ANNALES
DE L'INSTITUT
HENRI
POINCARÉ

PROBABILITÉS
ET STATISTIQUES

Fleming–Viot selects the minimal quasi-stationary distribution: The Galton–Watson case

Amine Asselah[a], Pablo A. Ferrari[b], Pablo Groisman[b] and Matthieu Jonckheere[c]

[a]*LAMA, Bat. P3/4, Université Paris-Est Créteil, 61 Av. General de Gaulle, 94010 Créteil Cedex, France. E-mail: amine.asselah@univ-paris12.fr*
[b]*Departamento de Matemática, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Pabellón 1, Ciudad Universitaria, 1428 Buenos Aires, Argentina. E-mail: pferrari@dm.uba.ar; pgroisma@dm.uba.ar*
[c]*Instituto de Investigaciones Matemáticas Luis Santaló, Pabellón 1, Ciudad Universitaria, 1428 Buenos Aires, Argentina. E-mail: mjonckhe@dm.uba.ar*

Abstract. Consider N particles moving independently, each one according to a subcritical continuous-time Galton–Watson process unless it hits 0, at which time it jumps instantaneously to the position of one of the other particles chosen uniformly at random. The resulting dynamics is called Fleming–Viot process. We show that for each N there exists a unique invariant measure for the Fleming–Viot process, and that its stationary empirical distribution converges, as N goes to infinity, to the minimal quasi-stationary distribution of the Galton–Watson process conditioned on non-extinction.

Résumé. Nous considérons N particules indépendantes. Chaque particule suit l'évolution d'un processus de Galton–Watson sous-critique jusqu'au moment où elle touche 0. À cet instant, cette particule choisit uniformément au hasard la position d'une des autres particules et y saute. Ce processus est appelé Fleming–Viot. Nous montrons que pour chaque entier N, il existe une unique mesure invariante pour le processus de Fleming–Viot, et que la mesure empirique stationnaire converge vers la loi quasi-stationnaire minimale d'un processus de Galton–Watson conditionné à ne pas mourir.

MSC: Primary 60K35; secondary 60J25

Keywords: Quasi-stationary distributions; Fleming–Viot processes; Galton–Watson processes; Selection principle

1. Introduction

The concept of *quasi-stationarity* arises in stochastic modeling of population dynamics. In 1947, Yaglom [25] considers subcritical Galton–Watson processes conditioned to survive long times. He shows that as time is sent to infinity, the conditioned process, started with one individual, converges to a law, now called a *quasi-stationary distribution*. For any Markov process, and a subset A of the state space, we denote by μT_t the law of the process at time t conditioned on not having hit A up to time t, with initial distribution μ. A probability measure on A^c is called *quasi-stationary distribution* if it is a fixed point of T_t for any $t > 0$.

In 1966, Seneta and Veres-Jones [23] realize that for subcritical Galton–Watson processes, there is a one-parameter family of quasi-stationary distributions and show that the Yaglom limit distribution has the minimal expected time of extinction among all quasi-stationary distributions. This unique *minimal* quasi-stationary distribution is denoted here ν_{qs}^*. They also show that with an initial distribution μ with finite first moment, μT_t converges to ν_{qs}^* as t goes to infinity.

In 1978, Cavender [10] shows that for Birth and Death chains on the non-negative integers absorbed at 0, the set of quasi-stationary measures is either empty or is a one parameter family. In the latter case, Cavender extends

the *selection principle* of Seneta and Veres-Jones. He also shows that the limit of the sequence of quasi-stationary distributions for truncated processes on $\{1, \ldots, L\}$ converges to ν_{qs}^* as L is sent to infinity. This picture holds for a class of irreducible Markov processes on the non-negative integers with 0 as absorbing state, as shown in 1995 by Ferrari, Kesten, Martinez and Picco [13]. The main idea in [13] is to think of the conditioned process μT_t as a mass transport with refeeding from the absorbing state to each of the transient states with a rate proportional to the transient state mass. More precisely, denoting $\mathbb{N}$ the set of positive integers, the Kolmogorov forward equation satisfied by $\mu T_t(x)$, for each $x \in \mathbb{N}$, reads

$$\frac{\partial}{\partial t} \mu T_t(x) = \sum_{y:y \neq x} \big(q(x, y) + q(x, 0)\mu T_t(y)\big)\big[\mu T_t(y) - \mu T_t(x)\big], \tag{1.1}$$

where $q(x, y)$ is the jump rate from x to y. The first term in the right-hand side represents the displacement of mass due to the jumps of the process and the second term represents the mass going from each x to 0 and then coming instantaneously to y.

In 1996, Burdzy, Holyst, Ingerman and March [9] introduce a *genetic* particle system called *Fleming–Viot* named after models proposed in [15], which can be seen as a particle system mimicking the evolution (1.1). The particle system can be built from a process with absorption Z_t called the *driving process*; the position Z_t is interpreted as a genetic trait, or fitness, of an individual at time t. In the N-particle Fleming–Viot system, each trait follows independent dynamics with the same law as Z_t except when one of them hits state 0, a lethal trait: at this moment the individual adopts the trait of one of the other individuals chosen uniformly at random. Leaving aside the genetic interpretation, the empirical distribution of the N particles at positions $\xi \in \mathbb{N}^N$ is defined as a function $m(\cdot, \xi): \mathbb{N} \to [0, 1]$ by

$$\forall x \in \mathbb{N}, \quad m(x, \xi) := \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}_{\{\xi(i)=x\}}. \tag{1.2}$$

The generator of the Fleming–Viot process with N particles applied to bounded functions $f : \mathbb{N}^N \to \mathbb{R}$ reads

$$\mathcal{L}^N f(\xi) = \sum_{i=1}^{N} \sum_{y=1}^{\infty} \bigg[q\big(\xi(i), y\big) + q\big(\xi(i), 0\big)\frac{N}{N-1}m(y, \xi)\bigg]\big[f\big(\xi^{i,y}\big) - f(\xi)\big], \tag{1.3}$$

where $\xi^{i,y}(i) = y$, and for $j \neq i$, $\xi^{i,y}(j) = \xi(j)$ and $q(x, y)$ are the jump rates of the driving process.

In a few special cases, where the driving process has a unique quasi-stationary distribution, denoted ν_{qs}, and where the associated N-particle Fleming–Viot system has a unique invariant measure, denoted λ^N, it is established that the law of the random measure $m(\cdot, \xi)$, where ξ has law λ^N, converges to ν_{qs}. These cases include diffusion processes on a bounded domain of $\mathbb{R}^d$, killed at the boundary [4,16,17,24], for jump processes under a Doeblin condition [14] and for finite state jump processes [1].

The subcritical Galton–Watson process has infinitely many quasi-stationary distributions. Our theorem proves that the stationary empirical distribution $m(\cdot, \xi)$ converges to ν_{qs}^*, the minimal quasi-stationary distribution. This phenomenon is a *selection principle*.

Theorem 1.1. *Consider a subcritical Galton–Watson process whose offspring law has some finite positive exponential moment. Let ν_{qs}^* be the minimal quasi-stationary distribution for the process conditioned on non-extinction. Then, for each $N \geq 1$, the associated N-particle Fleming–Viot system is ergodic. Furthermore, if we call its invariant measure λ^N, then*

$$\forall x \in \mathbb{N}, \quad \lim_{N \to \infty} \int \big|m(x, \xi) - \nu_{qs}^*(x)\big| \, d\lambda^N(\xi) = 0. \tag{1.4}$$

As a consequence, we get propagation of chaos:

Corollary 1.2. *For any finite set $S \subset \mathbb{N}$,*

$$\lim_{N \to \infty} \int \prod_{x \in S} m(x, \xi) \, d\lambda^N(\xi) = \prod_{x \in S} v_{qs}^*(x). \tag{1.5}$$

In the next section we show this corollary and the strategy for proving Theorem 1.1. Let us highlight two key steps in the proof with interest of their own. First, we control the position of the rightmost particle. Let

$$R(\xi) := \max_{i \in \{1, \ldots, N\}} \xi(i),$$

be the position of the rightmost particle of ξ and let $\xi_t^\xi(1), \ldots, \xi_t^\xi(N)$ be the positions at time t of the N Fleming–Viot particles, initially at $\xi(1), \ldots, \xi(N)$.

Proposition 1.3. *There is a time T and positive constants A, c_1, c_2, C and ρ, independent of N, such that for any $\xi \in \mathbb{N}^N$*

$$E\left(\exp\left(\rho R\left(\xi_T^\xi\right)\right)\right) - \exp\left(\rho R(\xi)\right) < -c_1 e^{\rho R(\xi)} \mathbf{1}_{\{R(\xi) > A\}} + N C e^{-c_2 R(\xi)}. \tag{1.6}$$

Corollary 1.4. *For each N there is a unique invariant measure λ^N for the Fleming–Viot system. Furthermore, there is a constant $\kappa > 0$ such that for any N,*

$$\int \exp\left(\rho R(\xi)\right) d\lambda^N(\xi) \leq \kappa N. \tag{1.7}$$

The proposition and its corollary imply that for each fixed N, the Fleming–Viot process converges to its invariant distribution λ^N exponentially fast in time. Let us mention that the control of the maximum is obtained through a coupling with a branching-type process whose population grows exponentially in time. The strong drift of the driving process allows us to use the coupling until a convenient fixed finite time.

The second result is that the ratio between the second and the first moment of the empirical distribution plays the role of a Lyapunov functional, given that the position of the rightmost particle is not too large. For a particle configuration ξ define

$$\psi(\xi) := \frac{\sum_{1 \leq i \leq N} \xi^2(i)}{\sum_{1 \leq i \leq N} \xi(i)}. \tag{1.8}$$

Recall $\mathcal{L}^N$ is the Fleming–Viot generator given by (1.3).

Proposition 1.5. *There are positive constants v, C_1 and C_2 independent of N such that*

$$\mathcal{L}^N \psi(\xi) \leq -v\psi(\xi) + C_1 \frac{R^2(\xi)}{N} + C_2. \tag{1.9}$$

Taking expectation in (1.9) under the invariant measure λ^N, and using (1.7) we obtain the following bound.

