Size-biased diffusion limits and the inclusion process

Paul Chleboun, Simon Gabriel, Stefan Grosskinsky

Size-biased diffusion limits and the inclusion process

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

Abstract

We consider the inclusion process on the complete graph with vanishing diffusivity, which leads to condensation of particles in the thermodynamic limit. Describing particle configurations in terms of size-biased and appropriately scaled empirical measures of mass distribution, we establish convergence in law of the inclusion process to a measure-valued Markov process on the space of probability measures. In the case where the diffusivity vanishes like the inverse of the system size, the derived scaling limit is equivalent to the well known Poisson-Dirichlet diffusion, offering an alternative viewpoint on these well-established dynamics. Moreover, our novel size-biased approach provides a robust description of the dynamics, which covers all scaling regimes of the system parameters and yields a natural extension of the Poisson-Dirichlet diffusion to infinite mutation rate. We also discuss in detail connections to known results on related Fleming-Viot processes.

1 Introduction and main results

The inclusion process $(\eta^{(L,N)}(t))_{t\geq 0}$ on the complete graph of L sites, describing the interaction of N particles, is characterised by the infinitesimal generator

$$\mathfrak{L}_{L,N} f(\eta) := \sum_{\substack{x,y=1 \\ x\neq y}}^{L} \eta_x(d+\eta_y)[f(\eta^{x,y}) - f(\eta)]. \tag{1.1}$$

Here $\eta \in \Omega_{L,N} := \{\eta \in \mathbb{N}_0^L : \sum_{x=1}^{L} \eta_x = N\}$ and $f : \Omega_{L,N} \mapsto \mathbb{R}$. Moreover, $\eta^{x,y}$ denotes the configuration η where one particle moved from x to y, i.e. $\eta^{x,y} = \eta + e^y - e^x$, with $(e^x)_z := \delta_{x,z}$, provided $\eta_x > 0$. The constant $d \geq 0$ is usually referred to as *diffusivity*.

---

[*]Department of Statistics, University of Warwick, Coventry, CV4 7AL, United Kingdom.
E-mail: paul.i.chleboun@warwick.ac.uk
[†]Institute for Analysis and Numerics, University of Münster, 48149 Münster, Germany.
E-mail: simon.gabriel@uni-muenster.de
[‡]Institute of Mathematics, University of Augsburg, 86135 Augsburg, Germany.
E-mail: Stefan.Grosskinsky@math.uni-augsburg.de

The dynamics of the inclusion process consists of two parts. Each particle performs a continuous-time random walk with rate d, and in addition particles attract each other at unit rate. The process was introduced in [28] as the dual of a model of energy transport and is the natural counterpart of the exclusion process, since interactions are attractive rather than repulsive. Moreover, it can be interpreted in the context of population genetics, describing a population of Moran type [42] where d corresponds to the mutation rate and resampling occurs at rate 1.

The present article concerns the study of the inclusion process in the thermodynamic limit: $N, L \to \infty$ such that $N/L \to \rho \geq 0$, usually abbreviated by simply writing $N/L \to \rho \geq 0$. We rescale the diffusivity $d = d(L)$ with the size of the system such that $d \to 0$ as $L \to \infty$. This leads to clustering of particles into chunks of diverging size [32, 33]. In the context of stochastic particle systems this phenomenon is known as condensation. The occurrence of condensation and the statistics of the condensed phase have been studied for various particle systems (see e.g. [24, 29, 12]), references for the inclusion process and related models include [48, 7, 10, 11] and [5, 38] in the context of metastable dynamics of a single condensate site. These results apply in different scaling regimes with fixed volume and diverging mass which are not addressed in this paper, and we discuss the connections to our results in detail in Section 4. In [35] the condensed phase in the stationary inclusion process was shown to exhibit clusters on scale L with a Poisson-Dirichlet size distribution in the case $dL \to \theta \in \mathbb{R}_+$, whereas for $dL \to \infty$ the clusters are of scale $1/d$ with independent exponential distribution. The aim of this paper is to characterise the dynamics of the mass distribution in the condensed phase in both regimes in terms of diffusion limits.

1.1 Main results

Our main results are twofold. On the one hand, we describe convergence of the inclusion process towards measure-valued diffusions, which are parameterised by $\theta := \lim_{N/L \to \rho} dL \in [0, \infty]$. In the case $\theta < \infty$, the determined scaling limit is equivalent to the well known infinitely-many-neutral-alleles diffusion model [22], also known as Poisson-Dirichlet diffusion. On the other hand, the new description of the infinitely-many-neutral-alleles diffusion model, we give in the following, allows us to construct a natural extension of the process, when $\theta = \infty$.

1.1.1 Scaling limits of the inclusion process

We distinguish between the cases $\theta < \infty$ and $\theta = \infty$. When $\theta < \infty$, we embed particle configurations into $\mathcal{M}_1([0,1])$ using the maps

$$\mu_{L,N}^{(\cdot)} : \Omega_{L,N} \to \mathcal{M}_1([0,1]) \quad \text{of the form} \quad \mu_{L,N}^{(\eta)} := \sum_{x=1}^{L} \frac{\eta_x}{N} \delta_{\frac{\eta_x}{N}} . \tag{1.2}$$

Here $\mathcal{M}_1([0,1])$ denotes the set of probability measures on $[0,1]$, equipped with the topology induced by weak convergence of measures. Note that this corresponds to a size-biased empirical measure on the space of mass fractions $[0,1]$. Not every measure in $\mathcal{M}_1([0,1])$ can be approximated by particle configurations using $\mu_{L,N}^{(\cdot)}$, e.g. every point mass $\alpha\delta_z$ in (1.2) satisfies $\alpha \geq z$. Instead, we restrict ourselves to the closed subspace of atomic measures

$$E := \mu^{(\overline{\nabla})} \subset \mathcal{M}_1([0,1]) , \tag{1.3}$$

defined as the range of the function

$$\mu^{(\cdot)} : \overline{\nabla} \to \mathcal{M}_1([0,1]) \quad \text{with} \quad \mu^{(p)} = (1 - \|p\|_1)\delta_0 + \sum_{i=1}^{\infty} p_i \delta_{p_i} , \tag{1.4}$$

where $\overline{\nabla}$ denotes the Kingman simplex

$$\overline{\nabla} = \{p \in [0,1]^{\mathbb{N}} : p_1 \geq p_2 \geq \ldots, \sum_{i=1}^{\infty} p_i \leq 1\}, \qquad (1.5)$$

equipped with the product topology induced by $[0,1]^{\mathbb{N}}$. We may often drop the subscripts in (1.2) and simply write $\mu^{(\cdot)}$, but the meaning will be clear from the context. The above mappings $\mu^{(\cdot)}$ and $\mu^{(\cdot)}_{L,N}$ do not preserve spatial information of particle configurations, but this also does not enter the dynamics on the complete graph. Note that the map (1.4) was already mentioned in [22], however, only to prove denseness of the domain of functions considered there.

Moreover, any *(unordered)* particle configuration $\eta \in \Omega_{L,N}$ can be mapped onto an element in $\overline{\nabla}$ using the *ranked mass embedding*

$$\tfrac{1}{N}\hat{\eta} := \tfrac{1}{N}(\hat{\eta}_1, \hat{\eta}_2, \ldots, \hat{\eta}_L), \qquad (1.6)$$

with $\hat{\eta}$ representing the entries of η in decreasing order. Another representation of η is by *size-biased sampling*. More precisely, we define

$$\tilde{\eta}_k := \eta_{\sigma(k)}, \qquad (1.7)$$

for some random permutation σ of $\{1, \ldots, L\}$ generated iteratively as follows: First

1. $\sigma(1) = x$ with probability $\frac{\eta_x}{N}$, $x \in \{1, \ldots, L\}$,

and for any following index $k = 2, \ldots, L$

(ii) $\sigma(k) = x$ with probability $\frac{\eta_x}{N - \sum_{j=1}^{k-1} \eta_{\sigma(j)}}$, $x \in \{1, \ldots, L\} \setminus \{\sigma(1), \ldots, \sigma(k-1)\}$.

We refer to [35, Definition 2] for details. The concept of size-biased sampling plays a central role in the present article, since (1.2) can be thought of as the empirical distribution of the first size-biased marginal $\tilde{\eta}_1$.

In order to describe the limiting dynamics, we consider the domain of functions

$$\mathcal{D}(\mathcal{L}_\theta) = \text{sub-algebra of } C(E) \text{ generated by functions } \mu \mapsto \mu(h), \ h \in C^3([0,1]). \quad (1.8)$$

The pre-generator of the corresponding superprocess acting on a function $H(\mu) = \mu(h_1) \cdots \mu(h_n) \in \mathcal{D}(\mathcal{L}_\theta)$ then reads

$$\mathcal{L}_\theta H(\mu) := 2 \sum_{1 \leq k < l \leq n} \left(\mu(Bh_k Bh_l) - \mu(Bh_k)\mu(Bh_l) \right) \prod_{j \neq k,l} \mu(h_j) \qquad (1.9)$$
$$+ \sum_{1 \leq k \leq n} \mu(A_\theta h_k) \prod_{j \neq k} \mu(h_j),$$

with $Bh(z) := h(z) + zh'(z) = (zh(z))'$. Here, the first part is usually referred to as interaction term and A_θ denotes the single-particle operator of the form

$$A_\theta h(z) := (1-z)(Bh)'(z) + \theta(Bh(0) - Bh(z)) \qquad (1.10)$$
$$= z(1-z)h''(z) + (2 - z(2+\theta))h'(z) + \theta(h(0) - h(z)).$$

The operator $Bh(z)$ should be thought of as a 'size-biased derivative', which appears due to our choice of embedding (1.2). For example, a single site containing a mass fraction $z \in [0,1]$ will be represented by a point-mass $z\delta_z$. Thus, change in z will result both in a change of the amount of mass and its position.

Our first result identifies the process described by $\mathcal{L}_\theta$ as the correct scaling limit of the inclusion process. Here and in the following we will sometimes write $\mu_{\#}\eta$ instead of $\mu^{(\eta)}$ in order to avoid overloaded notation, when keeping dependencies on the occurring parameters.

Theorem 1.1. *Let $\rho \in (0, \infty)$ and $d = d(L)$ such that $dL \to \theta \in [0, \infty)$. If $\eta^{(L,N)}(0)$ is such that $\mu_{\#}\eta^{(L,N)}(0) \xrightarrow{D} \mu_0 \in E$, then*

$$\left(\mu_{\#}\eta^{(L,N)}(t)\right)_{t \geq 0} \xrightarrow{D} (\mu_t)_{t \geq 0}, \quad \text{in } D([0, \infty), E), \text{ as } N/L \to \rho. \qquad (1.11)$$

Here $(\mu_t)_{t \geq 0}$ denotes the measure-valued process on E (1.3) generated by $\mathcal{L}_\theta$, cf. (1.9), with initial value μ_0.

Note that the limit process does not depend on the density ρ due to our choice of rescaling, which is discussed in more detail in Section 4. In Proposition 2.4 we prove that the closure of $(\mathcal{L}_\theta, \mathcal{D}(\mathcal{L}_\theta))$ is indeed the generator of a Feller process on the state space E.

For the case $\theta = \infty$, i.e. dL diverging, we expect clusters on the scale $1/d$, cf. [35] and the discussion around (1.31) below. Hence, in this case we consider the embedding

$$\hat{\mu}^{(\cdot)} = \hat{\mu}_{L,N}^{(\cdot)} : \Omega_{L,N} \to \mathcal{M}_1(\mathbb{R}_+) \quad \text{with} \quad \hat{\mu}_{L,N}^{(\eta)} := \sum_{x=1}^{L} \frac{\eta_x}{N} \delta_{dL \frac{\eta_x}{N}}, \qquad (1.12)$$

mapping particle configurations into the space of probability measures $\mathcal{M}_1(\mathbb{R}_+)$, again with the topology induced by weak convergence of measures. In contrast to $\theta < \infty$, any measure in $\mathcal{M}_1(\mathbb{R}_+)$ can be approximated by particle configurations using (1.12), see Lemma A.5, which is why we do not have to restrict ourselves to a strict subset of probability measures as above.

The lack of compactness of $\mathbb{R}_+$ now allows for diverging rescaled masses of particle configurations in the thermodynamic limit, thus when $dL \to \infty$, we expect the scaling limit to be a measure-valued process on $\mathcal{M}_1(\overline{\mathbb{R}}_+)$, with $\overline{\mathbb{R}}_+ = [0, \infty]$. We include ∞ to describe mass on larger scales than $1/d$. Indeed the correct limit turns out to be a process on $\mathcal{M}_1(\overline{\mathbb{R}}_+)$ without interaction and single-particle operator

$$\begin{aligned}
\hat{A}h(z) :=&(Bh)'(z) + (Bh(0) - Bh(z)) \\
=&zh''(z) + (2 - z)h'(z) + (h(0) - h(z)),
\end{aligned} \qquad (1.13)$$

acting on h in the domain

$$\mathcal{D}(\hat{A}) := \{h \,:\, h(\infty) = 0 \text{ and } h|_{\mathbb{R}_+} \in C_c^3(\mathbb{R}_+)\} \cup \{\text{constant functions}\} \subset C(\overline{\mathbb{R}}_+). \quad (1.14)$$

We refer to the proof of Theorem 1.6 for a derivation of $\hat{A}$ from A_θ in terms of a scaling argument when $\theta \to \infty$.

Slowing down the evolution of the inclusion process appropriately, we get the following result.

Theorem 1.2. *Let $\rho \in (0, \infty)$ and $d = d(L) \to 0$ such that $dL \to \infty$. If $\hat{\mu}_{\#}\eta^{(L,N)}(0) \xrightarrow{D} \hat{\mu}_0 \in \mathcal{M}_1(\overline{\mathbb{R}}_+)$, then*

$$\left(\hat{\mu}_{\#}\eta^{(L,N)}\left(\tfrac{t}{dL}\right)\right)_{t \geq 0} \xrightarrow{D} (\hat{\mu}_t)_{t \geq 0}, \quad \text{in } D([0, \infty), \mathcal{M}_1(\overline{\mathbb{R}}_+)), \text{ as } N/L \to \rho. \qquad (1.15)$$

Here $(\hat{\mu}_t)_{t \geq 0}$ denotes the measure-valued process on $\mathcal{M}(\overline{\mathbb{R}}_+)$ with initial value $\hat{\mu}_0$ generated by

$$\hat{\mathcal{L}}H(\mu) = \sum_{1 \leq k \leq n} \mu(\hat{A}h_k) \prod_{\substack{m=1 \\ m \neq k}}^{n} \mu(h_m) \quad \text{with } H(\mu) = \mu(h_1) \cdots \mu(h_n), \; h_k \in \mathcal{D}(\hat{A}). \quad (1.16)$$

The operator $\hat{\mathcal{L}}$ may be interpreted as a Fleming-Viot process without interaction. This is in contrast to the generator $\mathcal{L}_\theta$, which does not have a Fleming-Viot interpretation, since the interaction term is not of the form

$$2 \sum_{1 \leq k < l \leq n} \left(\mu(h_k h_l) - \mu(h_k)\mu(h_l) \right) \prod_{j \neq k, l} \mu(h_j) ,$$

see also Appendix B.

We will see that the limiting dynamics are deterministic with absorbing state $\hat{\mu} = \mathrm{Exp}(1) \in \mathcal{M}_1(\mathbb{R}_+)$. In fact, the statement of Theorem 1.2 can be reformulated into a hydrodynamic limit, cf. Proposition 3.5. Moreover, if $\hat{\mu}_0[\mathbb{R}_+] = 1$, then $\hat{\mu}_t[\mathbb{R}_+] = 1$ for every $t \geq 0$, cf. Corollary 3.7, i.e. mass does not escape to larger scales.

The two theorems fully determine the dynamics of the inclusion process, with vanishing diffusivity, on complete graphs in the thermodynamic limit with density $\rho \in (0, \infty)$. For a discussion of the boundary cases $\rho \in \{0, \infty\}$, we refer to Section 4. For measure valued processes with generators (1.9) and (1.16) the evolution w.r.t. a simple test function (in the appropriate domain) is given by

$$d\mu_t(h) = \mu_t(A_\theta h)\, dt + dM_t^{(h)} \quad \text{and} \quad d\hat{\mu}_t(h) = \hat{\mu}_t(\hat{A}h)\, dt , \qquad (1.17)$$

respectively. Here $t \mapsto M_t^{(h)}$ is a martingale with (predictable) quadratic variation

$$\left\langle M^{(h)} \right\rangle_t = \int_0^t \left(\mathcal{L}_\theta \big(\mu_s(h)^2 \big) - 2\mu_s(h)\mathcal{L}_\theta \mu_s(h) \right) ds = \int_0^t \left(\mu_s\big((Bh)^2 \big) - \mu_s(Bh)^2 \right) ds ,$$

given by the interaction term in $\mathcal{L}_\theta$ with $n = 2$ and $h_1 = h_2 = h$. Since $t \mapsto \hat{\mu}_t(h)$ solves a simple ODE without martingale part it is continuous for all $h \in \mathcal{D}(\hat{A})$, so that the process $(\hat{\mu}_t)_{t \geq 0}$ has continuous paths. Continuity of the process $(\mu_t)_{t \geq 0}$, and thus of the martingale $(M_t^{(h)})_{t \geq 0}$, follows from the equivalence with the Poisson-Dirichlet diffusion, cf. Proposition 1.3.

Taking expectations of the first term in (1.17), we see that $\bar{\mu}_t := \mathbb{E}_{\mu_0}[\mu_t]$ satisfies

$$\frac{d}{dt} \bar{\mu}_t(h) = \bar{\mu}_t(A_\theta h), \quad \text{for all } h \in C^2([0, 1]) ,$$

which agrees with the time evolution of $(\mathbf{E}_{\mu_0}[h(Z(t))])_{t \geq 0}$ for a process $(Z(t))_{t \geq 0}$ on $[0, 1]$ with generator A_θ and initial distribution μ_0. As a consequence, we have the following dualities

$$\mathbb{E}_{\mu_0}[\mu_t(h)] = \mathbf{E}_{\mu_0}[h(Z(t))] \quad \text{and} \quad \hat{\mu}_t(h) = \mathbf{E}_{\mu_0}[h(\hat{Z}(t))] \quad \forall t \geq 0 , \qquad (1.18)$$

where $(\hat{Z}(t))_{t \geq 0}$ is a process on $\overline{\mathbb{R}}_+$ with generator $\hat{A}$. In the latter case $\hat{\mu}_t$ itself is deterministic for fixed initial condition μ_0 since it solves the ODE (1.17). Both processes $(Z(t))_{t \geq 0}$ and $(\hat{Z}(t))_{t \geq 0}$ are one-dimensional diffusions with resetting to 0, which will be used in Sections 2 and 3 to study properties of the measure-valued processes.

