ALEA, Lat. Am. J. Probab. Math. Stat. **22**, 825–857 (2025)
DOI: 10.30757/ALEA.v22-33

Condensation in Fleming–Viot particle systems with fast selection mechanism

Lucas Journel, Tony Lelièvre and Julien Reygner

Institut de mathématique, Université de Neuchâtel, Suisse
E-mail address: lucas.journel@unine.ch

CERMICS, École des Ponts and Inria, Marne-La-Vallée, France
E-mail address: tony.lelievre@enpc.fr

CERMICS, École des Ponts and Inria, Marne-La-Vallée, France
E-mail address: julien.reygner@enpc.fr

Abstract. We study the Fleming–Viot particle system in a discrete state space, in the regime of a fast selection mechanism, namely with killing rates which grow to infinity. This asymptotics creates a time scale separation which results in the formation of a condensate of all particles, which then evolves according to a continuous-time Markov chain with jump rates depending in a nontrivial way on both the underlying mutation dynamics and the relative speed of growth to infinity of the killing rate between neighbouring sites. We prove the convergence of the particle system and completely describe the dynamics of the condensate in the case where the number of particles is kept fixed, and partially in the case when the number of particles goes to infinity together with (but slower than) the minimal killing rate.

1. Introduction

1.1. *Fleming–Viot particle systems.* *Fleming–Viot particle systems*, sometimes also referred to as *Moran processes*, are population dynamics models in which n individuals are subjected to two mechanisms:

- individuals evolve in some state space as independent copies of a given continuous-time stochastic process (called the *mutation* process);
- they are killed at random times with a rate which depends on their current position, and are then instantaneously duplicated at the position of a uniformly chosen individual among the $n-1$ remaining ones (this is the *selection*, or *sampling*, mechanism).

Received by the editors July 2nd, 2024; accepted June 4th, 2025.
2010 *Mathematics Subject Classification.* 60J27, 60F99, 60G70.
Key words and phrases. Fleming-Viot particle systems, slow/fast dynamic, Quasi-stationary measure.
The work of T.L. is partially funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (project EMC2, grant agreement No 810367), and also from the Agence Nationale de la Recherche through the grant ANR-19-CE40-0010-01 (QuAMProcs). The work of J.R. is partially supported by the Agence Nationale de la Recherche through the grants ANR-19-CE40-0010-01 (QuAMProcs) and ANR-23-CE40-0003 (Conviviality).

When the mutation process is a random walk on $\mathbb{Z}^d$ and the killing rate $\lambda > 0$ is uniform on this space, the empirical measure of this particle system is known to converge, in the $n \to \infty$ limit and under a diffusive space-time rescaling for the mutation process, to the *Fleming–Viot superprocess* (Fleming and Viot, 1979), which is a measure-valued process on $\mathbb{R}^d$ introduced to describe the distribution of alleles in a population with a large number of possible genetic states (Etheridge, 2000). In particular, in this regime, the selection mechanism is sped up by a factor n with respect to the mutation process.

The asymptotic behaviour of the system is also of interest without space-time rescaling. Indeed, denoting by $(X_t)_{t\geqslant 0}$ the mutation process and τ the killing time, it is known that the empirical measure of the system generally converges, when $n \to \infty$, to the deterministic quantity $\mathbb{P}(X_t \in \cdot | \tau > t)$, which is the law of X_t conditioned on survival up to time t. This observation also holds true for *hard* killing mechanism, that is to say instantaneous killing when the process reaches some subset of the state space. It has motivated numerous developments on the study of propagation of chaos and hydrodynamics limit (Burdzy et al., 2000; Grigorescu and Kang, 2004; Löbus, 2009; Asselah et al., 2011; Villemonais, 2014; Oçafrain and Villemonais, 2017; Journel and Monmarché, 2022) as well as fluctuations (Cérou et al., 2020; Lelièvre et al., 2018), for both soft and hard killing mechanisms, in both discrete and continuous spaces.

From a numerical point of view, Fleming–Viot particle systems can be seen as Interacting Particle Systems (Del Moral and Miclo, 2003; Del Moral, 2004; Rousset, 2006) allowing to approximate the conditional distribution $\mathbb{P}(X_t \in \cdot | \tau > t)$, whose evolution is given by a nonlinear Fokker–Planck equation and is therefore nontrivial to study. The $t \to \infty$ limit of this conditional distribution is called a *quasistationary distribution* (QSD) for the process $(X_t)_{t\geqslant 0}$ (Collet et al., 2013), and plays a prominent role in population dynamics (Méléard and Villemonais, 2012), molecular dynamics (Le Bris et al., 2012; Di Gesù et al., 2016) or Monte Carlo methods (Pollock et al., 2020). For recent works on the existence and uniqueness of QSDs, as well as convergence of the conditional distribution $\mathbb{P}(X_t \in \cdot | \tau > t)$ towards the QSD, we refer to Champagnat and Villemonais (2023, 2020); Bansaye et al. (2022).

While the numerical interest of Fleming–Viot particle systems lies in the consistency of their $n \to \infty$ limits, in practice they can only be implemented with finitely many particles. In this context, one may expect their efficiency to depend on the intensity of the killing rate. More precisely, to approximate the QSD of the mutation process by the stationary distribution of the Fleming–Viot particle process, one needs the latter to reach stationarity on a shorter time scale than the typical killing time. This time scale separation is referred to as the *metastability* of the system (Olivieri and Vares, 2005); in this situation, quantitative estimates for Fleming–Viot particle systems associated with diffusions with hard killing were recently obtained in Journel and Monmarché (2025). The primary motivation of the present article is to study the behaviour of Fleming–Viot particle systems in *non-metastable* systems, namely for which killing occurs on a faster time scale than convergence to stationarity for the mutation process. We therefore consider the Fleming–Viot particle system in the asymptotic regime where the mutation dynamics remains fixed, but the killing rates grow to infinity uniformly, with a minimal rate $\underline{\lambda}$.

1.2. *The fast selection asymptotics.* In the mathematical biology literature, and more precisely in the study of evolutionary models, this time scale separation between slow mutation and fast selection is known to induce the convergence of microscopic Individual Based Models toward models of *monomorphic* populations, namely populations in which all individuals share the same phenotype, or trait (Champagnat, 2006; Champagnat and Lambert, 2007). The dynamics of the dominant trait, on the slow time scale, is then driven by mutations, which after an infinitesimally short competition, invade the population and replace the previous dominant trait. From the mathematical point of view, this is expressed by the convergence of the empirical distribution of the microscopic system to

a Dirac mass. The limiting dynamics of this Dirac mass, which therefore describes the consecutive dominant traits of the population, is then the main object of interest.

We prove that, in the fast selection asymptotics, the same phenomenon occurs for the Fleming–Viot particle system: its empirical measure converges to a Dirac mass, which evolves randomly on the slow time scale. We establish this *condensation* phenomenon, and show that the dynamics of the Dirac mass (which we shall call *condensate*) is determined in a nontrivial way by both the jump rates of the microscopic mutation dynamics, and by the relative speed of growth to infinity of the killing rates between neighbouring sites. We shall first address the case of a fast selection and constant number of particles, and then the case of both fast selection and large number of particles. In the context of convergence of microscopic Individual Based Models, these two asymptotic regimes were addressed in Champagnat and Lambert (2007) and Champagnat (2006), respectively. The convergence of a Fleming–Viot system toward a Dirac mass evolving according to a deterministic ordinary differential equation was also recently derived in Champagnat and Hass (2025), however in the different asymptotic of a *small* (but not slow) mutation. More generally, the condensation phenomenon is also known to occur in various interacting particle systems in statistical mechanics, such as the Zero Range Process (Grosskinsky et al., 2003; Waclaw and Evans, 2012). Another example which is actually even more closely related (up to a time change) to our model is the Inclusion Process for which condensation phenomena are proven in Grosskinsky et al. (2013); Chleboun et al. (2024); Bianchi et al. (2017) in the limit of an infinitely small mutation rate and infinitely large number of particles. The main difference with our work is that we precisely describe the limiting behavior for a strongly heterogeneous in space selection rate.

1.3. *Outline of the article.* The article is organised as follows. In Section 2, the precise mathematical setting is introduced, and the results are presented. The proofs of Theorems 2.2 and 2.3 which give the asymptotic behaviour in the high killing rate regime for a fixed number of particles are respectively presented in Sections 3 and 4. Finally, the proof of Theorem 2.5 which describes the asymptotic dynamics in the limit where both the killing rate and the number of particles go to infinity is provided in Section 5.

2. Setting and main results

After presenting the mathematical setting in Section 2.1, we will present our results first for a fixed number of particles in Section 2.2 and then for a number of particles growing to infinity with the killing rate in Section 2.3.

2.1. *Mathematical setting.* We consider the Fleming–Viot particle process associated with a mutation process which is a continuous-time Markov chain in some finite or countably infinite state space D, with jump rates $(q(x,y))_{x,y\in D}$, and a selection mechanism induced by killing rates $(\lambda(x))_{x\in D}$. We suppose that

$$Q := \sup_{x\in D} \sum_{y\in D} q(x,y) < \infty, \qquad \sup_{x\in D} \lambda(x) < \infty. \tag{2.1}$$

We denote by $\mathbf{X} = (\mathbf{X}_t)_{t\geqslant 0}$ the Fleming–Viot n-particle process in D^n. We are interested in the evolution of the empirical measure $\pi(\mathbf{X}_t)$ of the system, where for any vector $\mathbf{x} = (x_1, \dots, x_n) \in D^n$,

$$\pi(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i}. \tag{2.2}$$

Since the particles of the Fleming-Viot process are exchangeable, the Markov dynamics can be expressed in terms of the empirical measure $\pi(\mathbf{X}_t)$. We write $\mathcal{M}^1(D)$ for the set of probability

measures on D and $\mathcal{M}_n^1(D) \subset \mathcal{M}^1(D)$ for the set of empirical measures of n particles in D:

$$\mathcal{M}_n^1(D) = \{\pi(\mathbf{x}), \mathbf{x} \in D^n\} = \left\{ \left(\frac{k_x}{n}\right)_{x\in D} \middle| \forall x \in D, k_x \in \mathbb{N} \text{ and } \sum_{x\in D} k_x = n \right\}.$$

This is a countable set, and in this space, the empirical measure $(\pi(\mathbf{X}_t))_{t\geq 0}$ of the Fleming–Viot process is the continuous-time Markov chain with infinitesimal generator

$$L = L_m + L_s,$$

where L_m and L_s respectively correspond to mutation and selection. More precisely, introducing the notation

$$\ell^\infty(\mathcal{M}_n^1(D)) = \left\{u : \mathcal{M}_n^1(D) \to \mathbb{R}, \|u\|_\infty < \infty\right\}, \quad \|u\|_\infty = \sup_{\xi\in\mathcal{M}_n^1(D)} |u(\xi)|,$$

the operator L_m is the infinitesimal generator of the empirical measure of n independent random walks in D with jump rates $q(x,y)$: for all $f \in \ell^\infty(\mathcal{M}_n^1(D))$, for all $\xi \in \mathcal{M}_n^1(D)$,

$$L_m f(\xi) = \sum_{x,y\in D} n\xi(x)q(x,y)\left(f\left(\xi + \frac{\delta_y - \delta_x}{n}\right) - f(\xi)\right). \tag{2.3}$$

On the other hand, the operator L_s describes the evolution of the empirical measure of n particles experiencing only the selection mechanism, where a particle located in x is killed at rate $\lambda(x)$ and duplicated at the location of one of the $n-1$ remaining particles, chosen uniformly: for all $f \in \ell^\infty(\mathcal{M}_n^1(D))$, for all $\xi \in \mathcal{M}_n^1(D)$,

$$L_s f(\xi) = \sum_{x,y\in D} \frac{n^2}{n-1}\xi(x)\lambda(x)\xi(y)\left(f\left(\xi + \frac{\delta_y - \delta_x}{n}\right) - f(\xi)\right).$$

Under (2.1), the Markov generators L, L_m and L_s are bounded operators of $\ell^\infty(\mathcal{M}_n^1(D))$. Moreover, because $\mathcal{M}_n^1(D)$ is countable, the latter space is put in duality with the space

$$\ell^1(\mathcal{M}_n^1(D)) = \left\{u : \mathcal{M}_n^1(D) \to \mathbb{R}, \|u\|_1 < \infty\right\}, \quad \|u\|_1 = \sum_{\xi\in\mathcal{M}_n^1(D)} |u(\xi)|,$$

through the duality bracket

$$\forall u \in \ell^1(\mathcal{M}_n^1(D)), v \in \ell^\infty(\mathcal{M}_n^1(D)), \qquad \langle u, v\rangle = \sum_{\xi\in\mathcal{M}_n^1(D)} u(\xi)v(\xi).$$

We denote by L^*, L_m^* and L_s^* the respective adjoint operators of L, L_m and L_s with respect to this duality bracket. They are bounded operators of $\ell^1(\mathcal{M}_n^1(D))$.

As explained in Section 1, we are interested in the regime of a fast selection mechanism, namely

$$\underline{\lambda} := \inf_{x\in D} \lambda(x) \to \infty. \tag{2.4}$$

A key remark in our approach is that, under the selection dynamics, particles can only move to a site which already contains other particles, so that Dirac masses $\delta_x \in \mathcal{M}_n^1(D)$, $x \in D$ are absorbing states for this dynamics. In the context of evolutionary models, the absorption time of the selection mechanism is called the *fixation time*. Therefore, if this time occurs faster than the typical time between two consecutive mutations, on the time scale of the mutation dynamics, all particles condense in a single site, and the motion of the condensate, in the set

$$\Delta = \{\delta_x, x \in D\} \tag{2.5}$$

of Dirac masses on D, is triggered by mutations. The goal of this article is precisely to describe the dynamics of this condensate. We first address the case where n is kept fixed while $\underline{\lambda}$ satisfies (2.4) in Section 2.2, and then the case where both n and $\underline{\lambda}$ go to infinity in Section 2.3. In the former

case, we importantly use the fact that with probability going to 1, the fixation time always (on a finite time horizon) occurs before a new mutation. This allows us to employ functional analytic tools of averaging for slow/fast systems (Pavliotis and Stuart, 2008), but in the somehow not so common case where the fast dynamics has several absorbing states rather than a unique stationary distribution. In the case where both n and $\underline{\lambda}$ go to infinity, we directly show that the order of magnitude of absorption under the selection mechanism is $\underline{\lambda}/n$, and therefore that, on the time scale of the mutation dynamics, the condition

$$\frac{\underline{\lambda}}{n} \gg 1 \tag{2.6}$$

ensures condensation.

2.2. *Fixed number of particles.* In this section, the number of particles n is fixed, and we show that, in the limit (2.4), the trajectories of the empirical process $(\pi(\mathbf{X}_t))_{t\geq 0}$ converge towards trajectories in the set Δ defined in (2.5). This limiting dynamics can be precisely described as a continuous-time Markov chain in D.

In order to make precise the limiting regime (2.4), let us consider a family $(\lambda_r)_{r\in\mathbb{N}}$ of functions $D \to \mathbb{R}_+$, and introduce the notation

$$\underline{\lambda}_r = \inf_{x\in D} \lambda_r(x). \tag{2.7}$$

We set

$$\alpha_{x,y,r} = \frac{\lambda_r(y)}{\lambda_r(x)}, \tag{2.8}$$

and we denote by $\mathbf{X}^r = (\mathbf{X}^r_t)_{t\geqslant 0}$ the underlying Fleming–Viot n-particle processes. We work in this section in the following setting:

Assumption 2.1. The number n of particles is fixed, $\lim_{r\to\infty} \underline{\lambda}_r = +\infty$, and for all $x, y \in D$, there exists $\alpha_{x,y,\infty} \in [0,\infty]$ such that:

$$\lim_{r\to\infty} \alpha_{x,y,r} = \alpha_{x,y,\infty}. \tag{2.9}$$

As explained above, it is expected that $(\pi(\mathbf{X}^r_t))_{t\geq 0}$ converges to a process $(\delta_{Y_t})_{t\geq 0}$ where $(Y_t)_{t\geq 0}$ is a continuous-time Markov chain on D. Intuitively, the evolution of $(Y_t)_{t\geq 0}$ is driven by the following process. If all the particles are in $x \in D$ at a given time t (i.e. $Y_t = x$), then the rate at which one particle goes from x to y is given by $nq(x,y)$. After such a move of one particle from x to y, since the selection mechanism occurs on a faster time scale than the mutation process, all the particles will condense on one of the two states x or y before any other jumps occurs. Computing the probability that the particles will condense in y boils down to the computation of the probability that a one-dimensional random walk (describing the evolution of the number of particles in x) reaches 0 before n, starting from 1. We shall show in Lemma 3.6 below that this probability can be explicitly computed and is given by $(\alpha_{x,y,r} - 1)/(\alpha^n_{x,y,r} - 1)$. This justifies the introduction of the new jump rates:

$$\tilde{q}_r(x,y) = \begin{cases} nq(x,y)\dfrac{\alpha_{x,y,r}-1}{\alpha^n_{x,y,r}-1} & \text{if } \alpha_{x,y,r} \neq 1, \\ q(x,y) & \text{otherwise,} \end{cases} \tag{2.10}$$

as well as their $r \to \infty$ limit:

$$\tilde{q}_\infty(x,y) = \begin{cases} nq(x,y)\dfrac{\alpha_{x,y,\infty}-1}{\alpha^n_{x,y,\infty}-1} & \text{if } \alpha_{x,y,\infty} \neq 1, \\ q(x,y) & \text{otherwise.} \end{cases} \tag{2.11}$$

In the above, $\tilde{q}_\infty(x,y) = 0$ in the case $\alpha_{x,y,\infty} = \infty$.

Our first theorem concerns the limit of the time-marginal distribution of the empirical measure process under Assumption 2.1. For any random variable π in $\mathcal{M}_n^1(D)$, we identify the probability measure $\mathcal{L}aw(\pi)$ as an element of $\ell^1(\mathcal{M}_n^1(D))$ and therefore use the ℓ^1 norm to measure distances between such probability measures.

