to \infty$, one considers the so-called *thermodynamic limit*, $\mu \searrow \mu_*$, where μ_* delimits the interval of convergence of the series $\sum_{k=1}^{\infty} k\alpha_k$ representing the expected number of monomers in the system. Notice, that as implied by formula (6), α_k-s are decreasing functions of μ, thus the limit $\mu \to \mu_*$ is always taken from above; see Section I.iv for more details regarding the value of μ_*.

Yu. Yakubovich [19] and A. Vershik [16] established that as far as the distributions of the (appropriately rescaled) p_k-s are concerned, the canonical and grand canonical measures are equivalent in respective limits as $M \to \infty$ and $\mu \searrow \mu_*$, provided the latter has a limit shape (see below). In this work we only study the grand canonical measures and whenever we omit the subscripts μ and β, we imply that the quantities in question are computed with respect to a grand canonical measure with corresponding parameters. Results regarding canonical measures may then be deduced whenever the equivalence of ensembles holds.

I.ii. Limit shapes

Following A. Vershik [16], we define the *size distribution function* of a partition as

$$f(x;\boldsymbol{p}) := \sum_{k \geq x} p_k; \qquad x \in \mathbb{R}^+ := [0,\infty). \tag{7}$$

The graph of this function is the boundary of a Young (Ferrers) diagram corresponding to the partition $\boldsymbol{p}$. The rescaled size distribution function is defined as

$$F_\mu(x;\boldsymbol{p}) := \frac{\varkappa}{\mathbb{E}\mathcal{M}} f(\varkappa x;\boldsymbol{p}) = \frac{\varkappa}{\mathbb{E}\mathcal{M}} \sum_{k \geq \varkappa x} p_k; \qquad x \in \mathbb{R}^+. \tag{8}$$

The factor $\mathbb{E}\mathcal{M}$ is chosen to make $\mathbb{E}F_\mu(x;\boldsymbol{p})$ integrate to one; the parameter $\varkappa = \varkappa(\mu)$ controls the scaling in the "horizontal" direction. Scaling in the "vertical" direction is then automatically determined as $\mathbb{E}\mathcal{M}/\varkappa$. The dependence $\varkappa(\mu)$ must be carefully chosen to obtain a sensible behavior for $F_\mu(x;\boldsymbol{p})$ in the thermodynamic limit. Whenever, with a proper choice of scaling, $F_\mu(x;\boldsymbol{p})$ converges (in an appropriate sense) to some nonzero deterministic function,

$$F(x) := \lim_{\mu \searrow \mu_*} F_\mu(x;\boldsymbol{p}), \tag{9}$$

the latter is called *limit shape* corresponding to the energies E_k. Its graph is the limiting curve (in probability) for boundaries of the rescaled Young diagrams for respective Gibbs partitions.

As p_k are independent Poisson random variables with parameters α_k, see formula (6), their size distribution functions are also Poisson-distributed, with parameters equal to the sums of the corresponding α_k-s. Therefore, such quantities as the expected mass of the partition, expectation

3

and variance of the rescaled size distribution function, are given by the following sums:

$$\mathbb{E}\mathcal{M} = \sum_{k=1}^{\infty} k\alpha_k; \qquad \mathbb{E}F_\mu(x;\boldsymbol{p}) = \frac{\varkappa}{\mathbb{E}\mathcal{M}} \sum_{k \geq \varkappa x} \alpha_k; \qquad \mathbb{V}F_\mu(x;\boldsymbol{p}) = \frac{\varkappa}{\mathbb{E}\mathcal{M}} \mathbb{E}F_\mu(x;\boldsymbol{p}). \qquad (10)$$

Thus this work is essentially a study of asymptotic behaviors of such sums as μ tends to μ_* — the boundary of their interval of convergence, see equation (13) below.

I.iii. Known results regarding limit shapes for partitions of sets

Most of the classical limit shape studies are motivated by problems in representation theory of the symmetric group, or by various combinatorial constructions; see e.g., [4]. In particular, one result relevant to our work concerns the measure induced on partitions by the uniform (Dirichlet-Haar) measure on the symmetric group via its cycle structure. This measure was analyzed in great detail by J. F. C. Kingman [13], and A. Vershik and A. Shmidt [17, 18]. It corresponds to setting $\alpha_k = \mathrm{e}^{-\mu k}/k$ in formula (6). There is no limit shape in this case. Instead, the distributions of cycle lengths converge (after appropriate transformation) to a Poisson-Dirichlet or a closely related Griffiths-Engen-McCloskey (GEM) distribution [12, 10, 1].

Ensembles represented by formula (6) may also be obtained by assigning length-dependent weights to the cycles of a permutation (rather than to the members of a partition). V. Betz, D. Ueltschi, and Y. Velenik [3]; or N. Ercolani and D. Ueltschi [7] computed various statistical quantities, such as the expected cycle length, or total number of (finite) cycles for several specific ensembles in this context.

Another result, obtained by Yu. Yakubovich [19], concerns the uniform measure on partitions of sets. This corresponds to $\alpha_k = \mathrm{e}^{-\mu k}/k!$ or $\beta = 0$ in our terminology. In this case the limit shape is given by the step function, $\mathbb{1}_{[0,1]}(x)$. The appropriate scaling is $\varkappa(\mu) = \mathrm{e}^{-\mu}$.

The so-called *expansive* case, corresponding in our notation to $\alpha_k \sim k^{d-1}\mathrm{e}^{-\mu k}$ with $d > 0$ was addressed directly in the context of canonical Gibbs ensembles by M. Erlihson and B. Granovsky [8] and also generalized by A. Cipriani and D. Zeindler [6]. In this case the thermodynamic limit is attained as $\mu \searrow 0$ and the appropriate scaling is $\varkappa = 1/\mu$. The limit shape is given by the incomplete gamma function, $\Gamma(x;d)$.

In this work we fill the gaps in existing results and establish a complete classification of possible asymptotic behaviors of the scaled size distribution functions for all possible asymptotic behaviors of the energies E_k under mild monotonicity assumptions. In particular, we determine for which E_k-s one gets a limit process rather than a limit shape, when the limit shape is given by the step function, or by the incomplete gamma function without any assumptions regarding specific functional form of the asymptotic behavior of E_k-s (or α_k-s). For scenarios in which the limit shape is given by the step function as in formula (15), we also study its local profile in the vicinity of the discontinuity point. Our methods are based on a mixture of direct probabilistic and analytical techniques similar to those utilized in [5, 9] and are somewhat simpler and more elementary than the methods based on saddle-point-type techniques for generating functions, as in the aforementioned studies.

I.iv. Informal statement of main results

In this section we summarize our main results in an informal manner. As formula (6) suggests, the $k!$ factor may be absorbed from the denominator of α_k-s into the exponential, and thus we introduce

$$u(k) := \beta E_k + \ln \Gamma(k+1). \tag{11}$$

It is more straightforward to work with $u(k)$ rather than E_k-s, and we employ it in most of our proofs and calculations. We assume that the energies may be extended from $k \in \mathbb{N}$ into $\mathbb{R}^+$ in a smooth and monotone manner:

Assumption 1   There exists a twice differentiable function $u : \mathbb{R}^+ \to \mathbb{R}$ satisfying equation (11), such that $\lim_{x \to \infty} x^2 u''(x)$ exists in $\mathbb{R} \cup \{-\infty, +\infty\}$.

This assumption is more technical than conceptual. The differentiability allows us to use the mean value theorem at various points in the proofs; it may always be achieved, as any function defined on integers can always be smoothly extended to reals. The condition on $u''(x)$ allows us to avoid dealing with lim-sup/infs at various stages in the proofs. It is not entirely essential, however: if it fails, there may exist subsequences $\{k_i\}$ along which the asymptotic behaviors of p_{k_i}-s differ. As all p_k-s are independent in the grand canonical ensembles, one could analyze all such subsequences separately using the same methods as presented here.

The thermodynamic limit as attained as μ approaches μ_*, a value such that $\lim_{\mu \searrow \mu_*} \mathbb{E}\mathscr{M} = \infty$. Employing formula (11), we can express the expected number of monomers as

$$\mathbb{E}\mathscr{M} = \sum_{k=1}^{\infty} k\, e^{-\mu k - u(k)}. \tag{12}$$

Using the root test, we can conclude that this sum converges for all $\mu > \mu_*$ and diverges for all $\mu < \mu_*$, where

$$\mu_* := -\lim_{x \to \infty} \frac{u(x)}{x} = -\lim_{x \to \infty} u'(x). \tag{13}$$

In particular, if $\mu_* = -\infty$, $\mathbb{E}\mathscr{M}$ is finite for all μ, while if $\mu_* = \infty$, $\mathbb{E}\mathscr{M}$ is always infinite. Notice that Assumption 1 allows us to use a regular limit instead of lim-inf in equation (13), while Lemmas A.1 and A.2 imply the second equality.

As discussed earlier in Section I.ii, the limit shapes appear provided that the sizes of Young diagram cells are appropriately rescaled as $\mu \searrow \mu_*$. In most cases, the required choice of the horizontal scaling factor is $\varkappa = \varkappa(\mu)$, the largest solution of

$$u'(\varkappa) = -\mu. \tag{14}$$

One can see from the formula (12), that this value of k corresponds to the maximum of the expression inside of the exponent, i.e., to the range of the most likely polymer sizes. In scenarios when $u''(x) > 0$ for all large enough x, e.g. when $x \in (X, \infty)$ for some value X, this implies that as $\mu \searrow \mu_*$, equation (14) has a unique solution in (X, ∞) and thus its largest solution is well-defined. In the other cases the appropriate scaling is $\varkappa = 1/(\mu - \mu_*)$ see below for details.

SUBCRITICAL REGIME, $\lim_{x\to\infty} x^2 u''(x) = -\infty.$ The concept of limit shape is not applicable here; there are two possible situations:

a) $\mu_* = \infty.$ In such systems the multiplicative measures defined by formula (6) are supported on sequences, p, for which $\mathscr{M}(p) = \infty$, i.e., they do not correspond to partitions of finite sets or integers. Consequently, the grand canonical probability measures cannot be defined via formula (5), while the sums representing the size distribution functions (7) diverge almost surely. One example when this scenario occurs is whenever $u(x)$ tends to negative infinity as fast or faster than $-x(\ln x)^p$, $p > 0$.

b) $\mu_* \in \mathbb{R}.$ In such systems the expected number of monomers remains finite for all values of μ, i.e., $\mathbb{E}\mathscr{M} < \infty$ when $\mu = \mu_*$. This implies that the thermodynamic limit cannot be achieved in the setting of grand canonical ensembles and the size distribution functions remain discrete as $\mu \searrow \mu_*$. One family of ensembles with this behavior is prescibed by $u(x) = (\ln x)^p$, $p > 1$.