1.1.2 A size-biased viewpoint on the Poisson-Dirichlet diffusion

The process described in Theorem 1.1 is a measure-valued process which provides an alternative description of the infinitely-many-neutral-alleles diffusion model introduced by Ethier and Kurtz in their seminal work [22]. Note that the process is also commonly referred to as Poisson-Dirichlet diffusion, which we will use throughout the paper. The

classical Poisson-Dirichlet diffusion is a Feller process on $\overline{\nabla}$ with pre-generator[1]

$$\mathcal{G}_\theta f = \sum_{i,j=1}^\infty p_i(\delta_{i,j} - p_j)\partial^2_{p_i p_j} f - \theta \sum_{i=1}^\infty p_i \partial_{p_i} f \,, \tag{1.19}$$

acting on functions in the domain

$$\mathcal{D}_{mon}(\mathcal{G}_\theta) := \text{sub-algebra of } C(\overline{\nabla}) \text{ generated by } 1, \varphi_2, \varphi_3, \dots, \tag{1.20}$$

where $\varphi_m(p) := \sum_{i=1}^\infty p_i^m$ for $m \geq 2$. $\mathcal{G}_\theta$ acts on such test functions with the convention that occurring sums on the r.h.s. of (1.19) are evaluated on $\nabla := \left\{ p \in \overline{\nabla} : \sum_{i=1}^\infty p_i = 1 \right\} \subset \overline{\nabla}$ and extended to $\overline{\nabla}$ by continuity. We stress that $\varphi_1(p) := \sum_{i=1}^\infty p_i$ is not a continuous function on $\overline{\nabla}$. Our size-biased approach circumvents such technical issues, which is one of its main advantages.

The name Poisson-Dirichlet diffusion is adequate, since its unique invariant distribution is the *Poisson-Dirichlet distribution* PD(θ). The Poisson-Dirichlet distribution is a one-parameter family of probability measures supported on ∇. It was first introduced by Kingman [39] as a natural limit of Dirichlet distributions. However, there is a more intuitive construction of the Poisson-Dirichlet distribution using a stick-breaking procedure, see for example [25]. Later, the distribution was identified as the unique stationary measure of the split-merge dynamics [49, 46] and the Poisson-Dirichlet diffusion [22]. Despite it being introduced in the field of population genetics, the Poisson-Dirichlet distribution has since then also appeared in statistical mechanics [30, 4, 34] and recently in interacting particle systems [35, 11].

Naturally, one can consider the mapping of the Poisson-Dirichlet diffusion under the isomorphism $\mu^{(\cdot)}$, cf. Lemma A.1, which yields a process on $E \subset \mathcal{M}_1([0,1])$. Indeed, this push-forward process agrees with the process generated by $\mathcal{L}_\theta$. The proof of this fact can be found in the appendix.

Proposition 1.3. *Let $(\mu_t)_{t \geq 0}$ be the measure-valued process generated by $\mathcal{L}_\theta$ (1.9) with initial data $\mu_0 \in E$, and $(X(t))_{t \geq 0}$ be the Poisson-Dirichlet diffusion generated by $\mathcal{G}_\theta$ (1.19) with $\mu^{(X(0))} = \mu_0$, then*

$$(\mu_t)_{t \geq 0} \stackrel{D}{=} \left(\mu^{(X(t))} \right)_{t \geq 0} . \tag{1.21}$$

In particular, the following properties translate immediately from $(X(t))_{t \geq 0}$ to $(\mu_t)_{t \geq 0}$:

1. *The process $(\mu_t)_{t \geq 0}$ has a unique stationary distribution, which is reversible. It is given by $\mathbf{P} = \mu_\# \mathrm{PD}(\theta)$, i.e. the law of*

$$\mu^{(X)} = \sum_{i=1}^\infty X_i \delta_{X_i} , \quad X \sim \mathrm{PD}(\theta) . \tag{1.22}$$

2. *The process $(\mu_t)_{t \geq 0}$ has continuous sample paths in E.*

3. *For any initial value $\mu_0 \in E$, we have*

$$\mathbb{P}(\mu_t(\{0\}) = 0 \,\forall t > 0) = 1 . \tag{1.23}$$

Together with Theorem 1.1, this yields the following corollary.

---

[1]The original formulation of the pre-generator in [22] includes a multiplicative factor of $\frac{1}{2}$ which we omitted here.

Corollary 1.4. *Let $\rho \in (0, \infty)$ and $d = d(L)$ such that $dL \to \theta \in [0, \infty)$. If $\eta^{(L,N)}(0)$ is such that $\frac{1}{N}\hat{\eta}^{(L,N)}(0) \xrightarrow{D} X(0) \in \overline{\nabla}$, then*

$$\frac{1}{N}\left(\hat{\eta}^{(L,N)}(t)\right)_{t \geq 0} \xrightarrow{D} (X(t))_{t \geq 0}, \quad \text{as } N/L \to \rho. \tag{1.24}$$

Here $(X(t))_{t \geq 0}$ denotes the Poisson-Dirichlet diffusion on $\overline{\nabla}$ with parameter θ and initial value $X(0)$, generated by $\mathcal{G}_\theta$, cf. (1.19), and $\hat{\eta}$ denotes the ordered particle configuration.

Remark 1.5. Proposition 1.3 yields a way to recover the measure–valued process $(\mu_t)_{t \geqslant 0}$ from the classical Poisson-Dirichlet diffusion $(X_t)_{t \geqslant 0}$. Since the map (1.4) is an isomorphism, see Lemma A.1, and has a continuous inverse $(\mu^{(\cdot)})^{-1} : E \to \overline{\nabla}$, the proposition also implies $(X_t)_{t \geqslant 0} \overset{D}{=} \left((\mu^{(\cdot)})^{-1}(\mu_t)\right)_{t \geqslant 0}$. The inverse map reconstructs elements in the Kingman simplex by counting and ordering mass occurrences, for example,

$$E \ni \tfrac{1}{2}\delta_{\frac{1}{4}} + \tfrac{3}{8}\delta_{\frac{1}{8}} + \tfrac{1}{8}\delta_0 \mapsto \left(\tfrac{1}{4}, \tfrac{1}{4}, \tfrac{1}{8}, \tfrac{1}{8}, \tfrac{1}{8}, 0, \dots\right) \in \overline{\nabla}.$$

In this article, we focus on a (joint) thermodynamic limit as $N, L \to \infty$ with $N/L \to \rho$, but our derived scaling limits do not actually depend on the density ρ. Therefore our approach also extends to different scaling regimes as is discussed in Section 4. The measure valued process generated by $\hat{\mathcal{L}}$ is the natural extension of the process $\mathcal{L}_\theta$ (and thus to the Poisson-Dirichlet diffusion generated by $\mathcal{G}_\theta$) when $\theta \to \infty$. A first indication for this relationship can already be observed on the level of stationary distributions. From Proposition 1.3(i), we recall that the stationary distribution w.r.t. $\mathcal{L}_\theta$ is given by the size-biased sample of $\mathrm{PD}(\theta)$. Consider $X^{(\theta)} \sim \mathrm{PD}(\theta)$ and sample an index $I \in \mathbb{N}$ such that

$$I = i \quad \text{with probability } X_i^{(\theta)}, \tag{1.25}$$

i.e. we pick the index I with size-bias. It is well known [25, Theorem 2.7] that $X_I^{(\theta)} \sim \mathrm{Beta}(1, \theta)$. Moreover, $\mathrm{Exp}(1)$ is the absorbing state of the deterministic dynamics induced by $\hat{\mathcal{L}}$, cf. Corollary 3.9. Now, the following connection between a Beta and an Exponential distribution holds:

$$\theta \, \mathrm{Beta}(1, \theta) \xrightarrow{D} \mathrm{Exp}(1), \quad \text{as } \theta \to \infty. \tag{1.26}$$

Hence, as $\theta \to \infty$, the rescaled size-biased sample $\theta \, X_I^{(\theta)}$ converges weakly to an $\mathrm{Exp}(1)$ random variable.

This relationship can also be made sense of on the level of processes, summarised in the following diagram:

$$
\begin{array}{ccc}
& \left(\mathcal{L}_\theta, E\right) \Longleftrightarrow \left(\mathcal{G}_\theta, \overline{\nabla}\right) & \quad \text{, if } \theta < \infty \\
& \nearrow^{\delta_{\frac{1}{N}\eta_x}} \quad \Big\downarrow {\scriptstyle\theta \to \infty \text{ when}} \begin{cases} z \mapsto \theta z, \\ t \mapsto t/\theta. \end{cases} & \\
\left(\mathfrak{L}_{L,N}, \Omega_{L,N}\right) \xrightarrow{N/L \to \rho} \Big\{ & & \\
& \searrow_{\delta_{\frac{dL}{N}\eta_x}} \quad \Big\downarrow & \\
& \left(\hat{\mathcal{L}}, \mathcal{M}_1(\overline{\mathbb{R}}_+)\right) & \quad \text{, if } \theta = \infty.
\end{array}
$$

We analyse the inclusion process $\left(\mathfrak{L}_{L,N}, \Omega_{L,N}\right)$ and consider the two cases $\theta < \infty$ (Theorem 1.1) and $\theta = \infty$ (Theorem 1.2), with appropriate embeddings of configurations in the space of probability measures. In the case $\theta < \infty$ the limiting process is equivalent to the Poisson-Dirichlet diffusion generated by $\mathcal{G}_\theta$ (Proposition 1.3). Furthermore, our

size-biased approach allows for a meaningful limit when $\theta = \infty$, identifying a natural extension for models with Poisson-Dirichlet diffusion limit, under appropriate rescaling of time and space. For this matter, we introduce the scaling operator $S_\theta : E \to \mathcal{M}_1(\mathbb{R}_+)$, which linearly scales measures on the unit interval to measures on the interval $[0, \theta]$, i.e.

$$S_\theta : \mu(dz) \mapsto \mu(d\tfrac{z}{\theta}) \,. \tag{1.27}$$

Theorem 1.6. *Let $(\mu_t^\theta)_{t\geq 0}$ be the process generated by $\mathcal{L}_\theta$ and $(\hat{\mu}_t)_{t\geq 0}$ be the process generated by $\hat{\mathcal{L}}$. If $S_\theta \mu_0^\theta \xrightarrow{D} \hat{\mu}_0 \in \mathcal{M}_1(\overline{\mathbb{R}}_+)$, then*

$$\left(S_\theta \mu_{t/\theta}^\theta \right)_{t\geq 0} \xrightarrow{D} (\hat{\mu}_t)_{t\geq 0} \,, \quad \text{in } C([0, \infty), \mathcal{M}_1(\overline{\mathbb{R}}_+)) \,, \quad \text{as } \theta \to \infty \,, \tag{1.28}$$

where we consider the topology induced by weak convergence on $\mathcal{M}_1(\overline{\mathbb{R}}_+)$.

Remark 1.7. In the above theorem we saw that scaling of space (of order θ) is necessary to observe a meaningful limit as $\theta \to \infty$. Similarly, one could scale the Kingman simplex to $\theta\overline{\nabla}$. However, in the limit we lack the property of distinguishing between separate scales. Consider for example $\theta = n^2$, $n \in \mathbb{N}$, and the sequence

$$p^{(\theta)} = \tfrac{1}{2} \big(\underbrace{\tfrac{1}{\sqrt{\theta}}, \ldots, \tfrac{1}{\sqrt{\theta}}}_{\sqrt{\theta} \text{ times}}, \underbrace{\tfrac{1}{\theta}, \ldots, \tfrac{1}{\theta}}_{\theta \text{ times}}, 0, \ldots \big) \in \nabla \,. \tag{1.29}$$

Then $\theta \, p_i^{(\theta)} \to \infty$ for every $i \in \mathbb{N}$. On the other hand,

$$S_\theta \mu^{(p^\theta)} = \tfrac{1}{2}\delta_{\frac{1}{2}} + \tfrac{1}{2}\delta_{\frac{1}{2}\sqrt{\theta}} \xrightarrow{D} \tfrac{1}{2}\delta_{\frac{1}{2}} + \tfrac{1}{2}\delta_\infty \in \mathcal{M}_1(\overline{\mathbb{R}}_+) \,, \tag{1.30}$$

which captures both the amount of diverging mass and information on scales $\frac{1}{\theta}$. This highlights the fact that considering the space E, instead of $\overline{\nabla}$, is essential for a detailed analysis of the Poisson-Dirichlet diffusion in the boundary case $\theta \to \infty$. In particular, this highlights that the gap in the schematic diagram above (at the bottom right) is not expected to be filled, by a naive scaling of $\overline{\nabla}$.

1.2 Comparison to the literature

1.2.1 Condensation and the inclusion process

After its introduction [28], the inclusion process has been subject to study as an interesting model of mass transport on its own [8]. In particular, in the context of condensation in stochastic particle systems it is a model of major interest. In short, a particle system exhibits condensation if a positive fraction of particles concentrates on sites with a vanishing volume fraction. Such sites with diverging occupation (and their occupying particles) are called the *condensate*, the remaining particles and sites are said to be the *background* or *bulk* of the system. In [35], the existence of a non-trivial condensate was proven; we refer to the same reference for an exact definition of the condensation phenomenon. While the dynamics of the bulk for occupation numbers of order one is covered by general results on the propagation of chaos for particle systems [31], the scope of the present article is to determine the dynamics of the mass distribution in the condensate. However, we want to stress that our analysis does not rely on previous results on condensation. In fact, the clustering of particles on diverging scales is an implicit consequence of the scaling limits, presented in Theorems 1.1 and 1.2.

For the inclusion process, the condensation phenomenon was first studied in [32], however for spatially inhomogeneous systems on a finite lattice with diverging number of particles. In homogeneous systems, condensation is a consequence of increasing particle interactions relative to diffusion as $d \to 0$. The finite lattice case has been subject to

further studies in [33, 5, 38, 36]. When considering a thermodynamic limit, i.e. diverging lattice size L and number of particles N with finite limiting density ρ, condensation was studied heuristically in [7]. They considered a one-dimensional periodic lattice with totally asymmetric dynamics and vanishing diffusion rate s.t. $\theta = 0$. A modified model with stronger particle interactions, leading to instantaneous condensation even in one spatial dimension has been considered in [48, 10].

On a rigorous level, the thermodynamic limit of stationary distributions has been treated in [35]. Under the assumption that $d \to 0$ as $L \to \infty$, we have the following cases:

- if $dL \to 0$, then the condensate is given by a single cluster, and if in addition $dL \log L \to 0$ this cluster is holding all the particles;
- if $dL \to \theta \in (0, \infty)$, then the condensate concentrates on macroscopic scales and is distributed according to a Poisson-Dirichlet distribution $\text{PD}(\theta)$;
- on the other hand if $dL \to \infty$, the condensate is located on mesoscopic scales and the clusters are independent, more precisely,

$$d(\tilde{\eta}_1, \ldots, \tilde{\eta}_n) \xrightarrow{D} \text{Exp}(\tfrac{1}{\rho})^{\otimes n} \,. \tag{1.31}$$

Here $\tilde{\eta}$ denotes a size-biased sample w.r.t. $\eta \in \Omega_{L,N}$, introduced in (1.7).

The above result holds for any irreducible and spatially homogeneous dynamics on diverging finite graphs, where the inclusion process has stationary product measures. Because condensation in homogeneous systems only occurs if $d \to 0$, [35] characterised the stationary condensates for the inclusion process, and in particular, the results hold for the complete graph dynamics we consider here. It was proven in [11] that perturbations of the transition rates (in the case $\theta \leq 1$) still give rise to a Poisson-Dirichlet distributed condensate.

Both Theorem 1.1 and Theorem 1.2 complete the picture of condensation behaviour of the inclusion process on complete graphs outside of stationarity. Moreover, our results link the inclusion process dynamics directly with the Poisson-Dirichlet diffusion, when $\theta < \infty$, which allows for enhanced understanding of the latter dynamics.

1.2.2 The Poisson–Dirichlet diffusion

In their work [22], Ethier and Kurtz proved that the closure of $(\mathcal{G}_\theta, \mathcal{D}_{mon}(\mathcal{G}_\theta))$, cf. (1.19), gives rise to a generator of a diffusion process on $\overline{\nabla}$, which is the natural scaling limit of a finite-dimensional Wright-Fisher model when sending the number of individuals and types to infinity separately. The restriction to test functions in $\mathcal{D}_{mon}(\mathcal{G}_\theta)$ turns out to be convenient, but makes it difficult to understand the precise dynamics of the infinite dimensional process. In [22], also an enlarged domain of test-functions of the form

$$p \mapsto \sum_{i=1}^{\infty} h(p_i)\,, \quad h \in C^2([0,1]) \text{ with } h(0) = h'(0) = 0\,, \tag{1.32}$$

was considered. However, this does not improve the understanding of the dynamics on an intuitive level, which is particularly due to the fact that 'sums are evaluated on ∇ and extended to $\overline{\nabla}$ by continuity'. In this paper, we instead propose to consider functions of the form

$$p \mapsto h(0) + \sum_{i=1}^{\infty} p_i (h(p_i) - h(0))\,, \quad h \in C^2([0,1])\,. \tag{1.33}$$

For a fixed $p \in \overline{\nabla}$, the r.h.s. is the expectation w.r.t. the probability measure $\mu^{(p)}$, recall (1.4). Moreover, note that the functions φ_m are of the form (1.33) with $h(p) = p^{m-1}$.

The usual approach in the literature, when constructing the Poisson-Dirichlet diffusion, is to take the large L-limit of an L-dimensional diffusion model. Alternatively, discrete models have been considered but then first convergence to the L-dimensional diffusion model, when $N \to \infty$, is proven. See for example [22, 14, 44]. To the authors best knowledge, the present article is the first to consider a thermodynamic limit, which is taking both size of the system and number of particles to infinity at the same time while keeping the density approximately constant. This makes sense both from a physical and population genetics perspective. In particular, taking a joint limit allows for interesting dynamics in the case $\theta = \infty$ which could otherwise not be considered, recall Theorem 1.6.