Theorem 2.2. *Under Assumption 2.1, for all $\eta \in \mathcal{M}_n^1(D)$, there exists a (deterministic) probability measure $\eta_\infty \in \mathcal{M}^1(D)$ such that if $(Y_t)_{t\geqslant 0}$ is a Markov process on D with jump rate $(\tilde{q}_\infty(x,y))_{(x,y)\in D}$ and initial distribution η_∞, then: for all $t > 0$,*

$$\lim_{r\to\infty} \|\mathcal{L}aw\left(\pi(\mathbf{X}_t^r)\right) - \mathcal{L}aw\left(\delta_{Y_t}\right)\|_1 = 0.$$

Moreover, if $\pi(\mathbf{X}_0^r) = \delta_z$ for some $z \in D$, then for all $T > 0$, there exists $C > 0$ such that if $(Y_t^r)_{t\geqslant 0}$ is a Markov process on D with jump rate $(\tilde{q}_r(x,y))_{(x,y)\in D}$ and initial condition $Y_0^r = z$, then:

$$\sup_{t\in[0,T]} \|\mathcal{L}aw\left(\pi(\mathbf{X}_t^r)\right) - \mathcal{L}aw\left(\delta_{Y_t^r}\right)\|_1 \leqslant \frac{C}{\underline{\lambda}_r}.$$

The probability measure η_∞ is described in Lemma 3.3. In the first statement of the theorem, the marginal distribution is taken at time $t > 0$. Indeed, if the Fleming–Viot particle system is initialized with a distribution that is not a Dirac distribution, namely with particles spread among several sites, then on the short time scale $1/\underline{\lambda}_r$, the selection mechanism makes all particles condense in a single site, whose distribution is precisely the probability measure η_∞. This shows that, in general, the convergence of $\mathcal{L}aw\left(\pi(\mathbf{X}_t^r)\right)$ to $\mathcal{L}aw\left(\delta_{Y_t}\right)$ cannot hold uniformly on $[0,T]$. However, if the initial distribution of the Fleming–Viot particle system is already condensed on a single site, then the second statement of the theorem shows that the distance between $\mathcal{L}aw\left(\pi(\mathbf{X}_t^r)\right)$ and $\mathcal{L}aw\left(\delta_{Y_t^r}\right)$ may be controlled uniformly on $[0,T]$, with a quantitative bound which is obtained by using averaging methods, as presented in Pavliotis and Stuart (2008, Chapter 16). The remaining distance, between $\mathcal{L}aw\left(\delta_{Y_t^r}\right)$ and $\mathcal{L}aw\left(\delta_{Y_t}\right)$, also vanishes (see Lemma 3.7), with a rate which depends on the convergence of $(\alpha_{x,y,r})_{x,y\in D}$ toward $(\alpha_{x,y,\infty})_{x,y\in D}$. The proof of Theorem 2.2 is given in Section 3.

We now state a trajectorial convergence result for the Fleming–Viot particle system towards the Δ-valued dynamics $(\delta_{Y_t})_{t\geqslant 0}$. For reasons which are detailed below, the Skorohod topologies for cadlag processes are not adapted to this system, and therefore we need to introduce another topology on the space of sample paths. To proceed, we first introduce the total variation norm $\|\cdot\|_{TV}$ defined by: for $\nu, \mu \in \mathcal{M}^1(D)$,

$$\|\mu - \nu\|_{TV} = \sum_{x\in D} |\mu(x) - \nu(x)|.$$

Note that, since D is a finite or countably infinite set, the weak topology on $\mathcal{M}^1(D)$ is metrized by the total variation norm. Next, for any $T > 0$, let $\mathcal{D}([0,T], \mathcal{M}^1(D))$ be the set of cadlag functions from $[0,T]$ to $\mathcal{M}^1(D)$. For $\pi, \tilde{\pi} \in \mathcal{D}([0,T], \mathcal{M}^1(D))$, we define the L^1_{TV} distance by:

$$\|\pi - \tilde{\pi}\|_{L^1_{TV}} = \int_0^T \|\pi_t(x) - \tilde{\pi}_t(x)\|_{TV} \mathrm{d}t = \int_0^T \sum_{x\in D} |\pi_t(x) - \tilde{\pi}_t(x)| \mathrm{d}t. \tag{2.12}$$

This defines a distance on $\mathcal{D}([0,T], \mathcal{M}^1(D))$, and the associated topology coincides in our case with the so-called Meyer-Zheng topology (Meyer and Zheng, 1984). We then have:

Theorem 2.3. *Under Assumption 2.1, for all $T > 0$, the process $(\pi(\mathbf{X}_t^r))_{t\in[0,T]}$ converges in distribution, in L^1_{TV}, to the process $(\delta_{Y_t})_{t\in[0,T]}$ introduced in Theorem 2.2.*

The proof of this theorem is given in Section 4.

Let us conclude this section with two remarks. First, due to the nature of the process, the convergence in Theorem 2.3 does not hold in the Skorokhod topology based on L^∞ norms, but in

the L^1_{TV} topology which is a weaker topology. Indeed, when a particle jumps from x to y, one observes motions from x to y and then back to x due to the selection mechanism, over a time scale of order $1/\underline{\lambda}_r$. This implies jumps of non vanishing heights (multiple of $1/n$) in the process $\pi(\mathbf{X}^r_t)$, and therefore, one cannot expect convergence to hold in the Skorokhod topology. Second, since the mapping $\pi \in \mathcal{D}([0,T], \mathcal{M}^1(D)) \mapsto \pi_t \in \mathcal{M}^1(D)$ is not continuous with respect to the L^1_{TV} distance, the convergence in Theorem 2.3 does not imply the convergence of the time-marginal distribution from Theorem 2.2.

2.3. *Infinitely many particles.* In this section, we are interested in the asymptotic regime where both the number of particles and the killing rates go to infinity. Therefore, in addition to the r-indexed family $(\lambda_r)_{r\in\mathbb{N}}$ of death rates, we introduce $(n_r)_{r\in\mathbb{N}} \subset \mathbb{N}^{\mathbb{N}}$ a sequence of numbers of particles. In order to keep the time scale separation (2.6) between the mutation process and the selection mechanism, we impose that $\underline{\lambda}_r/n_r \to \infty$, where $\underline{\lambda}_r$ remains defined by (2.7). Notice that this condition is necessary for the condensation phenomenon to occur: indeed, as already mentioned in Section 1, in the case $\underline{\lambda}_r = cn_r$, $\pi(\mathbf{X}^r_t)$ converges under a proper rescaling to the Fleming–Viot super-process and therefore exhibits a diffusive behaviour, see Etheridge (2000).

Last, we let $(\mathbf{X}^r_0)_{r\in\mathbb{N}}$ be a sequence of initial conditions, such that $\mathbf{X}^r_0 \in D^{n_r}$, and we will again denote by $\mathbf{X}^r = (\mathbf{X}^r_t)_{t\geqslant 0}$ the associated Fleming–Viot n_r-particle process.

2.3.1. *Nearly homogeneous killing rates.* Compared to the previous setting, we restrict ourselves to a situation where $\lambda_r(x)$ is a bounded perturbation of $\underline{\lambda}_r$. In this regime, the quantities $\alpha_{x,y,\infty}$ defined in the previous section are all equal to 1, and therefore the jump rates $\tilde{q}_\infty(x,y)$ in (2.11) are equal to $q(x,y)$ whatever the value of n. As a consequence, letting $n \to \infty$, one may expect the limiting jump rates to be simply $q(x,y)$, and this will indeed be the case as stated in Theorem 2.5 below.

Here is the precise setting of this section:

Assumption 2.4.

$$\sup_{r\in\mathbb{N}} \sup_{x,y\in D} |\lambda_r(x) - \lambda_r(y)| < \infty, \quad \lim_{r\to\infty} \frac{\underline{\lambda}_r}{n_r} = \infty, \quad \text{and} \quad \exists \eta \in \mathcal{M}^1(D),\ \pi(\mathbf{X}^r_0) \underset{r\to\infty}{\to} \eta \text{ weakly.}$$

Theorem 2.5. *Let us assume that Assumption 2.4 holds. Let $(Y_t)_{t\geqslant 0}$ be a Markov process on D with jump rate $(q(x,y))_{(x,y)\in D}$ and initial condition η. Then, the following convergence in law holds: for all $t > 0$,*

$$\pi(\mathbf{X}^r_t) \underset{r\to\infty}{\to} \delta_{Y_t},$$

where the convergence in law is considered with respect to the topology defined by the total variation distance on $\mathcal{M}^1(D)$. Moreover, if $\pi(\mathbf{X}^r_0) = \delta_z$ for some $z \in D$ and for all $r \in \mathbb{N}$, then for all $T > 0$, there exists $C > 0$ such that for all $F : \mathcal{M}^1(D) \to \mathbb{R}$ bounded and Lipschitz continuous, and for all $t \in [0,T]$:

$$|\mathbb{E}\left(F\left(\pi(\mathbf{X}^r_t)\right)\right) - \mathbb{E}\left(F\left(\delta_{Y_t}\right)\right)| \leqslant C\left(\|F\|_{\mathrm{Lip}} + \|F\|_\infty\right)\sqrt{\frac{n_r}{\underline{\lambda}_r}}.$$

Remark 2.6. Since the convergence in law of $\pi(\mathbf{X}^r_t)$ toward δ_{Y_t} in Theorem 2.5 is obtained for any sequence of initial conditions $\pi(\mathbf{X}^r_0)$ which converge weakly to η, it is easily seen that this result actually implies the convergence in the sense of the finite-dimensional distributions (in time) of the process $(\pi(\mathbf{X}^r_t))_{t\geqslant 0}$ to $(\delta_{Y_t})_{t\geqslant 0}$.

The proof is completely different in nature from the ones of Theorems 2.2 and 2.3, and relies on the Kolmogorov equation associated with the dynamics on $(\pi(\mathbf{X}^r_t))_{t\geq 0}$, see Section 5. Notice that the speed of convergence in Theorem 2.5 is not optimal. Indeed, choosing $n_r = n$ constant (which is allowed in Assumption 2.4), Theorem 2.5 yields a rate of convergence of order $1/\sqrt{\underline{\lambda}_r}$ whereas we know from Theorem 2.2 that it is of order $1/\underline{\lambda}_r$.

2.3.2. *A conjecture for general killing rates.* In the case where n_r and $\underline{\lambda}_r$ go to infinity in the regime (2.6), but without assuming that $\lambda_r(x)$ is a bounded perturbation of $\underline{\lambda}_r$ (namely without the first hypothesis in Assumption 2.4), we conjecture that the limiting behaviour of $(\pi(\mathbf{X}_t^r))_{t\geqslant 0}$ remains described by the $n\to\infty$ limit of the process $(Y_t)_{t\geqslant 0}$ identified in Theorem 2.2, which is obtained by taking the $\underline{\lambda}_r\to\infty$ limit of the Fleming–Viot particle system while keeping n fixed. Our proofs are not really tailored to show this statement: the method used in the proof of Theorem 2.2 fails in the case $n\to\infty$ as many estimates yield useless results in the limit of an infinite number of particles; similarly, the proof of Theorem 2.5 cannot deal with killing rates that are not of the same order of magnitude, as the proof of Lemma 5.1 requires $r\mapsto \sup_{x,y}|\lambda_r(x)-\lambda_r(y)|$ to be bounded. Nevertheless, should our conjecture hold true, the goal of this section is to derive a heuristic description of the process $(\overline{Y}_t)_{t\geqslant 0}$ obtained as the $n\to\infty$ limit of the process $(Y_t)_{t\geqslant 0}$.

Let us first recall that $(Y_t)_{t\geqslant 0}$ is the continuous-time Markov chain in D with jump rates

$$\tilde{q}_\infty(x,y)=\begin{cases} nq(x,y)\dfrac{\alpha_{x,y,\infty}-1}{\alpha_{x,y,\infty}^n-1} & \text{if } \alpha_{x,y,\infty}\neq 1,\\ q(x,y) & \text{otherwise,}\end{cases}$$

where

$$\alpha_{x,y,\infty}=\lim_{r\to\infty}\frac{\lambda_r(y)}{\lambda_r(x)}.$$

Assume that $x\in D$ is such that there exists $y\in D$ satisfying $q(x,y)>0$ and $\alpha_{x,y,\infty}<1$, so that (asymptotically, when r is large) the killing rate is strictly higher at x than at y. Then, when n is large, $\tilde{q}_\infty(x,y)\to\infty$, and the process Y should spend no time in x, making it disappear from the state space for $\overline{Y}$. Therefore, we define as the state space for $\overline{Y}$ the set

$$\overline{D}:=\{x\in D|\text{ for all } y\in D \text{ such that } q(x,y)>0, \text{ we have } \alpha_{x,y,\infty}\geqslant 1\}.$$

Notice that, when D is infinite, this set may be empty[1]: we discuss this point below.

Let us assume for the moment that $\overline{D}$ is nonempty. Then, starting from some $x\in\overline{D}$, the process Y can only perform a jump toward some $y\in D$ such that $q(x,y)>0$, and by the definition of $\overline{D}$, y must necessarily satisfy $\alpha_{x,y,\infty}\geqslant 1$. If $\alpha_{x,y,\infty}>1$, then when n is large, the rate $\tilde{q}_\infty(x,y)$ at which this jump is performed goes to 0, so in the $n\to\infty$ limit this jump will be discarded. Now if $\alpha_{x,y,\infty}=1$, the jump will be performed with rate $\tilde{q}_\infty(x,y)=q(x,y)$, independent from n. Then, two situations may occur:

- either $y\in\overline{D}$, in which case the dynamics restarts from y as is explained above;
- or $y\in D\setminus\overline{D}$, in which case Y will spend an infinitesimally small time in y and move to some $z\in D$ such that $q(y,z)>0$ and $\alpha_{y,z,\infty}<1$.

In the second case, there may be several z for which $q(y,z)>0$ and $\alpha_{y,z,\infty}<1$, and these z may themselves be in $D\setminus\overline{D}$: thus, the process Y will actually perform a sequence of infinitesimally short passages to states in $D\setminus\overline{D}$, which may be randomly chosen, until it finally reaches some state in $\overline{D}$ where it will stay a non-infinitesimal time. Let us denote by $\theta(y,z)$ the probability that, starting from $y\in D$, this sequence of infinitesimally short passages ends in the state $z\in\overline{D}$, with the convention that $\theta(y,z)=\mathbb{1}_{y=z}$ if $y\in\overline{D}$, that is to say when there is no infinitesimally short passage. Then, as an intermediate conclusion, the $n\to\infty$ limiting process $\overline{Y}$ should have jump rates

$$\overline{q}_\infty(x,z)=\sum_{y\in D}q(x,y)\mathbb{1}_{\alpha_{x,y,\infty}=1}\theta(y,z)$$

[1]Think of the example where $D=\{1,2,\dots\}$, $\lambda_r(x)=r(2-1/x)$, and $q(x,y)=\mathbb{1}_{y=x+1}$.

in $\overline{D}$. Actually, if D is infinite, it may happen that the sequence of infinitesimally short passages does not end: this is in particular the case for the example of Footnote 1, and more generally if

$$\sum_{z\in\overline{D}} \theta(y,z) < 1,$$

which includes the case where $\overline{D}$ is empty. In this case, the process Y actually leaves finite sets faster and faster when n grows, and therefore it does not have a limit when $n \to \infty$. From the point of view of the process $\overline{Y}$, this may be seen as an *explosion*, so an absorbing state $\{\infty\}$ may be added to the state space $\overline{D}$: this state is reached at rate

$$\overline{q}_\infty(x,\infty) = \sum_{y\in D} q(x,y)\mathbb{1}_{\alpha_{x,y,\infty}=1}\Big(1 - \sum_{z\in\overline{D}} \theta(y,z)\Big).$$

In the extreme case where $\overline{D}$ is empty, $\overline{Y}$ takes the value ∞ at all times.

To complete our description of the process $\overline{Y}$, we finally express the quantities $(\theta(y,z))_{y\in D, z\in\overline{D}}$ in terms of $(q(x,y))_{x,y\in D}$ and $(\alpha_{x,y,\infty})_{x,y\in D}$. Assume that the sequence of infinitesimally short passages starts at some state $y \in D \setminus \overline{D}$. Then the set

$$\mathcal{N}_y := \{z \in D : q(y,z) > 0 \text{ and } \alpha_{y,z,\infty} < 1\}$$

is nonempty. Let

$$\mathcal{N}_y^- := \operatorname*{argmin}_{z\in\mathcal{N}_y} \alpha_{y,z,\infty},$$

which represents the states which can be reached from y with lowest killing rates. In the $n \to \infty$ limit, the chain Y can only jump from y to states in $\mathcal{N}_y^-$, and the motion from y to $z \in \mathcal{N}_y^-$ occurs infinitesimally fast, with probability

$$w(y,z) := \frac{q(y,z)}{\sum_{z'\in\mathcal{N}_y^-} q(y,z')}.$$

Let us extend the notation $w(y,z)$ by setting $w(y,z) = 0$ if $z \notin \mathcal{N}_y^-$. Thus, if we denote by W the stochastic matrix on D given by

$$W(y,z) = \begin{cases} \mathbb{1}_{y=z} & \text{if } y \in \overline{D}, \\ w(y,z) & \text{if } y \in D \setminus \overline{D}, \end{cases}$$

we get that the sequence of infinitesimally short passages in $D \setminus \overline{D}$ has the law of $(Z_0, \dots, Z_\kappa)$ where $(Z_k)_{k\geqslant 0}$ is a Markov chain with transition matrix W and $\kappa = \inf\{k \geqslant 0 : Z_k \in \overline{D}\}$. We finally may conclude that, for any $y \in D$ and $z \in \overline{D}$,

$$\theta(y,z) = \mathbb{P}_y(\kappa < \infty, Z_\kappa = z) = \mathbb{1}_{y\in\overline{D}}\mathbb{1}_{y=z} + \mathbb{1}_{y\notin\overline{D}} \sum_{k=1}^{\infty} \left(W_{D\setminus\overline{D},D\setminus\overline{D}}^{k-1} W_{D\setminus\overline{D},\overline{D}}\right)(y,z),$$

where the notation $W_{D\setminus\overline{D},D\setminus\overline{D}}$, $W_{D\setminus\overline{D},\overline{D}}$ refers to obvious block decomposition. The probability for the sequence of infinitesimally short passages to not end is then exactly $\mathbb{P}_y(\kappa = \infty)$. See also Figure 2.1 for an example of how D may translate into $\overline{D}$.