SUPERCRITICAL REGIME, $\lim_{x\to\infty} x^2 u''(x) = \infty.$ In this regime most polymers have sizes in the vicinity of $k \sim \varkappa(\mu)$ as prescribed in equation (14). The limit shape is attained under the scaling of $\varkappa(\mu)$ and is given by the step function,

$$F(x) = \mathbb{1}_{[0,1]}(x). \tag{15}$$

This regime encompasses all cases which yield the same limit shape as the uniform measure on partitions of sets for which $u(x) = \ln\Gamma(x + 1)$, cf. [19]. Other examples in this category include $u(x) = x^p$, $p > 1$; $u(x) = x(\ln x)^p$, $p > 0$, etc.

CRITICAL REGIME, $\lim_{x\to\infty} x^2 u''(x) \in \mathbb{R}.$ In this regime the function $u(x)$ must have the form (Lemma A.1)

$$u(x) = -\mu_* x + (1 - d)\ln x + v(x), \quad \text{where} \quad d \in \mathbb{R}, \quad v(x) \ll \ln x. \tag{16}$$

The value of μ_* in this case is necessarily finite and without loss of generality may be assumed to be equal to zero. The limit shape behavior is determined by the finer (sublinear) features of $u(x)$. Let us relate these features back to the energies E_k that were used to introduce Gibbs ensembles in the first place. Using Stirling's approximation, $\ln\Gamma(k + 1) \sim k\ln k - k + 1/2 \ln k + \cdots$ in formula (11), we can see that unless $\beta E_k \sim -k\ln k$, the system is always (for all β) either in the sub or supercritical regime. The critical regime is only possible when the leading order asymptotics of βE_k and $\ln\Gamma(k + 1)$ cancel each other exactly, in which case the interplay between the lower order behavior of βE_k-s and $1/2 \ln k$ from $\ln\Gamma(k + 1)$ determines the outcome, i.e., the value of d and the asymptotics of $v(x)$ as $x \to \infty$. Different scenarios arise are possible:

a) **Either $d < 0$, or $d = 0$ and $v(x) \to \infty$.** If $d < -1$, $\mathbb{E}\mathscr{M}$ remains finite for all values of μ, as in the subcritical regime (b) above. In all other cases, the rescaled size distribution functions may only converge to a trivial function identically equal to zero. This regime generalizes the *convergent* case in [8] where the conditions $d < 0$ and $v(x) \sim$ const are assumed.

b) Either $d > 0$, or $d = 0$ and $v(x) \to -\infty$. The limit shape is given by the (rescaled) incomplete gamma function,

$$F(x) = \frac{\Gamma(x;d)}{\Gamma(d+1)}; \qquad \Gamma(x;d) := \int_x^\infty y^{d-1} e^{-y} \, dy. \tag{17}$$

It is attained with the scaling $\varkappa = 1/\mu$, asymptotically equivalent to (14). This regime generalizes the *expansive* case from [8] where $d > 0$ and $v(x) \sim$ const are assumed; and $d > 1$, $v(x) = -j \ln \ln x$ case from [6]. For example, our result covers cases such as $d = 0$, $v(x) = -(\ln x)^p$, $p \in (0,1)$; or $v(x) \sim -\ln \cdots \ln x$.

c) $d = 0$ and $v(x) \to C$, In this case the scaled size distribution functions (under the same scaling, $\varkappa = 1/\mu$) converge in distribution to a Poisson process which starts at 0 when $x = \infty$ and whose jumps are distributed with density e^{-C-x}/x. As $x \searrow 0$, this process tends to infinity almost surely. It is closely related to the Poisson-Dirichlet distribution [12, 10] which describes the distribution of the (scaled) cycle lengths in a random permutation; see Section II.iii.

<div align="center">—— Local profiles for the step function scenarios ——</div>

As discussed above, in the supercritical regime the limit shape is given by the step function. It is, therefore, natural to study the local behavior of the size distribution functions near the point of discontinuity. In order to do that, we consider the shifted size distribution function and its rescaled version, cf. equations (7) and (8):

$$g(x;\boldsymbol{p}) := \sum_{k \geq x+\varkappa} p_k \qquad \text{and} \qquad G_\mu(x;\boldsymbol{p}) := \frac{\varkappa}{\mathbb{E}\mathcal{M}} g(\zeta x;\boldsymbol{p}), \quad x \in \mathbb{R}. \tag{18}$$

Notice that the shift is given by $\varkappa(\mu)$, solution of equation (14); the "vertical" scaling factor, $\mathbb{E}\mathcal{M}/\varkappa$, remains as in equation (8), while the "horizontal" scaling factor, $\zeta = \zeta(\mu)$, is different and must be determined from further analysis, see Section III. Let

$$G(x) := \lim_{\mu \searrow \mu_*} G_\mu(x;\boldsymbol{p}) \tag{19}$$

be the local limit shape. The following three regimes are possible:

Gaussian regime, $\lim_{x \to \infty} u''(x) = 0$. The proper scaling is prescribed by $\zeta = 1/\sqrt{u''(\varkappa)}$, and the local shape is given by the Gaussian integral,

$$G(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-y^2/2} \, dy. \tag{20}$$

Discrete Gaussian regime, $\lim_{x \to \infty} u''(x) = C \in (0,\infty)$. No additional scaling is needed, i.e., $\zeta = 1$, and the local shape is a discrete Gaussian distribution,

$$G(x) = \frac{1}{Q} \sum_{k \geq x} e^{-Ck^2/2}, \quad \text{where} \quad Q = \sum_{k \in \mathbb{Z}} e^{-Ck^2/2}. \tag{21}$$

7

Hard step function regime $\lim_{x\to\infty} u''(x) = \infty$. In this regime the size distribution function asymptotically concentrates on finitely many values of k, and the local limit shape remains a step function under any rescaling such that $\zeta \to \infty$ as $\mu \searrow \mu_*$.

II. Limit shape theorems

In this section we present the formal statements and calculations regarding the limit shapes for the grand canonical Gibbs ensembles of set partitions. The Assumption 1 on the function $u(\cdot)$ introduced in section I.iv is enacted. As can be expected from Section I.ii, the calculations that we carry out in order to establish various limit shape statements involve various sums as in equation (10). Let us introduce some additional notation for these sums:

$$S_\mu(a,b) := \mathbb{E} \sum_{a \le k < b} p_k = \sum_{a \le k < b} e^{-\mu k - u(k)}; \qquad S_\mu(a) := S_\mu(a,\infty); \qquad S_\mu := S_\mu(1,\infty). \tag{22}$$

II.i. The step function regime

We start from the cases where the limit shape is the step function $F(x) = \mathbb{1}_{[0,1]}(x)$. This scenario occurs when $\lim_{x\to\infty} x^2 u''(x) = \infty$. In this regime the sums in (22) are dominated by the terms near the maximum of $-\mu k - u(k)$, i.e., in the vicinity of $k = \varkappa(\mu)$ determined by equation (14). Notice also that

$$\mathbb{E}\mathcal{M} = \sum_{k=1}^{\infty} k\, e^{-\mu k - u(k)} = \sum_{k=1}^{\infty} e^{-\mu k - [u(k) - \ln k]}. \tag{23}$$

Therefore, by Lemma A.2, we have $\lim_{\mu \searrow \mu_*} \mathbb{E}\mathcal{M} = \infty$, and the thermodynamic limit is achieved as $\mu \searrow \mu_*$.

Theorem 1
Suppose $u(\cdot)$ satisfies $\lim_{x\to\infty} x^2 u''(x) = \infty$. Let $F(x) = \mathbb{1}_{[0,1]}(x)$, and μ and $\varkappa$ be related as prescribed in equation (14). Then for each $\lambda_1 < 1$, $\lambda_2 > 1$, and $\epsilon > 0$,

$$\lim_{\mu \searrow \mu_*} \mathbb{P}\left\{ \sup_{x \in \mathbb{R}^+ \setminus (\lambda_1, \lambda_2)} \left| F_\mu(x; \boldsymbol{p}) - F(x) \right| > \epsilon \right\} = 0. \tag{24}$$

Proof As $F_\mu(x; \boldsymbol{p})$ is decreasing and $F(x) = \mathbb{1}_{[0,1]}(x)$, it is enough to show that for all $x \in [0,1) \cup (1, \infty)$,

$$\lim_{\mu \searrow \mu_*} \mathbb{P}\left\{ |F_\mu(x; \boldsymbol{p}) - F(x)| > \epsilon \right\} = 0. \tag{25}$$

Using the triangle inequality, we can split this into two parts:

$$\lim_{\mu \searrow \mu_*} \mathbb{P}\left\{ |F_\mu(x; \boldsymbol{p}) - \mathbb{E}F_\mu(x; \boldsymbol{p})| > \epsilon \right\} = 0; \qquad \lim_{\mu \searrow \mu_*} \mathbb{E}F_\mu(x; \boldsymbol{p}) = F(x). \tag{26}$$

For the first limit, by Chebyshev's inequality, we get

$$\mathbb{P}\left\{ |F_\mu(x; \boldsymbol{p}) - \mathbb{E}F_\mu(x; \boldsymbol{p})| > \epsilon \right\} \le \frac{1}{\epsilon^2} \mathbb{V}F_\mu(x; \boldsymbol{p}) = \frac{1}{\epsilon^2} \frac{\varkappa}{\mathbb{E}\mathcal{M}} \mathbb{E}F_\mu(x; \boldsymbol{p}). \tag{27}$$

By Lemma II.2, $\varkappa/\mathbb{E}\mathcal{M} \sim 1/S_\mu$, while by Lemma A.2, $\lim_{\mu \searrow \mu_*} S_\mu = \infty$. Thus the right-hand side of equation (27) tends to 0 as $\mu \searrow \mu_*$, whenever the second limit in equation (26) holds as well.