Corollary 1.6. *There is a positive constant C such that for all N,*

$$\int \psi(\xi) \, d\lambda^N(\xi) \leq C. \tag{1.10}$$

There are several related works motivated by genetics. Brunet, Derrida, Mueller and Munier [7,8] introduce a model of evolution of a population with selection. They study the genealogy of genetic traits, the empirical measure, and link the evolution of the barycenter with F-KPP equation $\partial_t u = \partial_{xx} u - u(1 - u)$ introduced in 1937 by R. A. Fisher to

describe the evolution of an advantageous gene in a population. These authors also discover an exactly soluble model whose genealogy is identical to those predicted by Parisi's theory of mean-field spin glasses. Durrett and Remenik [11] establish propagation of chaos for a related continuous-space and time model, and then show that the limit of the empirical measure is characterized as the solution of a free-boundary integro-differential equation. Bérard and Gouéré [2] establish a conjecture of Brunet and Derrida for the speed of the rightmost particle for still a third microscopic model of F-KPP equation introduced in [5,6]. Maillard [19] obtains the precise behavior of the empirical measure of an approximation of the same model, building on the results of Berestycki, Berestycki and Schweinsberg [3], which establish the genealogy picture described in [5,6].

We now mention two open problems. The first is to solve the analogue of Theorem 1.1 for a random walk with a constant drift toward the origin. The difficulty is that we have no candidate for the Lyapounov-type function. This is a delicate issue without which we cannot even bound the barycenter of the Fleming–Viot process under the stationary measure. The second problem is to obtain propagation of chaos directly on the stationary empirical measure, with a bound of order $1/N$.

In the next section, we describe our model, sketch the proof of our main result and describe the organization of the paper.

2. Notation and strategy

Let $\sigma > 0$ and p be a probability distribution on $\mathbb{N} \cup \{0\}$ such that

$$\sum_{\ell \geq 0} p(\ell) e^{\sigma \ell} < \infty. \tag{2.1}$$

Consider a Galton–Watson process $Z_t \in \mathbb{N} \cup \{0\}$ with offspring law p. Each individual lives an exponential time of parameter 1, and then gives birth to a random number of children with law p. The Galton–Watson is subcritical: we ask p to satisfy

$$-v := \sum_{\ell \geq -1} \ell p(\ell + 1) < 0. \tag{2.2}$$

In other words, the drift when $Z_t = x$ is $-vx < 0$. This constant $v > 0$ is the one that appears in Proposition 1.5. For distinct $x, y \in \mathbb{N} \cup \{0\}$, the rates of jump are given by

$$q(x, y) := \begin{cases} xp(0), & \text{if } y = x - 1 \geq 0, \\ xp(y - x + 1), & \text{if } y > x \geq 1, \\ 0, & \text{otherwise.} \end{cases} \tag{2.3}$$

The Galton–Watson process starting at x is denoted Z_t^x. For a distribution μ on $\mathbb{N}$, the law of the process starting with μ conditioned on non-absorption until time t is given by

$$\mu T_t(y) := \frac{\sum_{x \in \mathbb{N}} \mu(x) p_t(x, y)}{\sum_{x, z \in \mathbb{N}} \mu(x) p_t(x, z)}, \tag{2.4}$$

where $p_t(x, y) = P(Z_t^x = y)$.

Recall that ξ_t^ξ denotes the Fleming–Viot system with generator (1.3) and initial state ξ; $\xi_t(i)$ denotes the position of the ith particle at time t. For a real $\alpha > 0$ define $K(\alpha)$ as the subset of distributions on $\mathbb{N}$ given by

$$K(\alpha) := \left\{ \mu : \frac{\sum_{x \in \mathbb{N}} x^2 \mu(x)}{\sum_{x \in \mathbb{N}} x \mu(x)} \leq \alpha \right\}. \tag{2.5}$$

Observe that $\mu \in K(\alpha)$ implies $\sum x \mu(x) \leq \alpha$.

Proof of Theorem 1.1. The existence of the unique invariant measure λ^N for Fleming–Viot is given in Proposition 1.3 and Corollary 1.4.

To show (1.4) we use the invariance of λ^N and perform the following decomposition.

$$\int \left| m(x, \xi) - \nu_{\mathrm{qs}}^*(x) \right| \mathrm{d}\lambda^N(\xi)$$

$$= \int E \left| m(x, \xi_t^\xi) - \nu_{\mathrm{qs}}^*(x) \right| \mathrm{d}\lambda^N(\xi)$$

$$\leq \lambda^N(\psi > \alpha) + \int_{\psi \leq \alpha} E \left| m(x, \xi_t^\xi) - \nu_{\mathrm{qs}}^*(x) \right| \mathrm{d}\lambda^N(\xi)$$

$$\leq \lambda^N(\psi > \alpha) + \int_{\psi \leq \alpha} E \left| m(x, \xi_t^\xi) - m(\cdot, \xi) T_t(x) \right| \mathrm{d}\lambda^N(\xi) + \int_{\psi \leq \alpha} \left| m(\cdot, \xi) T_t(x) - \nu_{\mathrm{qs}}^*(x) \right| \mathrm{d}\lambda^N(\xi)$$

$$\leq \lambda^N(\psi > \alpha) + \sup_{\xi : \psi(\xi) \leq \alpha} \left| m(\cdot, \xi) T_t(x) - \nu_{\mathrm{qs}}^*(x) \right| + \sup_{\xi : \psi(\xi) \leq \alpha} E \left| m(x, \xi_t^\xi) - m(\cdot, \xi) T_t(x) \right|, \tag{2.6}$$

where ψ is defined in (1.8). We bound the three terms of the last line of (2.6).

First term

Corollary 1.6 and Markov's inequality imply that there is a constant $C > 0$ such that for any $\alpha > 0$

$$\lambda^N(\psi > \alpha) \leq \frac{C}{\alpha}. \tag{2.7}$$

Second term

Note that $\psi(\xi) \leq \alpha$ if and only if $m(\cdot, \xi) \in K(\alpha)$. The Yaglom limit converges to the minimal quasi-stationary distribution ν_{qs}^*, uniformly in $K(\alpha)$ as we show later in Proposition 7.2:

$$\lim_{t \to \infty} \sup_{\mu \in K(\alpha)} \left| \mu T_t(x) - \nu_{\mathrm{qs}}^*(x) \right| = 0. \tag{2.8}$$

Third term

We show in Proposition 8.1 that there exist positive constants C and c such that

$$\sup_{\xi \in \mathbb{N}^N} E \left[m(x, \xi_t^\xi) - m(\cdot, \xi) T_t \right]^2 \leq \frac{C e^{ct}}{N}, \quad x \in \mathbb{N} \tag{2.9}$$

for all N. The issue here is a uniform bound for the correlations of the empirical distribution of Fleming–Viot at sites $x, y \in \mathbb{N}$ at fixed time t. This was carried out in [1].

To bound the bottom line of (2.6) choose α large and use (2.7) to make the first term small (uniform in N). Use (2.8) to choose t large to make the second term small. For this fixed time, take N large and use (2.9) to make the third term small. $\qquad\square$

Proof of Corollary 1.2. Writing λf instead of $\int f(\xi) \mathrm{d}\lambda^N(\xi)$ and ν instead of ν_{qs}^*, we get $\left| \lambda \prod_{x \in S} m(x) - \prod_{x \in S} \nu(x) \right| \leq \sum_{x \in S} \lambda |m(x) - \nu(x)|$. It suffices now to apply (1.4). $\qquad\square$

The rest of the paper is organized as follows. In Section 3, we perform the graphical construction of Fleming–Viot jointly with a branching-type process (coupling). In Section 4 we obtain large deviation estimates for the Galton–Watson process. In Section 5 we prove Proposition 1.3. In Section 6 we study the Lyapunov-like functional ψ and prove Proposition 1.5 and Corollary 1.6. Convergence of the conditional evolution uniformly on $K(\alpha)$ is proved in Section 7. Finally, (2.9) is handled in Proposition 8.1 of Section 8.

3. Embedding Fleming–Viot on a branching-type process

In this section we construct a coupling between the Fleming–Viot process and a branching-type process. Let *FV-particles* refer to the N positions in the Fleming–Viot process, whereas *BT-particles* refer to a growing number of positions in the branching process. Each BT-particle has a position in $\mathbb{N}$ and a type $i \in \{1, \ldots, N\}$.

The i-FV-particle performs jumps governed by the rates $(\tilde{q}(x, y), x, y \in \mathbb{N})$ defined by $\tilde{q}(x, y) := q(x, y)\mathbf{1}_{\{y \neq 0\}}$ and "jumps to zero and then to the position of another FV-particle chosen uniformly at random" at rate $q(x, 0)/(N - 1)$.

The coupling has the following properties. There is always at least one i-BT-particle at the position of the i-FV-particle. When the i-FV-particle performs a jump governed by $\tilde{q}$, one of the i-BT-particles at the same position performs the same jump. When the i-FV-particle (is absorbed and immediately) jumps to the position of the j-FV-particle, a new i-BT-particle is created at the position of each j-BT-particle (that is, each j-BT-particle (dies and) branches into a j-BT-particle and an i-BT-particle, both at the same position).

We perform a joint Harris construction of the processes. The state of each process at time t is defined as a function of the initial configuration and a (multidimensional) Poisson process in the time interval $[0, t]$. The coupling emerges naturally by taking the same initial configuration and the same Poisson process for both processes. The coupling holds more generally when the driving process is a Markov process with rates $\{q(x, y), x, y \in \mathbb{N} \cup \{0\}\}$ with 0 being the absorbing state and $\bar{q} := \sup_x q(x, 0) < \infty$; the Galton–Watson is a particular case.

Spatial evolution

Each BT-particle has a position in $\mathbb{N}$ which evolves independently with transition rates $\tilde{q}$ so that there are no jumps to zero. The spatial evolution of new BT-particles born at branching times are independent and with the same rates $\tilde{q}$. Under our coupling, each spatial jump performed by the i-FV-particle is also performed by some i-BT-particle.

The refeeding and branching

At rate $\bar{q}/(N - 1)$, each j-BT-particle branches into two new BT-particles, one of type j and one of type i; consequently the total branching rate of each BT-particle is $\bar{q}N/(N - 1)$. Each new born i-BT-particle appears at the position of the corresponding j-BT-particle and then evolves independently with rates $\tilde{q}$. If the i-FV-particle is at x, then at rate $q(x, 0)/(N - 1)$ it jumps to the position of the j-FV-particle and – under our coupling – simultaneously each j-BT-particle branches into an i and a j-BT-particle. In this way, the i-FV-particle occupies always the site of some i-BT-particle. This can be done because since $q(x, 0) \leq \bar{q}$, the Poisson process of rate $q(x, 0)/(N - 1)$ governing the jumps from x to 0 can be set as a thinning of the Poisson process with rates $\bar{q}/(N - 1)$ governing the branchings.

