Poisson-Dirichlet asymptotics in condensing particle systems[*]

Paul Chleboun[†]   Simon Gabriel[‡]   Stefan Grosskinsky[§]

Abstract

We study measures on random partitions, arising from condensing stochastic particle systems with stationary product distributions. We provide fairly general conditions on the stationary weights, which lead to Poisson-Dirichlet statistics of the condensed phase in the thermodynamic limit. The Poisson-Dirichlet distribution is known to be the unique reversible measure of split-merge dynamics for random partitions, which we use to characterize the limit law. We also establish concentration results for the macroscopic phase, using size-biased sampling techniques and the equivalence of ensembles to characterize the bulk distribution of the system.

Keywords: Poisson-Dirichlet distribution; split-merge dynamics; random partitions; equivalence of ensembles; interacting particle systems; condensation.
MSC2020 subject classifications: Primary 60K35, Secondary 82C22; 82C26.
Submitted to EJP on June 29, 2021, final version accepted on November 9, 2022.

1 Introduction and results

1.1 Mathematical setting and motivation

The results presented in this paper are motivated by the study of interacting particle systems. We consider finite systems consisting of N particles on L sites indexed by the set Λ, where $|\Lambda| = L$. For simplicity we take $\Lambda = \{1, \ldots, L\}$ in the remainder of this paper. The space of such particle configurations $\eta = (\eta_1, \ldots, \eta_L)$ with $\eta_x \in \mathbb{N}_0$ is given by

$$\Omega_{L,N} := \left\{ \eta \in \mathbb{N}_0^L : \sum_{x=1}^L \eta_x = N \right\},$$

[†]Department of Statistics, University of Warwick, Coventry, CV4 7AL, United Kingdom.
 E-mail: paul.i.chleboun@warwick.ac.uk
[‡]Mathematics Institute, University of Warwick, Coventry, CV4 7AL, United Kingdom.
 E-mail: Simon.Gabriel@warwick.ac.uk
[§]Lehrstuhl für Stochastik und ihre Anwendungen Universität Augsburg, D-86135 Augsburg, Germany.
 E-mail: Stefan.Grosskinsky@math.uni-augsburg.de

which we equip with the discrete topology. The dynamics of systems we consider are assumed to be irreducible Markov processes on $\Omega_{L,N}$, conserving only the quantities N and L. Thus, there exists a unique invariant distribution $\pi_{L,N}$ of the system on $\Omega_{L,N}$, which is called the *canonical distribution*.

We will focus on models where the canonical distributions are of product form

$$\pi_{L,N}[d\eta] = \begin{cases} \dfrac{1}{Z_{L,N}} \displaystyle\prod_{x=1}^{L} w_L(\eta_x)\, d\eta & \text{if } \sum_{x=1}^{L} \eta_x = N\,, \\ 0 & \text{otherwise.} \end{cases} \tag{1.1}$$

Here $d\eta$ denotes the counting measure on $\Omega_{L,N}$ and $(w_L(n))_{n\in\mathbb{N}_0}$ is a sequence of positive weights, possibly depending on the system size L. The normalising constant, called *canonical partition function*, is given as

$$Z_{L,N} = \sum_{\eta\in\Omega_{L,N}} \prod_{x=1}^{L} w_L(\eta_x)\,.$$

Note that the weights $(w_L(n))_{n\in\mathbb{N}_0}$ are independent of the site x, thus the $\pi_{L,N}$ are permutation invariant and in particular spatially homogeneous, so that single-site marginals $\pi_{L,N}[\eta_x \in \cdot\,]$ do not depend on x.

We are primarily interested in the limiting behaviour of $\pi_{L,N}$ in the thermodynamic limit $N, L \to \infty$ such that $\frac{N}{L}$ converges to $\rho \geqslant 0$, which we will subsequently abbreviate by $N/L \to \rho$. Assume for now that the weak limit of the single-site marginals exists for all $\rho \geqslant 0$,

$$\pi_{L,N}[\eta_x \in \cdot\,] \to \nu_\rho[\cdot\,] \quad \text{as } N/L \to \rho\,, \tag{1.2}$$

and the limit is a probability measure on $\mathbb{N}_0$. This implies in particular convergence of the expectations for bounded functions $f : \mathbb{N}_0 \mapsto \mathbb{R}$, i.e.

$$\sum_{n=0}^{\infty} f(n)\pi_{L,N}[\eta_x = n] \to \nu_\rho(f)\,,$$

where we write $\nu_\rho(f)$ for the expectation of f under ν_ρ. Sometimes we will also write $\nu_\rho(f(\eta_x))$ for the corresponding expectation if needed for clarity. Looking at a single site's expected occupation number under $\pi_{L,N}$, we see that due to spatial homogeneity we have

$$\pi_{L,N}(\eta_x) = \frac{1}{L}\sum_{y=1}^{L} \pi_{L,N}(\eta_y) = \frac{1}{L}\pi_{L,N}\Big(\sum_{y=1}^{L} \eta_y\Big) = \frac{N}{L} \to \rho\,,$$

when taking the thermodynamic limit. However, because the identity $f(n) = n$ is an unbounded function, we cannot guarantee that the particle density of the system is conserved in the limit ν_ρ and $\nu_\rho(\eta_x)$ may be strictly smaller than ρ. This phenomenon is known as condensation.

Definition 1.1 (Condensation). *A system characterised by spatially homogeneous canonical distributions $(\pi_{L,N})_{L,N}$ exhibits **condensation** in the thermodynamic limit $N/L \to \rho$ if ν_ρ in (1.2) exists and*

$$\nu_\rho(\eta_x) < \rho = \lim_{N/L\to\rho} \pi_{L,N}(\eta_x)\,.$$

*Furthermore, we say that the system has a **condensation transition** with **critical density** $\rho_c \geqslant 0$ if*

$$\nu_\rho(\eta_x) \begin{cases} = \rho & \text{if } \rho < \rho_c\,, \\ < \rho & \text{if } \rho > \rho_c\,. \end{cases}$$

In the context of stochastic particle systems, condensation means that a positive fraction of the total density ρ is not observed in the thermodynamic limit, since it concentrates on sites with diverging occupation numbers called the *condensed phase*. Clearly, the number of such sites has a vanishing volume fraction and does not contribute to the weak limit ν_ρ, which describes the distribution of the *background* or *bulk phase*.

Condensation in homogeneous stochastic particle systems has been studied previously in great generality, partially reviewed e.g. in [6, 14, 20]. Early results are formulated in the context of zero-range processes in [10, 13, 19] and in [29, 26, 2, 1] on a rigorous level, where the condensed phase concentrates on a single lattice site. Our goal here is to understand details of the condensed phase when it extends over more than one site and exhibits a non-trivial structure. Such structures have previously been observed as a result of spatial correlations [40, 39, 38] and as a result of L-dependent stationary weights, with a soft cut-off for site occupation numbers under zero-range dynamics [37] or in the inclusion process [28].

On the level of particle configurations the condensed phase disappears in the thermodynamic limit due to its vanishing volume fraction. To study its structure, it is more useful to interpret a configuration as an ordered partition of the total mass. For models of type (1.1) partitions and particle configurations are equivalent, since $\pi_{L,N}$ is permutation invariant and the underlying lattice structure is irrelevant.

We will represent particle configurations rescaled by the total mass N as ordered partitions of the unit interval $[0,1]$ on the set

$$\overline{\nabla} := \left\{ p = (p_i)_i \in [0,1]^{\mathbb{N}} : \sum_{i=1}^{\infty} p_i \leqslant 1 \quad \text{and} \quad p_1 \geqslant p_2 \geqslant \cdots \right\}. \tag{1.3}$$

We use the map $T = T(L, N) : \Omega_{L,N} \to \overline{\nabla}$ with

$$T(\eta) := \frac{1}{N} \left(\widehat{\eta}_1, \ldots, \widehat{\eta}_L, 0, \ldots \right), \tag{1.4}$$

where $\widehat{\eta} = (\widehat{\eta}_1, \ldots, \widehat{\eta}_L)$ denotes the entries in η in decreasing order with $\widehat{\eta}_1 \geqslant \widehat{\eta}_2 \geqslant \ldots \geqslant \widehat{\eta}_L$. Since any permutation of entries in η yields the same partition in $\overline{\nabla}$, the map T is not injective. Thus, the push-forward measure of $\pi_{L,N}$ under T on $\overline{\nabla}$ is given by

$$\mu_{L,N}[dp] := \pi_{L,N} \circ T^{-1}[dp] = \pi_{L,N}[d(Np)] \left| T^{-1}(\{p\}) \right|, \tag{1.5}$$

with Np denoting any configuration in $\Omega_{L,N}$ inducing the finite ordered partition $p \in \overline{\nabla}$. Note that $\mu_{L,N}$ concentrates on finite partitions with at most L non-zero entries and $T^{-1}(\{p\}) = \emptyset$ otherwise. In fact, the $\mu_{L,N}$ further concentrate on the subset where $\sum_{i=1}^{\infty} p_i = 1$. However, this space is not compact, unlike $\overline{\nabla}$ which is compact w.r.t. the product topology by Tychonoff's theorem, ensuring existence of subsequential weak limits of $\mu_{L,N}$ in the thermodynamic limit $N/L \to \rho$. The objective of this article is to identify general assumptions on the weights $(w_L)_L$, such that for $\rho > 0$ large enough $\mu_{L,N}$ converges weakly to a Poisson-Dirichlet distribution as $N/L \to \rho$. Details on this distribution are introduced in Section 2.

The starting point of our analysis is the recent paper [28] in which weights of the form

$$w_L(n) = \frac{\Gamma(n+d)}{n!\,\Gamma(d)}, \tag{1.6}$$

with $d = d(L) \in \mathbb{R}$ such that $\lim_{L\to\infty} dL = \theta \in (0,\infty)$, are considered. Such weights emerge for example from the dynamics of the inclusion process introduced in [17, 5], which can also be applied in population genetics as a multi-species Moran model [35], with the above scaling of the parameter d corresponding to a small mutation rate. With

weights (1.6) the system (1.1) exhibits a condensation transition with critical density $\rho_c = 0$, leaving an empty bulk behind, i.e. $\nu_\rho(\eta_x) = 0$. For the condensed phase in this model we have

$$\mu_{L,N} \xrightarrow{d} \mathrm{PD}(\theta), \quad \text{as } N/L \to \rho \text{ for all } \rho > 0, \tag{1.7}$$

where $\mathrm{PD}(\theta)$ denotes the Poisson-Dirichlet distribution with parameter θ [28, Theorem 1]. The proof uses the fact that $\pi_{L,N}$ with weights (1.6) is a Dirichlet multinomial distribution which permits an exact, simple expression for the corresponding partition function $Z_{L,N}$. This leads to exact expressions for the distribution of size-biased marginals which characterize the Poisson-Dirichlet limit (cf. Section 2.1).

Our main result provides a generalization to models with more general weights that do not lead to exact expressions for $Z_{L,N}$, and with non-trivial bulk distribution where $0 < \nu_\rho(\eta_x) < \rho$. In our proof, we not only make use of the Poisson-Dirichlet distribution's characterisation via size-biased sampling, which was essential for the arguments in [28], but also use the characterisation as the unique reversible distribution under split-merge dynamics as explained in Section 2.1. This allows us to avoid explicit expressions or approximations of the partition function $Z_{L,N}$ which are not always at hand. Our approach is motivated by a recent paper by Ioffe and Tóth [27], where the embedding of integer configurations into partitions of $[0,1]$ was used to show convergence of cycle-length processes of stationary random stirring to the split-merge dynamics.

1.2 Main results

We recall from (1.1) that the $\pi_{L,N}$ are probability measures on $\Omega_{L,N}$ given by

$$\pi_{L,N}[d\eta] = \frac{1}{Z_{L,N}} \prod_{x=1}^{L} w_L(\eta_x)\, d\eta.$$

For our first result we fix a density $\rho > 0$ and choose $N, L \to \infty$ such that $N/L \to \rho$. We assume that

(A1) $w_L(n) \geqslant 0$ for all $n \in \mathbb{N}$ and the limit

$$\lim_{L \to \infty} w_L(n) =: w(n) \geqslant 0 \quad \text{exists for all fixed } n \in \mathbb{N}_0,$$

such that w is summable and non-trivial,

and a weak form of the equivalence of ensembles:

(A2) The limiting probability distribution (1.2) exists and is of the form

$$\pi_{L,N}[\eta_1 = n] \to \nu_\rho[n] = \frac{w(n)\phi^n}{\mathcal{Z}} \quad \forall n \geqslant 0,$$

for some $\phi > 0$, and $\mathcal{Z} = \sum_n w(n)\phi^n \in (0,\infty)$ is the corresponding normalising constant.

Remark 1.2. The equivalence of ensembles is the main mathematical framework to understand the large scale behaviour of statistical mechanics models, and in particular to show condensation as in Definition 1.1 (see Section 4 for details). Assumption (A2) has therefore been established for all homogeneous particle systems that are known to exhibit condensation (see citations above). We will see in our second result Theorem 1.7, that equivalence of ensembles and Assumption (A2) can be shown for a large class of models under slightly stronger assumptions on convergence of the stationary weights.

Remark 1.3. It is natural to ask if ν_ρ, cf. (1.2), can always be described by an exponential change of measure w.r.t. w, as is assumed in (A2). One possible sufficient condition to recover the postulated form in (A2) is to require that ν_ρ and the measure w defined by the limiting weights in (A1) are equivalent in the sense of measures, i.e. ν_ρ and w have the same support, and $w(0), w(1) > 0$. Then we can define

$$\phi := \frac{\nu_\rho[1]}{w(1)}\frac{w(0)}{\nu_\rho[0]} > 0 \quad \text{and} \quad \mathcal{Z} = \frac{w(0)}{\nu_\rho[0]} \in (0, \infty)$$

and use a telescopic product argument for the ratio of partition functions to see that

$$\frac{Z_{L-1,N-n}}{Z_{L,N}} = \frac{Z_{L-1,N}}{Z_{L,N}} \prod_{k=1}^{n} \frac{Z_{L-1,N-k}}{Z_{L-1,N-k+1}} \to \frac{1}{\mathcal{Z}}\phi^n\,.$$

Thus, we recover Assumption (A2) in this case, and a similar argument works in the degenerate case $w(0) > 0$ and $w(k) = 0$ for all $k \geqslant 1$, which applies for the inclusion process with weights (1.6). We believe that this connection between ν_ρ and w holds under more general conditions, but a more detailed discussion is out of the scope of this paper.

In order to formulate our main result with simple notation, we assume without loss of generality that $\phi = \mathcal{Z} = 1$, since we can absorb ϕ and $\mathcal{Z}$ into the weights by

$$\widetilde{w}_L(n) := \frac{w_L(n)\phi^n}{\mathcal{Z}} \quad \text{and} \quad \widetilde{w}(n) := \frac{w(n)\phi^n}{\mathcal{Z}}\,.$$

So w can be assumed to be the probability mass function of ν_ρ, i.e.

$$\nu_\rho[\eta_x = n] = w(n) \quad \text{for all } n \geqslant 0\,. \tag{1.8}$$

We are interested in the macroscopic part of the condensed phase, i.e. the distribution of occupation numbers η_x that scale linearly with the total mass N in the system when taking the thermodynamic limit. The structure of this macroscopic phase will depend on the asymptotic behaviour of the stationary weights. In order to see Poisson-Dirichlet statistics in this phase, the weights $w_L(n)$ must scale like $(nL)^{-1}$, for at least all n which are visible under macroscopic rescaling:

(A3) there exists $\theta > 0$ such that for all $\varepsilon \in (0, 1)$ we have

$$\sup_{\varepsilon N \leqslant n \leqslant N} |nw_L(n)L - \theta| \to 0 \qquad \text{as } N/L \to \rho\,.$$

For the second part of Theorem 1.4 we also impose a second moment condition:

(A4) The limit

$$\alpha^2 := \frac{1+\theta}{\rho} \lim_{N/L \to \rho} \frac{\pi_{L,N}(\eta_x^2)}{N} \in [0, 1] \quad \text{exists}\,. \tag{1.9}$$

In Section 3 we will see that

$$\alpha = \mu(\|p\|_1)\,, \quad \text{with} \quad \|p\|_1 = \sum_{j=1}^{\infty} p_j\,,$$

i.e. α coincides with the expected total mass fraction for each accumulation point μ of the measures $(\mu_{L,N})_{L,N}$, and that the variance of $\|p\|_1$ vanishes. Therefore, Assumption (A4) guarantees that the macroscopic phase is well defined in the thermodynamic limit, excluding fluctuations of mass towards other scales, and plays an important role when identifying the accumulation points μ. This leads to our first main result.