To prove the latter, we note that

$$\mathbb{E} F_\mu(x; \boldsymbol{p}) = \frac{\varkappa}{\mathbb{E}\mathcal{M}} S_\mu(\varkappa x) = \frac{\varkappa S_\mu}{\mathbb{E}\mathcal{M}} \frac{S_\mu(\varkappa x)}{S_\mu}. \tag{28}$$

The same Lemma II.2 implies that $\varkappa S_\mu \sim \mathbb{E}\mathcal{M}$, thus it remains to show that, for all $x \in [0,1) \cup (1,\infty)$, $\lim_{\mu \searrow \mu_*} S_\mu(\varkappa x)/S_\mu = \mathbb{1}_{[0,1]}(x)$. This follows from the fact that in this regime the principal contribution into the sums in equation (22) comes from the terms near $k = \varkappa(\mu)$, which is proven in Lemma II.1 below. ∎

Lemma II.1
Assume that a twice-differentiable function $u(\cdot)$ satisfies $\lim_{x \to \infty} x^2 u''(x) = \infty$. Fix arbitrary λ_1 and λ_2 satisfying $0 < \lambda_1 < 1 < \lambda_2$. Let $\varkappa(\mu)$ be the solution of (14). Then

$$\lim_{\mu \searrow \mu_*} S_\mu(\varkappa \lambda_1, \varkappa \lambda_2)/S_\mu = 1. \tag{29}$$

Proof   To prove (29), it is sufficient to show that $S_\mu(\varkappa \lambda_2)/S_\mu \to 0$ and $S_\mu(1, \varkappa \lambda_1)/S_\mu \to 0$ as $\mu \searrow \mu_*$. We split the proof into the following two steps:

Step 1. For the first assertion, let $\varepsilon = \lambda_2 - 1$; $\varkappa_n = (1 + n\varepsilon)\varkappa$. Then we have

$$S_\mu(\varkappa \lambda_2) = \sum_{n=1}^{\infty} S_\mu(\varkappa_n, \varkappa_{n+1}). \tag{30}$$

Let $\alpha_n = e^{-\mu \varkappa_n - u(\varkappa_n)}$, cf. equation (6). By definition of $\varkappa$, the quantity $-\mu k - u(k)$ is decreasing whenever $k \geq \varkappa$, therefore $S_\mu(\varkappa_n, \varkappa_{n+1}) \leq (\varkappa_{n+1} - \varkappa_n)\alpha_n = \varepsilon \varkappa \alpha_n$, and

$$S_\mu(\varkappa \lambda_2) \leq \varepsilon \varkappa \sum_{n=1}^{\infty} \alpha_n. \tag{31}$$

Observe that $\alpha_{n+1}/\alpha_n = e^{-\mu(\varkappa_{n+1} - \varkappa_n) - [u(\varkappa_{n+1}) - u(\varkappa_n)]} = e^{-\varepsilon \varkappa [\mu + u'(\xi)]}$ for some $\xi \in [\varkappa_n, \varkappa_{n+1}]$. As u' is increasing, $u'(\xi) \geq u'(\varkappa_n) \geq u'(\varkappa_1)$. Therefore $\alpha_{n+1}/\alpha_n \leq e^{-\varepsilon \varkappa (\mu + u'(\varkappa_1))}$, which implies that

$$\sum_{n=1}^{\infty} \alpha_n \leq \alpha_1 \sum_{n=0}^{\infty} e^{-n\varepsilon \varkappa [\mu + u'(\varkappa_1)]} = \frac{\alpha_1}{1 - e^{-\varepsilon \varkappa [\mu + u'(\varkappa_1)]}}. \tag{32}$$

Let $\tilde{\varkappa} = (1 + \varepsilon/2)\varkappa$ and $\tilde{\alpha}_0 = e^{-\mu \tilde{\varkappa} - u(\tilde{\varkappa})}$. We estimate $S_\mu(\varkappa, \tilde{\varkappa}) \geq (\tilde{\varkappa} - \varkappa)e^{-\mu \tilde{\varkappa} - u(\tilde{\varkappa})} = \varepsilon \varkappa \tilde{\alpha}_0/2$ for terms around $\varkappa$, and also observe that $\alpha_1/\tilde{\alpha}_0 \leq e^{-[\mu + u'(\tilde{\varkappa})]\varepsilon \varkappa/2}$. Then

$$S_\mu(\varkappa \lambda_2)/S_\mu(\varkappa, \tilde{\varkappa}) \leq \frac{2\alpha_1}{\tilde{\alpha}_0} \frac{1}{1 - e^{-\varepsilon \varkappa [\mu + u'(\varkappa_1)]}} \leq \frac{e^{-[\mu + u'(\tilde{\varkappa})]\varepsilon \varkappa/2}}{1 - e^{-\varepsilon \varkappa [\mu + u'(\varkappa_1)]}}. \tag{33}$$

Recall the relation (14) between μ and $\varkappa$, and that $x^2 u''(x) \to \infty$ by assumption. This implies that there exists some $\xi \in [\varkappa, \varkappa_1]$, such that

$$e^{-\varepsilon \varkappa [\mu + u'(\varkappa_1)]} = e^{-u''(\xi)\varepsilon^2 \varkappa^2} \to 0, \quad \text{as} \quad \mu \searrow \mu_* \quad \text{(or, equivalently, as } \varkappa \to \infty\text{)}. \tag{34}$$

9

Similarly,

$$e^{-[\mu+u'(\tilde{\varkappa})]\varepsilon\varkappa/2} \to 0, \quad \text{as } \mu \searrow \mu_*. \tag{35}$$

Then we conclude,

$$\lim_{\mu \searrow \mu_*} S_\mu(\varkappa\lambda_2)/S_\mu(\varkappa,\tilde{\varkappa}) = 0, \tag{36}$$

and therefore $S_\mu(\varkappa\lambda_2)/S_\mu \searrow 0$.

Step 2. For the second assertion, we first introduce a cut-off for k. Because $\lim_{x\to\infty} x^2 u''(x) = \infty$, we can find N such that $u''(x) > 0$ for $x \geq N$. In other words, $u'(x)$ is increasing on $[N, \infty)$. Let us show that

$$\lim_{\mu \searrow \mu_*} \frac{S_\mu(1,N)}{S_\mu} = 0. \tag{37}$$

If $\mu_* = -\infty$, let $r = \max_{1 \leq x \leq N-1} u'(x)$. Then for any $\mu < -r$, we have $-\mu k - u(k)$ is increasing on $1 \leq k \leq \varkappa$. As $\varkappa \to \infty$, equation (37) follows. On the other hand, if $\mu_* > -\infty$,

$$\lim_{\mu \searrow \mu_*} S_\mu(1,N) = \sum_{k=1}^{N-1} e^{-\mu_* k - u(k)} \leq (N-1) \max_{1 \leq k \leq N-1} e^{-\mu_* k - u(k)}. \tag{38}$$

As $\lim_{\mu \searrow \mu_*} S_\mu = \infty$, equation (37) follows in this case as well. Thus in order to show that the ratio $S_\mu(1,\varkappa\lambda_1)/S_\mu$ tends to zero, it suffices to have

$$S_\mu(N,\varkappa\lambda_1)/S_\mu \to 0. \tag{39}$$

To this end, we follow a similar idea as in Step 1. Let $\varepsilon = 1 - \lambda_1$ and $L = \min\{n \in \mathbb{N} : N \geq (1 - n\varepsilon)\varkappa\}$. Let $\varkappa_n = (1 - n\varepsilon)\varkappa$ for $x = 1,\ldots,L-1$ and $\varkappa_L = N$. Then we have

$$S_\mu(N,\varkappa\lambda_1) = \sum_{n=1}^{L-1} S_\mu(\varkappa_{n+1},\varkappa_n). \tag{40}$$

Let $\alpha_n = e^{-\mu\varkappa_n - u(\varkappa_n)}$. When $k \in [N,\varkappa]$, the quantity $-\mu k - u(k)$ is increasing, which implies that $S_\mu(\varkappa_{n+1},\varkappa_n) \leq (\varkappa_n - \varkappa_{n+1})\alpha_n \leq \varepsilon\varkappa\alpha_n$. Thus

$$S_\mu(N,\varkappa\lambda_1) \leq \varepsilon\varkappa \sum_{n=1}^{L-1} \alpha_n. \tag{41}$$

Observe that, when $n = 1,\ldots,L-2$, $\alpha_{n+1}/\alpha_n = e^{-\mu(\varkappa_{n+1}-\varkappa_n)-[u(\varkappa_{n+1})-u(\varkappa_n)]} = e^{\varepsilon\varkappa(\mu+u'(\xi))}$ for some $\xi \in [\varkappa_{n+1},\varkappa_n]$. As $u'(x)$ is increasing on $[N,\infty)$, $u'(\xi) \leq u'(\varkappa_n) \leq u'(\varkappa_1)$. $\alpha_{n+1}/\alpha_n \leq e^{\varepsilon\varkappa(\mu+u'(\varkappa_1))}$. This yields that

$$S_\mu(N,\varkappa\lambda_1) \leq \varepsilon\varkappa \sum_{n=1}^{L-1} \alpha_n \leq \varepsilon\varkappa\alpha_1 \sum_{n=0}^{\infty} e^{n\varepsilon\varkappa(\mu+u'(\varkappa_1))} = \frac{\varepsilon\varkappa\alpha_1}{1 - e^{\varepsilon\varkappa(\mu+u'(\varkappa_1))}}. \tag{42}$$

Let $\tilde{\varkappa} = (1 - \varepsilon/2)\varkappa$. Using the same argument as for validating formula (36) above, we conclude that $S_\mu(N,\varkappa\lambda_1)/S_\mu \to 0$. ∎

10

Lemma II.2
Assume that $u(\cdot)$ and $\varkappa(\mu)$ are as in Lemma II.1. Then

$$\lim_{\mu \searrow \mu_*} \varkappa S_\mu / \mathbb{E}\mathscr{M} = 1. \tag{43}$$

Proof Recall that $\mathbb{E}\mathscr{M}$ may be treated exactly as S_μ if we absorb the extra factor of k into $u(k)$, see equation (23). Thus as in the previous lemma, the principal contributions into $\mathbb{E}\mathscr{M}$ come from the terms in the vicinity of $\hat{\varkappa}$, the largest solution of

$$u'(\hat{\varkappa}) - 1/\hat{\varkappa} = -\mu. \tag{44}$$

Such $\hat{\varkappa}$ exists, as $u'(\hat{\varkappa}) - 1/\hat{\varkappa}$ is increasing for large $\hat{\varkappa}$. Let us show that $\lim_{\mu \searrow \mu_*}(\varkappa/\hat{\varkappa}) = 1$. Indeed, as $\mu = -u'(\varkappa)$,

$$u'(\varkappa) = u'(\hat{\varkappa}) - 1/\hat{\varkappa}. \tag{45}$$

Thus by the mean value theorem, there exists $\xi \in [\varkappa, \hat{\varkappa}]$ (note that $\hat{\varkappa} > \varkappa$), such that

$$u''(\xi)(\hat{\varkappa} - \varkappa) = 1/\hat{\varkappa}, \quad \text{or equivalently,} \quad \xi^2 u''(\xi)\left[\hat{\varkappa}(\hat{\varkappa} - \varkappa)/\xi^2\right] = 1. \tag{46}$$

By assumption, $\lim_{\xi \to \infty} \xi^2 u''(\xi) = \infty$, and therefore we must have

$$\lim_{\mu \searrow \mu_*}\left[\hat{\varkappa}(\hat{\varkappa} - \varkappa)/\xi^2\right] = 0. \tag{47}$$

As $\xi \le \hat{\varkappa}$, this implies that $\lim_{\mu \searrow \mu_*}(\varkappa/\hat{\varkappa}) = 1$, as claimed.
By Lemma II.1, for any $0 < \lambda_1 < 1 < \lambda_2$, we have

$$\mathbb{E}\mathscr{M} \sim \sum_{k \in [\hat{\varkappa}\lambda_1, \hat{\varkappa}\lambda_2)} k\,e^{-\mu k - u(k)}. \tag{48}$$

At the same time,

$$\hat{\varkappa}\lambda_1 \sum_{k \in [\hat{\varkappa}\lambda_1, \hat{\varkappa}\lambda_2)} e^{-\mu k - u(k)} \le \sum_{k \in [\hat{\varkappa}\lambda_1, \hat{\varkappa}\lambda_2)} k\,e^{-\mu k - u(k)} < \hat{\varkappa}\lambda_2 \sum_{k \in [\hat{\varkappa}\lambda_1, \hat{\varkappa}\lambda_2)} e^{-\mu k - u(k)}. \tag{49}$$