The actual construction of the coupling requires more notation and definitions. The branching-type process has state space

$$\mathcal{B} := \left\{ \zeta \in \mathbb{N}^{\{1,\ldots,N\} \times \mathbb{N}} : \sum_{i=1}^{N} \sum_{x \in \mathbb{N}} \zeta(i, x) < \infty \right\}.$$

For $i \in \{1, \ldots, N\}$, $x \in \mathbb{N}$, $\zeta_t(i, x)$ indicates the number of BT-particles of type i at site x at time t. Let $\delta_{(i,x)} \in \mathcal{B}$ be the delta function on (i, x) defined by $\delta_{(i,x)}(i, x) = 1$ and $\delta_{(i,x)}(j, y) = 0$ for $(j, y) \neq (i, x)$. The rates corresponding to the (independent) spatial evolution of the BT-particles at x are

$$b(\zeta, \zeta + \delta_{(i,y)} - \delta_{(i,x)}) = \zeta(i, x)\tilde{q}(x, y), \quad i \in \{1, \ldots, N\}, x, y \in \mathbb{N},$$

and those corresponding to the branching of all j-individuals into an individual of type j and an individual of type i are

$$b\left(\zeta, \zeta + \sum_{x \in \mathbb{N}} \zeta(j, x)\delta_{(i,x)} \right) = \frac{\bar{q}}{N - 1}, \quad i \neq j \in \{1, \ldots, N\}.$$

Note that the new born i-BT-particles get the spatial position of the corresponding j-BT-particle.

Harris construction of the branching process

Let $(\mathcal{N}(i, x, y, k), i \in \{1, \dots, N\}, x, y \in \mathbb{N}, k \in \mathbb{N})$ be a family of Poisson processes with rates $k\tilde{q}(x, y)$ such that $\mathcal{N}(i, x, y, k) \subset \mathcal{N}(i, x, y, k+1)$ for all k; we think of a Poisson process as a random subset of $\mathbb{R}$. The process $\mathcal{N}(i, x, y, k)$ is used to produce a jump of an i-BT-particle from x to y when there are k i-BT-particles at site x. The families $\{(\mathcal{N}(i, x, y, k), k \geq 1), i \in \{1, 2, \dots, N\}, x, y \in \mathbb{N}\}$ are taken independent. Let $(\mathcal{N}(i, j), i \neq j)$ be another family of independent Poisson processes of rate $\bar{q}/(N-1)$, these processes are used to branch each j-BT-particle into an i-BT-particle and a j-BT-particle. The two families are taken independent.

Fix $\zeta_0 = \zeta \in \mathcal{B}$, assume the process is defined until time $s \geq 0$ and proceed by recurrence.

1. Define $\tau(\zeta_s, s) := \inf\{t > s \colon t \in \bigcup_{i,x,y} \mathcal{N}(i, x, y, \zeta_s(i, x)) \cup \bigcup_{i,j} \mathcal{N}(i, j)\}$.
2. For $t \in [s, \tau)$ define $\zeta_t = \zeta_s$.
3. If $\tau \in \mathcal{N}(i, x, y, \zeta_s(i, x))$ then set $\zeta_\tau = \zeta_s + \delta_{(i,y)} - \delta_{(i,x)}$.
4. If $\tau \in \mathcal{N}(i, j)$ then set $\zeta_\tau = \zeta_s + \sum_{x \in \mathbb{N}} \zeta_s(j, x)\delta_{(i,x)}$.

The process is then defined until time τ. Put $s = \tau$ and iterate to define ζ_t for all $t \geq 0$. Denote ζ_t^ζ the process with initial state ζ. We leave the reader to prove that ζ_t^ζ so defined is the branching-type process, that is, a Markov process with rates b and initial state ζ.

Let $|\zeta| := \sum_{i,x} \zeta(i, x)$ be the total number of BT-particles in ζ. Let

$$R(\zeta) := \max\left\{x \colon \sum_i \zeta(x, i) > 0\right\}.$$

Let $\widetilde{Z}_t^z$ be the process on $\mathbb{N}$ with rates $\tilde{q}$ and initial position $z \in \mathbb{N}$.

Lemma 3.1. $E|\zeta_t^\zeta| = |\zeta| e^{\bar{q}t}$.

Proof. $E|\zeta_t|$ satisfies the equation

$$\frac{\mathrm{d}}{\mathrm{d}t} E|\zeta_t| = \frac{\bar{q}}{N-1} E\left(\sum_i \sum_{j:j \neq i} \sum_x \zeta_t(j, x)\right) = \frac{\bar{q}}{N-1}(N-1)E|\zeta_t| = \bar{q}E|\zeta_t|, \tag{3.1}$$

with initial condition $E|\zeta_0| = |\zeta|$. $\qquad\square$

Lemma 3.2. *Let* $g : \mathbb{N} \to \mathbb{R}^+$ *be non-decreasing. Then*

$$Eg(R(\zeta_t^\zeta)) \leq E|\zeta_t^\zeta| Eg(\widetilde{Z}_t^{R(\zeta)}). \tag{3.2}$$

Proof. Consider the following partial order on $\mathcal{B}$:

$$\zeta \prec \zeta' \quad \text{if and only if} \quad \sum_{y \geq x} \zeta(i, y) \leq \sum_{y \geq x} \zeta'(i, y), \quad \text{for all } i, x. \tag{3.3}$$

The branching process is attractive: the Harris construction with initial configurations $\zeta \prec \zeta'$ gives $\zeta_t^\zeta \prec \zeta_t^{\zeta'}$ almost surely; we leave the proof to the reader. Let $\zeta' := \sum_{i,x} \zeta(i, x)\delta_{(i, R(\zeta))}$ be the configuration having the same number of BT-particles of type i as ζ for all i, but all are located at $r := R(\zeta)$. Hence $\zeta \prec \zeta'$ and

$$Eg(R(\zeta_t^\zeta)) \leq \sum_{i,x} g(x) E\zeta_t^\zeta(i, x) \leq \sum_x g(x) \sum_i E\zeta_t^{\zeta'}(i, x), \tag{3.4}$$

because g is non-decreasing. Fix i and x and define

$$b_t(r, x) := \sum_i E\zeta_t^{\zeta'}(i, x), \qquad a_t := E|\zeta_t^\zeta|, \qquad \tilde{p}_t(r, x) := P(\widetilde{Z}_t^r = x).$$

Since $b_t(r, x)$ and $a_t \tilde{p}_t(r, x)$ satisfy the same Kolmogorov backwards equations and have the same initial condition, the right-hand side of (3.4) coincides with the right-hand side of (3.2). This can be seen as an analogue of the many-to-one lemma, see [18]. □

Harris construction of Fleming–Viot

Let $\mathcal{N}(i, j, x) \subset \mathcal{N}(i, j)$ be the Poisson process obtained by independently including each $\tau \in \mathcal{N}(i, j)$ into $\mathcal{N}(i, j, x)$ with probability $q(x, 0)/\bar{q}$ (≤ 1, by definition of $\bar{q}$). The processes $(\mathcal{N}(i, j, x), i, j \in \{1, \ldots, N\}, x \in \mathbb{N})$ are independent Poisson processes of rate $q(x, 0)/(N - 1)$.

Fix $\xi_0 = \xi \in \mathbb{N}^{\{1,\ldots,N\}}$, assume the process is defined until time $s \geq 0$ and proceed iteratively from $s = 0$ as follows.

1. Define $\tau(\xi_s, s) = \inf\{t > s : t \in \bigcup_{i,y} \mathcal{N}(i, \xi_s(i), y, 1) \cup \bigcup_{i,j} \mathcal{N}(i, j, \xi_s(i))\}$.
2. For $t \in [s, \tau)$ define $\xi_t = \xi_s$.
3. If $\tau \in \mathcal{N}(i, \xi_s(i), y, 1)$, then set $\xi_\tau(i) = y$ and for $i' \neq i$ set $\xi_\tau(i') = \xi_s(i')$.
4. If $\tau \in \mathcal{N}(i, j, \xi_s(i))$, then set $\xi_\tau(i) = \xi_s(j)$ and for $i' \neq i$ set $\xi_\tau(i') = \xi_s(i')$.

The process is then defined until time τ. Put $s = \tau$ and iterate to define ξ_t for all $t \geq 0$. We leave the reader to prove that ξ_t^ξ is a Markov process with generator $\mathcal{L}^N$ and initial configuration ξ and the following lemma.

Lemma 3.3. *The Fleming–Viot i-particle coincides with the position of a branching i-BT-particle at time t if this happens at time zero for all i. More precisely,*

$$\zeta_0\big(i, \xi_0(i)\big) \geq 1 \quad \text{for all } i \text{ implies} \quad \zeta_t\big(i, \xi_t(i)\big) \geq 1 \quad \text{for all } i, \text{ a.s.} \tag{3.5}$$

Corollary 3.4. *Assume $\zeta_0(i, \xi_0(i)) \geq 1$ for all i. Then,*

$$R(\xi_t) \leq R(\zeta_t), \quad \text{a.s.} \tag{3.6}$$

4. Galton–Watson estimates

We show now that for ρ small enough the functions $e^{\rho \cdot}$ belong to the domain of the generator of Galton–Watson, that is, the Kolmogorov equations hold for these functions. The total number of births of the Galton–Watson process $(Z_t^x, t \geq 0)$, is a random variable $H^x := x + \sum_{t>0}(Z_t^x - Z_{t-}^x)^+$. Theorem 2 in [20] says that (2.1) and (2.2) are equivalent to the existence of a $\sigma' > 0$ such that $E(\exp(\sigma' H^1)) < \infty$. We assume $\sigma' \leq \sigma$ (introduced just before (2.1)). Let

$$\mathbf{F} := \left\{ f : \mathbb{N} \cup \{0\} \to \mathbb{R} : \sum_{\ell \geq 0} e^{-\rho \ell} |f(\ell)| < \infty \text{ for some } \rho < \sigma' \right\}. \tag{4.1}$$

Note that if $f \in \mathbf{F}$, then there exist $\rho < \sigma'$ and $C > 0$ such that $|f(\ell)| \leq C e^{\rho \ell}, \ell \geq 0$. For $f \in \mathbf{F}$ define the Galton–Watson semigroup by $S_t f(x) := E(f(Z_t^x)) < \infty$, since $Z_t^x \leq H^x$ for all $t \geq 0$. The generator Q of Galton–Watson applied on functions f is given by

$$Qf(x) := \sum_{\ell=-1}^\infty x p(\ell + 1)\big(f(x + \ell) - f(x)\big), \quad x \geq 0, \tag{4.2}$$

if the right-hand side is well defined.