The Poisson-Dirichlet diffusion was treated previously as a measure-valued process in $\mathcal{M}_1([0,1])$ in [18, 19], where it was considered as a Fleming-Viot process with mutation operator

$$A_{FV} h(u) = \theta \int_0^1 [h(v) - h(u)] \, dv \,. \tag{1.34}$$

Here the elements in $[0,1]$ are interpreted as types, and uniform jumps at rate θ in the mutation operator correspond to mutation events. These dynamics can be derived in the thermodynamic limit from the inclusion process on a complete graph with $dL \to \theta \in (0, \infty)$, using the embedding

$$\eta \in \Omega_{L,N} \mapsto \sum_{x=1}^{L} \frac{\eta_x}{N} \delta_{\frac{x}{L}} \in \mathcal{M}([0,1]) \,. \tag{1.35}$$

This describes the spatial distribution of mass on the rescaled lattice and is different from the approach in the present paper, where we ignore spatial information and only keep track of the mass distribution, cf. (1.4). On the other hand, our approach is more robust and allows for an extended analysis of the model for $\theta = \infty$. The embedding (1.35) has been considered in [13] for the inclusion process on a complete graph of fixed size. They study convergence to equilibrium in the long-time limit and with diverging mass $N \to \infty$. The derivation of a Fleming-Viot process with mutation (1.34) in the thermodynamic limit is relatively straightforward if the inclusion process is formulated in terms of particle positions, which is presented briefly in Appendix B.

Applying our approach to other geometries may be possible for dense random graphs along the lines of [6], which have diverging degrees leading to a self-averaging effect similar to the complete graph. In general, spatial models are difficult to treat since the inclusion process after the embedding (1.4) is not Markovian. Consider for example nearest-neighbour dynamics on a regular lattice, then it is known that the random-walk and the inclusion part of the dynamics have two different time scales, see [3], and more sophisticated methods are necessary to treat this case.

1.3 Outline of the paper

In Section 2 we show that $\mathcal{L}_\theta$ generates a Feller process and prove Theorem 1.1. We make use of explicit approximations of the inclusion process generator and the Trotter-Kurtz approximation theorem. Moreover, we prove the equivalence of the Poisson-Dirichlet diffusion and our scaling limit in Section 2.2. Lastly, we discuss the advantages of considering size-biased dynamics in Section 2.3. In Section 3 we determine the scaling limit when $\theta = \infty$, following a similar approach as in the case $\theta < \infty$. We finish the section by proving the convergence $\frac{1}{\theta} \mathcal{L}_\theta \to \hat{\mathcal{L}}$ stated in Theorem 1.6. Lastly, we discuss boundary cases $\rho \in \{0, \infty\}$, fluctuations and open problems in Section 4.

2 Scaling limits in the case $dL \to \theta < \infty$

2.1 The measure-valued process

In this section, we will prove that the measure-valued process generated by $\mathcal{L}_\theta$ (1.9) is a Feller process on the state space E (1.3). Furthermore, we deduce weak convergence on the path space for the inclusion process configurations embedded in the space of probability measures on the unit interval.

2.1.1 Approximation of infinitesimal dynamics

The key result of this section is the following convergence result on the level of pre-generators

Proposition 2.1. *Let $\rho \in (0,\infty)$ and $d = d(L)$ such that $dL \to \theta \in [0,\infty)$. For every $H \in \mathcal{D}(\mathcal{L}_\theta)$, cf. (1.8), we have with $\mathfrak{L}_{L,N}$ defined in (1.1)*

$$\lim_{N/L \to \rho} \sup_{\eta \in \Omega_{L,N}} \left| \mathfrak{L}_{L,N} H(\mu^{(\cdot)})(\eta) - (\mathcal{L}_\theta H)(\mu^{(\eta)}) \right| = 0 \,. \tag{2.1}$$

We split the proof of Proposition 2.1 into two parts. First, we only consider test functions of elementary form $H(\mu) = \mu(h)$, which corresponds to measuring a single observable $h \in C^3([0,1])$. We then extend the convergence result to arbitrary test functions in the domain, which requires to understand correlations between several observables. As usual, it turns out that only pairwise correlations contribute to leading order.

Lemma 2.2. *Let $\rho \in (0,\infty)$ and $d = d(L)$ such that $dL \to \theta \in [0,\infty)$. Consider $H \in \mathcal{D}(\mathcal{L}_\theta)$ of the elementary form $H(\mu) = \mu(h)$, for some $h \in C^3([0,1])$. Then*

$$\lim_{N/L \to \rho} \sup_{\eta \in \Omega_{L,N}} \left| \mathfrak{L}_{L,N} H(\mu^{(\cdot)})(\eta) - \mu^{(\eta)}(A_\theta h) \right| = 0 \,, \tag{2.2}$$

where A_θ is the single-particle generator, introduced in (1.10).

Proof. Let $h \in C^3([0,1])$ and define $H(\mu) := \mu(h)$. For the sake of convenience we introduce the notation $\tilde{h}(z) := z\,h(z)$. Thus,

$$H(\mu^{(\eta)}) = H(\mu_\# \eta) = \mu^{(\eta)}(h) = \sum_{x=1}^{L} \frac{\eta_x}{N} h\left(\tfrac{\eta_x}{N}\right) = \sum_{x=1}^{L} \tilde{h}\left(\tfrac{\eta_x}{N}\right) \,, \tag{2.3}$$

which allows us to write

$$\begin{aligned}
H(\mu_\# \eta^{x,y}) - H(\mu_\# \eta) &= \tilde{h}\left(\tfrac{\eta_y+1}{N}\right) - \tilde{h}\left(\tfrac{\eta_y}{N}\right) + \tilde{h}\left(\tfrac{\eta_x-1}{N}\right) - \tilde{h}\left(\tfrac{\eta_x}{N}\right) \\
&= \frac{1}{N}\tilde{h}'\left(\tfrac{\eta_y}{N}\right) + \frac{1}{2}\frac{1}{N^2}\tilde{h}''\left(\tfrac{\eta_y}{N}\right) - \frac{1}{N}\tilde{h}'\left(\tfrac{\eta_x}{N}\right) \\
&\quad + \frac{1}{2}\frac{1}{N^2}\tilde{h}''\left(\tfrac{\eta_x}{N}\right) + \frac{1}{6}\frac{1}{N^3}\tilde{h}'''(\xi) \,,
\end{aligned} \tag{2.4}$$

using a second-order Taylor approximation of $\tilde{h}$, with $\xi \in [0,1]$. Therefore, we have uniformly over configurations η

$$\begin{aligned}
\mathfrak{L}_{L,N} H(\mu^{(\cdot)})(\eta) = \sum_{\substack{x,y=1 \\ x \neq y}}^{L} \eta_x(d + \eta_y) &\left[\frac{1}{N}\tilde{h}'\left(\tfrac{\eta_y}{N}\right) + \frac{1}{2}\frac{1}{N^2}\tilde{h}''\left(\tfrac{\eta_y}{N}\right) \right. \\
&\left. - \frac{1}{N}\tilde{h}'\left(\tfrac{\eta_x}{N}\right) + \frac{1}{2}\frac{1}{N^2}\tilde{h}''\left(\tfrac{\eta_x}{N}\right) \right] + o(1) \,,
\end{aligned}$$

where $o(1)$ denotes a (uniformly in $\Omega_{L,N}$) vanishing quantity as $N/L \to \rho$. We split the sum into two parts, by analysing terms with coefficients $d\,\eta_x$ and $\eta_x\eta_y$ separately. We begin with the latter:

- The contribution of inclusion rates $\eta_x\eta_y$ is limited to

$$\sum_{\substack{x,y=1 \\ x\neq y}}^{L} \frac{\eta_x}{N}\frac{\eta_y}{N}\tilde{h}''(\tfrac{\eta_x}{N}) = \sum_{x=1}^{L} \frac{\eta_x}{N}\Big(1 - \frac{\eta_x}{N}\Big)\tilde{h}''(\tfrac{\eta_x}{N})\,, \tag{2.5}$$

  due to exact cancellation of the first-order terms $\tilde{h}'$.

- On the other hand, contributions of the random-walk dynamics induced by rates $d\,\eta_x$ are given by

$$d\sum_{\substack{x,y=1 \\ x\neq y}}^{L} \frac{\eta_x}{N}\left[\widetilde{h}'(\tfrac{\eta_y}{N}) + \frac{1}{2}\frac{1}{N}\widetilde{h}''(\tfrac{\eta_y}{N}) - \widetilde{h}'(\tfrac{\eta_x}{N}) + \frac{1}{2}\frac{1}{N}\widetilde{h}''(\tfrac{\eta_x}{N})\right]$$

$$= dL\sum_{x=1}^{L} \frac{\eta_x}{N}\left[\frac{1}{L}\sum_{y\neq x}\widetilde{h}'(\tfrac{\eta_y}{N}) - \frac{L-1}{L}\widetilde{h}'(\tfrac{\eta_x}{N})\right] + o(1)\,, \tag{2.6}$$

  because second-order terms $\tilde{h}''$ vanish in the thermodynamic limit due to

$$\left|d\sum_{\substack{x,y=1 \\ x\neq y}}^{L} \frac{\eta_x}{2N^2}\big(\tilde{h}''(\tfrac{\eta_x}{N}) + \tilde{h}''(\tfrac{\eta_y}{N})\big)\right| \leq dL\sum_{x=1}^{L}\frac{\eta_x}{N^2}\|\tilde{h}''\|_\infty \leq \frac{dL}{N}\|\tilde{h}''\|_\infty \to 0\,. \tag{2.7}$$

Furthermore, we can absorb errors arising from replacing $\frac{L-1}{L}\tilde{h}'$ with $\tilde{h}'$, into $o(1)$.

Now, combining (2.5) and (2.6) yields

$$\mathfrak{L}_{L,N}H(\mu^{(\cdot)})(\eta) = \sum_{x=1}^{L}\frac{\eta_x}{N}\Big(1 - \frac{\eta_x}{N}\Big)\tilde{h}''(\tfrac{\eta_x}{N}) + dL\left[\tilde{h}'(0) - \sum_{x=1}^{L}\frac{\eta_x}{N}\tilde{h}'(\tfrac{\eta_x}{N})\right] + o(1)\,, \tag{2.8}$$

where we additionally used Lemma A.2 to conclude the uniform approximation $\frac{1}{L}\sum_{y=1,\,y\neq x}^{L}\tilde{h}'(\tfrac{\eta_y}{N}) = \tilde{h}'(0) + o(1)$. Rewriting (2.8) in terms of $\mu^{(\eta)}$, we have

$$\mathfrak{L}_{L,N}H(\mu^{(\cdot)})(\eta) = \mu^{(\eta)}\big(z(1-z)h''(z)$$
$$+ 2(1-z)h'(z) + dL(h(0) - h(z) - zh'(z))\big) + o(1)\,, \tag{2.9}$$

where we used that $\tilde{h}'(z) = h(z) + zh'(z) = Bh(z)$ and $\tilde{h}''(z) = 2h'(z) + zh''(z) = (Bh)'(z)$. Lastly, since $\|Bh\|_\infty < \infty$ and $dL \to \theta$, we indeed have

$$\mathfrak{L}_{L,N}H(\mu^{(\cdot)})(\eta) = \mu^{(\eta)}(A_\theta h) + o(1)\,, \tag{2.10}$$

uniformly over all $\eta \in \Omega_{L,N}$. This concludes the proof. $\qquad\square$

Remark 2.3. Note that the equivalence to the Poisson-Dirichlet diffusion can already be observed in (2.8) when considering h to be of the form $h(z) = z^{m-1}$, $m \geq 2$. In this case $H(\mu^{(\eta)}) = \mu^{(\eta)}(h) = \sum_{x=1}^{L}\tilde{h}(\tfrac{\eta_x}{N}) = \varphi_m(\tfrac{\eta}{N})$, cf. (1.20), and

$$\mathfrak{L}_{L,N}H(\mu^{(\cdot)})(\eta) \simeq \mathcal{G}_\theta\varphi_m(\tfrac{\hat{\eta}}{N})\,. \tag{2.11}$$

After having proved the statement of Proposition 2.1 for specific functions, we can now proceed with the proof of the full statement.

Proof of Proposition 2.1. Let $H \in \mathcal{D}(\mathcal{L}_\theta)$. Without loss of generality we may assume H has the form

$$H(\mu) = \mu(h_1) \cdots \mu(h_n), \quad h_k \in C^3([0,1]), \quad 1 \le k \le n, \tag{2.12}$$

since linear combinations of such functions can be treated by linearity of the operators and the triangle inequality. Thus, considering $\eta \in \Omega_{L,N}$ and the configuration after one particle jumped from x to y, we have

$$
\begin{aligned}
H\left(\mu_\# \eta^{x,y}\right) &= \prod_{k=1}^n \left(\mu_\# \eta^{x,y}\right)(h_k) \\
&= \prod_{k=1}^n \left[\widetilde{h}_k(\tfrac{\eta_y+1}{N}) - \widetilde{h}_k(\tfrac{\eta_y}{N}) + \widetilde{h}_k(\tfrac{\eta_x-1}{N}) - \widetilde{h}_k(\tfrac{\eta_x}{N}) + \mu^{(\eta)}(h_k) \right].
\end{aligned}
$$

Now, expanding the product yields

$$
\begin{aligned}
H\left(\mu_\# \eta^{x,y}\right) &= H\left(\mu_\# \eta\right) + \sum_{k=1}^n \left[\widetilde{h}_k(\tfrac{\eta_y+1}{N}) - \widetilde{h}_k(\tfrac{\eta_y}{N}) + \widetilde{h}_k(\tfrac{\eta_x-1}{N}) - \widetilde{h}_k(\tfrac{\eta_x}{N}) \right] \prod_{\substack{l=1 \\ l \ne k}}^n \mu^{(\eta)}(h_l) \\
&\quad + \sum_{1 \le k < l \le n} \left[\widetilde{h}_k(\tfrac{\eta_y+1}{N}) - \widetilde{h}_k(\tfrac{\eta_y}{N}) + \widetilde{h}_k(\tfrac{\eta_x-1}{N}) - \widetilde{h}_k(\tfrac{\eta_x}{N}) \right] \\
&\quad \times \left[\widetilde{h}_l(\tfrac{\eta_y+1}{N}) - \widetilde{h}_l(\tfrac{\eta_y}{N}) + \widetilde{h}_l(\tfrac{\eta_x-1}{N}) - \widetilde{h}_l(\tfrac{\eta_x}{N}) \right] \prod_{\substack{j=1 \\ j \ne k,l}}^n \mu^{(\eta)}(h_j) + r(\eta),
\end{aligned} \tag{2.13}
$$

with r denoting the remainder. This expansion allows us to split

$$\mathfrak{L}_{L,N} H(\mu^{(\cdot)})(\eta) = \sum_{\substack{x,y=1 \\ x \ne y}}^L \eta_x (d + \eta_y) \left[H(\mu_\# \eta^{x,y}) - H(\mu_\# \eta) \right] \tag{2.14}$$

into three parts:

- First, we make use of Lemma 2.2 which yields

$$
\begin{aligned}
\sum_{\substack{x,y=1 \\ x \ne y}}^L &\eta_x(d+\eta_y) \sum_{k=1}^n [\widetilde{h}_k(\tfrac{\eta_y+1}{N}) - \widetilde{h}_k(\tfrac{\eta_y}{N}) + \widetilde{h}_k(\tfrac{\eta_x-1}{N}) - \widetilde{h}_k(\tfrac{\eta_x}{N})] \prod_{\substack{l=1 \\ l \ne k}}^n \mu^{(\eta)}(h_l) \\
&= \sum_{k=1}^n \mathfrak{L}_{L,N}\left(\mu^{(\cdot)}(h_k)\right)(\eta) \prod_{\substack{l=1 \\ l \ne k}}^n \mu^{(\eta)}(h_l) = \sum_{k=1}^n \mu^{(\eta)}(A_\theta h_k) \prod_{\substack{l=1 \\ l \ne k}}^n \mu^{(\eta)}(h_l) + o(1).
\end{aligned}
$$

- Next, we prove that the remainder r has no contribution. More precisely, for any non-negative sequence a_N satisfying $N^2 a_N \to 0$, i.e. a_N lies in $o(\frac{1}{N^2})$, we have

$$a_N \sum_{\substack{x,y=1 \\ x \ne y}}^L \eta_x(d + \eta_y) \le a_N (dL + N)N \to 0. \tag{2.15}$$

This includes, in particular, the remainder $r(\eta)$ because each summand lies in $o(\frac{1}{N^3})$, recall that each square bracket in (2.13) vanishes uniformly like N^{-1}, cf. (2.4).

- Lastly, we derive the interaction part where two observables are affected by the transition of a particle. Again, we perform a Taylor approximation for each of the two square brackets appearing in (2.13). Due to (2.15), together with (2.4), it suffices to consider only products of first-order terms $\tilde{h}'$. Therefore, we are left with

$$
\frac{1}{N^2} \sum_{1 \le k < l \le n} \sum_{\substack{x,y=1 \\ x \ne y}}^{L} \eta_x(d + \eta_y) \left[\tilde{h}'_k(\tfrac{\eta_y}{N}) - \tilde{h}'_k(\tfrac{\eta_x}{N}) \right] \left[\tilde{h}'_l(\tfrac{\eta_y}{N}) - \tilde{h}'_l(\tfrac{\eta_x}{N}) \right] \prod_{\substack{j=1 \\ j \ne k,l}}^{n} \mu^{(\eta)}(h_j) + o(1) \, .
$$

For the same reason we include the random-walk interactions coming from $d\,\eta_x$ in $o(1)$, and finally arrive at

$$
\sum_{1 \le k < l \le n} \sum_{\substack{x,y=1 \\ x \ne y}}^{L} \frac{\eta_x}{N} \frac{\eta_y}{N} \left[\tilde{h}'_k(\tfrac{\eta_y}{N}) - \tilde{h}'_k(\tfrac{\eta_x}{N}) \right] \left[\tilde{h}'_l(\tfrac{\eta_y}{N}) - \tilde{h}'_l(\tfrac{\eta_x}{N}) \right] \prod_{\substack{j=1 \\ j \ne k,l}}^{n} \mu^{(\eta)}(h_j) + o(1)
$$

$$
= 2 \sum_{1 \le k < l \le n} \left(\mu^{(\eta)}(\tilde{h}'_k \tilde{h}'_l) - \mu^{(\eta)}(\tilde{h}'_k) \mu^{(\eta)}(\tilde{h}'_l) \right) \prod_{\substack{j=1 \\ j \ne k,l}}^{n} \mu^{(\eta)}(h_j) + o(1) \, ,
$$

where we expanded the product of square brackets and added the (non-contributing) diagonal $x = y$, before writing the expression in terms of $\mu^{(\eta)}$. Also, recall that $\tilde{h}' = Bh$.