Because $\varkappa \sim \hat{\varkappa}$, Lemma II.1 also implies that these upper and lower bounds are asymptotically equivalent to $\varkappa\lambda_1 S_\mu$ and $\varkappa\lambda_2 S_\mu$ respectively. As λ_1 and λ_2 may be chosen arbitrarily close to 1, the assertion follows. ∎

II.II. The incomplete gamma function regime

Let us now consider the critical regime, where $\lim_{x \to \infty} x^2 u''(x) \in \mathbb{R}$, see p. 7. Using the definition of μ_* (13) and Lemma A.1, we can represent the function $u(\cdot)$ as

$$u(x) = -\mu_* x + (1 - d)\ln x + v(x) \quad \text{with} \quad d \in \mathbb{R}, \quad \lim_{x \to \infty} xv'(x) = 0. \tag{50}$$

Notice that in this case the thermodynamic limit is achieved as $\mu \searrow \mu_* \in \mathbb{R}$. To simplify notation, let us assume, without loss of generality, that $\mu_* = 0$ and fix the scaling as $\varkappa = 1/\mu$. If $d > 1$,

this scaling is asymptotically equivalent to that determined by equation (14). Notice that when $0 \leq d \leq 1$, equation (14) has no solutions for $\mu > \mu_*$, however this scaling is still appropriate as may be seen from the analysis below. The scaled size distribution functions may now be defined as

$$F_\mu(x; \boldsymbol{p}) := \frac{1}{\mu \, \mathbb{E}\mathcal{M}} \sum_{\mu k \geq x} p_k. \tag{51}$$

Lemma II.3

Suppose that the condition (50) holds and $d \geq 0$. Then

$$\mathbb{E}\mathcal{M} \sim \mu^{-(d+1)} e^{-v(1/\mu)} \Gamma(d+1), \qquad S_\mu(x/\mu) \sim \mu^{-d} e^{-v(1/\mu)} \Gamma(x; d) \quad \text{for all } x > 0. \tag{52}$$

Proof We first compute

$$S_\mu(x/\mu) = \sum_{\mu k \geq x} k^{d-1} e^{-\mu k - v(k)} = \mu^{-d} e^{-v(1/\mu)} \sum_{\mu k \geq x} (\mu k)^{d-1} e^{-\mu k} e^{-v(k) + v(1/\mu)} \mu. \tag{53}$$

Take any $\varepsilon > 0$. As $\lim_{k \to \infty} k v'(k) = 0$, we have, $|v(x) - v(y)| \leq \varepsilon |\ln x - \ln y|$ for all large enough x and y. Therefore, we have $|-v(k) + v(1/\mu)| \leq \varepsilon |\ln(\mu k)|$ for all small enough μ. Sending $\mu \searrow 0$ and then $\varepsilon \to 0$, we obtain that the last sum in equation (53) converges to $\Gamma(x; d)$.

The asymptotics of $\mathbb{E}\mathcal{M}$ follows similarly. The difference is that the summation (23) for $\mathbb{E}\mathcal{M}$ starts from $k = 1$, and thus we must show that the contribution from the corresponding range of k is negligible in the limit of vanishing μ. Indeed, for any $\varepsilon \in (0, 1)$, take $N = N(\varepsilon)$ such that $|v(x) - v(y)| \leq \varepsilon |\ln x - \ln y|$ for all $x, y \geq N$. Split the sum producing $\mathbb{E}\mathcal{M}$ into two parts,

$$\sum_{k=1}^{N} k^d e^{-\mu k - v(k)} + \sum_{k=N+1}^{\infty} k^d e^{-\mu k - v(k)}. \tag{54}$$

As $v(k) \ll \ln k$, we have $\mu^{d+1} e^{v(1/\mu)} = \mathcal{O}(1)$. The first sum in equation (54) my be bounded by $N^d \exp\{\max_{k=1...N} |v(k)|\}$. Therefore

$$\lim_{\mu \to 0} \mu^{d+1} e^{v(1/\mu)} \mathbb{E}\mathcal{M} = \lim_{\varepsilon \to 0} \lim_{\mu \to 0} \mu^{d+1} e^{v(1/\mu)} \sum_{k=N+1}^{\infty} k^d e^{-\mu k - v(k)} = \Gamma(d+1), \tag{55}$$

concluding the proof. ∎

In the following theorem we prove that the limit shape is given by the incomplete gamma function if either $d > 0$; or $d = 0$ and $\lim_{x \to \infty} v(x) = -\infty$. Notice that when $d = 0$, we must require existence of the limit of $v(x)$ as $x \to \infty$, as it is not a consequence of our general assumptions that $\lim_{x \to \infty} x^2 v''(x) = 0$ and $\lim_{x \to \infty} x v'(x) = 0$. For example, consider $v(x) = \sin(\ln \ln x)$. (Alternatively, one could require sign-definiteness of $u''(x)$ or $v''(x)$ for all large enough x.)

Theorem 2

Let $u(\cdot)$ and $v(\cdot)$ be related as in equation (50), $F_\mu(x; \boldsymbol{p})$ be as defined in equation (51). Assume either that $d > 0$, or $d = 0$ and $\lim_{x \to \infty} v(x) = -\infty$. Then for each $a > 0$, $\epsilon > 0$,

$$\lim_{\mu \searrow 0} \mathbb{P}\left\{ \sup_{x \geq a} |F_\mu(x; \boldsymbol{p}) - F(x)| > \epsilon \right\} = 0. \tag{56}$$

Here $F(x) = \Gamma(x; d)/\Gamma(d+1)$; $\Gamma(x; d)$ is the incomplete gamma function, cf. (17).

Proof Notice that $F(x)$ is a decreasing function and $\lim_{x\to\infty} F(x) = 0$. By triangle inequalities, it is enough to show, for any fixed $x > 0$,

$$\lim_{\mu\searrow 0} \mathbb{P}_\mu\left\{|F_\mu(x;\boldsymbol{p}) - F(x)| > \epsilon\right\} = 0. \tag{57}$$

Lemma II.3 shows that

$$\lim_{\mu\searrow 0} \mathbb{E}\,F_\mu(x;\boldsymbol{p}) = \lim_{\mu\searrow 0} \frac{1}{\mu\,\mathbb{E}\,\mathcal{M}} S_\mu(x/\mu) = F(x). \tag{58}$$

It remains to show that

$$\lim_{\mu\searrow 0} \mathbb{P}_\mu\left\{|F_\mu(x;\boldsymbol{p}) - \mathbb{E}\,F_\mu(x;\boldsymbol{p})| > \epsilon\right\} = 0. \tag{59}$$

By Chebyshev's inequality and using that p_k-s are Poisson random variables,

$$\mathbb{P}_\mu\left\{|F_\mu(x;\boldsymbol{p}) - \mathbb{E}\,F_\mu(x;\boldsymbol{p})| > \epsilon\right\} \leq \epsilon^{-2}\,\mathbb{V}\,F_\mu(x;\boldsymbol{p}) = \frac{\epsilon^{-2}}{\mu\,\mathbb{E}\,\mathcal{M}}\,\mathbb{E}\,F_\mu(x;\boldsymbol{p}). \tag{60}$$

By Lemma II.3, $\mu\,\mathbb{E}\,\mathcal{M} \sim \mu^{-d}\,\mathrm{e}^{-v(1/\mu)}$, which tends to infinity as μ tends to zero. ∎

II.iii. Limit process regime

Formula (60) above and the accompanying argument show exactly why there is no limit shape if $d = 0$ and $\lim_{x\to\infty} v(x) \in \mathbb{R}$: the variance of the scaled distribution functions does not vanish in the thermodynamic limit. Instead of the limit shape, however, we get a limit process. This happens whenever the function $u(\cdot)$, see (11), may be represented as,

$$u(x) = -\mu_* x + \ln x + v(x), \qquad\text{with}\qquad v(x) \to C \in \mathbb{R} \quad\text{as}\quad x \to \infty. \tag{61}$$

As before, we may set $\mu_* = 0$ and consider the limit $\mu \searrow 0$ rather than $\mu \searrow \mu_*$. Computing the expected number of monomers in the system, we get,

$$\mathbb{E}\,\mathcal{M} = \sum_{\mu k \geq 0} \mathrm{e}^{-\mu k - v(k)} \sim \frac{1}{\mu} \int_0^\infty \mathrm{e}^{-t - C}\,\mathrm{d}t = \frac{\mathrm{e}^{-C}}{\mu}. \tag{62}$$

Thus in this case, formula (51) for the scaled size distribution functions may be replaced with a simpler equivalent,

$$F_\mu(x;\boldsymbol{p}) := \sum_{\mu k \geq x} p_k. \tag{63}$$

The "vertical" scaling in this case is μ-independent, as $\mu\,\mathbb{E}\,\mathcal{M} \to \mathrm{e}^{-C}$ in the limit.

As mentioned above, in this scaling the variance of $F_\mu(x;\boldsymbol{p})$ does not vanish as $\mu \searrow 0$, thus we are expecting that the latter converges to a stochastic process rather than to a deterministic limit shape. Let us define a random measure on $\mathbb{R}^+$ corresponding to $F_\mu(x;\boldsymbol{p})$ (its measure-valued derivative up to a minus sign):

$$\pi_\mu(\mathrm{d}x;\boldsymbol{p}) := \sum_{k=1}^\infty p_k \delta(x - \mu k)\,\mathrm{d}x, \quad\text{i.e.,}\quad \pi_\mu(\mathcal{A};\boldsymbol{p}) = \sum_{\mu k \in \mathcal{A}}^\infty p_k, \quad \mathcal{A} \subset \mathbb{R}^+. \tag{64}$$

13

The p_k-s are independent Poisson random variables with parameters $\mathrm{e}^{-\mu k - v(k)}/k$, consequently, for any finite sequence of disjoint intervals $[a_j, b_j)$, $j = 1 \ldots n$, the quantities $\pi_\mu([a_j, b_j); \boldsymbol{p})$ are also independent Poisson random variables with parameters $\sum_{a_j \le \mu k < b_j} \mathrm{e}^{-\mu k - v(k)}/k$. Computing

$$\lim_{\mu \searrow 0} \sum_{a \le \mu k < b} \frac{1}{k} \mathrm{e}^{-\mu k - v(k)} = \mathrm{e}^{-C} \int_a^b \frac{\mathrm{e}^{-x}}{x} \, \mathrm{d}x. \tag{65}$$

we conclude that π_μ converges in distribution to π, a Poisson point process on $\mathbb{R}^+$ with intensity given by e^{-C-x}/x, see [11]. Therefore we have just proved

Proposition II.1
Let $u(k)$ satisfy equation (61) and the scaled size distribution function be defined as in equation (63). Then its derivative, $\pi_\mu(x; \boldsymbol{p})$, converges in distribution to a Poisson point process with intensity e^{-C-x}/x as prescribed in equation (65). Alternatively,

$$F_\mu(x; \boldsymbol{p}) \xrightarrow{d} \pi([x, \infty)), \quad x \in \mathbb{R}^+. \tag{66}$$

Notice that the range of summation in formula (63) may be rewritten as $k \ge x \mathrm{e}^C \, \mathbb{E}\mathcal{M}$. If we remove the expectation and discard the factor of e^C, we get an alternative (random) scaling for the size distribution function:

$$\tilde{F}_\mu(x; \boldsymbol{p}) = \sum_{k \ge x \mathcal{M}(\boldsymbol{p})} p_k. \tag{67}$$

This scaling ensures that $\tilde{F}_\mu(x; \boldsymbol{p}) = 0$ whenever $x > 1$, and thus the size distribution function may be regarded as a random partition of the interval $[0, 1]$ in the sense that $\tilde{F}_\mu(a; \boldsymbol{p}) - \tilde{F}_\mu(b; \boldsymbol{p})$ is the number of subintervals with lengths in $[a, b)$ in such a partition. The limiting object in this case is a Poisson-Dirichlet distribution that describes cycle lengths in a random permutation [12, 10]. This, of course, is not a coincidence, as the functional form of the function $u(\cdot)$, given in equation (61), implies that in this regime the resulting grand canonical Gibbs measure on partitions is exactly the Poissonization of the uniform (Dirichlet-Haar) measures on symmetric groups.