Lemma 4.1. *Under the assumption* (2.1), *for $f \in \mathbf{F}$, $Qf(x)$ is well defined and the Kolmogorov equations hold*:

$$\frac{\mathrm{d}}{\mathrm{d}t} S_t f = Q S_t f = S_t Q f. \tag{4.3}$$

Proof. Since $|f(x)| \leq C \exp(\rho x)$ for all $x \in \mathbb{N}$,

$$|Qf(x)| \leq Cx\mathrm{e}^{\rho x}\left(\sum_{\ell \geq -1} p(\ell + 1)\mathrm{e}^{\rho \ell} + 1\right). \tag{4.4}$$

This shows the first part of the lemma. Consider $f \in \mathbf{F}$ and define the local martingale (see [22], Section IV-20, pp. 30–37)

$$M_t^x := f(Z_t^x) - f(x) - \int_0^t Qf(Z_s^x)\,\mathrm{d}s.$$

Using (4.4), for all $s \leq t$

$$|M_s^1| \leq \mathrm{e}^\rho + \exp(\rho H^1) + tCH^1 \exp(\rho H^1) \leq \tilde{C} \exp(\tilde{\rho}H^1),$$

with $\rho < \tilde{\rho} < \sigma'$. Hence $E \sup_{s \in [0,t]} |M_s^1| < \infty$ and M_t^1 is a martingale by dominated convergence. Since for $\rho \leq \sigma'$, $E\exp(\rho H^x) = (E\exp(\rho H^1))^x$, the same reasoning shows that M_t^x is a martingale and

$$Ef(Z_t^x) = f(x) + E\int Qf(Z_s^x)\,\mathrm{d}s,$$

which is equivalent to (4.3) for $f \in \mathbf{F}$. $\qquad\square$

The generator of the reflected Galton–Watson process $\widetilde{Z}_t$ reads

$$\widetilde{Q}f(x) := \sum_{\ell=-1}^{\infty} xp(\ell + 1)\mathbf{1}_{\{x+\ell \geq 1\}}(f(x+\ell) - f(x)), \quad x \in \mathbb{N}, \tag{4.5}$$

if the right-hand side is well defined. The reflected process can be thought of as an absorbed process regenerated at position 1 each time it gets extinct. Since the absorbed process can terminate only when it is at state 1 and jumps to 0 at rate $p(0)$, the number of regenerations until time t is dominated by a Poisson random variable $\mathcal{N}_t$ of mean $tp(0)$ and

$$E(\exp(\rho\widetilde{Z}_t^1)) \leq E\exp\left(\rho\sum_{n=1}^{\mathcal{N}_t} H_n^1\right),$$

where H_n^1 are i.i.d. random variables with the same distribution as H^1 and $\mathcal{N}_t$ is independent of $(H_n^1, n \geq 1)$. Hence, for some $C(\rho) > 0$

$$E(\exp(\rho\widetilde{Z}_t^1)) \leq \exp(tp(0)C(\rho)).$$

Let $\widetilde{S}_t$ be the semigroup of the reflected Galton–Watson process. Using the same reasoning as before, we obtain

Corollary 4.2. *Any $f \in \mathbf{F}$ satisfies the Kolmogorov equations for $\widetilde{Q}$:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{S}_t f = \widetilde{Q}\widetilde{S}_t f = \widetilde{S}_t\widetilde{Q}f. \tag{4.6}$$

Large deviations

We study $\widetilde{Z}_t$, the reflected Galton–Watson process with generator $\widetilde{Q}$ given by (4.5). Since p satisfies (2.1), for $\rho < \sigma' \leq \sigma$,

$$\Gamma(\rho) := p(0) + \sum_{\ell=1}^{\infty} p(\ell + 1)\ell^2\mathrm{e}^{\rho\ell} < \infty. \tag{4.7}$$

Recall that v is defined in (2.2) and define β as

$$\beta = \sup\{\rho > 0 \colon \rho \Gamma(\rho) \leq v\}, \tag{4.8}$$

which is well defined thanks to the exponential moment of p.

Lemma 4.3. *For any $\rho < \min\{\beta, \sigma'\}$, and $x \in \mathbb{N}$,*

$$E \exp(\rho \widetilde{Z}_t^x) \leq e^{-\rho v/2t} e^{\rho x} + t e^{\rho}. \tag{4.9}$$

Proof. Since $\rho < \sigma' \leq \sigma$, the reflected Galton–Watson generator (4.5) applied to $e^{\rho \cdot}$ is well defined and gives

$$\widetilde{Q}(e^{\rho \cdot})(x) = \sum_{\ell=-1}^{\infty} x p(\ell+1) e^{\rho x}(e^{\rho \ell} - 1) - p(0)\mathbf{1}_{\{x=1\}}(1 - e^{\rho})$$

$$= x e^{\rho x}\left(-\rho v + \sum_{\ell=-1}^{\infty} p(\ell+1)(e^{\rho \ell} - 1 - \rho \ell)\right) + p(0)\mathbf{1}_{\{x=1\}}(e^{\rho} - 1).$$

Since for $a \geq 0$, $e^a - (1+a) \leq \frac{a^2}{2}e^a$,

$$\widetilde{Q}(e^{\rho \cdot})(x) \leq \rho x e^{\rho x}\left(-v + \frac{\rho}{2}\Gamma(\rho)\right) + p(0)\mathbf{1}_{\{x=1\}}e^{\rho}$$

$$\leq -\frac{v\rho}{2}e^{\rho x} + e^{\rho}, \tag{4.10}$$

using $\rho < \beta$ and $\beta \Gamma(\beta) \leq v$. Since $\rho < \sigma'$, Corollary 4.2 and Gronwall's inequality give (4.9). $\qquad \square$

We obtain now a large deviation estimate.

Proposition 4.4. *Let $T \geq \frac{1}{4p(0)}$ and $\delta \geq 4T p(0) \geq 1$. Then, there is a constant κ, independent of x, such that*

$$P\left(\sup_{s < T}(\widetilde{Z}_s^x - e^{-vs}x) \geq \delta\right) \leq \exp\left(-\frac{\kappa}{T}\frac{\delta^2}{\max\{x, \delta\}}\right). \tag{4.11}$$

Proof. Set $z_t^x = e^{-vt}x$ and introduce the process

$$\epsilon_t^x := \widetilde{Z}_t^x - x + v\int_0^t \widetilde{Z}_s^x \, ds$$

$$= (\widetilde{Z}_t^x - z_t^x) + v\int_0^t (\widetilde{Z}_s^x - z_s^x) \, ds. \tag{4.12}$$

To stop $\widetilde{Z}_t^x$ when it crosses $2\max\{x, \delta\}$ define

$$\tau := \inf\{t \geq 0 \colon \widetilde{Z}_t^x \geq 2\max\{x, \delta\}\}. \tag{4.13}$$

Note that if $\tau < \infty$, then $\widetilde{Z}_\tau^x - z_\tau^x \geq 2\max\{x, \delta\} - x \geq \delta$. Thus,

$$\{\widetilde{Z}_t^x - z_t^x \geq \delta\} \subset \{\widetilde{Z}_{t \wedge \tau}^x - z_{t \wedge \tau}^x \geq \delta\}. \tag{4.14}$$

For functions $g_1, g_2 \colon \mathbb{R} \to \mathbb{R}$ verifying

$$g_1(t) = g_2(t) + v\int_0^t g_2(s) \, ds, \quad v \geq 0,$$

it holds

$$\sup_{t \leq T} |g_1(t)| \leq \frac{\delta}{2} \quad \Longrightarrow \quad \sup_{t \leq T} |g_2(t)| \leq \delta.$$

Hence,

$$\left\{ \sup_{t \leq T} |\widetilde{Z}^x_{t \wedge \tau} - z^x_{t \wedge \tau}| \geq \delta \right\} \subset \left\{ \sup_{t \leq T} |\epsilon^x_{t \wedge \tau}| \geq \frac{\delta}{2} \right\}. \tag{4.15}$$

Note that

$$\left\{ \sup_{t \leq T} |\epsilon^x_{t \wedge \tau}| \geq \frac{\delta}{2} \right\} = \left\{ \sup_{t \leq T} \epsilon^x_{t \wedge \tau} \geq \frac{\delta}{2} \right\} \cup \left\{ \inf_{t \leq T} \epsilon^x_{t \wedge \tau} \leq -\frac{\delta}{2} \right\}.$$

The treatment of the two terms on the right-hand side of the previous formula is similar, and we only give the simple argument for the first of them. For $\rho < \sigma'$, the following functional is a local martingale (see [12], page 66)

$$\mathcal{M}_t := \exp\left(\rho \widetilde{Z}^x_t - \rho x - \int_0^t (e^{-\rho \cdot} \widetilde{Q}(e^{\rho \cdot}))(\widetilde{Z}^x_s) \, ds \right). \tag{4.16}$$

Using the bounds of Lemma 4.1 we obtain that $\mathcal{M}_t$ is in fact a martingale. Observe that

$$e^{-\rho x} \widetilde{Q}(e^{\rho \cdot})(x) = x \sum_{\ell = -1}^\infty p(\ell+1)(e^{\rho \ell} - 1) + p(0) \mathbf{1}_{\{x=1\}}(e^\rho - 1)$$

$$\leq -\rho v x + \rho p(0) + \frac{\rho^2}{2} x \Delta(\rho) + p(0)(e^\rho - 1 - \rho), \tag{4.17}$$

with

$$\Delta(\rho) := \frac{2}{\rho^2} \sum_{\ell = -1}^\infty p(\ell+1)(e^{\rho \ell} - 1 - \rho \ell) \geq 0. \tag{4.18}$$

We have already seen that $\Delta(\rho) \leq \Gamma(\rho)$. Then, we bound the martingale $\mathcal{M}_t$ as follows

$$\mathcal{M}_t \geq \exp\left(\rho(\widetilde{Z}^x_t - x) - \left(-\rho v + \frac{\rho^2}{2} \Delta(\rho) \right) \int_0^t \widetilde{Z}^x_s \, ds - \rho p(0) t - p(0) t (e^\rho - 1 - \rho) \right)$$

$$\geq \exp\left(\rho \epsilon^x_t - \rho p(0) t - \frac{\rho^2}{2} \Gamma(\rho) \int_0^t \widetilde{Z}^x_s \, ds - p(0) t (e^\rho - 1 - \rho) \right). \tag{4.19}$$