Theorem 1.4. *Let $\rho > 0$ and $(w_L)_L$ be a sequence of weights satisfying (A1) – (A3) for some $\theta \in (0,1]$ and let μ be an accumulation point of the laws of mass partitions $(\mu_{L,N})_{L,N}$ defined in (1.5). Then*

$$\mu = PD_{[0,\alpha_\mu]}(\theta) \quad with \quad \alpha_\mu = \mu(\|p\|_1) \in [0,1] \,,$$

where $PD_{[0,\alpha']}(\theta)$ denotes the Poisson-Dirichlet distribution with parameter θ, concentrating on partitions of the interval $[0,\alpha']$, which depends on the accumulation point. If in addition (A4) holds then

$$(\mu_{L,N})_{L,N} \xrightarrow{d} PD_{[0,\alpha]}(\theta) \,, \quad as \quad N/L \to \rho \,.$$

Remark 1.5. (a) If $\theta = 0$, assumption (A3) does not specify the leading order limiting behaviour of the weights. Our proof can cover this case, if we assume in addition that $n \mapsto nw_L(n)$ is a regularly varying function[1] for large enough L (see (3.10) in the proof). This is consistent with choosing fixed weights of the form $w_L(n) = w(n) = n^{-b}/\mathcal{Z}$ for $b > 2$ whenever $n > 0$ and $w_L(0) = 1$. Indeed, for this choice, as part of a larger class of sub-exponential weights, it is a well known result that the condensed phase consists of a single cluster [13, 26, 2], which can be interpreted as a degenerate Poisson-Dirichlet distribution with $\theta = 0$.

(b) The result also covers the case $\alpha = 0$ where the macroscopic phase is empty and its limiting distribution is trivial. This includes models that do not condense at all, or where the condensed phase concentrates on sub-macroscopic scales such as for certain models with spatial correlations [40, 39, 38]. Of course our result does not say anything interesting in this case, since there is no mass on the macroscopic scale.

(c) All results in this paper extend to arbitrary $\theta > 0$ under the assumption that $PD(\theta)$ is the unique reversible distribution for the split-merge dynamics introduced in Section 2. It seems widely accepted that this is indeed the case, though to the authors' best knowledge no proof exists for $\theta > 1$.

Assumption (A3) is the core premise which guarantees the Poisson-Dirichlet limit of the macroscopic phase, and is consistent with the scaling of weights (1.6) for the inclusion process

$$w_L(n) = \frac{\Gamma(n+d)}{n!\Gamma(d)} \simeq \frac{\theta}{n\,L} \,, \quad \text{if we set} \quad d = \frac{\theta}{L} \,.$$

This scaling implies that a size-biased sample of a macroscopic cluster has the stationary weights $nw_L(n) \simeq \theta/L$ which are independent of n. That means, picking a particle uniformly at random, the size of its cluster is uniformly distributed. This is the trademark of Poisson-Dirichlet statistics, and the distribution can only be normalized in the scaling limit due to the factor $1/L$. Thus, to get a non-trivial macroscopic phase with $\theta > 0$, it is necessary that the weights w_L depend on the system size L. The role of θ and more details on size-biased sampling will be given in Section 2.

Remark 1.6. A system where (A1) – (A3) are satisfied but (A4) is difficult to verify, is for example given by weights of the form

$$w_L(n) = w(n)\mathbb{1}\{n \leqslant N^{1/2}\} + \frac{\theta}{n\,L}\mathbb{1}\{n > N^{1/2}\} \,, \tag{1.10}$$

with w being an arbitrary probability mass function on $\mathbb{N}$ (not necessarily of finite support). In this case, to determine if (A4) holds, careful evaluation of the second moment condition would be required that would depend on the choice of $w(n)$. Theorem 1.4 still allows to characterise the macroscopic part of the condensate as Poisson-Dirichlet, with the catch that it is possibly trivial with $\alpha = 0$.

---

[1]i.e. $nw_L(n)/\big(\lambda nw_L([\lambda n])\big) \to C \in (0,\infty)$ for all $\lambda > 0$ as $n \to \infty$

Recall that in Theorem 1.4 we have fixed the density $\rho > 0$ and it provides a very general result that also includes trivial cases without condensation. But we had to assume the equivalence of ensembles in (A2) and regularity of the macroscopic phase in (A4), which are not easy to check in general (if not established already for particular models). Strengthening the requirements (A1) and (A3) on the stationary weights $(w_L)_L$, we can use Theorem 1.4 to show a stronger but more specialized result, including the equivalence of ensembles and regularity of the macroscopic phase in the conclusion.

Theorem 1.7. *Assume $(w_L)_L$ is a sequence of non-negative weights satisfying the following two conditions:*

(B1) $(w_L)_L$ *converges in the sup-norm,* $\| \cdot \|_\infty$, *to a sequence w such that*

$$\sum_{n=0}^\infty w(n) = 1$$

and either $w(0) = 1$ or

$$w(0) > 0 \quad and \quad \sup_n [w(n-1) \wedge w(n)] > 0 \,. \tag{1.11}$$

(B2) There exists some $\theta \in (0,1]$ such that

$$\lim_{J \to \infty} \lim_{L \to \infty} \sup_{n > J} |n w_L(n) L - \theta| = 0 \,.$$

Then the system exhibits a condensation transition according to Definition 1.1 with critical density

$$\rho_c := \sum_{n=0}^\infty n w(n) \in [0, \infty) \,.$$

Furthermore, we have bulk density $\nu_\rho(\eta_x) = \rho_c$ for all $\rho \geqslant \rho_c$ and

$$(\mu_{L,N})_{L,N} \xrightarrow{d} PD_{[0,\alpha]}(\theta) \,, \quad as \quad N/L \to \rho \geqslant \rho_c \,,$$

with $\alpha = \alpha(\rho) = 1 - \frac{\rho_c}{\rho}$.

Remark 1.8. By assumption (B2), there exists an $A \in \mathbb{N}_0$ such that

$$\lim_{L \to \infty} \sup_{n > A} |n w_L(n) L - \theta| \leqslant C \quad \text{with an arbitrary constant } C > 0 \,.$$

So for each $n \geqslant A$ we have $w_L(n) \to 0$ as $L \to \infty$, and hence by assumption (B1)

$$w(n) = 0 \quad \text{for all } n > A \,, \tag{1.12}$$

and the limiting distribution w can only have finite support. In this sense, Theorem 1.4 is more general because it allows for arbitrary limiting distributions of possibly infinite support, at the cost of loosing control over intermediate scales. Recall (1.10) for an example.

Clearly, the restriction in (B1) that w can be interpreted as a probability mass function is for notational convenience, we could just assume summability. Furthermore, with $\theta > 0$ all models covered by this result have a macroscopic phase with non-trivial structure, excluding systems where the latter concentrates on a single site which have been studied previously (see citations above). Condition (1.11) is necessary to avoid lattice effects and establish the equivalence of ensembles for the bulk part of the distribution (see

Proposition A.1). In the special case that $w(0) = 1$ there is a simpler proof of the equivalence of ensembles result.

We want to stress that single-site-condensation in models is typically due to a strong enough attraction between particles, whereas for systems covered in Theorem 1.7 particle attraction alone is too weak, and condensation only occurs in combination with particle expulsion from the bulk as represented by condition (1.12) on the limiting weights. This strict exclusion condition for occupation numbers larger than A in the limiting weights prevents clustering of particles on sub-macroscopic scales. It should be possible to weaken this, but some form of bulk exclusion is essential for condensation with non-trivial macroscopic phase in models with stationary product measures. In Section 5 we provide an intuitive explanation of this in terms of dynamics of generic particle systems covered by our result.

1.3 Key steps of the proofs

The essential steps of the proof of Theorem 1.4 may be summarised as follows. By compactness we know that $(\mu_{L,N})_{L,N}$ has weak accumulation points. In order to determine the limit points' distributions, we prove that $(\mu_{L,N})_{L,N}$ is approximately reversible w.r.t. a discrete split-merge dynamics. These discrete dynamics converge to the generator of the coagulation-fragmentation process with split-merge dynamics, which we will introduce in Section 2, see (2.1). Lastly, we use size-biased sampling together with a disintegration argument to prove that the corresponding limit points concentrate and therefore have a Poisson-Dirichlet law. This is due to the fact that the Poisson-Dirichlet distribution is the unique distribution which concentrates and is reversible w.r.t. the limiting split-merge dynamics mentioned above [41, 36]. Here, we say that μ concentrates if $\|p\|_1 = \alpha$ μ-a.s. for some $\alpha \in [0, 1]$.

Essentially, Theorem 1.7 is a direct application of Theorem 1.4. Additionally, the stronger assumptions allow us to establish the equivalence of ensembles in Appendix A, which in our case implies the condensation transition. The proof is based on the application of a local central limit theorem (LCLT) which, together with a relative entropy bound, shows convergence of single-site marginals to a distribution independent of the particle density $\rho \geqslant \rho_c$.

The remainder of the paper will be structured as follows: in Section 2 we will give a short review on Poisson-Dirichlet distributions and size-biased sampling. Section 3 will focus on the proof of Theorem 1.4 whereas in Section 4 we state the proof of Theorem 1.7. Lastly, we discuss possible applications of Theorem 1.7 to a family of zero-range and generalized inclusion processes in Section 5. In Appendix A we give a brief introduction to grand-canonical ensembles before proving equivalence of ensembles for size-dependent weights under a sub-exponential growth condition.

2 Background on partitions

2.1 The Poisson-Dirichlet distribution

The Poisson-Dirichlet (PD) distribution is a one-parameter family of probability measures on the space of ordered partitions $\nabla = \nabla_{[0,1]}$ of the unit interval, where for any $\alpha > 0$ we denote

$$\nabla_{[0,\alpha]} := \left\{ p = (p_i)_i \in [0,1]^{\mathbb{N}} : \|p\|_1 = \alpha \;\; \text{and} \;\; p_1 \geqslant p_2 \geqslant \cdots \right\} .$$

Note that elements p in ∇ are not partitions themselves but induce partitions of the form $\{[0, p_1), [p_1, p_1 + p_2), \ldots\}$. The family of measures was first introduced by Kingman [31] in the study of random distributions on countably infinite sets, motivated by Bayesian inference and decision theory. Apart from the original construction as a limit of Dirichlet

distributions, the PD distribution can be more intuitively constructed via a stick-breaking procedure. Let $U_1, U_2, \ldots$ be independent Beta$(1, \theta)$-distributed random variables and define

$$V_1 := U_1, \quad V_2 := (1 - U_1)U_2, \quad V_3 := (1 - U_1)(1 - U_2)U_3, \ldots,$$

i.e. we start with a stick of unit length and continue by breaking a random fraction of U_1 apart. Then we do the same with the remaining part of the stick and iterate. The resulting random vector $V = (V_i)_{i \geqslant 1}$ is said to be GEM(θ)-distributed, named after Griffiths [23, 24] Engen [11] and McCloskey [34]. Reordering the entries of V in decreasing order yields $\widehat{V}$ which is known to be PD(θ)-distributed (see e.g. [16]).

Note the two special cases, $\theta = 1$ where the U_i's are uniformly distributed on the interval $[0, 1]$, and $\theta = 0$ where the U_i's are degenerated point-measures on one and hence $\widehat{V} = V = (1, 0, 0, \ldots)$. Clearly, the choice of the interval $[0, 1]$ is arbitrary and one can construct PD and GEM distributions on intervals $[0, \alpha]$ for arbitrary $\alpha > 0$ just by rescaling

$$p \sim PD_{[0,\alpha]}(\theta) \quad \Leftrightarrow \quad p/\alpha := (p_1/\alpha, p_2/\alpha, \ldots) \sim PD_{[0,1]}(\theta),$$

and analogously for $GEM_{[0,\alpha]}(\theta)$. Since its introduction in [31] the PD distribution emerged first in population biology [31, 12], before appearing in statistical mechanics [32, 25, 27] and interacting particle systems [28]. In particular, the statistics of cycle length distributions in spatial permutations can be linked to condensation phenomena in quantum-mechanical models, see e.g. [21, 3, 4] and references therein. While those models often involve weights that satisfy a decay condition similar to (A3), they consider a different scaling limit with only one diverging parameter, and our results are not directly applicable.

Besides its characterisation via the GEM-construction, the PD distribution was furthermore found to be the unique invariant (in fact reversible) measure on ∇ of the coagulation-fragmentation process with split-merge dynamics for $\theta \in (0, 1]$. This is a Markov process on the state space $\overline{\nabla}$ with infinitesimal generator given by

$$\mathcal{G}_\theta f(p) = \sum_{i \neq j} p_i p_j \left[f(\widehat{M}_{ij}p) - f(p) \right] + \theta \sum_i p_i^2 \left[\int_0^1 f(\widehat{S}_i^u p) du - f(p) \right]. \tag{2.1}$$

Here $\widehat{M}_{ij}p$ denotes the operator that merges the parts p_i and p_j to a single block of size $p_i + p_j$ and then reorders the partition to maintain the decreasing order. On the other hand, $\widehat{S}_i^u p$ defines the operation of splitting p_i into two blocks of size up_i and $(1 - u)p_i$ before reordering the resulting partition.

Proposition 2.1 ([41, 36]). *For $\theta \in [0, 1]$, the Poisson-Dirichlet distribution PD(θ) is the unique invariant measure on ∇ with respect to split-merge dynamics defined by $\mathcal{G}_\theta$, and it is also reversible.*

Since the generator in (2.1) conserves the total mass of partitions, it is clear that there exist stationary distributions for split-merge dynamics on $\nabla_{[0,\alpha]}$ for all $\alpha > 0$, which are unique and equal to $PD_{[0,\alpha]}(\theta)$ with the above result. In general, the set $C_b(\overline{\nabla})$ of bounded continuous functions is the natural domain for the (Feller) Markov semigroup associated to split-merge dynamics. Under the product topology on $\overline{\nabla}$, cf. (1.3), these include in particular bounded cylinder functions, which depend only on finitely many entries of a partition, and for all such functions the generator (2.1) is well defined (see also [33, Lemma 4]).

Originally, the split-merge process was constructed in discrete time, see [33], the extension to continuous time can be found in [21, Section 7.4]. The uniqueness of the invariant measure when $\theta = 1$ was proven in [41] by Zerner, Zeitouni, Mayer-Wolf and Diaconis. An alternative technique allowed Schramm to extend this result to $\theta \in (0, 1]$,

see [36] and Theorem 7.1 in [21]. Lastly, consider the case $\theta = 0$, then clearly $\delta_{(1,0,\dots)}$ is invariant because $\mathcal{G}_0$ only consists of the merge term and there cannot exist another invariant measure on ∇.

2.2 Size-biased sampling

Partitions in ∇ can be interpreted as probability mass functions themselves, which allows for a natural size-biased resampling of its elements. Given $p = (p_i)_{i \in \mathbb{N}} \in \nabla$ we sample an index $i \in \mathbb{N}$ at random according to $(p_i)_{i \in \mathbb{N}}$. Continuing this procedure, while renormalising the remaining partition to a total mass of one in each round, we construct a so-called *size-biased sample* $\widetilde{p}$ of p. For given $p \in \nabla$, $\widetilde{p}$ is a random element of the unordered set

$$\Delta := \left\{ p = (p_i)_i \in [0,1]^{\mathbb{N}} : \|p\|_1 = 1 \right\},$$

and we denote its distribution by σ_p. Of course $\widetilde{p}$ can also be defined in the same way for $p \in \Delta$.

The above procedure can be generalised to partitions in the compact space

$$\overline{\Delta} := \left\{ p = (p_i)_i \in [0,1]^{\mathbb{N}} : \|p\|_1 \leqslant 1 \right\},$$

which includes in particular $\overline{\nabla} = \bigcup_{\alpha \in [0,1]} \nabla_{[0,\alpha]}$ that we already introduced in (1.3). More precisely, fix an element $p \in \overline{\Delta}$. In the following, $q = \widetilde{p}$ will denote the size-biased sample with distribution σ_p on $\overline{\Delta}$ which is defined recursively:

- the first entry of q is assigned the value[2]

$$q_1 \leftarrow \begin{cases} p_j & \text{w.p. } p_j \text{ for all } j \in \mathbb{N}, \\ 0 & \text{w.p. } 1 - \|p\|_1, \end{cases} \tag{2.2}$$

- for $i > 1$, let V be the set of indices of p assigned to q_k for $k < i$, then

$$q_i \leftarrow \begin{cases} p_j & \text{w.p. } \dfrac{p_j}{1 - \sum_{k=1}^{i-1} q_k} \text{ for all } j \in \mathbb{N} \setminus V, \\ \\ 0 & \text{w.p. } \dfrac{1 - \|p\|_1}{1 - \sum_{k=1}^{i-1} q_k}. \end{cases} \tag{2.3}$$

In the case where p consists of finitely many non-zero components only, we sample zeros in each iteration after exhausting all non-zero components. In contrast to size-biased sampling on Δ with $\|p\|_1 = 1$, which is usually defined in terms of shuffling indices of the original sequence, size-biased sampling of $p \in \overline{\Delta}$ assumes a non-exhaustive reservoir of zeros from which we pick with probability proportional to $1 - \|p\|_1$ in each round. For given $p \in \overline{\Delta}$, the distribution $\sigma_p[dq]$ then denotes the law on $\overline{\Delta}$ of q defined above, and concentrates on partitions with $\|q\|_1 = \|p\|_1$.