II.iv. No limit shape regimes

Let us take a look at the regimes in which the limit shape does not exist. This may happen for several reasons: either because the number of monomers, $\mathbb{E}\mathcal{M}$, remains finite (as $\mu \searrow \mu_*$); because the grand canonical measures are supported on non-summable sequences $\boldsymbol{p}$ which do not correspond to partitions of finite integers/sets; or if the limit shape technically exists, but is given by a trivial function identically equal to zero.

- $x^2 u''(x) \to -\infty$ with $u'(x) \searrow -\infty$. We claim that the sum $\sum_{k=1}^\infty p_k$ is almost surely infinite for all μ (and thus so is $\mathcal{M}$). Indeed, as $\mu_* = \infty$, $S_\mu = \mathbb{E} \sum_{k=1}^\infty p_k = \infty$ for all $\mu \in \mathbb{R}$. The following proposition shows, that the grand canonical measures are concentrated on non-summable sequences, and thus there are no limit shapes in these scenarios:

14

Proposition II.2
Let $S_\mu = \infty$. Then

$$\mathbb{P}\left\{\sum_{k=1}^\infty p_k = \infty\right\} = 1, \quad \text{and consequently,} \quad \mathbb{P}\{\mathcal{M} = \infty\} = \mathbb{P}\left\{\sum_{k=1}^\infty k p_k = \infty\right\} = 1. \tag{68}$$

Proof   Recall that p_k-s are independent Poisson random variables, thus their sum, $\sum_{k=1}^m p_k$, is also Poisson with parameter $A_m := \sum_{k=1}^m e^{-\mu k - u(k)}$, see Section I.ii. For any fixed $N \in \mathbb{N}$ and all large enough m, we have

$$\mathbb{P}\left\{\sum_{k=1}^\infty p_k \le N\right\} \le \mathbb{P}\left\{\sum_{k=1}^m p_k \le N\right\} = e^{-A_m} \sum_{k=0}^N \frac{A_m^k}{k!} \le e^{-A_m} A_m^N. \tag{69}$$

Sending m to infinity, we obtain, $\mathbb{P}\left\{\sum_{k=1}^\infty p_k \le N\right\} = 0$ for all $N \in \mathbb{N}$, yielding the assertion.   ∎

• **$x^2 u''(x) \to -\infty$, $u'(x) \to$ const.** Recall formula (23) for the expected number of monomers, $\mathbb{E}\mathcal{M}$. By an argument similar to that in Lemma A.2, we can get that $\mu_* = -\lim_{x\to\infty} u'(x)$ and $(\mu_* x + u(x))/\ln x \to \infty$. This implies that $\mathbb{E}\mathcal{M} < \infty$ when $\mu = \mu_*$, i.e., the expected number of monomers remains finite as $\mu \searrow \mu_*$ and thus the proper thermodynamic limit cannot be achieved in the setting of grand canonical ensembles.

• **$\lim_{x\to\infty} x^2 u''(x) \in \mathbb{R}$.** In this regime, $u(\cdot)$ has the form $u(x) = -\mu_* x + (1 - d)\ln x + v(x)$, where $d \in \mathbb{R}$ and $\lim_{x\to\infty} x v'(x) = 0$, see equation (50). In Theorem 2 and Proposition II.1 we considered cases when either $d > 0$ or $d = 0$ with $\lim_{x\to\infty} v(x) \in \mathbb{R} \cup \{-\infty\}$. When $d < -1$, $\mathbb{E}\mathcal{M}$ is dominated by a converging series, $\propto \sum k^d$, and so is finite for all μ. Thus, similarly to the case above, there is no thermodynamic limit in this regime. It remains to discuss cases when $d \in [-1, 0)$, or $d = 0$ with $\lim_{x\to\infty} v(x) = \infty$.

As earlier, without loss of generality, we may set $\mu_* = 0$. Consider any scaling, such that $\lim_{\mu\to 0} \varkappa(\mu) = \infty$, which is necessary to get a continuous limit for the size distribution functions. For any $y > x > 0$, let

$$F_\mu(x, y; \boldsymbol{p}) := \frac{\varkappa}{\mathbb{E}\mathcal{M}} \sum_{\varkappa x \le k < \varkappa y} p_k. \tag{70}$$

Then we have $F_\mu(x; \boldsymbol{p}) = F_\mu(x, y; \boldsymbol{p}) + F_\mu(y; \boldsymbol{p})$, cf. formula (7). We know that $\sum_{\varkappa x \le k < \varkappa y} p_k$ is of Poisson distribution with parameter $S_\mu(\varkappa x, \varkappa y)$. Let us show that

$$\lim_{\mu\searrow 0} S_\mu(\varkappa x, \varkappa y) = 0 \tag{71}$$

for all x and y with $0 < x < y$. In fact, in the case $d < 0$, as $v(x) \ll \ln x$, we have

$$S_{\mu_*} = \sum_{k \ge 1} e^{(d-1)\ln k - v(k)} < \infty. \tag{72}$$

Notice that $S_\mu(\varkappa x, \varkappa y) \le S_{\mu_*}(\varkappa x, \varkappa y)$. As $\varkappa \to \infty$, equation (71) follows, cf. (22). We now turn to the case when $d = 0$ and $\lim_{x\to\infty} v(x) = \infty$. Note that

$$S_\mu(\varkappa x, \varkappa y) = \sum_{\varkappa x \le k < \varkappa y} \frac{1}{k} e^{-\mu k - v(k)} = e^{-v(\varkappa)} \sum_{\varkappa x \le k < \varkappa y} \frac{1}{k} e^{-\mu k - (v(k) - v(\varkappa))}. \tag{73}$$

15

Take any $\varepsilon > 0$. Notice that $|v(k_1) - v(k_2)| \leq \varepsilon |\ln(k_1/k_2)|$ for k_1, k_2 large, cf. Lemma A.1. Then for $\varkappa$ large, it holds

$$S_\mu(\varkappa x, \varkappa y) \leq c_{x,y,\varepsilon} \, e^{-v(\varkappa)} \sum_{\varkappa x \leq k < \varkappa y} \frac{1}{k} \leq c_{x,y,\varepsilon} \, e^{-v(\varkappa)} \ln\left(\frac{y}{x}\right) \tag{74}$$

where $c_{x,y,\varepsilon}$ is a constant depending only on x, y, and ε. By the assumption that $\lim_{x \to \infty} v(x) = \infty$, we obtain (71).

As a consequence of (71), we have

$$\lim_{\mu \searrow 0} \mathbb{P}\left\{ F_\mu(x, y; \boldsymbol{p}) = 0 \right\} = \lim_{\mu \searrow 0} \mathbb{P}\left\{ \sum_{\varkappa x \leq k < \varkappa y} p_k = 0 \right\} = \lim_{\mu \searrow 0} e^{-S_\mu(\varkappa x, \varkappa y)} = 1. \tag{75}$$

Assume now that $F_\mu(x; \boldsymbol{p})$ converges in distribution to some random variable ξ. Then $F_\mu(y; \boldsymbol{p})$ must converge in distribution to ξ as well. In other words, for all $x > 0$, $F_\mu(x; \boldsymbol{p})$ converges in distribution to ξ. However,

$$\int_0^\infty \mathbb{E}\xi \, dx \leq \int_0^\infty \liminf_{\mu \searrow 0} \mathbb{E} F_\mu(x; \boldsymbol{p}) \, dx \leq \liminf_{\mu \searrow 0} \int_0^\infty \mathbb{E} F_\mu(x; \boldsymbol{p}) \, dx = 1. \tag{76}$$

This implies that $\xi \equiv 0$, i.e, the only admissible limit shape is the zero function.

III. Local profiles of the step function shape

As discussed in Section II.i, whenever $x^2 u''(x) \to \infty$ as $x \to \infty$, the limit shape is given by the step function. In this section we investigate its local profile near the discontinuity point. We consider the shifted and rescaled size distribution function as defined in equation (18):

$$G_\mu(x; \boldsymbol{p}) = \frac{\varkappa}{\mathbb{E}\mathcal{M}} \sum_{k \geq \zeta x + \varkappa} p_k, \quad x \in \mathbb{R}. \tag{77}$$

The new parameter $\zeta = \zeta(\mu)$ controls the scaling in the vicinity of the discontinuity point.

III.i. Gaussian regime

We start from the case when the local limit shape is given by a Gaussian integral. This happens when $u''(x) \to 0$ with additional assumption that $x^2 u''(x)$ is non-decreasing and $u''(x)$ is non-increasing. New local limit shapes along subsequences of μ might show up if monotonicity is not assumed. See Section III.iv for a more detailed discussion.