By stopping the process at τ, and using that $\delta \geq 1$, we obtain for $t \leq T$

$$\exp(\rho \epsilon^x_{t \wedge \tau}) \leq \mathcal{M}_{t \wedge \tau} \exp\left(\rho p(0) T + \frac{\rho^2}{2} T p(0) \max\{x, \delta\} \left(\frac{\Gamma(\rho)}{p(0)} + \frac{e^\rho - 1 - \rho}{\rho^2/2} \right) \right).$$

Now, define $\alpha > 0$ (which is always possible) such that

$$\alpha \left(\frac{\Gamma(\alpha)}{p(0)} + \frac{e^\alpha - 1 - \alpha}{\alpha^2/2} \right) = 1, \tag{4.20}$$

and for $\rho \in]0, \alpha[$, and $p(0) T \leq \delta/4$, we have

$$\exp(\rho \epsilon^x_{t \wedge \tau}) \leq \mathcal{M}_{t \wedge \tau} \exp\left(\frac{\rho \delta}{4} + \frac{\rho^2}{2\alpha} T p(0) \max\{x, \delta\} \right). \tag{4.21}$$

We use (4.14), (4.15) and (4.21), with $\rho \in]0, \alpha[$, and Doob's martingale inequality.

$$P\left(\sup_{s \leq T}\left(\widetilde{Z}_s^x - z_s^x\right) \geq \delta\right) \leq P\left(\sup_{s \leq T} \mathcal{M}_{s \wedge \tau} \geq \exp\left(\frac{\rho\delta}{4} - \frac{p(0)T\rho^2}{2\alpha}\max\{x, \delta\}\right)\right)$$

$$\leq \exp\left(-\frac{\rho\delta}{8} + \frac{p(0)T\rho^2}{2\alpha}\max\{x, \delta\}\right). \tag{4.22}$$

We finally obtain an optimal $0 < \rho^* < \alpha$

$$\rho^* := \frac{\alpha}{4p(0)T}\frac{\delta}{\max\{x, \delta\}}. \tag{4.23}$$

The result follows now from (4.22) and (4.23) with $\kappa = \alpha/8p(0)$ which depends only on the offspring's law. □

5. Bounds for the rightmost Fleming–Viot-particle

In this section, we bound small exponential moments of the rightmost Fleming–Viot-particle. We first define a threshold A, such that with very small probability, the rightmost particle's position does not decrease when it is initially larger than A. Define

$$\gamma := \frac{1}{2}\left(1 - \exp\left(-\frac{v}{4p(0)}\right)\right) \in (0, 1).$$

Choose

$$\rho_0 := \frac{\min\{\beta, \sigma', \gamma\kappa p(0)\}}{4},$$

where κ is the constant given by Proposition 4.4. Define

$$A := \frac{2\kappa p(0)}{\rho_0} > 1 \quad \left(\text{and note that } \gamma \geq \frac{1}{A}\right). \tag{5.1}$$

Define the time and the error δ entering in the large deviation estimate of Proposition 4.4 as follows. For an arbitrary initial condition ξ,

$$T := \frac{1}{4p(0)} \quad \text{and} \quad \delta := \max\left\{1, \frac{R(\xi)}{A}\right\}, \tag{5.2}$$

recall here that $R(\xi) = \max_{i \leq N} \xi(i)$, and set $V_L(\xi) = \exp(\rho\min(R(\xi), L))$ for $L > A$ which will be taken to infinity later. We use the notation $[F(\xi_t)]_0^T := F(\xi_T) - F(\xi_0)$.

Proof of Proposition 1.3. We use the construction in Section 3 to couple the Fleming–Viot process ξ_t^ξ and the branching process ζ_t^ζ with $\zeta = \sum_i \delta_{(i, \xi(i))}$, so that $\zeta(i, \xi(i)) = 1$ for all i. Then, by (3.6) $R(\xi_t) \leq R(\zeta_t)$ and it is sufficient to prove an inequality like (1.6) for $R(\zeta_t)$. Notice that for the initial configurations ξ and ζ, $R(\xi) = R(\zeta)$. We drop the superscripts ξ and ζ in the remainder of this proof.

Define the event

$$\mathcal{G} = \mathcal{G}(\xi, T) := \left\{R(\zeta_T) - e^{-vT}R(\xi) \leq \delta\right\}, \tag{5.3}$$

and for a positive real c, we define the set

$$K_c := \left\{\xi: R(\xi) \leq c\right\}.$$

On K_A^c, $\delta = R/A < R$, and on $K_A^c \cap \mathcal{G}$,

$$R(\zeta_T) \leq \left(\frac{1}{A} + \mathrm{e}^{-vT}\right) R(\xi) \leq (1 - \gamma) R(\xi). \tag{5.4}$$

Hence,

$$\mathbf{1}_{K_A^c \cap \mathcal{G}} \big[V_L(\zeta_t)\big]_0^T \leq V_L(\xi)\big(\mathrm{e}^{-\gamma\rho R(\xi)} - 1\big)\mathbf{1}_{K_A^c \cap K_L \cap \mathcal{G}} \leq -V_L(\xi)\big(1 - \mathrm{e}^{-\gamma\rho A}\big)\mathbf{1}_{K_A^c \cap K_L \cap \mathcal{G}}. \tag{5.5}$$

Since $A > 1$, on $K_A \cap \mathcal{G}$, $R(\zeta_T) \leq A\mathrm{e}^{-vT} + 1 \leq 2A$ so that

$$\mathbf{1}_{K_A \cap \mathcal{G}} \big[\mathrm{e}^{\rho R(\zeta_t)}\big]_0^T \leq \mathrm{e}^{2\rho A} \mathbf{1}_{K_A \cap \mathcal{G}}.$$

Thus

$$\big[V_L(\zeta_t)\big]_0^T \leq -\big(1 - \mathrm{e}^{-\gamma\rho A}\big)\mathrm{e}^{\rho R(\xi)}\mathbf{1}_{K_A^c \cap K_L \cap \mathcal{G}} + \mathrm{e}^{2\rho A}\mathbf{1}_{K_A \cap \mathcal{G}} + \big[\mathrm{e}^{\rho R(\zeta)}\big]_0^T \mathbf{1}_{\mathcal{G}^c}$$

$$\leq -\big(1 - \mathrm{e}^{-\gamma\rho A}\big)V_L(\xi)\mathbf{1}_{K_A^c \cap K_L} + \mathrm{e}^{2\rho A}\mathbf{1}_{K_A} + 2\mathrm{e}^{\rho R(\zeta_T)}\mathbf{1}_{\mathcal{G}^c}, \tag{5.6}$$

where we used that

$$\mathbf{1}_{K_A^c \cap K_L} - \mathbf{1}_{K_A^c \cap K_L \cap \mathcal{G}} \leq \mathbf{1}_{\mathcal{G}^c}.$$

Choose $\rho := \min(\rho_0, \frac{\kappa}{4TA^2})$ and observe that by Lemma 3.2,

$$E\big[\mathrm{e}^{2\rho R(\zeta_T)}\big] \leq E|\zeta_T| E\big[\exp\big(2\rho \widetilde{Z}_T^{R(\xi)}\big)\big] \leq N\mathrm{e}^{p(0)T}\big(\mathrm{e}^{-\rho vT}\mathrm{e}^{2\rho R(\xi)} + T\mathrm{e}^{2\rho}\big),$$

by Lemma 3.1 for the bound of the first factor and Lemma 4.3 for the bound of the second factor. Also, Proposition 4.4 implies

$$P\big(\mathcal{G}^c\big) \leq E|\zeta_T| P\Big(\sup_{s<T}\big(\widetilde{Z}_s^{R(\xi)} - \mathrm{e}^{-vs}R(\xi)\big) > \delta\Big) \leq N\mathrm{e}^{p(0)T}\big(\mathrm{e}^{-(\kappa/TA)}\mathbf{1}_{R(\xi)\leq A} + \mathrm{e}^{-(\kappa R(\xi)/TA^2)}\mathbf{1}_{R(\xi)>A}\big).$$

Taking expectation on (5.6) we bound the last term as follows. For constants C_1, C_2, $\tilde{C}_1$, and $\tilde{C}_2$

$$E\big[\mathrm{e}^{\rho R(\zeta_T)}\mathbf{1}_{\{\mathcal{G}^c\}}\big] \leq \big(P\big(\mathcal{G}^c\big)E\big[\mathrm{e}^{2\rho R(\zeta_T)}\big]\big)^{1/2}$$

$$\leq N\mathrm{e}^{p(0)T}\left(C_1\mathbf{1}_{K_A} + C_2\exp\left(-\frac{\kappa R(\xi)}{2TA^2}\right)\mathbf{1}_{K_A^c}\right)^{1/2}$$

$$\leq \tilde{C}_1 N\mathbf{1}_{K_A} + \tilde{C}_2 N\exp\left(-\frac{\kappa R(\xi)}{4TA^2}\right)\mathbf{1}_{K_A^c}. \tag{5.7}$$

Gathering (5.7) and (5.6) we obtain, for any $L > A$,

$$EV_L\big(\xi_T^\xi\big) - V_L(\xi) < -c_1 V_L(\xi)\mathbf{1}_{\{L>R(\xi)>A\}} + C_1 N\mathbf{1}_{\{R(\xi)\leq A\}} + C_2 N\mathrm{e}^{-\rho\tilde{c}_2 R(\xi)}$$

$$\leq -c_1 V_L(\xi)\mathbf{1}_{\{L>R(\xi)>A\}} + C N\mathrm{e}^{-c_2 R(\xi)} \tag{5.8}$$

which completes the proof of inequality (1.6), as one takes L to infinity in (5.8). □

Proof of Corollary 1.4. Take $\tilde{C} > 0$ and observe that the set of ξ such that the right-hand side of (5.8) is larger than $-\tilde{C}$ is finite. Foster's criteria (Theorem A.1 in the Appendix) implies that the process ξ_t is ergodic with an invariant measure that we call λ^N.