3. Proof of Theorem 2.2

We let Assumption 2.1 hold in all this section. For all $x \in D$, write:

$$\alpha_r(x) = \frac{\lambda_r(x)}{\underline{\lambda}_r}. \tag{3.1}$$

For all $x \in D$, $\alpha_r(x) \geqslant 1$, and we write (let us recall the definition (2.3) of L_m):

$$L = L_m + \underline{\lambda}_r L_{s,0}^r, \tag{3.2}$$

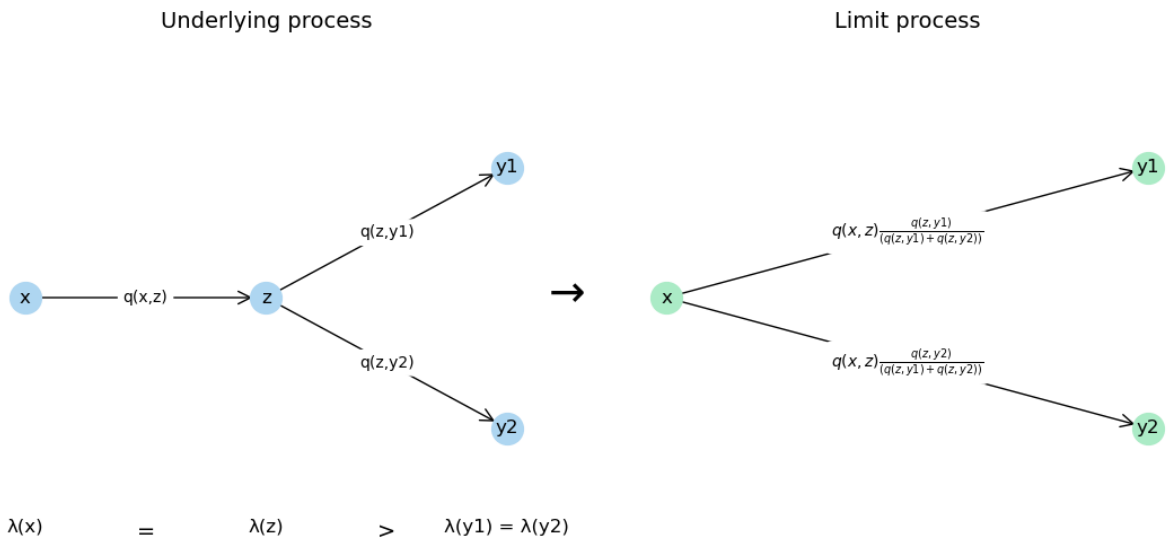

FIGURE 2.1. An example of D (left) and corresponding $\overline{D}$ (right).

where for $f \in \ell^\infty(\mathcal{M}_n^1(D))$ we denoted:

$$L_{s,0}^r f(\xi) = \sum_{x,y \in D} \frac{n^2}{n-1} \xi(x) \alpha_r(x) \xi(y) \left(f\left(\xi + \frac{\delta_y - \delta_x}{n} \right) - f(\xi) \right) = \frac{1}{\underline{\lambda}_r} L_s f(\xi). \tag{3.3}$$

For simplicity, we drop the dependence in r of $L_{s,0}^r$ and we simply write $L_{s,0}$. The subscript 0 refers to the fact that in the formal asymptotic expansion in powers of $\underline{\lambda}_r$ described below, $L_{s,0}$ is the operator appearing at the order $(\underline{\lambda}_r)^0$.

The proof of Proposition 2.2 uses averaging methods (Pavliotis and Stuart, 2008, Chapter 16). Let us first present the overall idea. Let us recall the notation (2.2) for the empirical measure. Denote for all $t \geq 0$ and $z \in D$

$$\pi_t(z) = \pi(\mathbf{X}_t^r)(z), \tag{3.4}$$

and let u be the law of the empirical measure process $(\pi_t)_{t \geq 0}$:

$$u(t, \xi) = \mathbb{P}(\pi_t = \xi), \qquad \forall \xi \in \mathcal{M}_n^1(D). \tag{3.5}$$

For all $t \geqslant 0$, $u(t, \cdot) \in \ell^1(\mathcal{M}_n^1(D))$, and we have the Kolmogorov equation:

$$\partial_t u = L^* u,$$

where we recall that L^* acts on $\ell^1(\mathcal{M}_n^1(D))$.

Let us present a formal argument in order to identify the limiting dynamics. If we assume an expansion of the form

$$u = u_0 + \frac{1}{\underline{\lambda}_r} u_1 + o\left(\frac{1}{\underline{\lambda}_r} \right),$$

with $\|u_0\|_1$ and $\|u_1\|_1$ bounded uniformly in $r \in \mathbb{N}$ and $t \geqslant 0$, then, using (3.2), a formal identification at order $(\underline{\lambda}_r)^1$ yields that u_0 satisfies

$$L_{s,0}^* u_0 = 0,$$

which shows that one expects the limiting dynamics to only take Dirac mass values. Indeed, for all $r \in \mathbb{N}$, $L_{s,0}$ is the generator of a Markov process that corresponds to a Fleming–Viot process with no mutation mechanism and only selection dynamics. We will refer in the sequel to this dynamics as the selection dynamics. Under this dynamics, each particle of the Fleming–Viot process is killed at rate α_r, and then branches uniformly at random over the surviving particles. Let $(\pi_t^0)_{t \geqslant 0}$ be

a Markov process with this generator $L_{s,0}$. Every Dirac measure is an absorbing point of such dynamics, and thus $\mathrm{Span}\left(\{\delta_{\delta_x}, x \in D\}\right) \subset \mathrm{Ker} L_{s,0}^*$, where δ_{δ_x} can also be seen as an element of $\ell^1(\mathcal{M}_n^1(D))$. Actually, writing (remember the notation (2.5))

$$\tau = \inf\left\{t \geqslant 0, \pi_t^0 \in \Delta\right\}, \tag{3.6}$$

we will show that $\tau < \infty$ almost surely (as an application of Lemma 3.1 below), and this will allow us to prove that, as a bounded operator from $\ell^1\left(\mathcal{M}_n^1(D)\right)$ to itself, $L_{s,0}^*$ satisfies

$$\overline{\mathrm{Span}\left(\{\delta_{\delta_x}, x \in D\}\right)} = \mathrm{Ker} L_{s,0}^*,$$

see Lemma 3.4.

We will then postulate the limiting dynamics on u_0 using our intuition on how Dirac masses evolve. Let us recall the intuition to derive these limiting dynamics: if the Dirac mass is at $x \in D$, at rate $nq(x,y)$ one particle will go from x to y. Since the selection dynamics is on a much faster time scale, we only have to consider this part of the dynamics, and the probability that the Dirac mass goes from x to y is given by the probability $p_{x,y}$ that a random walk on $\{1, \cdots, n\}$ (describing the evolution of the number of particles in x) reaches 0 before n, starting from 1, see Lemma 3.6. Therefore, the rates at which the Dirac mass will move from x to y is simply $nq(x,y)p_{x,y}$.

To formalize this reasoning, after having introduced u_0 following the expected limiting dynamics described above, we will define u_1 as a solution to the Poisson equation

$$L_{s,0}^* u_1 = \partial_t u_0 - L_m^* u_0 \tag{3.7}$$

which is the equation formally obtained at order $(\underline{\lambda}_r)^0$. The proof will then consist in proving that

$$u - u_0 - \frac{1}{\underline{\lambda}_r} u_1$$

goes to 0 as $r \to \infty$.

The remaining of this section is divided into four parts. In Section 3.1, we first present a general time-scale separation result, which will be used in Section 3.2 to provide a precise description of the early stage of the dynamics, when π_t^0 reaches Δ. Then, in Section 3.3, we will provide results on the well-posedness of Equation (3.7). We will then be in position to prove Theorem 2.2 in Section 3.4.

3.1. *A time-scale separation result.* In the selection/mutation mechanism, there exists a cascade of time-scales. The largest time scale is the mutation one, which is independent of $r \in \mathbb{N}$. Besides, we may define an equivalence relation on D by $x \sim y$ if and only if $0 < \alpha_{x,y,\infty} < \infty$. Then, each equivalence class associated with this equivalence relation gives a different time scale, as, if $\alpha_{x,y,\infty} = \infty$, then all particles in y will jump away from y to go to a site with a lower death rate before any particle in x have time to jump or move, with probability going to 1 as r goes to infinity. Lemma 3.1 below describes this phenomenon, and is a cornerstone in the proofs of the results under Assumption 2.1. In particular, if at a given time, the empirical process is not a Dirac mass, it will give us information on the time taken by the process to reach the set of Dirac measures and on how the particles will be redistributed to form a Dirac mass.

Lemma 3.1. *Let $a_r > 0$, $b_r \geqslant 0$, $L_r = a_r L_r^1 + b_r L_r^2$ be a jump Markov process generator on the state space $\mathcal{M}_n^1(D)$, such that L_r^1 is a selection dynamics with non-zero killing rates confined to some $\Theta \subset D$:*

$$\forall f \in \ell^\infty(\mathcal{M}_n^1(D)),\ L_r^1 f(\eta) = \sum_{x \in \Theta, y \in D} \frac{n^2}{n-1} \eta(x) \gamma_r(x) \eta(y) \left(f\left(\eta + \frac{\delta_y - \delta_x}{n}\right) - f(\eta)\right),$$

for some $\gamma_r \geqslant 1$, and such that the coefficients of L_r^2 are bounded from above uniformly in r:

$$\forall f \in \ell^\infty(\mathcal{M}_n^1(D)),\ L_r^2 f(\eta) = \sum_{\xi \in \mathcal{M}_n^1(D)} \beta_r^2(\eta, \xi)\left(f(\xi) - f(\eta)\right), \qquad \sum_{\xi \in \mathcal{M}_n^1(D)} \beta_r^2(\eta, \xi) \leqslant \bar{\mu},$$

for some $\bar{\mu} > 0$ and all $\eta \in \mathcal{M}_n^1(D)$. Let us assume that

$$\lim_{r\to\infty} \frac{b_r}{a_r} = 0,$$

and let us introduce $c_r > 0$ such that

$$\lim_{r\to\infty} \frac{c_r}{a_r} = \lim_{r\to\infty} \frac{b_r}{c_r} = 0.$$

Write

$$\mathbb{M}^1 = \left\{\xi \in \mathcal{M}_n^1(D), L_r^1 f(\xi) = 0,\, \forall f \in \ell^\infty(\mathcal{M}_n^1(D)),\, \forall r \in \mathbb{N}\right\},$$

and

$$\tau_1^r = \inf\left\{t \geqslant 0, \tilde{\pi}_t^r \in \mathbb{M}^1\right\},$$

where $\tilde{\pi}^r$ is a Markov process with generator L_r. Denote by S_1 the first jump of the process coming from the $b_r L_r^2$ part of the generator. Then the following holds:

$$\lim_{r\to\infty} \inf_{\xi\in\mathcal{M}_n^1(D)} \mathbb{P}_\xi\left(\tau_1^r < c_r^{-1} < S_1\right) = 1. \tag{3.8}$$

Moreover, if $L_r^2 = 0$, then there exist $C, c > 0$ such that for all $t \geqslant 0$ and $r \in \mathbb{N}$

$$\sup_{\xi\in\mathcal{M}_n^1(D),} \mathbb{P}_\xi\left(\tau_1^r \geqslant t\right) \leqslant Ce^{-ca_r t}. \tag{3.9}$$

This lemma will be applied to the Fleming–Viot generator (3.2), either with $\Theta = D$ or with Θ a subset of D such that the killing rates on Θ are much greater then the ones on $D \setminus \Theta$, in the sense that for all $x \in \Theta$, $y \notin \Theta$, $\alpha_{x,y,\infty} = 0$. In the latter case, the set $\mathbb{M}^1$ contains the Dirac measures, as well as the empirical measures of n particles with support in $D \setminus \Theta$.

Proof: Let $\bar{\pi}$ be a Markov process with generator L_r^1, defined using the same jump events as $\tilde{\pi}$, and

$$\bar{\tau}_1^r = \inf\left\{t \geqslant 0, \bar{\pi}_t^r \in \mathbb{M}^1\right\}.$$

On the event $\tau_1^r < S_1$, we have $\tau_1^r = a_r^{-1}\bar{\tau}_1^r$. Let us show that there exist $C, c > 0$ such that for all $t \geqslant 0$

$$\sup_{r\in\mathbb{N}} \sup_{\xi\in\mathcal{M}_n^1(D),} \mathbb{P}_\xi\left(\bar{\tau}_1^r \geqslant t\right) \leqslant Ce^{-ct}. \tag{3.10}$$

Fix $\xi \in \mathcal{M}_n^1(D)$, write

$$D_0 = \mathrm{supp}(\xi) := \{x \in D, \xi(x) > 0\},$$

and

$$\underline{\gamma}_r = \min\left\{\gamma_r(x), x \in D_0 \cap \Theta\right\}.$$

Since $\gamma_r(x) \geqslant 1$ for all $x \in D$, $\underline{\gamma}_r \geqslant 1$ for all $r \in \mathbb{N}$. Let $\underline{x}_r \in D$ be such that $\underline{\gamma}_r = \gamma_r(\underline{x}_r)$ if $D_0 \subset \Theta$, and let $\underline{x}_r \in D_0 \setminus \Theta$ otherwise. Let $\mathbf{X}$ be a process whose empirical measure is given by $\bar{\pi}$, and let $i_0 \in [\![1, n]\!]$ be such that $X_0^{i_0} = \underline{x}_r$. Denote

$$I = \left\{i,\, X_0^i \in D_0 \setminus \Theta\right\} \cup \{i_0\}, \quad n_0 = |I \setminus \{i_0\}|,$$

and write:

$$\mathcal{A}_r = \left\{\text{On } [0, 1/\underline{\gamma}_r], X^i,\, i \in I,\, \text{do not die, and all other die once and branch on } X^{i_0}\right\}.$$

We have that $\mathcal{A}_r \subset \{\bar{\tau}_1^r < 1\}$. Since the number of death events for $(X_s^i)_{0\leqslant s\leqslant 1/\underline{\gamma}_r}$ is upper (resp. lower) bounded by a Poisson random variable with parameter $\underline{\gamma}_r$ if $i \in I$ (resp. $i \notin I$), we have that:

$$\mathbb{P}_\xi\left(\mathcal{A}_r\right) \geqslant \frac{e^{-n}\left(1 - e^{-1}\right)^{n-n_0}}{n^{n-n_0}},$$

and hence, we get that

$$\inf_{r\in\mathbb{N}}\inf_{\xi\in\mathcal{M}_n^1(D)}\mathbb{P}_\xi\left(\bar\tau_1^r<1\right)\geqslant 1-\omega,$$

for some $\omega\in[0,1)$. Using the Markov property at time $t-1$, this implies that for all $r\in\mathbb{N}$, $\xi\in\mathcal{M}_n^1(D)$, and $t\geqslant 1$:

$$\mathbb{P}_\xi\left(\bar\tau_1^r\geqslant t\right)\leqslant\omega\mathbb{P}_\xi\left(\bar\tau_1^r\geqslant t-1\right),$$

and this implies

$$\sup_{r\in\mathbb{N}}\sup_{\xi\in\mathcal{M}_n^1(D)}\mathbb{P}_\xi\left(\bar\tau_1^r>t\right)\leqslant\omega^{\lfloor t\rfloor}\leqslant Ce^{-ct},$$

with $c=-\ln(\omega)>0$ and $C=\omega^{-1}$.

Writing $\beta_r^2(\eta)=\sum_{\xi\in\mathcal{M}_n^1(D)}\beta_r^2(\eta,\xi)$, S_1 is defined as:

$$S_1=\inf\left\{t\geqslant 0,E_1\leqslant\int_0^t b_r\beta_r^2(\tilde\pi_s^r)\mathrm{d}s\right\}\geqslant\inf\left\{t\geqslant 0,E_1\leqslant b_r\bar\mu t\right\}=b_r^{-1}\tilde S_1,$$

where E_1 is an exponential random variable with parameter 1. Hence, S_1 is bounded from below by $b_r^{-1}\tilde S_1$, where $\tilde S_1$ is an exponential random variable with parameter $\bar\mu$. Finally, using the bound (3.10) we have:

$$\begin{aligned}\mathbb{P}\left(\tau_1^r<c_r^{-1}<S_1\right)&=\mathbb{P}\left(a_r^{-1}\bar\tau_1^r<c_r^{-1}<S_1\right)\\&\geqslant 1-\mathbb{P}\left(\bar\tau_1^r\geqslant\frac{a_r}{c_r}\right)-\mathbb{P}\left(S_1<c_r^{-1}\right)\\&\geqslant 1-\mathbb{P}\left(\bar\tau_1^r\geqslant\frac{a_r}{c_r}\right)-\mathbb{P}\left(\tilde S_1<\frac{b_r}{c_r}\right)\underset{r\to\infty}{\to}1,\end{aligned}$$

and hence the limit (3.8). In the case where $L_r^2=0$, Equation (3.10) yields:

$$\mathbb{P}_\xi\left(\tau_1^r\geqslant t\right)=\mathbb{P}_\xi\left(\bar\tau_1^r\geqslant a_rt\right)\leqslant Ce^{-ca_rt},$$

which concludes the proof. □

3.2. *The initial condition.*

Now, we aim to address the issue of the initial condition. Given some initial condition $\eta\in\mathcal{M}_n^1(D)$, let us recall the definition

$$D_0=\mathrm{supp}(\eta)=\left\{x\in D,\eta(x)>0\right\},$$

and let us introduce

$$\Lambda=\left\{x\in D_0,\min_{y\in D_0}\alpha_{x,y,\infty}>0\right\}.\tag{3.11}$$

This set Λ corresponds to the subset of D_0 where, asymptotically in r, the order of magnitude of the killing rate is minimal, in the sense that for all $x,y\in\Lambda\times(D_0\setminus\Lambda)$,

$$\alpha_{x,y,\infty}=\lim_{r\to\infty}\frac{\lambda_r(y)}{\lambda_r(x)}=\infty.$$

In order to understand the initial condition of the limiting dynamics, two successive phenomena must be considered.

First, by applying Lemma 3.1, we will show that in the limit (2.4), all particles from $D_0\setminus\Lambda$ are redistributed onto Λ on a time scale of $1/\inf_{D_0\setminus\Lambda}\lambda_r$, with no movement from the particles in Λ. Moreover, the random configuration obtained in Λ after this redistribution is described as follows. Consider a Pólya's urn model with $|\Lambda|$ possible colors, which initially contains $n\eta(x)$ balls of the color $x\in\Lambda$. At each step, a ball is drawn uniformly at random and then returned to the urn along with an additional ball of the same color. After $n\eta(D_0\setminus\Lambda)$ draws, which corresponds to the number of particles originally in $D_0\setminus\Lambda$, one gets a random empirical measure $\nu\in\mathcal{M}_n^1(\Lambda)$.