Overall, combining the three bullets above, we derive

$$
\mathfrak{L}_{L,N} H(\mu^{(\cdot)})(\eta) = \mathcal{L}_\theta H(\mu^{(\eta)}) + o(1) \, , \tag{2.16}
$$

uniformly in $\eta \in \Omega_{L,N}$. This finishes the proof. $\qquad \square$

2.1.2 Convergence to the measure-valued process

The measure-valued process takes values in the space $E = \mu^{(\overline{\nabla})} \subset \mathcal{M}_1([0,1])$, cf. (1.3). Due to Lemma A.1 E itself is closed, thus, compact w.r.t. the topology induced by weak convergence of probability measures. Because this topology coincides with the subspace topology, the Hausdorff property of E is inherited from $\mathcal{M}_1([0,1])$.

In this section, we show that the dynamics described by $\mathcal{L}_\theta$ give rise to a Feller process and prove Theorem 1.1, which states that the process arises naturally as the scaling limit of the inclusion process.

Proposition 2.4. *For $\theta \in [0, \infty)$ the linear operator $(\mathcal{L}_\theta, \mathcal{D}(\mathcal{L}_\theta))$ is closable and its closure generates a Feller process on the state space $E \subset \mathcal{M}_1([0,1])$.*

The proof follows along the lines of [22, Theorem 2.5] where they proved existence of the Poisson-Dirichlet diffusion.

Proof. Throughout the proof we will make use of the sub-domain

$$
\mathcal{D}_{mon}(\mathcal{L}_\theta) := \Big\{ \text{sub-algebra of } C(E) \text{ generated by functions}
$$

$$
\mu \mapsto \mu(h) \text{ with } h(z) = z^m \, , m \in \mathbb{N}_0 \Big\} \subset \mathcal{D}(\mathcal{L}_\theta) \, . \tag{2.17}
$$

First note that, due to the Stone-Weierstrass theorem, $\mathcal{D}_{mon}(\mathcal{L}_\theta)$ (and therefore $\mathcal{D}(\mathcal{L}_\theta)$) is dense in $C(E)$ since it separates points: consider $\mu, \sigma \in E$ such that $\mu \ne \sigma$, then $\mu(z^m) \ne \sigma(z^m)$ for some $m \in \mathbb{N}$ since otherwise all moments, and hence μ and σ, agree.

Next, dissipativity of $(\mathcal{L}_\theta, \mathcal{D}(\mathcal{L}_\theta))$ follows from that of $(\mathfrak{L}_{L,N})_{L,N}$, since for any $H \in \mathcal{D}(\mathcal{L}_\theta)$ we have

$$\|(\lambda - \mathfrak{L}_{L,N})H(\mu^{(\cdot)})\|_{\Omega_{L,N},\infty} \geq \lambda \|H(\mu^{(\cdot)})\|_{\Omega_{L,N},\infty} \quad \forall \lambda > 0. \tag{2.18}$$

The left hand side is upper bounded by

$$\|(\lambda - \mathcal{L}_\theta)H\|_{E,\infty} + \|\mathcal{L}_\theta H(\mu^{(\cdot)}) - \mathfrak{L}_{L,N}H(\mu^{(\cdot)})\|_{\Omega_{L,N},\infty}, \tag{2.19}$$

with the second term vanishing due to Proposition 2.1. On the other hand, using Lemma A.4, we have

$$\sup_{\eta \in \Omega_{L,N}} |H(\mu^{(\eta)})| \to \sup_{p \in \overline{\nabla}} |H(\mu^{(p)})| = \|H\|_{E,\infty}. \tag{2.20}$$

In the remainder of the proof, we first conclude that $\mathcal{D}_{mon}(\mathcal{L}_\theta)$ is a core for $\mathcal{L}_\theta$, using the fact that $\mathcal{L}_\theta$ is triangulisable. The full statement then follows immediately by an extension argument. For that purpose, we define subspaces

$$D_n(\mathcal{L}_\theta) := \{H \in \mathcal{D}_{mon}(\mathcal{L}_\theta) : deg(H) \leq n\}, \tag{2.21}$$

where $deg(H) = m_1 + \cdots + m_k$ if H is of the form $\mu(z^{m_1}) \cdots \mu(z^{m_k})$, $m_j \in \mathbb{N}$ for $1 \leq j \leq k$. When H is given by linear combinations of such products, the degree denotes the maximum degree of the products. Note that $\left(D_n(\mathcal{L}_\theta)\right)_{n \geq 1}$ defines an increasing sequence with limit $\mathcal{D}_{mon}(\mathcal{L}_\theta)$. It is only left to show that $\mathcal{L}_\theta$ maps elements of $D_n(\mathcal{L}_\theta)$ back into itself. This is, however, immediate since both parts of the generator $\mathcal{L}_\theta$ (1.9) map polynomials of a certain degree back into polynomials of the same degree. Hence, using [23, Proposition I.3.5], we conclude that $(\mathcal{L}_\theta, \mathcal{D}_{mon}(\mathcal{L}_\theta))$ is indeed closable and gives rise to a strongly continuous contraction semigroup $(T_t)_{t \geq 0}$ on $C(E)$.

Now, we can easily verify that also $\mathcal{D}(\mathcal{L}_\theta)$ is a core, using [23, Proposition I.3.1], since

$$\mathcal{R}(\lambda - \mathcal{L}_\theta|_{\mathcal{D}(\mathcal{L}_\theta)}) \supset \mathcal{R}(\lambda - \mathcal{L}_\theta|_{\mathcal{D}_{mon}(\mathcal{L}_\theta)}), \tag{2.22}$$

is dense for some $\lambda > 0$. Since generators are maximal dissipative, we know that the closures w.r.t. both cores must agree, cf. [23, Proposition I.4.1] and hence give rise to the same semigroup $(T_t)_{t \geq 0}$. It is only left to show that the semigroup is positive and conservative, in particular E is invariant under the dynamics $\mathcal{L}_\theta$. In order to see this, we apply Trotter's theorem, see e.g. [23, Theorem I.6.1], which concludes that Proposition 2.1 implies

$$\lim_{N/L \to \rho} \sup_{\eta \in \Omega_{L,N}} \left|\mathfrak{T}_t^{(L,N)} H(\mu^{(\cdot)})(\eta) - (T_t H)(\mu^{(\eta)})\right| = 0, \quad \forall H \in C(E), t \geq 0, \tag{2.23}$$

where $\mathfrak{T}^{(L,N)}$ denotes the semigroup generated by $\mathfrak{L}_{L,N}$. Now, both positivity and conservation follow from those of $(\mathfrak{T}^{(L,N)})_{L,N}$. This conludes the proof. $\qquad\square$

Remark 2.5. It is natural to ask why one should go through the inconveniences of extending the core from $\mathcal{D}_{mon}(\mathcal{L}_\theta)$ to $\mathcal{D}(\mathcal{L}_\theta)$. However, we will see in the next subsection that the extended core allows for a better interpretation of the underlying dynamics in the Poisson-Dirichlet diffusion.

The proof of our first main result, namely, convergence of the inclusion process (when embedded in the space of probability measures) to the measure-valued process characterised by $\mathcal{L}_\theta$, is now an immediate consequence of a classical convergence theorem.

Proof of Theorem 1.1. We apply [23, Theorem IV.2.11] together with (2.23), which immediately concludes the desired convergence result. □

Finally, we can use Theorem 1.1 to prove convergence of the inclusion process to the Poisson-Dirichlet diffusion.

Proof of Corollary 1.4. Every function φ_m can be written in terms of an expectation

$$p \mapsto \mu^{(p)}(h_m) = \varphi_m(p),$$

with $h_m(z) := z^{m-1}$. Thus, we have

$$\left(\varphi_m \left(\tfrac{1}{N} \hat{\eta}^{(L,N)}(t) \right) \right)_{t \geq 0} = \left(\left(\mu_{\#} \eta^{(L,N)}(t) \right)(h_m) \right)_{t \geq 0} \xrightarrow{D} \left(\mu_t(h_m) \right)_{t \geq 0}, \tag{2.24}$$

using Theorem 1.1. This convergence can be extended to arbitrary elements in $\mathcal{D}_{mon}(\mathcal{G}_\theta)$, in particular such sequences are tight in $D([0,\infty), \mathbb{R})$. Thus, the sequence $\left(\left(\tfrac{1}{N} \hat{\eta}^{(L,N)}(t) \right)_{t \geq 0} \right)_{L,N}$ is tight in $D([0,\infty), \overline{\nabla})$ and has subsequential limits, see e.g. [23, Theorem III.9.1]. As convergence of finite dimensional marginals follows from (2.24), we conclude the statement together with Proposition 1.3, which is proved in the next subsection. □

2.2 Equivalence of the measure-valued process with PD-diffusion

In this section we prove Proposition 1.3 and investigate the equivalence of the measure-valued process generated by $\mathcal{L}_\theta$ (1.9) and the Poisson-Dirichlet diffusion on the simplex $\overline{\nabla}$, generated by $\mathcal{G}_\theta$ (1.19). We already saw in the proof of Lemma 2.2, cf. (2.8), the similarity of dynamics of $\mathcal{L}_\theta$ and $\mathcal{G}_\theta$. Indeed, a simple calculation shows that the two can be linked: Using the embedding (1.4) we get for all $p \in \overline{\nabla}$ and $H(\mu) = \mu(h)$, with $h \in \mathcal{D}(A)$,

$$\mathcal{L}_\theta H(\mu^{(p)}) = \mu^{(p)}(A_\theta h) = \left(1 - \|p\|_1 \right) A_\theta h(0)$$
$$+ \sum_{i=1}^{\infty} p_i \left(p_i(1 - p_i) h''(p_i) + \left(2(1 - p_i) - \theta p_i \right) h'(p_i) + \theta(h(0) - h(p_i)) \right).$$

Defining now $f(p) := \mu^{(p)}(h)$, we have

$$\mathcal{L}_\theta H(\mu^{(p)}) = 2h'(0)(1 - \|p\|_1) + \mathcal{G}_\theta f(p), \tag{2.25}$$

where we used that $\partial_{p_i} f(p) = -h(0) + p_i h'(p_i) + h(p_i))$ and $\partial_{p_i p_j} f(p) = \delta_{ij} \left(2h'(p_i) + p_i h''(p_i) \right)$ and $\mathcal{G}_\theta$ as defined in (1.19).

Remark 2.6. In [22], the authors extended the domain of $\mathcal{G}_\theta$ from $\mathcal{D}(\mathcal{G}_\theta)$ to the sub-algebra of $C(\overline{\nabla})$ generated by functions of the form $p \mapsto \sum_{i=1}^{\infty} g(p_i)$, with $g \in C^2([0,1])$ such that $g(0) = g'(0) = 0$. This yields a similar expression as (2.25), cf. [22, Display (2.17)]. However, the expression again only made sense with the convention that sums are evaluated on ∇ and extended by continuity, in which case the first summand in (2.25) disappears.

Proof of Proposition 1.3. In order to show the equivalence of the two processes, it suffices to restrict ourselves to the domains generated by monomials as defined in (1.20) and (2.17). Every function $H \in \mathcal{D}_{mon}(\mathcal{L}_\theta)$ can be mapped to $f_H \in \mathcal{D}_{mon}(\mathcal{G}_\theta)$ (and vise versa). Let $H \in \mathcal{D}_{mon}(\mathcal{L}_\theta)$ be of the form $H(\mu) = \mu(h_1) \cdots \mu(h_n)$, where $h_k(z) := z^{m_k - 1}$,

then $f_H = \varphi_{m_1} \cdots \varphi_{m_n}$ where we recall $\varphi_m(p) = \sum_{i=1}^\infty p_i^m$. Then

$$(\mathcal{L}_\theta H)(\mu^{(p)}) = 2 \sum_{1 \le k < l \le n} m_k m_l \big(\mu^{(p)}(h_k h_l) - \mu^{(p)}(h_k)\mu^{(p)}(h_l) \big) \prod_{j \ne k,l} \mu^{(p)}(h_j) \qquad (2.26)$$
$$+ \sum_{1 \le k \le n} \mu^{(p)}(A_\theta h_k) \prod_{j \ne k} \mu^{(p)}(h_j),$$

where we used that $Bh_k = m_k\, h_k$. Rewriting the r.h.s. in terms of φ's, we have

$$(\mathcal{L}_\theta H)(\mu^{(\cdot)}) = 2 \sum_{1 \le k < l \le n} m_k m_l \big(\varphi_{m_k + m_l - 1} - \varphi_{m_k}\varphi_{m_l} \big) \prod_{j \ne k,l} \varphi_{m_j}$$
$$+ \sum_{1 \le k \le n} \mathcal{G}_\theta \varphi_{m_k} \prod_{j \ne k} \varphi_{m_j},$$

where we used $\mu^{(p)}(A_\theta h_k) = \mathcal{G}_\theta \varphi_{m_k}$ from (2.25) and the fact that $\varphi_{m_k} \in \mathcal{D}_{mon}(\mathcal{G}_\theta)$ since $m_k \ne 1$. Thus, $(\mathcal{L}_\theta H)(\mu^{(\cdot)})$ agrees with $\mathcal{G}_\theta f_H$ on $\mathcal{D}_{mon}(\mathcal{G}_\theta)$, cf. [22, Display (2.13)]. Let $(X(t))_{t \ge 0}$ be the Poisson-Dirichlet diffusion, then for every $H \in \mathcal{D}_{mon}(\mathcal{L}_\theta)$

$$H(\mu^{(X(t))}) - \int_0^t \mathcal{L}_\theta H(\mu^{(X_s)})\, ds = f_H(X(t)) - \int_0^t \mathcal{G}_\theta f_H(X_s)\, ds \qquad (2.27)$$

defines a martingale in t. Thus, $(\mu^{(X(t))})_{t \ge 0}$ solves the martingale problem for $(\mathcal{L}_\theta, \mathcal{D}_{mon}(\mathcal{L}_\theta))$.

Now, it is almost immediate that properties (i) – (iii) in Proposition 1.3 hold for the measure valued process $(\mu_t)_{t \ge 0}$. First, let $G, H \in \mathcal{D}_{mon}(\mathcal{L}_\theta)$ and choose corresponding $f_G, f_H \in \mathcal{D}_{mon}(\mathcal{G}_\theta)$ as above. We know that $(\mathcal{L}_\theta G)(\mu^{(p)}) = \mathcal{G}_\theta f_G(p)$. Writing $\nu = \mathrm{PD}(\theta)$ for simplicity, we have $\mathbf{P} = \mu_\# \mathrm{PD}(\theta)$ with

$$\mathbf{P}(H\mathcal{L}_\theta G) = \nu\big(H(\mu^{(p)})(\mathcal{L}_\theta G)(\mu^{(p)}) \big) = \nu(f_H \mathcal{G}_\theta f_G). \qquad (2.28)$$

It is known that $\mathrm{PD}(\theta)$ is the unique invariant distribution w.r.t. $\mathcal{G}_\theta$ [22, Theorem 4.3]; which is also reversible. Together with the above display, this yields $\mathbf{P}(H\mathcal{L}_\theta G) = \mathbf{P}(G\mathcal{L}_\theta H)$.

Continuity of the trajectories in (ii) follows from the diffusion property of $(X(t)(\omega))_{t \ge 0}$ and continuity of the map $\mu^{(\cdot)}$, together with the fact $(\mu^{(X(t))})_{t \ge 0} \overset{d}{=} (\mu_t)_{t \ge 0}$.

Lastly, (iii) is a consequence of

$$\mathbb{P}\left[\mu_t(\{0\}) = 0 \quad \forall t > 0\right] = \mathbb{P}\left[\mu^{(X(t))}(\{0\}) = 0 \quad \forall t > 0\right] = \mathbb{P}\left[X(t) \in \nabla \quad \forall t > 0\right] = 1,$$

where we used [22, Theorem 2.6] in the last step. $\qquad \square$

Remark 2.7. Naturally, we could have proven convergence of the inclusion process to the Poisson-Dirichlet diffusion directly and then defined the measure-valued dynamics using the embedding via $\mu^{(\cdot)}$. This would have slightly shortened the exposition in the present section, since it would not have been necessary to verify existence of the limiting dynamics. We refrained from doing so for the sake of a better understanding of the underlying dynamics in the measure-valued process, in particular on the extended domain $\mathcal{D}(\mathcal{L}_\theta)$.

Recall the generator of the single-particle dynamic (1.10)

$$A_\theta h(z) := z(1-z)h''(z) + (2(1-z) - \theta z)h'(z) + \theta(h(0) - h(z)), \quad h \in C^2([0,1]),$$

which characterises a Feller process on the unit interval. The process evolves according to a diffusion with an additional renewal mechanism due to jumps to zero.

Lemma 2.8. *The Beta distribution* $\mathrm{Beta}(1,\theta)$ *is the unique invariant distribution with respect to* A_θ.

Proof. For $\theta = 0$, we interpret the degenerate distribution $\mathrm{Beta}(1,0)$ as the Dirac point mass δ_1. The statement is then clear since $A_\theta h(1) = 0$. Hence, in this case with $\theta = 0$ the point mass is even reversible.

Now, let $\theta > 0$ and consider $H(\mu) := \mu(h)$, then by Proposition 1.3(i)

$$0 = \mathbf{P}(\mathcal{L}_\theta H) = \int \mu(A_\theta h)\, \mathbf{P}(d\mu) = \mathbb{E}\Big[\sum_{i=1}^\infty X_i A_\theta h(X_i)\Big] = \mathbb{E}[A_\theta h(\tilde{X}_1)]\,, \qquad (2.29)$$

where $X \sim \mathrm{PD}(\theta)$. It is well known that the first size-biased marginal $\tilde{X}_1$ is $\mathrm{Beta}(1,\theta)$-distributed [25, Theorem 2.7].