For an arbitrary probability measure $\mu \in \mathcal{M}_1(\overline{\nabla})$ we then define its size-biased distribution $\sigma(\mu)$ as the law

$$\sigma(\mu)[dq] := \int_{\overline{\nabla}} \sigma_p[dq] \mu[dp]. \tag{2.4}$$

---

[2] We refrain from using an equal sign instead of '$\leftarrow$', since this could lead to mathematically wrong statements. For example, if $p = (\frac{1}{2}, \frac{1}{2}, 0, \dots)$ then $q_1 = \frac{1}{2}$ with probability $\frac{1}{2} + \frac{1}{2} = 1$ which is reflected by our notation in (2.2), but would read $q_1 = \frac{1}{2}$ with probability $\frac{1}{2}$ when replacing '$\leftarrow$' with '$=$'. In a fully rigorous construction of size-biased samples we actually sample the index j at random and not the value p_j, see [18] for the full construction and more details. Because our analysis does not differentiate between entries of the same size, we omit this step to significantly simplify notation and assign the value p_j directly.

One interesting result regarding size-biased distributions with reservoirs of zeros, is that weak convergence of measures on $\overline{\nabla}$ implies weak convergence of the corresponding size-biased distributions.

Lemma 2.2. *If a sequence $(\mu_n)_{n\in\mathbb{N}}$ of probability measures on $\overline{\nabla}$ converges weakly to a measure μ, then also $\sigma(\mu_n) \xrightarrow{d} \sigma(\mu)$ on $\overline{\Delta}$.*

Originally this result was stated in [9, Theorem 1] with a flawed construction and proof. A correct proof of the lemma can be found in [18, Theorem 1] along with a nice exposition on size-biased sampling.

From the stick-breaking construction of the PD distribution it is easy to see (e.g. in [16]) that the size-biased distribution of $\text{PD}(\theta)$ on $\nabla_{[0,1]}$ is precisely the $\text{GEM}(\theta)$ distribution. Considering on the other hand $\text{PD}_{[0,\alpha]}(\theta)$ on $\nabla_{[0,\alpha]}$, its size-biased distribution, as defined in (2.2) and (2.3), contains 0-elements and does not coincide with $\text{GEM}_{[0,\alpha]}(\theta)$ whenever $\alpha < 1$. However, this connection still holds for a modified (positive) size-biasing $\tilde{p}'$ without 0-elements, defined again via scaling. For $p \in \overline{\Delta} \setminus \{0\}$ (i.e. we have $\|p\|_1 > 0$), the positive size-biased sample is defined as

$$\tilde{p}' := \|p\|_1 q \quad \text{where} \quad q \sim \sigma_{p/\|p\|_1} \,. \tag{2.5}$$

Therefore the law of $\tilde{p}'$ on $\overline{\Delta} \setminus \{0\}$ is $\sigma'_p[dq] = \sigma_{p/\|p\|_1} \circ \frac{1}{\|p\|_1}[dq]$, the pushforward measure under rescaling of the size-biasing $\sigma_{p/\|p\|_1}$ on Δ, which does not contain any 0-elements. As for σ_p, $\sigma'_p[dq]$ concentrates on partitions with $\|q\|_1 = \|p\|_1$, and for a distribution μ on $\overline{\Delta}$ we define $\sigma'(\mu)$ analogously to (2.4).

Note that $\sigma(\mu) = \sigma'(\mu)$ on $\mathcal{M}_1(\nabla)$ since $\|p\|_1 = 1$ μ-a.s., and in this case we will write $\mu[\tilde{p} \in \cdot]$ instead to ease notation. Furthermore, we can recover the finite dimensional marginals of $\sigma'(\mu)$ from $\sigma(\mu)$ by ignoring zeros sampled w.r.t. the latter measure [18]. We will only use this fact for the first marginal and state its derivation for completeness: for fixed $p \in \overline{\nabla} \setminus \{0\}$

$$\sigma_p[q_1 \in \cdot \mid q_1 > 0] = \frac{\sigma_p[q_1 \in \cdot, \ q_1 > 0]}{\sigma_p[q_1 > 0]} = \sum_{j=1}^{\infty} \frac{p_j \mathbb{1}\{p_j > 0, \ p_j \in \cdot\}}{\|p\|_1} \tag{2.6}$$

$$= \sum_{j=1}^{\infty} \frac{p_j}{\|p\|_1} \mathbb{1}\{p_j \in \cdot\} = \sigma_{p/\|p\|_1}[q_1 \in \tfrac{\cdot}{\|p\|_1}] = \sigma'_p[q_1 \in \cdot] \,,$$

where we started with the definition of σ_p, cf. (2.2), and rewrote the conditional probability in terms of σ'_p. The statement for general μ follows by averaging.

Returning to the case of the PD distribution $\mu \sim PD_{[0,\alpha]}(\theta)$, we can retrieve the first marginal of the corresponding GEM distribution on $[0,\alpha]$:

$$\sigma'(\mu)[q_1 \in \cdot] = \sigma(\mu)[q_1 \in \cdot \mid q_1 > 0] = \text{Beta}_{[0,\alpha]}(1,\theta) \,.$$

Note also that for positive size-biased distributions the equivalent statement of Lemma 2.2 does not hold, because loss of mass of the corresponding sequences $p \sim \mu_n$ may occur. This would correspond to a positive probability of sampling a block size of zero in the limit, which is precisely the case for distributions $\mu_{L,N}$ we study in this paper.

We will work with size-biased sampling not only on $\overline{\Delta}$ but also on $\Omega_{L,N}$, which corresponds to uniformly picking a particle and sampling the occupation number on its site. Using our definition from above (cf. [28, Section 2.3] for a more detailed construction), the size-biased distribution of $\pi_{L,N}$ is given by

$$\sigma(\pi_{L,N})[d\eta] := \sigma(\mu_{L,N})\left[d\left(\frac{\eta}{N}\right)\right] \,,$$

where $\mu_{L,N}$ (1.5) is the distribution of the ordered partition corresponding to a particle configuration η. Because $\mu_{L,N} \in \mathcal{M}_1(\nabla)$, we will write $\pi_{L,N}[\widetilde{\eta} \in \cdot]$ instead of $\sigma(\pi_{L,N})$, and note that due to the product structure of (1.1) we have e.g. for the first marginal

$$\pi_{L,N}[\widetilde{\eta}_1 = n] = \frac{L}{N} n\, w_L(n) \frac{Z_{L-1,N-n}}{Z_{L,N}} \quad \text{for all } n = 0, \dots, N\,. \tag{2.7}$$

Here, we used the notation $\widetilde{\eta}$ for the size-biased configuration, in the same manner as for partitions on $\overline{\Delta}$.

3 Proof of Theorem 1.4

Recall the canonical distributions $\pi_{L,N}$ on $\Omega_{L,N}$ given by

$$\pi_{L,N}[d\eta] = \frac{1}{Z_{L,N}} \prod_{z=1}^{L} w_L(\eta_z) d\eta$$

and their macroscopic counterparts $\mu_{L,N}$ on $\nabla \subset \overline{\nabla}$, see (1.5). The proof of Theorem 1.4 can be broken down into two main steps. We begin by showing that every accumulation point of $(\mu_{L,N})_{L,N}$ is reversible with respect to the infinitesimal generator $\mathcal{G}_\theta$ of the split-merge process (2.1). The second step consists of proving that each limiting measure on $\overline{\nabla}$ in fact concentrates on $\nabla_{[0,\alpha]}$ for some α. Together with Proposition 2.1 this implies that the limiting measure must be a Poisson-Dirichlet distribution on the interval $[0, \alpha]$.

3.1 Reversibility of weak accumulation points

The space $\overline{\nabla}$ of ordered (sub-)partitions is compact w.r.t. the product topology on $[0,1]^{\mathbb{N}}$, which implies compactness of the space $\mathcal{M}_1(\overline{\nabla})$ w.r.t. the topology induced by weak convergence. Therefore, every subsequence of $(\mu_{L,N})_{L,N}$ has a weakly convergent subsequence and $(\mu_{L,N})_{L,N}$ has at least one accumulation point.

We prepare the proof of Theorem 1.4 which is given at the end of this Section by the following intermediate result.

Proposition 3.1. *Consider weights $(w_L)_L$ satisfying assumptions (A1) – (A3) with $\theta > 0$. All weak accumulation points of $(\mu_{L,N})_{L,N}$ are reversible w.r.t. $\mathcal{G}_\theta$.*

In the remainder of this subsection, we set the stage for the proof of the proposition, which can be found after the statement of Lemma 3.4.

Because $\mu_{L,N}$ concentrates on discrete partitions and by Assumption (A3) we only control the weights on the macroscopic scale, it will be convenient to consider a discrete version $\mathcal{G}_\theta^{(N,\varepsilon)} : C_b(\overline{\nabla}) \to C_b(\overline{\nabla})$ of the split-merge process corresponding to $\mathcal{G}_\theta$ (2.1), which only acts on the parts of the partition exceeding a fixed size of $\varepsilon \in (0,1)$,

$$\mathcal{G}_\theta^{(N,\varepsilon)} f(p) := \frac{N}{N-1} \sum_{i \neq j} p_i p_j \mathbb{1}\{p_i, p_j \geqslant \varepsilon\} \left[f(\widehat{M}_{ij} p) - f(p) \right]$$

$$+ \frac{\theta}{N-1} \sum_i p_i \mathbb{1}\{p_i \geqslant 2\varepsilon\} \sum_{k=\varepsilon N}^{N(p_i - \varepsilon)} \left[f(\widehat{S}_i^{\frac{k}{Np_i}} p) - f(p) \right]\,.$$

As we will see Lemma 3.3, it suffices to control the weights on macroscopic scales larger than ε in order to observe a PD limit.

Recall that under $\mu_{L,N}$ partitions consist of (at most) L blocks. Since we will lift the split-merge dynamics to the space of particle configurations $\Omega_{L,N}$, the resulting partitions should not exceed length L either. However, a priori it is not clear that there is always an empty site to split on and the split-merge dynamics have to be slightly

adapted to achieve this: recall from (1.8) that without loss of generality we can assume that $\nu_\rho[\eta_x = n] = w(n)$. Thus, Assumption (A2) guarantees that there exists $m \in \mathbb{N}$ such that $w(m) = \lim_{L \to \infty} w_L(m) > 0$. This implies that under $\mu_{L,N}$ a positive fraction of blocks has size m/N, see Lemma 3.2 below. Hence, instead of leaving empty sites behind when merging, and splitting onto empty sites only, we can impose to leave a fraction of m/N behind when merging and only split onto blocks of size m/N, cf. Figure 1. This will not affect the statistics on a macroscopic scale, where microscopic blocks are indistinguishable; which is why henceforward we assume $m = 0$ for notational convenience.
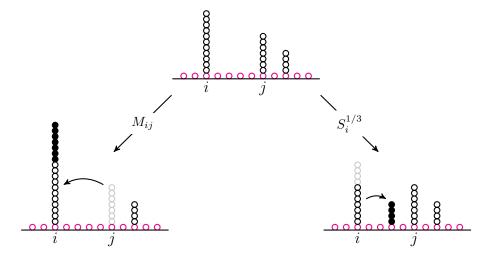

Figure 1: For example, consider the case $m = 1$ where $\lim_{L \to \infty} w_L(0) = w(0) = 0$ but $w(1) > 0$. Configuration $\eta \sim \pi_{L,N}$ will not include any empty sites with high probability. Thus, when splitting we instead impose to split onto sites with single-occupation. Similarly, we leave a single particle behind when merging occupations of two sites. Here, S_i and M_{ij} denote the corresponding split- and merge-operations on configurations; we ignored the subsequent order-operation $\widehat{\cdot}$ for the sake of a clearer picture.

Lemma 3.2. *Let $\#_0(\eta)$ be the number of sites with zero occupation in the configuration η. We have that*

$$\pi_{L,N}\left(\left(\frac{\#_0(\eta)}{L} - w(0) \right)^2 \right) \to 0, \quad \text{as } N/L \to \rho.$$

In particular, we have a weak law of large numbers and for every $\varepsilon > 0$

$$\pi_{L,N}\left[\left| \frac{\#_0(\eta)}{L} - w(0) \right| > \varepsilon \right] \to 0, \quad \text{as } N/L \to \rho.$$

Proof. By direct calculation of the second moment, it suffices to show that (using the product structure of $\pi_{L,N}$)

$$\pi_{L,N}[\eta_x = 0, \eta_y = 0] = \pi_{L-1,N}[\eta_x = 0] \, \pi_{L,N}[\eta_y = 0]$$

converges to $w(0)^2$ for all $1 \leqslant x, y \leqslant L$, $x \neq y$, which holds due to (A2). The second statement follows immediately by Chebyshev's inequality. □

The operator $\mathcal{G}_\theta^{(N,\varepsilon)}$ approximates the corresponding continuous process acting on

blocks of size larger than ε, which is characterised by the generator

$$\mathcal{G}_\theta^{(\varepsilon)} f(p) := \sum_{\substack{i,j=1 \\ i \neq j}}^\infty p_i p_j \mathbb{1}\{p_i, p_j \geqslant \varepsilon\} \left[f(\widehat{M}_{ij}p) - f(p) \right]$$

$$+ \theta \sum_{i=1}^\infty p_i^2 \mathbb{1}\{p_i \geqslant 2\varepsilon\} \left[\int_\varepsilon^{1-\varepsilon} f(\widehat{S}_i^u p) du - f(p) \right].$$

We summarise the corresponding convergence behaviour of the generators in the following lemma:

Lemma 3.3. *We have*

$$\mathcal{G}_\theta^{(N,\varepsilon)} \to \mathcal{G}_\theta^{(\varepsilon)}, \quad \text{as } N \to \infty, \qquad \text{and} \qquad \mathcal{G}_\theta^{(\varepsilon)} \to \mathcal{G}_\theta, \quad \text{as } \varepsilon \to 0,$$

in the strong operator topology on bounded continuous functions $C_b(\overline{\nabla})$.

Proof. Let $g \in C_b(\overline{\nabla})$. To prove the first part of the statement, we need to show that $\|(\mathcal{G}_\theta^{(N,\varepsilon)} - \mathcal{G}_\theta^{(\varepsilon)})g\|_\infty$ vanishes in the limit $N \to \infty$. We will compare the two parts of each operator separately. We start with the merge term:

$$\left| \sum_{\substack{i \neq j}}^L p_i p_j \mathbb{1}\{p_i, p_j \geqslant \varepsilon\} \left[g(\widehat{M}_{ij}p) - g(p) \right] \right.$$

$$\left. - \frac{N}{N-1} \sum_{\substack{i \neq j}}^L p_i p_j \mathbb{1}\{p_i, p_j \geqslant \varepsilon\} \left[g(\widehat{M}_{ij}p) - g(p) \right] \right|$$

$$\leqslant \frac{2\|g\|_\infty}{N-1} \sum_{\substack{i \neq j}}^L p_i p_j \mathbb{1}\{p_i, p_j \geqslant \varepsilon\} \leqslant \frac{2\|g\|_\infty}{N-1}.$$

And similarly for the split term:

$$\left| \sum_{i=1}^L p_i^2 \mathbb{1}\{p_i \geqslant 2\varepsilon\} \left[\int_\varepsilon^{1-\varepsilon} g(\widehat{S}_i^u p) du - g(p) \right] \right.$$

$$\left. - \frac{1}{N-1} \sum_{i=1}^L p_i \mathbb{1}\{p_i \geqslant 2\varepsilon\} \sum_{k=\varepsilon N}^{N(p_i-\varepsilon)} \left[g(\widehat{S}_i^{\frac{k}{Np_i}} p) - g(p) \right] \right|$$

$$\leqslant \sum_{i=1}^L p_i^2 \left| \int_\varepsilon^{1-\varepsilon} g(\widehat{S}_i^u p) du - \frac{1}{(N-1)p_i} \sum_{k=\varepsilon N}^{N(p_i-\varepsilon)} g(\widehat{S}_i^{\frac{k}{Np_i}} p) \right| + \frac{\|g\|_\infty}{N-1} \sum_{i=1}^L p_i(1-p_i)$$

$$\leqslant \max_i \omega(g(\widehat{S}_i^\bullet p), \tfrac{1}{Np_i}) + \frac{\|g\|_\infty}{N-1},$$

where ω is the modulus of continuity of the function $u \mapsto g(\widehat{S}_i^u p)$. The last inequality holds since the sum inside the absolute value approximates the Riemann-sum which in turn converges to the corresponding integral, where the error is controlled by the modulus of continuity. By [27, Lemma 1], we know that

$$\|\widehat{S}_i^u p - \widehat{S}_i^v p\|_1 \leqslant 2|u-v|p_i,$$

which in particular implies

$$\omega(\widehat{S}_i^\bullet p, \tfrac{1}{Np_i}) = \sup_{\substack{u,v \in [0,1] \\ |u-v| \leqslant (Np_i)^{-1}}} \|\widehat{S}_i^u p - \widehat{S}_i^v p\|_2 \leqslant \sup_{\substack{u,v \in [0,1] \\ |u-v| \leqslant (Np_i)^{-1}}} \|\widehat{S}_i^u p - \widehat{S}_i^v p\|_1^{1/2} \leqslant \frac{\sqrt{2}}{\sqrt{N}}.$$

This allows us to uniformly bound the moduli of continuity above, in the sense that

$$
\begin{aligned}
\omega(g(\widehat{S}_i^\bullet p), \tfrac{1}{Np_i}) &= \sup_{|u-v| \leqslant \frac{1}{Np_i}} \left| g(\widehat{S}_i^u p) - g(\widehat{S}_i^v p) \right| \\
&\leqslant \sup_{\|p'-q'\|_2 \leqslant \omega(\widehat{S}_i^\bullet p, \frac{1}{Np_i})} \left| g(p') - g(q') \right| = \omega\big(g, \omega(\widehat{S}_i^\bullet p, \tfrac{1}{Np_i})\big) \\
&\leqslant \omega\big(g, \tfrac{\sqrt{2}}{\sqrt{N}}\big) \,,
\end{aligned}
\tag{3.1}
$$

where $\omega(g, \cdot)$ denotes the modulus of continuity w.r.t. g on $(\overline{\nabla}, \|\cdot\|_2)$. The r.h.s. of (3.1) vanishes due to uniform continuity of g. Here we used that the topology on $\overline{\nabla}$ induced by $\|\cdot\|_2$ coincides with the product topology.