Second, once this initial redistribution has occurred, the particle system will form a Dirac mass on a time scale of $1/\inf_{D_0}\lambda_r$ whose law is then given by the committor function defined below,

corresponding to a stable state of π^0 (recall that it corresponds to a Fleming-Viot dynamic with only selection) with initial condition ν.

Let us recall that π^0 denotes a Markov process with generator $L_{s,0}$, and the definition (3.6) of τ. Let us apply Lemma 3.1, with $L_r^2 = 0$, $a_r = 1$, $\Theta = \mathrm{supp}(\pi_0^0)$, $\gamma_r = \alpha_r$, so that $L_r^1 = L_{s,0}$ and $\tau_r^1 = \tau$. Equation (3.9) then yields that $\tau < \infty$ almost surely. Denote for $\eta \in \mathcal{M}_n^1(D)$:

$$\psi_x^r(\eta) = \mathbb{P}_\eta\left(\pi_\tau^0 = \delta_x\right). \tag{3.12}$$

The functions ψ_x^r are the so called committor functions, and Norris (1998, Theorem 3.3.1.) yields that they satisfy:

$$\psi_x^r(\delta_x) = 1,\ \psi_x^r(\delta_y) = 0\,\forall y \neq x,\ L_{s,0}\psi_x^r(\xi) = 0,\ \forall \xi \in \mathcal{M}_n^1(D) \setminus \Delta. \tag{3.13}$$

Remark that for all $\xi \in \Delta$, $f \in \ell^\infty\left(\mathcal{M}_n^1(D)\right)$, $L_{s,0}f(\xi) = 0$, so that we actually have $L_{s,0}\psi_x^r(\xi) = 0$, for all $\xi \in \mathcal{M}_n^1(D)$. The next lemma proves that the committor functions are the unique solution to (3.13) in $\ell^\infty\left(\mathcal{M}_n^1(D)\right)$.

Lemma 3.2. *Let $\Theta \subset D$ and $\gamma : \Theta \to [1, +\infty)$. Let L^1 be defined by: $\forall f \in \ell^\infty\left(\mathcal{M}_n^1(\Theta)\right)$,*

$$L^1 f(\eta) = \sum_{x,y\in\Theta} \frac{n^2}{n-1}\eta(x)\gamma(x)\eta(y)\left(f\left(\eta + \frac{\delta_y - \delta_x}{n}\right) - f(\eta)\right).$$

For all $x \in \Theta$, if $f \in \ell^\infty\left(\mathcal{M}_n^1(\Theta)\right)$ satisfies

$$f(\delta_x) = 1,\ f(\delta_y) = 0\,\forall y \neq x,\ L^1 f(\xi) = 0,\ \forall \xi \in \mathcal{M}_n^1(\Theta) \setminus \Delta, \tag{3.14}$$

then for all $\xi \in \mathcal{M}_n^1(\Theta)$

$$f(\xi) = \mathbb{P}_\xi\left(\pi_{\tau_1}^1 = \delta_x\right),$$

where $\tau_1 = \inf\left\{t \geqslant 0, \pi_t^1 \in \Delta\right\}$ and π^1 is a Markov process with generator L^1 and values in $\mathcal{M}_n^1(\Theta)$.

In particular, ψ_x^r defined by (3.12) are the unique solution to (3.13) in $\ell^\infty\left(\mathcal{M}_n^1(D)\right)$.

Proof: Let f be a solution to Equation (3.14). Let $\xi \in \mathcal{M}_n^1(\Theta)$ and π^1 be a Markov process with generator L^1 and initial condition ξ. The map $t \mapsto f(\pi_{\tau_1\wedge t}^1)$ is a martingale, and hence for all $t \geqslant 0$,

$$\mathbb{E}_\xi\left(f(\pi_{\tau_1\wedge t}^1)\right) = f(\xi). \tag{3.15}$$

Equation (3.9) from Lemma 3.1 yields that $\tau_1 < \infty$ almost surely, and since f is bounded, letting t go to infinity in (3.15) yields that $f(\xi) = \mathbb{P}_\xi\left(\pi_{\tau_1}^1 = \delta_x\right)$. □

We are now able to prove

Lemma 3.3. *For all $\eta \in \mathcal{M}_n^1(D)$, there exists $\eta_\infty \in \mathcal{M}^1(D)$ such that if π^0 is a Markov process with generator $L_{s,0}$ and initial condition η, and Y_0 is a random variable with law η_∞, then*

$$\lim_{r\to\infty} \|\mathcal{L}aw(\pi_\tau^0) - \mathcal{L}aw(\delta_{Y_0})\|_{TV} = 0,$$

where, we recall, $L_{s,0}$ defined by (3.3) implicitly depends on r, and τ is defined by (3.6).

In the statement of Theorem 2.2, η_∞ is the law of the initial condition Y_0.

Proof: Fix $\eta \in \mathcal{M}_n^1(D)$ and write $\bar{\eta}_r = \mathcal{L}aw(\pi_\tau^0) \in \mathcal{M}^1(\Delta)$. Denote $D_0(\eta) = \mathrm{supp}(\eta)$ and

$$\underline{\alpha}_r = \inf_{x\in D_0} \alpha_r(x),$$

where $\alpha_r(x)$ is defined in (3.1). Let us recall the definition (3.11) of Λ. The proof is divided into two parts. If $\Lambda = D_0$, then we can find the limit of $\bar{\eta}_r$ when $r \to \infty$ using the fact that $\bar{\eta}_r(\delta_x) = \psi_x^r(\eta)$ and that $(\psi_x^r(\eta))_{\eta\in\mathcal{M}_n^1(\Lambda)}$ is the unique solution to a Dirichlet equation on a finite set (see Lemma 3.2). If $\Lambda \neq D_0$, then using Lemma 3.1, we show that all particles from $D_0 \setminus \Lambda$ will first be distributed on Λ, without any particles from Λ moving, and then we can apply the argument from the case $D_0 = \Lambda$.

Let us suppose that $\Lambda = D_0$. Write $L_{s,0}^{D_0,r}$ the re-normalised generator of the selection dynamics

$$L_{s,0}^{D_0,r} f(\xi) = \sum_{x,y \in D_0} \frac{\alpha_r(x)}{\underline{\alpha}_r} \xi(x)\xi(y) \left(f\left(\xi + \frac{\delta_y - \delta_x}{n}\right) - f(\xi) \right),$$

defined for $f : \mathcal{M}_n^1(D_0) \to \mathbb{R}$. Denote also

$$\alpha_{D_0,\infty}(x) = \lim_{r\to\infty} \frac{\alpha_r(x)}{\underline{\alpha}_r} = \frac{1}{\min_{y\in D_0} \alpha_{x,y,\infty}},$$

which is well defined and lies in $[1,\infty)$ since $D_0 = \Lambda$, as well as

$$L_{s,0}^{D_0,\infty} f(\xi) = \sum_{x,y \in D_0} \alpha_{D_0,\infty}(x) \xi(x)\xi(y) \left(f\left(\xi + \frac{\delta_y - \delta_x}{n}\right) - f(\xi) \right).$$

The operator $L_{s,0}^{D_0,\infty}$ is the generator of a selection dynamics with death rates $\alpha_{D_0,\infty}$. We have that

$$\bar{\eta}_r(\delta_x) = \psi_x^r(\eta),$$

and for all $x \in D_0$ and $\xi \in \mathcal{M}_n^1(D_0)$,

$$L_{s,0}^{D_0,r} \psi_x^r(\xi) = 0. \tag{3.16}$$

Since $D_0 \times \mathcal{M}_n^1(D_0)$ is a finite set, we may consider $(\psi_x^\infty(\xi))_{x\in D_0, \xi \in \mathcal{M}_n^1(D_0)}$ a sub-sequential limit of $(\psi_x^r(\xi))_{x\in D_0,\xi\in\mathcal{M}_n^1(D_0)}$. Passing to the limit $r \to \infty$ in Equation (3.16), we get that for all $x \in D_0$ and $\xi \in \mathcal{M}_n^1(D_0)$:

$$L_{s,0}^{D_0,\infty} \psi_x^\infty(\xi) = 0,$$

as well as $\psi_x^\infty(\delta_y) = \mathbb{1}_{x\neq y}$. Lemma 3.2 then implies that

$$\psi_x^\infty(\xi) = \mathbb{P}_\xi \left(\pi_\tau^{0,\infty} = \delta_x \right),$$

where $\pi^{0,\infty}$ is a Markov process with generator $L_{s,0}^{D_0,\infty}$. Hence there is a unique sub-sequential limit, and $\psi_x^r(\xi)$ converge for all $\xi \in \mathcal{M}_n^1(D_0)$ and $x \in D_0$, and this is particularly true for $\xi = \eta$.

We now suppose that $\Lambda \neq D_0$. In this case, for all $x \in D_0 \setminus \Lambda$, $\alpha_{D_0,\infty} = \infty$ and $L^{D_0,\infty}$ is ill-defined. To resort to the previous case, write

$$S_1^\Lambda = \inf \{ t \geqslant 0, \text{ one particle starting from } \Lambda \text{ dies} \},$$

and

$$\tau_{\mathcal{M}_n^1(\Lambda)} = \inf \left\{ t \geqslant 0, \pi_t^0 \in \mathcal{M}_n^1(\Lambda) \right\}.$$

Using the notation of Lemma 3.1, write

$$\Theta = D_0 \setminus \Lambda, \quad a_r = \inf \{ \alpha_r(x)/\underline{\alpha}_r, x \in D_0 \setminus \Lambda \}, \quad b_r = 1, \quad L_r^2 = L_{s,0}^{D_0,r}.$$

In this case, we have $\mathbb{M}^1 = \Delta \cup \mathcal{M}_n^1(\Lambda)$, and because $\Lambda \neq D_0$, $a_r \underset{r\to\infty}{\to} \infty$, and Lemma 3.1 (with $c_r = \sqrt{a_r}$) yields that:

$$\lim_{r\to\infty} \mathbb{P}_\eta \left(S_1^\Lambda \leqslant \tau_{\mathcal{M}_n^1(\Lambda)} \right) = 0.$$

Now, using the strong Markov property at time $\tau_{\mathcal{M}_n^1(\Lambda)}$, write:

$$\mathbb{P}_\eta \left(\pi_\tau^0 = \delta_x \right) = \sum_{\xi \in \mathcal{M}_n^1(\Lambda)} \mathbb{P}_\eta \left(S_1^\Lambda > \tau_{\mathcal{M}_n^1(\Lambda)} \right) \mathbb{P}_\eta \left(\pi^0_{\tau_{\mathcal{M}_n^1(\Lambda)}} = \xi \middle| S_1^\Lambda > \tau_{\mathcal{M}_n^1(\Lambda)} \right) \mathbb{P}_\xi \left(\pi_\tau^0 = \delta_x \right) + R_r,$$

where a rest satisfying

$$R_r \leqslant \mathbb{P}_\eta \left(S_1^\Lambda \leqslant \tau_{\mathcal{M}_n^1(\Lambda)} \right) \underset{r\to\infty}{\to} 0.$$

From the case $D_0 = \Lambda$ we get that for all $\xi \in \mathcal{M}_n^1(\Lambda)$, $\mathbb{P}_\xi \left(\pi_\tau^0 = \delta_x \right)$ admits a limit as r goes to infinity. In addition,

$$\mathbb{P}_\eta \left(\pi^0_{\tau_{\mathcal{M}_n^1(\Lambda)}} = \pi \middle| S_1^\Lambda > \tau_{\mathcal{M}_n^1(\Lambda)} \right)$$

corresponds to the law of a Polya urn Markov chain with colors indexed by Λ, initial condition $n\eta|_\Lambda$, after $|D_0 \setminus \Lambda|$ draws, and is hence independent of r. This concludes the convergence of $\mathbb{P}_\eta\left(\pi^0_\tau = \delta_x\right)$ towards some $\bar\eta_\infty \in \mathcal{M}^1(\Delta)$. This concludes the proof since if $Y_0 \in D$ is such that $\mathcal{L}aw(\delta_{Y_0}) = \bar\eta_\infty$, then η_∞ is the law of Y_0. □

3.3. *The Poisson equation* (3.7). Let us now consider the Poisson equation (3.7). Before showing existence and uniqueness of a solution to this equation, let us describe the kernel of $L^*_{s,0}$ and $L_{s,0}$. Although this is not strictly necessary for the proof of Theorem 2.2, this makes the functional setting clearer. To proceed, let us recall that $L_{s,0}$ is defined as an operator of $\ell^\infty\left(\mathcal{M}^1_n(D)\right)$, and $L^*_{s,0}$ as an operator of $\ell^1\left(\mathcal{M}^1_n(D)\right)$. Define the weak-$*$ topology on $\ell^\infty\left(\mathcal{M}^1_n(D)\right)$ by the fact that a sequence of elements $(w_k)_{k\geqslant 1}$ weak-$*$ converges to w if $\langle v, w_k\rangle \to \langle v, w\rangle$ for any $v \in \ell^1(\mathcal{M}^1_n(D))$. Then $\ell^1\left(\mathcal{M}^1_n(D)\right)$ equipped with the (strong) topology induced by the norm $\|\cdot\|_1$, and $\ell^\infty(\mathcal{M}^1_n(D))$ equipped with the weak-$*$ topology, are each other's topological dual. Let us introduce

$$(\mathrm{Ker}L_{s,0})^\perp = \left\{v \in \ell^1\left(\mathcal{M}^1_n(D)\right) | \langle v, w\rangle = 0,\ \forall w \in \mathrm{Ker}L_{s,0}\right\}.$$

The following Lemma provides an explicit description of $(\mathrm{Ker}L_{s,0})^\perp$ and related vector spaces in terms of the committor functions ψ^r_x introduced above in (3.12) (in this Lemma and in the following $\overline{A}$ denotes the closure of a set A for the topology at hand).

Lemma 3.4. *We have*

$$(\mathrm{Ker}L_{s,0})^\perp = \{v \in \ell^1\left(\mathcal{M}^1_n(D)\right),\ \langle v, \psi^r_x\rangle = 0,\ \forall x \in D\},$$

as well as

$$\mathrm{Ker}L^*_{s,0} = \overline{\mathrm{Span}(\delta_{\delta_x}, x \in D)}.$$

In the case where D is infinite, the inclusion

$$\mathrm{Span}(\psi_x, x \in D) \subset \mathrm{Ker}L_{s,0}$$

is strict, and the kernel of $L_{s,0}$ is given by

$$\mathrm{Ker}L_{s,0} = \overline{\mathrm{Span}(\psi_x, x \in D)}^{w*},$$

where $\overline{A}^{w*}$ denotes the closure for the weak-$*$ topology of a set $A \subset \ell^\infty(\mathcal{M}^1_n(D))$.

Proof: For all $x \in D$, Lemma 3.2 yields that $\psi^r_x \in \mathrm{Ker}L_{s,0}$, so that:

$$(\mathrm{Ker}L_{s,0})^\perp \subset \{v \in \ell^1\left(\mathcal{M}^1_n(D)\right),\ \langle v, \psi^r_x\rangle = 0,\ \forall x \in D\}.$$

Let us now prove the reverse inclusion. Let $v \in \ell^1\left(\mathcal{M}^1_n(D)\right)$ such that $\langle v, \psi^r_x\rangle = 0$ for all $x \in D$, and $w \in \mathrm{Ker}L_{s,0}$ (and thus $w \in \ell^\infty\left(\mathcal{M}^1_n(D)\right)$. Let us recall that τ is the stopping time defined by (3.6). The map $t \mapsto w(\pi^0_{t\wedge\tau})$ is a martingale, and hence for all $t \geqslant 0$,

$$w(\xi) = \mathbb{E}_\xi\left(w(\pi^0_{\tau\wedge t})\right).$$

Since w is bounded, and $\tau < \infty$ almost surely, by letting t go to infinity, the dominated convergence theorem yields that for all $\xi \in \mathcal{M}^1_n(D)$:

$$w(\xi) = \sum_{x\in D} w(\delta_x)\psi_x(\xi), \tag{3.17}$$

and the sum is actually finite because $\psi_x(\xi) = 0$ if $x \notin \mathrm{supp}(\xi)$. We have:

$$\sum_{\xi\in\mathcal{M}^1_n(D)}\sum_{x\in D} |w(\delta_x)v(\xi)\psi_x(\xi)| = \sum_{\xi\in\mathcal{M}^1_n(D)} |v(\xi)| \sum_{x\in D} |w(\delta_x)|\psi^r_x(\xi) \leqslant \|v\|_1\|w\|_\infty < \infty.$$

Hence, the following computation holds:

$$\langle v, w\rangle = \sum_{\xi\in\mathcal{M}^1_n(D)} v(\xi)\sum_{x\in D} w(\delta_x)\psi_x(\xi) = \sum_{x\in D} w(\delta_x) \sum_{\xi\in\mathcal{M}^1_n(D)} v(\xi)\psi_x(\xi) = \sum_{x\in D} w(\delta_x)\langle v, \psi^r_x\rangle = 0,$$

which yields that $v \in (\mathrm{Ker} L_{s,0})^{\perp}$, and concludes the first point.