Uniqueness of the Beta distribution is due to Harris recurrence of the process, see e.g. [41]. For the case $\theta > 0$, the resetting mechanism guarantees that the process returns to zero infinitely often almost surely. On the other hand for $\theta = 0$, A_θ agrees with a Jacobi diffusion, see (2.32) below. The corresponding process runs into the absorbing state $z = 1$ in finite time, independent of the initial condition. $\qquad \square$

Due to the jumps to zero, one does not expect that $\mathrm{Beta}(1,\theta)$, $\theta > 0$, is reversible w.r.t. A_θ. Indeed, this can be verified easily by considering the example $h(x) = x$ and $g(x) = x^2$, in which case

$$\mathrm{Beta}(1,\theta)(gA_\theta h) = \frac{8\theta\Gamma(\theta+1)}{\Gamma(\theta+4)} \neq \frac{6\theta\Gamma(\theta+1)}{\Gamma(\theta+4)} = \mathrm{Beta}(1,\theta)(hA_\theta g)\,, \quad \forall \theta > 0\,. \qquad (2.30)$$

2.3 The advantage of a size-biased evolution

Throughout the previous sections, we have seen two viewpoints of the same dynamics. The classical Poisson-Dirichlet diffusion considers a ranked configuration space. However, this obscures the dynamics on microscopic scales, which results e.g. into defining the r.h.s. of the generator $\mathcal{G}_\theta$ (1.19) to be *evaluated on ∇ and extended to $\overline{\nabla}$ by continuity*, because $\varphi_1 = \|\cdot\|_1$ does not lie in the domain $\mathcal{D}_{mon}(\mathcal{G}_\theta)$. Alternatively, one can consider unordered dynamics, i.e. observing the evolution from a fixed position, or with a size-biased viewpoint. The Poisson-Dirichlet diffusion concentrates immediately on configurations consisting of macroscopic-sized fragments, which can only concentrate on a vanishing fraction of the volume. Hence, an unordered state space can only describe dynamics up to a certain point when the mass present at the observed positions disappears. The goal of this section is to emphasise that a size-biased viewpoint allows for both, a complete description of the macroscopic dynamics while observing interaction with the microscopic scale.

2.3.1 Time evolution on fixed sites

First, we look at arbitrary finite positions and observe the evolution of masses on them. As the inclusion process is spatially homogeneous, we may choose for simplicity $\eta \mapsto (\eta_1, \dots, \eta_n)$.

We start by only considering the evolution on the first site. Performing similar approximations as in Section 2, we can see for an arbitrary function $h \in C^3([0,1])$

$$\mathfrak{L}_{L,N} h(\tfrac{(\cdot)_1}{N})(\eta) = A_{Jac(\theta)} h(\tfrac{\eta_1}{N}) + o(1)\,, \qquad (2.31)$$

where

$$A_{Jac(\theta)} h(z) := z(1-z)h''(z) - \theta z\, h'(z)\,. \qquad (2.32)$$

In fact, we can see $A_{Jac(\theta)}$ emerging in (2.8) when fixing a position x. The operator $A_{Jac(\theta)}$ is the generator of a Jacobi-diffusion, cf. [27], and describes the evolution of a single chunk of mass located at a given position.

To describe the evolution on the first n positions we introduce for $i = 0, 1, \ldots, n$

$$\xi_i = \xi_i(\eta) := \begin{cases} \eta_i & \text{if } 1 \leq i \leq n \\ N - \sum_{j=1}^n \xi_j & \text{if } i = 0 \,, \end{cases} \tag{2.33}$$

where ξ_0 is the remaining mass in the system outside sites $1, \ldots, n$. Thus, the rescaled vector $\frac{1}{N}\xi$ lies in $\Delta_{n+1} := \{p \in [0,1]^{n+1} : \sum_{i=0}^n p_i = 1\}$. Again, by approximation of the generators, one can show that

$$\tfrac{1}{N}(\xi(t))_{t \geq 0} \xrightarrow{D} \mathrm{WF}_{n+1}(\theta, 0, \ldots, 0) \,. \tag{2.34}$$

Here, $\mathrm{WF}_{n+1}(\theta, 0, \ldots, 0)$ denotes the Wright-Fisher diffusion on Δ_{n+1} which is characterised by the generator

$$A_{WF_{n+1}(\theta,\mathbf{0})} h(z_{0,\ldots,n}) = \sum_{i,j=0}^n z_i(\delta_{i,j} - z_j)\partial^2_{z_i z_j} h(z_{0,\ldots,n}) \tag{2.35}$$

$$+ \theta \sum_{i=1}^n z_i(\partial_{z_0} h - \partial_{z_i} h)(z_{0,\ldots,n}) \,,$$

acting on those h that have an extension to $\mathbb{R}^{n+1}$ which is twice continuously differentiable.

However, we can already see for a single observable that the Jacobi-diffusion has an absorbing state at $z = 0$, which it will run into in finite time almost surely [17, Section 7.10]. Similarly, the Wright-Fisher diffusion will be absorbed at $(1, 0, \ldots 0) \in \Delta_{n+1}$, after which the process does not capture the dynamics of the infinite-dimensional process anymore as all the mass has moved away from the first n sites.

For the Poisson-Dirichlet diffusion, the relationship to the Jacobi and Wright-Fisher diffusion has been studied in greater generality in the two-parameter setting [27]. They use a Fleming-Viot construction of the process, cf. (1.35). Because they start from the Poisson-Dirichlet diffusion on $\overline{\nabla}$, there is no underlying graph structure and instead of placing mass at a fixed position, they choose a uniform random variable on $[0,1]$ which determines the position of the point mass.

Moreover, it is interesting to note that the boundary behaviour of the Jacobi diffusion agrees with the one of the PD-diffusion. More precisely, in the case of the Jacobi diffusion the state 1 can only be reached if and only if $\theta < 1$, see e.g. [47, Theorem 4.1] or [17, Section 7.10]. Similarly, the PD-diffusion $(X(t))_{t \geq 0}$ hits the finite dimensional sub-simplices $\nabla \cap \{\sum_{i=1}^n p_i = 1\}$, for any $n \geq 1$, if and only if $\theta < 1$ [45].

2.3.2 Duality and size-biased time evolution

We recall from Dynkin's formula, i.e. taking expectations of the first term in (1.17),

$$\frac{d}{dt}\mathbb{E}_{\mu_0}\big[\mu_t(h)\big] = \mathbb{E}_{\mu_0}\big[\mu_t(A_\theta h)\big] \,, \quad \forall h \in \mathcal{D}(A) \,,$$

which implies the duality (1.18)

$$\mathbb{E}_{\mu_0}[\mu_t(h)] = \mathbf{E}_{\mu_0}[h(Z(t))] \,, \quad \forall h \in \mathcal{D}(A) \,, \tag{2.36}$$

where $(Z(t))_{t \geq 0}$ is a process on $[0,1]$ with generator A_θ, cf. (1.13), and initial distribution μ_0. The identity can be extended to all $h \in C([0,1])$ by standard arguments, see e.g. [20,

Section 6]. In analogy to known duality properties of the microscopic particle system [28, 9] this can for example be used to get closed evolution equations for moments. Due to size-biasing $h(z) = z$ describes the expected second moment of the mass distribution and we get

$$\frac{d}{dt}\mathbb{E}_{\mu_0}[\mu_t(z)] = \mathbf{E}_{\mu_0}[A_\theta \mathrm{Id}(Z(t))] = \mathbf{E}_{\mu_0}\Big[2\big(1 - (1+\theta)Z_t\big)\Big] = 2 - 2(1+\theta)\mathbb{E}_{\mu_0}[\mu_t(z)] \,.$$

This has an exponential solution which converges to the stationary point $\frac{1}{1+\theta}$, the expected second moment of the GEM (θ) distribution. Dualities of this form were previously considered in [15] and [16], see also [20, Section 6] for a summary. In Proposition 3.4 we will see that for $dL \to \infty$ dualities of the form (2.36) extend directly to nonlinear test functions $H(\mu)$, due to the absence of an interaction part in the generator $\hat{\mathcal{L}}$, cf. (1.16). In the present case it may be possible to establish higher dimensional dual processes evaluated at Bh rather than h, with Fleming-Viot-type resampling. This is not relevant for the aim of this paper but could be an interesting question for future studies.

We stress once more the difference in point of view: whereas in previous works, see [20] and references therein, the dual particles encode the position of clusters on the underlying lattice, in our size-biased approach the state of dual particles characterises the fragmentation of mass in a given configuration/partition. This allows for observing the dynamics of macroscopic cluster size distributions, while tracking only a finite number of dual particles.

The duality in (2.36) is also interesting from a computational point of view, as it allows to continuously track the expected behaviour of the infinite dimensional process using only a finite-dimensional diffusion, without running into any absorbing states as is the case when observing a fixed set of lattice sites. A simple example is the second moment of cluster sizes in the Poisson-Dirichlet diffusion at time t which is given by $\mathbf{E}_{\mu_0}[Z(t)]$ as mentioned above.

3 The diffusion limit in the case $dL \to \infty$

The case of $dL \to \infty$ may be considered as an interpretation of the Poisson-Dirichlet diffusion with infinite mutation rate $\theta = \infty$. Clearly, this corresponds to an infinite drift towards zero in the single particle operator A_θ, cf. (1.10). Thus, in order to see non-trivial dynamics, we have to rescale time appropriately. Recalling (2.9), we see that

$$\frac{1}{dL}\mathfrak{L}_{L,N}H(\mu^{(\cdot)})(\eta) = \mu^{(\eta)}\Big((h(0) - h(z) - zh'(z))\Big) + o(1)\,, \tag{3.1}$$

where $H(\mu) = \mu(h)$, $h \in C^3([0,1])$. The time-change also eradicates the interaction term in the corresponding limiting measure-valued process. We are left with a process that pushes mass (deterministically) from the interval $(0,1]$ onto zero. Hence, mass will not accumulate on the macroscopic scale, instead we need to consider an appropriate mesoscopic scale to see the actual dynamics of the fast mixing mechanism.

In [35] it was proven that, at stationarity, mass accumulates on the mesoscopic scale of order d^{-1}, when $\rho \in (0, \infty)$, cf. (1.31). Thus the embedding of particle configurations into $\mathcal{M}_1(\mathbb{R}_+)$ via (1.12) with $\hat{\mu}^{(\eta)} = \sum_{x=1}^{L} \frac{\eta_x}{N} \delta_{\frac{dL}{N}\eta_x}$ is an appropriate a-priori choice.[2] In order to take particle configurations with mass lying on larger scales than $N/(dL)$ into account, we will consider probability measures $\mathcal{M}_1(\overline{\mathbb{R}}_+)$ on the one-point compactification, instead of restricting ourselves to the positive real line. We equip $\mathcal{M}_1(\overline{\mathbb{R}}_+)$ with the topology induced by weak convergence, thus, $\mathcal{M}_1(\overline{\mathbb{R}}_+)$ is compact.

---

[2]For the case $\rho \in (0, \infty)$, also the choice $\delta_{d\eta_x}$ is appropriate and leads to a ρ dependent limit. However, for the boundary cases $\rho \in \{0, \infty\}$ the given choice turns out to be the correct one.

3.1 Deriving the diffusion limit

Once more we rely on the Trotter-Kurtz approximation to conclude the scaling limit in Theorem 1.2. We follow the same steps as in Section 2, carried out below for completeness.

Proposition 3.1. *Let $H \in \mathcal{D}(\hat{\mathcal{L}})$, then*

$$\lim_{N/L \to \rho} \sup_{\eta \in \Omega_{L,N}} \left| \frac{1}{dL} \mathfrak{L}_{L,N} H(\hat{\mu}^{(\cdot)})(\eta) - \hat{\mathcal{L}} H(\hat{\mu}^{(\eta)}) \right| = 0. \tag{3.2}$$

Below, we will show that the interaction term of the limiting measure-valued process indeed vanishes. First, we only consider test functions of the form $\mu \mapsto \mu(h)$.

Lemma 3.2. *Let $\rho \in (0, \infty)$ and $H(\mu) = \mu(h)$, with $h \in \mathcal{D}(\hat{A})$ (1.14). Then*

$$\lim_{N/L \to \rho} \sup_{\eta \in \Omega_{L,N}} \left| \frac{1}{dL} \mathfrak{L}_{L,N} H(\hat{\mu}^{(\cdot)})(\eta) - \hat{\mu}^{(\eta)}(\hat{A}h) \right| = 0, \tag{3.3}$$

where $\hat{A}$ is the single-particle generator defined in (1.13).

Proof. Without loss of generality, we assume that $h \in C_c^3(\mathbb{R}_+)$. For simplicity of notation we will write $p_x = dL \frac{\eta_x}{N}$. Following the same steps as in the proof of Lemma 2.2, we have

$$\frac{1}{dL} \mathfrak{L}_{L,N} H(\hat{\mu}^{(\cdot)})(\eta) = \frac{1}{2N^2} \sum_{\substack{x,y=1 \\ x \neq y}}^{L} \eta_x \eta_y \left(\widetilde{h}''(p_x) + \widetilde{h}''(p_y) \right) \tag{3.4}$$

$$+ \frac{1}{LN} \sum_{\substack{x,y=1 \\ x \neq y}}^{L} \eta_x \left(\widetilde{h}'(p_y) - \widetilde{h}'(p_x) \right) + o(1),$$

where we gained an additional factor $(dL)^{-1}$ by rewriting $\frac{\eta_x}{N} h(d\eta_x) = \frac{1}{dL} \tilde{h}(p_x)$. Here we used again the fact that second-order terms in the second sum have a vanishing contribution and first-order terms in the first sum cancel exactly, cf. proof of Lemma 2.2. Display (3.4) can be written as

$$\frac{1}{dL} \mathfrak{L}_{L,N} H(\mu^{(\cdot)})(\eta) = \sum_{x=1}^{L} \frac{\eta_x}{N} \left(1 - \frac{p_x}{dL} \right) \tilde{h}''(p_x) \tag{3.5}$$

$$+ \left[\frac{1}{L} \sum_{y=1}^{L} \tilde{h}'(p_y) - \sum_{x=1}^{L} \frac{\eta_x}{N} \tilde{h}'(p_x) \right] + o(1).$$

Using again Lemma A.2, we have $\frac{1}{L} \sum_{y=1}^{L} \tilde{h}'(p_y) = \tilde{h}'(0) + o(1)$. Hence,

$$\frac{1}{dL} \mathfrak{L}_{L,N} H(\mu^{(\cdot)})(\eta) = \hat{\mu}^{(\eta)} \left(\tilde{h}''(p) + [\tilde{h}'(0) - \tilde{h}'(p)] \right) + o(1) = \hat{\mu}^{(\eta)} \left(\hat{A}h \right) + o(1), \tag{3.6}$$

where we additionally used the fact that $\tilde{h}''$ is bounded and of compact support, thus, $\frac{p}{dL} \tilde{h}''(p)$ vanishes in the thermodynamic limit because $dL \to \infty$. □

Remark 3.3. Considering the state space $\mathcal{M}_{\leq 1}(\mathbb{R}_+)$ (positive measures on $\mathbb{R}_+$ with mass ≤ 1) instead of $\mathcal{M}_1(\overline{\mathbb{R}}_+)$, (3.5) suggests that $\hat{A}$ should act via

$$H \mapsto \left(\mu \mapsto \mu(\hat{A}h) + (1 - \mu(1))h(0) \right), \tag{3.7}$$

on functions $H : \mathcal{M}_{\leq 1}(\mathbb{R}_+) \mapsto \mathbb{R}$ of the form $H(\mu) = \mu(h)$. The extra term takes into account the transfer of mass from larger scales, i.e. above $N/(dL)$, which is pushed onto

microscopic scales, cf. Corollary 3.7. However, $\mu \mapsto \mu(1)$ is not a continuous function on $\mathcal{M}_{\leq 1}(\mathbb{R}_+)$. Instead the mass transport from larger scales is implicit in the generator $\hat{\mathcal{L}}$, as we will see below.

Proof of Proposition 3.1. It suffices to consider functions $H \in \mathcal{D}(\hat{\mathcal{L}})$ of the form $H(\mu) = \mu(h_1) \cdots \mu(h_n)$, with $h_k \in C_c^3(\mathbb{R})$. The interaction term of the operator $\hat{\mathcal{L}}$ is again given by the second-order term of the following expansion

$$
\begin{aligned}
H(\hat{\mu}_\# \eta^{x,y}) = {}& H(\hat{\mu}^{(\eta)}) \\
&+ \frac{1}{dL} \sum_{k=1}^n [\widetilde{h}_k(p_y + \tfrac{dL}{N}) - \widetilde{h}_k(p_y) + \widetilde{h}_k(p_x - \tfrac{dL}{N}) - \widetilde{h}_k(p_x)] \prod_{\substack{l=1 \\ l \neq k}}^n \hat{\mu}^{(\eta)}(h_l) \\
&+ \frac{1}{(dL)^2} \sum_{1 \leq k < l \leq n} [\widetilde{h}_k(p_y + \tfrac{dL}{N}) - \widetilde{h}_k(p_y) + \widetilde{h}_k(p_x - \tfrac{dL}{N}) - \widetilde{h}_k(p_x)] \quad (3.8) \\
&\qquad \times [\widetilde{h}_l(p_y + \tfrac{dL}{N}) - \widetilde{h}_l(p_y) + \widetilde{h}_l(p_x - \tfrac{dL}{N}) - \widetilde{h}_l(p_x)] \prod_{\substack{j=1 \\ j \neq k,l}}^n \hat{\mu}^{(\eta)}(h_j) \\
&+ \frac{1}{(dL)^3} r(\eta).
\end{aligned}
$$

In the first order term each summand can be treated individually using Lemma 3.2, it only remains to check that both second-order term and remainder have no contribution.