Since the sum of the first estimate and second estimate multiplied by θ bounds $\|(\mathcal{G}_\theta^{(N,\varepsilon)} - \mathcal{G}_\theta^{(\varepsilon)})g\|_\infty$ from above, we take the thermodynamic limit $N/L \to \rho$ and conclude the first part of the lemma.

The second statement requires us to show that $\|(\mathcal{G}_\theta - \mathcal{G}_\theta^{(\varepsilon)})g\|_\infty \to 0$ as $\epsilon \to 0$. First, note that

$$
\begin{aligned}
(\mathcal{G}_\theta - \mathcal{G}_\theta^{(\varepsilon)})g(p) &= \sum_{\substack{i,j=1 \\ i \neq j}}^\infty p_i p_j \mathbb{1}\{p_i < \varepsilon\} \cup \{p_j < \varepsilon\} \left[g(\widehat{M}_{ij} p) - g(p) \right] \\
&\quad + \theta \sum_{i=1}^\infty p_i^2 \left[\int_0^1 g(\widehat{S}_i^u p)du - g(p) \right] \\
&\quad - \theta \sum_{i=1}^\infty p_i^2 \mathbb{1}\{p_i \geqslant 2\varepsilon\} \left[\int_\varepsilon^{1-\varepsilon} g(\widehat{S}_i^u p)du - g(p) \right] \\
&= \sum_{\substack{i,j=1 \\ i \neq j}}^\infty p_i p_j \mathbb{1}\{p_i < \varepsilon\} \cup \{p_j < \varepsilon\} \left[g(\widehat{M}_{ij} p) - g(p) \right] \\
&\quad + \theta \sum_{i=1}^\infty p_i^2 \mathbb{1}\{p_i < 2\varepsilon\} \left[\int_0^1 g(\widehat{S}_i^u p)du - g(p) \right] \\
&\quad + \theta \sum_{i=1}^\infty p_i^2 \mathbb{1}\{p_i \geqslant 2\varepsilon\} \left[\int_0^1 g(\widehat{S}_i^u p)du - \int_\varepsilon^{1-\varepsilon} g(\widehat{S}_i^u p)du \right] \,.
\end{aligned}
$$

The first two terms vanish by dominated convergence and for the last term we have the estimate

$$
\left| \sum_{i=1}^\infty p_i^2 \mathbb{1}\{p_i \geqslant 2\varepsilon\} \left[\int_0^1 g(\widehat{S}_i^u p)du - \int_\varepsilon^{1-\varepsilon} g(\widehat{S}_i^u p)du \right] \right| \leqslant 2\varepsilon \|g\|_\infty \sum_{i=1}^\infty p_i^2 \leqslant 2\varepsilon \|g\|_\infty \,.
$$

This concludes the proof. $\qquad\square$

The proof of Proposition 3.1, requires the following key observation which states that $\mu_{L,N}$ is approximately reversible w.r.t. the dynamics corresponding to $\mathcal{G}_\theta^{(N,\varepsilon)}$.

Lemma 3.4. *For every $f, g \in C_b(\overline{\nabla})$ and $\varepsilon \in (0,1)$ we have*

$$
\left| \mu_{L,N}(f \, \mathcal{G}_\theta^{(N,\varepsilon)} g) - \mu_{L,N}(g \, \mathcal{G}_\theta^{(N,\varepsilon)} f) \right| \to 0 \,,
\tag{3.2}
$$

in the thermodynamic limit $N/L \to \rho \geqslant 0$.

We postpone the proof of Lemma 3.4 until after the one of Proposition 3.1. Now we have everything to state the proof of this section's main finding.

Proof of Proposition 3.1: Due to compactness of the space $\mathcal{M}_1(\overline{\nabla})$ w.r.t. the topology induced by weak convergence, the sequence $(\mu_{L,N})$ has weak accumulation points. Let μ be such an accumulation point and $(\mu_{L_j,N_j})_j$ a subsequence converging to it. Then for each $\epsilon \in (0,1)$ we have

$$
\begin{aligned}
|\mu(f\mathcal{G}_\theta g) - \mu(g\mathcal{G}_\theta f)| \;\leqslant\; & \left| \mu(f\mathcal{G}_\theta g) - \mu_{L_j,N_j}(f\mathcal{G}_\theta^{(N_j,\varepsilon)}g) \right| \\
& + \left| \mu_{L_j,N_j}(f\mathcal{G}_\theta^{(N_j,\varepsilon)}g) - \mu_{L_j,N_j}(g\mathcal{G}_\theta^{(N_j,\varepsilon)}f) \right| \\
& + \left| \mu_{L_j,N_j}(g\mathcal{G}_\theta^{(N_j,\varepsilon)}f) - \mu(g\mathcal{G}_\theta f) \right| .
\end{aligned}
$$

The middle term on the r.h.s. vanishes due to Lemma 3.4, whereas the first term can be estimated by

$$
\begin{aligned}
\left| \mu(f\mathcal{G}_\theta g) - \mu_{L_j,N_j}(f\mathcal{G}_\theta^{(N_j,\varepsilon)}g) \right| \;\leqslant\; & \left| \mu(f\mathcal{G}_\theta g) - \mu_{L_j,N_j}(f\mathcal{G}_\theta g) \right| \\
& + \left| \mu_{L_j,N_j}(f\mathcal{G}_\theta g) - \mu_{L_j,N_j}(f\mathcal{G}_\theta^{(N_j,\varepsilon)}g) \right| .
\end{aligned}
\tag{3.3}
$$

Since $(\mu_{L_j,N_j})_j$ converges in distribution to μ and $\mathcal{G}_\theta g$ is bounded and continuous, see [33, Lemma 4], the first term on the r.h.s. of (3.3) vanishes as L diverges. Also, by Lemma 3.3, the second term on the right vanishes after taking the limit $L, N \to \infty$ before $\varepsilon \to 0$, since

$$
\left| \mu_{L_j,N_j}(f(\mathcal{G}_\theta - \mathcal{G}_\theta^{(N_j,\varepsilon)})g) \right| \;\leqslant\; \|f\|_\infty \|(\mathcal{G}_\theta - \mathcal{G}_\theta^{(N_j,\varepsilon)})g\|_\infty .
$$

The same steps hold when applied to $\left| \mu_{L_j,N_j}(g\mathcal{G}_\theta^{(N_j,\varepsilon)}f) - \mu(g\mathcal{G}_\theta f) \right|$. Overall, this yields $\mu(f\mathcal{G}_\theta g) = \mu(g\mathcal{G}_\theta f)$ which finishes the proof. $\qquad\square$

It only remains to state the proof of Lemma 3.4.

Proof of Lemma 3.4: First, we note that we can write $\mu_{L,N}(f\mathcal{G}^{(N,\varepsilon)}g)$ in terms of the canonical distribution $\pi_{L,N}$:

$$
N(N-1)\mu_{L,N}(f\mathcal{G}^{(N,\varepsilon)}g) \tag{3.4}
$$

$$
= \sum_{\eta \in \Omega_{L,N}} f\left(\frac{\widehat{\eta}}{N}\right) \sum_{i \neq j}^{L} \widehat{\eta}_i \widehat{\eta}_j \mathbb{1}\{\widehat{\eta}_i, \widehat{\eta}_j \;\geqslant\; \varepsilon N\} \left[g\left(\widehat{M}_{ij}\left(\frac{\widehat{\eta}}{N}\right)\right) - g\left(\frac{\widehat{\eta}}{N}\right) \right] \pi_{L,N}(\eta)
$$

$$
+ \theta \sum_{\eta \in \Omega_{L,N}} f\left(\frac{\widehat{\eta}}{N}\right) \sum_{i=1}^{L} \widehat{\eta}_i \mathbb{1}\{\widehat{\eta}_i \;\geqslant\; 2\varepsilon N\} \sum_{k=\varepsilon N}^{\widehat{\eta}_i - \varepsilon N} \left[g\left(\widehat{S}_i^{k/\widehat{\eta}_i}\left(\frac{\widehat{\eta}}{N}\right)\right) - g\left(\frac{\widehat{\eta}}{N}\right) \right] \pi_{L,N}(\eta) ,
$$

where we multiplied with $N(N-1)$ for convenience.

In the following it will be easier not to work with discrete partitions, but with corresponding particle configurations without worrying about the order of the corresponding sites. Hence, we require a new notation to lift split and merge operations to the space of particle configurations: for $\eta \in \Omega_{L,N}$ we define

$$
\begin{aligned}
M_{xy}\eta &:= \eta + \eta_y(e^{(x)} - e^{(y)}) , \\
S_{xy}^k\eta &:= \eta + k(e^{(y)} - e^{(x)}) , \\
S_x^k\eta &:= \eta + k(e^{(L+1)} - e^{(x)}) ,
\end{aligned}
$$

where $e^{(x)} \in \Omega_{L,1}$ denotes the configuration with a single particle at site x, i.e. $(e^{(x)})_z = \delta_{x,z}$. The operator S_x^k is only necessary for the case of a full particle configuration η, i.e.

$\#_0(\eta) = 0$. Then, we simply append the additional block of particles at the end of the configuration, which then has length $(L + 1)$. This arbitrary but convenient choice of course does not change the projection of the dynamics on the level of partitions. As such we can rewrite (3.4) as

$$
N(N-1)\mu_{L,N}(f\mathcal{G}^{(N,\varepsilon)}g) = \pi_{L,N}\left(\widehat{f}(\eta)\sum_{x\neq y}^{L}\eta_x\eta_y\mathbb{1}\{\eta_x,\eta_y \geqslant \varepsilon N\}\left[\widehat{g}(M_{xy}\eta) - \widehat{g}(\eta)\right]\right)
$$

$$
+ \theta\,\pi_{L,N}\left(\mathbb{1}\{\#_0(\eta) > 0\}\widehat{f}(\eta)\sum_{x=1}^{L}\eta_x\mathbb{1}\{\eta_x \geqslant 2\varepsilon N\}\right.
$$

$$
\times \sum_{k=\varepsilon N}^{\eta_x-\varepsilon N}\sum_{y=1}^{L}\frac{\mathbb{1}\{\eta_y = 0\}}{\#_0(\eta)}\left[\widehat{g}\left(S_{xy}^k\eta\right) - \widehat{g}(\eta)\right]\right) \tag{3.5}
$$

$$
+ \theta\,\pi_{L,N}\left(\mathbb{1}\{\#_0(\eta) = 0\}\widehat{f}(\eta)\sum_{x=1}^{L}\eta_x\mathbb{1}\{\eta_x \geqslant 2\varepsilon N\}\right.
$$

$$
\times \sum_{k=\varepsilon N}^{\eta_x-\varepsilon N}\left[\widehat{g}\left(S_{x}^k\eta\right) - \widehat{g}(\eta)\right]\right),
$$

where $\widehat{f} := f \circ T$ and $\widehat{g} := g \circ T$ with the ordering map T given in (1.4). Furthermore, we can see that the terms that depend on $\widehat{f}$ and $\widehat{g}$ only through the product $\widehat{f}(\eta)\widehat{g}(\eta)$ cancel when taking the difference between $\mu_{L,N}(f\mathcal{G}^{(N,\varepsilon)}g)$ and $\mu_{L,N}(g\mathcal{G}^{(N,\varepsilon)}f)$ in (3.2). Additionally, the very last term in (3.5) is negligible since, using $(\eta_x - 1)/(N-1) \leqslant 1$ and $\sum_{x=1}^{L}\eta_x/N = 1$ $\pi_{L,N}$-almost surely, we have

$$
\frac{1}{N(N-1)}\pi_{L,N}\left(\mathbb{1}\{\#_0(\eta) = 0\}\sum_{x=1}^{L}\eta_x(\eta_x - 1)\right) \leqslant \pi_{L,N}\left[\#_0(\eta) = 0\right],
$$

which vanishes due to Lemma 3.2 because $w(0) > 0$.

To simplify notation we introduce

$$
V_{f,g}^{(L,N,\varepsilon)} := \pi_{L,N}\left(\widehat{f}(\eta)\sum_{x\neq y}^{L}\eta_x\eta_y\mathbb{1}\{\eta_x,\eta_y \geqslant \varepsilon N\}\widehat{g}(M_{xy}\eta)\right) \tag{3.6}
$$

and

$$
U_{f,g}^{(L,N,\varepsilon)} := \theta\,\pi_{L,N}\left(\mathbb{1}\{\#_0(\eta) > 0\}\widehat{f}(\eta)\sum_{x=1}^{L}\eta_x\mathbb{1}\{\eta_x \geqslant 2\varepsilon N\}\right.
$$

$$
\times \sum_{k=\varepsilon N}^{\eta_x-\varepsilon N}\sum_{y=1}^{L}\frac{\mathbb{1}\{\eta_y = 0\}}{\#_0(\eta)}\widehat{g}\left(S_{xy}^k\eta\right)\right). \tag{3.7}
$$

We are then left to analyse

$$
N(N-1)(\mu_{L,N}(f\mathcal{G}^{(N,\varepsilon)}g) - \mu_{L,N}(g\mathcal{G}^{(N,\varepsilon)}f))
$$

$$
= V_{f,g}^{(L,N,\varepsilon)} + U_{f,g}^{(L,N,\varepsilon)} - V_{g,f}^{(L,N,\varepsilon)} - U_{g,f}^{(L,N,\varepsilon)} + o(N^2).
$$

The goal is to compare $V_{f,g}^{(L,N,\varepsilon)} - U_{g,f}^{(L,N,\varepsilon)}$ and $U_{f,g}^{(L,N,\varepsilon)} - V_{g,f}^{(L,N,\varepsilon)}$, respectively, and show that these differences vanish in the limit if divided by $N(N-1)$. In order to prove this, we perform a change of measure, since both $V^{(L,N,\varepsilon)}$ and $U^{(L,N,\varepsilon)}$ are expectations with respect to $\pi_{L,N}$.