Now, consider again (5.8) for a fixed L. Note that V_L is bounded, so that by integrating (5.8) with this invariant measure, and then taking L to infinity, we obtain (1.7). □

6. The empirical moments of Fleming–Viot

In this section we prove Proposition 1.5 and Corollary 1.6. Introduce the occupation numbers $\eta : \mathbb{N} \times \mathbb{N}^N \to \mathbb{N}$ defined as

$$\eta(x, \xi) := \sum_{i=1}^{N} \mathbf{1}_{\{\xi(i)=x\}},$$

for which we often drop the coordinate ξ. Notice that $m(x, \xi) = \eta(x, \xi)/N$.

For any integer k, define the kth moment of the N particles' positions as

$$M_k(\xi) := \sum_{i=1}^{N} \xi^k(i) = \sum_{x=1}^{\infty} x^k \eta(x, \xi).$$

As there are only N particles, M_k is well defined. Observe that $\psi = M_2/M_1$. Note the inequalities

$$1 \le \frac{M_1(\xi)}{N} \le \psi(\xi) \le R(\xi). \tag{6.1}$$

The function ψ is not compactly supported (nor bounded). Even though $\mathcal{L}^N \psi$ is well defined, we need to use later that $\int \mathcal{L}^N \psi \, d\lambda^N = 0$. We do so by approximating ψ by a function ψ^L for which we have

$$\int \mathcal{L}^N \psi^L \, d\lambda^N = 0 \quad \text{and} \quad \lim_{L \to \infty} \mathcal{L}^N \psi^L = \mathcal{L}^N \psi \quad \text{pointwise}, \tag{6.2}$$

where

$$\psi^L(\xi) := \frac{M_2^L(\xi)}{M_1^L(\xi)}, \quad \text{with } M_k^L(\xi) := \sum_{i=1}^{N} \min(\xi^k(i), L^k) = \sum_{x=1}^{L} x^k \eta(x, \xi) + L^k \sum_{x>L} \eta(x, \xi). \tag{6.3}$$

Observe that M_k^L and ψ^L depend only on $(\eta(1), \dots, \eta(L))$. It is easy, and we omit the proof, to see that there exist a positive constant C such that

$$\left| \mathcal{L}^N \psi - \mathcal{L}^N \psi^L \right| \le \left| \mathcal{L}^N \psi \right| + \left| \mathcal{L}^N \psi^L \right| \le C\psi \le CR, \tag{6.4}$$

where we recall that $R(\xi) = \max_i \xi(i)$. We have established in Proposition 1.3 that $R(\xi)$ is integrable with respect to λ^N, so that (6.2) implies that

$$\int \mathcal{L}^N \psi \, d\lambda^N = 0. \tag{6.5}$$

Proof of Proposition 1.5. We decompose the generator (1.3) into two generators, one governing the refeed part and the other the spatial evolution of the particles: $\mathcal{L}^N = \mathcal{L}_{\text{drift}}^N + \mathcal{L}_{\text{refeed}}^N$, which applied to functions depending on ξ only through $\eta(\cdot, \xi)$, read

$$\mathcal{L}_{\text{refeed}}^N = p(0)\eta(1) \sum_{x=1}^{\infty} \frac{\eta(x)}{N-1} \left(A_1^- A_x^+ - \mathbf{1} \right), \quad \text{with } A_x^{\pm}(\eta)(y) = \begin{cases} \eta(y), & y \ne x, \\ \eta(x) \pm 1, & y = x, \end{cases} \tag{6.6}$$

$$\mathcal{L}_{\text{drift}}^N = \sum_{x=2}^{\infty} x\eta(x) p(0) \left(A_x^- A_{x-1}^+ - \mathbf{1} \right) + \sum_{x=1}^{\infty} x\eta(x) \sum_{i=1}^{\infty} p(i+1) \left(A_x^- A_{x+i}^+ - \mathbf{1} \right). \tag{6.7}$$

It is convenient to introduce a boundary term

$$B = -\eta(1) p(0) \left(A_1^- A_0^+ - \mathbf{1} \right) \quad \text{and call} \quad \mathcal{L}_0^N = \mathcal{L}_{\text{drift}}^N - B, \tag{6.8}$$

which applied on ψ yield

$$B\psi = -p(0)\eta(1)\left(\frac{M_2 - M_1}{M_1(M_1 - 1)}\right); \tag{6.9}$$

$$\mathcal{L}_0^N \psi = \sum_{x=1}^{\infty} x\eta(x) \sum_{i=-1}^{\infty} p(i+1)\left(\frac{M_2 + 2ix + i^2}{M_1 + i} - \frac{M_2}{M_1}\right)$$

$$= \sum_{x=1}^{\infty} x\eta(x) \left\{ \sum_{i=-1}^{\infty} ip(i+1)\left(\frac{2xM_1 - M_2 + iM_1}{M_1(M_1 + i)}\right) \right\}$$

$$= -p(0)\frac{M_2 - M_1}{M_1 - 1} + \left(\sum_{i=1}^{\infty} p(i+1)i\frac{M_1}{M_1 + i}\right) \times \frac{M_2}{M_1} + \sum_{i=1}^{\infty} p(i+1)i^2\frac{M_1}{M_1 + i}$$

$$\leq -v\psi + p(0)\frac{M_1}{M_1 - 1} + \sum_{i=1}^{\infty} p(i+1)i^2 \leq -v\psi + C_0, \tag{6.10}$$

for some positive constant C_0. Finally, for the jump term

$$\mathcal{L}_{\text{refeed}}^N \psi = p(0)\eta(1) \sum_{x=1}^{\infty} \frac{\eta(x)}{N-1}\left(\frac{M_2 + x^2 - 1}{M_1 + x - 1} - \frac{M_2}{M_1}\right)$$

$$= p(0)\eta(1) \sum_{x=1}^{\infty} \frac{\eta(x)}{N-1} \frac{M_1(x^2 - 1) - M_2(x-1)}{M_1(M_1 - 1)} \times \frac{1}{1 + x/(M_1 - 1)}. \tag{6.11}$$

If we set $\Delta(x) = 1/(1 + x) - (1 - x)$, for $x \in [0, 1]$, then

$$\Delta(x) = \frac{x^2}{1 + x} \quad \text{and} \quad 0 \leq \Delta(x) \leq x^2. \tag{6.12}$$

We apply (6.12) to expand the last term in (6.11), with $x/(M_1 - 1) \leq 1$ for $x \leq R(\xi)$, and obtain

$$\mathcal{L}_{\text{refeed}}^N \psi = p(0)\eta(1) \sum_{x=1}^{\infty} \frac{\eta(x)}{N-1} \frac{M_1(x^2 - 1) - M_2(x-1)}{M_1(M_1 - 1)} \times \left(1 - \frac{x}{M_1 - 1} + \Delta\left(\frac{x}{M_1 - 1}\right)\right). \tag{6.13}$$

Note that

$$\sum_{x=1}^{\infty} \eta(x)(M_1 x^2 - M_2 x) = 0 \quad \text{and} \quad \sum_{x=1}^{\infty} \eta(x)(M_1 x^2 - M_2 x)(-x) = -M_3 M_1 + (M_2)^2.$$

Also,

$$\sum_{x=1}^{\infty} \frac{\eta(x)}{N-1}(M_2 - M_1)\left(1 - \frac{x}{M_1 - 1}\right) = \left(N - \frac{M_1}{M_1 - 1}\right)\frac{(M_2 - M_1)}{N-1}$$

$$= \left(1 - \frac{1}{(N-1)(M_1 - 1)}\right)(M_2 - M_1)$$

$$= (M_2 - M_1) - \frac{M_2 - M_1}{(N-1)(M_1 - 1)}.$$

Thus

$$\mathcal{L}_{\text{refeed}}^N \psi = -p(0)\frac{\eta(1)}{N-1}\frac{M_3 M_1 - (M_2)^2}{M_1(M_1 - 1)^2} + p(0)\eta(1)\frac{M_2 - M_1}{M_1(M_1 - 1)} + \text{Rest},$$

where

$$\text{Rest} = -\frac{p(0)\eta(1)(M_2 - M_1)}{(N-1)M_1(M_1-1)^2} + p(0)\eta(1)\sum_{x=1}^{\infty}\frac{\eta(x)}{N-1}\frac{M_1(x^2-1) - M_2(x-1)}{M_1(M_1-1)} \times \Delta\left(\frac{x}{M_1-1}\right). \quad (6.14)$$

Using that $M_2 - M_1 \geq 0$,

$$\text{Rest} \leq p(0)\frac{\eta(1)}{N-1}\sum_{x=1}^{\infty}\eta(x) \times \frac{x^2}{(M_1-1)^2}\left(\frac{M_1 x^2 + M_2 x}{(M_1-1)^2} + \frac{M_2 - M_1}{(M_1-1)^2}\right) \quad (6.15)$$

$$\leq 3p(0)\left(\frac{M_1}{M_1-1}\right)^4\frac{\eta(1)}{N-1}\frac{M_2}{(M_1)^2}\frac{R^2}{M_1} \leq 24p(0)\frac{R^2}{N}. \quad (6.16)$$

Observe that

$$\mathcal{L}_{\text{refeed}}^N\psi + B\psi = -p(0)\frac{\eta(1)}{N-1}\frac{M_3 M_1 - (M_2)^2}{M_1(M_1-1)^2} + \text{Rest} \leq \text{Rest}.$$

Thus, we reach that for C_0 independent of N,

$$\mathcal{L}^N\psi \leq -v\psi + 24p(0)\frac{R^2}{N} + C_0. \quad (6.17)$$

This is Proposition 1.5. We now integrate (6.17) with respect to the invariant measure, and use that $\int \mathcal{L}^N\psi \, d\lambda^N = 0$ to obtain for constants C_1, and C_2 (independent of N)

$$\int \psi \, d\lambda^N \leq C_1 + C_2\frac{\int R^2 \, d\lambda^N}{N}. \quad (6.18)$$

$\square$

7. Uniform convergence to the Yaglom limit

In this section we show a uniform convergence for $\mu \in K(\alpha)$ of the generating functions of μT_t to the generating function of the quasi-stationary distribution v, where the generating function of a distribution μ on $\mathbb{N}$ is defined by

$$G(\mu; z) := \sum_{x \in \mathbb{N}}\mu(x)z^x, \quad z \in \mathbb{R}, |z| < 1. \quad (7.1)$$

We invoke a key result of Yaglom [25]. The continuous time version can be found in Zolotarev [26].