Theorem 3
Assume that $\lim_{x \to \infty} u''(x) = 0$, $\lim_{x \to \infty} x^2 u''(x) = \infty$ and both $u''(x)$ and $x^2 u''(x)$ are monotone. Let

$$G_\mu(x; \boldsymbol{p}) := \frac{\varkappa}{\mathbb{E}\mathcal{M}} \sum_{k - \varkappa \geq x/\sqrt{u''(\varkappa)}} p_k. \tag{78}$$

16

Then, for each $\epsilon > 0$, we have

$$\lim_{\mu \searrow \mu_*} \mathbb{P} \left\{ \sup_{x \in \mathbb{R}} |G_\mu(x; \boldsymbol{p}) - G(x)| > \epsilon \right\} = 0 \tag{79}$$

where

$$G(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} \, dt \tag{80}$$

PROOF   The theorem will follow if we show, for all $x \in \mathbb{R}$,

$$\lim_{\mu \searrow \mu_*} \mathbb{E} \, G_\mu(x; \boldsymbol{p}) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} \, dt. \tag{81}$$

Recall the notation in equation (22). By Lemma II.2, $\varkappa / \mathbb{E} \mathcal{M} \sim 1/S_\mu$. Therefore

$$\mathbb{E} \, G_\mu(x; \boldsymbol{p}) \sim S_\mu \left(\varkappa + x/\sqrt{u''(\varkappa)} \right) / S_\mu. \tag{82}$$

Thus it suffices to show that

$$\lim_{\varkappa \to \infty} \sqrt{u''(\varkappa)} \, e^{\mu\varkappa + u(\varkappa)} S_\mu = \int_{-\infty}^\infty e^{-t^2/2} \, dt, \tag{83}$$

$$\lim_{\varkappa \to \infty} \sqrt{u''(\varkappa)} \, e^{\mu\varkappa + u(\varkappa)} S_\mu \left(\varkappa + x/\sqrt{u''(\varkappa)} \right) = \int_x^\infty e^{-t^2/2} \, dt.$$

We will prove the first limit only; the second follows by a similar argument. By Lemma II.1, for any $\varepsilon > 0$, $S_\mu \sim S_\mu((1-\varepsilon)\varkappa, (1+\varepsilon)\varkappa)$. Thus

$$\sqrt{u''(\varkappa)} \, e^{u(\varkappa) + \mu\varkappa} S_\mu \sim \sum_{(1-\varepsilon)\varkappa \leq k \leq (1+\varepsilon)\varkappa} e^{-\mu(k-\varkappa) - (u(k)-u(\varkappa))} \sqrt{u''(\varkappa)} = \sum_{-\varepsilon\varkappa \leq k \leq \varepsilon\varkappa} e^{-\mu k - (u(k+\varkappa)-u(\varkappa))} \sqrt{u''(\varkappa)}. \tag{84}$$

By the assumption, $\lim_{x \to \infty} x^2 u''(x) = \infty$, and thus $\varepsilon\varkappa\sqrt{u''(\varkappa)} \to \infty$. As $u''(\varkappa) \to 0$,

$$\lim_{\varepsilon \searrow 0} \lim_{\mu \searrow \mu_*} \sum_{-\varepsilon\varkappa \leq k \leq \varepsilon\varkappa} e^{-u''(\varkappa)k^2/2} \sqrt{u''(\varkappa)} = \int_{-\infty}^\infty e^{-x^2/2} \, dx = \sqrt{2\pi}. \tag{85}$$

We now show that the exponential function in the last summation of formula (84) may be replaced by $e^{-u''(\varkappa)k^2/2}$. Using Taylor expansion, we can find $\xi = \xi(k, \varkappa)$ in between $\varkappa$ and $k + \varkappa$, such that

$$-\mu k - (u(k+\varkappa) - u(\varkappa)) = -[u(k+\varkappa) - u(\varkappa) - u'(\varkappa)k] = -\frac{u''(\xi)}{2} k^2 = -\frac{u''(\xi)}{u''(\varkappa)} \frac{1}{2} u''(\varkappa)k^2. \tag{86}$$

As $u''(x)$ is decreasing, we must have $u''((1+\varepsilon)\varkappa)/u''(\varkappa) \leq u''(\xi)/u''(\varkappa) \leq 1$ for $0 \leq k \leq \varepsilon\varkappa$; and $1 \leq u''(\xi)/u''(\varkappa) \leq u''((1-\varepsilon)\varkappa)/u''(\varkappa)$ for $-\varepsilon\varkappa \leq k < 0$. Therefore,

$$\sum_{0 \leq k \leq \varepsilon\varkappa} e^{-u''(\varkappa)k^2/2} \leq \sum_{0 \leq k \leq \varepsilon\varkappa} e^{-\mu k - (u(k+\varkappa)-u(\varkappa))} \leq \sum_{0 \leq k \leq \varepsilon\varkappa} e^{-\frac{u''((1+\varepsilon)\varkappa)}{u''(\varkappa)} u''(\varkappa)k^2/2}$$

$$\sum_{-\varepsilon\varkappa \leq k < 0} e^{-u''(\varkappa)k^2/2} \geq \sum_{-\varepsilon\varkappa \leq k < 0} e^{-\mu k - (u(k+\varkappa)-u(\varkappa))} \geq \sum_{-\varepsilon\varkappa \leq k < 0} e^{-\frac{u''((1-\varepsilon)\varkappa)}{u''(\varkappa)} u''(\varkappa)k^2/2} \tag{87}$$

17

As $x^2 u''(x)$ is increasing, $(1-\varepsilon)^2 u''((1-\varepsilon)\varkappa) \leq u''(\varkappa) \leq (1+\varepsilon)^2 u''((1+\varepsilon)\varkappa)$. Thus we conclude that

$$
\lim_{\varepsilon \searrow 0} \lim_{\mu \searrow \mu_*} \sum_{0 \leq k \leq \varepsilon \varkappa} e^{-\mu k - (u(k+\varkappa) - u(\varkappa))} \sqrt{u''(\varkappa_m)} = \int_0^\infty e^{-x^2/2}\, dx,
$$

$$
\lim_{\varepsilon \searrow 0} \lim_{\mu \searrow \mu_*} \sum_{-\varepsilon \varkappa \leq k < 0} e^{-\mu k - (u(k+\varkappa) - u(\varkappa))} \sqrt{u''(\varkappa)} = \int_{-\infty}^0 e^{-x^2/2}\, dx.
$$
(88)

And the limit in equation (83) follows. ∎

III.ii. Discrete Gaussian regime

When $\lim_{x \to \infty} u''(x) = c$ for some $0 < c < \infty$, for the shifted limit shape, no rescaling is needed in the "horizontal" direction and the shape is a discrete Gaussian.

Theorem 4
Assume that $\lim_{x \to \infty} u''(x) = c$ for some $0 < c < \infty$. Let

$$
G_\mu(x; \boldsymbol{p}) := \frac{\varkappa}{\mathbb{E}\mathcal{M}} \sum_{k - \varkappa \geq x} p_k.
$$
(89)

Then for each $\epsilon > 0$

$$
\lim_{\mu \searrow \mu_*} \mathbb{P}\left\{ \sup_{x \in \mathbb{R}} |G_\mu(x; \boldsymbol{p}) - G(x)| > \epsilon \right\} = 0
$$
(90)

where $G(x) = \frac{1}{M_c} \sum_{k \geq x} e^{-ck^2/2}$ and $M_c = \sum_{k \in \mathbb{Z}} e^{-ck^2/2}$.

Proof Similar to proof of Theorem 3, it suffices to show that for all $x \in \mathbb{R}$

$$
\mathbb{E}\, G_\mu(x; \boldsymbol{p}) = \frac{\varkappa}{\mathbb{E}\mathcal{M}} S_\mu(\varkappa + x) \to \frac{1}{M_c} \sum_{k \geq x} e^{-ck^2/2} \quad \text{as } \mu \searrow \mu_*.
$$
(91)

Recall that $\mu = -u'(\varkappa)$. Then for some $\xi = \xi(k, \varkappa)$ in between $\varkappa$ and k, we have

$$
e^{u(\varkappa) + \mu \varkappa} S_\mu(\varkappa + x) = e^{u(\varkappa) + \mu \varkappa} \sum_{k - \varkappa \geq x} e^{-\mu k - u(k)} = \sum_{k - \varkappa \geq x} e^{-u''(\xi)(k - \varkappa)^2/2}.
$$
(92)

Notice that $\varkappa \nearrow \infty$ as $\mu \searrow \mu_*$. Then, for μ close to μ_*, we can relabel $k - \varkappa$ as k so that

$$
\sum_{k - \varkappa \geq x} e^{-u''(\xi)(k - \varkappa)^2/2} = \sum_{k \geq x} e^{-u''(\xi)k^2/2} \to \sum_{k \geq x} e^{-ck^2/2}.
$$
(93)

In the last convergence we have used $\xi \geq \min\{\varkappa, \varkappa + x\}$ and $\lim_{k \to \infty} u''(k) = c$. It remains to consider the contribution from $\varkappa/\mathbb{E}\mathcal{M}$. By Lemmas II.1 and II.2, for any fixed $\varepsilon \in (0, 1)$, we have

$$
e^{u(\varkappa) + \mu \varkappa} \frac{\mathbb{E}\mathcal{M}}{\varkappa} \sim e^{u(\varkappa) + \mu \varkappa} \sum_{k \geq (1-\varepsilon)\varkappa} e^{-\mu k - u(k)} = \sum_{k \geq -\varepsilon \varkappa} e^{-u''(\xi)k^2/2}.
$$
(94)

18

Here we have $\xi \geq (1 - \varepsilon)\varkappa$. Notice that $u''(k) \to c > 0$. Sending $\mu \searrow \mu_*$, we obtain

$$e^{u(\varkappa)+\mu\varkappa} \frac{\mathbb{E}\mathcal{M}}{\varkappa} \sim \sum_{k > -\infty} e^{-ck^2/2} =: M_c. \tag{95}$$

Formula (91) follows. ∎

III.iii. Hard step function regime

In the case when $\lim_{x \to \infty} u''(x) = \infty$, the distribution of p_k is so concentrated around $\varkappa$ that, for any scaling in x direction, the local limit shape function is always the step function $\mathbb{1}_{(-\infty,0]}(x)$.

Theorem 5
Assume that $\lim\limits_{x \to \infty} u''(x) = \infty$. For any $\zeta = \zeta(\mu) \nearrow \infty$ as $\mu \searrow \mu_*$, let

$$G_\mu(x;\boldsymbol{p}) := \frac{\varkappa}{\mathbb{E}\mathcal{M}} \sum_{k - \varkappa \geq \zeta x} p_k. \tag{96}$$

Then for each or each $\lambda_1 < 0$, $\lambda_2 > 0$, and $\epsilon > 0$,

$$\lim_{\mu \searrow \mu_*} \mathbb{P}\left\{ \sup_{x \in \mathbb{R} \setminus (\lambda_1, \lambda_2)} |G_\mu(x;\boldsymbol{p}) - G(x)| > \epsilon \right\} = 0. \tag{97}$$

where $G(x) = \mathbb{1}_{(-\infty,0]}(x)$.

Proof It suffices to show that $\lim_{\mu \searrow \mu_*} \mathbb{E} G_\mu(x;\boldsymbol{p}) = G(x)$ for all $x \neq 0$. Following the proof of Theorem 1, we show that $\sum_{k=1}^\infty e^{-\mu k - u(k)}$ concentrates on $k = \lfloor \varkappa \rfloor - 1, \lfloor \varkappa \rfloor, \lfloor \varkappa \rfloor + 1$, and $\lfloor \varkappa \rfloor + 2$ where $\lfloor \varkappa \rfloor$ is the integer part of $\varkappa$.