First, we note that the restriction of the merge map

$$M_{xy}\Big|_{\{\eta_y=k\}} : \Omega_{L,N} \cap \{\eta_y = k\} \to \Omega_{L,N}$$

is injective and therefore defines a bijection between the set $\Omega_{L,N} \cap \{\eta_y = k\}$ and its image $M_{xy}(\Omega_{L,N} \cap \{\eta_y = k\}) = \Omega_{L,N} \cap \{\eta_y = 0, k \leqslant \eta_x\}$ with inverse S_{xy}^k. Therefore, the change of measure of $\pi_{L,N}$ and its pushforward measure under M_{xy} for fixed x, y, k on the set $\Omega_{L,N} \cap \{\eta_y = k\}$ is given by:

$$\frac{\pi_{L,N}}{\pi_{L,N} \circ M_{xy}}[\eta] = \frac{\prod_{z=1}^L w_L(\eta_z)}{\prod_{z=1}^L w_L((M_{xy}\eta)_z)} = \frac{w_L((M_{xy}\eta)_x - k)w_L(k)}{w_L((M_{xy}\eta)_x)w_L(0)}. \tag{3.8}$$

This will allow us to perform a change of measure in the following sense. Fix x, y and k, furthermore let $h_{y,k}$ be a real valued function on $\Omega_{L,N}$ with support in $\{\eta_y = k\}$. By definition of M_{xy}, we can define $\widetilde{h}_{x,y,k} = h_{y,k} \circ M_{xy}^{-1}$ which is zero outside of $\Omega_{L,N} \cap \{\eta_y = 0, k \leqslant \eta_x\}$. We then have

$$\pi_{L,N}(h_{y,k}(\eta)) = \pi_{L,N}(\widetilde{h}_{x,y,k}(M_{xy}\eta)) = \int_{\{\eta_y=k\}} \widetilde{h}_{x,y,k}(M_{xy}\eta)\, \pi_{L,N}[d\eta]$$

$$= \int_{\{\xi_y=0, k \leqslant \xi_x\}} \widetilde{h}_{x,y,k}(\xi)\, \pi_{L,N} \circ M_{xy}^{-1}[d\xi]$$

$$= \int_{\{\xi_y=0, k \leqslant \xi_x\}} \widetilde{h}_{x,y,k}(\xi) \frac{w_L(\xi_x - k)w_L(k)}{w_L(\xi_x)w_L(0)}\, \pi_{L,N}[d\xi].$$

Before we apply the change of measure, we divide and multiply by $(M_{xy}\eta)_x$, decompose over $\{\eta_y = k\}$, and interchange the order of integration:

$$V_{f,g}^{(L,N,\varepsilon)} = \sum_{x \neq y}^L \sum_{k=1}^N \pi_{L,N} \left((M_{xy}\eta)_x \widehat{f}(\eta)\, \widehat{g}(M_{xy}\eta) \frac{\eta_x \eta_y}{(M_{xy}\eta)_x} \mathbb{1}\{k, \eta_x \geqslant \varepsilon N\} \right.$$

$$\left. \times \mathbb{1}\{\eta_y = k\} \mathbb{1}\{(M_{xy}\eta)_y = 0\} \right)$$

$$= \sum_{x \neq y}^L \sum_{k=1}^N \pi_{L,N} \left((M_{xy}\eta)_x \widehat{f}\left(S_{xy}^k M_{xy}\eta\right) \widehat{g}(M_{xy}\eta) \frac{((M_{xy}\eta)_x - k)k}{(M_{xy}\eta)_x} \mathbb{1}\{k \geqslant \varepsilon N\} \right.$$

$$\left. \times \mathbb{1}\{k \leqslant (M_{xy}\eta)_x - \varepsilon N\} \mathbb{1}\{(S_{xy}^k M_{xy}\eta)_y = k\} \mathbb{1}\{(M_{xy}\eta)_y = 0\} \right),$$

where in the last step we expressed η in terms of $M_{xy}\eta$. Recall now (3.8) and note that we can recover η from $M_{xy}\eta$ and $k = \eta_y$. Therefore, the change of measure yields

$$V_{f,g}^{(L,N,\varepsilon)} = \sum_{x \neq y}^L \sum_{k=1}^N \pi_{L,N} \left(\mathbb{1}\{\#_0(\eta) > 0\} \eta_x \widehat{f}\left(S_{xy}^k \eta\right) \widehat{g}(\eta) \frac{(\eta_x - k)k}{\eta_x} \frac{w_L(\eta_x - k)w_L(k)}{w_L(\eta_x)w_L(0)}. \right.$$

$$\left. \times \mathbb{1}\{k \leqslant \eta_x - \varepsilon N\} \mathbb{1}\{k \geqslant \varepsilon N\} \mathbb{1}\{(S_{xy}^k \eta)_y = k\} \mathbb{1}\{\eta_y = 0\} \right)$$

$$= \pi_{L,N} \left(\mathbb{1}\{\#_0(\eta) > 0\} \widehat{g}(\eta) \sum_{x=1}^L \eta_x \mathbb{1}\{\eta_x \geqslant 2\varepsilon N\} \sum_{k=\varepsilon N}^{\eta_x - \varepsilon N} \sum_{y=1}^L \frac{\mathbb{1}\{\eta_y = 0\}}{\#_0(\eta)} \right.$$

$$\left. \times \widehat{f}\left(S_{xy}^k \eta\right) \#_0(\eta) \frac{(\eta_x - k)k}{\eta_x} \frac{w_L(\eta_x - k)w_L(k)}{w_L(\eta_x)w_L(0)} \right).$$

Comparing now $V_{f,g}^{(L,N,\varepsilon)}$ to $U_{g,f}^{(L,N,\varepsilon)}$ in (3.7), the only discrepancy is the term

$$\#_0(\eta)\frac{(\eta_x - k)k}{\eta_x}\frac{w_L(\eta_x - k)w_L(k)}{w_L(\eta_x)w_L(0)} = \frac{\#_0(\eta)}{L}\frac{k\,w_L(k)L}{w_L(0)}\frac{(\eta_x - k)\,w_L(\eta_x - k)}{\eta_x\,w_L(\eta_x)}\,. \tag{3.9}$$

First, note that by Assumption (A3)

$$k\,w_L(k)L\frac{(\eta_x - k)w_L(\eta_x - k)}{\eta_x w_L(\eta_x)} \to \theta\,, \quad \text{as } L \to \infty\,, \tag{3.10}$$

uniformly in k and η, since $\varepsilon N \leqslant k, \eta_x - k, \eta_x$. Therefore, we have

$$\left|\mu_{L,N}(f\mathcal{G}^{(N,\varepsilon)}g) - \mu_{L,N}(g\mathcal{G}^{(N,\varepsilon)}f)\right|$$

$$= \frac{1}{N(N-1)}\left|\left(V_{f,g}^{(L,N,\varepsilon)} - U_{g,f}^{(L,N,\varepsilon)}\right) + \left(U_{f,g}^{(L,N,\varepsilon)} - V_{g,f}^{(L,N,\varepsilon)}\right)\right| + o(1)$$

$$\leqslant \frac{2\|f\|_\infty\|g\|_\infty\theta}{N(N-1)}\pi_{L,N}\left(\sum_{x=1}^{L}\eta_x\sum_{k=\varepsilon N}^{\eta_x - \varepsilon N}\left|\frac{\#_0(\eta)}{L}\frac{1}{w_L(0)} - 1\right|\right) + o(1)$$

$$\leqslant 2\|f\|_\infty\|g\|_\infty\theta\,\pi_{L,N}\left(\left|\frac{\#_0(\eta)}{L}\frac{1}{w_L(0)} - 1\right|\right) + o(1)\,,$$

where we used (3.10) before dropping all indicator functions in the first inequality. Lastly, by Lemma 3.2 and Assumption (A1) we know that

$$\lim_{N/L\to\rho}\pi_{L,N}\left(\left|\frac{\#_0(\eta)}{L}\frac{1}{w_L(0)} - 1\right|\right) = 0\,,$$

which completes the proof. $\qquad\square$

3.2 Concentration and uniqueness of the limit

As already mentioned in the introduction, the only invariant distribution w.r.t. $\mathcal{G}_\theta$ which concentrates on full partitions $\nabla_{[0,\alpha]}$, for some $\alpha \in [0,1]$, is the Poisson-Dirichlet distribution $\mathrm{PD}_{[0,\alpha]}(\theta)$. Hence, in order to prove that weak accumulation points of measures $(\mu_{L,N})_{L,N}$ coincide, it is enough to show that each such weak limit μ from Proposition 3.1 satisfies

$$\mathrm{Var}_\mu\left(\sum_{i=1}^{\infty}p_i\right) = 0$$

and there exists $\alpha_\mu \in [0,1]$ such that $\mu\left(\sum_{i=1}^{\infty}p_i\right) = \alpha_\mu$ for every limit μ.

In the following we consider not only the ℓ^1-norm but also the ℓ^2-norm as a function on $\overline{\nabla}$ and write

$$\|p\|_1 = \sum_{i=1}^{\infty}p_i \quad \text{and} \quad \|p\|_2^2 = \sum_{i=1}^{\infty}|p_i|^2\,,$$

respectively. Note that the latter has the advantage, in contrast to $\|\cdot\|_1$, of being a continuous function on $\overline{\nabla}$ with respect to the product topology.

First, we recall the first part of Lemma 5 in [33] where Mayer-Wolf et al. proved the following relationship between $\mu(\|p\|_1^2)$ and $\mu(\|p\|_2^2)$, which is a natural consequence of the balance between expected split and merge rates for the stationary distribution μ.

Lemma 3.5 (Mayer-Wolf et al.)**.** *Let $\theta \in (0,1]$ and let $\mu \in \mathcal{M}_1(\overline{\nabla})$ be invariant w.r.t. $\mathcal{G}_\theta$, then*

$$\mu(\|p\|_1^2) = (1 + \theta)\mu(\|p\|_2^2)\,.$$

Together with its counterpart in the following lemma, this immediately proves that each subsequential limit of $(\mu_{L,N})_{L,N}$ concentrates on some $\nabla_{[0,\alpha_\mu]}$.

Lemma 3.6. *Let $\theta \in (0,1]$ and let $\mu \in \mathcal{M}_1(\overline{\nabla})$ be an accumulation point of $(\mu_{L,N})_{L,N}$. Then*

$$\mu(\|p\|_1)^2 = (1+\theta)\mu(\|p\|_2^2).$$

Proof. Since the statement is trivial for $\mu = \delta_{(0,0,\dots)}$, we assume without loss of generality $\mu \neq \delta_{(0,0,\dots)}$. Let $(\mu_{L_j,N_j})_{j \in \mathbb{N}}$ be a sequence converging weakly to μ. For such μ we first note that

$$\mu(\|p\|_2^2) = \lim_{j \to \infty} \mu_{L_j,N_j}(\|p\|_2^2) = \lim_{j \to \infty} \frac{L_j}{N_j^2} \pi_{L_j,N_j}(\eta_1^2), \tag{3.11}$$

since μ_{L_j,N_j} converges weakly to μ. Rewriting the r.h.s. as size-biased expectation, cf. (2.7), yields in particular

$$\mu(\|p\|_2^2) = \lim_{j \to \infty} \frac{1}{N_j} \pi_{L_j,N_j}(\widetilde{\eta}_1) = \lim_{j \to \infty} \mu_{L_j,N_j}(\widetilde{p}_1) = \lim_{j \to \infty} \sigma(\mu_{L_j,N_j})(q_1). \tag{3.12}$$

In the second to last step we used that the size-biased distribution is invariant under reordering of the configuration. Furthermore, weak convergence of μ_{L_j,N_j} implies weak convergence of the size biased distributions by Lemma 2.2, i.e.

$$\sigma(\mu_{L_j,N_j}) \to \sigma(\mu)$$

in the topology induced by weak convergence. Recall that under $\sigma(\mu)$, the first component in q is zero with probability

$$\sigma(\mu)[q_1 = 0] = \mu(1 - \|p\|_1). \tag{3.13}$$

Since the projection map on the first component is continuous w.r.t. the product topology, we have $\sigma(\mu_{L,N})(q_1) \to \sigma(\mu)(q_1)$ which yields with (3.12) and (2.6)

$$\mu(\|p\|_2^2) = \sigma(\mu)(q_1) = \sigma(\mu)(q_1 \mid q_1 > 0)\sigma(\mu)[q_1 > 0] = \sigma'(\mu)(q_1)\mu(\|p\|_1), \tag{3.14}$$

where $\sigma'(\mu)[q_1 \in \cdot]$ is the size-biased distribution conditioned on positive components defined. Now, using (2.5) and disintegration of measures we write

$$\sigma'(\mu)(q_1) = \int_{\overline{\nabla}} \|p\|_1 \int_{\nabla} q_1\, \sigma_{p/\|p\|_1}[dq]\, \mu[dp]$$

$$= \int_0^1 \bar{\alpha} \int_{\overline{\nabla}} \int_{\nabla} q_1\, \sigma_{p/\|p\|_1}[dq]\, \mu^{(\bar{\alpha})}[dp]\, \mu[\|p\|_1 \in d\bar{\alpha}],$$

where $\mu^{(\bar{\alpha})} = \mu[\,\cdot\,|\|p\|_1 = \bar{\alpha}]$ denotes the measure μ conditioned on $\nabla_{[0,\bar{\alpha}]}$, which is well defined almost surely w.r.t. $\mu[\|p\|_1 \in d\bar{\alpha}]$. Since $\mathcal{G}_\theta$ (2.1) conserves $\|p\|_1$, $\mu^{(\bar{\alpha})}$ is also invariant for $\mathcal{G}_\theta$ and thus equal to $\mathrm{PD}_{[0,\bar{\alpha}]}(\theta)$ by Proposition 2.1. Hence, $\sigma(\mu^{(\bar{\alpha})})[dq] = \int_{\overline{\nabla}} \sigma_{p/\|p\|_1}[dq]\, \mu^{(\bar{\alpha})}[dp]$, as in (2.4), is the GEM distribution and therefore

$$\int_{\overline{\nabla}} \int_{\nabla} q_1\, \sigma_{p/\|p\|_1}[dq]\, \mu^{(\bar{\alpha})}[dp] = \frac{1}{1+\theta}, \quad \text{for a.e. } \bar{\alpha} \in (0,1].$$

Thus,

$$\sigma'(\mu)(q_1) = \frac{1}{1+\theta} \int_0^1 \bar{\alpha}\, \mu[\|p\|_1 \in d\bar{\alpha}] = \frac{1}{1+\theta} \mu(\|p\|_1)$$

which, with (3.14), yields

$$\mu(\|p\|_2^2) = \frac{1}{1+\theta} \mu(\|p\|_1)^2,$$

and concludes the proof. $\qquad \square$

Remark 3.7. It is interesting to note that Lemma 3.6 holds in greater generality, because all k-norms are continuous whenever $k > 1$. We have

$$\alpha_\mu^k = \mu(\|p\|_1)^k = \frac{1}{(k-1)!}\left[\prod_{j=1}^{k-1}(j+\theta)\right]\mu(\|p\|_k^k) \qquad \text{for all } k \geqslant 1\,.$$

This can be proven in exactly the same way as Lemma 3.6, with the only difference that we replace $\mu(\|p\|_2^2)$ with $\mu(\|p\|_k^k)$ when $k > 1$. The case $k = 1$ is trivial.

Lemmas 3.5 and 3.6 imply $\mu(\|p\|_1^2) = \mu(\|p\|_1)^2$, which is equivalent to

$$\text{Var}_\mu(\|p\|_1) = \mu(\|p\|_1^2) - \mu(\|p\|_1)^2 = 0\,. \tag{3.15}$$

Therefore, accumulation points μ concentrate on $\nabla_{[0,\alpha_\mu]}$ with $\alpha_\mu := \mu(\|p\|_1)$. To conclude the proof of Theorem 1.4, it is left to show that α_μ is independent of the choice of μ if assumption (A4) is satisfied.

Proof of Theorem 1.4. Let $\theta \in (0,1]$. By Proposition 3.1, we know that under Assumptions (A1) – (A3) all weak accumulation points of $(\mu_{L,N})_{L,N}$ are reversible w.r.t. $\mathcal{G}_\theta$. Furthermore, every such limit μ of a subsequence $(\mu_{L_j,N_j})_{j\geqslant 1}$ concentrates on $\nabla_{[0,\alpha_\mu]}$ for some $\alpha_\mu \in [0,1]$ as follows from (3.15). Thus, Proposition 2.1 implies that μ must be the Poisson-Dirichlet distribution on $[0,\alpha_\mu]$ with parameter θ, which proves the first statement.

The only control we have on α_μ from (A1) – (A3) is that

$$\alpha_\mu \leqslant 1 - \frac{\nu_\rho(\eta_1)}{\rho}\,, \quad \text{where} \quad \nu_\rho(\eta_1) = \sum_{n=0}^\infty nw(n)$$

is the expected density in the bulk. However, Assumption (A4) together with Lemma 3.6 implies uniqueness of the limit, since for every accumulation point μ we have

$$\mu(\|p\|_1) = \sqrt{(1+\theta)\mu(\|p\|_2^2)} = \sqrt{\frac{1+\theta}{\rho}\lim_{N/L\to\rho}\frac{\pi_{L,N}(\eta_x^2)}{N}} = \alpha\,, \tag{3.16}$$

where we used the identity (3.11) in the second equality. This implies the second statement and concludes the proof of the theorem. □

4 Proof of Theorem 1.7

In this section we prove Theorem 1.7, a specialized version of Theorem 1.4 which is better suited for application to condensation in particle systems. In contrast to assumptions (A1) – (A3), we require uniform convergence of the weights as well as stronger control on the weights in sub-L scales in (B2). However, this allows us to drop assumption (A4) that was needed to guarantee the concentration of the macroscopic phase in Theorem 1.4. Thanks to the stronger assumptions, we can explicitly calculate the form of limiting single-site marginals, which will imply the equivalance of ensembles (A2) and the condensation transition.

In Appendix A we prove the equivalence of ensembles and deduce that the system defined by weights $(w_L)_L$ satisfying (B1) (and a growth condition on the weights which is weaker than (B2)) exhibits condensation for ρ large enough. We summarize this result here, a more general and detailed version is given in Proposition A.1.