We already saw in the introduction of Section 3 that $\mathrm{Span}(\delta_{\delta_x}, x \in D) \subset \mathrm{Ker} L_{s,0}^*$. Since $L_{s,0}^*$ is a bounded operator, $\mathrm{Ker} L_{s,0}^*$ is closed and

$$\overline{\mathrm{Span}(\delta_{\delta_x}, x \in D)} \subset \mathrm{Ker} L_{s,0}^*.$$

Let us now prove the reverse inclusion. For $v \in \mathrm{Ker} L_{s,0}^*$, the map

$$t \mapsto \sum_{\xi \in \mathcal{M}_n^1(D)} v(\xi) \mathbb{P}_\xi \left(\pi_t^0 = \eta\right) \tag{3.18}$$

is constant equal to $v(\eta)$ (its value at $t = 0$), for all $\eta \in \mathcal{M}_n^1(D)$. Indeed, its time derivative is given by

$$\sum_{\xi \in \mathcal{M}_n^1(D)} v(\xi) L_{s,0} \mathbb{P}_\xi \left(\pi_t^0 = \eta\right) = \sum_{\xi \in \mathcal{M}_n^1(D)} L_{s,0}^* v(\xi) \mathbb{P}_\xi \left(\pi_t^0 = \eta\right) = 0.$$

Write for all $\eta \in \mathcal{M}_n^1(D)$ and $t \geqslant 0$,

$$\begin{aligned} \mathbb{P}_\xi \left(\pi_t^0 = \eta\right) - \mathbb{P}_\xi \left(\pi_\tau^0 = \eta\right) &= \mathbb{P}_\xi \left(\pi_t^0 = \eta, \tau > t\right) + \mathbb{P}_\xi \left(\pi_t^0 = \eta, \tau < t\right) - \mathbb{P}_\xi \left(\pi_\tau^0 = \eta\right) \\ &= \mathbb{P}_\xi \left(\pi_t^0 = \eta, \tau > t\right) - \mathbb{P}_\xi \left(\pi_\tau^0 = \eta, \tau > t\right), \end{aligned}$$

which yields

$$\sum_{\eta \in \mathcal{M}_n^1(D)} \left|\mathbb{P}_\xi \left(\pi_t^0 = \eta\right) - \mathbb{P}_\xi \left(\pi_\tau^0 = \eta\right)\right| \leqslant 2 \mathbb{P}_\xi \left(\tau > t\right). \tag{3.19}$$

Using the notations of Lemma 3.1, write:

$$\Theta = D, \quad a_r = 1, \quad \gamma_r = \alpha_r, \quad L_r^2 = 0.$$

Lemma 3.1 then yields that there exist $C, c > 0$ such that for all $t \geqslant 0$

$$\sup_{r \in \mathbb{N}} \sup_{\xi \in \mathcal{M}_n^1(D),} \mathbb{P}_\xi \left(\tau > t\right) \leqslant C e^{-ct}. \tag{3.20}$$

The bounds (3.19) and (3.20) yields that we may consider the $t \to \infty$ limit in the equality $v(\eta) = \sum_{\xi \in \mathcal{M}_n^1(D)} v(\xi) \mathbb{P}_\xi \left(\pi_t^0 = \eta\right)$ (see the map defined by (3.18)) and get:

$$v = \sum_{x \in D} v(\delta_x) \delta_{\delta_x}.$$

Since this sum converges in $\ell^1\left(\mathcal{M}_n^1(D)\right)$ as soon as $v \in \ell^1\left(\mathcal{M}_n^1(D)\right)$, $v \in \overline{\mathrm{Span}(\delta_{\delta_x}, x \in D)}$, which concludes the proof of the second point.

Define for all $\xi \in \mathcal{M}_n^1(D)$,

$$f(\xi) = \sum_{x \in D} \psi_x(\xi).$$

In this sum, there are at most n non-vanishing terms, and thus $f \in \ell^\infty\left(\mathcal{M}_n^1(D)\right)$. It is easily seen that $L_{s,0} f = 0$. If $g = \sum_{x \in D_0} a_x \psi_x(\xi)$ for some finite subset $D_0 \subset D$ and $a_x \in \mathbb{R}$, because D is infinite, $D_0 \neq D$ and we have that $\|f - g\|_\infty \geqslant 1$, yielding that

$$f \in \mathrm{Ker} L_{s,0} \setminus \overline{\mathrm{Span}(\psi_x, x \in D)}.$$

Finally, for all $w \in \mathrm{Ker} L_{s,0}$, we have shown that w satisfies (3.17), and the sum in the right-hand side of (3.17) converges for the weak-$*$ topology, which proves the last point. □

We are now ready to solve the Poisson equation (3.7), relying on the next lemma. For any $v \in \ell^1(\mathcal{M}_n^1(D))$ and $s \geqslant 0$, let us define $v e^{s L_{s,0}}$ by:

$$\forall \eta \in \mathcal{M}_n^1(D), \qquad v e^{s L_{s,0}}(\eta) = \sum_{\xi \in \mathcal{M}_n^1(D)} v(\xi) \mathbb{P}_\xi(\pi_s^0 = \eta), \tag{3.21}$$

which defines an element of $\ell^1(\mathcal{M}_n^1(D))$.

Lemma 3.5. *For all $v \in (\mathrm{Ker} L_{s,0})^\perp (\subset \ell^1(\mathcal{M}_n^1(D)))$, there exists a unique $w \in (\mathrm{Ker} L_{s,0})^\perp$ such that*

$$L_{s,0}^* w = v, \tag{3.22}$$

given by:

$$w = -\int_0^\infty v e^{sL_{s,0}} \mathrm{d}s.$$

Moreover, there exists $C > 0$ independent of r and v such that

$$\|w\|_1 \leqslant C \|v\|_1.$$

Proof: Let us start with the existence. Write, for $\eta \in \mathcal{M}_n^1(D)$

$$w(\eta) = -\left(\int_0^\infty v e^{sL_{s,0}} \mathrm{d}s\right)(\eta) = -\int_0^\infty \sum_{\xi \in \mathcal{M}_n^1(D)} v(\xi) \mathbb{P}_\xi\left(\pi_s^0 = \eta\right) \mathrm{d}s.$$

We first show that w is well-defined. Let us recall that the stopping time τ is defined in (3.6). Since $v \in (\mathrm{Ker} L_{s,0})^\perp$, Lemma 3.4 yields that $\langle v, \psi_x^r \rangle = 0$ for all $x \in D$, and we have

$$\sum_{\xi \in \mathcal{M}_n^1(D)} v(\xi) \mathbb{P}_\xi\left(\pi_\tau^0 = \delta_x\right) = \sum_{\xi \in \mathcal{M}_n^1(D)} v(\xi) \psi_x^r(\xi) = 0.$$

Since, if $\eta \notin \Delta$, $\mathbb{P}_\xi\left(\pi_\tau^0 = \eta\right) = 0$, we have that

$$\sum_{\xi \in \mathcal{M}_n^1(D)} v(\xi) \mathbb{P}_\xi\left(\pi_s^0 = \eta\right) = \sum_{\xi \in \mathcal{M}_n^1(D)} v(\xi) \left(\mathbb{P}_\xi\left(\pi_s^0 = \eta\right) - \mathbb{P}_\xi\left(\pi_\tau^0 = \eta\right)\right).$$

The bounds (3.19) and (3.20) imply that w is well-defined, because $v \in \ell^1(\mathcal{M}_n^1(D))$, and we have

$$\|w\|_1 \leqslant 2C \int_0^\infty e^{-cs} \mathrm{d}s \|v\|_1 = \frac{2C}{c} \|v\|_1.$$

The Kolmogorov equations yield that $L_{s,0}^* \mathbb{P}_\xi\left(\pi_s^0 = \cdot\right) = \partial_s \mathbb{P}_\xi\left(\pi_s^0 = \cdot\right)$, and hence for all $\eta \in \mathcal{M}_n^1(D)$:

$$\begin{aligned} L_{s,0}^* w(\eta) &= -\int_0^\infty \sum_{\xi \in \mathcal{M}_n^1(D)} v(\xi) L_{s,0}^* \mathbb{P}_\xi\left(\pi_s^0 = \cdot\right)(\eta) \, \mathrm{d}s \\ &= -\int_0^\infty \sum_{\xi \in \mathcal{M}_n^1(D)} v(\xi) \partial_s \mathbb{P}_\xi\left(\pi_s^0 = \eta\right) \mathrm{d}s \\ &= -\int_0^\infty \partial_s \left(\sum_{\xi \in \mathcal{M}_n^1(D)} v(\xi) \mathbb{P}_\xi\left(\pi_s^0 = \eta\right)\right) \mathrm{d}s. \end{aligned}$$

We already saw that

$$\lim_{s \to \infty} \sum_{\xi \in \mathcal{M}_n^1(D)} v(\xi) \mathbb{P}_\xi\left(\pi_s^0 = \eta\right) = 0,$$

and hence $L_{s,0}^* w(\eta) = v(\eta)$.

Let $f \in \mathrm{Ker} L_{s,0}$. We have that

$$\langle w, f \rangle = -\int_0^\infty \sum_{\xi, \eta \in \mathcal{M}_n^1(D)} v(\xi) \mathbb{P}_\xi\left(\pi_s^0 = \eta\right) f(\eta) \mathrm{d}s = -\int_0^\infty \sum_{\xi \in \mathcal{M}_n^1(D)} v(\xi) \left(e^{sL_{s,0}} f\right)(\xi) \mathrm{d}s = 0,$$

so that $w \in (\mathrm{Ker} L_{s,0})^\perp$, which concludes the existence.

To prove uniqueness, fix $w \in \mathrm{Ker} L_{s,0}^* \cap (\mathrm{Ker} L_{s,0})^\perp$. The fact that $L_{s,0}^* w = 0$ yields that for all $\eta \in \mathcal{M}_n^1(D)$, $t \geqslant 0$:

$$w(\eta) = \sum_{\xi \in \mathcal{M}_n^1(D)} w(\xi) \mathbb{P}_\xi \left(\pi_t^0 = \eta\right).$$

Letting $t \to \infty$, because $w \in \ell^1\left(\mathcal{M}_n^1(D)\right)$, Inequalities (3.19), and (3.20) then imply that

$$w(\eta) = \sum_{\xi \in \mathcal{M}_n^1(D)} w(\xi) \mathbb{P}_\xi \left(\pi_\tau^0 = \eta\right) \ \forall \eta \in \mathcal{M}_n^1(D).$$

In particular, $w(\eta) = 0$ for all $\eta \in \mathcal{M}_n^1(D) \setminus \Delta$, and thus we may write $w(\eta) = \sum_{x \in D} w(\delta_x) \delta_{\delta_x}(\eta)$, where the equality holds in $\mathcal{M}_n^1(D)$. We have for all $x \in D$:

$$\sum_{\eta \in \mathcal{M}_n^1(D)} \sum_{y \in D} |w(\delta_y) \delta_{\delta_y}(\eta) \psi_x^r(\eta)| = \sum_{y \in D} |w(\delta_y)| \psi_x^r(\delta_y) \leqslant \|w\|_1 < \infty,$$

and hence the following computation holds:

$$\begin{aligned} 0 = \langle w, \psi_x^r \rangle &= \sum_{\eta \in \mathcal{M}_n^1(D)} w(\eta) \psi_x^r(\eta) \\ &= \sum_{\eta \in \mathcal{M}_n^1(D)} \left(\sum_{y \in D} w(\delta_y) \delta_{\delta_y}(\eta) \right) \psi_x^r(\eta) = \sum_{y \in D} w(\delta_y) \psi_x^r(\delta_y) = w(\delta_x). \end{aligned}$$

Thus, $w = 0$, which concludes for the uniqueness. □

In the proof of Theorem 2.2, in order to apply Lemma 3.5, we need to show that for all $x \in D$,

$$\langle \partial_t u_0 - L_m^* u_0, \psi_x^r \rangle = 0.$$

To this end, let us introduce

$$\delta_{x;y} = \frac{n-1}{n} \delta_x + \frac{1}{n} \delta_y \in \mathcal{M}_n^1(D).$$

The next lemma provides an explicit formula for $\psi_x^r(\delta_{x;y})$. The fact that we only need the values of the committor functions on the set $\Delta \cup \{\delta_{x;y}; x, y \in D\}$ is a consequence of the fact that, as explained above, in the limiting regime, the selection dynamics is much faster than the mutation dynamics, so that the support of π_t is concentrated on $\{x, y\}$ in the transition from δ_x to δ_y. After one particle jumped from x to y, the probability that all the particles go back to x is $\psi_x^r(\delta_{x;y})$, and the probability that all the particles jumps to y is $\psi_y^r(\delta_{x;y})$.

Lemma 3.6. *For all $x, y \in D$, $x \neq y$, the committor functions ψ_x^r satisfy:*

$$\psi_x^r(\delta_{x;y}) = \alpha_{x,y,r} \frac{\alpha_{x,y,r}^{n-1} - 1}{\alpha_{x,y,r}^n - 1}, \qquad \psi_x^r(\delta_{y;x}) = \alpha_{x,y,r}^{n-1} \frac{\alpha_{x,y,r} - 1}{\alpha_{x,y,r}^n - 1},$$

where $\alpha_{x,y,r}$ is defined in (2.8).

Proof: The selection dynamics cannot extend the support of a measure, hence if $\mathrm{supp}\, \pi_0^0 = \{x, y\}$, then $\mathrm{supp}\, \pi_t^0 \subset \{x, y\}$ for all $t \geqslant 0$ and $Z_t = n\pi_t^0(x)$ completely determines π_t^0. The process Z is a Markov process on $\{0, ..., n\}$ with jump rates

$$p(k, k+1) = \alpha_r(y) \frac{k(n-k)}{n-1}, \quad p(k, k-1) = \alpha_r(x) \frac{k(n-k)}{n-1},$$

where α_r is given by (3.1). Denote by R its generator,

$$\tau_i(Z) = \inf\{t \geqslant 0, Z_t = i\}, \quad g(k) = \mathbb{P}_k(\tau_n(Z) < \tau_0(Z)).$$

Then, g is solution to the problem:

$$Rg = 0,\ g(0) = 0,\ g(n) = 1.$$

We have:

$$Rg(k) = \alpha_r(y)\frac{k(n-k)}{n-1}\left(g(k+1)-g(k)\right) + \alpha_r(x)\frac{k(n-k)}{n-1}\left(g(k-1)-g(k)\right).$$

Hence:

$$g(k+1)-g(k) = \alpha_{x,y,r}^{-1}\left(g(k)-g(k-1)\right),$$

$\Leftrightarrow$

$$g(k)-g(k-1) = \alpha_{x,y,r}^{-(k-1)}\left(g(1)-g(0)\right) = \alpha_{x,y,r}^{-(k-1)}g(1),$$

$\Leftrightarrow$

$$g(k) = \frac{\alpha_{x,y,r}^{-k}-1}{\alpha_{x,y,r}^{-1}-1}g(1).$$

The condition $g(n)=1$ yields:

$$g(1) = \frac{\alpha_{x,y,r}^{-1}-1}{\alpha_{x,y,r}^{-n}-1},$$

and finally:

$$\psi_x^r(\delta_{x;y}) = g(n-1) = \frac{\alpha_{x,y,r}^{-(n-1)}-1}{\alpha_{x,y,r}^{-n}-1}, \quad \psi_x^r(\delta_{y;x}) = g(1) = \frac{\alpha_{x,y,r}^{-1}-1}{\alpha_{x,y,r}^{-n}-1},$$

which concludes the proof. □

3.4. *Proof of Theorem 2.2.* The proof of Theorem 2.2 requires a last technical result on the convergence of jump Markov processes when their jump rates converge. We recall that the rates $\tilde{q}_r$ and $\tilde{q}_\infty$ are respectively defined by (2.10) and (2.11).

Lemma 3.7. *Let Y^r and Y be jump Markov processes on D with rates $\tilde{q}_r$ and $\tilde{q}_\infty$ respectively. Under Assumption 2.1*

$$\lim_{r\to\infty} \|\mathcal{L}aw(Y_0^r) - \mathcal{L}aw(Y_0)\|_{TV} = 0 \implies \forall t \geqslant 0,\ \lim_{r\to\infty} \|\mathcal{L}aw(Y_t^r) - \mathcal{L}aw(Y_t)\|_{TV} = 0.$$

Proof: We have that for all $x, y \in D$,

$$\lim_{r\to\infty} \tilde{q}_r(x,y) = \tilde{q}_\infty(x,y),$$

and for all $x \in D$,

$$\sum_{y\in D} \sup_{r\in\mathbb{N}} \tilde{q}_r(x,y) < \infty.$$

Hence, Kallenberg (2002, Theorem 19.25, p. 385) yields the result. □

We are now in position to prove Theorem 2.2.

Proof of Theorem 2.2: Let us recall the definition of the jump rates:

$$\tilde{q}_r(x,y) = \begin{cases} nq(x,y)\dfrac{\alpha_{x,y,r}-1}{\alpha_{x,y,r}^n-1} & \text{if } \alpha_{x,y,r} \neq 1, \\ q(x,y) & \text{otherwise,} \end{cases}$$

and let Y^r be the Markov process with jump rates $(\tilde{q}_r(x,y))_{x,y\in D}$ and initial condition Y_0^r such that the law of $\delta_{Y_0^r}$ is $\mathcal{L}aw(\pi_\tau^0)$. Write $p_x^r(t) = \mathbb{P}\left(Y_t^r = x\right)$, for $x \in D$. The $(p_x^r)_{x\in D}$ satisfy:

$$\sum_{x\in D} p_x^r = 1,$$

and the Kolmogorov equations:

$$\partial_t p_x^r = \sum_{y\neq x} \left(\tilde{q}_r(y,x)p_y^r - \tilde{q}_r(x,y)p_x^r\right).$$

Denote by $u(t) = (\mathbb{P}(\pi_t = \xi))_{\xi \in \mathcal{M}_n^1(D)}$ the law of the empirical measure of the Fleming–Viot process (which depends on r, even though this is not indicated explicitly for the sake of notation) and define the expected limit of u (see Equation (3.5)) by:

$$u_0(t) = \sum_{x \in D} p_x^r(t) \delta_{\delta_x},$$

which corresponds to the law of $\delta_{Y_t^r}$. Lemma 3.4 yields that that for all $t \geqslant 0$:

$$L_{s,0}^* u_0(t) = 0.$$

The goal now is to define u_1 as a solution to (3.7) using Lemma 3.5. In order to do so, let us show that

$$\partial_t u_0 - L_m^* u_0 \in (\mathrm{Ker} L_{s,0})^\perp .$$

Fix $x \in D$. First:

$$\langle \partial_t u_0, \psi_x^r \rangle = \partial_t p_x^r.$$

For all $f \in \ell^1(\mathcal{M}_n^1(D))$:

$$L_m^* f(\xi) = \sum_{x \neq y} \left((n\xi(y) + 1)\, q(y,x) f\left(\xi + \frac{\delta_y - \delta_x}{n}\right) - n\xi(x) q(x,y) f\left(\xi\right) \right),$$

hence for all $x \neq y$, $n \geqslant 3$:

$$L_m^* u_0(\eta) = \begin{cases} -n \sum_{z \neq x} q(x,z) p_x^r & \text{if } \eta = \delta_x, \\ n q(x,y) p_x^r & \text{if } \eta = \delta_{x;y}, \\ 0 & \text{otherwise.} \end{cases}$$