Step 1. Recall notation in equation (22). Let us first show that $\lim_{\mu \searrow \mu_*} S_\mu(\varkappa + 3)/S_\mu = 0$. Take any two integers a, b such that $\lfloor \varkappa \rfloor + 1 < a < b$. Recall that $\mu = -u'(\varkappa)$. Then for some ξ, ξ' with $a \leq \xi \leq b$ and $\varkappa \leq \xi' \leq \xi$,

$$\frac{e^{-\mu b - u(b)}}{e^{-\mu a - u(a)}} = e^{u'(\varkappa)(b-a) - u'(\xi)(b-a)} = e^{-(\xi - \varkappa)(b-a)u''(\xi')} \leq e^{-(a-\varkappa-1)(b-a)u''(\xi')} \leq e^{-(b-a)u''(\xi')}. \tag{98}$$

Here we used $\xi - \varkappa \geq a - (\lfloor \varkappa \rfloor + 1) \geq 1$. Take some $A > 0$. As $u''(x) \to \infty$, $e^{-\mu b - u(b)}/e^{-\mu a - u(a)} \leq e^{-(b-a)A}$. Therefore for all large enough $\varkappa$-s, $S_\mu(\varkappa + 3)/S_\mu \leq S_\mu(\varkappa + 3)/S_\mu(2) \leq e^{-A}$. Sending $A \to \infty$, we establish the desired limit.

Step 2. It remains to show that $\lim_{\mu \searrow \mu_*} S_\mu(1, \varkappa - 1)/S_\mu = 0$. By Lemma II.1, we may ignore first finitely many terms in the sum $S_\mu(1, \varkappa - 1)$, i.e., for any fixed N, $\lim_{\mu \searrow \mu_*} S_\mu(1, N)/S_\mu = 0$. As $u''(x) \to \infty$, we may find $A(N)$ such that $\lim_{N \to \infty} A(N) = \infty$ and $u''(x) \geq A(N)$ for $x \geq N$. By a similar argument as in Step 1, $S_\mu(1, \varkappa - 1)/S_\mu \leq S_\mu(1, \varkappa - 1)/S_\mu(N + 1, \varkappa) \leq e^{-A(N)}$. The proposition follows after sending $\mu \to \mu_*$ and then $N \to \infty$. ∎

III.iv. Remark on the monotonicity requirement

In section III.i we required that $x^2 u''(x)$ be non-decreasing and $u''(x)$ be non-increasing to get the Gaussian limit for the local shape function. These monotonicity assumptions are essential in

the sense that if either of them is not satisfied, either the Gaussian limit is not ensured or different limit shapes may occur along different subsequences of μ. Let us construct an example illustrating this phenomenon.

First of all, using the scaling relation $u'(\varkappa) = -\mu$, we can deduce that

$$\int_{\varkappa}^{\varkappa+k} (\varkappa + k - x) u''(x) \, dx = \mu k + u(k + \varkappa) - u(\varkappa). \tag{99}$$

Define $u''(\cdot)$ in the following way: first let $u''(x) = 1$ for $0 < x \le 1$, $u''(x) = 1/2^n$ for $2^{n-1} < x \le 2^n$, $n \in \mathbb{N}$. (One could modify u'' in $[2^n, 2^n + 2^{-n}]$ by interpolating between the end-point values to obtain a continuous function, but it's not required.) It is not hard to check that $x^2 u''(x)$ tends to infinity, however not in a monotone manner. Recall the local shape function $G_\mu(x; \boldsymbol{p})$ as in (78). We will see that the resulting limit shape is only determined up to subsequences.

First, pick the sequence $\varkappa_n = 3 \cdot 2^{n-1}$, $n \in \mathbb{N}$. The corresponding μ_n is then determined by the relation (14). Assume $\varepsilon < 1/3$. Note that $u''(x) = u''(\varkappa_n)$ for $(1 - \varepsilon)\varkappa_n \le x \le (1 + \varepsilon)\varkappa_n$. Then, using $\mu_n = -u'(\varkappa_n)$, we get,

$$\sum_{-\varepsilon\varkappa_n \le k \le \varepsilon\varkappa_n} \exp\left\{-\mu_n k - \left[u(k + \varkappa_n) - u(\varkappa_n)\right]\right\} = \sum_{-\varepsilon\varkappa_n \le k \le \varepsilon\varkappa_n} \exp\left\{-\frac{k^2}{2} u''(\varkappa_n)\right\} \tag{100}$$

Repeating the arguments of Theorem 3, we obtain the standard Gaussian $G(x) = \int_x^\infty e^{-t^2/2} \, dt / \sqrt{2\pi}$ as the limit of $G_\mu(x; \boldsymbol{p})$.

We now show that a different limit shape may appear if we follow another sequence of $\varkappa$. Let $\tilde{\varkappa}_n = 2^n$ and, accordingly, $\tilde{\mu}_n = -u'(\tilde{\varkappa}_n)$. Assume $\varepsilon < 1/2$. Now $u''(x)$ behaves differently on the different sides of $\varkappa_n$: $u''(x) = u''(\varkappa_n)$ for $(1 - \varepsilon)\varkappa_n \le k \le \varkappa_n$ and $u''(x) = u''(\varkappa_n)/2$ for $\varkappa_n < k \le (1 + \varepsilon)\varkappa_n$:

$$\sum_{-\varepsilon\varkappa_n \le k \le 0} \exp\left\{-\mu_n k - \left[u(k + \varkappa_n) - u(\varkappa_n)\right]\right\} = \sum_{-\varepsilon\varkappa_n \le k \le 0} \exp\left\{-\frac{k^2}{2} u''(\varkappa_n)\right\};$$

$$\sum_{0 < k \le \varepsilon\varkappa_n} \exp\left\{-\mu_n k - \left[u(k + \varkappa_n) - u(\varkappa_n)\right]\right\} = \sum_{0 < k \le \varepsilon\varkappa_n} \exp\left\{-\frac{k^2}{4} u''(\varkappa_n)\right\}. \tag{101}$$

Repeating the arguments of Theorem 3 again, we now find that $G_\mu(x; \boldsymbol{p})$ converges to

$$\tilde{G}(x) = \frac{\sqrt{2}}{(1 + \sqrt{2})\sqrt{\pi}} \int_x^\infty h(t) \, dt, \tag{102}$$

where $h(t) = e^{-t^2/4}$ for $t > 0$ and $h(t) = e^{-t^2/2}$ for $t \le 0$.

Notice, however, that a sequence of $\varkappa_n$ realizing the Gaussian limit as in Theorem 3 always exists if we require monotonicity of $u''(x)$. Indeed, by Lemma A.3, we may find positive sequences $\varepsilon_n \downarrow 0$ and $\varkappa_m \uparrow \infty$, such that

$$\lim_{m \to \infty} \varepsilon_m \varkappa_m \sqrt{u''(\varkappa_m)} = \infty, \quad \limsup_{m \to \infty} \frac{u''((1 - \varepsilon_m)\varkappa_m)}{u''(\varkappa_m)} = 1, \quad \lim_{n \to \infty} \liminf_{m \to \infty} \frac{u''((1 + \varepsilon_n)\varkappa_m)}{u''(\varkappa_m)} = 1. \tag{103}$$

20

Thus we have $u''((1 + \varepsilon_n)\varkappa_m)/u''(\varkappa_m) \leq u''(\xi)/u''(\varkappa_m) \leq 1$ for all $\xi \in [\varkappa_m, (1 + \varepsilon_n)\varkappa_m)]$; and also $1 \leq u''(\xi)/u''(\varkappa_m) \leq u''((1 - \varepsilon_m)\varkappa_m)/u''(\varkappa_m)$ for all $\xi \in [(1 - \varepsilon_m)\varkappa_m, \varkappa_m)$. Therefore, using Taylor expansion, cf. (86),

$$\sum_{0 \leq k \leq \varepsilon_n \varkappa_m} e^{-u''(\varkappa_m)k^2/2} \leq \sum_{0 \leq k \leq \varepsilon_n \varkappa_m} e^{-\mu k - (u(k+\varkappa_m) - u(\varkappa_m))} \leq \sum_{0 \leq k \leq \varepsilon_n \varkappa_m} e^{-\frac{u''((1+\varepsilon_n)\varkappa_m)}{u''(\varkappa_m)} u''(\varkappa_m)k^2/2};$$

$$\sum_{-\varepsilon_n \varkappa_m \leq k < 0} e^{-u''(\varkappa_m)k^2/2} \geq \sum_{-\varepsilon_n \varkappa_m \leq k < 0} e^{-\mu k - (u(k+\varkappa_m) - u(\varkappa_m))} \geq \sum_{-\varepsilon_m \varkappa_m \leq k < 0} e^{-\frac{u''((1-\varepsilon_m)\varkappa_m)}{u''(\varkappa_m)} u''(\varkappa_m)k^2/2}. \qquad (104)$$

Taking $m \to \infty$ and then $n \to \infty$, we obtain the Gaussian limit by (103).

A. Appendix

The following lemmas establish a few asymptotic properties of functions satisfying Assumption 1:

Lemma A.1
Assume that $u(\cdot)$ is a twice-differentiable function satisfying

$$\lim_{x \to \infty} x^2 u''(x) = \gamma \in \mathbb{R}. \qquad (105)$$

Then there exist $c \in \mathbb{R}$ and $v(x)$ satisfying $\lim_{x \to \infty} x v'(x) = 0$, such that

$$\lim_{x \to \infty} u'(x) = c, \quad u(x) = cx - \gamma \ln x + v(x). \qquad (106)$$

Notice that we also have, $v(x) \ll \ln x$ as $x \to \infty$.

Proof   Equation (105) implies that for any $\varepsilon > 0$ there exists x_0, such that $\gamma - \varepsilon < x^2 u''(x) < \gamma + \varepsilon$ for all $x > x_0$. Dividing these inequalities by x^2 and integrating, we conclude that

$$|u'(x_1) - u'(x_2)| \leq (|\gamma| + \varepsilon)|1/x_1 - 1/x_2| \qquad (107)$$

for all $x_1, x_2 > x_0$. Thus $\lim_{x \to \infty} u'(x)$ exists and is finite. Let us denote this limit by c and define

$$v(x) := u(x) - cx + \gamma \ln x. \qquad (108)$$

Directly from this definition we can calculate that $\lim_{x \to \infty} x^2 v''(x) = \lim_{x \to \infty} v'(x) = 0$. Then for any $\varepsilon > 0$ and x, x_0 large enough, we also have $|v'(x) - v'(x_0)| \leq \varepsilon |1/x - 1/x_0|$. Sending x_0 to infinity, we obtain $-\varepsilon/x \leq v'(x) \leq \varepsilon/x$, from which we conclude that $\lim_{x \to \infty} x v'(x) = 0$ and $v(x) \ll \ln x$.   ■

Lemma A.2
Assume that $u(\cdot)$ is a twice-differentiable function satisfying

$$\lim_{x \to \infty} x^2 u''(x) = \infty. \qquad (109)$$

Then the following holds:

1. $\mu_* := -\lim_{x \to \infty} \dfrac{u(x)}{x} = -\lim_{x \to \infty} u'(x).$

21

2. *If $\mu_* = -\infty$, then $u(x) \gg x$. If $\mu_* \in \mathbb{R}$, then $\ln x \ll u(x) + \mu_* x \ll x$ and $\lim\limits_{x\to\infty} [u(x) + \mu_* x] = -\infty$.*

3. $\lim\limits_{\mu \searrow \mu_*} S_\mu = \infty$ *(refer to p.8 for the definition of S_μ).*

Proof   *1.* The assumption (109) implies that $u''(x) > 0$ for all large enough x. Thus both limits $\lim_{x\to\infty} u'(x)$ and $\lim_{x\to\infty} u(x)$ exist. Observe that if the latter is finite, we must also have $\lim_{x\to\infty} u'(x) = 0$, which implies $\mu_* = -\lim_{x\to\infty} u'(x)$. If $\lim_{x\to\infty} u(x) = \pm\infty$, $\mu_* = -\lim_{x\to\infty} u'(x)$ as well by L'Hospital's rule. Note that this assertion also holds by similar argument if the limit in equation (109) were negative infinity.