Proposition 4.1 (Equivalence of ensembles)**.** *Consider weights $(w_L)_L$ and w satisfying (B1), (B2) with corresponding canonical measures $(\pi_{L,N})_{L,N}$ as defined in (1.1). Then*

the system exhibits a condensation transition in the thermodynamic limit $N/L \to \rho \geqslant 0$ (cf. Definition 1.1) with critical density $\rho_c := \sum_{n=1}^{\infty} n w(n)$. More precisely, we have convergence of single-site marginals $\pi_{L,N}[\eta_x \in \cdot] \to \nu_\rho$ such that

$$\nu_\rho(\eta_x) = \begin{cases} \rho & \text{if } \rho < \rho_c, \\ \rho_c & \text{if } \rho \geqslant \rho_c. \end{cases} \tag{4.1}$$

and $\nu_\rho = w$ for $\rho \geqslant \rho_c$.

This establishes existence of the condensed phase for $\rho > \rho_c$, and the following result guarantees that it agrees with the macroscopic phase, and there is no mass on intermediate scales.

Proposition 4.2. *Consider weights $(w_L)_L$ and w satisfying (B1) and (B2) with corresponding canonical measures $(\pi_{L,N})_{L,N}$. Then, for every $\rho > \rho_c$*

$$\lim_{\varepsilon \to 0} \lim_{N/L \to \rho} \pi_{L,N}[\widetilde{\eta}_1 > \varepsilon N] = 1 - \frac{\rho_c}{\rho}.$$

Proof. We fix a density $\rho > \rho_c$. Now, by definition of the first size-biased marginal, cf. (2.7), we have for every $\varepsilon \in (0, 1]$

$$\pi_{L,N}[J < \widetilde{\eta}_1 \leqslant \varepsilon N] = \frac{L}{N} \sum_{n=J+1}^{\varepsilon N} n \, w_L(n) \frac{Z_{L-1,N-n}}{Z_{L-1,N}},$$

for N large enough. By Assumption (B2), there exists a sufficiently large $J = J(\varepsilon)$ such that

$$\lim_{L \to \infty} \sup_{n \geqslant J} |n \, w_L(n) L - \theta| < \varepsilon.$$

On the other hand, applying Lemma 4.3 below yields

$$\limsup_{N/L \to \rho} \sup_{J < n \leqslant \varepsilon N} \frac{Z_{L-1,N-n}}{Z_{L-1,N}} \leqslant \left(1 - \frac{\varepsilon}{1 - \rho_c/\rho}\right)^{-1}.$$

Altogether,

$$\lim_{N/L \to \rho} \pi_{L,N}[J < \widetilde{\eta}_1 \leqslant \varepsilon N] \leqslant \varepsilon (\theta + \varepsilon) \left(1 - \frac{\varepsilon}{1 - \rho_c/\rho}\right)^{-1},$$

which vanishes as we take the limit $\varepsilon \to 0$. This completes the proof, since

$$\lim_{N/L \to \rho} \pi_{L,N}[\widetilde{\eta}_1 \leqslant J] = \lim_{N/L \to \rho} \frac{L}{N} \sum_{n=0}^{J} n \, \pi_{L,N}[\eta_1 = n] = \frac{\rho_c}{\rho},$$

which is a direct implication of Proposition 4.1. $\qquad \square$

We now prove the key estimate used in Proposition 4.2, which guarantees that the ratio of partition functions $Z_{L-1,N-n}/Z_{L,N}$ does not blow up for $n = o(N)$.

Lemma 4.3. *Assume that (B1) and (B2) are both satisfied. Then we have*

$$\limsup_{N/L \to \rho} \frac{Z_{L-1,(1-\kappa)N}}{Z_{L,N}} \leqslant \left(1 - \frac{\overline{\kappa}}{1 - \rho_c/\rho}\right)^{-1},$$

for every $\kappa = \kappa(L) = O(1)$ such that $\kappa \leqslant 1 - \frac{\rho_c}{\rho}$, where $\overline{\kappa} := \limsup_{L \to \infty} \kappa(L)$.

Proof. Fix $\varepsilon \in (0,1)$, by (B2) and Remark 1.8 there exists $J \in \mathbb{N}_0$ such that

$$\frac{m\, w_L(m)}{n\, w_L(n)} \geqslant 1 - \varepsilon \quad \text{for all } m, n \geqslant J, \tag{4.2}$$

for all L sufficiently large (depending on J and ε). Let $K = \lceil \kappa N \rceil$, by definition of the canonical measures we have

$$
\pi_{L,N}(\eta_1 \,;\, \eta_1 > J) \;\geqslant\; \pi_{L,N}(\eta_1 \,;\, \eta_1 > J + K) = \sum_{n > J+K}^{N} n\, w_L(n) \frac{Z_{L-1,N-n}}{Z_{L,N}}
$$

$$
= \sum_{n > J}^{N-K} (n+K)\, w_L(n+K) \frac{Z_{L-1,N-K-n}}{Z_{L,N}}
$$

$$
\geqslant \big(1 - \varepsilon\big) \sum_{n > J}^{N-K} n\, w_L(n) \frac{Z_{L-1,N-K-n}}{Z_{L,N}} \,,
$$

where the last inequality follows from (4.2). The sums above are all non-empty because $N - K > J$ for L, N sufficiently large since $\kappa < 1$. Multiplying and dividing by $Z_{L,N-K}$ we have

$$
\pi_{L,N}(\eta_1 \,;\, \eta_1 > J) \;\geqslant\; \big(1 - \varepsilon\big) \frac{Z_{L,N-K}}{Z_{L,N}} \pi_{L,N-K}(\eta_1 \,;\, \eta_1 > J) \,. \tag{4.3}
$$

By equivalence of ensembles (Proposition 4.1), in the thermodynamic limit $N/L \to \rho \geqslant \rho_c$ the single site marginals of $\pi_{L,N}$ converge weakly to ν_ρ which has mean ρ_c, so

$$
\pi_{L,N}(\eta_1 \,;\, \eta_1 > J) = \frac{N}{L} - \pi_{L,N}(\eta_1 \,;\, \eta_1 \leqslant J) \to \rho - \rho_J, \quad \text{as } N/L \to \rho \,,
$$

where ρ_J is given by $\nu_\rho(\eta_1 \,;\, \eta_1 \leqslant J)$. Furthermore, by dominated convergence, $\rho_J \to \rho_c$ as $J \to \infty$. Now taking the thermodynamic limit in (4.3), followed by the limit $J \to \infty$, we have

$$
\limsup_{N/L \to \rho} \frac{Z_{L,(1-\kappa)N}}{Z_{L,N}} \;\leqslant\; \frac{\rho - \rho_c}{(1 - \overline{\kappa})\rho - \rho_c} \frac{1}{1 - \varepsilon} \,.
$$

From the equivalence of ensembles in Proposition 4.1, and the identity $\pi_{L,N}[\eta_1 = 0] = w_L(0) Z_{L-1,N}/Z_{L,N}$, we observe that $\lim_{N/L \to \rho} Z_{L-1,N}/Z_{L,N} = 1$, which completes the proof, since we can choose ε arbitrarily small after taking $J \to \infty$. $\qquad \square$

We are now ready to state the full proof of Theorem 1.7 which follows by putting together the statements of Theorem 1.4, Proposition 4.1 and Proposition 4.2.

Proof of Theorem 1.7. The equivalence of ensembles implies the condensation transition with critical density ρ_c, and with its formulation in Proposition A.1, also Assumption (A2) is satisfied. Next, we recover assumptions (A1), (A3) and (A4) for a fixed choice of $\rho > \rho_c$. Clearly, (A1) is a direct implication of (B1), also (A3) follows immediately from (B2). This already yields that the subsequential limits of the corresponding measures $\mu_{L,N}$ are Poisson-Dirichlet distributions. It is only left to show that the macroscopic phase is non-trivial and indeed agrees with the condensed phase. Recall from (3.13) that

$$
\mu(\|p\|_1) = \sigma(\mu)[q_1 > 0] = \lim_{\varepsilon \to 0} \sigma(\mu)[q_1 > \varepsilon]
$$

and, as $\varepsilon \to 0$,

$$\sigma(\mu)[q_1 > \varepsilon] = \lim_{N/L \to \rho} \mu_{L,N}[\widetilde{p}_1 > \varepsilon]$$

$$= \lim_{N/L \to \rho} \pi_{L,N}[\widetilde{\eta}_1 > \varepsilon N] \to 1 - \frac{\rho_c}{\rho} \,,$$

which follows from Proposition 4.2. Hence, with (3.16) we conclude (A4) with $\alpha = \mu(\|p\|_1) = 1 - \frac{\rho_c}{\rho}$. Altogether, we verified (A1) – (A4) and so can apply Theorem 1.4. $\quad\square$

5 Application to interacting particle systems and conclusion

As mentioned in the introduction, condensation transitions occur naturally and have been studied extensively for interacting particle systems, more precisely for stochastic lattice gases which model transport phenomena and conserve the number of particles. A large class of such models with state space $\Omega_{L,N}$ has been introduced in [7] with infinitesimal generator of the form

$$\mathcal{L}^{(L)} f(\eta) = \sum_{x,y=1}^{L} p(x,y) u(\eta_x, \eta_y) \left[f(\eta^{xy}) - f(\eta) \right] \,, \tag{5.1}$$

where $f \in C_b(\Omega_{L,N})$. Whenever $\eta_x > 0$, η^{xy} denotes the configuration $\eta - e^{(x)} + e^{(y)}$ where one particle moved from site x to site y. Recall that $e_z^{(x)} = \delta_{x,z}$. The jump rate $u(n,m) \geqslant 0$ is a non-negative function of the occupation numbers n on the departure and m on the target site of a particle jump, and to avoid degeneracies we assume that $u(n,m) = 0$ if and only if $n = 0$. $p(x,y)$ denotes an irreducible probability kernel on $\{1, \dots, L\}$ and models the geometry of the underlying lattice. Systems have been studied, e.g. on regular lattices in various dimensions and with different boundary conditions, here we assume that the system is closed and conserves the total number of particles $\sum_{x=1}^{L} \eta_x$.

For any fixed number of particles $N \in \mathbb{N}$, the operator $\mathcal{L}^{(L)}$ defines an irreducible, continuous-time Markov process on the finite state space $\Omega_{L,N}$, which therefore has a unique invariant distribution $\pi_{L,N}$. It has been established in [7, 15] that this distribution is indeed of product form and spatially homogeneous, cf. (1.1), under the conditions:

$$\frac{u(n+1,m)}{u(m+1,n)} = \frac{u(n+1,0)}{u(1,n)} \frac{u(1,m)}{u(m+1,0)} \qquad \forall n, m \geqslant 0 \,, \tag{5.2}$$

and at least one of the following

- $p(\cdot, \cdot)$ is symmetric,

or

- $p(\cdot, \cdot)$ is doubly stochastic, i.e. $\sum_y \big(p(x,y) - p(y,x) \big) = 0$, and

$$u(n,m) - u(m,n) = u(n,0) - u(m,0) \qquad \forall n, m \geqslant 0 \,. \tag{5.3}$$

Then the stationary weights are

$$w(n) = \prod_{k=1}^{n} \frac{u(1,k-1)}{u(k,0)} \qquad \forall n \geqslant 0 \,, \tag{5.4}$$

which depend only on the jump rates u but not on the kernel p.

For the special case of zero-range dynamics, (5.4) simplifies further since $u(k, \cdot) = u(k)$, and (5.2) and (5.3) are fulfilled. In this case (5.4) leads to the simple identification between stationary weights and rates

$$u(n) = \frac{w(n-1)}{w(n)} \qquad \forall n \geqslant 1 \,. \tag{5.5}$$

Due to this simple one-to-one correspondence between weights and transition rates, zero-range processes provide a generic framework of studying condensation transitions in interacting particle systems.

Another less restrictive simplification is to assume that $u(n, m) = u_1(n) u_2(m)$ is of product form, which automatically satisfies (5.2) and always leads to factorized stationary measures under symmetric dynamics, i.e. $p(\cdot, \cdot)$ is a symmetric transition kernel. The weights from (5.4) now take the form

$$w(n) = \left(\frac{u_1(1)}{u_2(0)}\right)^n \prod_{k=1}^n \frac{u_2(k-1)}{u_1(k)} \qquad \forall n \geqslant 0 \,. \tag{5.6}$$

Note that due to the conservation law the exponential factor $\left(\frac{u_1(1)}{u_2(0)}\right)^n$ is usually omitted, since it cancels in the definition of $\pi_{L,N}$ (1.1). One particular example is the inclusion process with rates

$$u(n, m) = n\,(d + m) \quad \text{with } d > 0 \,, \tag{5.7}$$

which leads to the stationary weights (1.6). If we set

$$d = d(L) > 0 \quad \text{with} \quad dL \to \theta > 0 \text{ as } L \to \infty \,, \tag{5.8}$$

the system exhibits a condensation transition with $\rho_c = 0$ and a Poisson-Dirichlet structure with $\mathrm{PD}_{[0,1]}(\theta)$, see [28, Theorem 1] or (1.7) above. Theorem 1.7 recovers this result for $\theta \in (0, 1]$. The restriction of θ is solely due to the fact that the Poisson-Dirichlet distribution is (so far) only proven to be the unique invariant distribution for the split-merge process if $\theta \in (0, 1]$.

Using the size-dependent parameter d as above, we will provide a few instructive examples of particle systems with size-dependent jump rates u_L of product form, where Theorem 1.7 applies with a non-trivial critical density $\rho_c > 0$. We start by fixing the weights

$$w_L(n) = \begin{cases} w(n) & \text{if } n \leqslant A \,, \\ \frac{d}{n} & \text{if } n > A \,, \end{cases} \quad \text{for some fixed } A \in \mathbb{N} \,, \tag{5.9}$$

where $d = d(L)$ as in (5.8), $\theta \in (0, 1]$ and w is a probability mass function on the set $\{0, \dots, A\}$. Let us fix for simplicity the uniform distribution with $w(n) = (1 + A)^{-1} \mathbb{1}\{n \leqslant A\}$. Using (5.5), the corresponding rates of a zero-range process are given by $u_2(m) = 1$ and

$$u_L(n) = u_{1,L}(n) = \begin{cases} 1 & \text{if } 1 \leqslant n \leqslant A \,, \\ d^{-1} \simeq L/\theta & \text{if } n = A + 1 \,, \\ \frac{n}{n-1} & \text{if } n > A + 1 \,. \end{cases} \tag{5.10}$$

This underlines the mechanism that leads to condensation in such systems: Sites with occupation numbers different from $A + 1$ are stable and eject particles at rates of order 1. Sites with occupation number $A + 1$ are unstable and eject a particle at diverging rate of order L, creating a sharp threshold between bulk sites ($\eta_x \leqslant A$) and cluster sites ($\eta_x > A + 1$). Note that with (5.6) the same effect could be achieved by a vanishing rate of arrival onto sites with occupation number A. If in the thermodynamic limit

$$N/L \to \rho > \rho_c = \frac{A}{2} \,,$$

i.e. the total density exceeds the expectation of the uniform bulk distribution, the system exhibits a condensation transition with mass fraction $1 - \rho_c/\rho$ in the condensate. Heuristically, the excess mass is expelled from the bulk and accumulates in stable clusters with occupation numbers larger than $A + 1$. The particular form of the rates for those clusters leads to a macroscopic phase with Poisson-Dirichlet statistics $\mathrm{PD}_{[0,1-\rho_c/\rho]}(\theta)$, which follows from Theorem 1.7 and is illustrated in Figure 2. Clearly, the weights $(w_L)_L$ converge uniformly to w and therefore satisfy both Assumption (B1) and (B2). The same process with rates $du_L(n)$ corresponds to asymptotically vanishing exit rates from stable sites, which is simply a time change and leads of course to the same stationary behaviour.
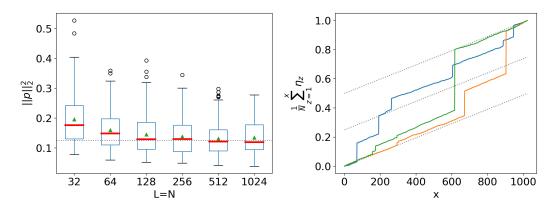

Figure 2: Simulation results at stationarity for a zero-range process with rates (5.10), parameters $\theta = A = \rho = 1$ and $\rho_c = 1/2$. The left boxplot displays $\|p\|_2^2$ where $p = \frac{\hat{\eta}}{N}$ was sampled 200-times for every system size, indicating convergence to $\frac{1}{8}$ which agrees with $\mu_{L,N}(\|p\|_2^2) \simeq (1+\theta)^{-1}(1 - \frac{\rho_c}{\rho})^2$ from Lemma 3.6. The right plot shows three samples of accumulated configurations in a system of size $N = L = 1024$, where we clearly see occurance of large clusters, with a background density of $\rho_c = 1/2$ indicated by dotted lines.