Lemma 7.1 (Yaglom [25], Zolotarev [26]). *There is a probability measure v such that*

$$\lim_{t \to \infty} G(\delta_1 T_t; z) = G(v; z), \quad (7.2)$$

and the generating function of v is given by

$$G(v; z) = 1 - \exp\left(-v\int_0^z \frac{du}{\sum_{\ell \geq 0}p(\ell)u^\ell - z}\right), \quad z \in [0, 1). \quad (7.3)$$

The measure v is in fact v_{qs}^*, the minimal QSD. We do not use the explicit expression (7.3) of the generating function of v; we only use (7.2). Recall that μT_t is the law of Z_t with initial distribution μ conditioned on survival until t and

that $K(\alpha)$ is defined in (2.5). The next result says that the Yaglom limit holds uniformly for all initial measures in $K(\alpha)$.

Proposition 7.2. *For any $\alpha > 0$, and $z \in [0, 1]$*

$$\lim_{t \to \infty} \sup_{\mu \in K(\alpha)} \left| G(\mu T_t; z) - G\left(\nu_{\mathrm{qs}}^*; z\right) \right| = 0. \tag{7.4}$$

As a consequence, for each $x \in \mathbb{N}$, we obtain (2.8).

Proof. Recall that S_t is the semigroup of the Galton–Watson process and observe that for any $\ell \in \mathbb{N}$, $G(\delta_\ell S_t; z) = G^\ell(\delta_1 S_t; z)$. We set, for simplicity,

$$g(z) := 1 - G(\delta_1 S_t; z) \in [0, 1],$$

for $z \in [0, 1]$. The following inequalities are useful. For $z \in [0, 1]$,

$$1 - \ell g(z) \le \left(1 - g(z)\right)^\ell \le 1 - \ell g(z) + \ell^2 g^2(z). \tag{7.5}$$

The generating function of μT_t reads as follows.

$$\begin{aligned}
G(\mu T_t; z) &= \frac{G(\mu S_t; z) - G(\mu S_t; 0)}{1 - G(\mu S_t; 0)} \\
&= \frac{\sum_{\ell \ge 1} \mu(\ell)(G(\delta_\ell S_t; z) - G(\delta_\ell S_t; 0))}{\sum_{\ell \ge 1} \mu(\ell)(1 - G(\delta_\ell S_t; 0))} \\
&= \frac{\sum_{\ell \ge 1} \mu(\ell)((1 - g(z))^\ell - (1 - g(0))^\ell)}{\sum_{\ell \ge 1} \mu(\ell)(1 - (1 - g(0))^\ell)}.
\end{aligned} \tag{7.6}$$

Also,

$$1 - G(\delta_1 T_t; z) = \frac{1 - G(\delta_1 S_t; z)}{1 - G(\delta_1 S_t; 0)} = \frac{g(z)}{g(0)}.$$

We now produce upper and lower bounds for $G(\mu T_t; z) - G(\nu_{\mathrm{qs}}^*; z)$. We start with the upper bound. Using first (7.6) and then (7.5),

$$\begin{aligned}
G(\mu T_t; z) - G\left(\nu_{\mathrm{qs}}^*; z\right) &= \frac{\sum_{\ell \ge 1} \mu(\ell)((1 - g(z))^\ell - 1 + (1 - (1 - g(0))^\ell)(1 - G(\nu_{\mathrm{qs}}^*; z)))}{\sum_{\ell \ge 1} \mu(\ell)(1 - (1 - g(0))^\ell)} \\
&\le \frac{\sum_\ell \ell \mu(\ell)(-g(z) + \ell g^2(z) + g(0)(1 - G(\nu_{\mathrm{qs}}^*; z)))}{\sum_\ell \ell \mu(\ell)(g(0) - \ell g^2(0))} \\
&\le \frac{\sum_\ell \ell \mu(\ell)((1 - G(\nu_{\mathrm{qs}}^*; z)) - g(z)/g(0)) + \sum_\ell \ell^2 \mu(\ell)(g(z)/g(0))g(z)}{\sum_\ell \ell \mu(\ell)(1 - \ell g(0))} \\
&\le \frac{G(\delta_1 T_t; z) - G(\nu_{\mathrm{qs}}^*; z) + (M_2(\mu)/M_1(\mu))(1 - G(\delta_1 T_t; z))g(z)}{1 - (M_2(\mu)/M_1(\mu))g(0)},
\end{aligned} \tag{7.7}$$

where $M_k(\mu) := \sum_\ell \ell^k \mu(\ell)$, $k \in \mathbb{N}$. Thus,

$$\sup_{\mu \in K(\alpha)} G(\mu T_t; z) - G\left(\nu_{\mathrm{qs}}^*; z\right) \le \frac{|G(\delta_1 T_t; z) - G(\nu_{\mathrm{qs}}^*; z)| + (1 - G(\delta_1 T_t; z))g(z)\alpha}{1 - \alpha g(0)}. \tag{7.8}$$

Now, for the lower bound, we use similar arguments to reach

$$G(\mu T_t; z) - G(v_{qs}^*; z) \geq \frac{\sum_\ell \ell \mu(\ell)(-g(z) + g(0)(1 - G(v_{qs}^*; z)) - \ell g^2(0)(1 - G(v_{qs}^*; z)))}{\sum_\ell \ell \mu(\ell) g(0)}$$

$$\geq G(\delta_1 T_t; z) - G(v_{qs}^*; z) - \frac{M_2(\mu)}{M_1(\mu)} g(0)(1 - G(v_{qs}^*; z)). \tag{7.9}$$

Thus,

$$\inf_{\mu \in K(\alpha)} G(\mu T_t; z) - G(v_{qs}^*; z) \geq -|G(\delta_1 T_t; z) - G(v_{qs}^*; z)| - \alpha g(0)(1 - G(v_{qs}^*; z)). \tag{7.10}$$

Since $g(z)$ goes to 0 as the implicit t goes to infinity, both (7.8) and (7.10) go to 0. This proves (7.4). The proof of (2.8) follows from (7.4) and Lemma 7.3 below on convergence of probability measures. □

Lemma 7.3. *Let $\{\mu_n^\gamma, n \in \mathbb{N}, \gamma \in \Gamma\}$ be a family of probability measures. Assume that for each $z \in [0, 1]$ we have*

$$\lim_{n \to \infty} \sup_{\gamma \in \Gamma} |G(\mu_n^\gamma, z) - G(v, z)| = 0. \tag{7.11}$$

Then, for each $x \in \mathbb{N}$ we have

$$\lim_{n \to \infty} \sup_{\gamma \in \Gamma} |\mu_n^\gamma(x) - v(x)| = 0.$$

Proof. Let $f = \mathbf{1}_{\{x\}}$. We consider the one-point compactification of $\mathbb{N}$, which we denote $\bar{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$ and extend $f : \bar{\mathbb{N}} \to \mathbb{R}$ by $f(\infty) = 0$. Since f is a continuous function on $\bar{\mathbb{N}}$, the Stone–Weierstrass approximation theorem yields a function h, which is a linear combination of functions of the form $\{y \mapsto a^y, 0 \leq a \leq 1\}$ (finite linear combinations of these functions form an algebra that separates points and contains the constants), such that for any $\epsilon > 0$, $\sup_{y \in \mathbb{N}} |f(y) - h(y)| < \epsilon$. Then

$$\sup_\gamma |\mu_n^\gamma(x) - v(x)| = \sup_\gamma |\mu_n^\gamma f - v f| \leq \sup_\gamma |\mu_n^\gamma f - \mu_n^\gamma h| + \sup_\gamma |\mu_n^\gamma h - v h| + |v h - v f|.$$

The first and the third term on the right-hand side are smaller than ϵ while the second one goes to zero as n goes to infinity by assumption. □

8. Closeness of the two semi-groups

In this section we show how propagation of chaos implies the closeness of $Em(x, \xi_t^\xi)$ and $m(\cdot, \xi) T_t$ uniformly in $\xi \in \Lambda^N$. The arguments are similar to those used in [1,14]. The key is a control of the correlations that we state below. For a signed measure μ in $\mathbb{N}$ we will need to work with the ℓ_2 norm given by $\|\mu\|^2 = \sum_{x \in \mathbb{N}} (\mu(x))^2$.

Proposition 8.1. *There exist constants c and C such that,*

$$\sup_{\xi \in \mathbb{N}^N} \|E[m(x, \xi_t^\xi)] - m(\cdot, \xi) T_t\| \leq \frac{Ce^{ct}}{N}. \tag{8.1}$$

As a consequence,

$$\sup_{\xi \in \mathbb{N}^N} |E[m(x, \xi_t^\xi)] - m(\cdot, \xi) T_t(x)| \leq \frac{Ce^{ct}}{N}, \quad x \in \mathbb{N}. \tag{8.2}$$

Furthermore

$$\sup_{\xi \in \mathbb{N}^N} E\big(\big|m\big(x, \xi_t^\xi\big) - m(\cdot, \xi) T_t\big|^2\big) \leq \frac{Ce^{ct}}{N}, \quad x \in \mathbb{N}. \tag{8.3}$$

Proposition 8.2 (Proposition 2 of [1]). *For each $t > 0$, and any $x, y \in \mathbb{N}$*

$$\sup_{\xi \in \mathbb{N}^N} \big|E\big[m\big(x, \xi_t^\xi\big)m\big(y, \xi_t^\xi\big)\big] - E\big[m\big(x, \xi_t^\xi\big)\big]E\big[m\big(y, \xi_t^\xi\big)\big]\big| \leq \frac{2p(0)e^{2p(0)t}}{N}. \tag{8.4}$$

This proposition is proved in [1] for processes with bounded rates; the extension to our case is straightforward, and we omit its proof.