2. If $\mu_* = -\infty$, i.e., $u'(x) \to \infty$ as $x \to \infty$, $u(x)$ must be superlinear, i.e., $u(x) \gg x$. If $\mu_* \in \mathbb{R}$, $u(x) + \mu_* x$ must be sublinear. At the same time, property (109) implies that for any fixed $M > 0$ and all sufficiently large x, $u''(x) \geq M/x^2$. Integrating from x to infinity, we get $u'(x) + \mu_* \leq -M/x$. Another integration yields that $\lim_{x\to\infty}[u(x) + \mu_* x] \leq -M \ln x + C \to -\infty$ as $x \to \infty$.

3. If $\mu_* = -\infty$, $S_\mu \to \infty$, because every term in the series (22) diverges as $\mu \searrow \mu_*$. If $\mu_* \in \mathbb{R}$, the monotone convergence implies that $S_\mu \to S_{\mu_*} = \infty$. The latter series diverges, as its terms grow unboundedly by assertion (2). ∎

The following lemma is used in Section III.iv to demonstrate existence of non-Gaussian local shapes when the monotonicity assumption is omitted:

Lemma A.3
Consider a non-increasing function, $f : (0,\infty) \mapsto (0,\infty)$. Assume $\lim_{x\to\infty} x^2 f(x) = \infty$. Then there exist sequences $\varepsilon_n \searrow 0$ and $x_k \nearrow \infty$, such that

$$\lim_{k\to\infty} \varepsilon_k x_k \sqrt{f(x_k)} = \infty, \quad \limsup_{k\to\infty} \frac{f(x_k - \varepsilon_k x_k)}{f(x_k)} = 1, \quad \lim_{n\to\infty} \liminf_{k\to\infty} \frac{f(x_k + \varepsilon_n x_k)}{f(x_k)} = 1. \tag{110}$$

Proof   Fix any sequence $\tau_n \searrow 0$ and assume $\tau_n < 1$ for all n. Let α_n to be a sequence such that $0 < (1 + \tau_n)^2 \alpha_n < 1$ for all n and $\lim_{n\to\infty} \alpha_n = 1$. Since $x^2 f(x) \to \infty$, we can find a sequence z_k such that $\tau_k x \sqrt{f(x)} \geq k$ whenever $x \geq z_k$. We first show that there exist $y_k \nearrow \infty$, such that

$$y_k \geq z_k, \quad f(y_k + \tau_n y_k)/f(y_k) \geq \alpha_n, \text{ for all } n \leq k. \tag{111}$$

We first show that there exists $y_1 \geq z_1$ such that

$$f(y_1 + \tau_1 y_1)/f(y_1) \geq \alpha_1. \tag{112}$$

Suppose it does not hold, i.e. for all $x \geq z_1$ we have

$$f(x + \tau_1 x) < \alpha_1 f(x). \tag{113}$$

Let $u_k = (1 + \varepsilon_1)^k z_1$. Then $f(u_k) \leq \alpha_1{}^k f(z_1)$. Therefore, as $0 < (1 + \tau_1)^2 \alpha_1 < 1$

$$\lim_{k\to\infty} u_k^2 f(u_k) \leq \lim_{k\to\infty} \left[(1 + \tau_1)^2 \alpha_1\right]^k z_1^2 f(z_1) = 0 \tag{114}$$

which contradicts the assumption that $\lim_{x\to\infty} x^2 f(x) = \infty$.
We now show that we may find $y_2 \geq z_2$ such that

$$f(y_2 + \tau_1 y_2)/f(y_2) \geq \alpha_1, \quad f(y_2 + \tau_2 y_2)/f(y_2) \geq \alpha_2. \tag{115}$$

22

Assume the opposite. Then for all $x \geq z_2$,

$$f(x + \tau_1 x) < \alpha_1 f(x) \quad \text{or} \quad f(x + \tau_2 x) < \alpha_2 f(x). \tag{116}$$

Let $u_0 = z_2$ and $u_k = (1+\tau_1)^{m_k}(1+\tau_2)^{n_k} u_{k-1}$ where $m_k + n_k = 1$, $m_k = 1$ if $f(u_{k-1} + \tau_1 u_{k-1}) < \alpha_1 f(u_{k-1})$, $m_k = 0$ otherwise. When $n_k = 1$, we have $f(u_{k-1} + \tau_2 u_{k-1}) < \alpha_2 f(u_{k-1})$. Therefore,

$$f(u_k) < \alpha_1^{\sum_{j=1}^{k} m_j} \alpha_2^{\sum_{j=1}^{k} n_j} f(u_0), \tag{117}$$

and

$$u_k^2 f(u_k) < \left[(1+\tau_1)^2 \alpha_1\right]^{\sum_{j=1}^{k} m_j} \left[(1+\tau_2)^2 \alpha_2\right]^{\sum_{j=1}^{k} n_j} u_0^2 f(u_0). \tag{118}$$

As $\sum_{j=1}^{k} m_j + \sum_{j=1}^{k} n_j = k$, at least one of the two sums goes to infinity as $k \to \infty$. Therefore we get $\lim_{k \to \infty} u_k^2 f(u_k) = 0$ which contradicts the assumption that $\lim_{k \to \infty} x^2 f(x) = \infty$. Similarly, for any $k \geq 3$ we can find y_k such that (111) holds.

Define $\varepsilon_n := \tau_n/4$ and $x_k := (1 + \tau_k/2)y_k$. We now show that (111) implies the lemma. As $x_k \geq y_k \geq z_k$, it follows that $\varepsilon_k x_k \sqrt{f(x_k)} \to \infty$. Note that, as $\tau_k < 1$,

$$y_k = \left(1 - \frac{\tau_k/2}{1 + \tau_k/2}\right) x_k \leq \left(1 - \frac{\tau_k}{4}\right) x_k = (1 - \varepsilon_k) x_k. \tag{119}$$

Therefore

$$\limsup_{k \to \infty} \frac{f(x_k - \varepsilon_k x_k)}{f(x_k)} \leq \limsup_{k \to \infty} \frac{f(y_k)}{f(x_k)} \leq \limsup_{k \to \infty} \frac{f(y_k)}{f(y_k + \tau_k y_k)} = 1. \tag{120}$$

Finally, notice that for any fixed n and all large enough k, $(1 + \tau_n/4)(1 + \tau_k/2) \leq 1 + \tau_n$. Thus

$$\liminf_{k \to \infty} \frac{f(x_k + \varepsilon_n x_k)}{f(x_k)} = \liminf_{k \to \infty} \frac{f((1 + \varepsilon_n)(1 + \tau_k/2)y_k)}{f((1 + \tau_k/2)y_k)} \geq \liminf_{k \to \infty} \frac{f((1 + \tau_n)y_k)}{f(y_k)} \geq \alpha_n. \tag{121}$$

Take $n \to \infty$ and note that $\lim_{n \to \infty} \alpha_n = 1$ to complete the proof. ∎

References

[1] R. Arratia, A. D. Barbour, and S. Tavaré. A tale of three couplings: Poisson-Dirichlet and GEM approximations for random permutations. *Combinatorics, Probability & Computing*, 15(1-2):31–62, 2006.

[2] V. Betz, H. Schäfer, and D. Zeindler. Random permutations without macroscopic cycles. *Annals of Applied Probability*, 30(3):1484–1505, 2020.

[3] V. Betz, D. Ueltschi, Y Velenik, et al. Random permutations with cycle weights. *The Annals of Applied Probability*, 21(1):312–331, 2011.

[4] L. V. Bogachev. Unified derivation of the limit shape for multiplicative ensembles of random integer partitions with equiweighted parts. *Random Structures & Algorithms*, 47(2):227–266, 2015.

[5] S. Chatterjee and P. Diaconis. Fluctuations of the Bose-Einstein condensate. *Journal of Physics A: Mathematical and Theoretical*, 47(8):085201, 2014.

[6] A. Cipriani and D. Zeindler. The limit shape of random permutations with polynomially growing cycle weights. *ALEA, Lat. Am. J. Probab. Math. Stat.*, 12(2):971–999, 2015.

[7] N. M. Ercolani and D. Ueltschi. Cycle structure of random permutations with cycle weights. *Random Structures & Algorithms*, 44(1):109–133, 2014.

[8] M. M. Erlihson and B. L. Granovsky. Limit shapes of Gibbs distributions on the set of integer partitions: the expansive case. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 44(5):915–945, 2008.

[9] I. Fatkullin and V. Slastikov. Limit shapes for Gibbs ensembles of partitions. *Journal of Statistical Physics*, 172(6):1545–1563, 2018.

[10] L. Holst. The Poisson-Dirichlet distribution and its relatives revisited. *Preprint of the Royal Institute of Technology, Stockholm, Sweden*, 2001.

[11] J. Kerstan, K. Matthes, and J. Mecke. *Infinitely divisible point processes*. Wiley series in probability and mathematical statistics. Wiley, 1978.

[12] J. F. C. Kingman. Random discrete distributions. *Journal of the Royal Statistical Society: Series B (Methodological)*, 37(1):1–15, 1975.

[13] J. F. C. Kingman. The population structure associated with the Ewens sampling formula. *Theoretical Population Biology*, 11(2):274–283, 1977.

[14] J. Pitman. Combinatorial stochastic processes. *Lecture Notes in Mathematics; UC Berkeley*, 2006.

[15] N. Robles and D. Zeindler. Random permutations with logarithmic cycle weights. *Annales de l'Institut Henri Poincare, Probabilites et Statistiques*, 56(3):1991–2016, 2020.

[16] A. Vershik. Statistical mechanics of combinatorial partitions, and their limit shapes. *Functional Analysis and Its Applications*, 30(2):90–105, 1996.

[17] A. Vershik and A. Shimdt. Limit measures arising in the asymptotic theory of symmetric groups. i. *Theory of Probability & Its Applications*, 22(1):70–85, 1977.

[18] A. Vershik and A. Shmidt. Limit measures arising in the asymptotic theory of symmetric groups, ii. *Theory of Probability & Its Applications*, 23(1):36–49, 1978.

[19] Yu. V. Yakubovich. Asymptotics of random partitions of a set. *Zapiski Nauchnykh Seminarov POMI*, 223:227–250, 1995.