We can also generalize the inclusion process dynamics (5.7) to stationary weights of the form (5.9). Consider a process with rates of product form $u_L(n,m) = u_{1,L}(n)u_{2,L}(m)$ with

$$u_{1,L}(n) = u_{2,L}(n) = \begin{cases} d \simeq \theta/L & \text{if } 1 \leqslant n \leqslant A, \\ n & \text{if } n > A, \end{cases} \qquad (5.11)$$

and $u_{2,L}(0) = d$. It is easy to see from (5.6) that the stationary weights for this process are given by (5.9) and Theorem 1.7 applies. Heuristically, sites with occupation number up to A eject and attract particles at a slow rate d, and particles on sites with higher occupation numbers become "free" and leave independently at the same rate and also attract other particles, as a simple generalization of the standard inclusion interaction. As a result, the dynamics in the condensed phase happen at a much higher rate than in the bulk. This separation of time scales leads to completely different dynamics than in the zero-range example above, even though both models share the same stationary distributions. We want to stress that due to the general nature of Assumptions (B1) and (B2), Theorem 1.7 applies also to modifications of these examples and the particular form of the weights (5.9) is not important.

Understanding the dynamics of the condensed phase in these models is a very interesting question for future research, in particular the coarsening regime, where macroscopic

clusters emerge from homogeneous initial conditions and approach stationarity by exchanging particles. Heuristically, their stationary mass partition can be understood as a balance between aggregation and fragmentation of macroscopic clusters. Note that these dynamics are not described by split-merge processes, which we only use as an auxiliary tool to characterize PD distributions, but are rather of a diffusive nature. A diffusive model on partitions that has stationary PD distribution has been introduced in [12], and it would be very interesting to study hydrodynamic scaling limits in this context.

As we have seen, the Poisson-Dirichlet structure in the macroscopic phase arises due to uniform stationary weights under size-biased sampling, which leads to particular rates in the zero-range process (5.10) or the generalized inclusion process (5.11) for large occupation numbers. In the context of the dynamics of interacting particle systems this is only one particular case, and it would be interesting to study the statistics of the condensed phase beyond Poisson-Dirichlet under a different scaling behaviour of the weights.

A Equivalence of ensembles

In this section we prove that, under assumption (B1) and if the weights $(w_L)_L$ decay sub exponentially, i.e.

$$\frac{1}{L}\log w_L(aL) \to 0, \quad \forall a > 0, \quad \text{and} \quad \sum_{n=0}^{\infty} n^2 w(n) < \infty, \tag{A.1}$$

we have equivalence of ensembles and condensation in the sense of weak convergence of finite dimensional marginals. Note that this includes condensation in the sense of Definition 1.1, and (A.1) is weaker than Assumption (B2) in Theorem 1.7. The result is in the same spirit as previous results on equivalence of ensembles and condensation in stochastic particle systems with stationary product measures (see for example [6]). However, as far as we know, this is the first general result in this direction for models with size-dependent weights.

In order to state the result in more generality we first introduce some extra notation. For a sequence of non-negative, non-trivial weights, $(w_L(n))_{n \in \mathbb{N}_0}$, possibly depending on the system size L, we define a family of probability measures on $\mathbb{N}_0$ by tilting the weights by a non-negative *fugacity* parameter $\phi \geqslant 0$:

$$\bar{\nu}_{\phi,L}[dn] := \frac{1}{z_L(\phi)} w_L(n)\phi^n dn \quad \text{with} \quad z_L(\phi) := \sum_{n=0}^{\infty} w_L(n)\phi^n,$$

which is well defined for each $\phi \in D_L := \{\phi : z_L(\phi) < \infty\}$. The corresponding family of *grand-canonical distributions* is given by the product measures

$$\bar{\nu}_{\phi,L}^{\otimes L}[d\eta] = \frac{1}{z_L(\phi)^L} \prod_{x=1}^{L} w_L(\eta_x)\phi^{\eta_x} d\eta, \tag{A.2}$$

which are defined on the configuration space $\Omega_L = \mathbb{N}_0^L = \bigcup_{N=0}^{\infty} \Omega_{L,N}$, where the total number of particles is arbitrary. The expected number of particles per site (density) $R_L : D_L \to [0, \infty)$ is a strictly increasing function of ϕ, with

$$R_L(\phi) := \bar{\nu}_{\phi,L}(\eta_x) = \phi\, \partial_\phi \log z_L(\phi), \quad \text{for } \phi \in D_L. \tag{A.3}$$

We denote the inverse of R_L by Φ_L. Furthermore, we define the variance of $\bar{\nu}_{\phi,L}$

$$\sigma_L^2(\phi) = \bar{\nu}_{\phi,L}(\eta_x^2) - R_L(\phi)^2,$$

which is finite for each ϕ in the interior of D_L. By construction, the canonical measures (1.1) on $\Omega_{L,N}$ are given by conditioning any grand-canonical measure on the total number of particles, i.e.

$$\pi_{L,N}[d\eta] := \bar{\nu}_{\phi,L}^{\otimes L}\left[d\eta \,\bigg|\, \sum_{x=1}^{L} \eta_x = N\right] = \frac{1}{Z_{L,N}} \prod_{x=1}^{L} w_L(\eta_x) d\eta\,,$$

which is independent of $\phi \in D_L$.

By Assumption (B1), the weights $(w_L(n))_{n,L}$ converge uniformly in n as $L \to \infty$ to a probability measure $(w(n))_{n \in \mathbb{N}_0}$. We define the limiting grand-canonical measures by

$$\bar{\nu}_\phi[dn] := \frac{1}{z(\phi)} w(n)\phi^n dn \quad \text{with} \quad z(\phi) := \sum_{n=0}^{\infty} w(n)\phi^n\,.$$

Since $(w(n))_{n \in \mathbb{N}_0}$ is normalised, these measures must exist at least for each $\phi \in [0,1]$, and $\bar{\nu}_1$ corresponds to the weights w. By analogy with (A.3), we define the function $R(\phi) = \bar{\nu}_\phi(\eta_1)$, which is a strictly increasing function $R: [0,1] \to [0,\rho_c]$ with

$$\rho_c = R(1) = \sum_{n=0}^{\infty} nw(n)\,, \quad \text{as given in Theorem 1.7}\,.$$

We denote the inverse of R by Φ, so that the average particle density under $\bar{\nu}_{\Phi(\rho)}^{\otimes L}$ is ρ for all $\rho \leqslant \rho_c$. Further, we denote the variance of $\bar{\nu}_\phi$ by $\sigma^2(\phi)$ which is finite for $\phi \in [0,1]$ by the second moment condition in (A.1). Note that it may be possible that $R(\phi) < \infty$ and $\bar{\nu}_\phi$ is well defined also for $\phi > 1$, but such measures are not accessible as limits of $\bar{\nu}_{\phi,L}$ and do not play a role in the following.

Under Assumption (B1) and sub-exponential weights (A.1), it turns out that there is a condensation transition according to Definition 1.1 with critical density $\rho_c = R(1)$. In this case, in the thermodynamic limit $N/L \to \rho$, all finite dimensional marginals of the canonical measures converge weakly to the limiting grand-canonical measures with density ρ if $\rho \leqslant \rho_c$, and with density ρ_c if $\rho \geqslant \rho_c$. This implies that for $\rho > \rho_c$ the excess mass must condense on a vanishing volume fraction in the thermodynamic limit.

Proposition A.1 (Equivalence of ensembles). *Consider non-negative weights $(w_L)_L$ satisfying (B1) with limit w. Furthermore, assume that w_L's have sub-exponential tails, in the sense of (A.1). Then for each $M \subset \mathbb{N}_0$, with $|M| = m < \infty$, denoting the marginal of $\pi_{L,N}$ on M by $\pi_{L,N}^M$, we have*

$$\pi_{L,N}^M \xrightarrow{d} \begin{cases} \bar{\nu}_{\Phi(\rho)}^{\otimes m} & \text{if } \rho < \rho_c\,, \\ \bar{\nu}_1^{\otimes m} & \text{if } \rho \geqslant \rho_c\,, \end{cases}$$

where $\rho_c = R(1) < \infty$.

Throughout the proof we assume further that condition (1.11) is satisfied, i.e.

$$w(0) > 0 \quad \text{and} \quad \sup_n[w(n-1) \wedge w(n)] > 0\,,$$

which implies that the variance, $\sigma^2(\phi)$, given by $\sum n^2 w(n)\phi^n$, is positive for each $\phi \in (0,1]$. The special case of $w(0) = 1$ is covered at the end of the proof.

We firstly observe that for each $\rho \in [0,\rho_c]$ we can construct a sequence of size-dependent fugacities such that the mean of the size-dependent grand-canonical measures converges to ρ and the variance remains bounded. The typical behaviour of $R_L(\phi)$ is illustrated in Figure 3.
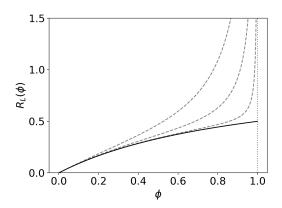
Figure 3: The dashed lines represent $R_L(\phi)$ for systems of sizes $L \in \{4, 16, 128\}$ and weights (5.9) with $\theta = A = 1$. The black line denotes the limit $R(\phi) = \frac{\phi}{1+\phi}$.

Lemma A.2. *Under Assumption (B1) and if the weights $(w_L)_L$ are sub-exponential in the sense of* (A.1)*, then for each $\varphi \in [0, 1]$ there exists a sequence $(\varphi_L)_L$ in $[0, 1)$, with limit point φ, such that*

$$\|w_L - w\|_\infty \frac{1}{(1 - \varphi_L)^3} \to 0 \,, \quad \text{as } L \to \infty \,, \tag{A.4}$$

and

$$R_L(\varphi_L) \to R(\varphi) \in [0, \rho_c] \,, \quad \sigma_L^2(\varphi_L) \to \sigma^2(\varphi) > 0 \,, \quad \text{and } z_L(\varphi_L) \to z(\varphi) \,. \tag{A.5}$$

Proof. For $\varphi \in [0, 1)$ we may choose $\varphi_L = \varphi$ for each L. For $\varphi = 1$, we let $\varphi_L = 1 - \|w_L - w\|_\infty^{1/4}$. Then $\varphi_L \to 1$ as $L \to \infty$ and condition (A.4) is satisfied since $\|w_L - w\|_\infty \to 0$ by assumption (B1).

It is only left to prove the convergence in (A.5), which is equivalent to showing that

$$\lim_{L \to \infty} \sum_{n=0}^\infty n^p w_L(n) \varphi_L^n = \sum_{n=0}^\infty n^p w(n) \varphi^n \,,$$

for $p \in \{0, 1, 2\}$. Since the weights converge uniformly, we have

$$\left| \sum_{n=0}^\infty n^p (w_L(n) - w(n)) \varphi_L^n \right| \leqslant \|w_L(n) - w(n)\|_\infty \sum_{n=0}^\infty n^p \varphi_L^n \,.$$

Term-by-term differentiation of the geometric series yields for $p \in \{0, 1, 2\}$

$$\sum_{n=0}^\infty n^p \varphi_L^n = \frac{(p \vee 1)\varphi_L^p}{(1 - \varphi_L)^{p+1}} + \frac{\varphi_L}{(1 - \varphi_L)^2} \mathbb{1}\{p = 2\} \,.$$

Hence, using again (A.4), we have

$$\sum_{n=0}^\infty n^p w_L(n) \varphi_L^n \simeq \sum_{n=0}^\infty n^p w(n) \varphi_L^n \to \sum_{n=0}^\infty n^p w(n) \varphi^n \,,$$

where we used dominated convergence in the last step. $\qquad\square$

We will prove Proposition A.1 by showing that the relative entropy between the single site marginal of $\pi_{L,N}$ and $\bar{\nu}_{\varphi_L,L}$ vanishes, where the limit density $R(\varphi)$ is equal to ρ in the sub-critical case, and ρ_c in the super critical case. Optimally, in the super-critical case, we would like to measure the relative entropy w.r.t. the limiting measure directly, however this is not possible since $\pi_{L,N}[\eta_x \in \cdot\,] \ll w$ is in general not satisfied and the relative entropy would be infinite.

The main tool we rely on in the proof of Proposition A.1 is a local central limit theorem (see for example [8, Theorem 1.2]) which allows us to estimate the decay of the relative entropy. For completeness we include the local limit theorem here. To state it, we first introduce the Bernoulli part decomposition q of a probability measure $\mathbb{P}_{x,L}$ on $\mathbb{Z}$ as

$$q(\mathbb{P}_{x,L}) := \sum_{n \in \mathbb{Z}} (\mathbb{P}_{x,L}[n] \wedge \mathbb{P}_{x,L}[n+1]) \,.$$

Moreover, for a family of measures $(\mathbb{P}_{x,L})_{1 \leqslant x \leqslant L}$ we define $Q_L := \sum_{x=1}^{L} q(\mathbb{P}_{x,L})$.

Lemma A.3 ([8, Theorem 1.2] Local central limit theorem)**.** *Consider a triangular array of independent integer valued random variables $\eta_{x,L}$, for $1 \leqslant x \leqslant L$, and $L \in \mathbb{N}$, where $\eta_{x,L}$ has law $\mathbb{P}_{x,L}$. Suppose there exist sequences a_L and b_L, $L \geqslant 1$, such that $b_L \to \infty$, $\limsup_{L \to \infty} b_L^2 / Q_L < \infty$ and*

$$\frac{1}{b_L} \sum_{x=1}^{L} (\eta_{x,L} - a_L) \xrightarrow{d} \mathcal{N}(0,1) \,. \tag{A.6}$$

Then

$$\sup_{n \in \mathbb{Z}} \left| b_L \, \mathbb{P}_L \Big[\sum_{x=1}^{L} \eta_{x,L} = n \Big] - g\Big(\frac{n - a_L}{b_L} \Big) \right| \to 0 \,,$$

where $\mathbb{P}_L$ denotes the product measure $\bigotimes_{x=1}^{L} \mathbb{P}_{x,L}$ and g the density of a standard normal.

To apply Lemma A.3 we consider independent random variables

$$(\eta_{x,L})_{1 \leqslant x \leqslant L} \quad \text{with law} \quad \bar{\nu}_{\varphi_L,L}^{\otimes L} \quad \text{for each } L \geqslant 1 \,, \tag{A.7}$$

where $(\varphi_L)_{L \in \mathbb{N}}$ is a sequence in $[0,1)$ satisfying (A.4). To apply Lemma A.3 in the proof of Proposition A.1, we first verify the central limit theorem (A.6) for the $\eta_{x,L}$'s.

Lemma A.4. *Consider $(\varphi_L)_L$ to be a sequence in $[0,1)$ with limit point $\varphi \in [0,1]$ satisfying (A.4) and (A.5). Furthermore, let $(\eta_{x,L})_{1 \leqslant x \leqslant L, L \in \mathbb{N}}$ be as in (A.7) and define*

$$\zeta_{x,L} := \frac{\eta_{x,L} - R_L(\varphi_L)}{\sqrt{L \, \sigma_L^2(\varphi_L)}} \,.$$

Then $\sum_{x=1}^{L} \zeta_{x,L}$ converges weakly to a standard normal, as L tends to infinity.

Proof. We want to apply the Lindeberg-Feller central limit theorem, see [30, Theorem 5.12]: because the $\zeta_{x,L}$'s are centered and normalised, it suffices confirm that the following Lindeberg condition holds:

$$\text{For every } \varepsilon > 0, \quad \lim_{L \to \infty} \sum_{x=1}^{L} \bar{\nu}_{\varphi_L,L}(\zeta_{x,L}^2 \mathbb{1}\{|\zeta_{x,L}| > \varepsilon\}) = 0 \,.$$

Since $R_L(\varphi_L)$ and $\sigma_L^2(\varphi_L)$ converge to positive numbers, we have that for L large enough

$$
\sum_{x=1}^{L} \bar{\nu}_{\varphi_L, L}(\zeta_{x,L}^2 \mathbb{1}\{|\zeta_{x,L}| > \varepsilon\})
$$

$$
= \frac{1}{z_L(\varphi_L)\,\sigma_L^2(\varphi_L)} \sum_{n=0}^{\infty} (n - R_L(\varphi_L))^2 w_L(n)\varphi_L^n \mathbb{1}\Big\{|n - R_L(\varphi_L)| > \varepsilon\sqrt{L\,\sigma_L^2(\varphi_L)}\Big\}
$$

$$
\leqslant \frac{1}{z_L(\varphi_L)\,\sigma_L^2(\varphi_L)} \sum_{n=K_{\varepsilon,L}}^{\infty} n^2 w_L(n)\varphi_L^n\,,
$$

where $K_{\varepsilon,L} = \varepsilon\sqrt{L\,\sigma_L^2(\varphi_L)} + R_L(\varphi_L)$, which diverges like $\sqrt{L}$. The denominator in the above expression converges to a positive constant. For the numerator, we observe

$$
\Big| \sum_{n=K_{\varepsilon,L}}^{\infty} n^2(w_L(n) - w(n))\varphi_L^n \Big| \leqslant \|w_L(n) - w(n)\|_\infty \sum_{n=K_{\varepsilon,L}}^{\infty} n^2\varphi_L^n \to 0\,, \qquad \text{(A.8)}
$$

where we used property (A.4) of the sequence $(\varphi_L)_L$. Thus, using the second-moment assumption on w in (A.1), we have, for each ϵ

$$
\sum_{n=K_{\varepsilon,L}}^{\infty} n^2 w_L(n)\varphi_L^n \simeq \sum_{n=K_{\varepsilon,L}}^{\infty} n^2 w(n)\varphi_L^n \leqslant \sum_{n=K_{\varepsilon,L}}^{\infty} n^2 w(n) \to 0\,,
$$

as $L \to \infty$, since $K_{\epsilon,L} \to \infty$. This concludes the Lindeberg condition. $\qquad\square$

Proof of Proposition A.1. We first consider the sub-critical and critical case together, fix $\rho \in (0, \rho_c]$. Let $(\varphi_L)_L$ be a sequence converging to $\varphi := \Phi(\rho) \in (0,1]$ satisfying (A.4) and (A.5). We will measure the relative entropy between single-site marginals of $\pi_{L,N}[\eta_x \in \cdot]$ and $\bar{\nu}_{\varphi_L,L}$.