In the case $n = 2$, the formula is adapted by $L_m^* u_0(\delta_{x;y}) = nq(x,y)p_x^r + nq(y,x)p_y^r$. Using Lemma 3.6, this yields that:

$$\begin{aligned}
&\langle L_m^* u_0, \psi_x^r \rangle \\
&= L_m^* u_0(\delta_x)\psi_x^r(\delta_x) + \sum_{y \neq x} \left(L_m^* u_0(\delta_y)\psi_x^r(\delta_y) + L_m^* u_0(\delta_{x;y})\psi_x^r(\delta_{x;y}) + L_m^* u_0(\delta_{y;x})\psi_x^r(\delta_{y;x}) \right) \\
&= -n \sum_{z \neq x} q(x,z) p_x^r + \sum_{y \neq x} n q(x,y) p_x^r \alpha_{x,y,r} \frac{\alpha_{x,y,r}^{n-1} - 1}{\alpha_{x,y,r}^n - 1} + n q(y,x) p_y^r \alpha_{x,y,r}^{n-1} \frac{\alpha_{x,y,r} - 1}{\alpha_{x,y,r}^n - 1} \\
&= \sum_{y \neq x} \left(-\tilde{q}_r(x,y) p_x^r + \tilde{q}_r(y,x) p_y^r \right).
\end{aligned}$$

Finally we get that for all $x \in D$:

$$\langle \partial_t u_0 - L_m^* u_0, \psi_x^r \rangle = \partial_t p_x^r - \sum_{y \neq x} \left(\tilde{q}_r(y,x) p_y^r - \tilde{q}_r(x,y) p_x^r \right) = 0.$$

Lemma 3.4 then yields that $\partial_t u_0 - L_m^* u_0 \in (\mathrm{Ker} L_{s,0})^\perp$. Thus, Lemma 3.5 yields that, for any $t \geqslant 0$, there exists $u_1(t) \in (\mathrm{Ker} L_{s,0})^\perp$ such that

$$\partial_t u_0(t) - L_m^* u_0(t) = L_{s,0}^* u_1(t),$$

with the explicit representation

$$u_1(t) = \int_0^\infty \left(\partial_t u_0(t) - L_m^* u_0(t) \right) e^{s L_{s,0}} \mathrm{d}s.$$

From this representation and the regularity of $u_0(t)$ provided by its definition, one gets that u_1 is a $\mathcal{C}^1$ function of t, and that $\partial_t u_1$ solves the Poisson equation

$$\partial_t^2 u_0(t) - L_m^* \partial_t u_0(t) = L_{s,0}^* \partial_t u_1(t).$$

As a conclusion, using Lemma 3.5 again, we deduce that

$$\|u_1(t)\|_1 \leqslant C\|\partial_t u_0(t) - L_m^* u_0(t)\|_1, \quad \|\partial_t u_1(t)\|_1 \leqslant C\|\partial_t^2 u_0(t) - L_m^* \partial_t u_0(t)\|_1.$$

Let us recall the definition (2.1) of Q. The Kolmogorov equations yield that $\|\partial_t u_0\|_1 \leqslant Q$, and $\|\partial_t^2 u_0\|_1 \leqslant Q^2$, and since L_m^* is a bounded operator, we have that:

$$\sup_{r\in\mathbb{N}} \sup_{t\geqslant 0} \|u_1(t)\|_1 < \infty, \quad \sup_{r\in\mathbb{N}} \sup_{t\geqslant 0} \|\partial_t u_1(t)\|_1 < \infty. \tag{3.23}$$

Let us define v by:

$$u = u_0 + \frac{1}{\underline{\lambda}_r} u_1 + v. \tag{3.24}$$

Then v satisfies the following equation:

$$\partial_t v = L^* v + \frac{1}{\underline{\lambda}_r} \left(L_m^* u_1 - \partial_t u_1\right), \; v(0) = u(0) - u_0(0) - \frac{1}{\underline{\lambda}_r} u_1(0).$$

In the case where there exists $z \in D$ such that $\pi(\mathbf{X}_0) = \delta_z$, then $\tau = 0$ so $Y_0^r = z$, and we have $v(0) = -u_1(0)/\underline{\lambda}_r$. Thus we can write:

$$v(t) = -\frac{1}{\underline{\lambda}_r} e^{tL^*} u_1(0) + \frac{1}{\underline{\lambda}_r} \int_0^t e^{(t-s)L^*} \left(L_m^* u_1 - \partial_t u_1\right) \mathrm{d}s, \tag{3.25}$$

where for any $w \in \ell^1(\mathcal{M}_n^1(D))$, $\eta \in \mathcal{M}_n^1(D)$

$$e^{tL^*} w(\eta) = \sum_{\xi \in \mathcal{M}_n^1(D)} w(\xi) \mathbb{P}_\xi \left(\pi_t = \eta\right).$$

Thanks to the bounds on $\|u_1\|_1$ and $\|\partial_t u_1\|_1$, Equation (3.25) yields that for any $T > 0$, there exists $C > 0$ such that:

$$\sup_{t\in[0,T]} \|v(t)\|_1 \leqslant C/\underline{\lambda}_r. \tag{3.26}$$

As a conclusion, from (3.24), for any $t \in [0, T]$,

$$\|\mathcal{L}aw\left(\pi(\mathbf{X}_t)\right) - \mathcal{L}aw\left(\delta_{Y_t^r}\right)\|_1 \leqslant \|v(t)\|_1 + \frac{\|u_1(t)\|_1}{\underline{\lambda}_r},$$

and the bounds (3.23) and (3.26) thus yield the second statement of Theorem 2.2.

To prove the first statement, with an initial condition which is not necessarily a Dirac mass, we first write

$$\|\mathcal{L}aw\left(\pi(\mathbf{X}_t)\right) - \mathcal{L}aw\left(\delta_{Y_t}\right)\|_1 \leqslant \|\mathcal{L}aw\left(\pi(\mathbf{X}_t)\right) - \mathcal{L}aw\left(\delta_{Y_t^r}\right)\|_1 + \|\mathcal{L}aw\left(\delta_{Y_t^r}\right) - \mathcal{L}aw\left(\delta_{Y_t}\right)\|_1,$$

where Y is the continuous-time Markov process on D with jump rates $\tilde{q}_\infty(x, y)$ and initial condition η_∞ given by Lemma 3.3. For the first term of the right-hand side, we integrate in time the equation satisfied by v between $1/\sqrt{\underline{\lambda}_r}$ and t to get:

$$v(t) = e^{(t-1/\sqrt{\underline{\lambda}_r})L^*} v\left(\frac{1}{\sqrt{\underline{\lambda}_r}}\right) + \frac{1}{\underline{\lambda}_r} \int_{\frac{1}{\sqrt{\underline{\lambda}_r}}}^t e^{(t-s)L^*} \left(L_m^* u_1 - \partial_t u_1\right) \mathrm{d}s.$$

The second term of this last equality still goes to 0 at speed $1/\underline{\lambda}_r$, and we are just left with showing that:

$$v\left(\frac{1}{\sqrt{\underline{\lambda}_r}}\right) = u\left(\frac{1}{\sqrt{\underline{\lambda}_r}}\right) - u_0\left(\frac{1}{\sqrt{\underline{\lambda}_r}}\right) - \frac{1}{\underline{\lambda}_r} u_1\left(\frac{1}{\sqrt{\underline{\lambda}_r}}\right) \underset{r\to\infty}{\to} 0,$$

in the ℓ^1 distance. Since u_1 is bounded in $\ell^1(\mathcal{M}_n^1(D))$, uniformly in $t \geqslant 0$ and $r \in \mathbb{N}$, we have that:

$$\lim_{r\to\infty} \left\| \frac{1}{\underline{\lambda}_r} u_1\left(\frac{1}{\sqrt{\underline{\lambda}_r}}\right) \right\|_1 = 0.$$

Moreover, Y^r is a continuous-time Markov chain with initial condition Y_0^r, where Y_0^r is such that $\mathcal{L}aw(\delta_{Y_0^r}) = \mathcal{L}aw(\pi_\tau^0)$, and jump rates uniformly bounded with respect to r, hence $\underline{\lambda}_r \to \infty$ yields that:

$$\lim_{r\to\infty} \left\| \mathcal{L}aw(\pi_\tau^0) - u_0\left(\frac{1}{\sqrt{\underline{\lambda}_r}}\right) \right\|_1 = 0.$$

Let us now show that

$$\lim_{r\to\infty} \|u\left(1/\sqrt{\underline{\lambda}_r}\right) - \mathcal{L}aw(\pi_\tau^0)\|_1 = 0,$$

even though $u(0) = \eta$ for all $r \in \mathbb{N}$. Write

$$\tau_1 = \inf\{t \geqslant 0, \pi_t \in \Delta\}, \tag{3.27}$$

and S_1 for the first jump of a particle due to the mutation mechanism. Using the notation of Lemma 3.1, write $\Theta = \mathrm{supp}(\eta)$, $b_r = 1$, $L_r^2 = L_m$, $L_r^1 = L_{s,0}^{\mathrm{supp}(\eta)}$, where

$$L_{s,0}^{\mathrm{supp}(\eta)} f(\xi) = \sum_{x,y\in\mathrm{supp}(\eta)} \alpha_r(x)\xi(x)\xi(y)\left(f\left(\xi + \frac{\delta_y - \delta_x}{n}\right) - f(\xi)\right)$$

as well as $a_r = \underline{\lambda}_r$, and $c_r = \sqrt{\underline{\lambda}_r}$. In this case, we have $\mathbb{M}^1 = \Delta$, and Lemma 3.1 yields that:

$$\lim_{r\to\infty} \mathbb{P}_\eta\left(\tau_1 < \frac{1}{\sqrt{\underline{\lambda}_r}} < S_1\right) = 1. \tag{3.28}$$

For all $\xi \in \mathcal{M}_n^1(D)$, we may write:

$$\begin{aligned} u\left(\frac{1}{\sqrt{\underline{\lambda}_r}}\right)(\xi) &= \mathbb{P}_\eta\left(\pi_{\frac{1}{\sqrt{\underline{\lambda}_r}}} = \xi, \tau_1 < \frac{1}{\sqrt{\underline{\lambda}_r}} < S_1\right) + \mathbb{P}_\eta\left(\pi_{\frac{1}{\sqrt{\underline{\lambda}_r}}} = \xi, \left(\tau_1 < \frac{1}{\sqrt{\underline{\lambda}_r}} < S_1\right)^c\right) \\ &= \mathbb{P}_\eta\left(\pi_\tau^0 = \xi, \tau_1 < \frac{1}{\sqrt{\underline{\lambda}_r}} < S_1\right) + \mathbb{P}_\eta\left(\pi_{\frac{1}{\sqrt{\underline{\lambda}_r}}} = \xi, \left(\tau_1 < \frac{1}{\sqrt{\underline{\lambda}_r}} < S_1\right)^c\right), \end{aligned}$$

and thus

$$\|u\left(1/\sqrt{\underline{\lambda}_r}\right) - \mathcal{L}aw(\pi_\tau^0)\|_1 \leqslant 2\mathbb{P}_\eta\left(\left(\tau_1 < \frac{1}{\sqrt{\underline{\lambda}_r}} < S_1\right)^c\right) \underset{r\to\infty}{\to} 0.$$

Finally, Lemma 3.3 and Lemma 3.7 imply the convergence of $\|\mathcal{L}aw\left(\delta_{Y_t^r}\right) - \mathcal{L}aw\left(\delta_{Y_t}\right)\|_1$ towards 0, and conclude the proof of Theorem 2.2. □

4. Proof of Theorem 2.3

We are now interested in the convergence of the trajectories, in the space $\mathcal{D}([0,T], \mathcal{M}^1(D))$ of $\mathcal{M}^1(D)$-valued cadlag paths. As stated in the introduction, there is no hope of having convergence in the Skorohod topology. Instead, we use the L_{TV}^1 topology. Let $F : \mathcal{D}([0,T], \mathcal{M}^1(D)) \to \mathbb{R}$ be bounded and Lipschitz continuous for the L_{TV}^1 distance (2.12). We decompose

$$\begin{aligned} &\mathbb{E}\left(F\left((\pi(\mathbf{X}_t^r))_{t\in[0,T]}\right)\right) - \mathbb{E}\left(F\left((\delta_{Y_t})_{t\in[0,T]}\right)\right) \\ &= \left(\mathbb{E}\left(F\left((\pi(\mathbf{X}_t^r))_{t\in[0,T]}\right)\right) - \mathbb{E}\left(F\left(\left(\delta_{Y_t^r}\right)_{t\in[0,T]}\right)\right)\right) \\ &\quad + \left(\mathbb{E}\left(F\left(\left(\delta_{Y_t^r}\right)_{t\in[0,T]}\right)\right) - \mathbb{E}\left(F\left((\delta_{Y_t})_{t\in[0,T]}\right)\right)\right), \end{aligned}$$

with $\mathcal{L}aw(Y_0^r) = \mathcal{L}aw(\pi_\tau^0)$ and $\mathcal{L}aw(Y_0) = \eta_\infty$. Thus, the conclusion of Theorem 2.3 follows from the combination of Lemmas 4.1 and 4.2.

Lemma 4.1. *The process $(\delta_{Y_t^r})_{t\in[0,T]}$ converges in distribution, in $\mathcal{D}([0,T],\mathcal{M}^1(D))$, to the process $(\delta_{Y_t})_{t\in[0,T]}$.*

Proof: Let us endow the set D with the distance $\rho(x,y) = \mathbb{1}_{x\neq y}$, and the set $\mathcal{D}([0,T],D)$ of cadlag trajectories $[0,T]\to D$ with the associated Skorohod J_1-topology. By Kallenberg (2002, Theorem 19.25, p. 385) (see the proof of Lemma 3.7), the process $(Y_t^r)_{t\in[0,T]}$ converges in distribution to $(Y_t)_{t\in[0,T]}$ in this space. To conclude, we now check that the mapping

$$\begin{array}{ccc}\mathcal{D}([0,T],D) & \to & \mathcal{D}([0,T],\mathcal{M}^1(D)) \\ (y_t)_{t\in[0,T]} & \mapsto & (\delta_{y_t})_{t\in[0,T]}\end{array}$$

is continuous. Let $((y_t^k)_{t\in[0,T]})_{k\geqslant 1}$ be a sequence of elements of $\mathcal{D}([0,T],D)$ which converges to $(y_t)_{t\in[0,T]}$ in the J_1-topology. We have

$$\|(\delta_{y_t^k})_{t\in[0,T]} - (\delta_{y_t})_{t\in[0,T]}\|_{L^1_{TV}} = 2\int_0^T \mathbb{1}_{y_t^k\neq y_t}\mathrm{d}t = 2\int_0^T \rho(y_t^k,y_t)\mathrm{d}t.$$

By Billingsley (1999, p. 125), the convergence of $(y_t^k)_{t\in[0,T]}$ in the J_1-topology implies the convergence of $\rho(y_t^k,y_t)$ to 0 for almost every $t\in[0,T]$, and therefore the conclusion follows from the dominated convergence theorem. □

Lemma 4.2. *There exists a coupling of the processes $(\pi(\mathbf{X}_t^r))_{t\in[0,T]}$ and $\left(\delta_{Y_t^r}\right)_{t\in[0,T]}$ such that*

$$\lim_{r\to\infty}\mathbb{E}\left[\int_0^T \|\pi(\mathbf{X}_t^r) - \delta_{Y_t^r}\|_{TV}\,\mathrm{d}t\right] = 0.$$

Proof: Fix $T>0$. The proof of Lemma 4.2 is divided into five steps. The first one is dedicated the definition of the coupling.

Step 1. Let us recall the notation (3.4) and introduce $\sigma_0 = 0$,

$$\tau_k = \inf\{t\geqslant\sigma_k, \pi_t\in\Delta\},$$

and

$$\sigma_{k+1} = \inf\{t\geqslant\tau_k, \pi_t\notin\Delta\},$$

for all $k\in\mathbb{N}$. We define an intermediate process, which is not Markov:

$$\bar\eta_t = \pi_{\tau_{N(t)}}, \quad N(t) = \max\{k, \tau_k\leqslant t\},$$

and for $t\in[0,\tau_0]$, write $\bar\eta_t = \pi_{\tau_0}$. This is a process living in Δ, and we will get a Markov process with jump rates $\tilde q_r$ by a time re-scaling and conditioning. Indeed, let us define the change of time:

$$s(t) = \inf\left\{s\geqslant 0, \int_0^s \mathbb{1}_{\pi_u\in\Delta}\mathrm{d}u > t\right\}.$$

Notice that $(\bar\eta_{s(t)})_{t\geqslant 0} = (\pi_{s(t)})_{t\geqslant 0}$ is a Markov process known as the trace process of $(\pi_t)_{t\geqslant 0}$ on Δ (see Beltrán and Landim (2011); Seo (2020); Berglund (2025)). Write $\tilde N(t) = \max\{k,\sigma_k\leqslant t\}$, and

$$\mathcal{A}_r(T) = \left\{\text{For all } k\in[\![0,\tilde N(T)]\!], \text{there are no jumps from the mutation mechanism between } \sigma_k \text{ and } \tau_k\right\}.$$

If $\pi_{\tau_k} = \delta_x$, then $\sigma_{k+1}-\tau_k$ is an exponential random variable with rate $n\sum_{y\neq x} q(x,y)$. Additionally, on $\mathcal{A}_r(T)$, Lemma 3.6 yields that the Dirac mass goes from x to y with probability

$$\frac{q(x,y)}{\sum_{y\neq x} q(x,y)}\alpha_{y,x,r}^{n-1}\frac{\alpha_{y,x,r}-1}{\alpha_{y,x,r}^n-1}.$$

Hence, writing

$$\delta_{Y_t^{r,1}} = \bar\eta_{s(t)},$$

we have that on $\mathcal{A}_r(T)$, $Y^{r,1}$ is exactly a continuous-time Markov chain with jump rates $\tilde{q}_r$ and initial condition π_{τ_0}. Let $Y^{r,2}$ be another continuous-time Markov chain on D with jump rates $\tilde{q}_r$, initial condition with distribution $\mathcal{L}aw(\pi_\tau^0)$, independent of the Fleming–Viot process, and write:

$$Y^r = Y^{r,1}\mathbb{1}_{\mathcal{A}_r(T)} + Y^{r,2}\mathbb{1}_{\mathcal{A}_r^c(T)}.$$

The process Y^r is a jump Markov process with jump rates $\tilde{q}_r$ and initial condition with distribution $\mathcal{L}aw(\pi_\tau^0)$. In Step 2, we show that the probability of $\mathcal{A}_r(T)$ converges to 1. In Step 3, we show that π is close to $\bar{\eta}$, and finally in Steps 4 and 5, we show that $\bar{\eta}$ is close to $Y^{r,1}$ in the L^1_{TV} distance defined in (2.12), and thus to Y^r.