Using a first-order Taylor expansion yields the following bound

$$
\begin{aligned}
\Big| \widetilde{h}(p_y + \tfrac{dL}{N}) &- \widetilde{h}(p_y) + \widetilde{h}(p_x - \tfrac{dL}{N}) - \widetilde{h}(p_x) \Big| \\
&= \frac{dL}{N} \Big| \tilde{h}'(p_y) - \tilde{h}'(p_x) + \frac{1}{2} \frac{dL}{N}(\tilde{h}''(\xi_y) + \tilde{h}''(\xi_x)) \Big| \\
&\leq 2 \frac{dL}{N}(\|\tilde{h}'\|_\infty + \|\tilde{h}''\|_\infty),
\end{aligned}
$$

where $\xi_x, \xi_y \in [0, dN]$ corresponding to the associated remainder term. Hence, the second-order term in (3.8), after applying $\frac{1}{dL} \mathfrak{L}_{L,N}$, is upper bounded (up to a constant) by

$$
\frac{1}{dL} \sum_{1 \leq k < l \leq n} \frac{1}{N^2} \sum_{\substack{x,y=1 \\ x \neq y}}^L \eta_x(d + \eta_y)(\|h_k'\|_\infty + \|h_k''\|_\infty)(\|h_l'\|_\infty + \|h_l''\|_\infty) \prod_{\substack{j=1 \\ j \neq k,l}}^n \|h_j\|_\infty,
$$

which vanishes as $dL \to \infty$. For the same reason, higher order terms in the expansion (3.8) have no contribution either. □

The closure of $(\hat{A}, \mathcal{D}(\hat{A}))$ generates a Feller semigroup on $\overline{\mathbb{R}}_+$, thus, the closure of $(\hat{\mathcal{L}}, \mathcal{D}(\hat{\mathcal{L}}))$ generates a Fleming-Viot process on the compact space $\mathcal{M}_1(\overline{\mathbb{R}}_+)$, with trajectories in $C([0, \infty), \mathcal{M}_1(\overline{\mathbb{R}}_+))$, in the absence of interaction [19, Theorem 2.3]. We now have everything at hand to prove our second main result, Theorem 1.2.

Proof of Theorem 1.2. Once more, we apply [23, Theorem IV.2.11] together with Proposition 3.1, which immediately concludes the desired convergence (1.15) in $D([0, \infty), \mathcal{M}_1(\overline{\mathbb{R}}_+))$. Note that any fixed initial condition $\mu \in \mathcal{M}_1(\overline{\mathbb{R}}_+)$ can be approximated by particle configurations using the embedding $\hat{\mu}^{(\cdot)}$, cf. Lemma A.5. This completes the proof. □

3.2 Duality and the hydrodynamic limit

The absence of interaction in $\hat{\mathcal{L}}$ leads to a deterministic evolution of $(\hat{\mu}_t(h))_{t \geq 0}$, $h \in \mathcal{D}(\hat{A})$. This is a consequence of Dynkin's formula as derived in (1.17), since the process solves the ODE $d\hat{\mu}_t(h) = \hat{\mu}_t(\hat{A}h)\, dt$. Hence, the evolution of $(\hat{\mu}_t(h))_{t \geq 0}$ can be described by a single particle evolving according to the process generated by $\hat{A}$, averaged over its initial condition $\hat{\mu}_0$. See also the duality mentioned in (1.18) which we extend in the next result. Unlike in the case of $dL \to \theta < \infty$, we can fully characterise the semigroup of $\hat{\mathcal{L}}$ by considering only the evolution w.r.t. the single particle generator $\hat{A}$.

Proposition 3.4. *Let $g \in C(\overline{\mathbb{R}}_+^n)$ and define $G(\mu) := \mu^{\otimes n}(g)$, then for any $\hat{\mu}_0 \in \mathcal{M}_1(\overline{\mathbb{R}}_+)$*

$$G(\hat{\mu}_t) = \mathbf{E}_{\hat{\mu}_0^{\otimes n}}\left[g(\hat{Z}(t))\right], \quad \forall t \geq 0,\tag{3.9}$$

where $(\hat{\mu}_t)_{t \geqslant 0}$ evolves w.r.t. $\hat{\mathcal{L}}$ and $(\hat{Z}(t))_{t \geq 0}$ is the process consisting of n independent copies generated by the single-particle generator $\hat{A}$, cf. (1.13). In particular, for $n = 1$ we have $\hat{\mu}_t = Law(\hat{Z}(t))$ whenever the initial conditions agree in the sense that $\hat{Z}(0) \sim \hat{\mu}_0$.

Proof. Following precisely the same steps as in [20, Section 6] using the resolvent operator, one can conclude the duality

$$\mathbb{E}_{\hat{\mu}_0}[G(\hat{\mu}_t)] = \mathbf{E}_{\hat{\mu}_0^{\otimes n}}\left[g(\hat{Z}(t))\right], \quad \forall t \geq 0,\tag{3.10}$$

with $g \in C(\overline{\mathbb{R}}_+^n)$ and $G(\mu) := \mu^{\otimes n}(g)$. This is essentially a direct consequence of the absence of an interaction term in the generator $\hat{\mathcal{L}}$, cf. (1.16), implying for $H(\mu) = \mu(h_1) \cdots \mu(h_n)$, $h_i \in \mathcal{D}(\hat{A})$,

$$\hat{\mathcal{L}}H(\mu) = \sum_{1 \leq k \leq n} \mu(\hat{A}h_k) \prod_{\substack{l=1 \\ l \neq k}}^{n} \mu(h_l).$$

Now, let us consider the case $n = 1$ for which the identity reads

$$\hat{\mu}_t(h) = \mathbb{E}_{\hat{\mu}_0}[\hat{\mu}_t(h)] = \mathbf{E}_{\hat{\mu}_0}\left[h(\hat{Z}(t))\right], \quad \forall t \geq 0,\ h \in C(\overline{\mathbb{R}}_+),\tag{3.11}$$

where we additionally used the fact that $\hat{\mu}_t(h)$ is deterministic, cf. (1.17). In particular, (3.11) implies that $\hat{\mu}_t = Law(\hat{Z}(t))$ and the measure-valued evolution $(\hat{\mu}_t)_{t \geqslant 0}$ is indeed deterministic. Hence, the expected value on the left-hand side of (3.10) has no affect and can be dropped. $\square$

The duality result in Proposition 3.4, and equivalently Theorem 1.2, can be interpreted in the sense of a hydrodynamic limit.

Proposition 3.5 (Hydrodynamic limit)**.** *Consider the process $(\hat{\mu}_t)_{t \geq 0}$ generated by $\hat{\mathcal{L}}$ with initial data $\hat{\mu}_0 \in \mathcal{M}_1(\overline{\mathbb{R}}_+)$. Then for every $t > 0$, $\hat{\mu}_t$ has a Lebesgue-density $f(t, \cdot)$ on $\mathbb{R}_+$. The evolution of the density $(f(t, \cdot))_{t > 0}$ solves*

$$\begin{cases} \partial_t f(t, z) = z\, \partial_z^2 f(t, z) + z\, \partial_z f(t, z) \\ \lim_{z \to 0} f(t, z) = 1 \end{cases}\tag{3.12}$$

with $\lim_{t \to 0} \int_0^\infty h(z) f(t, z)\, dz = \hat{\mu}_0(h)$ for every $h \in C_c(\mathbb{R}_+)$.

Remark 3.6. The diffusion part of the generator $\hat{A}$, given by

$$\hat{A}_D h(z) := z h''(z) + (2 - z) h'(z),\tag{3.13}$$

generates the so called Cox-Ingersoll-Ross model [40], which is a well studied diffusion process in mathematical finance and population genetics.

Proof. Consider first the case for an initial condition that has no atom at infinity, i.e. $\|f_0\|_{L^1(\mathbb{R}_+)} = 1$, and let $z_0 \in [0, \infty)$. The Cox-Ingersoll-Ross model generated by $\hat{A}_D$ is known to have a density $g(t, \cdot|z_0)$ for any positive time and initial data [40, Display below (3.2)]. In fact, it is explicitly given and for $z_0 = 0$ it evaluates to

$$g(t, z|0) = \frac{z}{(2\ell_t)^2} e^{-z(2\ell_t)^{-1}} \quad \text{with} \quad \ell_t := \tfrac{1}{2}(1 - e^{-t}). \tag{3.14}$$

Furthermore, for any $t > 0$ we have $g_t(\cdot|z_0)\big|_{(0,\infty)} \in C^\infty((0, \infty))$ [40, Proposition 3.2]. The resetting mechanism is given by a Poisson jump process, thus, [2, Theorem 1] guarantees that also the process $(\hat{Z}(t))_{t \geq 0}$ generated by $\hat{A}$ has a density which is given by

$$f(t, z|z_0) = e^{-t}g(t, z|z_0) + \int_0^t e^{-s} g(s, z|0)\, ds. \tag{3.15}$$

We note that $f(t, \cdot|z_0)$ inherits the regularity properties of $g(t, \cdot|z_0)$ on $(0, \infty)$. This follows from the change of variable $r = z/(2\ell_s)$ with $\frac{dr}{ds} = -\frac{z}{(2\ell_s^2)}e^{-s}$ which yields for every $z > 0$

$$f(t, z|z_0) = e^{-t}g(t, z|z_0) + \int_{\frac{z}{2\ell_t}}^\infty e^{-r}\, dr = e^{-t}g(t, z|z_0) + e^{-z(2\ell_t)^{-1}}. \tag{3.16}$$

Thus, $f(t, \cdot|z_0) \in C^\infty((0, \infty))$.

Now, it is only left to verify that $(f(t, \cdot))_{t \geq 0}$ indeed solves the given PDE. Using integration by parts, we see that for any $h \in C_c^2(\mathbb{R}_+)$, we have

$$\mu_t(\hat{A}h) = \int_0^\infty f(t, z)\, \hat{A}h(z)\, dz = \int_0^\infty \hat{A}^* f(t, z)\, h(z)\, dz \tag{3.17}$$

with the adjoint action defined as

$$\hat{A}^* f(z) := zf''(z) + zf'(z) + \delta_0(z)\big(1 - f(z)\big). \tag{3.18}$$

Hence, the density $f = \big(f(t, \cdot)\big|_{(0,\infty)}\big)_{t \geq 0}$ in (3.16) solves the PDE

$$\partial_t f = \hat{A}^* f, \quad f(0, \cdot) = \delta_{z_0}. \tag{3.19}$$

It is easy to see from (3.16) that $\lim_{z \to 0} f(t, z) = 1$ for any $t > 0$ since $g(t, 0|z_0) = 0$ for all $z_0 \geq 0$, thus, the boundary term in (3.18) vanishes and we are left with the PDE in the statement.

Consider now the case of $f_0 = \delta_\infty$. The only way $(\hat{Z}(t))_{t \geq 0}$ can escape infinity is via the resetting mechanism. Thus, (3.16) still applies with $g(t, z|z_0)$ replaced by δ_∞, since $\mathbb{P}(\hat{Z}(t) = \infty)$ is equal to the probability that the process has not jumped yet. One can check that also $f(\cdot|\infty)$ solves the given PDE on $(0, \infty)$ with the correct boundary condition. The result for arbitrary initial conditions now follows by integrating over the densities w.r.t. μ_0 and Leibniz rule. $\square$

Given the explicit form of the density (3.16), we can read of the evolution of the mass process $(\mu_t[\mathbb{R}_+])_{t \geq 0}$.

Corollary 3.7. *We have*

$$\hat{\mu}_t[\mathbb{R}_+] = 1 - (1 - \hat{\mu}_0[\mathbb{R}_+])e^{-t}. \tag{3.20}$$

Proof. We integrate the PDE from Proposition 3.5 in space, which yields

$$\partial_t \|f(t, \cdot)\|_{L^1(\mathbb{R}_+)} = \int_0^\infty \big(z\, \partial_z^2 f(t, z) + z\, \partial_z f(t, z)\big)\, dz. \tag{3.21}$$
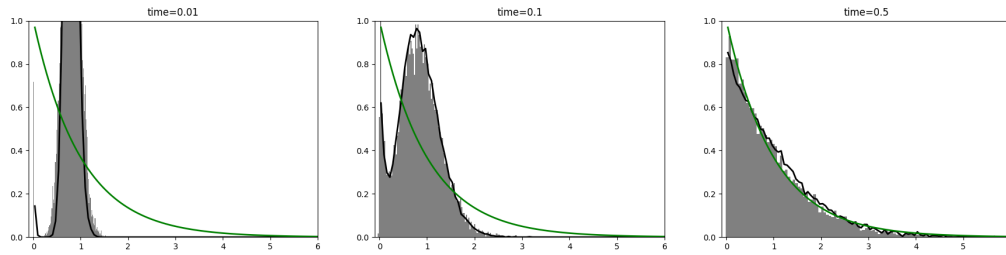
Figure 1: Simulations for both the inclusion process ($N = L = 1024$, $d = L^{-1/2} = \frac{1}{32}$) and the jump diffusion generated by $\hat{A}$ agree in accordance with Proposition 3.4. The black graph shows the density profile of embedded inclusion process (1.12) (1000 samples), whereas the grey histogram represents the density of the jump diffusion (10000 samples). Both profiles converge rapidly to the unit exponential density (green line), cf. Lemma 3.8. We considered an initial condition $\hat{\mu}_0 = \delta_z$, $z \simeq \frac{25}{32}$.

The right-hand side simplifies to the differential equation

$$d\alpha_t = (1 - \alpha_t)\, dt\,, \quad \alpha_0 = \mu_0[\mathbb{R}_+]\,, \tag{3.22}$$

using integration by parts, its solution is given by (3.20). $\qquad\square$

Furthermore, we summarise invariance and exponential ergodicity of $(\hat{\mu}_t)_{t\geq 0}$ in the following lemma:

Lemma 3.8. *The process $(\hat{Z}(t))_{t\geq 0}$ satisfies the following properties:*

1. *The exponential distribution $\mathrm{Exp}(1)$ is the unique invariant probability measure.*

2. *We have*

$$\left\| Law(\hat{Z}(t)) - \mathrm{Exp}(1)\right\|_{TV} \leq e^{-t}\,, \quad \forall t \geq 0\,. \tag{3.23}$$

Proof. The fact that the exponential distribution is invariant can be explicitly proven using integration by parts, but also follows directly from [2, Corollary 1]. The exponential ergodicity is a consequence of [2, Theorem 2]. The same result yields that the process is Harris recurrent, thus, the Exponential distribution is the unique invariant distribution. $\qquad\square$

Lastly, it is an easy consequence that the point mass on the Exponential distribution is an absorbing point for the measure valued process.

Corollary 3.9. $\hat{\mathbf{P}} = \delta_{\mathrm{Exp}(1)}$ *is reversible w.r.t. the dynamics induced by $\hat{\mathcal{L}}$.*

Proof. Let $H \in \mathcal{D}(\hat{\mathcal{L}})$ be of the form $H(\mu) = \mu(h_1)\cdots\mu(h_n)$. For simplicity we write $\nu := \mathrm{Exp}(1) \in \mathcal{M}_1(\mathbb{R}_+)$. Then

$$\hat{\mathcal{L}}H(\nu) = \sum_{j=1}^{n} \nu(\hat{A}h_i)\prod_{l\neq j}\nu(h_l) = 0\,, \tag{3.24}$$

as ν is invariant w.r.t. $\hat{A}$, cf. Lemma 3.8. Thus, $\hat{\mathbf{P}}(F\hat{\mathcal{L}}H) = 0 = \hat{\mathbf{P}}(H\hat{\mathcal{L}}F)$. $\qquad\square$

3.3 A natural extension of the Poisson-Dirichlet diffusion

In this section we prove Theorem 1.6 which states that $\mathcal{L}_\theta$ (which is equivalent to the Poisson-Dirichlet diffusion) has a limit as $\theta \to \infty$ under appropriate rescaling.

Proof of Theorem 1.6. The statement is once more a conclusion of the Trotter-Kurtz approximation. We state the essential steps for completeness. Let $H \in \mathcal{D}(\hat{\mathcal{L}})$ be of the form $\mu \mapsto \mu(h_1) \cdots \mu(h_n)$, $h_k \in \mathcal{D}(\hat{A})$, and define $H_\theta \in \mathcal{D}(\mathcal{L}_\theta)$ by

$$E \ni \mu \mapsto \mu(h_1(\theta \cdot)) \cdots \mu(h_n(\theta \cdot)), \tag{3.25}$$

where we interpreted the $h_k(\theta \cdot)$'s to be elements of $C^3([0,1])$. We have

$$\hat{A} h_k(\theta z) - \frac{1}{\theta} A_\theta h_k(\theta \cdot)(z) = z\big(\theta z\, h_k''(\theta z) + 2\, h_k'(\theta z)\big). \tag{3.26}$$

If h_k is a constant function, the r.h.s. vanishes. On the other hand, if $h_k \in C_c^3(\mathbb{R}_+)$, we can write

$$\sup_{\mu \in E} \big|\mu\big((\hat{A} h_k)(\theta \cdot)\big) - \tfrac{1}{\theta}\mu\big(A_\theta h_k(\theta \cdot)\big)\big| \leq C_{h_k} \sup_{\mu \in E} \mu\big(Z\, \mathbb{1}_{\theta Z \in \text{supp}(h_k)}\big)$$

$$= C_{h_k} \sup_{z \in [0,1]} \big\{z\, \mathbb{1}_{\theta z \in \text{supp}(h_k)}\big\},$$

where $Z \sim \mu$ and C_{h_k} is a finite constant, depending on h_k. The right hand side vanishes as $\theta \to \infty$. Along the same lines, we can show that the interaction term of $\frac{1}{\theta}\mathcal{L}_\theta$ disappears. Overall we conclude

$$\lim_{\theta \to \infty} \sup_{\mu \in E} \big|\tfrac{1}{\theta}\mathcal{L}_\theta H_\theta(\mu) - (\hat{\mathcal{L}} H)(S_\theta \mu)\big| = 0. \tag{3.27}$$

Again, the convergence of generators suffices to conclude weak convergence on the process level. $\qquad \square$

4 Discussion and outlook

We conclude this paper with a discussion of boundary cases in the setting considered and outline future directions as well as work in progress.

Throughout the paper we assumed that $\rho = \lim_{N,L \to \infty} N/L \in (0, \infty)$. However, the derived scaling limits do not depend on the actual value of ρ, as we study the distribution of mass after renormalising by N. As long as $N, L \to \infty$, our results extend to the boundary cases $\rho \in \{0, \infty\}$, up to a regime around $\rho = 0$ in the case $dL \to \infty$. In this regime we see an interesting transition of the clustering behaviour, cf. Lemma 4.1 below.