We start with an expression for the relative entropy between $\pi_{L,N}$ and $\bar{\nu}_{\varphi_L,L}^{\otimes L}$ which is used frequently in the proof of similar equivalence of ensembles results (see for example [6]),

$$
H\left(\pi_{L,N} \mid \bar{\nu}_{\varphi_L,L}^{\otimes L}\right) = \sum_{\eta \in \Omega_{L,N}} \pi_{L,N}[\eta] \log\left(\frac{\pi_{L,N}[\eta]}{\bar{\nu}_{\varphi_L,L}^{\otimes L}[\eta]}\right) = \log\left(\frac{z_L(\varphi_L)^L}{Z_{L,N}\varphi_L^N}\right)
$$

$$
= -\log \bar{\nu}_{\varphi_L,L}^{\otimes L}\left[\sum_{x=1}^{L} \eta_x = N\right].
$$

Then, by subadditivity of the relative entropy we have for marginals

$$
H\left(\pi_{L,N}[\eta_x \in \cdot] \mid \bar{\nu}_{\varphi_L,L}\right) \leqslant -\frac{1}{L}\log \bar{\nu}_{\varphi_L,L}^{\otimes L}\left[\sum_{x=1}^{L} \eta_x = N\right]. \qquad \text{(A.9)}
$$

We estimate the right-hand side using the local limit theorem in Lemma A.3 with the specific choices of

$$
a_L := L\,R_L(\varphi_L) \quad \text{and} \quad b_L := \sqrt{L\,\sigma_L^2(\varphi_L)}\,.
$$

It follows from Lemma A.4 that

$$
\sum_{x=1}^{L} \zeta_{x,L} = \frac{\sum_{x=1}^{L} \eta_{x,L} - a_L}{b_L} \xrightarrow{d} \mathcal{N}(0,1)\,.
$$

Moreover, b_L diverges in the large L limit because $\lim_{L \to \infty} \sigma_L^2(\varphi_L) = \sigma^2(\varphi) > 0$. Also,

$$\limsup_{L \to \infty} \frac{b_L^2}{Q_L} = \limsup_{L \to \infty} \frac{\sigma_L^2(\varphi_L)}{q(\bar{\nu}_{\varphi_L,L})} \leqslant \frac{\sigma^2(\varphi)}{\sup_{n \in \mathbb{N}_0} [w(n) \wedge w(n+1)]} \,,$$

where the final inequality follows by dominated convergence. The right hand side is finite by assumption (B1). Therefore, we may apply the local limit theorem stated in Lemma A.3 which yields

$$\bar{\nu}_{\varphi_L,L}^{\otimes L} \left[\sum_{x=1}^{L} \eta_x = N \right] = O\left(L^{-1/2} \right) \,,$$

and coming back to (A.9)

$$H\left(\pi_{L,N}[\eta_x \in \cdot\,] \mid \bar{\nu}_{\varphi_L,L} \right) \sim \frac{\log L}{L} \to 0 \quad \text{as } L \to \infty \,.$$

With Pinsker's inequality (see e.g. [22, Lemma 6.2]) this implies for the total variation distance

$$d_{TV}\left(\pi_{L,N}[\eta_x \in \cdot\,], \bar{\nu}_{\varphi_L,L} \right) \to 0 \,. \tag{A.10}$$

Also, by (A.5) and uniform convergence of the weights, $\bar{\nu}_{\varphi_L,L}$ converges weakly to $\bar{\nu}_{\varphi}$, which together with (A.10) implies $\pi_{L,N}[\eta_x \in \cdot\,] \xrightarrow{d} \bar{\nu}_{\Phi(\rho)}$ as $N/L \to \rho \leqslant \rho_c$.

Finally, we conclude the super-critical case $\rho > \rho_c$ using a large deviation estimate. Now let φ_L be a sequence converging to 1 and satisfying (A.4) and (A.5) so that $R_L(\varphi_L) \to \rho_c$. Then

$$-\frac{1}{L} \log \bar{\nu}_{\varphi_L,L}^{\otimes L} \left[\sum_{x=1}^{L} \eta_x = N \right] \leqslant -\frac{1}{L} \log \bar{\nu}_{\varphi_L,L}^{\otimes L \setminus \{1\}} \left[\sum_{x=2}^{L} \eta_x = N - \lfloor (\rho - \rho_c) L \rfloor \right]$$
$$- \frac{1}{L} \log \bar{\nu}_{\varphi_L,L} \left[\eta_1 = \lfloor (\rho - \rho_c) L \rfloor \right] \,,$$

where the first term on the r.h.s. converges to zero by the local central limit theorem, since $(N - \lfloor (\rho - \rho_c) L \rfloor)/L \to \rho_c$. For the second term,

$$-\frac{1}{L} \log \bar{\nu}_{\varphi_L,L} \left[\eta_1 = \lfloor (\rho - \rho_c) L \rfloor \right]$$
$$= -\frac{1}{L} \log w_L(\lfloor (\rho - \rho_c) L \rfloor) - (\rho - \rho_c) \log \varphi_L + \frac{1}{L} \log z_L(\varphi_L) \to 0 \,,$$

where convergence follows from the sub-exponential assumption (A.1), and since $\varphi_L \to 1$ and $z_L(\varphi_L) \to z(1) \in (0, \infty)$. It follows that $H\left(\pi_{L,N}[\eta_x \in \cdot\,] \mid \bar{\nu}_{\varphi_L,L} \right)$ vanishes and, for the same reason as in the sub-critical case, $\pi_{L,N}[\eta_x \in \cdot\,]$ converges weakly to $w = \bar{\nu}_1$.

Finally, to establish weak convergence of finite dimensional marginals; fix $n_1, \ldots, n_m \in \mathbb{N}_0$ and $x_1, \ldots, x_m$ distinct indices, then

$$\pi_{L,N}[\eta_{x_1} = n_1, \eta_{x_2} = n_2]$$
$$= \frac{w_L(n_1) Z_{L-1,N-n_1}}{Z_{L,N}} \frac{1}{Z_{L-1,N-n_1}} w_L(n_2) \sum_{\xi \in \Omega_{L-2,N-n_1-n_2}} \left(\prod_z w_L(\xi_z) \right)$$
$$= \pi_{L,N}[\eta_{x_1} = n_1] \, \pi_{L-1,N-n_1}[\eta_{x_2} = n_2] \,.$$

This identity immediately generalises to

$$\pi_{L,N}[\eta_{x_1} = n_1, \ldots, \eta_{x_m} = n_m] = \prod_{j=1}^{m} \pi_{L-j+1,N-\sum_{k=1}^{j-1} n_k}[\eta_{x_j} = n_j] \,,$$

which, by taking the thermodynamic limit on the right hand side, completes the proof.

It is possible to drop the assumption that the variance of the limiting weights is positive in the case $w(n) = \mathbb{1}\{n = 0\}$. If we fix $\rho > 0$, we can use the same argument as in the super-critical case above and put all particles on a single site. In this case the right hand side of (A.9) vanishes since

$$-\frac{1}{L} \log \bar{\nu}_{\varphi_L, L}^{\otimes L} \left[\sum_{x=1}^{L} \eta_x = 0 \right] = -\log \nu_{\varphi_L, L} \left[\eta_x = 0 \right] \to \log w(0) = 0 \,,$$

as $L \to \infty$. In this case we do not use the local central limit theorem. Otherwise, the proof remains unchanged. □

References

[1] Inés Armendáriz, Stefan Grosskinsky, and Michail Loulakis, *Zero-range condensation at criticality*, Stochastic Processes and their Applications **123** (2013), no. 9, 3466–3496. MR3071386

[2] Inés Armendáriz and Michail Loulakis, *Thermodynamic limit for the invariant measures in supercritical zero range processes*, Probability Theory and Related Fields **145** (2008), no. 1-2, 175–188. MR2520125

[3] Volker Betz and Daniel Ueltschi, *Spatial Random Permutations and Poisson-Dirichlet Law of Cycle Lengths*, Electronic Journal of Probability **16** (2011), 1173–1192. MR2820074

[4] Leonid V. Bogachev and Dirk Zeindler, *Asymptotic statistics of cycles in surrogate-spatial permutations*, Communications in Mathematical Physics **334** (2014), no. 1, 39–116. MR3304271

[5] Gioia Carinci, Cristian Giardinà, Claudio Giberti, and Frank Redig, *Duality for stochastic models of transport*, Journal of Statistical Physics **152** (2013), no. 4, 657–697. MR3092998

[6] Paul Chleboun and Stefan Grosskinsky, *Condensation in stochastic particle systems with stationary product measures*, Journal of Statistical Physics **154** (2013), no. 1-2, 432–465. MR3162548

[7] Christiane Cocozza-Thivent, *Processus des misanthropes*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete **70** (1985), no. 4, 509–523. MR0807334

[8] Burgess Davis and David McDonald, *An elementary proof of the local central limit theorem*, Journal of Theoretical Probability **8** (1995), no. 3, 693–701. MR1340834

[9] Peter Donnelly and Paul Joyce, *Continuity and weak convergence of ranked and size-biased permutations on the infinite simplex*, Stochastic Processes and their Applications **31** (1989), no. 1, 89–103. MR0996613

[10] Jean-Michel Drouffe, Claude Godrèche, and Federico Camia, *A simple stochastic model for the dynamics of condensation*, Journal of Physics A: Mathematical and General **31** (1998), no. 1, L19–L25. MR1483194

[11] Steinar Engen, *Stochastic abundance models*, Springer Netherlands, 1978. MR0515721

[12] Stewart N. Ethier and Thomas G. Kurtz, *The infinitely-many-neutral-alleles diffusion model*, Advances in Applied Probability **13** (1981), no. 3, 429–452. MR0615945

[13] Martin R. Evans, *Phase transitions in one-dimensional nonequilibrium systems*, Brazilian Journal of Physics **30** (2000), no. 1, 42–57.

[14] Martin R. Evans and Bartek Waclaw, *Condensation in stochastic mass transport models: beyond the zero-range process*, Journal of Physics A: Mathematical and Theoretical **47** (2014), no. 9, 095001.

[15] Lucie Fajfrová, Thierry Gobron, and Ellen Saada, *Invariant measures of mass migration processes*, Electron. J. Probab. **21** (2016), 52 pp. MR3563888

[16] Shui Feng, *The Poisson-Dirichlet distribution and related topics*, Springer Berlin Heidelberg, 2010. MR2663265

[17] Cristian Giardinà, Jorge Kurchan, and Frank Redig, *Duality and exact correlations for a model of heat conduction*, Journal of mathematical physics **48** (2007), no. 3, 033301. MR2314497

[18] Alexander V. Gnedin, *On convergence and extensions of size-biased permutations*, Journal of Applied Probability **35** (1998), no. 3, 642–650. MR1659532

[19] Claude Godrèche, *Dynamics of condensation in zero-range processes*, Journal of Physics A: Mathematical and General **36** (2003), no. 23, 6313–6328. MR1987434

[20] Claude Godrèche, *Condensation for random variables conditioned by the value of their sum*, Journal of Statistical Mechanics: Theory and Experiment **2019** (2019), no. 6, 063207. MR4007056

[21] Christina Goldschmidt, Daniel Ueltschi, and Peter Windridge, *Quantum Heisenberg models and their probabilistic representations*, Contemporary Mathematics (2011), 177–224. MR2868048

[22] Robert M. Gray, *Entropy and information theory*, Springer US, 2011. MR3134681

[23] Robert C. Griffiths, *Lines of descent in the diffusion approximation of neutral Wright-Fisher models*, Theoretical Population Biology **17** (1980), no. 1, 37–50. MR0568666

[24] Robert C. Griffiths, *On the distribution of points in a Poisson Dirichlet process*, Journal of Applied Probability **25** (1988), no. 2, 336–345. MR0938197

[25] Stefan Grosskinsky, Alexander A. Lovisolo, and Daniel Ueltschi, *Lattice permutations and Poisson-Dirichlet distribution of cycle lengths*, Journal of Statistical Physics **146** (2012), no. 6, 1105–1121. MR2903040

[26] Stefan Großkinsky, Gunter M. Schütz, and Herbert Spohn, *Condensation in the zero range process: Stationary and dynamical properties*, Journal of Statistical Physics **113** (2003), no. 3/4, 389–410. MR2013129

[27] Dmitry Ioffe and Bálint Tóth, *Split-and-merge in stationary random stirring on lattice torus*, Journal of Statistical Physics **180** (2020), no. 1-6, 630–653. MR4131005

[28] Watthanan Jatuviriyapornchai, Paul Chleboun, and Stefan Grosskinsky, *Structure of the condensed phase in the inclusion process*, Journal of Statistical Physics **178** (2019), no. 3, 682–710. MR4059956

[29] Intae Jeon, Peter March, and Boris Pittel, *Size of the largest cluster under zero-range invariant measures*, The Annals of Probability **28** (2000), no. 3. MR1797308

[30] Olav Kallenberg, *Foundations of modern probability*, second ed., Probability and its Applications (New York), Springer-Verlag, New York, 2002. MR1876169

[31] John F. C. Kingman, *Random discrete distributions*, Journal of the Royal Statistical Society: Series B (Methodological) **37** (1975), no. 1, 1–15. MR0368264

[32] Florent Krzakala, Andrea Montanari, Federico Ricci-Tersenghi, Guilhem Semerjian, and Lenka Zdeborova, *Gibbs states and the set of solutions of random constraint satisfaction problems*, Proceedings of the National Academy of Sciences **104** (2007), no. 25, 10318–10323. MR2317690

[33] Eddy Mayer-Wolf, Ofer Zeitouni, and Martin Zerner, *Asymptotics of certain coagulation-fragmentation processes and invariant Poisson-Dirichlet measures*, Electronic Journal of Probability **7** (2002), no. 0. MR1902841

[34] John W. McCloskey, *A model for the distribution of individuals by species in an environment*, Ph.D. thesis, Michigan State University, 1965. MR2615013

[35] Patrick A. P. Moran, *Random processes in genetics*, Mathematical Proceedings of the Cambridge Philosophical Society **54** (1958), no. 1, 60–71. MR0127989

[36] Oded Schramm, *Compositions of random transpositions*, Israel Journal of Mathematics **147** (2005), no. 1, 221–243. MR2166362

[37] Yonathan Schwarzkopf, Martin R. Evans, and David Mukamel, *Zero-range processes with multiple condensates: statics and dynamics*, Journal of Physics A: Mathematical and Theoretical **41** (2008), no. 20, 205001. MR2450504

[38] Alasdair G. Thompson, Julien Tailleur, Michael E. Cates, and Richard A. Blythe, *Zero-range processes with saturated condensation: the steady state and dynamics*, Journal of Statistical Mechanics: Theory and Experiment **2010** (2010), no. 02, P02013.

[39] Bartek Waclaw, Julien Sopik, Wolfhard Janke, and Hildegard Meyer-Ortmanns, *Pair-factorized steady states on arbitrary graphs*, Journal of Physics A: Mathematical and Theoretical **42** (2009), no. 31, 315003. MR2521301

[40] Bartlomiej Waclaw and Martin R. Evans, *Explosive condensation in a mass transport model*, Physical Review Letters **108** (2012), no. 7.

[41] Martin P. W. Zerner, Ofer Zeitouni, Eddy Mayer-Wolf, and Persi Diaconis, *The Poisson-Dirichlet law is the unique invariant distribution for uniform split-merge transformations*, The Annals of Probability **32** (2004), no. 1B, 915–938. MR2044670

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4], ProjectEuclid[5]

- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1]LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2]EJMS: Electronic Journal Management System: https://vtex.lt/services/ejms-peer-review/
[3]IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4]BS: Bernoulli Society http://www.bernoulli-society.org/
[5]Project Euclid: https://projecteuclid.org/
[6]IMS Open Access Fund: https://imstat.org/shop/donation/