Step 2. Let us show that for all $T > 0$,

$$\lim_{r\to\infty} \mathbb{P}\left(\mathcal{A}_r(T)\right) = 1.$$

For any $k \geqslant 0$, the two events $\left\{\tilde{N}(T) \geqslant k\right\}$ and $\{$there is a mutation jump between σ_k and $\tau_k\}$ are independent. Thus we have:

$$\begin{aligned}
&\mathbb{P}\left(\mathcal{A}_r(T)^c\right) \\
&\leqslant \sum_{k=0}^{\infty} \mathbb{P}\left(\tilde{N}(T) \geqslant k, \text{ there is a mutation jump between } \sigma_k \text{ and } \tau_k\right) \\
&\leqslant \sup_{k\in\mathbb{N}} \mathbb{P}\left(\text{there is a mutation jump between } \sigma_k \text{ and } \tau_k\right) \sum_{k=0}^{\infty} \mathbb{P}\left(\tilde{N}(T) \geqslant k\right).
\end{aligned}$$

Because $\sigma_{k+1} - \tau_k$ is an exponential random variable with parameter bounded by nQ (let us recall that Q is defined in (2.1)), $\tilde{N}(T)$ is bounded by a Poisson random variable with parameter nQT, and hence:

$$\sum_{k=0}^{\infty} \mathbb{P}\left(\tilde{N}(T) \geqslant k\right) = \mathbb{E}\left(\tilde{N}(T)\right) \leqslant nQT.$$

The convergence (3.8) and the strong Markov property at time σ_k yield that

$$\lim_{r\to\infty} \sup_{k\in\mathbb{N}} \mathbb{P}\left(\text{there is a mutation jump between } \sigma_k \text{ and } \tau_k\right) = 0,$$

which concludes the proof.

Step 3. Let us now prove that the expected L^1_{TV} distance between π and $\bar{\eta}$ goes to 0 when r goes to infinity. We have:

$$\|\pi - \bar{\eta}\|_{L^1_{TV}} \mathbb{1}_{\mathcal{A}_r(T)} \leqslant 2\mathbb{1}_{\mathcal{A}_r(T)} \sum_{k=0}^{\tilde{N}(T)} \bar{\tau}_k - \sigma_k \leqslant 2\sum_{k=0}^{\infty} \left(\bar{\tau}_k - \sigma_k\right) \mathbb{1}_{\sigma_k \leqslant T},$$

where

$$\bar{\tau}_k = \inf\left\{t \geqslant \sigma_k, \bar{\pi}_t^k \in \Delta\right\}, \tag{4.1}$$

and $\bar{\pi}^k$ is equal to π on $[0, \sigma_k]$, and on $[\sigma_k, \infty]$, $\bar{\pi}^k$ is a Markov process with generator L_s, constructed with the same exponential variables as the ones used for the selection mechanism in the dynamics of π. The application of Lemma 3.1 to $(\bar{\pi}^k_{t+\sigma_k})_{t\geqslant 0}$ yields that there exist $C, c > 0$, independent of $k \in \mathbb{N}$, such that for all $t \geqslant 0$:

$$\mathbb{P}\left(\bar{\tau}_k - \sigma_k > t | \sigma_k \leqslant T\right) \leqslant Ce^{-c\underline{\lambda}_r t},$$

yielding that

$$\lim_{r\to\infty} \sup_{k\in\mathbb{N}} \mathbb{E}\left(\bar{\tau}_k - \sigma_k | \sigma_k \leqslant T\right) = 0.$$

Using again that $\tilde{N}(T)$ is bounded by a Poisson random variable with parameter nQT we get:

$$\begin{aligned}\mathbb{E}\left(\|\pi-\bar{\eta}\|_{L^1_{TV}}\right) &\leqslant 2T\mathbb{P}\left(\mathcal{A}_r(T)^c\right)+2\sup_{k\in\mathbb{N}}\mathbb{E}\left(\bar{\tau}_k-\sigma_k|\sigma_k\leqslant T\right)\sum_{k\geqslant 0}\mathbb{P}\left(\tilde{N}(T)\geqslant k\right)\\ &= 2T\mathbb{P}\left(\mathcal{A}_r(T)^c\right)+2\sup_{k\in\mathbb{N}}\mathbb{E}\left(\bar{\tau}_k-\sigma_k|\sigma_k\leqslant T\right)\mathbb{E}(\tilde{N}(T))\\ &\leqslant 2T\mathbb{P}\left(\mathcal{A}_r(T)^c\right)+2nQT\sup_{k\in\mathbb{N}}\mathbb{E}\left(\bar{\tau}_k-\sigma_k|\sigma_k\leqslant T\right)\underset{r\to\infty}{\to}0,\end{aligned}$$

where we used the result of Step 2 to control the term $\mathbb{P}\left(\mathcal{A}_r(T)^c\right)$.

Step 4. In order to show that the distance between $\bar{\eta}$ and $(\delta_{Y^{r,1}_t})_{0\leqslant t\leqslant T}$ goes to zero as well, let us define the event

$$\mathcal{B}_r(T)=\left\{\sum_{k=0}^{\tilde{N}(T)}\tau_k-\sigma_k<\min_{0\leqslant k\leqslant\tilde{N}(T)}\sigma_{k+1}-\tau_k\right\},$$

and show that

$$\lim_{r\to\infty}\mathbb{P}\left(\mathcal{A}_r(T),\mathcal{B}_r(T)\right)=1.$$

Fix $\varepsilon>0$, and $N_0\in\mathbb{N}$ such that $\mathbb{P}\left(\tilde{N}(T)>N_0\right)<\varepsilon$. Then

$$\mathbb{P}\left(\mathcal{B}_r(T)^c,\mathcal{A}_r(T)\right)\leqslant\varepsilon+\mathbb{P}\left(\sum_{k=0}^{N_0}\tau_k-\sigma_k>\min_{0\leqslant k\leqslant N_0}\sigma_{k+1}-\tau_k,\mathcal{A}_r(T)\right).$$

The random variable $\min_{0\leqslant k\leqslant N_0}\sigma_{k+1}-\tau_k$ is lower bounded by an exponential random variable E with parameter $(N_0+1)nQ$, independent of the family of random variables $(\tau_k-\sigma_k)_{0\leqslant k\leqslant N_0}$. This implies that

$$\begin{aligned}\mathbb{P}\left(\sum_{k=0}^{N_0}\tau_k-\sigma_k>\min_{0\leqslant k\leqslant N_0}\sigma_{k+1}-\tau_k,\mathcal{A}_r(T)\right)&\leqslant\mathbb{P}\left(\exists k\in[\![0,N_0]\!],\,\bar{\tau}_k-\sigma_k>E/(N_0+1)\right)\\ &\leqslant(N_0+1)\sup_{k\in\mathbb{N}}\mathbb{P}\left(\bar{\tau}_k-\sigma_k>E/(N_0+1)\right).\end{aligned}$$

By conditioning on E, Equation (3.9) then yields

$$\sup_{k\in\mathbb{N}}\mathbb{P}\left(\bar{\tau}_k-\sigma_k>E/(N_0+1)\right)\leqslant C\mathbb{E}\left(e^{-c\underline{\lambda}_rE/(N_0+1)}\right)=C\frac{nQ(N_0+1)^2}{nQ(N_0+1)^2+c\underline{\lambda}_r}\underset{r\to\infty}{\to}0,$$

so that

$$\limsup_{r\to\infty}\mathbb{P}\left(\mathcal{A}_r(T),\mathcal{B}_r(T)^c\right)\leqslant\varepsilon,$$

for all $\varepsilon>0$, which concludes Step 4.

Step 5. By the result of Step 3, all is left to do is to show that

$$\lim_{r\to\infty}\|\bar{\eta}-(\delta_{Y^r_t})_{0\leqslant t\leqslant T}\|_{L^1_{TV}}=0.$$

To proceed we shall use the result of Step 4. First notice that for all $s\geqslant 0$

$$\int_0^s\mathbb{1}_{\pi_u\in\Delta}\mathrm{d}u=s-\sum_{k=0}^{\tilde{N}(s)}s\wedge\tau_k-\sigma_k,$$

so that

$$s(t)=t+\sum_{k=0}^{\tilde{N}(s(t))}\tau_k-\sigma_k\geqslant t.$$

Additionally, $\delta_{Y_t^{r,1}} \neq \bar{\eta}_t$ only if there is $p \in \mathbb{N}$ such that $t \leqslant \sigma_p < s(t)$, and on the event $\mathcal{B}_r(T)$, $t \leqslant \sigma_p < s(t)$ implies that $\tilde{N}(s(t)) = p$. Hence

$$\sigma_p - \sum_{k=0}^{p} \tau_k - \sigma_k \leqslant t \leqslant \sigma_p.$$

Hence the total amount of time where $\delta_{Y^{r,1}} \neq \bar{\eta}$ is bounded as follows

$$\mathbb{1}_{\mathcal{A}_r(T),\mathcal{B}_r(T)} \int_0^T \mathbb{1}_{\delta_{Y_t^{r,1}} \neq \bar{\eta}_t} \mathrm{d}t \leqslant \mathbb{1}_{\mathcal{A}_r(T),\mathcal{B}_r(T)} \sum_{p=0}^{\tilde{N}(T)} \sum_{k=0}^{p} \tau_k - \sigma_k \leqslant \sum_{k=0}^{\tilde{N}(T)} \left(\tilde{N}(T) - k\right) \left(\bar{\tau}_k - \sigma_k\right),$$

where $\bar{\tau}_k$ was defined in (4.1). We have:

$$\|\delta_{Y^r} - \bar{\eta}\|_{L^1_{TV}} \mathbb{1}_{\mathcal{A}_r(T)} \mathbb{1}_{\mathcal{B}_r(T)} = \|\delta_{Y^{r,1}} - \bar{\eta}\|_{L^1_{TV}} \mathbb{1}_{\mathcal{A}_r(T)} \mathbb{1}_{\mathcal{B}_r(T)} \leqslant 2 \sum_{k=0}^{\infty} \tilde{N}(T) \left(\bar{\tau}_k - \sigma_k\right) \mathbb{1}_{\sigma_k \leqslant T}.$$

Using Cauchy-Schwarz inequality, one gets:

$$\begin{aligned}\mathbb{E}\left(\sum_{k=0}^{\infty} \tilde{N}(T) \left(\bar{\tau}_k - \sigma_k\right) \mathbb{1}_{\sigma_k \leqslant T}\right) &\leqslant \sum_{k=0}^{\infty} \mathbb{E}\left(\tilde{N}(T)^2\right)^{\frac{1}{2}} \mathbb{E}\left(\left(\bar{\tau}_k - \sigma_k\right)^2 \mathbb{1}_{\sigma_k \leqslant T}\right)^{\frac{1}{2}} \\ &\leqslant \mathbb{E}\left(\tilde{N}(T)^2\right)^{\frac{1}{2}} \left(\sup_{k \in \mathbb{N}} \mathbb{E}\left(\left(\bar{\tau}_k - \sigma_k\right)^2 \middle| \sigma_k \leqslant T\right)\right)^{\frac{1}{2}} \sum_{k \in \mathbb{N}} \sqrt{\mathbb{P}\left(\tilde{N}(T) \geqslant k\right)}.\end{aligned}$$

Write $\mathfrak{N}$ for a Poisson random variable with parameter nQT. Since $\tilde{N}(T)$ is bounded by $\mathfrak{N}$, by the Markov inequality we have that

$$\sup_{r \in \mathbb{N}} \sum_{k \in \mathbb{N}} \sqrt{\mathbb{P}\left(\tilde{N}(T) \geqslant k\right)} \leqslant \sqrt{\mathbb{E}\left(\mathfrak{N}^4\right)} \sum_{k \in \mathbb{N}} \frac{1}{k^2} < \infty,$$

and hence:

$$\begin{aligned}&\mathbb{E}\left(\|\delta_{Y^r} - \bar{\eta}\|_{L^1_{TV}}\right) \\ &\qquad \leqslant 2T\mathbb{P}\left(\left(\mathcal{A}_r(T) \cap B_r(T)\right)^c\right) + 2 \sup_{k \in \mathbb{N}} \mathbb{E}\left(\left(\bar{\tau}_k - \sigma_k\right)^2 \middle| \sigma_k \leqslant T\right)^{\frac{1}{2}} \mathbb{E}\left(\mathfrak{N}^2\right)^{\frac{1}{2}} \mathbb{E}\left(\mathfrak{N}^4\right)^{\frac{1}{2}} \pi^2/6.\end{aligned}$$

This yields that

$$\lim_{r \to \infty} \mathbb{E}\left(\|\delta_{Y^r} - \bar{\eta}\|_{L^1_{TV}}\right) = 0,$$

where we used Step 4 to control the first term, which concludes the proof. □

5. Proof of Theorem 2.5

The aim of this section is to prove Theorem 2.5, and we therefore suppose that Assumption 2.4 holds. In this case, we can write:

$$\lambda_r(x) = \underline{\lambda}_r + m_r(x),$$

where $m_r \geqslant 0$ and

$$\|m\|_\infty := \sup_{r \in \mathbb{N}} \sup_{x \in D} m_r(x) < \infty.$$

We start by showing that π_t converges to a Dirac mass. Let us recall the notation (3.4): for all $t \geqslant 0$ and $z \in D$,

$$\pi_t(z) = \pi(\mathbf{X}_t^r)(z). \tag{5.1}$$

Lemma 5.1. *Under Assumption 2.4, we have for all $t > 0$:*

$$\lim_{r\to\infty} \sum_{x\neq y\in D} \mathbb{E}(\pi_t(x)\pi_t(y)) \to 0, \qquad \lim_{r\to\infty} \mathbb{E}(\|\pi_t\|_\infty) = 1.$$

Moreover, if $\pi(\mathbf{X}_0) = \delta_z$ for some $z \in D$ and all $r \in \mathbb{N}$, then:

$$\sum_{x\neq y\in D} \mathbb{E}(\pi_t(x)\pi_t(y)) \leqslant \left(Q + \frac{n_r}{2(n_r-1)}\|m\|_\infty\right)\frac{n_r}{\underline{\lambda}_r},$$

and

$$\mathbb{E}\left(\|\pi_t\|_\infty\right) \geqslant 1 - \left(Q + \frac{n_r}{2(n_r-1)}\|m\|_\infty\right)\frac{n_r}{\underline{\lambda}_r}.$$

Proof: Write, for a given probability measure ξ on D,

$$g_2(\xi) = \sum_{x\in D} \xi(x)^2. \tag{5.2}$$

We have:

$$\begin{aligned}
Lg_2(\xi) &= \sum_{x\neq y} n_r\xi(x)\left(q(x,y) + \lambda_r(x)\frac{n_r}{n_r-1}\xi(y)\right) \\
&\qquad \left(\left(\xi(x) - \frac{1}{n_r}\right)^2 - \xi(x)^2 + \left(\xi(y) + \frac{1}{n_r}\right)^2 - \xi(y)^2\right) \\
&= 2\sum_{x\neq y} n_r\xi(x)\left(q(x,y) + \lambda_r(x)\frac{n_r}{n_r-1}\xi(y)\right)\left(\frac{1}{n_r^2} + \frac{\xi(y)-\xi(x)}{n_r}\right) \\
&\geqslant \frac{2}{n_r}\sum_{x\neq y}\lambda_r(x)\xi(x)\xi(y) - 2\sum_{x\neq y}\xi(x)^2 q(x,y) + \frac{2n_r}{n_r-1}\sum_{x\neq y}\lambda_r(x)\xi(x)\xi(y)(\xi(y)-\xi(x)) \\
&\geqslant \frac{2\underline{\lambda}_r}{n_r}(1-g_2(\xi)) - 2Qg_2(\xi) - \frac{n_r}{n_r-1}\sum_{x\neq y}(\lambda_r(y)-\lambda_r(x))\xi(x)\xi(y)(\xi(y)-\xi(x)) \\
&\geqslant \frac{2\underline{\lambda}_r}{n_r}(1-g_2(\xi)) - \left(2Q + \frac{n_r}{n_r-1}\|m\|_\infty\right).
\end{aligned}$$

Using the Kolmogorov equation, this yields that for all $t \geqslant 0$:

$$\partial_t\mathbb{E}(g_2(\pi_t)) \geqslant \frac{2\underline{\lambda}_r}{n_r}(1-\mathbb{E}(g_2(\pi_t))) - \left(2Q + \frac{n_r}{n_r-1}\|m\|_\infty\right)$$

and hence

$$\mathbb{E}(g_2(\pi_t)) \geqslant 1 - e^{-\frac{2\underline{\lambda}_r}{n_r}t}(1-\mathbb{E}(g_2(\pi_0))) - \left(2Q + \frac{n_r}{n_r-1}\|m\|_\infty\right)\frac{n_r}{2\underline{\lambda}_r}. \tag{5.3}$$

Now, since $\sum_{x\in D}\xi(x) = 1$ for all $\xi \in \mathcal{M}^1(D)$, we have:

$$g_2(\pi_t) \leqslant \|\pi_t\|_\infty,$$

which implies the result on $\|\pi_t\|_\infty$. We also have:

$$\sum_{x\neq y\in D}\mathbb{E}(\pi_t(x)\pi_t(y)) = 1 - \mathbb{E}(g_2(\pi_t)),$$

which concludes the proof. □

The preceding Lemma showed that π_t converges towards a Dirac measure as $r \to \infty$. The goal of the next Lemma is to determine the law of this limiting Dirac mass.