First consider $\theta < \infty$ and $d \to 0$, for which both cases $\rho \in \{0, \infty\}$ are covered by our proof. For $\rho = 0$, i.e. $N \ll L$, this is clear intuitively, as an increasing number of empty sites does not affect the dynamics since the total diffusivity per particle is $dL \to \theta$. On the other hand, if $\rho = \infty$ it may seem surprising that the number of sites L does not play a role (as long they are divergent). Here, the core lies in Lemma A.4, which states that for any thermodynamic limit, we can approximate configurations in $\overline{\nabla}$ (equivalently measures in E) by a sequence of particle configurations. Indeed, having a closer look at the proof of Lemma A.4, we see it is only necessary that a macroscopic excess mass of order $\sim N$ can be distributed uniformly over sites such that it is not visible under the macroscopic rescaling $\frac{1}{N}$. This is always the case as we can put $\sim N/L$ particles on sites (which might itself diverge), however, under macroscopic rescaling we have $\sim \frac{1}{N}\frac{N}{L} \to 0$.

Now assume $\theta = \infty$. In the case $\rho \in (0, \infty)$, the results in the present paper, and also [35], yield that a size-biased chosen chunk (at equilibrium) is approximately exponentially

distributed with mean $\simeq \frac{N}{dL}$. In fact, looking at the proof of Theorem 1.2, the result remains true as long as $dL/N \to 0$. This trivially holds when $\rho = \infty$, in which case cluster sizes live on the scale of order

$$\frac{N}{dL} \gg \frac{1}{d} . \tag{4.1}$$

On the other hand, if $N/L \to 0$ we have no control over dN (in contrast to $N/L \to \infty$, which implies $dN \to \infty$ since $dL \to \infty$). Therefore, for $\rho = 0$ the convergence in Theorem 1.2 remains true only if $dL/N \to 0$, i.e. $d \ll N/L$. Assume on the other hand that $dL/N \to \gamma \in (0, \infty]$, then we don't expect any clustering of particles on diverging scales. This is indeed the case, in fact, we see a finer structure emerging in the limit on scales of order one. Note that the following is independent of the underlying graph structure and holds more generally for irreducible and spatially homogeneous dynamics, cf. [35].

Lemma 4.1. *Assume that $N/L \to 0$, as $N, L \to \infty$, $d \to 0$ and $dL \to \infty$ such that $N/(dL) \to \gamma \in [0, \infty)$. Then*

$$\lim_{N/L \to 0} \pi_{L,N}[\tilde{\eta}_1 \in \cdot] = \mathrm{Geom}\left(\frac{1}{1+\gamma}\right) . \tag{4.2}$$

Here $\pi_{L,N}$ denotes the unique invariant distribution w.r.t. $\mathfrak{L}_{L,N}$.

Hence, for $\rho = 0$ and $dL \to \infty$, there is a critical scaling $N \sim dL$ below which the equilibrium measure does not exhibit clustering of particles on diverging scales.
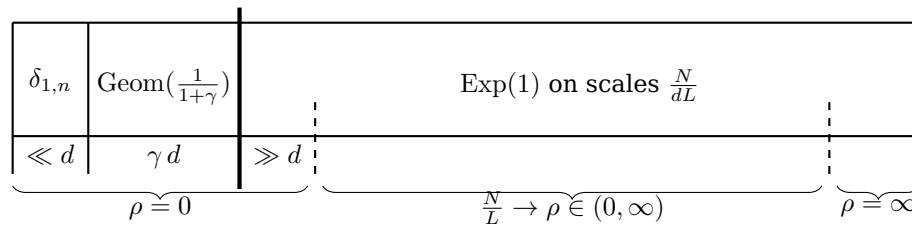

Figure 2: Graphical summary of the clustering of particles at equilibrium for the inclusion process when $dL \to \infty$. The distributions displayed describe the first size-biased marginal $\tilde{\eta}_1$ on the appropriate scale. Note particularly the transition from diverging scales to scales of order 1, when moving from the regime $\rho \gg d$ into $\rho \sim \gamma d$.

Before proving the above lemma, we require some notation and representations, which can be found in more detail in [35]. Recall that $\pi_{L,N}$ denotes the unique invariant distribution w.r.t. $\mathfrak{L}_{L,N}$ supported on $\Omega_{L,N}$, which is given explicitly by

$$\pi_{L,N}[d\eta] = \frac{1}{Z_{L,N}} \prod_{x=1}^{L} w_L(\eta_x) \, d\eta , \tag{4.3}$$

where $d\eta$ denotes the counting measure, $Z_{L,N}$ the appropriate partition function and w_L describing weights of the form

$$w_L(n) = \frac{\Gamma(n+d)}{n!\Gamma(d)} , \tag{4.4}$$

arising from the choice of transition rates in $\mathfrak{L}_{L,N}$. Moreover, the partition function can be explicitly written in terms of

$$Z_{L,N} = \frac{\Gamma(N+dL)}{N!\Gamma(dL)} . \tag{4.5}$$

Proof of Lemma 4.1. Recall from [35, Display (14)] that

$$\pi_{L,N}[\tilde{\eta}_1 = n] = \frac{L}{N} n \, w_L(n) \frac{Z_{L-1,N-n}}{Z_{L,N}} \,, \tag{4.6}$$

which is equal to zero if $n = 0$. Thus, without loss of generality let $n > 0$. We replace the terms in the previous display with the corresponding expressions in (4.4) and (4.5), which yields

$$\pi_{L,N}[\tilde{\eta}_1 = n] \simeq \frac{dL}{N} \frac{\Gamma(n+d)}{(n-1)! \, d\Gamma(d)} \frac{\Gamma(N-n+dL)}{\Gamma(N+dL)} N^n \,. \tag{4.7}$$

We analyse the remaining terms individually and conclude

$$\pi_{L,N}[\tilde{\eta}_1 = n] \simeq \frac{dL}{N} \left(\frac{\frac{N}{dL}}{1 + \frac{N}{dL}} \right)^n \to \frac{\gamma^{n-1}}{(1+\gamma)^n} \,, \quad \text{if } \gamma \in [0, \infty) \,. \tag{4.8}$$

This finishes the proof. $\qquad\square$

Our approach can further be adapted, with minimal changes in Lemma A.2, to cover the situation of fixed L and $d = d_N \to 0$ as $N \to \infty$. In this limit, the full mass fraction condenses on a single cluster site in a typical stationary configuration. This scaling has been considered in [5, 38] with additional assusmption $d_N \log N \to 0$ to study the metastable dynamics of the cluster location on a large time scale, and in [33] under the additional assumption $d_N N \to \infty$ to study emergence of the single cluster site on a slower time scale. The interaction and metastable motion of several spatially separated clusters on arbitrary finite lattices has been understood in full detail by now (see [37] and references therein). In our model clusters can exchange particles directly due to the complete graph geometry. Our results in both regimes imply that their rescaled sizes fluctuate diffusively due to mass exchanges on the time scale $1/d_L$. (Note that the generator (1.1) has already been speeded up by a factor $L - 1$ for complete graph dynamics). Only in the case $d_L L \to 0$ does a single stable cluster emerge, whose location should exhibit a uniform jump process on a time scale $1/d_L$. This is not covered by our results and could be approached by a combination of techniques in [5, 38] and [1].

In summary, our approach not only fully determines the clustering of particles on diverging scales in the inclusion process, but also describes the corresponding dynamics of the limiting Markov process in case of the complete graph. In this paper we focused on a (joint) thermodynamic limit $N, L \to \infty$ with $N/L \to \rho$ which is a natural approach in the context of interacting particle systems. It is equally possible to consider a two-step limit $N \to \infty$ towards a Wright–Fisher diffusion on a finite dimensional simplex, then taking $L \to \infty$ in a second step, arriving at the same scaling limits. This approach is often used in the context of population genetics [22, 14, 44]. Following the discussion above, this should be possible with the same size-biased techniques used in this paper, using a suitable scaling of d.

A natural next step, in view of the hydrodynamic limit Proposition 3.5, when $dL \to \infty$, is to study fluctuations around the equilibrium. A second moment calculation w.r.t. the stationary measure yields

$$\pi_{L,N} \left(\left(\hat{\mu}^{(\eta)}(h) - \text{Exp}(1)(h) \right)^2 \right) \simeq \frac{1}{dL} \pi_{L,N} \left(\tilde{\eta}_1 \frac{dL}{N} h(\tilde{\eta}_1 \frac{dL}{N})^2 \right) \to 0 \,, \tag{4.9}$$

for any $h \in C_b(\mathbb{R}_+)$. Hence, in order to see a non-trivial limit, we should investigate fluctuations of order $\sqrt{dL}$ by studying the limiting behaviour of

$$\sqrt{dL} \left(\left(\hat{\mu}_\# \eta^{(L,N)}(\tfrac{t}{dL}) \right)(h) - \text{Exp}(1)(h) \right) \,, \quad t \geq 0 \,, \tag{4.10}$$

where $(\eta^{(L,N)}(t))_{t\geq 0}$ denotes the inclusion process generated by $\mathfrak{L}_{L,N}$. Note that we slowed down time of the process, as indicated in Theorem 1.2. Due to decoupling of the size-biased marginals, the fluctuations are expected to be Gaussian.

Our approach should be robust towards perturbation of transition rates, as we do not require the explicit form of the partition function of the canonical distribution, cf. (4.5). However, the compact form of the limit dynamics and in particular the duality (1.18) are not expected to extend to more general models. Throughout the paper, we have focused on the one parameter family of Poisson-Dirichlet diffusions. There exists a two-parameter extension of the process, which was introduced in [43]. This process has gained a lot of attention over the past years [44, 26, 21, 14], just to name a few. It would be interesting to investigate the size-biased approach in this setting. To the best of the authors' knowledge, the two parameter process has only been studied when fixing finitely many locations/sites and observing the evolution of mass on them, see for example [27] and also the discussion in Section 2.3.1.

Furthermore, it would be interesting to investigate diffusion limits of the generalised version of the inclusion process with non-trivial bulk, studied in [11]. Numerical simulations and heuristic arguments suggest that the macroscopic phase evolves under the dynamics described in Theorem 1.1. At the same time one can observe a transfer of mass between the bulk and the condensate whose evolution is described by a system of ODEs, similar to Corollary 3.7.

A   Embeddings and approximations of particle configurations

First, we show that size-biased probability measures E are isomorphic to the Kingman simplex.

Lemma A.1. *The map $\mu^{(\cdot)} : \overline{\nabla} \to E$, cf. (1.4), is an isomorphism.*

Proof. First note that surjectivity is trivial due to the definition of E. Now, consider $p, q \in \overline{\nabla}$ such that $p \neq q$. Then there exists an index $i \in \mathbb{N}$ such that $p_i \neq q_i$ and $p_j = q_j$ for all $j < i$, without loss of generality assume $p_i > q_i$. Then

$$\mu^{(p)}([p_i, 1]) \geq \sum_{j=1}^{i} p_j > \sum_{j=1}^{i-1} q_j = \mu^{(q)}([p_i, 1]). \tag{A.1}$$

thus, $\mu^{(p)} \neq \mu^{(q)}$.

In order to show that the map $\mu^{(\cdot)}$ is continuous, consider a sequence of partitions $(p^{(n)})_{n\in\mathbb{N}}$ converging to p in $\overline{\nabla}$. Then for every $h \in C([0,1])$ (uniformly in n)

$$\left| \sum_{i=M}^{\infty} p_i^{(n)}(h(p_i^{(n)}) - h(0)) \right| \leq \sup_{0 \leq z \leq \frac{1}{M}} |h(z) - h(0)| \to 0, \quad \text{as } M \to \infty, \tag{A.2}$$

where we used the fact that $p_i \leq \frac{1}{i}$ for any $i \in \mathbb{N}$. This implies in particular $\mu^{(p^{(n)})} \xrightarrow{D} \mu^{(p)}$, recall that $\mu^{(p)}(h) = h(0) + \sum_{i=1}^{\infty} p_i(h(p_i) - h(0))$.

Continuity of the inverse is now immediate: let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence in E weakly converging to $\mu \in \mathcal{M}_1([0,1])$. Then we can identify each μ_n with a unique $p^{(n)}$ satisfying $\mu^{(p^{(n)})} = \mu_n$. Due to compactness of $\overline{\nabla}$, it suffices to consider convergent subsequences, say $(p^{(n_j)})_{j\in\mathbb{N}}$ with limit p. Thus, by assumption and continuity of $\mu^{(\cdot)}$

$$\mu_\# p^{(n_j)} \xrightarrow{D} \mu^{(p)} = \mu, \tag{A.3}$$

which particularly implies that $\mu \in E$. This implies that each accumulation point must agree with $(\mu^{(\cdot)})^{-1}(\mu) = p$. $\qquad\square$

Lemma A.2. *Let $h \in C(\mathbb{R}_+)$ and $\rho \in [0, \infty)$. Then for any $\zeta_L \to 0$*

$$\lim_{N/L \to \rho} \sup_{\eta \in \Omega_{L,N}} \left| \frac{1}{L} \sum_{x=1}^{L} h(\zeta_L \eta_x) - h(0) \right| = 0. \tag{A.4}$$

Proof. Let $\varepsilon > 0$ and $\eta \in \Omega_{L,N}$, then

$$\left| \frac{1}{L} \sum_{x=1}^{L} h(\zeta_L \eta_x) - h(0) \right| \leq \left| \frac{1}{L} \sum_{x=1}^{L} \left(h(\zeta_L \eta_x) - h(0) \right) \mathbb{1}_{\zeta_L \eta_x > \varepsilon} \right|$$

$$+ \left| \frac{1}{L} \sum_{x=1}^{L} \mathbb{1}_{\zeta_L \eta_x \leq \varepsilon} \left(h(\zeta_L \eta_x) - h(0) \right) \right|$$

$$\leq 2\|h\|_\infty \frac{1}{L} \sum_{x=1}^{L} \mathbb{1}_{\zeta_L \eta_x > \varepsilon} + \frac{1}{L} \sum_{x=1}^{L} \mathbb{1}_{\zeta_L \eta_x \leq \varepsilon} \left| h(\zeta_L \eta_x) - h(0) \right|$$

$$\leq 2\|h\|_\infty \frac{1}{L} \sum_{x=1}^{L} \mathbb{1}_{\zeta_L \eta_x > \varepsilon} + \sup_{0 \leq v \leq \varepsilon} \left| h(v) - h(0) \right| \frac{1}{L} \sum_{x=1}^{L} \mathbb{1}_{\zeta_L \eta_x \leq \varepsilon}.$$

The first term on the r.h.s. vanishes because the number of sites satisfying $\zeta_L \eta_x > \varepsilon$ is upper bounded by $\zeta_L N \varepsilon^{-1}$ (otherwise the total mass exceeds N). Thus,

$$\frac{1}{L} \sum_{x=1}^{L} \mathbb{1}_{\eta_x > \varepsilon \zeta_L^{-1}} \leq \frac{\zeta_L N}{\varepsilon L}. \tag{A.5}$$

The second term, on the other hand, is upper bounded by $\sup_{0 \leq v \leq \varepsilon} \left| h(v) - h(0) \right|$, which vanishes in the small ε-limit. Note that both upper bounds are uniform in $\Omega_{L,N}$. Now, taking first the thermodynamic limit $N/L \to \rho$ before taking $\varepsilon \to 0$, finishes the proof. $\square$

Remark A.3. Note that Lemma A.2 remains true if $\rho = \infty$ with a choice ζ_L satisfying $\zeta_L L / N \to 0$, cf. (A.5). In particular, the case $\zeta_L = dL/N$ is covered.

Indeed, we can show that, independently of the thermodynamic limit taken, any element in $\overline{\nabla}$ can be approximated by particle configurations:

Lemma A.4. *Let $N/L \to \rho \in [0, \infty]$, then for any $p \in \overline{\nabla}$ there exist $\eta^{(L,N)} \in \Omega_{L,N}$ such that*

$$\frac{1}{N} \hat{\eta}^{(L,N)} = \frac{1}{N} \eta^{(L,N)} \to p, \quad in \ \overline{\nabla}. \tag{A.6}$$

Proof. Consider $p \in \overline{\nabla}$ with $\|p\|_1 = 1 - \gamma$, we then define $\bar{\eta}_i^{(L,N)} := \lfloor p_i N \rfloor$, for $i \in \{1, \ldots, L\}$. Hence, there are

$$M_{L,N}(p) := N - \sum_{x=1}^{L} \bar{\eta}_x^{(L,N)} = \gamma N + \sum_{i=1}^{L} (p_i N - \lfloor p_i N \rfloor) \leq \gamma N + L \tag{A.7}$$

particles to spare. Thus, defining $\eta^{(L,N)} \in \Omega_{L,N}$ via

$$\eta_x^{(L,N)} := \bar{\eta}_x^{(L,N)} + \left\lfloor \frac{M_{L,N}(p)}{L} \right\rfloor + \mathbb{1}_{x \leq (M_{L,N}(p) \mod L)}, \quad x \in \{1, \ldots, L\}, \tag{A.8}$$

yields the desired approximation, since $M_{L,N}(p)/(N L) \to 0$, as $N, L \to \infty$. $\square$

Similarly, the embedding via $\hat{\mu}$, cf. (1.12), allows to approximate any probability measure on $\overline{\mathbb{R}}_+$ by particle configurations.

Lemma A.5. *Let $\rho \in [0, \infty]$ and $d = d(L)$ such that*

$$\frac{N}{L} \to \rho, \quad dL \to \infty \text{ and } \frac{dL}{N} \to 0.$$

Then for any $\mu \in \mathcal{M}_1(\overline{\mathbb{R}}_+)$ there exist $\eta^{(L,N)} \in \Omega_{L,N}$ such that

$$\hat{\mu}_\# \eta^{(L,N)} \xrightarrow{D} \mu. \tag{A.9}$$

Proof. We will see that it suffices to approximate discrete measures of the form

$$\alpha_0 \delta_0 + \alpha_\infty \delta_\infty + \sum_{i=1}^m \alpha_i \delta_{p_i} \in \mathcal{M}_1(\overline{\mathbb{R}}_+), \tag{A.10}$$

with $p_i \in (0, \infty)$, $1 \le i \le m$. Let ν be such a probability measure.