Proof of Proposition 8.1. Fix $\xi \in \mathbb{N}^N$ and introduce the simplifying notations

$$u(t, x) := Em\big(x, \xi_t^\xi\big) \quad \text{and} \quad v(t, x) := m(\cdot, \xi) T_t(x). \tag{8.5}$$

Define $\delta(t, x) = u(t, x) - v(t, x)$. We want to show that for any $t > 0$,

$$\frac{\partial}{\partial t}\big\|\delta(t)\big\|^2 \leq \frac{5}{2}\big\|\delta(t)\big\|^2 + \frac{4p(0)e^{2p(0)t}}{N}. \tag{8.6}$$

Recall the definition (2.3) of the rates q and the evolution equations satisfied by $v(t, x)$ and $u(t, x)$. The equation for v is written in (1.1), while the one for u comes from $\partial_t u(t, x) = E[\mathcal{L}^N m(x, \xi_t^\xi)]$ with $\mathcal{L}^N$ given in (1.3).

$$\frac{\partial}{\partial t}v(t, x) = \sum_{z \neq x, z > 0} q(z, x)v(t, z) - \bigg(\sum_{z \neq x} q(x, z)\bigg)v(t, x) + p(0)v(t, 1)v(t, x), \tag{8.7}$$

$$\frac{\partial}{\partial t}u(t, x) = \sum_{z \neq x, z > 0} q(z, x)u(t, z) - \bigg(\sum_{z \neq x} q(x, z)\bigg)u(t, x) + p(0)u(t, 1)u(t, x) + W(\xi; t, x). \tag{8.8}$$

Here,

$$W(\xi; t, x) = p(0)\bigg(\frac{N}{N-1}E\big[m\big(x, \xi_t^\xi\big)m\big(1, \xi_t^\xi\big)\big] - E\big[m\big(1, \xi_t^\xi\big)\big]E\big[m\big(x, \xi_t^\xi\big)\big]\bigg). \tag{8.9}$$

Proposition 8.2 implies that

$$\sup_\xi\big|W(\xi; t, x)\big| \leq \frac{2p(0)e^{2p(0)t}}{N}. \tag{8.10}$$

Observe two simple facts. First, set $D = \{(x, z): x \geq 1, z \geq 1, x \neq z\}$, and for any function $f: \mathbb{N} \to \mathbb{R}$

$$\sum_{(x,z) \in D}\big(q(x, z) + q(z, x)\big)f^2(x) - 2\sum_{(x,z) \in D} q(x, z)f(x)f(z) = \sum_{(x,z) \in D} q(z, x)\big(f(x) - f(z)\big)^2. \tag{8.11}$$

The second observation is specific to our rates. For $x > 0$

$$\sum_{z \neq x} q(z, x) \leq \sum_{z \neq x} q(x, z) + p(0). \tag{8.12}$$

Observation (8.11) is obvious and we omit its proof. Observation (8.12) is done in details.

$$
\begin{aligned}
\sum_{z \neq x} q(z, x) &= \sum_{z \geq 0, z \neq x} z p(x - z + 1) = x \sum_{z \geq 0, z \neq x} p(x - z + 1) + \sum_{z \geq 0, z \neq x} (z - x) p(x - z + 1) \\
&= x \big(p(0) + p(1) + \cdots + p(x + 1) \big) + \big(p(0) - p(2) - \cdots - x p(x + 1) \big) \\
&\leq x \sum_{i \geq 0} p(i) + p(0) = \sum_{z \neq x} q(x, z) + p(0).
\end{aligned}
\tag{8.13}
$$

Now, we have

$$
\begin{aligned}
\sum_{x > 0} \delta(t, x) \frac{\partial}{\partial t} \delta(t, x) = \sum_{(x, z) \in D} \big(q(z, x) \delta(t, x) \delta(t, z) - q(x, z) \delta^2(t, x) \big) \\
+ p(0) \sum_{x > 0} \big(u(t, x) u(t, 1) - v(t, x) v(t, 1) \big) \delta(t, x) + \sum_{x > 0} \delta(t, x) W(\xi; t, x).
\end{aligned}
\tag{8.14}
$$

Let us deal with each term of the right-hand side of (8.14). For the first term we use (8.11) and (8.12).

$$
\begin{aligned}
\sum_{(x, z) \in D} &\big(q(z, x) \delta(t, x) \delta(t, z) - q(x, z) \delta^2(t, x) \big) \\
&\leq \sum_{(x, z) \in D} q(z, x) \delta(t, x) \delta(t, z) - \frac{1}{2} \sum_{x > 0} \bigg(\sum_{z \neq x} q(x, z) + \sum_{z \neq x} q(z, x) - p(0) \bigg) \delta^2(t, x) \\
&\leq -\frac{1}{2} \sum_{(x, z) \in D} q(z, x) \big(\delta(t, x) - \delta(t, z) \big)^2 + \frac{p(0)}{2} \big\| \delta(t) \big\|^2 \\
&\leq \frac{p(0)}{2} \big\| \delta(t) \big\|^2.
\end{aligned}
\tag{8.15}
$$

To deal with the second term, first note that

$$
\sup_{x > 0} \big| \delta(t, x) \big| \leq \sqrt{\sum_{x > 0} \delta^2(t, x)} = \big\| \delta(t) \big\|.
$$

Then,

$$
\begin{aligned}
\sum_{x > 0} \big(u(t, x) u(t, 1) - v(t, x) v(t, 1) \big) \delta(t, x) &\leq \sum_{x > 0} \big(\delta(t, x) u(t, 1) + v(t, x) \delta(t, 1) \big) \delta(t, x) \\
&\leq \sum_{x > 0} \delta^2(t, x) + \big| \delta(t, 1) \big| \sup_{x > 0} \big| \delta(t, x) \big| \sum_{x > 0} v(t, x) \leq 2 \big\| \delta(t) \big\|^2.
\end{aligned}
\tag{8.16}
$$

For the last term, we have

$$
\bigg| \sum_{x > 0} \delta(t, x) W(\xi; t, x) \bigg| \leq \sup_{x > 0} \big| W(\xi; t, x) \big| \times \sum_{x > 0} \big| \delta(t, x) \big| \leq 2 \sup_{x > 0} \big| W(\xi; t, x) \big|.
\tag{8.17}
$$

Thus, we obtain (8.6). Gronwall's inequality allows to conclude (8.1), which implies (8.2) easily. Statement (8.3) is obtained using triangle inequality, (8.2) and (8.4) with $y = x$. $\qquad \square$

Appendix: Foster's criterion

Theorem A.1 (Foster's criteria, [21], Thm 8.13). *Let (ξ_t) be a Markov process with countable state space Λ. If there exist a function $V : \Lambda \to \mathbb{R}_+$ and constants $T, K, \epsilon > 0$ such that for $V(\xi) > K$,*

$$E\big(V(\xi_T)\big) - V(\xi) \leq -\epsilon,$$

with $\{\xi : V(\xi) \leq K\}$ a finite subset and $E(\xi_T^\xi) < \infty$ for $\xi \in \Lambda$, then (ξ_t) is ergodic.

Acknowledgements

References

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

[2] J. Bérard and J. B. Gouéré. Brunet–Derrida behavior of branching-selection particles systems on the line. *Comm. Math. Phys.* **298** (2) (2010) 323–342. MR2669438

[3] J. Berestycki, N. Berestycki and J. Schweinsberg. The genealogy of branching Brownian motion with absorption. *Ann. Probab.* **41** (2) (2013) 527–618. MR3077519

[4] M. Bieniek, K. Burdzy and S. Finch. Non-extinction of a Fleming–Viot particle model. *Probab. Theory Related Fields* **153** (1–2) 293–332. MR2925576

[5] E. Brunet and B. Derrida. Effect of microscopic noise on front propagation. *J. Stat. Phys.* **103** (1–2) (2001) 269–282. MR1828730

[6] E. Brunet and B. Derrida. Shift in the velocity of a front due to a cutoff. *Phys. Rev. E (3)* **56** (3) (1997), part A, 2597–2604. MR1473413

[7] E. Brunet, B. Derrida, A. H. Mueller and S. Munier. Noisy traveling waves: Effect of selection on genealogies. *Europhys. Lett.* **76** (1) (2006) 1–7. MR2299937

[8] E. Brunet, B. Derrida, A. H. Mueller and S. Munier. Effect of selection on ancestry: An exactly soluble case and its phenomenological generalization. *Phys. Rev. E (3)* **76** (4) (2007) 1–20. MR2365627

[9] K. Burdzy, R. Holyst, D. Ingerman and P. March. Configurational transition in a Fleming–Viot-type model and probabilistic interpretation of Laplacian eigenfunctions. *J. Phys. A: Math. Gen.* **29** (1996) 2633–2642.

[10] J. A. Cavender. Quasi-stationary distributions of birth-and-death processes. *Adv. in Appl. Probab.* **10** (3) (1978) 570–586. MR0501388

[11] R. Durrett and D. Remenik. Brunet–Derrida particles systems, free boundary problems and Wiener–Hopf equations. *Ann. Probab.* **39** (6) (2011) 2043–2078. MR2932664

[12] S. N. Ethier and T. G. Kurtz. *Markov Processes. Characterization and Convergence. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. Wiley, New York, 1986. MR0838085

[13] P. A. Ferrari, H. Kesten, S. Martinez and P. Picco. Existence of quasi-stationary distributions. A renewal dynamical approach. *Ann. Probab.* **23** (2) (1995) 501–521. MR1334159

[14] P. A. Ferrari and N. Maric. Quasi-stationary distributions and Fleming–Viot processes in countable spaces. *Electron. J. Probab.* **12** (24) (2007) 684–702. MR2318407

[15] W. H. Fleming and M. Viot. Some measure-valued Markov processes in population genetics theory. *Indiana Univ. Math. J.* **28** (5) (1979) 817–843. MR0542340

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

[17] I. Grigorescu and M. Kang. Immortal particle for a catalytic branching process. *Probab. Theory Related Fields* **153** (1–2) (2011) 333–361. MR2925577

[18] S. C. Harris and M. I. Roberts. The many-to-few lemma and multiple spines. Available at arXiv:1106.4761.

[19] P. Maillard. Branching Brownian motion with selection of the *N* right-most particles: An approximate model. Available at arXiv:1112.0266v2.

[20] M. K. Nakayama, P. Shahabuddin and K. Sigman. On finite exponential moments for branching processes and busy periods for queues. *J. Appl. Probab.* **41** (2004). MR2057579

[21] P. Robert. *Stochastic Networks and Queues. Stochastic Modelling and Applied Probability. Applications of Mathematics* **52**. Springer, New York, 2003. MR1996883

[22] L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales, 1: Foundations,* 2nd edition. Wiley, Chichester, 1994. MR1331599

[23] E. Seneta and D. Vere-Jones. On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states. *J. Appl. Probab.* **3** (1966) 403–434. MR0207047

[24] D. Villemonais. Interacting particle systems and Yaglom limit approximation of diffusions with unbounded drift. *Electron. J. Probab.* **16** (2011) 1663–1692. MR2835250

[25] A. M. Yaglom. Certain limit theorems of the theory of branching random processes. *Dokl. Akad. Nauk SSSR (N.S.)* **56** (1947) 795–798. MR0022045

[26] V. M. Zolotarev. More exact statements of several theorems in the theory of branching processes. *Theory Probab. Appl.* **2** (3) (1957) 245–253. MR0096321