Lemma 5.2. *Under Assumption 2.4, we have for all $t > 0$:*

$$\lim_{r\to 0} \sum_{x\in D} |\mathbb{E}(\pi_t(x)) - \mathbb{P}(Y_t = x)| = 0,$$

where $(Y_t)_{t\geqslant 0}$ is the Markov process with jump rates $(q(x,y))_{x,y\in D}$ and initial condition η. Moreover, if $\pi(\mathbf{X}_0^r) = \delta_z$ for some $z \in D$ and all $r \in \mathbb{N}$, then for all $T > 0$, there exists $C > 0$ such that for all $t \in [0,T]$:

$$\sum_{x\in D} |\mathbb{E}(\pi_t(x)) - \mathbb{P}(Y_t = x)| \leqslant C\frac{n_r}{\underline{\lambda}_r}.$$

Proof: The Kolmogorov equation yields that for all $x \in D$:

$$\begin{aligned}\partial_t \mathbb{E}(\pi_t(x)) = &\sum_{y\neq x} \mathbb{E}(\pi_t(y))q(y,x) - \mathbb{E}(\pi_t(x)) \sum_{y\neq x} q(x,y) \\ &+ \frac{n_r}{n_r - 1} \sum_{y\neq x} \mathbb{E}(\pi_t(x)\pi_t(y)) (\lambda_r(y) - \lambda_r(x)).\end{aligned}$$

Let $(\nu_t)_{t\geqslant 0}$ be the unique solution to the following system:

$$\partial_t \nu_t(x) = \sum_{y\neq x} \nu_t(y)q(y,x) - \nu_t(x) \sum_{y\neq x} q(x,y) = L_m^*(\nu_t)(x),$$

with initial condition $\nu_0 = \eta$. We can identify $\nu_t(x) = \mathbb{P}_\eta(Y_t = x)$. Besides, L_m^* is a linear operator, and we have:

$$\partial_t (\mathbb{E}(\pi_t) - \nu_t) = L_m^* (\mathbb{E}(\pi_t) - \nu_t) + g_t,$$

where

$$g_t(x) = \sum_{y\neq x} \frac{n_r}{n_r - 1} \mathbb{E}(\pi_t(x)\pi_t(y)) (\lambda_r(y) - \lambda_r(x)),$$

for all $x \in D$. We get:

$$\mathbb{E}(\pi_t) - \nu_t = e^{tL_m^*} (\pi(\mathbf{X}_0^r) - \eta) + \int_0^t e^{sL_m^*} g_{t-s} \mathrm{d}s.$$

We have that for all $t \geqslant 0$:

$$\|g_t\|_1 \leqslant \frac{n_r}{n_r - 1} \|m\|_\infty \sum_{x\neq y\in D} \mathbb{E}(\pi_t(x)\pi_t(y)) \leqslant 2\|m\|_\infty (1 - \mathbb{E}(g_2(\pi_t))),$$

where g_2 is defined in (5.2). Since L_m is a Markov generator, Lemma 5.1 (and more precisely Equation (5.3)) yields that

$$\lim_{r\to\infty} \|\mathbb{E}(\pi_t) - \nu_t\|_1 = 0,$$

as well as, if $\eta = \pi(X_0)$:

$$\|\mathbb{E}(\pi_t) - \nu_t\|_1 \leqslant C\frac{n_r}{\underline{\lambda}_r},$$

for all $0 < t < T$ and some $C > 0$, which concludes the proof. □

We are now in position to conclude the proof of Theorem 2.5.

Proof of Theorem 2.5: Let $F : \mathcal{M}^1(D) \to \mathbb{R}$ bounded and Lipschitz continuous with respect to $\|\cdot\|_{TV}$, and $0 < \varepsilon < 1/2$. Write

$$\left|\mathbb{E}(F(\pi_t)) - \sum_{z\in D} F(\delta_z)\mathbb{P}(\|\pi_t - \delta_z\|_{TV} \leqslant \varepsilon)\right| \leqslant \sum_{z\in D} \varepsilon\|F\|_{\mathrm{Lip}}\mathbb{P}(\|\pi_t - \delta_z\|_{TV} \leqslant \varepsilon) + R_{r,t},$$

where

$$R_{r,t} = \|F\|_\infty \mathbb{P}(d(\pi_t, \Delta) > \varepsilon).$$

Using the fact that $\|\pi_t - \delta_z\|_{TV} = 2\,(1 - \pi_t(z))$, we get

$$R_{r,t} \leqslant \|F\|_\infty \mathbb{P}\left(\|\pi_t\|_\infty < 1 - \varepsilon/2\right) \leqslant \frac{2\|F\|_\infty}{\varepsilon} \mathbb{E}(1 - \|\pi_t\|_\infty),$$

and finally

$$\left|\mathbb{E}(F(\pi_t)) - \sum_{z\in D} F(\delta_z)\mathbb{P}\left(\|\pi_t - \delta_z\|_{TV} \leqslant \varepsilon\right)\right| \leqslant \varepsilon\|F\|_{\mathrm{Lip}} + \frac{2}{\varepsilon}\|F\|_\infty \mathbb{E}\left(1 - \|\pi_t\|_\infty\right). \tag{5.4}$$

Using Markov inequality, we have:

$$\mathbb{P}\left(\|\pi_t - \delta_z\|_{TV} \leqslant \varepsilon\right) = \mathbb{P}\left(\pi_t(z) \geqslant 1 - \varepsilon/2\right) \leqslant (1 - \varepsilon/2)^{-1}\mathbb{E}(\pi_t(z)) \leqslant (1 + \varepsilon)\mathbb{E}(\pi_t(z)),$$

so that

$$\mathbb{P}\left(\|\pi_t - \delta_z\|_{TV} \leqslant \varepsilon\right) - \mathbb{E}(\pi_t(z)) \leqslant \varepsilon\mathbb{E}(\pi_t(z)).$$

Besides, we have that

$$\mathbb{E}(\pi_t(z)) = \mathbb{E}\left(\pi_t(z)\left(\mathbb{1}_{\pi_t(z)\geqslant 1-\varepsilon/2} + \mathbb{1}_{\pi_t(z)\leqslant 1-\varepsilon/2}\right)\right) \leqslant \mathbb{P}\left(\pi_t(z) \geqslant 1 - \varepsilon/2\right) + \mathbb{E}\left(\pi_t(z)\mathbb{1}_{\pi_t(z)\leqslant 1-\varepsilon/2}\right),$$

so that

$$\mathbb{E}(\pi_t(z)) - \mathbb{P}\left(\|\pi_t - \delta_z\|_{TV} \leqslant \varepsilon\right) \leqslant \mathbb{E}\left(\pi_t(z)\mathbb{1}_{\pi_t(z)\leqslant 1-\varepsilon/2}\right).$$

Moreover,

$$\begin{aligned}\sum_{z\in D} \mathbb{E}\left(\pi_t(z)\mathbb{1}_{\pi_t(z)\leqslant 1-\varepsilon/2}\right) &= \sum_{z\in D} \mathbb{E}\left(\pi_t(z)\mathbb{1}_{\pi_t(z)\leqslant 1-\varepsilon/2}\left(\mathbb{1}_{\|\pi\|_\infty\leqslant 1-\varepsilon/2} + \mathbb{1}_{\|\pi\|_\infty > 1-\varepsilon/2}\right)\right)\\ &\leqslant \mathbb{P}\left(\|\pi(\mathbf{X}_t)\|_\infty < 1 - \varepsilon/2\right) + \frac{\varepsilon}{2},\end{aligned}$$

and therefore

$$\begin{aligned}&\left|\sum_{z\in D} F(\delta_z)\mathbb{P}\left(\|\pi_t - \delta_z\|_{TV} \leqslant \varepsilon\right) - \sum_{z\in D} F(\delta_z)\mathbb{E}\left(\pi_t(z)\right)\right|\\ &\qquad\leqslant \|F\|_\infty \sum_{z\in D} \left|\mathbb{P}\left(\|\pi_t - \delta_z\|_{TV} \leqslant \varepsilon\right) - \mathbb{E}\left(\pi_t(z)\right)\right|\end{aligned}$$

yielding that

$$\left|\sum_{z\in D} F(\delta_z)\mathbb{P}\left(\|\pi_t - \delta_z\|_{TV} \leqslant \varepsilon\right) - \sum_{z\in D} F(\delta_z)\mathbb{E}\left(\pi_t(z)\right)\right| \leqslant \|F\|_\infty\left(\frac{3}{2}\varepsilon + \frac{2}{\varepsilon}\mathbb{E}(1 - \|\pi_t\|_\infty)\right). \tag{5.5}$$

Lastly, we have that:

$$\left|\sum_{z\in D} F(\delta_z)\mathbb{E}\left(\pi_t(z)\right) - \sum_{z\in D} F(\delta_z)\mathbb{P}\left(Y_t = z\right)\right| \leqslant \|F\|_\infty \sum_{z\in D} \left|\mathbb{E}(\pi_t(z)) - \mathbb{P}\left(Y_t = z\right)\right|. \tag{5.6}$$

Equations (5.4), (5.5) and (5.6) together yield

$$\begin{aligned}&\left|\mathbb{E}(F(\pi_t)) - \sum_{z\in D} F(\delta_z)\mathbb{P}\left(Y_t = z\right)\right|\\ &\qquad\leqslant \varepsilon\|F\|_{\mathrm{Lip}} + \|F\|_\infty\left(\frac{3}{2}\varepsilon + \frac{4}{\varepsilon}\mathbb{E}(1 - \|\pi_t\|_\infty) + \sum_{z\in D}\left|\mathbb{E}(\pi_t(z)) - \mathbb{P}\left(Y_t = z\right)\right|\right),\end{aligned} \tag{5.7}$$

and Lemma 5.1 and Lemma 5.2 yield, for $t > 0$,

$$\limsup_{r\to\infty}\left|\mathbb{E}(F(\pi_t)) - \sum_{z\in D} F(\delta_z)\mathbb{P}\left(Y_t = z\right)\right| \leqslant \left(\|F\|_{\mathrm{Lip}} + \frac{3}{2}\|F\|_\infty\right)\varepsilon.$$

Letting $\varepsilon \to 0$ yields the first part of the theorem. In the case where the initial condition is a Dirac mass, plugging $\varepsilon = \sqrt{n_r/\underline{\lambda}_r}$ in equation (5.7) and Lemma 5.2 yields for all $T > 0$ and $0 < t < T$:

$$\begin{aligned}
&\left| \mathbb{E}(F(\pi_t)) - \sum_{z \in D} F(\delta_z) \mathbb{P}\left(Y_t = z\right) \right| \\
&\quad \leqslant \|F\|_{\mathrm{Lip}} \sqrt{\frac{n_r}{\underline{\lambda}_r}} + \|F\|_\infty \left(\frac{3}{2} \sqrt{\frac{n_r}{2\underline{\lambda}_r}} + 4 \sqrt{\frac{\underline{\lambda}_r}{n_r}} \left(Q + \frac{n_r}{2(n_r - 1)} \|m\|_\infty \right) \frac{n_r}{\underline{\lambda}_r} + C \frac{n_r}{\underline{\lambda}_r} \right) \\
&\quad \leqslant C' \left(\|F\|_{\mathrm{Lip}} + \|F\|_\infty \right) \sqrt{\frac{n_r}{\underline{\lambda}_r}},
\end{aligned}$$

for some $C' > 0$, and this concludes the proof. □

Acknowledgements

The author would like to thank an anonymous referee for pointing out several mistakes in a preliminary version.

References

Asselah, A., Ferrari, P. A., and Groisman, P. Quasistationary distributions and Fleming-Viot processes in finite spaces. *J. Appl. Probab.*, **48** (2), 322–332 (2011). MR2840302.

Bansaye, V., Cloez, B., Gabriel, P., and Marguet, A. A non-conservative Harris ergodic theorem. *J. Lond. Math. Soc. (2)*, **106** (3), 2459–2510 (2022). MR4498558.

Beltrán, J. and Landim, C. Metastability of reversible finite state Markov processes. *Stochastic Process. Appl.*, **121** (8), 1633–1677 (2011). MR2811018.

Berglund, N. Reducing metastable continuous-space Markov chains to Markov chains on a finite set. *Ann. Inst. Henri Poincaré Probab. Stat.*, **61** (3), 2036 – 2073 (2025). DOI: 10.1214/24-AIHP1468.

Bianchi, A., Dommers, S., and Giardinà, C. Metastability in the reversible inclusion process. *Electron. J. Probab.*, **22**, Paper No. 70, 34 (2017). MR3698739.

Billingsley, P. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition (1999). ISBN 0-471-19745-9. MR1700749.

Burdzy, K., Hołyst, R., and March, P. A Fleming-Viot particle representation of the Dirichlet Laplacian. *Comm. Math. Phys.*, **214** (3), 679–703 (2000). MR1800866.

Cérou, F., Delyon, B., Guyader, A., and Rousset, M. A central limit theorem for Fleming-Viot particle systems. *Ann. Inst. Henri Poincaré Probab. Stat.*, **56** (1), 637–666 (2020). MR4059003.

Champagnat, N. A microscopic interpretation for adaptive dynamics trait substitution sequence models. *Stochastic Process. Appl.*, **116** (8), 1127–1160 (2006). MR2250806.

Champagnat, N. and Hass, V. Convergence of population processes with small and frequent mutations to the canonical equation of adaptive dynamics. *Ann. Appl. Probab.*, **35** (1), 1–63 (2025). MR4870757.

Champagnat, N. and Lambert, A. Evolution of discrete populations and the canonical diffusion of adaptive dynamics. *Ann. Appl. Probab.*, **17** (1), 102–155 (2007). MR2292582.

Champagnat, N. and Villemonais, D. Practical criteria for R-positive recurrence of unbounded semigroups. *Electron. Commun. Probab.*, **25**, Paper No. 6, 11 (2020). MR4066299.

Champagnat, N. and Villemonais, D. General criteria for the study of quasi-stationarity. *Electron. J. Probab.*, **28**, Paper No. 22, 84 (2023). MR4546021.

Chleboun, P., Gabriel, S., and Grosskinsky, S. Size-biased diffusion limits and the inclusion process. *Electron. J. Probab.*, **29**, Paper No. 65, 36 (2024). MR4739837.

Collet, P., Martínez, S., and San Martín, J. *Quasi-stationary distributions. Markov chains, diffusions and dynamical systems*. Probability and its Applications (New York). Springer, Heidelberg (2013). ISBN 978-3-642-33130-5; 978-3-642-33131-2. MR2986807.

Del Moral, P. *Feynman-Kac formulae. Genealogical and interacting particle systems with applications*. Probability and its Applications (New York). Springer-Verlag, New York (2004). ISBN 0-387-20268-4. MR2044973.

Del Moral, P. and Miclo, L. Particle approximations of Lyapunov exponents connected to Schrödinger operators and Feynman-Kac semigroups. *ESAIM Probab. Stat.*, **7**, 171–208 (2003). MR1956078.

Di Gesù, G., Lelièvre, T., Le Peutrec, D., and Nectoux, B. Jump Markov models and transition state theory: the quasi-stationary distribution approach. *Faraday Discuss.*, **195**, 469–495 (2016). DOI: 10.1039/C6FD00120C.

Etheridge, A. M. *An introduction to superprocesses*, volume 20 of *University Lecture Series*. American Mathematical Society, Providence, RI (2000). ISBN 0-8218-2706-5. MR1779100.

Fleming, W. H. and Viot, M. Some measure-valued Markov processes in population genetics theory. *Indiana Univ. Math. J.*, **28** (5), 817–843 (1979). MR542340.

Grigorescu, I. and Kang, M. Hydrodynamic limit for a Fleming-Viot type system. *Stochastic Process. Appl.*, **110** (1), 111–143 (2004). MR2052139.

Grosskinsky, S., Redig, F., and Vafayi, K. Dynamics of condensation in the symmetric inclusion process. *Electron. J. Probab.*, **18**, no. 66, 23 (2013). MR3078025.

Grosskinsky, S., Schütz, G. M., and Spohn, H. Condensation in the zero range process: stationary and dynamical properties. *J. Statist. Phys.*, **113** (3-4), 389–410 (2003). MR2013129.

Journel, L. and Monmarché, P. Convergence of a particle approximation for the quasi-stationary distribution of a diffusion process: uniform estimates in a compact soft case. *ESAIM Probab. Stat.*, **26**, 1–25 (2022). MR4363453.

Journel, L. and Monmarché, P. Uniform convergence of the Fleming-Viot process in a hard killing metastable case. *Ann. Appl. Probab.*, **35** (2), 1019–1048 (2025). MR4897753.

Kallenberg, O. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition (2002). ISBN 0-387-95313-2. MR1876169.

Le Bris, C., Lelièvre, T., Luskin, M., and Perez, D. A mathematical formalization of the parallel replica dynamics. *Monte Carlo Methods Appl.*, **18** (2), 119–146 (2012). MR2926765.

Lelièvre, T., Pillaud-Vivien, L., and Reygner, J. Central limit theorem for stationary Fleming-Viot particle systems in finite spaces. *ALEA Lat. Am. J. Probab. Math. Stat.*, **15** (2), 1163–1182 (2018). MR3860819.

Löbus, J.-U. A stationary Fleming-Viot type Brownian particle system. *Math. Z.*, **263** (3), 541–581 (2009). MR2545857.

Méléard, S. and Villemonais, D. Quasi-stationary distributions and population processes. *Probab. Surv.*, **9**, 340–410 (2012). MR2994898.

Meyer, P.-A. and Zheng, W. A. Tightness criteria for laws of semimartingales. *Ann. Inst. H. Poincaré Probab. Statist.*, **20** (4), 353–372 (1984). MR771895.

Norris, J. R. *Markov chains*, volume 2 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge (1998). ISBN 0-521-48181-3. MR1600720.

Oçafrain, W. and Villemonais, D. Convergence of a non-failable mean-field particle system. *Stoch. Anal. Appl.*, **35** (4), 587–603 (2017). MR3651131.

Olivieri, E. and Vares, M. E. *Large deviations and metastability*, volume 100 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge (2005). ISBN 0-521-59163-5. MR2123364.

Pavliotis, G. A. and Stuart, A. M. *Multiscale methods. Averaging and homogenization*, volume 53 of *Texts in Applied Mathematics*. Springer, New York (2008). ISBN 978-0-387-73828-4. MR2382139.

Pollock, M., Fearnhead, P., Johansen, A. M., and Roberts, G. O. Quasi-stationary Monte Carlo and the ScaLE algorithm. *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, **82** (5), 1167–1221 (2020). MR4176340.

Rousset, M. On the control of an interacting particle estimation of Schrödinger ground states. *SIAM J. Math. Anal.*, **38** (3), 824–844 (2006). MR2262944.

Seo, I. Analysis of metastable behavior via solutions of Poisson equations. In *Stochastic analysis on large scale interacting systems*, RIMS Kôkyûroku Bessatsu, B79, pp. 1–17. Res. Inst. Math. Sci. (RIMS), Kyoto (2020). MR4279008.

Villemonais, D. General approximation method for the distribution of Markov processes conditioned not to be killed. *ESAIM Probab. Stat.*, **18**, 441–467 (2014). MR3333998.

Waclaw, B. and Evans, M. R. Explosive Condensation in a Mass Transport Model. *Phys. Rev. Lett.*, **108**, 070601 (2012). DOI: 10.1103/PhysRevLett.108.070601.