We explicitly construct configurations in $\Omega_{L,N}$ that converge to ν under the map

$$\hat{\mu}^{(\eta)} = \hat{\mu}_{L,N}^{(\eta)} = \sum_{x=1}^L \frac{\eta_x}{N} \delta_{dL \frac{\eta_x}{N}},$$

cf. (1.12), when considering the thermodynamic limit $N/L \to \rho \in [0, \infty]$.

First, we consider the point masses lying in $(0, \infty)$. For convenience, let us introduce

$$k_i := \left\lfloor \frac{N}{dL} p_i \right\rfloor \quad \text{and} \quad \#_i := \left\lfloor \frac{\alpha_i N}{k_i} \right\rfloor.$$

Note that $k_i \to \infty$, as $N/L \to \rho$, by assumption.

Now, let η' be the vector given by gluing together vectors $(k_i, \ldots, k_i) \in \mathbb{N}^{\#_i}$, $1 \le i \le m$. Note that

$$\frac{1}{L} \sum_{i=1}^m \#_i \le \sum_{i=1}^m \frac{\alpha_i N}{\frac{1}{2} k_i L} \le 2d \sum_{i=1}^m \frac{\alpha_i}{p_i},$$

and the r.h.s. vanishes because the sum is finite, recall $p_i > 0$. Hence, without loss of generality, we assume that the resulting vector η' lies in $\mathbb{N}^L$, as we can append zeros to the constructed vector until η' has length L. In particular, the relative number of empty sites converges to one.

It only remains to distribute the remaining particles to create the point masses at zero and infinity. Thus, far only $\#_\Sigma = \sum_{i=1}^m \#_i$ sites are occupied in η'. We start by adding $k_\infty := \lfloor \alpha_\infty N \rfloor$ particles onto the $\#_\Sigma + 1$-th position of η' (which is empty). This corresponds to the point mass at infinity.

Now, the number of allocated particles to η' is upper bounded by

$$k_\infty + \sum_{i=1}^m k_i \#_i \le \alpha_\infty N + \sum_{i=1}^m \alpha_i N = (1 - \alpha_0) N.$$

We distribute the remaining $k_0 := N - k_\infty - \sum_{i=1}^m k_i \#_i$ particles as uniform as possible between all empty sites of η' (there are $\#_0 := L - \#_\Sigma - 1$ many). This yields the configuration

$$\eta_x := \begin{cases} \eta'_x & \text{if } 1 \le x \le \#_\Sigma + 1 \\ \left\lfloor \frac{k_0}{\#_0} \right\rfloor + \mathbb{1}_{x \in \{\#_\Sigma + 2, \ldots, \#_\Sigma + 1 + (k_0 \mod \#_0)\}} & \text{otherwise.} \end{cases}$$

Because the number of non-empty sites in η' was relatively vanishing, the particles distributed on previously empty sites will correspond to the point mass at zero. Note that $\eta \in \Omega_{L,N}$ by construction.

Indeed, the constructed particle configuration η approximates the discrete measure arbitrarily well in the thermodynamic limit, as for every $f \in C_b(\bar{\mathbb{R}}_+)$ we have

$$\hat{\mu}^{(\eta)}(f) = \sum_{x=\#\Sigma+2}^{L} \frac{\eta_x}{N} f\left(\frac{dL}{N}\eta_x\right) + \frac{k_\infty}{N} f\left(\frac{dL}{N}k_\infty\right) + \sum_{i=1}^{m} \frac{k_i}{N} \#_i\, f\left(\frac{dL}{N}k_i\right), \qquad \text{(A.11)}$$

which converges to

$$\lim_{N/L \to \rho} \hat{\mu}^{(\eta)}(f) = \alpha_0 f(0) + \alpha_\infty f(\infty) + \sum_{i=1}^{m} \alpha_i f(p_i) = \nu(f).$$

Now, for every $\mu \in \mathcal{M}_1(\bar{\mathbb{R}}_+)$ there exists a sequence $(\nu_n)_{n\in\mathbb{N}}$ of measures of the form (A.10) such that $\nu_n \xrightarrow{D} \mu$. Moreover, each ν_n can be approximated by a sequence $(\eta_n^{(L,N)})_{L,N}$ following the above approach. Hence, we can construct a sequence of configurations with the desired property using a diagonal argument. For the sake of clarity we write out the details w.r.t the thermodynamic limit explicitly: we consider $N_j, L_j \to \infty$ such that $\lim_{j\to\infty} N_j/L_j = \rho > 0$. Then for every j we choose $n_j > n_{j-1}$ such that

$$d(\nu_{n_j}, \hat{\mu}_\# \eta_{n_j}^{(L_j,N_j)}) \leq 2^{-j} \qquad \text{(A.12)}$$

where $d(\cdot,\cdot)$ denotes an appropriate metric, e.g. the Lévy-Prokhorov metric. Thus,

$$d(\mu, \hat{\mu}_\# \eta_{n_j}^{(L_j,N_j)}) \leq d(\mu, \nu_{n_j}) + 2^{-j} \to 0, \quad \text{as } j \to \infty, \qquad \text{(A.13)}$$

which completes the proof. $\qquad\square$

B Convergence to a Fleming-Viot process

In this appendix we outline briefly how to prove convergence of the inclusion process to a Fleming-Viot process with with mutation operator A_{FV} (1.34), recall

$$A_{FV}h(u) = \theta \int_0^1 [h(v) - h(u)]\, dv.$$

The generator of the Fleming-Viot process, when applied to a cylindrical test function of the form $H(\nu) = \nu(h_1) \cdots \nu(h_n)$, is given by

$$\mathcal{L}_{FV} H(\nu) = 2 \sum_{1 \leq k < l \leq n} \left(\nu(h_k h_l) - \nu(h_k)\nu(h_l)\right) \prod_{j \neq k,l} \nu(h_j) \qquad \text{(B.1)}$$
$$+ \sum_{1 \leq k \leq n} \nu(A_{FV} h_k) \prod_{j \neq k} \nu(h_j).$$

Instead of considering the generator $\mathfrak{L}_{L,N}$ (1.1) on $\Omega_{L,N}$, it is more convenient to define the inclusion process on a state space keeping track of particle positions. More precisely, we define

$$S_{L,N} := \Lambda_L^N, \quad \text{with} \quad \Lambda_L := \{1, \ldots, L\}.$$

Now, the (labelled) inclusion process is described by the infinitesimal generator

$$\mathfrak{G}_{L,N} g(\sigma) = \sum_{i,j=1}^{N} [g(\sigma^{i\to\sigma_j}) - g(\sigma)] + d \sum_{i=1}^{N} \sum_{x\in\Lambda} [g(\sigma^{i\to x}) - g(\sigma)], \qquad \text{(B.2)}$$

where

$$
\sigma_j^{i \to x} := \begin{cases} x & \text{if } i = j\,, \\ \sigma_j & \text{if } i \neq j\,, \end{cases}
\tag{B.3}
$$

denotes the updated position after the i-th particle jumped onto site x. The generator $\mathfrak{G}_{L,N}$ characterises a Markov process on $S_{L,N}$ which we denote by $(\sigma^{(L,N)}(t))_{t \geqslant 0}$. We can recover the corresponding unlabelled particle configuration by

$$
\iota : S_{L,N} \to \Omega_{L,N} \quad \text{with} \quad \iota(\sigma)_x := \sum_{i=1}^{N} \mathbb{1}_{\sigma_i = x}\,,
$$

and in particular we have $\mathfrak{L}_{L,N} f(\iota(\sigma)) = \mathfrak{G}_{L,N} f(\iota(\cdot))(\sigma)$.

Again, we interpret particle configurations as probability measures on $[0,1]$. However, now we consider the embedding

$$
\nu^{(\cdot)} : \sigma \mapsto \frac{1}{N} \sum_{i=1}^{N} \delta_{\frac{\sigma_i}{L}} \in \mathcal{M}_1([0,1])\,,
\tag{B.4}
$$

where rescaled particle locations are encoded on the 'type space' $[0,1]$. Now the convergence of processes under the embedding (B.4) follows from approximation of generators in analogy to our main result. Again, we start with test functions of the form

$$
\nu^{(\sigma)}(h) = \frac{1}{N} \sum_{i=1}^{N} h\left(\frac{\sigma_i}{L}\right)\,, \quad h \in C^3([0,1])\,,
$$

in which case the action of $\mathfrak{G}_{L,N}$ reads

$$
\mathfrak{G}_{L,N} \nu^{(\sigma)}(h) = \sum_{i,j=1}^{N} \frac{1}{N} [h(\tfrac{\sigma_j}{L}) - h(\tfrac{\sigma_j}{L})] + d \sum_{i=1}^{N} \sum_{x \in \Lambda} \frac{1}{N} [h(\tfrac{x}{L}) - h(\tfrac{\sigma_i}{L})]\,.
$$

Because the first sum on the r.h.s. vanishes by symmetry, we are only left with

$$
\mathfrak{G}_{L,N} \nu^{(\sigma)}(h) = d \sum_{i=1}^{N} \sum_{x \in \Lambda} \frac{1}{N} [h(\tfrac{x}{L}) - h(\tfrac{\sigma_i}{L})] = dL \cdot \nu^{(\sigma)} \left(\frac{1}{L} \sum_{x \in \Lambda} h(\tfrac{x}{L}) - h \right)\,,
$$

which implies the uniform convergence

$$
\lim_{N/L \to \rho} \sup_{\sigma \in \Lambda^N} \left| \mathfrak{G}_{L,N}(\nu^{(\cdot)}(h))(\sigma) - \nu^{(\sigma)}(A_{FV} h) \right| = 0\,.
$$

By considering cylindrical test-functions, this convergence can be extended to a core of the Fleming-Viot process with generator $\mathcal{L}_{FV}$ (B.1) in full analogy to our main results in Section 2. We leave out further details.

Overall, this yields convergence of the (labelled) inclusion process in the following sense: if $\nu_{\#} \sigma^{(L,N)}(0) \xrightarrow{D} \nu_0$ then

$$
\left(\nu_{\#} \sigma^{(L,N)}(t) \right)_{t \geqslant 0} \xrightarrow{D} (\nu_t)_{t \geqslant 0}\,, \quad \text{in } D([0,\infty), \mathcal{M}_1([0,1]))\,,
$$

where $(\nu_t)_{t \geqslant 0}$ denotes the Fleming-Viot process generated by (B.1) with initial condition ν_0.

References

[1] Inés Armendáriz, Stefan Grosskinsky, and Michail Loulakis, *Metastability in a condensing zero-range process in the thermodynamic limit*, Probability Theory and Related Fields **169** (2017), 105–175. MR3704767

[2] Konstantin Avrachenkov, Alexey Piunovskiy, and Yi Zhang, *Markov processes with restart*, Journal of Applied Probability **50** (2013), no. 4, 960 – 968. MR3161367

[3] Mario Ayala, Gioia Carinci, and Frank Redig, *Condensation of SIP Particles and Sticky Brownian Motion*, Journal of Statistical Physics **183** (2021), no. 3, 1–42. MR4266161

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

[5] Alessandra Bianchi, Sander Dommers, and Cristian Giardinà, *Metastability in the reversible inclusion process*, Electronic Journal of Probability **22** (2017), 1–34. MR3698739

[6] Anton Bovier, Frank den Hollander, Saeda Marello, Elena Pulvirenti, and Martin Slowik, *Metastability of glauber dynamics with inhomogeneous coupling disorder*, arXiv preprint arXiv:2209.09827 (2022), 1–30.

[7] Jiarui Cao, Paul Chleboun, and Stefan Grosskinsky, *Dynamics of Condensation in the Totally Asymmetric Inclusion Process*, Journal of Statistical Physics **155** (2014), no. 3, 523–543. MR3192172

[8] 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

[9] Gioia Carinci, Cristian Giardinà, and Frank Redig, *Consistent particle systems and duality*, Electronic Journal of Probability **26** (2021), 1–31. MR4320949

[10] Yu-Xi Chau, Colm Connaughton, and Stefan Grosskinsky, *Explosive condensation in symmetric mass transport models*, Journal of Statistical Mechanics: Theory and Experiment **2015** (2015), no. 11, P11031.

[11] Paul Chleboun, Simon Gabriel, and Stefan Grosskinsky, *Poisson-Dirichlet asymptotics in condensing particle systems*, Electronic Journal of Probability **27** (2022), 1 – 35. MR4516309

[12] Paul Chleboun and Stefan Grosskinsky, *Condensation in stochastic particle systems with stationary product measures*, Journal of Statistical Physics **154** (2014), 432–465. MR3162548

[13] Bertrand Cloez and Marie-Noémie Thai, *Fleming-Viot processes: two explicit examples*, Latin American Journal of Probability and Mathematical Statistics **13** (2016), no. 1, 337. MR3487076

[14] Cristina Costantini, Pierpaolo De Blasi, Stewart N. Ethier, Matteo Ruggiero, and Dario Spanò, *Wright–Fisher construction of the two-parameter Poisson–Dirichlet diffusion*, The Annals of Applied Probability **27** (2017), no. 3, 1923–1950. MR3678488

[15] Donald A. Dawson and Kenneth J. Hochberg, *Wandering Random Measures in the Fleming-Viot Model*, The Annals of Probability **10** (1982), no. 3, 554–580. MR0659528

[16] Peter Donnelly and Thomas G. Kurtz, *Particle Representations for Measure-Valued Population Models*, The Annals of Probability **27** (1999), no. 1, 166–205. MR1681126

[17] Richard Durrett, *Probability Models for DNA Sequence Evolution*, Springer New York, 2008. MR2439767

[18] Stewart N. Ethier and Robert C. Griffiths, *The Infinitely-Many-Sites Model as a Measure-Valued Diffusion*, The Annals of Probability **15** (1987), no. 2, 515–545. MR0885129

[19] Stewart N. Ethier and Thomas G. Kurtz, *The Infinitely-Many-Alleles Model with Selection as a Measure-Valued Diffusion*, Lecture Notes in Biomathematics, Springer Berlin Heidelberg, 1987, pp. 72–86. MR0893637

[20] Stewart N. Ethier and Thomas G. Kurtz, *Fleming–Viot Processes in Population Genetics*, SIAM Journal on Control and Optimization **31** (1993), no. 2, 345–386. MR1205982

[21] Stewart N. Ethier, *A property of Petrov's diffusion*, Electronic Communications in Probability **19** (2014), 1–4. MR3262071

[22] 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

[23] Stewart N. Ethier and Thomas G. Kurtz, *Markov processes: characterization and convergence*, John Wiley & Sons, 2009. MR0838085

[24] Martin R. Evans and Tom Hanney, *Nonequilibrium statistical mechanics of the zero-range process and related models*, Journal of Physics A: Mathematical and General **38** (2005), no. 19, R195. MR2145800

[25] Shui Feng, *The Poisson-Dirichlet Distribution and Related Topics*, Springer Berlin Heidelberg, 2010. MR2663265

[26] Shui Feng, Wei Sun, Feng-Yu Wang, and Fang Xu, *Functional inequalities for the two-parameter extension of the infinitely-many-neutral-alleles diffusion*, Journal of Functional Analysis **260** (2011), no. 2, 399–413. MR2737405

[27] Noah Forman, Soumik Pal, Douglas Rizzolo, and Matthias Winkel, *Ranked masses in two-parameter fleming–viot diffusions*, Transactions of the American Mathematical Society **376** (2023), no. 02, 1089–1111. MR4531670

[28] 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

[29] Claude Godrèche and Jean-Marc Luck, *Condensation in the inhomogeneous zero-range process: an interplay between interaction and diffusion disorder*, Journal of Statistical Mechanics: Theory and Experiment **2012** (2012), no. 12, P12013. MR3036143

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

[31] Stefan Grosskinsky and Watthanan Jatuviriyapornchai, *Derivation of mean-field equations for stochastic particle systems*, Stochastic Processes and their Applications **129** (2019), no. 4, 1455–1475. MR3926563

[32] Stefan Grosskinsky, Frank Redig, and Kiamars Vafayi, *Condensation in the Inclusion Process and Related Models*, Journal of Statistical Physics **142** (2011), no. 5, 952–974. MR2781716

[33] Stefan Grosskinsky, Frank Redig, and Kiamars Vafayi, *Dynamics of condensation in the symmetric inclusion process*, Electronic Journal of Probability **18** (2013), 1–23. MR3078025

[34] 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

[35] 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

[36] Seonwoo Kim, *Second time scale of the metastability of reversible inclusion processes*, Probability Theory and Related Fields **180** (2021), no. 3-4, 1135–1187. MR4288339

[37] Seonwoo Kim, *Hierarchical structure of metastability in the reversible inclusion process: third time scale and complete characterization*, arXiv preprint arXiv:2308.13842 (2023), 1–59.

[38] Seonwoo Kim and Insuk Seo, *Condensation and metastable behavior of non-reversible inclusion processes*, Communications in Mathematical Physics **382** (2021), 1343–1401. MR4227174

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

[40] Stefano De Marco, *Smoothness and asymptotic estimates of densities for SDEs with locally smooth coefficients and applications to square root-type diffusions*, The Annals of Applied Probability **21** (2011), no. 4, 1282 – 1321. MR2857449

[41] Sean P. Meyn and R. L. Tweedie, *Stability of Markovian processes III: Foster–Lyapunov criteria for continuous-time processes*, Advances in Applied Probability **25** (1993), no. 3, 518–548. MR1234295

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

[43] Leonid A. Petrov, *Two-parameter family of infinite-dimensional diffusions on the Kingman simplex*, Functional Analysis and Its Applications **43** (2009), no. 4, 279–296. MR2596654

[44] Matteo Ruggiero and Stephen Walker, *Countable representation for infinite dimensional diffusions derived from the two-parameter Poisson-Dirichlet process*, Electronic Communications in Probability **14** (2009), 501–517. MR2564485

[45] Byron Schmuland, *A result on the infinitely many neutral alleles diffusion model*, Journal of Applied Probability **28** (1991), no. 2, 253–267. MR1104564

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

[47] Tokuzo Shiga, *Diffusion processes in population genetics*, Kyoto Journal of Mathematics **21** (1981), no. 1, 133–151. MR0606316

[48] Bartlomiej Waclaw and Martin R. Evans, *Explosive Condensation in a Mass Transport Model*, Physical Review Letters **108** (2012), no. 7.

[49] 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